


ВСЕРОССИЙСКАЯ КОНФЕРЕНЦИЯ
С МЕЖДУНАРОДНЫМ УЧАСТИЕМ

# ОБРАБОТКА ПРОСТРАНСТВЕННЫХ ДАННЫХ В ЗАДАЧАХ МОНИТОРИНГА ПРИРОДНЫХ И АНТРОПОГЕННЫХ ПРОЦЕССОВ

26-30 августа 2019 г.
г. Бердск

**Сборник трудов**

УДК 004

**Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов (SDM-2019)** [Электронный ресурс]: Сборник трудов всероссийской конференции (26-30 августа 2019 г., г. Бердск). Новосибирск: ИВТ СО РАН, 2019. 492 с. ISBN 978-5-905569-11-1.

Сборник трудов подготовлен по результатам работы всероссийской конференции «Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов» (SDM-2019), которая проходила с 26 по 30 августа 2019 г. на базе ООО «Санаторий Рассвет» (г. Бердск, Новосибирская область). В сборнике представлены результаты исследований по следующим направлениям: интегрированные геоинформационные технологии и системы в задачах мониторинга; оперативный региональный спутниковый мониторинг окружающей среды; моделирование и мониторинг экологических и техногенных процессов и систем. Сборник будет полезен для научных и инженерных работников, аспирантов и студентов вузов, занимающихся проблемами мониторинга окружающей среды. Материалы сборника публикуются в авторской редакции.

## ПРОГРАММНЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

Председатель:

**Шокин Ю.И.** академик РАН, Институт вычислительных технологий СО РАН

Заместители председателя:

**Альт В.В.** академик РАН, Сибирский физико-технический институт аграрных проблем СФНЦА РАН

**Бычков И.В.** академик РАН, Институт динамики систем и теории управления СО РАН

**Потатуркин О.И.** д.т.н., профессор, Институт автоматики и электрометрии СО РАН

Члены программного комитета:

**Борзов С.М.** к.т.н., Институт автоматики и электрометрии СО РАН

**Винокуров Ю.И.** д.г.н., профессор, Институт водных и экологических проблем СО РАН

**Вологжина С.Ж.** к.г.н., Иркутский государственный университет

**Гордов Е.П.** д.ф.-м.н., профессор, Институт мониторинга климатических и экологических систем СО РАН

**Гук А.П.** д.т.н., профессор, Сибирский государственный университет геосистем и технологий

**Добрецов Н.Н.** к.г.-м.н., Институт геологии и минералогии им. В.С. Соболева СО РАН

**Жайнаков А.Ж.** академик НАН Кыргызской Республики

**Калимолдаев М.Н.** академик НАН Республики Казахстан, Институт информационных и вычислительных технологий КН МОН РК

**Лагутин А.А.** д.ф.-м.н., профессор, Алтайский государственный университет

**Москвичев В.В.** д.т.н., профессор, Красноярский филиал ИВТ СО РАН – СКТБ «Наука» СО РАН

**Ноженкова Л.Ф.** д.т.н., профессор, Институт вычислительного моделирования СО РАН

**Потапов В.П.** д.т.н., профессор, Кемеровский филиал ИВТ СО РАН

**Пузанов А.В.** д.б.н., профессор, Институт водных и экологических проблем СО РАН

**Ротанова И.Н.** к.г.н., профессор, Алтайский государственный университет

**Ружников Г.М.** д.т.н., Институт динамики систем и теории управления СО РАН

**Смагин С.И.** чл.-корр. РАН, Вычислительный центр ДВО РАН

**Сойфер В.А.** академик РАН, Институт систем обработки изображений РАН

**Стемпковский А.Л.** академик РАН, Институт проблем проектирования в микроэлектронике РАН

**Суторихин И.А.** д.ф.-м.н., профессор, Институт водных и экологических проблем СО РАН

**Сысолятин С.В.** д.х.н., Институт проблем химико-энергетических технологий СО РАН

**Тулохонов А.К.** академик РАН, Байкальский институт природопользования СО РАН

**Турчановский И.Ю.** к.ф.-м.н., Томский филиал ИВТ СО РАН

**Федотов А.М.** чл.-корр. РАН, Институт вычислительных технологий СО РАН

**Цибульский Г.М.** д.т.н., профессор, Институт космических и информационных технологий СФУ

**Чимитдоржиев Т.Н.** д.ф.-м.н., Институт физического материаловедения СО РАН

**Шайдуров В.В.** чл.-корр. РАН, Институт вычислительного моделирования СО РАН

**Якубайлик О.Э.** к.ф.-м.н., Институт вычислительного моделирования СО РАН

ISBN 978-5-905569-11-1

© Институт вычислительных технологий СО РАН, 2019
© Коллектив авторов, 2019

# СОДЕРЖАНИЕ

## ПЛЕНАРНЫЕ ДОКЛАДЫ

Цифровой мониторинг экосистемы озера Байкал
*Бычков И.В., Ружников Г.М., Хмельнов А.Е., Фёдоров Р.К., Маджара Т.И.* ..... 8
Анализ систематических и случайных ошибок определения высоты соснового леса по данным радиолокационной интерферометрии
*Кирбижекова И.И., Чимитдоржиев Т.Н., Дмитриев А.В., Балтухаев А.К., Дагуров П.Н.* . 15
Оценка объема сжигаемого природного газа в факельных установках на территории Западной Сибири по спутниковым данным в видимом и ближнем ИК-диапазонах в ночное время
*Лагутин А.А., Мордвин Е.Ю., Волков Н.В., Тучина Н.В.* ..... 22
Мониторинг, оценка рисков и безопасность объектов техносферы и урбанизированных территорий
*Москвичев В.В., Чернякова Н.А.* ..... 27
Выявление малых цветовых различий устройствами компьютерного зрения
*Пальчикова И.Г., Смирнов Е.С.* ..... 33
Текущие возможности Центра коллективного пользования «ИКИ-Мониторинг»
*Прошин А.А., Лупян Е.А., Кашницкий А.В., Балашов И.В., Бурцев М.А.* ..... 39
Цифровое маркирование видеоданных аэрофотосъемки
*Фаворская М.Н., Буряченко В.В., Гусев К.А.* ..... 45
Система обработки спутниковых данных в Региональном центре ДЗЗ ФИЦ КНЦ СО РАН
*Якубайлик О.Э., Кадочников А.А., Токарев А.В.* ..... 50

## ИНТЕГРИРОВАННЫЕ ГЕОИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И СИСТЕМЫ В ЗАДАЧАХ МОНИТОРИНГА

Онтолого-ориентированный подход к организации web-среды по активной сейсмологии
*Брагинская Л.П., Григорюк А.П., Ковалевский В.В.* ..... 56
Интерактивная среда моделирования процессов обработки и анализа данных ДЗЗ
*Брежнев Р.В., Маглинец Ю.А., Раевич К.В.* ..... 61
Применение комплекса программного обеспечения в прикладных селекционных задачах
*Гребенникова И.Г., Чешкова А.Ф., Чанышев Д.И., Стёпочкин П.И.* ..... 68
Использование метода последовательного квадратичного программирования для разделения смесей объектов на гиперспектральных изображениях
*Григорьева О.В., Саидов А.Г., Чапурский Л.И.* ..... 74
Алгоритмы последовательного анализа многоспектральных аэрокосмических снимков на основе использования структурно-статистического подхода при дешифрировании природных объектов
*Гук А.П., Алтынцев М.А., Евстратова Л.Г., Алтынцева М.А.* ..... 77
Оценка параметров снежного покрова на земной поверхности без рельефа и с рельефом методом радарной интерферометрии
*Дагуров П.Н., Дмитриев А.В., Добрынин С.И., Чимитдоржиев Т.Н.* ..... 84
Информативность текстурных признаков в задаче классификации почвенно-растительного покрова
*Дмитриев Е.В., Козодеров В.В., Соколов А.А.* ..... 91
Интегрированная геоинформационная система мониторинга состояния внутриконтинентальных водных объектов
*Донцов А.А., Суторихин И.А., Коломейцев А.А.* ..... 98
Обработка сигналов в приемной системе с пространственно-распределенными приемными элементами
*Зима Д.Н., Соколова Д.О., Спектор А.А.* ..... 103

Геоинформационные технологии и программное обеспечение для оперативной оценки загрязнения атмосферы
*Кадочников А.А., Токарев А.В., Якубайлик О.Э.* ..... *108*
Прогноз изменения климата Алтае-Саянской горной страны до 2030 года
*Кирста Ю.Б., Ловцкая О.В., Пузанов А.В.* ..... *113*
Методы обработки пространственных данных в задачах мониторинга природных явлений, основанные на байесовском подходе
*Климова Е.Г.* ..... *117*
Картографическое моделирование температурных полей почв Среднесибирского трансекта на основе совместного анализа наземных и спутниковых данных
*Кудряшова С.Я., Чумбаев А.С., Пестунов И.А., Синявский Ю.Н., Чубаров Д.Л., Безбородова А.Н., Ермаков Н.Б.* ..... *123*
Радиолокационное обнаружение сосредоточенных объектов при использовании широкополосного излучения
*Лозовский И.Ф., Мурасев А.А., Спектор А.А.* ..... *131*
Непараметрическое обнаружение сейсмического сигнала по опорному и рабочему участку амплитудного спектра
*Морозов Ю.В., Райфельд М.А., Спектор А.А.* ..... *136*
Нейросетевая классификация крупноразмерных многосегментных полигонов с формированием признаков путем преобразования Гильберта – Хуанга гиперспектральных данных
*Нежевенко Е.С.* ..... *141*
Виртуальная исследовательская среда в фундаментальных и прикладных задачах климатологии
*Гордов Е.П., Окладников И.Г., Титов А.Г., Рязанова А.А.* ..... *146*
UNISAT. Технология построения унифицированных систем ведения сверхбольших распределенных архивов разнородных спутниковых данных
*Прошин А.А., Лупян Е.А., Кашницкий А.В., Балашов И.В.* ..... *153*
Построение зависимости размера кроны сосны обыкновенной от диаметра ствола
*Пятаев А.С., Вайс А.А.* ..... *159*
Метод распознавания лиц по визуальным данным
*Пятаева А.В., Верхотурова М.В.* ..... *165*
Рассмотрение подходов к представлению знаний об оперировании пространственными данными в задачах мониторинга
*Раевич К.В., Маглинец Ю.А., Брежнев Р.В.* ..... *172*
Локализация точечных источников со случайным пространственным расположением и случайной дисциплиной генерации импульсов
*Резник А.Л., Тузиков А.В., Торгов А.В., Соловьев А.А., Ковалев В.А.* ..... *178*
Каркасный подход к реализации библиотек высокопроизводительной обработки изображений
*Русин Е.В.* ..... *183*
Алгоритм восстановления коэффициентов отражения земной поверхности по данным MODIS с учетом неоднородности земной поверхности
*Тарасенков М.В., Белов В.В., Зимовая А.В., Энгель М.В.* ..... *189*
Восстановление коэффициентов отражения земной поверхности по данным спутников Ресурс-П и Метеор-М
*Тарасенков М.В., Белов В.В., Энгель М.В., Гриднев Ю.В., Зимовая А.В.* ..... *197*
Многомерное представление и анализ данных стационарных наземных наблюдений за концентрациями атмосферных примесей для верификации гидродинамических моделей прогноза
*Токарев В.М., Здерева М.Я.* ..... *204*

Автоматизация формирования базы данных ледников на основе ДЗЗ
*Фёдоров Р.К., Китов А.Д., Авраменко Ю.В.* ..... *207*
Оценка газообмена в Иркутском водохранилище
*Хмельнов А.Е., Гаченко А.С., Михайлов А.А., Минаев В.В., Минаев А.В.* ..... *212*
Комбинация нейронной сети и линейной фильтрации для обнаружения объектов
*Шакенов А.К.* ..... *219*
Программный комплекс атмосферной коррекции данных российских спутниковых приборов
*Белов В.В., Гриднев Ю.В., Зимовая А.В., Тарасенков М.В., Энгель М.В.* ..... *224*
Информационная система сбора, хранения и обработки спутниковых данных ИВТ СО РАН как инструмент для решения задач мониторинга природных и антропогенных объектов
*Шокин Ю.И., Пестунов И.А., Чубаров Д.Л., Кихтенко В.А., Мамаш Е.А., Добрецов Н.Н., Воронина П.В., Синявский Ю.Н.* ..... *229*
О географической привязке контента текстовых документов
*Жижимов О.Л., Леонова Ю.В.* ..... *241*
Применение искусственных нейронных сетей для задачи распознавания теста
*Пятаева А.В., Генза С.А.* ..... *248*

**ОПЕРАТИВНЫЙ РЕГИОНАЛЬНЫЙ СПУТНИКОВЫЙ МОНИТОРИНГ ОКРУЖАЮЩЕЙ СРЕДЫ**

Сервисы анализа серии космоснимков для оценки динамики изменения объектов
*Авраменко Ю.В., Ружников Г.М., Фёдоров Р.К.* ..... *254*
Сопоставление двух методик определения параметров облачного покрова по данным сканера МСУ-МР (КА Метеор-М №2) и AVHRR (КА NOAA-18, 19, METOP-B)
*Волкова Е.В., Косторная А.А., Голомолзин В.В., Кузьминых С.А., Амикишиева Р.А.* ..... *260*
Особенности вариаций температуры и озона в атмосфере во время сейсмособытий на Ближнем Востоке
*Кашкин В.Б., Рублева Т.В., Одинцов Р.В.* ..... *264*
Обнаружение малоразмерных объектов в двухцветных изображениях с пространственно-нестационарным фоном
*Косых В.П., Громилин Г.И., Яковенко Н.С.* ..... *271*
Оценка загрязнения атмосферы по спутниковым данным различного пространственного разрешения
*Краснощеков К.В., Якубайлик О.Э.* ..... *275*
Астроклиматические условия Республики Алтай по данным спутниковых наблюдений
*Мордвин Е.Ю., Лагутин А.А., Волков Н.В., Райкин Р.И., Ревякин А.И., Клёпов М.А.* ..... *280*
Наблюдение кольцевых структур на льду Байкала с помощью спутников Sentinel 1 и Sentinel 2 весной 2018-2019
*Родионова Н.В.* ..... *284*
Определение талого/мерзлого состояния почв Якутии по радарным данным Sentinel 1
*Родионова Н.В.* ..... *289*
Программно-технологическое обеспечение геоинформационной веб-системы оперативной обработки и визуализации спутниковых данных
*Кадочников А.А.* ..... *295*

**МОДЕЛИРОВАНИЕ И МОНИТОРИНГ ЭКОЛОГИЧЕСКИХ И ТЕХНОГЕННЫХ ПРОЦЕССОВ И СИСТЕМ**

Критерии качества атмосферного воздуха городской среды для веществ канцерогенного действия
*Бельская Е.Н., Иванова У.С., Тасейко О.В.* ..... *302*

Обработка данных натурных наблюдений поверхностной скорости и температуры в озере Шира
*Володько О.С., Компаниец Л.А., Толомеев А.П., Дроботов А.В., Гаврилова Л.В.* ..... *307*
Исследование температурного режима Новосибирской агломерации по данным спутникового зондирования
*Воронина П.В., Чубаров Д.Л., Добрецов Н.Н.* ..... *313*
Биогеохимические особенности накопления Zn, Cu, Pb и Cd растениями постпирогенных болот
*Гашкова Л.П.* ..... *320*
Информативность NDVI в изучении пространственной изменчивости содержания магния в почвах и растениях
*Гопп Н.В., Нечаева Т.В., Савенков О.А., Смирнова Н.В., Тархова В.П., Смирнов А.В.* ..... *324*
Описание урбанизированной поверхности с использованием геопространственных данных для задач моделирования и оценки качества воздуха
*Гочаков А.В., Пененко А.В., Колкер А.Б., Антохин П.Н.* ..... *330*
Анализ климатических характеристик территории распространения сибирского шелкопряда
*Дергунов А.В., Якубайлик О.Э.* ..... *335*
Дифференцированный подход к оценке индивидуального риска ЧС и происшествий
*Иванова У.С. ,Черных Д.А., Тасейко О.В., Никулина Т.Н.* ..... *340*
Анализ данных о предвестниках сильных землетрясений
*Сибгатулин В.Г., Перетокин С.А., Кабанов А.А., Симонов К.В.* ..... *345*
Пространственно-временные градиенты парметров состояния поверхности Земли для спутникового мониторинга природных систем
*Картушинский А.В. [(1)(2)], Огородов Н.А.[(3)], Ларько А.А.[(1)]* ..... *352*
Прогнозирование распространения болезней в тропических зонах с помощью методов машинного обучения 357
*Колесников А.А.[(1)], Кикин П.М.[(2)]* ..... *357*
Построение аналитической модели для специализированной модельно-ориентированной системы сбора данных научных исследований 363
*Коробко А.В., Коробко А.А.* ..... *363*
Место геоинформатики в цифровой экономике России 370
*Кошкарев А.В.* ..... *370*
Использование спутниковой информации в исследовании процессов распространения примесей в пограничном слое атмосферы 375
*Леженин А.А., Рапута В.Ф., Ярославцева Т.В.* ..... *375*
Методы автоматизированной обработки много- и гиперспектральных данных дистанционного зондирования Земли в интересах решения прикладных задач оценки состояния природных и антропогенных объектов
*Марков А.В., Григорьева О.В., Жуков Д.В., Иванец М.О.* ..... *381*
Поиск изменений температуры городской среды по спутниковым данным на примере города Красноярска
*Гостева А.А., Матузко А.К., Якубайлик О.Э.* ..... *387*
Интегральное аналитическое оценивание состояния природно-техногенной безопасности территорий Красноярского края
*Метус А.М., Пенькова Т.Г.* ..... *392*
Вероятностный анализ сейсмической опасности площадок особо ответственных объектов
*Миронов В.А., Перетокин С.А., Симонов К.В.* ..... *399*
Метан в атмосфере Западной Сибири: результаты спутниковых наблюдений и моделирования
*Мордвин Е.Ю., Лагутин А.А., Волков Н.В., Макушев К.М., Прилипкова А.С.* ..... *404*

Чувствительность уходящего длинноволнового потока к вариациям коэффициента излучения подстилающей поверхности
*Мордвин Е.Ю., Лагутин А.А., Кравченко Р.В.* ..... *409*
Метод использования критериев опасностей для идентификации рискоопасных ситуаций
*Ничепорчук В.В.* ..... *413*
Информационная технология обработки спутниковых данных для вычисления градиентных характеристик поверхности мирового океана
*Огородов Н.А., Картушинский А.В.* ..... *420*
Веб-сервис классификации сейсмических событий на базе системы распределенных вычислений Apache Spark
*Попов С.Е., Замараев Р.Ю.* ..... *426*
Оценка снегозапасов по данным микроволнового радиометра AMSR-2
*Ромасько В.Ю., Гордеев И.Н., Бураков Д.А.* ..... *437*
Сегментация оливковых деревьев с использованием архитектуры Mask R-CNN
*Сафонова А.Н., Табик С., Маглинец Ю.А.* ..... *442*
Классификация пространственных структур отрогов Большого Васюганского болота с использованием спутниковых данных
*Синюткина А.А.* ..... *447*
Сигналы GPS, ГЛОНАСС и GALILEO в мониторинге состояния древостоя
*Сорокин А.В., Иванов В.А., Пономарев Е.И., Макаров Д.С., Харламов Д.В., Голубев Д.В.* *452*
Влияние антропогенных факторов на качество воды в реке в районах с интенсивной угледобычей
*Счастливцев Е.Л., Быков А.А., Юкина Н.И.* ..... *457*
Отслеживание перемещений диких животных по визуальным данным фотоловушек на основе сетей глубокого обучения
*Фаворская М.Н., Пахирка А.И., Зотин А.Г.* ..... *461*
Анализ влияния факторов окружающей среды на смертность населения г. Красноярска
*Черных Д.А., Тасейко О.В., Иванова У.С.* ..... *466*
Использование беспилотных летательных аппаратов для съемки микрорельефа при мониторинге агроценозов
*Шевырногов А.П., Мальчиков Н.О., Ботвич И.Ю., Емельянов Д.В., Ларько А.А.* ..... *471*
Пространственный анализ загрязнения атмосферы Красноярска
*Якубайлик О.Э., Заворуев В.В., Малимонов М.И., Пушкарев А.А.* ..... *477*
Обнаружения болезней на посевах пшеницы по данным гиперспектральной съемки
*Дубровская О.А., Гурова Т.А., Пестунов И.А.* ..... *483*
Вычисление сочетания генетических маркеров в определении идентичности образцов древесины
*Пятаев А.С., Ибе А.А., Шилкина Е.А.* ..... *487*

# ПЛЕНАРНЫЕ ДОКЛАДЫ

# ЦИФРОВОЙ МОНИТОРИНГ ЭКОСИСТЕМЫ ОЗЕРА БАЙКАЛ

*Бычков И.В.[(1)], Ружников Г.М.[(1)], Хмельнов А.Е.[(1)(2)], Фёдоров Р.К.[(1)(2)], Маджара Т.И.[(1)]*
[(1)] Институт динамики систем и теории управления имени В.М. Матросова СО РАН, Иркутск
[(2)] Иркутский научный центр СО РАН, Иркутск

Рассмотрены актуальность и направления создания и развития комплексного цифрового мониторинга и оценки изменений экосистемы озера Байкал и его прибрежной территории с использованием больших объёмов разноформатных пространственно-временных тематических данных и знаний, сервис-ориентированной парадигмы, стандартов OGC, Web-сервисов, а также внедрения современных информационно-телекоммуникационных технологий.
*Ключевые слова: пространственные и временные данные цифрового мониторинга, сервис-ориентированная парадигма, сервисы*

**Введение.** Рассмотрены вопросы организации системы цифрового экологического мониторинга озера Байкал и его прибрежной территории с использованием современных цифровых технологий и сервисов обработки больших объёмов распределённых пространственно-временных данных.

**Мониторинг озера Байкал.** Исследования уникальных экосистем, проводимые в мире и России, базируются на комплексном мониторинге, хранении и обработке больших объёмов научных данных и знаний о системе, которые имеют пространственно-временной характер, а также на использовании распределённых информационно-вычислительных технологий и их приложений и современных сетей передачи данных [1-3].

Сохранение озера Байкала, как мирового источника чистой пресной воды и как природного участка с уникальной фауной и флорой, является главной природоохранной задачей России и важнейшим условием устойчивого развития региона. Байкальская природная территория (БПТ) представлена уникальными ландшафтами, требующими своего сохранения в естественном виде согласно «Конвенции об охране всемирного культурного и природного наследия ЮНЕСКО» и законодательства Российской Федерации.

Основные направления исследований озера отмечены в Указе Президента РФ №204 от 7 мая 2018 года «Сохранение уникальных водных объектов, в том числе реализации задач по сохранению озера Байкал», а также определены в рамках Национального проекта (программы) «Экология» (Федеральный проект «Сохранение озера Байкал»).

Постановлением Правительства РФ от 2 февраля 2015 г. № 85 «Об утверждении Положения о государственном экологическом мониторинге уникальной экологической системы озера Байкал» отмечено, что он является частью государственного экологического мониторинга (государственного мониторинга окружающей среды).

Под государственным экологическим мониторингом понимается комплексная система регулярных наблюдений в пространстве и во времени за состоянием окружающей среды и её изменениями под воздействием природных и антропогенных факторов [4]. Для обеспечения унификации проводится мониторинг строго регламентированных наблюдений по утверждённому списку параметров состояния окружающей среды, а также по заданным требованиям к используемым средствам, приёмам измерений, частоте отбора проб и методам исследований. Система экологического мониторинга включает сбор данных о фактическом состоянии загрязнения окружающей среды, тщательную обработку, анализ этих данных с последующим выявлением динамики состояния и выработкой рекомендаций по развитию экономики региона на основе научно обоснованных экологических прогнозов.

Одной из целей экологического мониторинга является информационное обеспечение управленческих и инвестиционных решений на основе достоверных, своевременных, полных данных о состоянии окружающей природной среды.

Государственный мониторинг экосистемы озера Байкал осуществляется уполномоченными федеральными органами исполнительной власти: Министерством природных ресурсов

и экологии РФ, Министерством сельского хозяйства РФ, Федеральной службой по гидрометеорологии и мониторингу окружающей среды, Федеральной службой государственной регистрации, кадастра и картографии, Федеральным агентством лесного хозяйства, Федеральным агентством по недропользованию, Федеральным агентством водных ресурсов и Федеральным агентством по рыболовству, а также органами исполнительной власти Республики Бурятия, Забайкальского края и Иркутской области в соответствии с их компетенцией в порядке, установленном постановлением Правительства РФ от 9 августа 2013 г. № 681 «О государственном экологическом мониторинге и государственном фонде данных государственного экологического мониторинга». Основные требования к осуществлению мониторинга: территориальный охват в пределах БПТ; детализация подходов – с учетом экологического зонирования БПТ, установленного статьей 2 Федерального закона «Об охране озера Байкал»; учет правовых режимов – с учетом особого режима хозяйственной и иной деятельности, осуществляемой на БПТ. К сожалению, действующие нормативные акты противоречивы и не позволяют реализовать специальные полномочия по мониторингу уникальной экосистемы озера Байкал.

Междисциплинарные научные исследования, экспедиции и частичный мониторинг экосистемы озера Байкала и его прибрежной территории также проводят институты РАН, университеты Минобрнауки [6, 9].

Государственный экологический мониторинг экосистемы озера Байкал включает:

Регулярными наблюдениями не охвачена прибрежная мелководная зона Байкала, которая испытывает наибольшую антропогенную нагрузку. Особенно это касается о. Ольхон и района Малого моря, где летом в разы возрастает поток туристов.

За последнее десятилетие озеро Байкал и его прибрежная территория претерпевают значительные экологические и ландшафтные изменения. К наиболее существенным из них можно отнести: снижение водности притоков, увеличение эмиссии метана в водную толщу, аномальное развитие нитчатых водорослей; вымирание байкальских губок; изменение структуры и таксономического состава макрозообентоса вблизи городов и поселков в период массового развития спирогиры с вселением видов, несвойственных Байкалу, развитие токсин-продуцирующих цианобактерий в количествах, близких к критическим; снижение биомассы байкальского омуля и т.д. Также нельзя не отметить изменение ландшафтов водосборной площади озера Байкал и БПТ из-за массовых пожаров и хозяйственной деятельности человека, роста выпадения кислых атмосферных осадков, эпидемиологического поражения хвойных пород деревьев таежных массивов, загрязнения подземной гидросферы. Кроме этого, озеро Байкал и БПТ подвержены опасным эндогенным и экзогенным геологическим процессам, таким как сейсмичность, обвалы, сели и т.п. Не исключена опасность значительных колебаний уровня озера в результате цунами.

В настоящее время к особенностям проводимого мониторинга экосистемы озера Байкал можно отнести:




Рис. 1. Государственный мониторинг экосистемы озера Байкал.

- Использование не скоординированных между собой сетей передачи данных наблюдений.
- Ведение ведомственных схем мониторинга, формирующих и использующих большие объёмы пространственных и тематических данных, которые, как правило, локализованы и не скоординированы между собой, в параметрическом, хронологическом и других аспектах. В открытый доступ выложены только разрозненные данные. Это осложняет проведение комплексных оценок, прогнозирования, принятие управленческих решений на базе имеющихся данных.
- Отсутствие оценки системной достаточности выбора «информативных показателей» мониторинга экосистемы озера и его прибрежной территории, а также единой системы идентификации пространственных объектов, как универсальных элементов связи ведомственных пространственных и тематических баз данных.
- Отсутствие, в режиме реального времени, полноценного государственного мониторинга гидрохимических, гидрофизических и биологических параметров по всей акватории озера, а проведение только сезонных наблюдений.
- Остается вне системных наблюдений большая часть мелководной зоны Байкала, характеризующаяся высокой антропогенной нагрузкой (Малое море, Баргузинский и Чивыркуйский заливы, Селенгинское мелководье, южная оконечность Байкала).
- Существующие ведомственные системы мониторинга экосистемы озера Байкал не позволяют оперативно реагировать на изменения природного и антропогенного характера, выявлять в них компоненты локального или глобального генезиса.

Всё это неблагоприятно воздействуют на экологическую и социально-экономическую ситуацию в Байкальском регионе, а также на имидж его туристического потенциала.

В целях повышения эффективности экологического мониторинга озера необходимо усилить три составляющие:

- *Технологическую.* Дополнить существующие ведомственные сети станций мониторинга природных и антропогенных факторов автоматическими станциями, логгерами и оборудованием низковысотного дистанционного зондирования, БПЛА и т.д., позволяющими регистрировать параметры водной среды и атмосферы в квазинепрерывном режиме, с онлайн передачей информации в ЦОД.
- *Научную.* Внедрить в существующие методологию и схемы мониторинга природных и антропогенных факторов озера Байкал и БПТ современные молекулярно-генетические и информационные методы, технологии и т.д. В настоящее время, ведомствами созданы большие объёмы данных о процессах, проходящих в экосистеме озера Байкал, изучены формы этих процессов и механизмы формирования. По отдельным видам мониторинга, разработаны уникальные модели, а также информационные системы обработки, анализа, прогноза, оптимизации и оценки влияния природно-климатических факторов и процессов жизнедеятельности на озере Байкал.
- *Интеграционную.* Для интеграции существующего опыта необходимо модернизировать компьютерные и программное обеспечение центров коллективного пользования (ИИВС ИРНОК, ИСКЦ) и алгоритмы, модели, информационные базы данных и методы их обработки для комплексного мониторинга, создать уникальную сетевую инфраструктуру, дополнив существующие объекты новым оборудованием, новыми аппаратными мощностями для сбора, хранения и передачи данных, применяя новые способы обработки, интерпретации и прогноза.

**Цифровой мониторинг экосистемы озера Байкал.** Исследования по формированию комплексный цифрового мониторинга экосистем полностью соответствуют приоритетам Стратегии научно-технологического развития Российской Федерации - переход к передовым цифровым, интеллектуальным производственным технологиям, роботизированным системам, новым материалам и способам конструирования, создание систем обработки больших объемов данных, машинного обучения и искусственного интеллекта.

Необходимо создать информационно-телекоммуникационную платформу Цифрового экологического мониторинга (ЦЭМ) обеспечивающего координацию межведомственного (Росгидромет, Минприроды, Рослесхоз, Росводресурсов, Росрыболовства РАН, Минобрнауки) и межрегионального (Республика Бурятия, Забайкальский край и Иркутская область) взаимодействия. На базе ЦЭМ должно осуществляться сбор, хранение и обработки пространственно-временных тематических данных цифрового экологического мониторинга, с возможностью прямого информационного доступа. Интеграция распределённых междисциплинарных данных (Big Data) в ЦЭМ позволит значительно повысить качество аналитических и прогнозных моделей по сохранению и развитию экосистемы Байкала. Подобный дата центр реализован в рамках Национального управления океанических и атмосферных исследований США (NOAA).

Ключевыми составляющими информационно-телекоммуникационной платформы ЦЦЭМ цифрового мониторинга экосистемы озера Байкала и БПТ являются: инфраструктуры центров коллективного пользования ИДСТУ СО РАН «Интегрированная информационно-вычислительная сеть Иркутского научно-образовательного комплекса» и «Иркутский суперкомпьютерный центр СО РАН», а также «Информационно-аналитическая среда» (ИАС), которые обеспечат сбор, передачу, поиск, хранение, и параллельную обработку больших объемов данных, возможность онлайн-доступа к данным, каталогам, сервисам и информационно-аналитическим системам, возможность проведения на основе полученных данных оценки, моделирования и прогноза экологических и климатических изменений Байкала и прилегающих территорий с применением средств суперкомпьютерного моделирования и облачных вычислений.

Для проведения цифрового мониторинга озера Байкал разрабатываются эффективные методы и технологии, а также распределенная сервис-ориентированная информационно-аналитическая среда (ИАС) геопортального типа, включающая подсистемы сбора, передачи, хранения, поиска и обработки больших объемов разноформатных пространственно-временных данных и знаний [5].

Методология исследований базируется на комплексном применении сервис-ориентированной парадигмы и современных технологий распределенной обработки данных, использовании декларативных спецификаций и интеллектуализации с использованием методов и технологий глубокого обучения. При этом декларативные спецификации обеспечивают компактность, выразительность и предметную ориентированность, включая возможность интерпретации трансформационными и другими процедурами. В свою очередь, использование сервис-ориентированного подхода позволяет проводить полноценный учет распределенных информационных ресурсов в сочетании с легкостью тестирования, масштабируемостью и возможностью повторного использования создаваемых сервисов [5].

Для организации комплексного мониторинга экологических систем озера Байкал сервисы тематического экологического мониторинга интегрируются с помощью логических конструкций для решения задач обработки данных, управления потоком выполнения и т.д. Разнообразие коллекций разрабатываемых сервисов тематического экологического мониторинга позволит передавать данные между ними, согласовывать форматы данных, запускать асинхронные вычислительные процессы.

ИАС должна с заданной степенью надежности сохранять любые виды данных экологического мониторинга: временные ряды с результатами измерений различных датчиков, материалы экспедиций, космические снимки, векторные карты и т.д. (рис. 2). Все эти данные, характеризующиеся пространственной и временной привязкой, могут иметь ряд дополнительных атрибутов, специфичных для конкретного вида информации.

Геопортальная сервис-ориентированная информационно-аналитическая среда комплексного цифрового мониторинга озера Байкал позволит обеспечивать:

- онлайн-доступ к распределенным сенсорам;
- доступ к архивным данным сенсоров;
- высокую скорость обработки данных;

– доступ к высокопроизводительным вычислительным ресурсам и ресурсам хранения данных центров коллективного пользования;
– масштабирование вычислительных ресурсов и ресурсов хранения и обработки данных центров коллективного пользования с учетом роста числа задач и объемов данных мониторинга;
– возможность использования различных методов и технологий распределённой обработки данных.

В качестве программно-аппаратной платформы, обеспечивающей непрерывную работу ИАС, используется схема построения центров обработки данных (рис. 3), в основе которой лежат следующие основополагающие принципы:
– полное резервирование инженерной инфраструктуры и аппаратного комплекса;
– эффективное использование оборудования за счет организации пулов вычислительных ресурсов;
– виртуализация ресурсов и приложений;
– системы резервного копирования и восстановления.



Рис. 2 Структура ИАС.

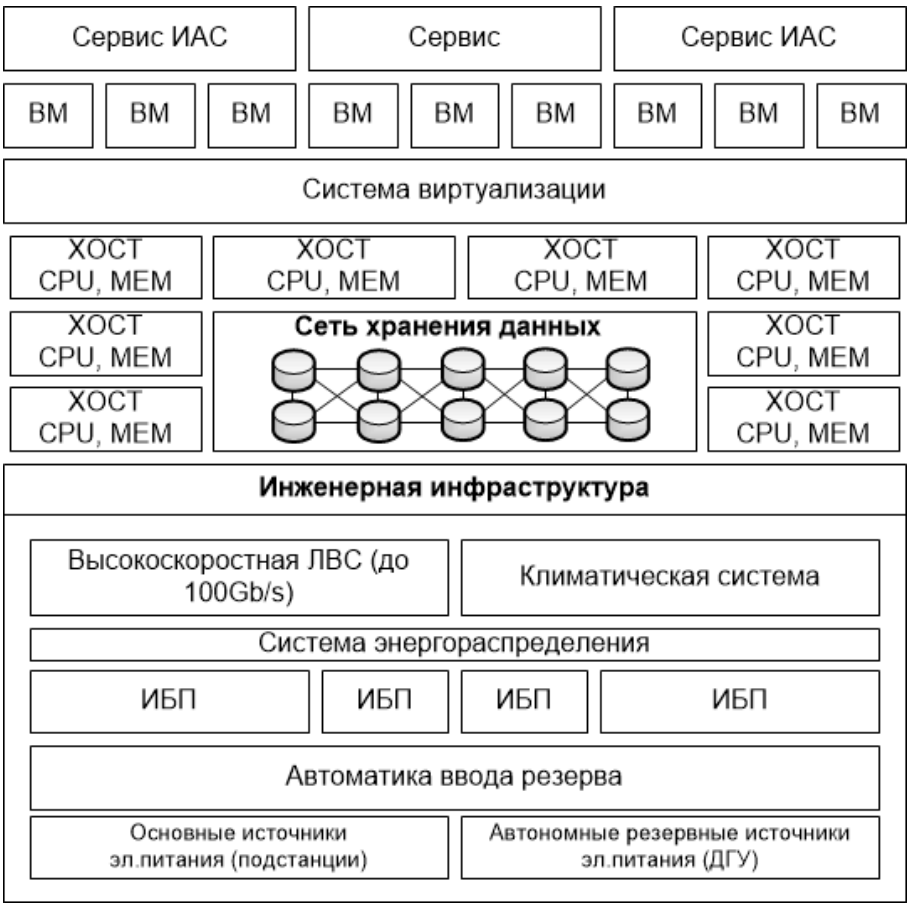



Рис. 3. Инфраструктура ЦОД.

Создание ЦОД на начальном этапе включает:

- реконструкцию инженерной инфраструктуры – систем бесперебойного питания и охлаждения;
- модернизацию сетевой инфраструктуры;
- развертывание серверов и систем хранения и обработки данных на базе оборудования HPE и программного обеспечения виртуализации VMWare vSphere.
- масштабирование вычислительной инфраструктуры - дооснащение вычислительного кластера ЦКП ИСКЦ узлами на базе:
  - процессоров Intel Xeon и графических ускорителей NVidia Tesla последних поколений (+ 300 TFlops);
  - высокопроизводительной системы хранения данных с параллельной файловой системой Panasas ActiveStor 18 (+ 120 TB)

В качестве стартовой площадки для развертывания ЦОД будут использованы существующие и успешно функционирующие центры коллективного пользования ИИВС ИрНОК и ИСКЦ, что обеспечит повышение доступности высокопроизводительных вычислительных ресурсов для обработки данных цифрового мониторинга, в том числе, с применением средств суперкомпьютерного моделирования.

**Заключение.** Предложенная информационно-телекоммуникационная платформа цифрового мониторинга экосистемы озера Байкала и БПТ, позволит перейти на качественно новый уровень проведения междисциплинарных научных исследований, а также обеспечит оперативный анализ и принятие решений по проблемам озера и его природной территории.

*Исследования проводятся при частичной финансовой поддержке РФФИ (проекты №№ 18-07-00758-а, 17-57-44006-монг-а, 17-47-380007-р), программ президиума РАН (№ 2 и № 23), интеграционной программы ИНЦ СО РАН и центров коллективного пользования ИДСТУ СО РАН.*

## ЛИТЕРАТУРА


[1] NOAA`s Nartional Centers for Environmental Information. **https://www.ngdc.noaa.gov**.

[2] Center for International Earth Science Information Network (CIESIN). **http://www.ciesin.org**.

[3] Биоразнообразие и динамика экосистем: информационные технологии и моделирование. Отв. ред. акад. В.К. Шумный, акад. РАН Ю.И. Шокин. Новосибирск: Изд-во СО РАН, 2006. 648 с.

[4] *Израэль Ю.А.* Экология и контроль состояния природной среды. Л.: Гидрометеоиздат, 1979. 376 с.

[5] *Бычков И.В., Ружников Г.М., Хмельнов А.Е. и др.* Инфраструктура информационных ресурсов и технологии создания информационно-аналитических систем территориального управления. Новосибирск: Изд-во СО РАН, 2016. 240 с.

[6] *Сутурин А.Н., Чевыркин Е.П., Мальник В.В. и др.* Роль антропосферных факторов в развитии экологического стресса в литорали озера Байкал (акватория пос. Листвянка) // География и природные ресурсы. 2016. № 6. С.43-54.

[7] *Кравцова Л.С. Ижболдина Л.А., Ханаев И.В. и др.* Нарушение вертикальной зональности зелёных водорослей в открытом Лиственничном заливе озера Байкал, как следствие антропогенного воздействия // Докл. РАН. 2012. Т. 447, № 2. С. 227-229.

[8] *Грачов М.А.* О современном состоянии экологической системы озера Байкал. Новосибирск: Изд-во СО РАН, 2002. 156 с.

[9] *Самсонов Д.П., Кочетков А.И., Пасынкова Е.М. и др.* Содержание стойких органических загрязнителей в компонентах уникальной экологической системы озера Байкал // Метеорология и гидрология. 2017. № 5. С. 105-115.

# АНАЛИЗ СИСТЕМАТИЧЕСКИХ И СЛУЧАЙНЫХ ОШИБОК ОПРЕДЕЛЕНИЯ ВЫСОТЫ СОСНОВОГО ЛЕСА ПО ДАННЫМ РАДИОЛОКАЦИОННОЙ ИНТЕРФЕРОМЕТРИИ

*Кирбижекова И.И., Чимитдоржиев Т.Н., Дмитриев А.В., Балтухаев А.К., Дагуров П.Н.*
Институт физического материаловедения СО РАН, Улан-Удэ

Представлены результаты исследования систематических и случайных ошибок определения высоты соснового леса, полученной на основе разности интерферометрических фаз радиолокационных данных X- и L-диапазона.
*Ключевые слова: радиолокационная интерферометрия, высота леса, статистический анализ.*

**Введение.** В настоящее время для Российской Федерации актуальной является разработка технологий определения таксационных показателей леса на основе данных дистанционного зондирования Земли (ДЗЗ). Согласно Лесному кодексу Российской Федерации (ЛК РФ) с 2006 г. используется система лесоучетных работ – государственная инвентаризация лесов (ГИЛ). ГИЛ должна проводиться на основе статистической выборки и базироваться на результатах измерений и расчетов на пробных площадях, заложенных в пределах кластеров – участков лесного фонда [1]. При реализации ее задач должны использоваться как наземные, так и дистанционные методы оценки. Разработка методов ДЗЗ имеет стратегическое значение в организации и ведении лесного хозяйства, в частности для изучения и мониторинга изменений, происходящих в лесах в процессе проведения лесохозяйственных мероприятий и соответствие результатов этих мероприятий лесным планам субъектов РФ.

К таксационным показателям относят количественные и качественные параметры, оценки биологических и физических особенностей структуры лесной среды, производительности насаждения в пределах занимаемой им территории: происхождение, форма, состав, средняя высота, средний диаметр, возраст, класс бонитета, полнота, запас древесины, класс товарности лесного насаждения, подрост, подлесок, тип леса и др. [2]. Характеристики леса можно разделить на две группы: первичные (непосредственно физически измеряемые) и производные рассчитываемые на основе эмпирических соотношений. Спутниковые определения являются в большинстве случаев производными на основе исследований закономерностей и связей параметров лесной среды со спектральными и текстурными характеристиками оптических и радиолокационных изображений объектов исследования. Методы радиолокационной интерферометрии и поляриметрии позволяют количественно оценить ключевые характеристики лесной среды такие как высота и биомасса.

В ИФМ СО РАН в настоящее время проводятся исследования возможности комбинирования оптических и радиолокационных интерферометрических и поляриметрических данных для определения характеристик хвойных и смешанных лесов Байкальского региона и создания технологии дистанционного определения количественных таксационных характеристик [3]. В данной статье представлены промежуточные результаты исследований по определению высоты лесного полога на основе разности интерферометрических фаз X- и L-диапазонов (TanDEM-TerraSAR-X и ALOS PALSAR). Основной целью данного блока исследований является анализ систематических и случайных ошибок интерферометрических измерений высоты.

**Характеристика территории и объекта исследования.** Территория исследования Верхняя Березовка располагается на севере г. Улан-Удэ и представляет собой пригородную зеленую зону в отрогах горного хребта Улан-Бургасы. Географические координаты полигона: 51°54′ с.ш. и 106°11′ в.д. Местность в основном покрыта сосновым лесом, с небольшими включениями лиственницы, березы, осины, ольхи и др. на дерново-лесных и супесчаных почвах. Средние высоты в данной зоне не превышают значений 800-1000 м, превышение высот над

долинами р. Селенга и Уда составляет до 310 м. Поверхность склонов, впадин и гребней холмов относительно гладкая, местами присутствуют противопожарные рвы, изрезанные овраги глубиной 3-5 м и более.

Климат является умеренным резко континентальным, засушливым. Диапазон значений температуры за период метеонаблюдений 1847-2018 гг. составил 95°С: от –54,4 °С в 1931 г., до + 40,6°С в 2016 г. Среднемесячные температуры варьируются от –23,3°С в январе до +19,8°С в июле [4]. Среднегодовая температура варьируется от –3,5° в 1850 г. до +1,7° в 2007 г., в среднем –1,23°С, наблюдается положительный тренд, прирост составляет +0,0215°С/год. Количество осадков невелико. Среднемесячные значения варьируются от 3 мм в феврале-марте до 68 мм в августе, среднегодовое количество осадков – 265 мм. Снежный покров устанавливается на 130 дней/год, среднемесячная толщина в ноябре-марте достигает 4-12 см, максимально – 30 см. Влажность воздуха варьируется от 49% в мае до 77-78% в декабре-январе. Уровень солнечной радиации высок, варьируется от 93 ч/мес. в декабре до 288 ч/мес. в мае–июле, что составляет более 2400 часов/год.

Подспутниковые и спутниковые измерения произведены сначала на 36 тестовых участках площадью 900-1200 кв.м. Затем это число было доведено до 84. Из них $n_1$=21 пункт на относительно плоских участках с уклоном менее 5°, $n_2$=42 – на восточных склонах с уклоном 10-15°, и $n_3$=21 на западных и юго-западных склонах с уклоном 10-15°. На рис. 1 представлена трек Garmin и схема расположения 36 тестовых участков во время первых экспедиций в апреле-мае 2018 г. В качестве подложки использован снимок Ресурс-П Геотон от 30 марта 2018 г. с пространственным разрешением 2 м. Породный состав преимущественно представлен сосной, на отдельных участках редкие вкрапления березы, ольхи, акации, кустарников. По данным полевых измерений: для спелых деревьев диаметр стволов варьируется от 15 до 41 см, высота от 9 до 26 м. Для молодняка диаметр стволов 6–10 см, высота 6–10 м. Высота кустарников менее 2 м.

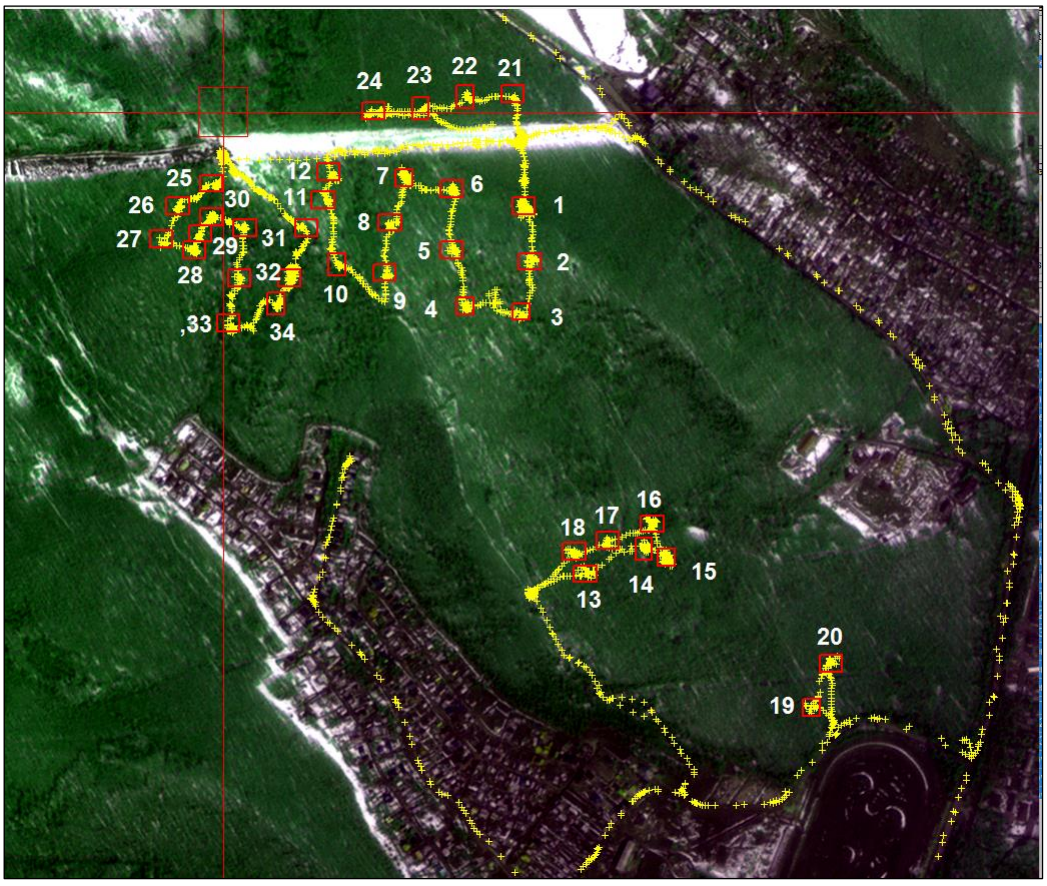



Рис. 1. Трек Garmin и схема расположения тестовых участков подспутниковых и спутниковых измерений на полигоне верхняя Березовка во время двух первых экспедиций 2018 г.

Табл. 1. Результаты подспутниковых измерений характеристик соснового леса на тестовых участках Верхней Березовки в апреле-мае 2018 г.

| № п/п | Высота, м | Запас древостоя, м³/га | полнота | Диаметр, см | № п/п | Высота, м | Запас древостоя, м³/га | полнота | Диаметр, см |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 17,67 | 170 | 0,6 | 28,67 | 19 | 19,67 | 310 | 1 | 28,00 |
| 2 | 14,33 | 210 | 1 | 23,67 | 20 | 16,50 | 260 | 1 | 26,67 |
| 3 | 22,67 | 260 | 0,7 | 39,00 | 21 | 14,33 | 160 | 0,8 | 20,33 |
| 4 | 24,33 | 350 | 0,9 | 28,67 | 22 | 16,33 | 190 | 0,8 | 18,50 |
| 5 | 19,00 | 240 | 0,8 | 28,00 | 23 | 13,17 | 130 | 0,7 | 14,50 |
| 6 | 18,67 | 290 | 1 | 27,00 | 24 | 12,83 | 150 | 0,8 | 16,33 |
| 7 | 17,00 | 260 | 1 | 25,00 | 25 | 15,67 | 120 | 0,5 | 32,67 |
| 8 | 18,83 | 210 | 0,7 | 33,67 | 26 | 13,00 | 90 | 0,5 | 17,67 |
| 9 | 17,00 | 210 | 0,8 | 20,33 | 27 | 14,50 | 110 | 0,5 | 22,00 |
| 10 | 17,67 | 220 | 0,8 | 30,33 | 28 | 14,67 | 130 | 0,6 | 25,00 |
| 11 | 17,33 | 260 | 1 | 21,33 | 29 | 16,17 | 140 | 0,6 | 21,33 |
| 12 | 19,50 | 250 | 0,8 | 25,33 | 30 | 19,33 | 180 | 0,6 | 24,00 |
| 13 | 19,33 | 290 | 1 | 25,33 | 31 | 21,00 | 230 | 0,7 | 29,33 |
| 14 | 20,00 | 310 | 1 | 28,33 | 32 | 18,17 | 190 | 0,7 | 30,00 |
| 15 | 15,33 | 200 | 0,9 | 20,67 | 33 | 18,67 | 210 | 0,7 | 35,33 |
| 16 | 19,67 | 250 | 0,8 | 33,00 | 34 | 19,00 | 240 | 0,8 | 31,67 |
| 17 | 21,00 | 300 | 0,9 | 26,00 | 35 | 16,67 | 150 | 0,6 | 30,00 |
| 18 | 18,67 | 290 | 1 | 31,33 | 36 | 13,00 | 130 | 0,7 | 33,67 |
| **Объем выборки n=36** | | | | | | | | | |
| **Среднее значение (mean)** | | | | | | 17,52 | 213,61 | 0,79 | 26,46 |
| **Среднеквадратичное отклонение (ско или σ)** | | | | | | 2,76 | 65,82 | 0,16 | 5,67 |
| **Вариация (Var)** | | | | | | 16% | 31% | 21% | 21% |

**Методика и результаты подспутниковых измерений.** Для верификации спутниковых определений высоты леса, анализа и последующего учета ошибок помимо высоты деревьев также были проведены полевые измерения диаметров стволов на высоте 1,2 м и суммы площадей сечений стволов в древостое S. Использованы следующие средства измерения:

1) GPS-навигатор Garmin – для определения местоположения (широты, долготы и высоты над уровнем моря);
2) цепной полнотометр Биттерлиха – для измерения суммы площадей сечений стволов в древостое, S кв.м./га;
3) высотомер-эклиметр ЭТ-1П и рулетка – для измерения высоты $H_{полев}$ (м) и расстояния до деревьев r (м), а также уклона участков $\alpha$ (в градусах);
4) линейка – для измерения диаметра деревьев (d, см).

На каждом тестовом участке произведено по три измерения каждой характеристики с последующим усреднением. Затем по стандартной таблице ЦНИИЛХА 1933 г. (Орлов М.М., 1931) для соснового леса данного региона на основе средней высоты и S определены полнота П, видовое число $H_f$ и запас насаждений. Результаты сведены в таблицу. В табл. 1 для примера представлены основные характеристики: высота (x1), запас древостоя (x2), полнота (x3) и диаметр деревьев (x4) первых n=36 измерений. В конце таблице для каждого параметра также представлены средние значения $\bar{x}$, среднеквадратичные отклонения (ско σ) и вариации (Var) по данной выборке (n=36):

$$\bar{x} = \frac{1}{n}\sum_i^n x_i\,;\ \ \sigma = \sqrt{\frac{1}{n} * \sum_i^n (x_i - \bar{x})^2}\,:\ Var = \frac{\sigma}{\bar{x}} * 100\%\,; \qquad (1)$$

**Методика и результаты спутниковых измерений высоты на основе разности интерферометрических фаз X- и L-диапазонов.** Для определения высоты лесного полога используется комбинация радиолокационных интерферометрических данных L-диапазона (25 см) ALOS PALSAR (интерферометрическая пара 20090728-20090912 на восходящей орбите) и X-

диапазона (3 см) TANDEM-X от 27 октября 2012 г. и 26 сентября 2016 г. На основе данных X- и L-диапазонов рассчитана разность интерферометрическая фаз с пространственным разрешением R = 5 м/пс или 0,000083° = 0,005′ = 0,3″ пс$^{-1}$ до конвертации снимков в картографическую систему UTM WGS84. Предварительно была выполнена частичная коррекция рельефа местности с помощью данных SRTM (с разрешением 50 м). Разность интерферометрических фаз в двух выбранных диапазонах соответствует отражению радиолокационных сигналов от нижней подстилающей (L-диапазон) и верхней огибающей (X-диапазон) поверхностей лесной среды или толщине лесной среды. Данные представлены в комплексной форме, которые были конвертированы и разложены на 5 составляющих: 1) амплитуда (Pw); 2) амплитуда в логарифмическом виде (Mg); 3) реальная составляющая (Re); 4) мнимая составляющая (Im); 5) фазовая составляющая (ΔPh).

Для типовых объектов (водная поверхность рек Селенга и Уда, участки городской и сельской застройки, участки без растительности, участки леса и т.п.) были определены статистические и текстурные характеристики растровых изображений Pw, Mg, Re, Im, ΔPh, такие как среднее значение, диапазон значений DN, вариация, коэффициент асимметрии (skew). Для них были установлены пороговые значения и комбинации условий, построено дерево решений и выполнена классификация. На рис. 2 приведен фрагмент результата классификации для полигона Верхняя Березовка. В частности, класс 9 соответствует участкам лесной растительности. Условия кластеризация лесных массивов соответствуют средним значениям яркости, незначительными вариациями DN всех показателей, отрицательными значениями реальной части: {{(–2.35<Pw<–2.00) OR (0.14<Mg<0.40)} AND (Re<0)}.

Последняя компонента разложения ΔPh дает возможность оценить толщину лесного покрова $\Delta H_{int}$. Для оценки высоты леса $\Delta H_{int}$ производится инверсия интерферометрической разности фаз и преобразование с учетом того, что разности фаз $\Delta\varphi=2\pi$ соответствует разность высот равная для тандемных съемок 2012 г. и 2016 г. соответственно $\Delta H_{TDX\text{-}2012}$ =18,667 м и $\Delta H_{TDX\text{-}2016}$ =26,676 м:

$$\Delta H_{int} = \frac{\pi - Ph}{2\pi} * \Delta H_{TDX} \text{ м.} \qquad (2)$$

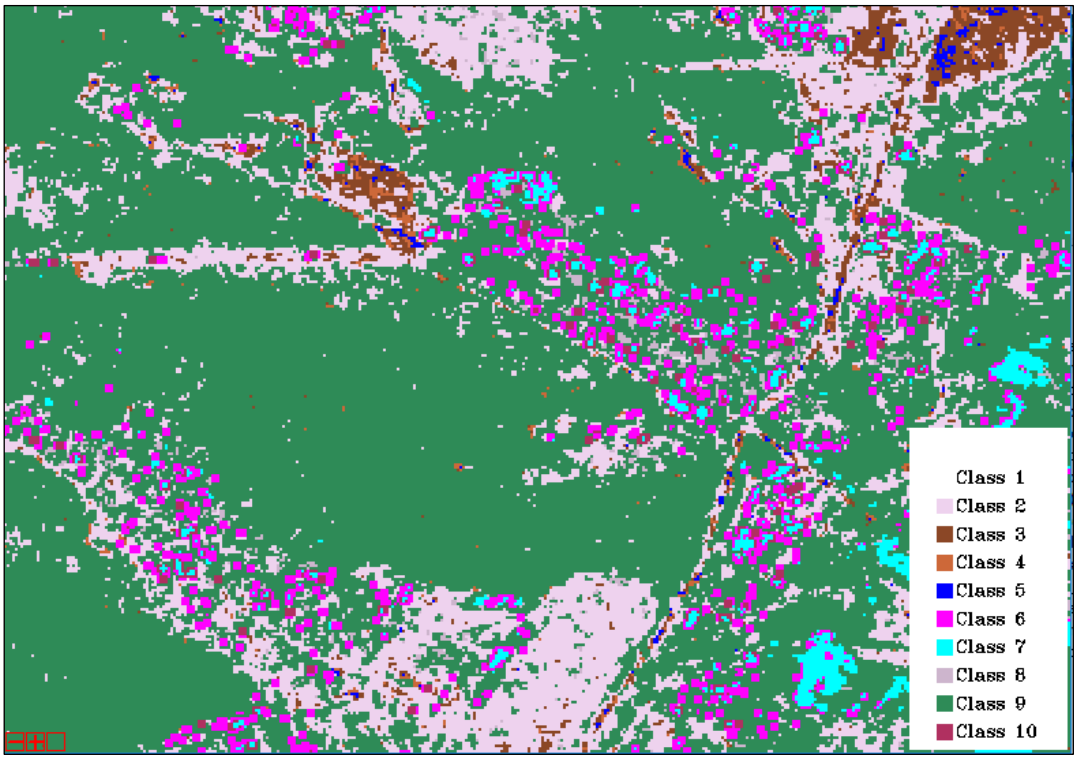
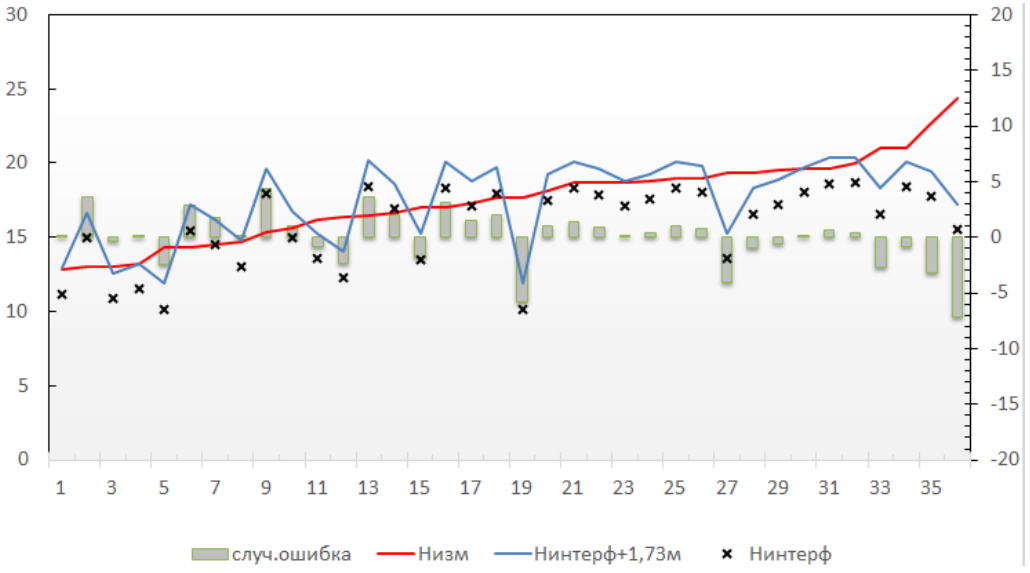
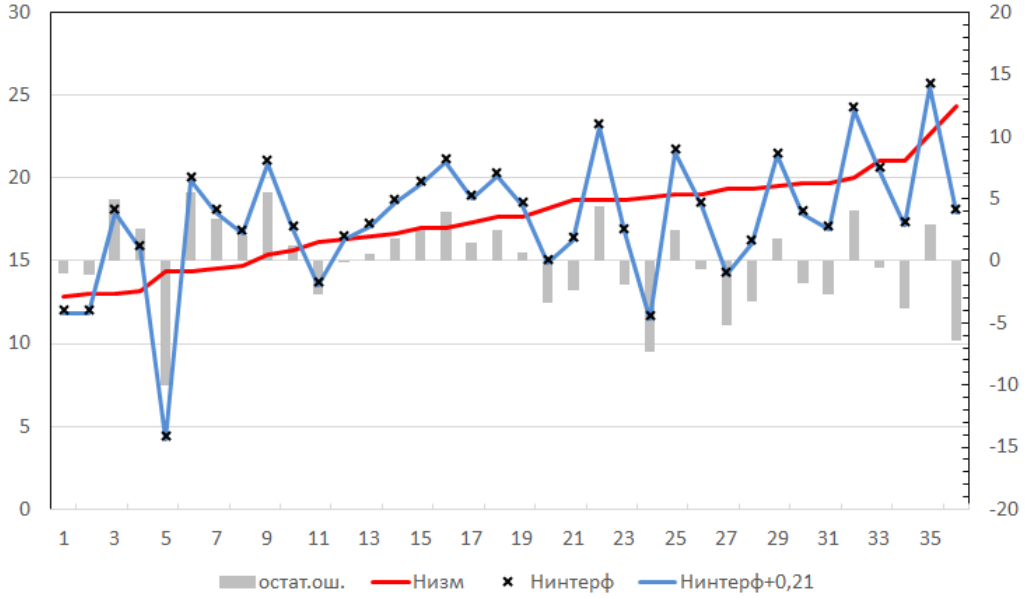



Рис. 2. Результат классификации с применением дерева решений.
Классы: фон – 1; водные объекты – 5; участки открытой почвы – 2, 3, 4, 8; лес – 9; сельская застройка – 6; городская застройка – 10; горы (передние склоны) – 7.




*а*




*б*

Рис 3. Сравнение подспутниковых и спутниковых измерений высоты леса на полигоне Верхняя Березовка по данным TDX-2012 (а) и TDX-2016 г. (б).

При этом диапазон значений ΔHint $\in$[0; $\Delta H_{TDX}$] м. Очевидно, при достижении максимального значения $\Delta H_{TDX}$, происходит обнуление высоты, что необходимо учитывать при наличии участков леса с более высокими деревьями. На полигоне Верхняя Березовка высота деревьев менее 25 м (по выборке n=84).

На рис. 3 демонстрируется сравнение подспутниковых (Низм – линии красного цвета) и спутниковых измерений высоты леса в 36 пунктах (Нинтерф – крестики) по данным TDX-2012 (а) и TDX-2016 (б). В каждом пункте производился расчет усредненного диапазона значений высот (преобразованной разности интерферометрических фаз) на участках 30х30 или 30х40 м. При усреднении брались только пиксели, соответствующие верхней границе лесного полога с помощью маски {skew($\Delta H_{int}$ )<0}. Без применения маски ошибки более значительны, в частности на относительно ровных участках занижение составляет до 5-7 м. При построении графиков данные ранжированы в соответствии с ростом подспутниковых высот.

Ошибки определения высоты леса вызваны систематическими и случайными причинами:

$$\varepsilon = \varepsilon_{сист} + \varepsilon_{случ} \quad (3)$$

Табл. 2. Корреляционная матрица и уравнения линейной парной регрессии ошибок спутниковых определений высоты леса.

| R | Коэффициенты линейной корреляции | | | | | $R^2$ | Уравнение регрессии |
|---|---|---|---|---|---|---|---|
| | x1 | x2 | x3 | x4 | y | | |
| **x1** | 1 | 0,7962 | 0,3012 | 0,5696 | -0,4753 | 0.2259 | Y=7.5404-0.4304*x1 |
| **x2** | 0,7962 | 1 | 0,8067 | 0,3371 | -0,2419 | 0,0585 | Y=1.9622-0.0092*x2 |
| **x3** | 0,3012 | 0,8067 | 1 | 0,0295 | 0,0448 | 0.002 | Y=0.6893*x3-0.5419 |
| **x4** | 0,5696 | 0,3371 | -0,0295 | 1 | -0,1356 | 0.0184 | Y=1.5807-0.0597*x4 |

Систематические ошибки обусловлены технологией производства интерферометрических и подспутниковых измерений высоты и определяются как средние значения по выборке ошибок $\varepsilon_i$ или как разность средних спутниковых и подспутниковых значений высот:

$$\varepsilon_{\text{сист}} = \frac{1}{n}\sum \varepsilon_i = \overline{H_{int}} - \overline{H_{\text{изм}}} \quad (4)$$

Систематические ошибки $\varepsilon_{\text{сист}}$ для TDX-2012 и TDX-2016 составили соответственно –1,73 и +0,21 м. Скорректированные с учетом систематических ошибок спутниковые значения высоты леса приведены на рисунках 3 линиями синего цвета. Разница в значениях $\varepsilon_{\text{сист}}$ вызвана отчасти четырехлетним интервалом между тандемными съемками 2012 и 2016 г., отчасти разницей $\Delta H_{TDX}$, обусловленной различиями базового расстояния. Последний фактор проявляется на рис. 3. На рис.3а для участков леса с высотой более 18 м наблюдается систематическое занижение спутниковых измерений, в то время как на рис.3б этот эффект «потолка» отсутствует.

Остаточные случайные ошибки обусловлены локальными различиями как характеристик лесной среды, так и физических и геометрических условий съемки. Они представлены на графиках прямоугольниками серого цвета. Справа приведена отдельная шкала значений ошибок определения высоты. СКО($\varepsilon_{\text{случ}}$), значения которых характеризуют абсолютную точность спутниковых измерений высоты, равны соответственно 2,50 и 3,70 м. Относительная точность соответственно равна 14 и 21 %, при нормативной точности определения высоты при дешифровочном способе таксации 15 % [5].

Проведен корреляционно-регрессионный анализ зависимости ошибок (y) от четырех факторов: высоты деревьев (полевые измерения) – x1; запаса древостоя – x2; полноты – x3; диаметра деревьев – x4. Были составлены линейные уравнения парной и множественной регрессии, вычислены коэффициенты корреляции и детерминации (см. табл. 2).

С учетом эмпирических поправок к высоте наблюдается улучшение точности спутниковых определений высоты: за счет фактора x1 – на 30 см (12 %), x2 – на 7 см (3 %), x3 – на 0,3 см (0,1 %), x4 – на 2 см (0,9 %)

**Заключение.** В результате исследований разработаны основы методики определения и коррекции высоты лесного полога на основе разности интерферометрических фаз X- и L-диапазонов (TanDEM-TerraSAR-X и ALOS PALSAR). В частности, по подспутниковому полигону Верхняя Березовка установлено, что:

- предварительная классификация и формирование маски для пикселей соответствующих верхней границе леса значительно уменьшают ошибки определения усредненных высот;
- определены систематические ошибки (–1,73 и +0,21 м), абсолютные (2,5 и 3,7 м) и относительные (14 и 21 %) показатели точности определения высот по данным TanDEM-X 2012 и 2016 г; различия обусловлены временной разницей и разницей размеров базовой линии тандемных съемок;
- корреляционно-регрессионный анализ показал слабую возможность коррекции спутниковых определений высоты за счет эмпирических линейных уравнений зависимости случайной ошибки от таких факторов как высота (x1), запас древостоя (x2), полнота (x3) и диаметр стволов (x4).

Полученные результаты будут использованы для продолжения исследований, в частности предполагается изучение влияния особенностей рельефа посредством привлечения дополнительных спутниковых данных на восходящих и нисходящих орбитах, получение эмпирических поправок для коррекции высот и др.

*Работа выполнена в рамках госзадания и при частичной финансовой поддержке гранта РФФИ № 18-47-030001 р_а.*

ЛИТЕРАТУРА

[1] Лесной кодекс Российской Федерации от 04.12.2006 N 200-ФЗ (с изменениями и дополнениями, вступившими в силу с 01.01.2019).

[2] *Мартынов А.Н., Мельников Е.С., Ковязин В.Ф., Аникин А.С.* Основы лесного хозяйства и таксация леса. СПб: Лань, 2012. 3-е изд. 432 с.

[3] *Чимитдоржиев Т.Н., Дмитриев А.В., Кирбижекова И.И., Шерхоева А.А., Балтухаев А.К., Дагуров П.Н.* Дистанционные оптико-микроволновые измерения параметров леса: современное состояние исследований и экспериментальная оценка возможностей // Современные проблемы дистанционного зондирования Земли из космоса. 2018. Т. 15. № 4. С. 9-26.

[4] Справочно-информационный портал «Погода и климат». **http://www.pogodaiklimat.ru/history/30823.htm** (дата обращения 22.06.2019).

[5] Технический регламент по проведению контроля за качеством лесоучетных работ ФГУП «Рослесинфорг». Приложение №1 к приказу от 13 марта 2013 г.

# ОЦЕНКА ОБЪЕМА СЖИГАЕМОГО ПРИРОДНОГО ГАЗА В ФАКЕЛЬНЫХ УСТАНОВКАХ НА ТЕРРИТОРИИ ЗАПАДНОЙ СИБИРИ ПО СПУТНИКОВЫМ ДАННЫМ В ВИДИМОМ И БЛИЖНЕМ ИК-ДИАПАЗОНАХ В НОЧНОЕ ВРЕМЯ

*Лагутин А.А., Мордвин Е.Ю., Волков Н.В., Тучина Н.В.*
Алтайский государственный университет, Барнаул

Представлены результаты оценки объема сжигаемого природного газа в факельных установках на территории Западной Сибири для ряда периодов 2013-2019 годов. Оценки получены в рамках технологии, включающей восстановление характеристик факельных установок по спутниковым данным в видимом и ближнем ИК-диапазонах в ночное время, а также регрессионное соотношение между мощностью источника и объемом сжигаемого газа. Установлено, что объем сжигаемого газа в факельных установках на территории Западной Сибири в этот период менялся от 15,7 млрд. м$^3$ в 2013 г. до 14.8 млрд. м$^3$ в 2019 г.
*Ключевые слова: Западная Сибирь, факельные установки, природный газ, объем сжигаемого газа, Suomi-NPP, VIIRS*

**Введение.** Добыча нефти на каждом месторождении сопровождается выделением попутного газа. Часть этого газа сжигается в факельных установках (ФУ) в местах добычи нефти, а также на других этапах ее переработки. По оценкам Национальной администрации по атмосфере и океану США, основанных на спутниковых наблюдениях, в 2015 г. функционировало 13605 факельных установок в 88 странах мира [1]. Лидерами по объему сжигаемого газа являются Россия, Ирак, Иран, Венесуэла, США (см. рис. 2 статьи [1]). По оценкам [1] в ФУ ежегодно сжигается около 140 миллиардов кубометров газа, что приводит к выбросу порядка 300 миллионов тонн $CO_2$.

Следует отметить, что сжигание попутного газа приводит к выбросам в атмосферу, помимо $CO_2$ и метана, окиси азота, сернистого ангидрида, сажи и других токсичных компонентов, а также к тепловому воздействию на подстилающую поверхность в районе расположения ФУ. Эти выбросы оказывают влияние на климат, окружающую среду и состояние здоровья людей как на региональном уровне, так и в глобальном масштабе.

В силу отсутствия регулярных и полных данных по расположению факельных установок и режиму их функционирования для оценки воздействия ФУ на биосферу Земли необходимо проведение мониторинга характеристик этих установок, а также системы «атмосфера – подстилающая поверхность» в зоне их расположения. Ключевой характеристикой этого мониторинга является объем сжигаемого попутного газа.

Целью данной работы является оценка объема сжигаемого попутного газа в факельных установках на территории Западной Сибири по спутниковым данным в видимом и ближнем ИК-диапазонах в ночное время.

**Информационная база.** Информационной основой проводимого в работе исследования являются данные радиометра VIIRS (Visible Infrared Imaging Radiometer Suite) [2] спутника Suomi-NPP [3].

Радиометр VIIRS имеет 22 спектральных канала, регистрирующих уходящее из атмосферы электромагнитное излучение в диапазоне 0.41-12.5 мкм. В отличие от большинства приборов этого класса, 4 спектральных каналов VIIRS регистрируют поток уходящего в космическое пространство излучения в видимом и ближнем ИК-диапазонах в ночное время суток. Характеристики этих каналов показаны в табл. 1.

Схема обработки принимаемой в режиме прямого вещания наземными станциями Алтайского госуниверситета спутниковой информации VIIRS/Suomi-NPP состоит из следующих основных этапов:

- распаковка спутниковых данных;
- проведение геопривязки данных и их калибровка;
- построение маски облачности;
- поиск термальных аномалий;

– определение измеряемых величин (продукты уровня 2).

**Восстановление параметров ФУ.** Физической основой метода восстановления температуры высокотемпературных термальной аномалии, которые включают ФУ, является утверждение о том, что в безлунную ночь весь поток электромагнитной энергии, уходящей в космическое пространство в видимом и ближнем ИК-диапазоне, в основном является собственным излучением этого источника.

Регистрация излучения в каналах VIIRS в видимом и ближнем ИК-диапазонах позволяет восстановить распределение Планка для анализируемого высокотемпературного источника.

Понятно, что пиксел, включающий аномалию, является серым телом. В этом случае регистрируемая в этом пикселе радиометром VIIRS интенсивность равна произведению функции Планка при температуре источника $T$ и масштабирующего множителя $\varepsilon$, который определяется как отношение наблюдаемой спутником интенсивности к интенсивности пиксела при этой температуре $T$ при условии, что весь пиксел имеет эту температуру [4]. Площадь источника $S$ находится путем умножения площади пиксела на $\varepsilon$.

В случае сжигания чистого метана, температура горения которого составляет 2223 К, из закона смещения Вина следует, что длина волны, на которую приходится максимум интенсивности излучения, равна ~1.3 мкм, т.е. между каналами M8 (1.2 мкм) и M10 (1.6 мкм) (см. рис. 1). Этот факт позволяет использовать показания канала M10 в качестве пороговых значений при идентификации термальных аномалий.

Табл. 1. Характеристики каналов VIIRS, регистрируюших уходящее излучение в ночное время суток в видимом и ближнем ИК-диапазонах.

| **Канал** | **Диапазон длин волн (мкм)** | **Максимум функции чувствительности (мкм)** | **Полоса пропускания (мкм)** | **Размер пикселя (км×км)** |
|---|---|---|---|---|
| **DNB** | 0.500 – 0.900 | 0.700 | 0.400 | 0.742×0.742 |
| **M7** | 0.843– 0.881 | 0.862 | 0.042 | 0.742×0.776 |
| **M8** | 1.225 – 1.252 | 1.240 | 0.030 | 0.742×0.776 |
| **M10** | 1.571 – 1.631 | 1.601 | 0.060 | 0.742×0.776 |

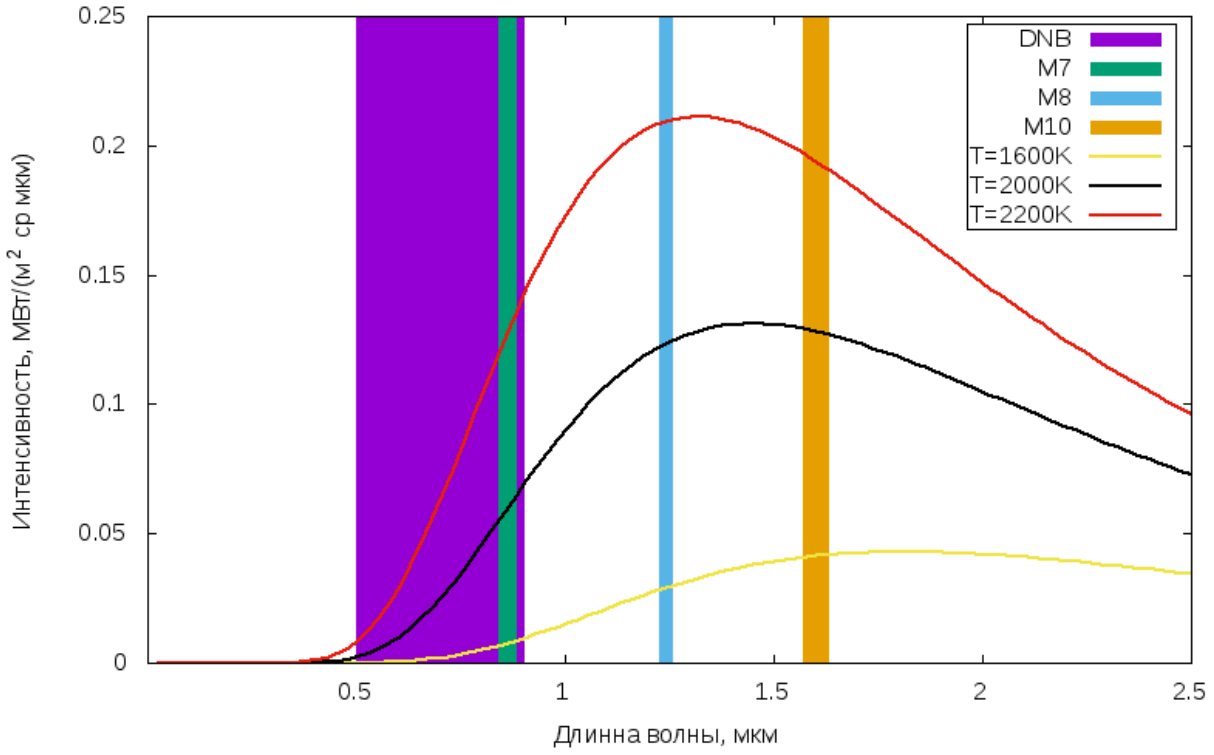



Рис. 1. Распределение Планка для различных температур источника и ширина каналов VIIRS, регистрирующих уходящее излучение в ночное время суток в видимом и ближнем ИК-диапазонах.

В используемом в работе алгоритме определения очагов термальных аномалий считается, что в конкретном пикселе VIIRS идентифицирована такая аномалия, если показания канала M10 для этого пиксела превышает пороговое значение $sr + 6std$, где $sr$ – среднее значение шума в канале M10, а $std$ – его стандартное отклонение. Для пиксела сцены, идентифицированного как термальная аномалия, восстанавливалась температура $T$ источника и коэффициент масштабирования $\varepsilon$. Для решения этой задачи с использованием алгоритма Левенберга-Марквардта производилась оптимизация целевой функции переменных $T$ и $\varepsilon$

$$\min_{T,\varepsilon} F(T,\varepsilon) = \sum_{j=\{DNB,7,8,10,12,13\}} \left[L(\lambda_j) - \varepsilon B(T,\lambda_j)\right]^2. \quad (1)$$

В (1) суммирование квадратов невязок производится для показаний каналов DNB, M7, M8, M10, M12 и M13 радиометра VIIRS. В качестве начальных значений температуры источника и коэффициента масштабирования использовались $T_0 = 600$ K и $\varepsilon_0 = 0.01$, соответственно.

Площадь источника находилась с использованием выражения $S = \Delta A \Delta D \varepsilon$ [5], в котором $\Delta A$ есть размер пиксела вдоль направления сканирования, а $\Delta D$ – вдоль движения спутника (см. выражения (4) и (5) работы [5]).

**Высокотемпературные источники излучения на территории Западной Сибири.** Предложенный метод был реализован в центре космического мониторинга АлтГУ для поиска высокотемпературных источников излучения с $T > 1500$ K на территории Западной Сибири в период с 10 апреля по 10 мая для 2013, 2015, 2017 и 2019 годов. Очаги термальных аномалий, выявленные для серии наблюдений только один раз, исключались из дальнейшего анализа.

В результате работы для каждого из исследуемых периодов было обнаружено ~300 аномалий. На рис. 2 в качестве примера показано пространственное распределение источников, найденных в 2013 году.

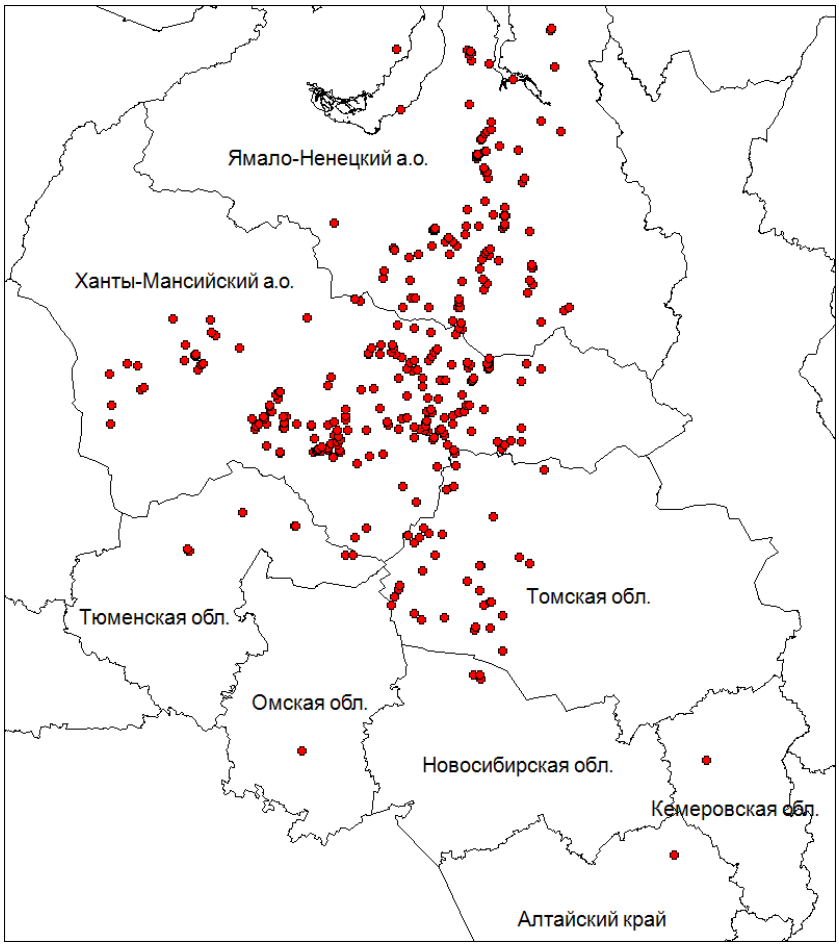



Рис. 2. Пространственное распределение высокотемпературных термальных аномалий в Западной Сибири, найденных в 2013 г.

Табл. 2. Результаты оценки объема сжигаемого газа в ФУ Западной Сибири.

| Год | Количество найденных источников | $V$, млрд. м$^3$ |
|---|---|---|
| 2013 | 336 | 15,7 |
| 2015 | 315 | 14,1 |
| 2017 | 291 | 14,9 |
| 2019 | 311 | 14,8 |

Исследование вариаций температуры найденных источников показало, что за все периоды, данные для которых обрабатывались в работе, значения $T$ попадают в диапазон 1500-2300 К. Этот результат свидетельствует о том, что эти источники имеют антропогенное происхождение, а диапазон полученных температур указывает на горение газовой смеси на основе легких углеводородов.

Для визуальной идентификации/классификации источников использовались открытые геопространственные данные геоинформационной системы Google Earth [6], Яндекс.Карты [7], а также данные спутниковой съемки высокого пространственного разрешения Landsat-8 [8]. Совокупный анализ этих данных показал, что большинство установленных источников являются факельными установками для сжигания природного газа.

**Оценка объема сжигаемого газа.** В основе оценки объема сжигаемого газа лежит регрессионное соотношение между измеряемой спутниковым прибором мощностью источника и объемом сжигаемого газа, установленное в [4] по данным о выбросах ФУ 47 стран. Это соотношение имеет вид

$$V = 0.0274 \cdot W, \quad (2)$$

где $V$ – объем сжигаемого газа (в млрд. м$^3$), а $W$ – мощность источника (в МВт). Мощность источника находилась с использованием закона Стефана-Больцмана

$$W = \sigma T^4 S. \quad (3)$$

В работе [4] было показано, что существует нелинейная зависимость между $W$ и объемом сжигаемого факельного газа $V$. Учет этой нелинейности возможен введением показателя степени площади источника при оценке мощности источника $W$ [4]

$$W = \sigma T^4 S^{0.7}. \quad (4)$$

С использованием выражений (2) и (4) и результатов идентификации очагов термальных аномалий была выполнена оценка объема сжигаемого газа для каждой ФУ в Западной Сибири. Суммарные оценки объемов сжигаемого газа в ФУ региона представлены в табл. 2.

Полученные результаты показывают, что максимальный объем сжигаемого попутного газа в ФУ Западной Сибири по данным VIIRS, равный 15,7 млрд. м$^3$, наблюдается в 2013 году.

**Заключение.** Главной целью данной работы являлась оценка объема сжигаемого попутного газа в факельных установках на территории Западной Сибири по спутниковым данным в видимом и ближнем ИК-диапазонах в ночное время. В результате решения задач, направленных на достижение заявленной цели, получены следующие результаты.

1. Предложен и верифицирован метод идентификации очагов термальных аномалий с температурой $T > 1500$ K в условиях безоблачной атмосферы с использованием ночных данных радиометра VIIRS спутника Suomi-NPP.
2. Созданный алгоритм был использован для поиска термальных аномалий на территории Западной Сибири с 10 апреля по 10 мая для 2013, 2015, 2017 и 2019 годов. Для каждого из периодов найдено порядка 300 источников, идентифицированных как факельные установки сжигания природного газа.
3. С использованием регрессионного соотношения между объемом сжигаемого газа и мощностью источника, получены оценки объемов сжигаемого попутного газа в факельных установках на территории Западной Сибири. Установлено, что в исследуемый период эти объемы изменялись от 15,7 млрд. м$^3$ в 2013 г. до 14.8 млрд. м$^3$ в 2019 г.

ЛИТЕРАТУРА


[1] *Elvidge C.D., Bazilian M.D., Zhizhin M. et al.* The potential role of natural gas flaring in meeting greenhouse gas mitigation targets // Energy Strategy Reviews. 2018. Vol. 20. P. 156-162.

[2] *Hillger D., Kopp T, Lee T. et al.* First-light imagery from SUOMI NPP VIIRS // BAMS. 2013. Vol. 94. P. 1019-1029.

[3] *Powell A.M., Jr., Weng F.* Introduction to special section on Suomi National Polar-Orbiting Partnership satellite calibration, validation, and applications // J. Geophys. Res. Atmos. 2013. Vol. 118. P. 12216-12217.

[4] *Elvidge C.D., Zhizhin M., Baugh K. et al.* Methods for global survey of natural gas flaring Visible Infrared Imaging Radiometer Suite data // Energies. 2016. Vol. 9. 14 p. DOI: 10.3390/en9010014.

[5] *Elvidge C.D., Zhizhin M., Hsu F-C. et al.* VIIRS Nightfire: Satellite pyrometry at night // Remote Sens. 2013. Vol. 5. P. 4423-4449.

[6] Google Earth. **https://www.google.com/earth** (дата обращения 15.06.2019).

[7] Яндекс.Карты. **https://yandex.ru/maps** (дата обращения 15.06.2019).

[8] USGS Earth Explorer. **https://earthexplorer.usgs.gov** (дата обращения 15.06.2019).

# МОНИТОРИНГ, ОЦЕНКА РИСКОВ И БЕЗОПАСНОСТЬ ОБЪЕКТОВ ТЕХНОСФЕРЫ И УРБАНИЗИРОВАННЫХ ТЕРРИТОРИЙ

*Москвичев В.В., Чернякова Н.А.*
Красноярский филиал Института вычислительных технологий СО РАН – Специальное конструкторско-технологическое бюро «Наука», Красноярск

Рассмотрены общие постановки региональных проблем оценки природно-техногенной безопасности на основе риск-ориентированного подхода. Приведены базовые соотношения для расчетов территориальных рисков, получены фактические значения техногенных рисков для критически важных объектов техносферы. Исследования территориальных рисков выполнены для субъектов Сибирского федерального округа и муниципальных образований Красноярского края.
*Ключевые слова: риск-анализ, чрезвычайные ситуации, территориальный риск, мониторинг, объекты техносферы, модель риска, защищенность.*

**Введение.** При стратегическом планировании и прогнозном анализе социально-экономического развития территориальных образований (субъект РФ, регион, муниципальное образование, промышленная агломерация и т.д.) в качестве основных показателей устойчивого развития принимаются количественные оценки техногенных, природных, экологических, технологических, социальных и других видов рисков. Территориальное образование рассматривается как взаимосвязанная социально-природно-техногенная система (С-П-Т система) [1-3], включающая элементы техносферы (стратегически и критически важные объекты, потенциально опасные объекты и т.д.), экосферы (атмосфера, геосфера, биосфера, гидросфера), социосферы (личности, коллективы, сообщества). Устойчивое функционирование С-П-Т систем определяется эффективностью мониторинга состояния объектов техносферы и экосферы, организацией системы управления антропогенными, экологическими, природными и территориальными рисками. Мониторинг природно-техногенной безопасности выступает как фактор стабилизации кризисных явлений в социально-экономическом развитии, обеспечивающий сохранность и функционирование основных производственных фондов и защиту населения и территорий от чрезвычайных ситуаций (ЧС) природного и антропогенного характера. Дестабилизация устойчивости С-П-Т систем происходит в результате игнорирования требований концепции устойчивого развития и обострения дилеммы научно-технического прогресса – высокие темпы развития с одной стороны, и возникновение новых угроз человеку, обществу, природной среде со стороны объектов техносферы.

Данные постановки проблемы природно-техногенной безопасности обуславливают необходимость оценки уровня природных и антропогенных рисков как основы экономических механизмов управления территориальными образованиями. Следует отметить, что снижение рисков ЧС обеспечивает более устойчивое функционирование экономического потенциала и повышает конкурентные (инвестиционные) преимущества соответствующей территории.

**Мониторинг и оценка рисков.** Основные задачи в области оценки территориальных рисков:

1. Разработка моделей и технологий оценки состояния, прогнозирования и управления С-П-Т системами и территориальными образованиями на основе данных мониторинга и с учетом возможных рисков развития и возникновения ЧС природного и техногенного характера, определяющих безопасность и устойчивое развитие территорий.
2. Определение базовых и нормативных уровней рисков, характеризующих допустимое воздействие на элементы С-П-Т систем с учетом региональных социо-техно-экосферных особенностей и специфики антропогенного воздействия.
3. Ранжирование территорий по степени риска с использованием ГИС-технологий и методов пространственного интеллектуального анализа; формирование программ и разработка рекомендаций, нацеленных на снижение уровня рисков и повышение эффективности управления территориальными образованиями.

В настоящее время существующие системы территориального мониторинга охватывают федеральный и региональный уровни, методы и технологии которых регламентируются соответствующими нормативными документами (ФЗ, ПП РФ, ГОСТ и др.):

1. Мониторинг техногенных ЧС.
2. Мониторинг природных ЧС (гидрология, метеорология, сейсмология, природные пожары).
3. Государственный экологический мониторинг (состояние окружающей среды, биологические ресурсы, недра, лесопатологический, радиационный и др.).
4. Биосферный мониторинг.
5. Аэрокосмический мониторинг.
6. Социально-гигиенический.
7. Производственный экологический.

Каждый компонент указанных систем мониторинга представлен в регионе соответствующими структурами, способствующими решению задач территориального управления и прогнозных оценок состояния С-П-Т систем. Следует отметить существование в этой области значительного числа проблемных вопросов организационно-управленческого и научно-методического характера, при этом сети наблюдения и мониторинга получают огромные массивы информации, часто избыточной, но недостаточной для оценки стратегических рисков регионального развития. Для решения ряда научно-методических проблем разработан проект нормативного документа «Методические рекомендации по оценке территориальных рисков муниципальных образований», позволяющего дать комплексную оценку рисков развития С-П-Т систем по двум группам:

Группа 1. Риски чрезвычайных ситуаций (аварийные):
- индивидуальный риск;
- материальный риск;
- коллективный риск.

Группа 2. Медико-экологические риски (перманентные):
- коллективный риск;
- индивидуальный канцерогенный риск;
- популяционный канцерогенный риск;
- относительный риск;
- неканцерогенный риск;
- относительный риск смертности.

Основные соотношения для различных моделей рисков приведены в таблице.

Исходной информацией для расчетов рисков являются:
- данные государственных докладов «О состоянии защиты населения и территорий Российской Федерации от чрезвычайных ситуаций природного и техногенного характера», «О состоянии и об охране окружающей среды Российской Федерации»;
- статистические данные региональных систем мониторинга;
- нормативные документы федерального и регионального уровня.

Количественная оценка рисков развития С-П-Т систем требует сбора, обработки и анализа огромных массивов разнородных данных. Значительное число параметров, характеризующих состояние таких систем, их зависимость от многочисленных факторов, сложность комплексного анализа предопределили разработку информационной системы территориального управления рисками и безопасностью (ИСТУ РБ) для отдельных промышленных агломераций, субъектов и регионов [2]. Блок-схема ИСТУ РБ представлена на рис. 1. Создание данной системы позволяет:

- обеспечить информационную поддержку территориального управления, научно-технологической базы мониторинга источников опасностей и чрезвычайных ситуаций, принятия решений по снижению рисков и прогнозу развития территориальных образований;

Основные соотношения для различных моделей рисков.

| № | Определение риска | Соотношение для расчета риска | Пояснения, параметры | Источник |
|---|---|---|---|---|
| **1** | Уровень индивидуального пожизненного канцерогенного риска при воздействии вещества в дозе LADD | $CR = LADD \cdot SF$ | CR – индивидуальный пожизненный канцерогенный риск;<br>LADD – среднесуточная доза в течение жизни, мг/(кг х день);<br>SF – фактор наклона [мг/(кг х день)][1] | МосМР 2.1.9.001-03;<br>МР 2.1.4.0032-11.2.1.4;<br>Р 2.1.10.1920-04 |
| **2** | Популяционный канцерогенный риск характеризует дополнительное число случаев злокачественных образований при воздействии в течение всей жизни | $PCR = LADD \cdot SF \cdot POP$ | PCR – популяционный канцерогенный риск;<br>POP – численность исследуемой популяции;<br>70 лет – средняя продолжительность жизни | |
| **3** | Риск развития неканцерогенных эффектов (оценивается по коэффициенту опасности) | $HQ = {AD}/{RfD}$<br>или<br>$HQ = {AC}/{RfC}$ | HQ – коэффициент опасности;<br>AD – средняя суточная доза за период воздействия, мг/кг;<br>AC – средняя концентрация, мг/ м$^3$;<br>RfD – референтная (безопасная) доза, мг/кг;<br>RfC – референтная (безопасная) концентрация, мг/ м$^3$ | |
| **4** | Модель оценки неканцерогенного риска беспороговым методом | $Risk = 1 - \exp\frac{\ln(0{,}84)}{(ПДК \cdot К_з) \cdot C}$ | Risk – неканцерогенный беспороговый риск от химических веществ для питьевой воды, поступивших пероральным путем;<br>$К_з$ – коэффициент запаса;<br>C – средняя ежедневная концентрация вещества, поступающего в организм человека с питьевой водой в течение его жизни;<br>ПДК – предельно-допустимые концентрации веществ для питьевой воды | МР 2.1.4.0032-11.2.1.4 |
| **5** | Социальный риск (или риск поражения группы людей) – зависимость частоты возникновения сценариев аварий F, в которых пострадало на определенном уровне не менее N человек, от этого числа N | $F(x) = \sum_{i=1}^{I(x)} Q_i^x$<br>$F(N_j) = F(N_j) \cdot \frac{N_j}{[N_j]}$ | Социальный риск рекомендуется представлять в виде графика ступенчатой функции F(x). $Q_i^x$ –ожидаемые частоты реализаций аварийных ситуаций $C_i$, при которых гибнет не менее х человек; N(x) – число сценариев $C_i$, при которых гибнет не менее х человек;<br>$F(N_j)$ – сумма частот сценариев с ожидаемым числом погибших не менее $N_j$ | Приказ Ростехнадзора от 11.04.2016 N 144 «Об утверждении Руководства по безопасности «Методические основы по проведению анализа опасностей и оценки риска аварий на опасных производственных объектах» |
| **6** | Модель оценки индивидуального риска ЧС | $R = \frac{N_п}{N_н}$ | *R* - индивидуальный риск;<br>$N_п$ – среднее количество погибших в год за последние 5 лет при определённом виде ЧС и происшествии на заданной территории;<br>$N_н$ – количество населения, проживающего на данной территории | Методические рекомендации МЧС РФ №2-4-71-40 |

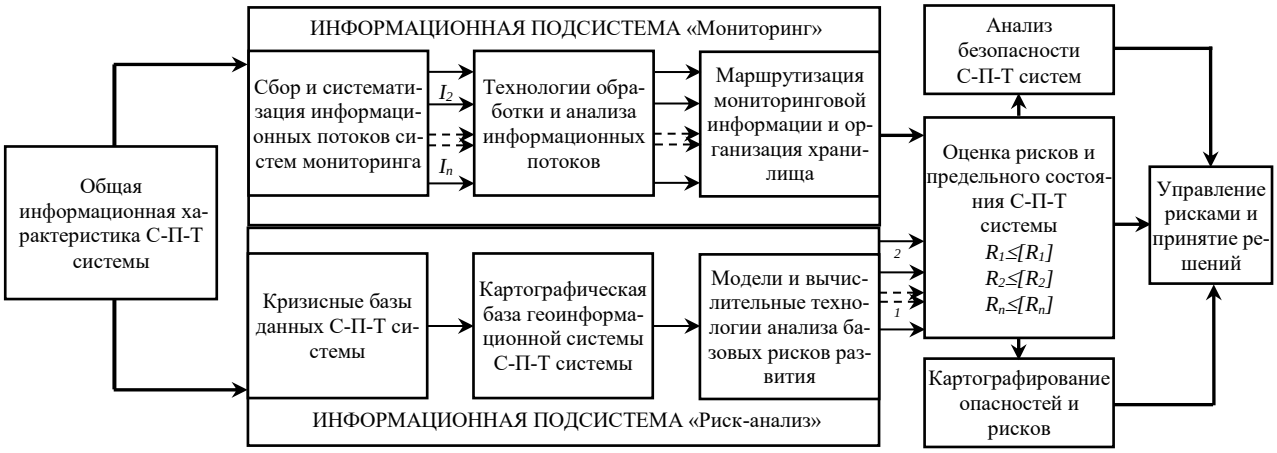



Рис. 1. Блок-схема информационной системы территориального управления рисками и безопасностью (ИСТУ РБ «Регион»).

- исследовать особенности территориального управления, состояния и развития С-П-Т систем конкретных промышленных регионов страны и составляющих их элементов;
- проводить комплексную оценку безопасности и экологического состояния по данным мониторинга различных природно-техногенных систем, критических и стратегических объектов;
- на единой методологической основе осуществлять сбор, хранение, обработку и анализ неоднородной пространственной информации, характеризующей состояние С-П-Т систем, включая новейшие методы интеллектуальной обработки пространственных данных с целью получения новых знаний о процессах, происходящих в С-П-Т системах;
- обеспечить внедрение нового поколения прикладных информационно-аналитических систем на основе технологий BigData и программных комплексов с использованием технологий облачных сервисов как универсальных элементов мониторинговых систем.

**Основные концептуальные положения и примеры риск-анализа объектов техносферы.**

1. Результатом аварийных и катастрофических ситуаций объектов техносферы является поражение персонала, окружающей среды, объектов и систем с образованием прямых и косвенных ущербов.
2. Причинно-следственные связи в процессе эксплуатации, возникновение предельных состояний и аварийных ситуаций (АС), сценарное развитие АС и реализации катастрофического разрушения носят вероятностный характер.
3. Расчетный риск-анализ технических систем (ТС) проводится для следующих основных процедур и этапов:
   - проектная экспертиза технических решений по результатам моделирования АС;
   - оперативная диагностика технического состояния в процессе эксплуатации;
   - экспертиза последствий аварийных ситуаций;
   - оценка остаточного ресурса для предельных состояний ТС на базе критериев механики деформирования и разрушения.
4. Расчетный риск-анализ базируется на применении вероятностных моделей прочности, ресурса и безопасности: структуры и свойств конструкционных материалов, нагрузок и воздействий, технологической и эксплуатационной дефектности, накопления и развития повреждений, моделей анализа предельных и напряженно-деформированных состояний, причинно-следственного комплекса отказов и сценариев АС, остаточной прочности, ресурса и живучести, риск-анализа и безопасности ТС.

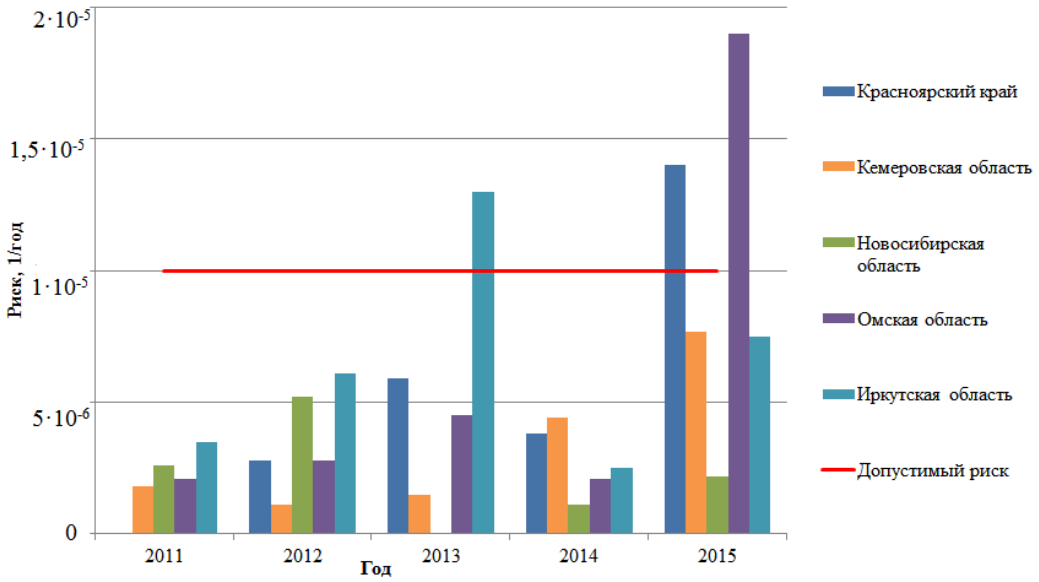



Рис. 2. Индивидуальный риск ЧС техногенного характера регионов центральной Сибири.

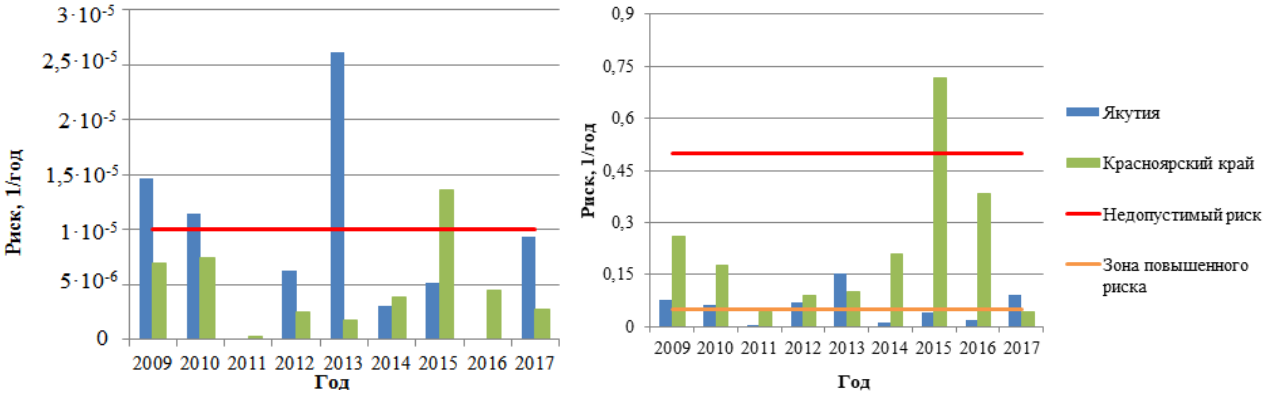



Рис. 3. Индивидуальный (слева) и материальный (справа) риски ЧС техногенного характера Красноярского края и Республики Саха (Якутия).

Расчетно-экспериментальные оценки риск-анализа выполнены для следующих объектов техносферы: гидротехнические сооружения и оборудование ГЭС, сварные элементы корпуса реактора ВВЭР-1000, элементы ракетно-космической техники (металлокомпозитные баки, рефлектора, ферменные конструкции и т.д.), потенциально опасные объекты (трубопроводные системы, крановые конструкции, карьерная техника).

**Оценка территориальных рисков развития.** Исследования территориальных рисков развития выполнены для ряда субъектов Сибирского федерального округа и муниципальных образований Красноярского края, Иркутской области, Республика Саха (Якутия) [3-5]. Проведены расчеты индивидуальных и материальных рисков ЧС техногенного характера (рис. 2 и 3), канцерогенных и неканцерогенных рисков загрязнения воздушной среды и водных ресурсов, популяционных и социальных рисков. Ранжирование муниципальных образований по уровням рисков проведено с разделением их на три группы (опасные, пограничные, безопасные) в зависимости от числа погибших и пострадавших, величины ущерба, ожидаемой продолжительности жизни. Необходимо отметить значительное, в ряде случаев, повышение установленных уровней риска нормативных значений в ретроспективе последнего десятилетия. Полученные результаты позволили сформулировать требования и необходимые мероприятия по повышению защищенности населения и территорий от ЧС природного и техногенного характера.

## ЛИТЕРАТУРА


[1] *Махутов Н.А.* Безопасность и риски: системные исследования и разработки. Новосибирск: Наука, 2017. 724 с.

[2] *Москвичев В.В., Бычков И.В., Потапов В.П., Тасейко О.В., Шокин Ю.И.* Информационная система территориального управления рисками развития и безопасностью // Вестник Российской академии наук. 2017. № 8. С. 696-705.

[3] *Левкевич А.М., Москвичев В.В., Шокин Ю.И. и др.* Безопасность и риски устойчивого развития территорий. Красноярск: Сиб. федер. ун-т, 2014. 222 с.

[4] *Москвичев В.В., Тасейко О.В., Иванова У.С., Черных Д.А.* Базовые региональные риски развития территорий Сибирского федерального округа // Вычислительные технологии. 2018. № 4. С. 95-109.

[5] *Москвичев В.В., Тасейко О.В., Иванова У.С., Черных Д.А.* Базовые риски природно-техногенной безопасности Красноярской промышленной агломерации // Проблемы анализа риска. 2018. № 1. С. 42-47.

# ВЫЯВЛЕНИЕ МАЛЫХ ЦВЕТОВЫХ РАЗЛИЧИЙ УСТРОЙСТВАМИ КОМПЬЮТЕРНОГО ЗРЕНИЯ

*Пальчикова И.Г., Смирнов Е.С.*
Конструкторско-технологический институт
научного приборостроения СО РАН, Новосибирск

Экспериментально показано, что цифровой атлас монохроматических цветовых стимулов с контролируемой шириной спектра возможно построить с помощью монохроматора УМ-2. Представлен метод тестирования спектральной чувствительности *RGB*-сенсоров камеры и измерения ее цветового охвата. Предлагается для обнаружения и количественной характеризации малых цветовых различий элементов изображения разделить объективный анализ изображения на блоки цветовой сегментации и оценки цветового тона (и цветового различия).
*Ключевые слова: цифровое изображение, цвет, доминирующая длина волны, насыщенность.*

**Введение.** В настоящее время для передачи и оценивания цвета и цветовых различий чаще всего используются трехцветные колориметры [1], в которых реализованы оптические системы освещения и построения изображения, что обусловливает их высокое пространственное разрешение. В таких колориметрах регистраторами излучения являются фотоматрицы с трехцветными фильтрами Байера. В этом качестве могут использоваться популярные изображающие системы, такие как цифровые камеры или сканеры. Выявление и количественное выражение малых цветовых различий необходимо в полиграфии [2], при разработке устройств с использованием термокраски [3], в криминалистике [4-6] при экспертизе документов и дифференциации штрихов реквизитов документа, и др.

Задачи, которые решает здесь компьютерное зрение, настолько разнообразны, что практически в каждом конкретном случае требуется разработка специализированных алгоритмов обработки цифровых изображений [7, 8].

Рассмотрение особенностей построения осветительной части колориметров, изучение цветопередачи цифровых камер (или фотоматриц) и качества их цветового профилирования, а также разработка соответствующих алгоритмов и программного обеспечения являются актуальными.

В настоящей работе рассматривается метод определения спектральной чувствительности *RGB*-сенсоров камеры с помощью монохроматических стимулов. Обсуждаются экспериментальная установка и способ создания цифрового атласа монохроматических стимулов. Объективную оценку цветовых различий близких стимулов (или штрихов реквизитов документа) предлагается выполнять путем математической обработки цифровых изображений стимулов по разработанным алгоритмам, включающих расчет карты цветовых характеристик изображений стимулов, ее цветовую сегментацию и статистический анализ данных по областям интереса.

**Методы и результаты исследования.** Для создания монохроматических стимулов и изучения спектральной зависимости отклика каналов камеры была создана экспериментальная установка, состоящая из источника света – лампы накаливания, конденсора, универсального монохроматора УМ-2 (с элементами в составе: линза, призма сравнения, входная щель переменной ширины, объектив коллиматора, диспергирующая призма Аббе со связанным с ней барабаном длин волн, объектив зрительной трубы, выходная щель переменной ширины), малогабаритного многоканального волоконного спектрометра «Колибри-2» (ООО «ВМК-Оптоэлектроника», Россия, Новосибирск), и приемника излучения – камеры *Canon EOS 500D* (*Canon Inc., Japan*) с объективом.

Свет от источника конденсором через линзу направляется на входную щель монохроматора, находящуюся в фокусе объектива коллиматора, после которого свет проходит через призму Аббе и разлагается в спектр. Объектив зрительной трубы строит изображение входной

щели конфокально с выходной щелью монохроматора, которое рассматривается объективом камеры и сохраняется в виде цифрового изображения. Выходная щель УМ-2 является источником монохроматического излучения. Спектральный состав этого излучения контролировался волоконным спектрометром «Колибри-2», спектр снимался при повороте барабана и изменении длины волны на 1 нм.

При получении цифрового атласа цветов одновременно со снятием спектров «Колибри-2» на неподвижную относительно установки цифровую зеркальную камеру брались изображения выходной щели УМ-2. Режим работы камеры «Пользовательский», который подробно охарактеризован в [5]. Этот режим исключает предобработку изображения встроенным процессором камеры по ее усмотрению.

В видимом спектре были получены 168 изображений спектральных стимулов – образцов цифрового атласа с шагом 1 нм. В каждом элементе атласа область полутени удалена. Полученные монохроматические стимулы позволяют протестировать камеру и построить спектральные отклики фотоматрицы. Спектральный отклик фотоматрицы является сверткой функции спектральной чувствительности самой фотоматрицы с функцией спектрального распределения энергии в источнике излучения. На рис. 1 приведена спектральная чувствительность *RGB* элементов фотоматрицы после деконволюции со спектром лампы. Полученные кривые являются спектрами пропускания фильтра Байера на фотоматрице.

По монохроматическим стимулам производились вычисления соответствующих им доминирующих длин волн (ДДВ), величина ДДВ усреднялась по всему изображению стимула. По полученным данным координаты стимулов были нанесены на диаграмму цветности (*x*,*y*) и таким образом получен цветовой охват камеры, как показано на рис. 2, где сплошная жирная линия приведена для средних значений координат (*x*,*y*) на каждом стимуле, а точечная и пунктирная линии – для величин их среднеквадратичных отклонений (СКО) в отрицательную и положительную стороны, соответственно. Полученная фигура заметно отличается от треугольника. Цветовой охват камеры расположен в пределах локуса, на котором указаны длины волн монохроматических стимулов.

Таким образом, предложен метод тестирования цветопередачи цифровых камер с помощью экспериментально получаемых монохроматических стимулов. Апробирование метода выполнено на примере камеры *Canon EOS 500D*.

Правильно откалиброванная камера позволяет выявлять и количественно определять малые цветовые разницы. Цветоразличение и количественные характеристики цветового различия определяются [2] в цветовом пространстве *CIE LAB*. Полное цветовое различие $\Delta E^*_{ab}$ рассчитывается как евклидово расстояние между двумя точками цветами:

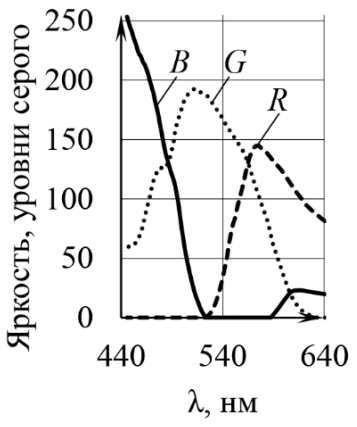



Рис. 1. Спектральный отклик каналов *RGB* после деконволюции со спектром лампы. Графики построены для средних значений яркости.

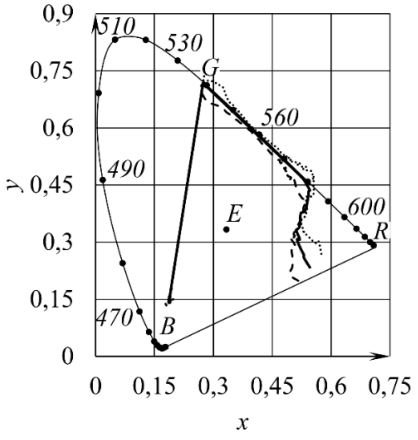



Рис. 2. Экспериментально определенный цветовой охват цифровой камеры на диаграмме (*x*,*y*). Длины волн монохроматических стимулов (в нм) указаны на локусе.

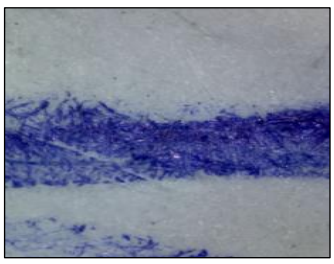


Рис. 3. Микрофотография штриха на бумаге.

$$(\Delta E^*_{ab})^2 = (L^*_T - L^*_R)^2 + (a^*_T - a^*_R)^2 + (b^*_T - b^*_R)^2,$$

где индексы обозначают $T$ – тестируемый стимул, $R$ – стимул сравнения.

Во многих приложениях [9] величина $\Delta E^*_{ab} = 3$ принимается за порог визуального цветоразличения, т.е. цвета, для которых полное цветовое различие меньше трех, при визуальной оценке наблюдателем будут достоверно неразличимы.

В работе [4] представлены результаты экспериментов по выявлению дописок в надписях с помощью трехцветного колориметра *VideoTool-M* [1]. Проведенные [4] эксперименты показали, что при установлении факта изменения первоначального содержания документа возможно, не разрушая документ, дифференцировать однородные, визуально сходные по цветовым характеристикам красящие вещества, используя в качестве количественной характеристики как разницу насыщенности цвета (НЦ) (или ДДВ), так и полное цветовое различие между двумя цветами $\Delta E^*_{ab}$ выделенных областей изображения. Сложность такого подхода заключается в необходимости выделения именно окрашенных областей. Как правило, красящие вещества ложатся на поверхность бумаги неравномерно, и на изображении штриха всегда есть не прокрашенные пятна, как показано на рис. 3. Кроме того, не всегда удается выделить область интереса именно в пределах штриха, захватывается и изображение бумаги. Все это вносит шум, искажает оценку разницы ДДВ, НЦ и $\Delta E^*_{ab}$, и в пограничных случаях ($\Delta E^*_{ab}$~4) затрудняет выделение дописки.

Мы разработали и апробировали новый алгоритм для обнаружения и количественной характеризации дописок по цифровому фотоизображению сделанной надписи. Объективный анализ изображения предлагается разделить на блоки цветовой сегментации и оценки цветового тона и цветовой разницы. В такой постановке блок цветовой сегментации выполняет функции предобработки, задавая карту границ разноокрашенных классов для последующих вычислений.

Чтобы исключить влияние неравномерности освещения и виньетирование объектива, первоначально выполняется фотографирование белого фона и вычисление матрицы коэффициентов для нормировки изображений.

**Цветовая сегментация.** В рассматриваемых задачах необходимо отделить области изображения, относящиеся к бумаге, от окрашенных областей, причем цвета окраски очень близки между собой. То есть все пиксели необходимо разделить на два класса по цвету. Методы пороговой сегментации изображений по яркости известны и достаточно хорошо разработаны [8, 10]. В настоящей работе метод Оцу [8] оптимального глобального порогового преобразования модифицирован применительно к задаче сегментации изображения по цвету.

Сегментация производится по одному из количественных параметров цвета, а именно – по значениям НЦ (либо значениям ДДВ). Из (*R,G,B*) матриц цветного изображения рассчитывается матрица значений НЦ (либо значений ДДВ), которая может содержать значения от 0 до 1. Каждый пиксель имеет уникальное вещественное значение НЦ, поэтому гистограмма НЦ будет представлять собой множество столбцов единичной высоты, причем вдоль оси абсцисс в диапазоне от 0 до 1 столбцы расположатся неравномерно. Алгоритм Оцу не позволяет найти порог сегментации для такой гистограммы уникальных значений НЦ. Обычно алгоритм Оцу применяется к гистограмме яркостей монохромного изображения, в которой величины яркостей – это целые числа в диапазоне от 0 до 255 уровней серого. Предлагается преобразовать гистограмму НЦ путем округления значений или разбиения диапазона НЦ (от 0 до 1) на равномерные части. В нашем случае диапазон НЦ разбивается на 200 частей и вычисляется нормированная гистограмма НЦ для изображения надписи. Предлагаемое построение гистограммы НЦ является первым пунктом модифицированного алгоритма Оцу для определения порога сегментации изображения по НЦ. Последующие пункты алгоритма не изменяются [8, стр. 863].

Порог сегментации выбирается путем анализа гистограммы распределения значений НЦ в пикселях цифрового изображения. Если на гистограмме выделяются две области так, что каждая из этих областей была наиболее компактна, то граница между этими областями на гистограмме будет наилучшим порогом между фоном и надписью на изображении. Алгоритм

ищет порог так, чтобы межклассовая дисперсия двух областей была максимальной. По выбранному порогу НЦ проводится сегментация – отделяются штрихи, имеющие бóльшую НЦ от бумаги, имеющей меньшую НЦ.

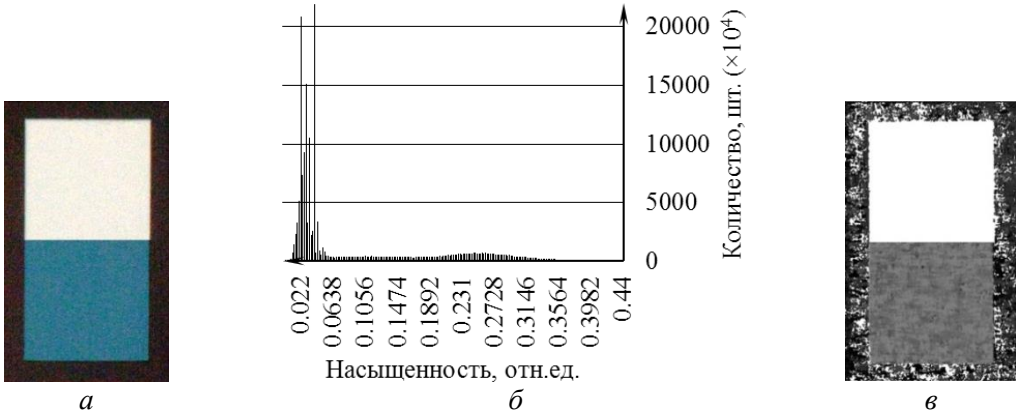
Предлагаемый метод цветовой сегментации программно реализован и корректность работы проверена на контрольном примере – сегментации изображений элементов цветового атласа *Macbeth ColorChecker chart* [10].

Основные этапы цветовой сегментации иллюстрирует рис. 4.

**Оценка цветового тона и цветовой разницы.** Выбранный порог НЦ задает границы разноокрашенных классов изображения для последующих вычислений. Выделенные пиксели, относящиеся к красящим веществам, служат данными для вычислений значений (*R,G,B*) и координат (*x,y*) в каждом пикселе, а так же средних значений ⟨*R,G,B*⟩ по области интереса (по штриху) и величин СКО. Из значений (*R,G,B*) рассчитываются матрицы значений НЦ и ДДВ, цветовые различия, а так же могут быть вычислены матрицы любых пересчётных значений параметров. Такая количественная характеризация позволяет сравнивать области интереса на изображении штрихов между собой. Можно предположить, что если некоторые штрихи выполнены различными двумя красителями, то генеральная совокупность значений цветовых углов всей надписи будет содержать в себе два нормальных распределения (по числу красителей). Для оценки соответствия распределения данных стандартному нормальному, блок оценки цветовой разницы был дополнен построением нормального графика *QQ-plot*.

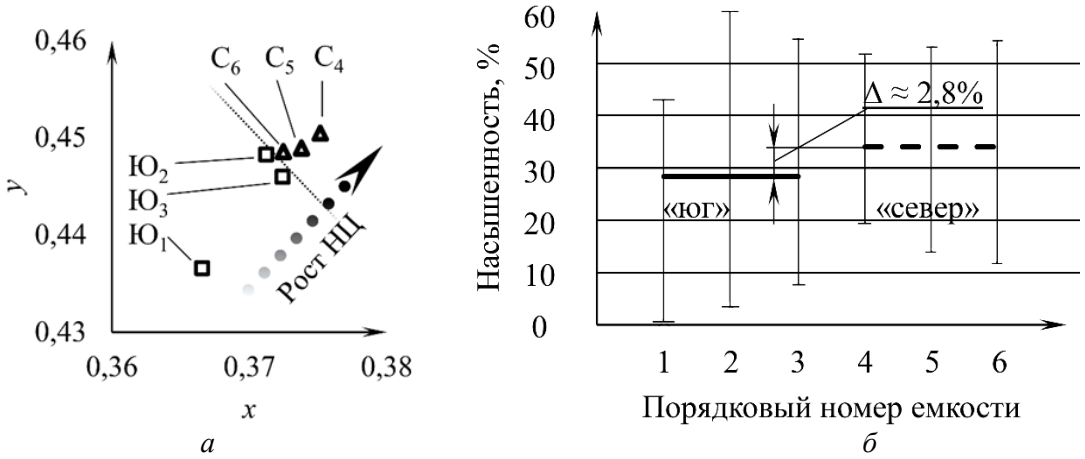
Апробация алгоритма на примерах обнаружения дописок, красящие вещества которых по определителю цвета [12] Н.А. Селиванова не различались, подтвердила его высокую эффективность в решении экспертных задач.

При обработке данных цифровых фотографий в задачах мониторинга природных и антропогенных процессов, одной из важных характеристик является окраска растений. Разработанные программы были апробированы для выявления и количественной оценки малых цветовых различий листьев георгинов, выращенных в весенний период на южном («Ю») и северном («С») окнах помещения, по три емкости, соответственно. Были получены и обработаны 74 фотографии растений на фоне белой бумаги. При съемке использовались одинаковые настройки камеры и одинаковый алгоритм обработки цифровых изображений. Алгоритм включает в себя нормировку исходного снимка на изображение белого фона вблизи изображения растения. Для каждой емкости вычислялись усредненные цветовые координаты (*x,y*) и средние значения ДДВ и НЦ. Координаты шести точек (по числу емкостей) были нанесены на диаграмму цветностей (*x,y*). (рис. 5,*а*). На рис. 5,*а* точечная линия, разделяющая точки «Ю» и «С», параллельна линии локуса, направление увеличения насыщенности обозначено линией со стрелкой. По цвету образцы практически не отличались. Отличие по средним групповым значениям ДДВ составило чуть более 0,5 нм, в то время как СКО было ±1,3 нм.




*а* *б* *в*

Рис. 4. Основные этапы цветовой сегментации: *а* – изображение двух элементов атласа; *б* – гистограмма НЦ изображения; *в* – карта НЦ изображения, полученная в результате обработки модифицированным методом Оцу.




*а* *б*

Рис. 5. Цветовой контраст между группами растений: *а* – расположение цветовых точек емкостей на диаграмме цветности (x,y); *б* – средние величины насыщенностей и их СКО (обозначенных как « I») в зависимости от номера емкости.

На рис. 5,*а* видно, что точки цветности «Ю» и «С» располагаются практически вдоль одной линии НЦ последовательно: вначале Южные «Ю», а затем – Северные «С» по мере увеличения НЦ. По величине НЦ удалось разделить группы с южного и северного окон – разница составила около 2,8% (рис. 5,*б*). Северные образцы обнаружили более насыщенные оттенки. Величины СКО для южных образцов были на 7% больше, чем для северных. Между сравниваемыми группами вычисленное среднее значение полной цветовой разницы составило $\Delta E^*_{ab} = 4$ . То есть визуально группы практически не различимы. Использование камеры в качестве трехцветного колориметра позволило выявить разницу в условиях выращивания растений одного вида и оценить тонкое цветовое отличие в окраске листьев георгинов.

**Заключение.** Создана экспериментальная установка на основе монохроматора УМ-2, которая позволяет получать монохроматические цветовые стимулы с контролируемой шириной спектра. Показано, что спектральная чувствительность *RGB*-сенсоров и цветовой охват камеры могут быть протестированы с помощью монохроматических стимулов.

Разработан новый алгоритм для обнаружения и количественной характеризации дописок по фотоизображению сделанной надписи. Объективный анализ изображения предлагается разделить на блоки цветовой сегментации и оценки цветового тона и цветового различия. В такой постановке блок цветовой сегментации выполняет функции предобработки, задавая карту границ разноокрашенных классов для последующих вычислений цветовых различий.

*Исследование выполнено при финансовой поддержке РФФИ и Правительства Новосибирской области в рамках научного проекта № 17-47-540269 и частично – проекта РФФИ № 19-08-00874.*

ЛИТЕРАТУРА

[1] *Пальчикова И.Г, Алейников А.Ф., Чугуй Ю.В. и др*. Видеоанализатор количественных цветовых характеристик образцов // Приборы. 2014. Т.12. С. 38-44.

[2] ГОСТ Р 52490 – 2005 (ИСО 7724-3:1984). Материалы лакокрасочные. Колориметрия. Часть 3. Введ. 2007-01-01. М.: Стандартинформ, 2006. 6 с.

[3] ГОСТ Р 53698-2009. Контроль неразрушающий. Методы тепловые. Термины и определения. Введ. 2011-01-01. М.: Стандартинформ, 2010. 8 с.

[4] *Баринова О.А., Пальчикова И.Г.* Возможность цветового анализа красящих веществ при производстве технико-криминалистической экспертизы документов // Судебная экспертиза. 2017. № 4(52). С. 75-82.

[5] *Пальчикова И.Г. Латышов И.В., Васильев В.А. и др*. Цветовой анализ цифровых изображений при производстве экспертных исследований следов выстрела // Доклады Академии наук высшей школы Российской Федерации. 2015. № 2 (27). С. 88-101.

[6] *Купин А.Ф.* Криминалистическое исследование рукописей, выполненных с подражанием почерку другого лица: дис. канд. юрид. наук. М., 2012. 234 с.
[7] *Грузман И.С., Киричук В.С., Косых В.П. и др.* Цифровая обработка изображений в информационных системах: Учебное пособие. Новосибирск: Изд-во НГТУ, 2000. 168 с.
[8] *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. М.: Техносфера, 2012. 1104 с.
[9] *Джадд Д., Вышецки Г.* Цвет в науке и технике. М.: Мир, 1978. 592 с.
[10] *Чочиа П.А.* Сегментация изображений на основе анализа расстояний в пространстве признаков // Автометрия. 2014. Т.50, № 6. С. 97-110.
[11] *Mc Camy C.S., Marcus H., Davidson J.G.* A color rendition chart // J. Appl. Photogr. Eng. 1976. Vol. 11. P. 95-99.
[12] *Селиванов Н. А.* Криминалистический определитель цвета. М., 1977. 29 с.

# ТЕКУЩИЕ ВОЗМОЖНОСТИ ЦЕНТРА КОЛЛЕКТИВНОГО ПОЛЬЗОВАНИЯ «ИКИ-МОНИТОРИНГ»

*Прошин А.А., Лупян Е.А., Кашницкий А.В., Балашов И.В., Бурцев М.А.*
Институт космических исследований Российской академии наук, Москва

Для решения многих научных и прикладных задач необходим доступ к многолетним архивам спутниковых данных, а также значительные вычислительные ресурсы для их обработки и анализа, а создание необходимой для этого инфраструктуры в рамках отдельных проектов бывает не всегда возможным. Статья посвящена описанию текущих возможностей центра коллективного пользования «ИКИ-Мониторинг», позволяющего решать такие задачи на основе высокопроизводительных ресурсов центра предоставления данных.
*Ключевые слова: дистанционное зондирование Земли, архивы спутниковых данных, система доступа к спутниковым данным, обработка спутниковых данных, сверхбольшие массивы данных.*

**Введение.** На протяжении последних нескольких десятилетий наблюдается практически экспоненциальный рост объемов данных, поступающих со спутниковых систем дистанционного зондирования Земли [1]. По состоянию на начало 2018 года на орбите Земли действовало более 400 космических аппаратов ДЗЗ, данные нескольких десятков из которых являются общедоступными. Одновременно с этим растет и качество данных, получаемых от современных систем ДЗЗ, которые в основном уже превратились из наблюдательных систем в измерительные. В результате, существенно возрос спрос на данные дистанционного зондирования Земли, которые в настоящее время используются для решения самых разных исследовательских и прикладных задач, связанных с мониторингом природной среды и антропогенных объектов. Существенно, что для решения многих из этих задач необходим доступ к большим многолетним архивам спутниковых данных и различным информационным продуктам, получаемым на их основе, а также значительные вычислительные ресурсы для их обработки и анализа. Однако, в рамках реализации отдельных тематических проектов создание зачастую дорогостоящей инфраструктуры для хранения и обработки больших массивов спутниковых данных далеко не всегда возможно.

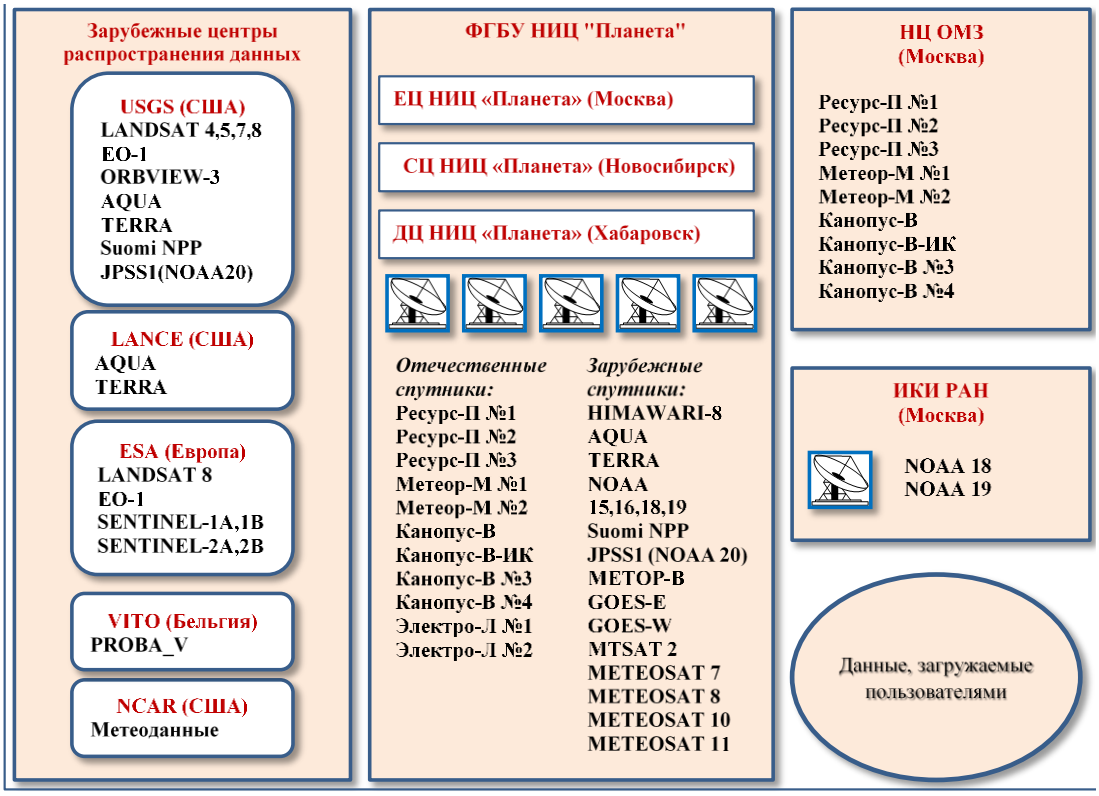
Развитие современных информационных технологий позволило реализовывать принципиально новые подходы к организации работы со спутниковой информацией, обеспечивающие эффективное коллективное использование дорогостоящих вычислительных ресурсов центров сбора, обработки и представления данных дистанционного зондирования. Они позволяют не только обеспечить пользователям доступ к интересующим их спутниковым данным, но предоставить им разнообразные сервисы для их обработки и анализа. Именно на базе таких новых технологий и подходов в 2012 году был создан Центр коллективного пользования системами архивации, обработки и анализа данных спутниковых наблюдений ИКИ РАН – ЦКП «ИКИ-Мониторинг» [2]. Основной его целью является обеспечение распределенной работы со сверхбольшими архивами спутниковых данных в интересах различных научных и прикладных проектов. Центр создан на основе технологий и программного обеспечения, разработанных в ИКИ РАН [3-6]. На текущий момент возможностями этого воспользовалось более 70 различных научных организаций. А различные ресурсы, предоставляемые ЦКП «ИКИ-Мониторинг», использовались при выполнении нескольких десятков научных проектов, поддерживаемых, в том числе РНФ, РФФИ и Министерством науки и образования Российской федерации.

Настоящая статья посвящена рассмотрению текущих возможностей ЦКП «ИКИ-Мониторинг». В ней приводятся общие сведения о накопленных архивах спутниковых данных и тематических продуктов, полученных на основе их обработки. Далее рассматриваются основные механизмы обеспечения доступа к архивам спутниковых данных, при этом особое внимание уделено описанию реализованных на настоящий момент инструментов для удаленного

анализа и обработки данных. В статье также приводятся примеры информационных систем, реализованных на основе возможностей, предоставляемых ЦКП «ИКИ-Мониторинг».

**Архивы.** В настоящее время архивы ЦКП «ИКИ-Мониторинг» содержат данные более чем 35 различных приборов наблюдения, установленных как на отечественных, так и на зарубежных спутниках дистанционного зондирования Земли. Благодаря сотрудничеству ИКИ РАН и ФГБУ «Научно-исследовательский центр космической гидрометеорологии «Планета» пользователям центра также доступны архивы единой системы работы со спутниковыми данными ФГБУ «НИЦ «Планета» [7]. Суммарный объем данных, непосредственно доступных пользователям, в настоящее время составляет около 2,5 петабайтов, а каждые сутки в архивы поступает около 3 терабайт новых данных. Одним из ключевых преимуществ ЦКП «ИКИ-Мониторинг» является уникальное для отечественных ресурсов временное и пространственное покрытие, которое в настоящее время составляет более 27% площади поверхности суши Земли. В настоящее время в зону интересов, по которой производится постоянный сбор и обработка спутниковых данных, входит практически вся территория Северной Евразии, включая арктические территории, приграничные моря России, а также ряд регионов в Африке, Азии, Северной и Южной Америке. Существенно, что в архивах центра на текущий момент накоплены достаточно длинные ряды данных. Например, архивы данных спутников серии Landsat начинаются с марта 1984 года.

Основные источники поступления спутниковых данных в архивы ЦКП «ИКИ-Мониторинг» представлены на рисунке. Большая часть данных поступает из зарубежных центров распространения спутниковых данных, в частности, из различных подразделений американской научно-исследовательской правительственной организации USGS, и европейского космического агентства ESA. Данные отечественных спутниковых систем поступают в архивы в основном из подразделений ФГБУ «НИЦ «Планета», в которых также осуществляется прием и обработка данных многих зарубежных спутников дистанционного зондирования Земли. Кроме этого данные ряда российских спутников поступают из научного центра оперативного мониторинга Земли (НЦ ОМЗ). Отдельно стоит упомянуть поступление в архивы данных, загружаемых пользователями ЦКП «ИКИ-Мониторинг».




Основные источники поступления данных в архивы ЦКП «ИКИ-Мониторинг».

Актуальная информация о доступных пользователям данных доступна на информационном сервере центра в разделе «Архивы данных» (http://ckp.geosmis.ru/default.aspx?page=6). Следует иметь в виду, что в этом разделе приводится информация об объемах, физически хранящихся в архивах данных, большую часть из которых составляют исходные спутниковые данные, используемые для динамического формирования широкого спектра различных «виртуальных» информационных продуктов.

**Доступ к данным в архивах.** В рамках ЦКП «ИКИ-Мониторинг» реализуется три основных механизма доступа к данным:

- спутниковый информационный сервис «ВЕГА-SCIENCE» (http://sci-vega.ru) обеспечивающий для удаленных пользователей интерактивный доступ к архивам данных ЦКП «ИКИ-Мониторинг», а также проведение их обработки и анализа [8, 9];
- программные интерфейсы для доступа к данным из информационных систем;
- программный шлюз, обеспечивающий доступ к данным, физически располагаемым во внешних архивах спутниковых данных.

Спутниковый информационный сервис «ВЕГА-SCIENCE» обеспечивает доступ пользователей к данным архивов ЦКП «ИКИ-Мониторинг» на базе использования многофункционального картографического интерфейса. В рамках этого интерфейса реализован широкий спектр различных инструментов для удаленного анализа и обработки спутниковых данных [10]. Многие из них до недавнего времени были доступны только в специализированных локальных (настольных) и достаточно дорогостоящих комплексах анализа спутниковой информации. Кроме этого пользователям предоставляется доступ к разнообразным отчетным формам, графикам и гистограммам, содержащим количественную информацию, полученную на основе анализа и обработки спутниковых данных, а также необходимых данных из других источников. На основе применения BI-технологии (Business intelligence) реализуются различные интерактивные инструменты анализа (динамические отчетные формы, гистограммы, графики и карты), содержащие как информацию, полученную на основе обработки спутниковых данных, так и данные из других источников [11].

Программные интерфейсы для доступа к данным предназначены для использования возможностей ЦКП «ИКИ-Мониторинг» в рамках различных специализированных информационных систем. Кроме базового функционала по работе с данными в таких системах, как правило, реализуются специальные возможности, предназначенные для решения конкретных научных или прикладных задач.

Программные шлюзы позволяют обеспечивать доступ пользователей центра к данным, физически располагаемым во внешних архивах спутниковых данных. На основе такого механизма, в частности, реализован доступ к данным Геопортала Роскосмоса [12].

**Возможности интерактивного анализа и обработки данных.** В настоящее время специалистами ИКИ РАН реализован широкий спектр различных процедур для удаленного интерактивного анализа и обработки спутниковых данных. При этом, большинство из разработанных на текущий момент инструментов доступно пользователям спутникового информационного сервиса «Вега-Science» (http://sci-vega.ru), там же можно получить достаточно детальное описание имеющихся инструментов для анализа обработки данных и ознакомиться с возможными сценариями их использования. Ниже приводятся основные функциональные группы реализованных на текущий момент инструментов.

1. *Инструменты для доступа и выбора данных* позволяют пользователям однотипно формировать интересующие его наборы данных из различных архивов спутниковых данных, используя при этом такие критерии поиска, как: географическая область, диапазон дат, спутник, прибор, продукт, облачность и другие. Пользователю предоставляется возможность просмотра данных в различном пространственном разрешение и проекции с наложением выбранных картографических слоев. При необходимости пользователь может осуществить дополнительную пространственную привязку данных. Имеются также возможности формирования наборов данных для проведения их дальнейшей обработки и/или экспорта.

2. *Инструменты для обработки спутниковых* данных реализуют широкий спектр различных операций по обработке данных. Ниже приведены примеры таких инструментов:
   - Инструменты расчета спектральных индексов и алгебра изображений, позволяющие проводить арифметические, логические операции и различные математические преобразования над данными, а также вычислять спектральные индексы с произвольными выбранными каналами.
   - Обучаемая и необучаемая классификация спутниковых данных, позволяющая разделять спутниковые изображения на отдельные классы по определенным параметрам.
   - Цветовая коррекция изображения и синтез различных изображений, включая разновременные.
   - Работа с палитрой изображения.
   - Коррекция данных. Инструмент позволяет провести фильтрацию или топографическую коррекцию, вырезать произвольную область выбранных каналов любых выбранных спутниковых изображений из архива.
   - Структурный анализ изображений. Инструмент реализует технологию LESSA (Linement Extraction and Stripe Statistical Analysis). Эта технология предназначена для автоматизации анализа в геологических исследованиях данных различного типа - изображений, схем, цифровых моделей рельефа (ЦМР).
3. *Инструменты для работы с рядами данных* позволяют в графическом виде анализировать временные, пространственные и спектральные ряды данных. В том числе они позволяют проводить анализ:
   - временных рядов данных: значений в заданных точках, индексов по объектам (полям), метеопараметров и т.п.;
   - спектральных профилей в выбранных точках или выделенных объектов;
   - метеоданных, включая их вертикальные профили;
   - пространственных профилей вдоль произвольно заданных маршрутов.
4. *Инструменты анализа различных характеристик данных* позволяют проводить измерение длин и площадей различных объектов, анализ одномерных и двумерных гистограмм, расчет различных статистических характеристик изображений или произвольно выбранных областей изображений.
5. *Специализированные инструменты для анализа и мониторинга* различных объектов. Примеры задач, решаемых при помощи таких инструментов:
   - мониторинг и анализ состояния сельскохозяйственных полей и контроля динамики развития посевов с/х культур;
   - выявление лесных вырубок на основе на использовании временной серии чувствительных к изменениям растительного покрова спектральных каналов спутниковых изображений;
   - выделение пепловых шлейфов и их характеристик;
   - оконтуривание гарей от пожаров по данным высокого разрешения.
6. *Инструменты моделирования* предназначены для прогнозирования изменений тех или иных объектов наблюдения во времени:
   - Моделирование поведения пепловых шлейфов от вулканов.
   - Моделирование динамики развития пожаров.
7. *Инструменты, обеспечивающие работу с метеорологической информацией.*
8. *Инструмент для подготовки презентационных web-интерфейсов* для иллюстрации различных явлений и процессов. С помощью данного инструмента вся полученная и анализируемая информация из информационной системы может быть предоставлена в сокращенном виде, удобном для широкого представления и обсуждения.

**Примеры реализованных информационных систем.** В настоящее время на основе возможностей, предоставляемых центром коллективного пользования ЦКП «ИКИ-Мониторинг», функционирует большое количество различных систем научного и прикладного назначения

(http://smiswww.iki.rssi.ru/default.aspx?page=11). Ниже приведены несколько примеров реализованных на базе возможностей центра информационных систем:

- Спутниковый сервис See The Sea (http://ocean.smislab.ru) [13], предназначенный для решения междисциплинарных задач исследования Мирового океана.
- Информационная система «Дистанционный мониторинг активности вулканов Камчатки и Курил» VolSatView (http://volcanoes.smislab.ru) [14].
- Информационная система дистанционного мониторинга Федерального агентства лесного хозяйства ИСДМ-Рослесхоз (https://aviales.ru) [15], предназначенная для обеспечения оперативного дистанционного мониторинга лесных пожаров и их последствий.

**Заключение.** ЦКП «ИКИ-Мониторинг» предоставляет уникальные возможности для эффективной работы с данными дистанционного мониторинга Земли для решения различных научных и исследовательских задач, позволяя избежать необходимости создания дорогостоящей инфраструктуры для каждого из проектов. По реализованному функционалу, обеспечивающему распределенный анализ и обработку данных, он является ведущим центром в России и конкурирует с наиболее передовыми мировыми системами этого класса.

В перспективе, в первую очередь, планируется продолжить работы по совершенствованию и расширению предоставляемого пользователям функционала для интерактивного анализа и обработки спутниковых данных.

## ЛИТЕРАТУРА


[1] *Лупян Е.А., Бурцев М.А., Прошин А.А., Кобец Д.А.* Развитие подходов к построению информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2018. Т. 15, № 3. С. 53-66. DOI: 10.21046/2070-7401-2018-15-3-53-66.

[2] *Лупян Е.А., Прошин А.А., Бурцев М.А., Балашов И.В., Барталев С.А., Ефремов В.Ю., Кашницкий А.В., Мазуров А.А., Матвеев А.М., Суднева О.А., Сычугов И.Г., Толпин В.А., Уваров И.А.* Центр коллективного пользования системами архивации, обработки и анализа спутниковых данных ИКИ РАН для решения задач изучения и мониторинга окружающей среды // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12, № 5. С. 263-284.

[3] *Прошин А.А., Лупян Е.А., Балашов И.В., Кашницкий А.В., Бурцев М.А.* Создание унифицированной системы ведения архивов спутниковых данных, предназначенной для построения современных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13, № 3. С. 9-27. DOI: 10.21046/2070-7401-2016-13-3-9-27.

[4] *Толпин В.А., Балашов И.В., Ефремов В.Ю., Лупян Е.А., Прошин А.А., Уваров И.А., Флитман Е.В.* Создание интерфейсов для работы с данными современных систем дистанционного мониторинга (система GEOSMIS) // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т.8, № 3. С. 93-108.

[5] *Кобец Д.А., Матвеев А.М., Мазуров А.А., Прошин А.А.* Организация автоматизированной многопотоковой обработки спутниковой информации в системах дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12, № 1. С. 145-155.

[6] *Кашницкий А.В., Лупян Е.А., Балашов И.В., Константинова А.М.* Технология создания инструментов обработки и анализа данных сверхбольших распределенных спутниковых архивов // Оптика атмосферы и океана. 2016. Т. 29, № 9. С. 772-777. DOI: 10.15372/AOO20160908.

[7] *Лупян Е.А., Милехин О.Е., Антонов В.Н., Крамарева Л.С., Бурцев М.А., Балашов И.В., Толпин В.А., Соловьев В.И.* Система работы с объединенными информационными ресурсами, получаемыми на основе спутниковых данных в центрах НИЦ «Планета» // Метеорология и гидрология. 2014. № 12. С. 89-97.

[8] *Лупян Е.А., Савин И.Ю., Барталев С.А., Толпин В.А., Балашов И.В., Плотников Д.Е.* Спутниковый сервис мониторинга состояния растительности («Вега») // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т.8, № 1. С. 190-198.

[9] *Толпин В.А., Лупян Е.А., Барталев С.А., Плотников Д.Е., Матвеев А.М.* Возможности анализа состояния сельскохозяйственной растительности с использованием спутникового сервиса «ВЕГА» // Оптика атмосферы и океана. 2014. Т. 27, № 7 (306). С. 581-586.

[10] *Кашницкий А.В., Лупян Е.А., Балашов И.В., Константинова А.М.* Технология создания инструментов обработки и анализа данных сверхбольших распределенных спутниковых архивов // Оптика атмосферы и океана. 2016. Т. 29, № 9. С. 772-777. DOI: 10.15372/AOO20160908.

[11] *Кобец Д.А., Балашов И.В., Сычугов И.Г., Толпин В.А.* Построение на основе BI-технологий инструментов анализа информации о состоянии лесов, получаемой на основе данных спутниковых наблюдений // Аэрокосмические методы и геоинформационные технологии в лесоведении, лесном хозяйстве и экологии: Доклады VI Всероссийской конференции (Москва, 20-22 апреля 2016 г.) М.: ЦЭПЛ РАН, 2016. С. 24-28.

[12] *Тохиян О.О., Васильев А.Ю., Гладков А.П.* Расширение возможностей сервисов Геопортала Роскосмоса // Геоматика. 2015. № 2. С. 22–25.

[13] *Mityagina M.I., Lavrova O.Yu., Uvarov I.A.* «See the Sea»: Multi-user information system for investigating processes and phenomena in coastal zones via satellite remotely sensed data, particularly hyperspectral data // Remote Sensing of the Osean, Sea Ice, Coastal Waters, and Large WaterRegions Location, Book Series: Proceedings of SPIE, Netherlands, Amsterdam, Article number: 92401C, 2014. Vol. 9240. DOI: 10.1117/12.2067300.

[14] *Гирина О.А., Лупян Е.А., Гордеев Е.И., Сорокин А.А., Крамарева Л.С., Мельников Д.В., Маневич А.Г., Уваров И.А., Кашницкий А.В., Романова И.М., Константинова А.М., Королев С.П.* Информационная система VolSatView для комплексного анализа активности вулканов Камчатки и Курил // IV Всероссийская научно-практическая конференция «Информационные технологии и высокопроизводительные вычисления», Хабаровск, 11-14 сентября 2017 г. С. 36-39.

[15] *Лупян Е.А., Барталев С.А., Ершов Д.В., Котельников Р.В., Балашов И.В., Бурцев М.А., Егоров В.А., Ефремов В.Ю., Жарко В.О., Ковганко К.А., Колбудаев П.А., Крашенинникова Ю.С., Прошин А.А., Мазуров А.А., Уваров И.А., Стыценко Ф.В., Сычугов И.Г., Флитман Е.В., Хвостиков С.А., Шуляк П.П.* Организация работы со спутниковыми данными в информационной системе дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ-Рослесхоз) // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т.12, № 5. С.222-250.

# ЦИФРОВОЕ МАРКИРОВАНИЕ ВИДЕОДАННЫХ АЭРОФОТОСЪЕМКИ

*Фаворская М.Н., Буряченко В.В., Гусев К.А.*
ФГБОУ ВО «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева», Красноярск

Приводится классификация методов цифрового маркирования видеопоследовательностей, а также классификация Интернет-атак, которые подразделяются на преднамеренные и случайные атаки. Предложен способ организации многоуровневой защиты с использованием хрупкого цифрового водяного знака, встраивания информативного водяного знака, содержащего полетную информацию, в текстурированные регионы с его возможным шифрованием. Разработан метод встраивания и извлечения цифровых водяных знаков, инвариантный к глобальным и локальным геометрическим искажениям.
*Ключевые слова: цифровое маркирование, видеоданные аэрофотосъемки, защита авторских прав, безопасность, емкость, устойчивость.*

**Введение.** В последние годы цифровое маркирование видеоматериалов приобретает все большее значение в связи с резким возрастанием мультимедийных данных по незащищенным каналам связи. Цифровое маркирование подразумевает внедрение в изображение или видеопоследовательность скрытых цифровых водяных знаков (ЦВЗ), текстовой или аудио информации в зависимости от решаемой задачи. Цели цифрового маркирования также могут быть различными, и в зависимости от цели используются разные алгоритмы внедрения и извлечения ЦВЗ. В докладе рассматривается задача цифрового маркирования видеоданных аэрофотосъемки актуальной полетной информацией, защищенной авторским правом. Более сложной задачей является цифровое маркирование в виде результатов аннотирования, например, траектории объекта сопровождения, количество движущихся объектов, обнаружение задымления в результате лесного пожара и т.д.

**Классификация методов цифрового маркирования.** Информационная безопасность передачи данных по каналам связи играет первостепенную роль в защите информации. Как правило, системы информационной безопасности разделяются на две основные категории: шифрование (криптография) и сокрытие информации (стеганография и маркирование) [3]. Считается, что методы стеганографии являются дополнительной защитой для криптографических методов. При этом стеганография не изменяет формат данных или сообщений, как и цифровое маркирование. Первоначально, основной функцией цифрового маркирования являлись такие популярные методы, как метод аутентификации авторского права. Основной целью цифрового маркирования являлось создание безопасных, надежных и эффективных водяных знаков, видимых или невидимых, которые внедрялись в медиафайлы или документы, и были постоянными для посторонних лиц [1]. Однако в более широком смысле цифровое маркирование имеет отношение к стеганографии, т. к. оба подхода скрывают информацию в видеоданных. Оба подхода должны соответствовать таким критериям, как безопасность, емкость, устойчивость и незаметность, но приоритеты расставлены по-разному. Так, незаметность имеет первостепенное значение для стеганографии, в то время как цифровое маркирование в первую очередь отвечает за надежность. Однако в настоящее время разработаны комплексные технологии, которые затрудняют явное отнесение того или иного метода к криптографии, стеганографии или цифровому маркированию.

Три противоречивых критерия – надежность сокрытия, объем встраиваемой информации и устойчивость к атакам, – учитываются при разработке алгоритмов встраивания/извлечения водяных знаков. В последнее время водяные знаки рассматриваются в более широком аспекте, чем визуальные или текстовые логотипы, встраиваемые для защиты авторских прав. В качестве водяных знаков используются небольшие текстовые сообщения или область интереса на изображении. Водяные знаки могут быть хрупкими (разрушающимися при любых типах атак),

полухрупкими (при атаке разрушается часть водяного знака) и устойчивыми (в целом, сохраняют свой первоначальный вид, а степень равномерного разрушения зависит от неизвестных параметров атаки).

Водяные знаки внедряются в пространственную или частотную область изображения или кадра. Пространственные методы являются более простыми в реализации и, в то же время, более заметными для человеческого глаза и менее устойчивыми к атакам относительно частотных методов встраивания. К пространственным методам относятся: метод наименьшего значащего бита, разность значений пикселов, метод смещения гистограммы, метод на основе битовых плоскостей, метод на основе квантования, метод на основе шаблона, метод на основе модуляции и т.д. Частотные методы проявляют большую устойчивость к атакам и надежность сокрытия встраиваемой информации, однако объем встраиваемой информации существенно ниже. Частотные методы встраивания скрытой информации основаны на преобразованиях (дискретное преобразование Фурье, дискретное косинусное преобразование, полярное гармоническое преобразование, дискретное вейвлет-преобразование, комплексные вейвлет-преобразования, дискретное курвлет-преобразование, дискретное шиарлет-преобразование), а также моментах (моменты Зернике, моменты псевдо-Зернике, моменты Чебышева).

Также возможны «слепое» маркирование и не «слепое» маркирование. При «слепом» маркировании исходное изображение или кадр не передаются по каналам связи, передается только секретный ключ. В этом случае алгоритмы извлечения водяных знаков и оценки качества восстановленного после атак изображения или кадра существенно усложняются.

**Организация многоуровневой защиты видеоданных аэрофотосъемки.** Концепция многоуровневой защиты все чаще применяется на практике как реакция на более изощренные Интернет-атаки, применяемые к мультимедийному контенту. Следует отметить большое разнообразие типов атак в отношении видеоматериалов, которые могут быть как преднамеренными, так и случайными. Преднамеренные атаки направлены на искажение видеопоследовательности в целом или одного кадра. Преднамеренные атаки подразделяются на функциональные и геометрические атаки, в то время как случайные (непреднамеренные) атаки являются только функциональными атаками.

Удаление кадров означает удаление одного или нескольких кадров из маркированной видеопоследовательности. Усреднение кадров искажает движение в сцене. Изменение порядка кадров означает изменение последовательности отображения кадров. Если количество удаленных, усредненных или замененных кадров велико, то качество маркированной видеопоследовательности будет низким. Ряд атак, таких как MPEG/JPEG сжатие, искажение цвета, изменение контраста и добавление шума, применяются как ко всей видеопоследовательности, так и к отдельным кадрам. Атаки копирования используются для подделки кадра на основе текстурного анализа [4]. Кадр может быть произвольно искажен аффинными преобразованиями (повороты, масштабирование, сдвиги вдоль координатных осей), а также переворотами, удалением части кадра или локальным изгибом. Композиционные атаки подразумевают одновременное применение к кадру нескольких видов атак. Также геометрические атаки могут быть глобальными и локальными. Наличие бликов, MPEG сжатие, изменение частоты кадров или разрешения являются типичными непреднамеренными атаками.

Следует отметить, что любые манипуляции с видеопоследовательностями – очень простой процесс редактирования. В то же время восстановление искаженной маркированной видеопоследовательности является большой проблемой, т.к. параметры искажений неизвестны. В настоящее время существующие методы «слепого» маркирования не могут предотвратить большинство распространенных видов атак, а использование метода не «слепого» маркирования требует передачи исходной видеопоследовательности. Поэтому развитие методов «слепого» маркирования происходит в направлении многоуровневой защиты и применении инвариантных к ряду атак преобразований видеоконтента.

В качестве первого уровня защиты применяется внедрение хрупкого (видимого, полувидимого или невидимого) ЦВЗ $WM_{FR}$ на установленное заранее местоположение. В качестве

хрупкого ЦВЗ можно использовать логотип организации. Встраивание хрупкого ЦВЗ целесообразно выполнять с использованием дискретного преобразования Адамара, не требующего больших вычислительных ресурсов [7]. Подавляющее большинство манипуляций с контентом приводят к частичному или полному разрушению данного ЦВЗ.

Второй уровень защиты связан с применением инвариантных к большинству видов атак методов внедрения, основанных, в основном, на частотных преобразованиях. Если видеоданные аэрофотосъемки не подвергаются сжатию, то целесообразно для встраивания полетной информации $WM_{FI}$ использовать дискретное вейвлет-преобразование или дискретное шиарлет-преобразование с применением сигнулярного разложения [8]. Также немаловажную роль играет выбор регионов для встраивания скрытой информации. Основные принципы выбора регионов для встраивания заключаются в нахождении сильно текстурированных регионов, регионов, не привлекающих внимание человека, и регионов с превалированием компоненты синего цвета как цвета, к которому человеческий глаз менее чувствителен [5].

Третий уровень защиты заключается в возможности шифрования основного ЦВЗ до встраивания в видеоконтент. Если основной ЦВЗ является изображением, то для этого, как правило, используются обратимые хаотические преобразования, самым известным из которых является преобразование Арнольда [2]. Преобразование Арнольда является периодическим обратимым отображением. Число итераций, приводящих к появлению исходного изображения, называется периодом Арнольда. Заранее выбранное количество итераций записывается в секретный ключ. Определенный повтор зашифрованного изображения (период Арнольда минус значение секретного ключа) приводит к полному восстановлению исходного изображения. Такая процедура называется процедурой скремблирования. Если основной ЦВЗ представляет собой установленный формат числовых данных, то можно применить типовые процедуры шифрования текстовых данных (подстановки, перестановки), параметры которых также записываются в секретный ключ.

**Метод встраивания и извлечения ЦВЗ.** Основными процессами цифрового маркирования являются:

- подготовка ЦВЗ (включая преобразование к требуемому формату и шифрование при необходимости) двух типов – ЦВЗ, содержащее полетную информацию $WM_{FI}$, и хрупкий ЦВЗ $WM_{FR}$, а также секретного ключа $K$ – процесс $GN$. Следует отметить, что при обнаружении события в сцене (например, объекта слежения, признаков экологической катастрофы, лесного пожара с помощью дополнительных программных средств) формируется ЦВЗ события $WM_{EV}$;
- внедрение ЦВЗ обоих типов в предварительно выбранные области изображения-контейнера (кадра) – процесс $EM$;
- извлечение всех имеющихся ЦВЗ из изображения-контейнера после передачи по сетям с использованием секретного ключа – процесс $EX$;
- оценка качества извлеченной информации, восстановление ЦВЗ и изображения-контейнера при необходимости – процесс $RC$.

Каждый процесс имеет свои особенности и заслуживает отдельного рассмотрения. При этом процессы внедрения и извлечения ЦВЗ являются взаимообратными. Однако процесс внедрения ЦВЗ играет определяющую роль как в плане качества сокрытия информации, так и устойчивости к Интернет-атакам. Авторами предложен оригинальный метод адаптивного цифрового маркирования, устойчивый к глобальным и локальным геометрическим атакам, основанный на использовании точечных дескрипторов. Дело в том, что точечные особенности являются устойчивыми к аффинным преобразованиям и если функцию, описывающую окрестность точечной особенности на единичном круге, преобразовать в инвариантную к поворотам изображения функцию (например, с помощью экспоненциальных моментов), то в нее можно внедрить информацию о координатах данной точечной особенности на исходном изображении. Такая обработка применяется к ограниченному числу точечных особенностей (не более 10), равномерно распределенных по кадру. Это дает возможность вычислить параметры

аффинного преобразования и нормализовать изображение до процесса извлечения информативных ЦВЗ. Отметим, что хрупкий ЦВЗ используется на самой первой стадии процесса извлечения. Если он не претерпел изменений, то атак применено не было. Для встраивания хрупкого ЦВЗ применяется дискретное преобразование Адамара [7], информативные ЦВЗ встраиваются с помощью дискретного вейвлет-преобразования. Используется метод «слепого» маркирования. После компенсации глобальных геометрических искажений соответствующие точечные особенности анализируются на наличие локальных геометрических искажений. Если установлено наличие локальных геометрических искажений, то после извлечения ЦВЗ применяется бикубическая интерполяция для повышения качества кадра.

**Экспериментальные результаты.** Эффективность работы алгоритма встраивания водяного знака в кадры видеопоследовательности исследовалась с использованием базы данных, полученной с беспилотных летательных аппаратов при различных условиях съемки [9]. База включает 12 видеопоследовательностей, полученных с использованием дрона DJI Mavic Pro.

В процессе работы осуществляется внедрение двух графических водяных знаков, которые могут отображать информацию о компании или дополнительные данные, например, о параметрах съемки. Один водяной знак является разрушаемым и видимым. Он подвергается негативным эффектам от различных атак. Второй водяной знак является скрытым и обладает повышенной устойчивостью к преднамеренным и случайным атакам на видеоматериал.

Эффективность алгоритма внедрения водяного знака оценивалась по нескольким параметрам. Одним из способов является сравнение качества оригинального изображения, и изображения со встроенным водяным знаком. Для этого используются метрики PSNR (Peak Signal-to-Noise ratio) и метрика NCC (Normalized correlation coefficient) [6].

Оценка качества изображений.

| **Виды атак** | **PSNR (Оригинальное и маркированное изображения), %** | | | **NCC (Оригинальный водяной знак и водяной знак, подвергнутый атаке)** | | |
|---|---|---|---|---|---|---|
| | **Видеопоследовательности** | | | | | |
| | *Creux du Van Flight* | Bluemlisal Flyover | Berg-house Leopard | *Creux du Van Flight* | Bluemlisal Flyover | Berghouse Leopard |
| Нет | 30,68 | 31,23 | 32,28 | 1,00 | 1,00 | 1.00 |
| Поворот (15º) | 29,27 | 29,88 | 31,71 | 0,89 | 0,93 | 0,96 |
| Шум вида Соль и перец (0.01) | 30,33 | 30,89 | 31,33 | 0,97 | 0,99 | 0,98 |
| Гауссов шум (0, 0.01) | 30,41 | 30,91 | 31,45 | 0,95 | 0.97 | 0,96 |
| Коррекция яркости (1.2) | 30,27 | 30,.45 | 31,29 | 0,95 | 0,98 | 0,97 |
| Размытие по движению (10, 45º) | 30,21 | 30,37 | 31,10 | 0,96 | 0,91 | 0,95 |
| Коррекция яркости, гауссов шум, размытие | 30,13 | 30,22 | 30,89 | 0,91 | 0,85 | 0,92 |
| Медианный фильтр (3×3) | 30,43 | 31,08 | 31,99 | 0,93 | 0,94 | 0,97 |
| Гауссов шум (0, 0.01) и медианный фильтр (3×3) | 30,15 | 30,68 | 31,81 | 0,91 | 0,93 | 0,92 |
| Масштабирование (1.15) | 29,86 | 30,76 | 31,11 | 0,83 | 0,86 | 0,89 |
| Обрезка изображения (25%) | 29,91 | 30,51 | 31,48 | 0,91 | 0,87 | 0,84 |
| JPEG-сжатие | 30,21 | 30,89 | 31,05 | 0,85 | 0,93 | 0,91 |

В таблице представлены результаты внедрения водяного знака на видеопоследовательности различного качества, полученные с беспилотных летательных аппаратов. Наилучшие результаты получены для видеопоследовательности Berghouse Leopard.avi, что связано с хорошим качеством съемки и простой структурой сцены. Также, в данной видеопоследовательности фон содержит большую часть травы с отслеживаемой текстурой, что повышает эффективность алгоритма встраивания водяного знака.

Наиболее сильное влияние на качество восстановления водяного знака оказывают геометрические преобразования – масштабирование, обрезка кадра и поворот, при этом теряется до 19% информации. Алгоритм показывает высокую устойчивость к типичным видам атак, которые могут происходить непреднамеренно при передаче информации – зашумление, размытие и сжатие данных, потери составляют не более 10% данных, содержащихся в водяном знаке.

**Заключение.** В работе предложен метод встраивания скрытых и хрупких водяных знаков в кадры видеопоследовательностей, который может применяться для цифрового маркирования при аэрофотосъемке с беспилотных летательных аппаратов и дронов для защиты информации либо встраивания дополнительных данных в изображения, например содержащих координаты или высоту полета. Разработан алгоритм, позволяющий обеспечить высокий уровень защиты данных за счет шифрования ЦВЗ и встраивания хрупкого водяного знака, который позволяет получить сведения о наличии атак на видеопоследовательность. Проведенные эксперименты по имитации преднамеренных и случайных атак показывают высокую устойчивость водяного знака к геометрическим преобразованиям и другим видам атак, которые могут случаться при передаче видеоданных.

*Работа выполнена при финансовой поддержке РФФИ (грант № 19-07-00047-а).*

ЛИТЕРАТУРА

[1] *Shih F.Y.* Digital Watermarking and steganography: Fundamentals and Techniques. 2nd edn., Boca Raton, London, New York: CRC Press, 2017. 270 p.

[2] *Arnol'd, V.I., Avez, A*. Ergodic problems of classical mechanics. Mathematical physics monograph series. New York, Benjamin, 1968. 286 p.

[3] *Cheddad A., Condell J., Curran K., Mc Kevitt P.* Digital image steganography: survey and analysis of current methods // Signal Processing. 2010. V. 90, no. 3. P. 727-752.

[4] *Lu C.S., Hsu C.Y.* Near-optimal watermark estimation and its countermeasure: antidisclosure watermark for multiple watermark embedding // IEEE Transactions on Circuits and Systems for Video Technology. 2007. V. 17, no. 4. P. 454-467.

[5] *Favorskaya, M., Pyataeva, A., Popov, A.* // Texture analysis in watermarking paradigms. Procedia Computer Science. 2017. V. 112. P. 1460-1469.

[6] *Sachin, G., Vinay, K.* // A RDWT and Block-SVD based Dual Watermarking Scheme for Digital Images. International Journal of Advanced Computer Science and Applications (IJACSA). 2017. V. 8 Issue 4, P. 211-219.

[7] *Favorskaya M., Savchina E., Popov A.* Adaptive visible image watermarking based on Hadamard transform // IOP Conference Series: Materials Science and Engineering, MIST Aerospace, 2018. 2018450:052003.

[8] *Favorskaya, M.N., Jain, L.C. Savchina E.I.* Perceptually tuned watermarking using non-subsampled shearlet transform // Computer Vision in Control Systems-3: Springer International Publishing Switzerland. 2018. ISRL, V. 136. P. 41-69.

[9] Drone Videos DJI Mavic Pro Footage in Switzerland. **https://www.kaggle.com/kmader/drone-videos** (дата обращения 05.05.2019).

# СИСТЕМА ОБРАБОТКИ СПУТНИКОВЫХ ДАННЫХ В РЕГИОНАЛЬНОМ ЦЕНТРЕ ДЗЗ ФИЦ КНЦ СО РАН

*Якубайлик О.Э., Кадочников А.А., Токарев А.В.*
Федеральный исследовательский центр «Красноярский научный центр СО РАН»,
Институт вычислительного моделирования СО РАН, Красноярск

Обсуждаются методы создания многослойных картографических изображений для систем оперативного спутникового мониторинга, основанных на технологиях веб-ГИС. Формируемая карта может содержать набор растровых и векторных тематических слоев пространственных данных. Настройка параметров представления предусматривает возможность выбора формата геоданных на клиенте, управления параметрами отображения (интерактивная легенда, перечень отображаемых атрибутивных полей), использования настраиваемых шаблонов для вывода табличной информации.
*Ключевые слова: картографический веб-сервис, геопортал, пространственные метаданные, система спутникового мониторинга, веб-картография, каталог геоданных, веб-приложение, веб-ГИС.*

**Введение.** Развитие систем дистанционного зондирования Земли в настоящее время обусловлено совокупностью ряда факторов – увеличением числа космических аппаратов и улучшением их технических характеристик, повышением доступности спутниковой информации, связанной с развитием сети Интернет: ростом скорости и снижением стоимости доступа, развитием веб-сервисов и стандартов передачи геопространственных данных [1]. Новые технологии автоматизации процессов получения и обработки данных дистанционного зондирования Земли, создания сверхбольших архивов в специализированных дата-центрах позволили организовать принципиально новые схемы работы со спутниковыми данными, создать новое поколение систем оперативного мониторинга природной среды. Их ключевым элементом стали геоинформационные веб-технологии, интерактивные онлайновые картографические системы с прямым доступом к спутниковой информации, обеспеченные возможностями настройки различных параметров визуализации геопространственных данных и построения сложных аналитических запросов.

Технологии геоинформационных веб-систем и сервисов динамично развиваются в настоящее время [2-4]. В Интернете можно найти уже более сотни российских геопорталов, десятки геосервисов, веб-ГИС, картографических сервисов и визуализаторов данных. В том числе – специализированных решений для работы со спутниковыми данными. Наряду с уже традиционными задачами простой визуализации карт местности и мозаик спутниковых снимков, современные системы предоставляют пользователю возможности гибкой настройки различных параметров картографической визуализации, сервисы обработки пространственных данных [5].

Опыт собственных исследований и разработок в рассматриваемой предметной области показывает, что системообразующим элементом в разработке программного обеспечения для решения задач мониторинга оценки состояния окружающей природной среды на основе данных дистанционного зондирования Земли является геоинформационная платформа, концептуально определяющая архитектуру и возможности реализуемых решений. Основным трендом в настоящее время стал переход к использованию в качестве рабочего места конечного пользователя ГИС-решения на основе веб-браузера, а также встраивания необходимого набора скриптов для работы с системой в геоинформационные Интернет-порталы. В некоторых случаях данные решения являются вспомогательными и выполняют в основном функции просмотра пространственных данных.

**Геоинформационная платформа.** Современные веб-ГИС строятся в так называемой сервис-ориентированной архитектуре, и их можно рассматривать как комплекс взаимосвязанных программных средств для управления пространственными данными – импорта/экспорта,

каталогизации, визуализации, создания, обработки, распространения, и т.д. [6, 7]. Технологической основой подобных решений обычно выступают библиотеки программных интерфейсов типа Google Maps API, Mapserver Mapscript, Сканекс GeoMixer, и проч.; они обеспечивают доступ к функциям и контексту картографических элементов веб-страниц – средствам визуализации карты, пространственным метаданным, например – дорожная сеть с данными о пробках, рельеф, точки на карте, и т.д. [8]. Программно-технологическое обеспечение системы спутникового мониторинга также создается в соответствии с указанным подходом, т.к. она тоже представляет из себя веб-ГИС.

Одной из первых задач на этапе проектирования любой прикладной системы является выбор технологической схемы (геоинформационной платформы) построения ГИС – программного обеспечения, технологий для хранения пространственных данных и их обработки. В настоящее время традиционно рассматриваются две альтернативы в выборе базового программного обеспечения – на основе коммерческих программных продуктов или свободных бесплатных (opensource) программ. Оставляя за скобками финансовый, философский и конъюнктурный аспекты этого выбора, хотелось бы отметить, что сегодня на практике чаще всего нет противопоставления двух рассматриваемых подходов. И причина в том, что сейчас коммерческие и свободные ГИС хорошо дополняют друг друга – благодаря совместимости форматов данных, стандартам информационного обмена, основанным на веб-сервисах, и т.д. Можно, например, выполнять анализ пространственных данных в ESRI ArcGIS, конвертировать их в MapInfo для передачи заказчику, и при этом использовать свободное программное обеспечение Mapserver для представления на веб-страницах, а каталог пространственных метаданных формировать средствами GeoNetwork Open Source. При этом для хранения пространственных данных использовать открытую СУБД PostgreSQL с модулем расширения PostGIS, что для подавляющего большинства задач практически не уступает по производительности и функциональным возможностям ESRI ArcGIS Server или СУБД Oracle с расширением для работы с пространственными данными Oracle Spatial. Определяющим фактором становятся компетенции специалистов, которые осуществляют техническую поддержку выбранных решений. Коммерческие продукты обычно проще в первоначальной настройке, сразу работают «из коробки». Свободное программное обеспечение тоже имеет свои преимущества, связанные с открытостью архитектуры, возможностями расширения.

Отмечая специфику систем спутникового мониторинга, здесь следует отметить, что, например, большинство алгоритмов обработки спутниковых данных, созданных в NASA, доступны как раз в исходных кодах. Специалисты со всего мира проявляют к подобным ресурсам значительный интерес, т.к. получают возможность непосредственного анализа существующих методов обработки данных и создания собственных модифицированных алгоритмов. NASA заинтересовано в конструктивной критике и оперативном исправлении замеченных ошибок.

**Система визуализации спутниковых данных.** В качестве программной основы каталога спутниковых данных использовался геоинформационный веб-сервер, модульная архитектура которого позволяет развивать систему уже в процессе её эксплуатации, путем добавления новых средств доступа к данным, а также совершенствования уже существующих средств. Вся разработка выполнена на основе свободного и открытого программного обеспечения (free & open source) – как в части ГИС, так и во всех других компонентах разработки. Рассматриваемое веб-приложение обеспечивает пользователя средствами наполнения и редактирования данных и метаданных системы, поиска и классификации картографических ресурсов, веб-сервисами прямого доступа к данным на основе стандартных протоколов WMS/WFS, возможностями аналитической обработки [9].

Серверная часть системы реализована с применением шаблона проектирования MVC (модель-представление-контроллер), технологий «Веб 2.0» [10]. Использование данной архитектуры предполагает разделение данных приложения, пользовательского интерфейса и управляющей логики на три отдельных компонента: модель, представление и контроллер – таким образом, что модификация каждого компонента может осуществляться независимо. В

условиях постоянной модернизации системы, непрерывного уточнения технических требований и постановки задачи, эти возможности становятся очень актуальными.

В процессе разработки рассматриваемой системы было создано несколько новых программных компонентов и библиотек, которые можно использовать (тиражировать) в других проектах [11]. Это элементы пользовательского интерфейса, сервисы работы с геопространственной СУБД, прикладные картографические веб-сервисы, и т.д.

Базовый функционал веб-ГИС обеспечивается средствами программного обеспечения UMN Mapserver и MapGuide Open Source. Также используются геопространственная СУБД PostgreSQL/PostGIS, системы управления веб-контентом на основе CMS Drupal и вики-системы DokuWiki, целое семейство библиотек построения пользовательского интерфейса и прочих компонентов системы для языков программирования JavaScript, PHP – JQuery, Fusion, TinyMCE, ExtJS, MapScript, и многие другие.

**Особенности реализации каталога спутниковых данных.** Приоритетной задачей исследований и разработок на данном этапе было создание программных средств для работы с каталогами спутниковых данных. Разработанные программные инструменты обеспечивают решение первоочередных задач оперативной обработки данных ДЗЗ, поступающих с нового спутникового приемного комплекса УниСкан в ФИЦ КНЦ СО РАН, который был введен в эксплуатацию весной 2017 г.

Базовое программное обеспечение спутникового приемного комплекса ФИЦ КНЦ СО РАН формирует набор стандартных продуктов 1-го уровня – поканальные изображения (яркости спектральных каналов). Дальнейшие задачи, связанные с извлечением полезной информации, тематической обработкой данных, решаются отдельно, средствами специального программного обеспечения.

Система веб-визуализации спутниковой информации основана на наборах специально формируемых коллекций мультимасштабных снимков – с возможностью выбора в веб-интерфейсе комбинаций отображаемых каналов на мелких масштабах и в то же время – наличием детальных данных на максимально доступном пространственном разрешении. Ограничение в выборе комбинаций каналов на детальном уровне связано с экономией дискового пространства. Таким образом реализуется компромисс между системой класса «работаем с грубыми квик-луками» и системой уровня «все что угодно с максимальной детализацией». Созданный веб-интерфейс обеспечивает просмотр архива спутниковых снимков. В нем реализована возможность выбора сенсора космического аппарата, выбора предопределенного набора каналов и продуктов у каждого снимка для простого анализа данных. Возможности созданного программного модуля позволяют комбинировать любые сочетания каналов, доступных в изображении без дополнительной настройки серверного программного обеспечения.

При формировании цветных изображений используется LUT (Look Up Table), это своеобразная «таблица поправок» для внесения изменений в каждый из трех каналов. Ранее для улучшения качества изображения использовалось спектральное преобразование, которое строится на работе со спектральной диаграммой, показывающей зависимость между количеством пикселей изображения и значениями спектральной яркости. При спектральных преобразованиях изменяется такой параметр как контрастность. Для повышения контрастности использовалось линейное растягивание гистограммы, заключающееся в том, что всем значениям яркости присваиваются новые значения с целью охватить весь возможный диапазон от 0 до 255. LUT позволяет изменять значение яркости точек изображения при преобразовании снимка к 8 битам из 16 бит с помощью линейной интерполяции между ними. Это позволяет сделать снимок более ярким. Пример веб-интерфейса представлен на рисунке.

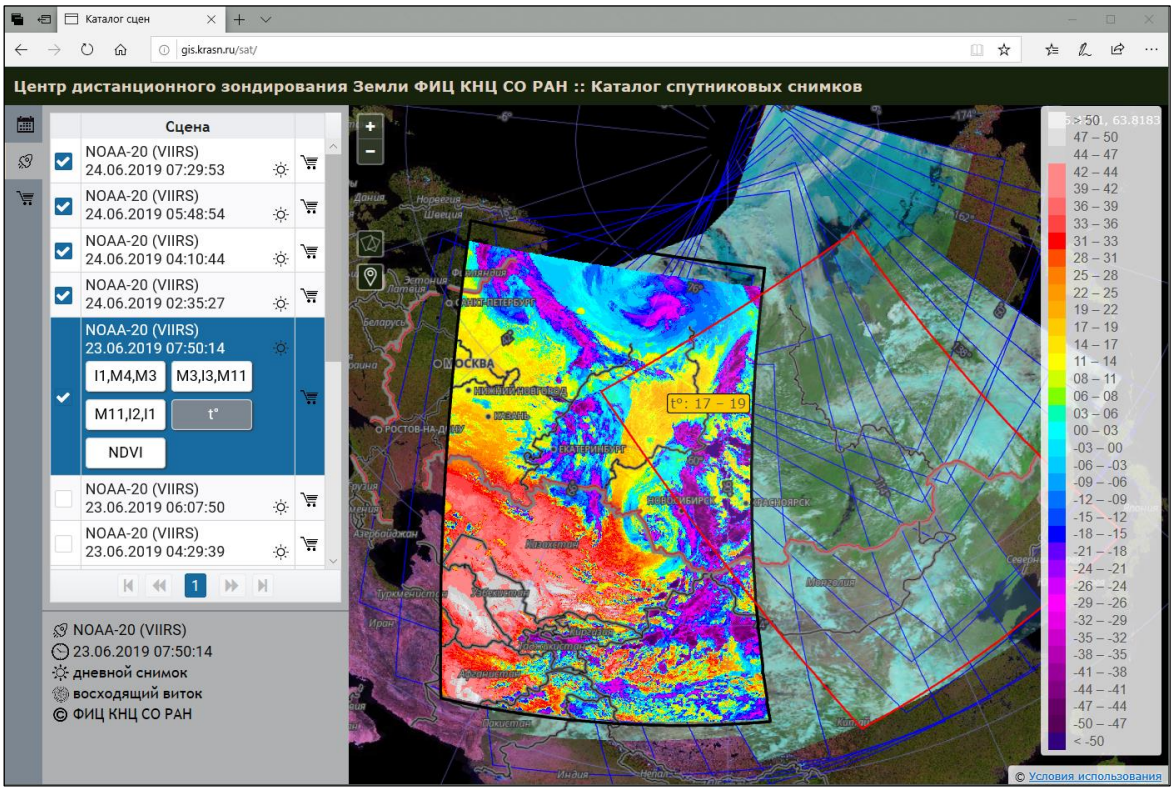
Основные задачи разработки на данном этапе связаны с реализацией методов и алгоритмов обработки спутниковых данных. Технологические особенности работы с растровыми данными предопределяются необходимостью быстрого отображения файлов больших объемов. Эксперименты с различными форматами и их параметрами показали, что оптимальным выбором с точки зрения минимизации времени отображения в веб-браузере является формат TIFF с геопривязкой (GeoTIFF), с использованием структуры TILES (изображение состоит из

набора независимо хранящихся однотипных фрагментов), предварительно вычисленными пирамидальными (обзорными) слоями OVERVIEW. Все перечисленные этапы обработки выполняются утилитами библиотеки GDAL.

Разработанные программные инструменты обеспечивают решение первоочередных задач оперативной обработки поступающих данных ДЗЗ, их каталогизации, интерактивной визуализации с помощью веб-приложения. Рассматривались следующие данные: MODIS TERRA/AQUA, Suomi NPP/NOAA-20.

Для быстрого и удобного поиска в каталоге спутниковых данных, для минимизации нагрузки на серверное программное и аппаратное обеспечение подготовлен набор серверных приложений для предварительной обработки спутниковых данных, включающих следующие этапы обработки:

1. Преобразование исходных растровых данных в формат GeoTIFF с преобразованием исходной проекции в азимутальную равновеликую проекцию Ламберта (Lambert Azimuthal Equal Area) (код EPSG:3576), являющуюся обязательной для работы сервисов Open Geospatial Consortium.
2. Создание для всех сцен каталога цветных изображений в грубом разрешении («квик-луков») в растровом формате PNG, с прозрачностью. Такие изображения используются для быстрого обзорного отображения спутниковых данных, зарегистрированных в каталоге, с одновременной визуализацией набора таких «квик-луков» в окне веб-браузера, с масштабированием на стороне клиента, без перезагрузки/докачки данных через Интернет.
3. Создание мультимасштабного набора многоканальных изображений формата GeoTIFF. Каждое такое изображение, созданное в определенном пространственном разрешении, содержит несколько «значимых» каналов, из которых в веб-приложении «на лету» генерируются некоторый набор цветных композитных изображений. Например, для сенсора TERRA/MODIS мультиканальное изображение содержит 1, 2, 3, 4, 7, 31 каналы, на основе которых динамически формируются набор обычно используемых на практике композитных R-G-B изображений: 1-4-3 («естественные цвета»), 7-2-1 («пожары»), 3-6-7 («снег и лед»), а также одноканальные изображения с палитрами – 31 («температура»), 2-1/2+1 (NDVI).




Интерфейс каталога спутниковых снимков ФИЦ КНЦ СО РАН.

**Заключение.** Создание эффективного программно-технологического инструментария для задач регионального спутникового мониторинга на основе технологий геоинформационных веб-систем имеет значительные перспективы. Развиваемый подход может стать основой в решении актуальных прикладных задач, в основе которых – использование данных дистанционного зондирования Земли, современной инфраструктуры пространственных данных – для эффективного социально-экономического и инновационного развития, повышения конкурентоспособности и обеспечения безопасной жизнедеятельности.

## ЛИТЕРАТУРА


[1] *Kashnitskii A.V., Lupyan E.A., Balashov I.V., Konstantinova A.M.* Technology for designing tools for the process and analysis of data from very large scale distributed satellite archives // Atmospheric and Oceanic Optics, 2017, V. 30, № 1, P. 84-88.

[2] *Pinde Fu, Jiulin Sun* Web GIS: principles and applications, ESRI Press, 2011.

[3] *Li Songnian, Suzana Dragicevic, Bert Veenendaal.* Advances in Web-based GIS, Mapping Services and Applications, CRC Press, 2011.

[4] *J.T. Sample, K. Shaw, S. Tu, et al* Geospatial Services and Applications for the Internet, Springer-Verlag US, 2008.

[5] *Mari R., Bottai L, Busillo C, Calastrini F. et al.* A GIS-based interactive web decision support system for planning wind farms in Tuscany (Italy) // Renewable Energy, 2011, 36, pp. 754-763.

[6] *David B. Lobell, David Thau, Christopher Seifert, Eric Engle, Bertis Little* A scalable satellite-based crop yield mapper // Remote Sensing of Environment, 2015, 164, pp. 324-333.

[7] *L. Bastin, G. Buchanan, A. Beresford, J. F. Pekel, G. Dubois* Open-source mapping and services for Web-based land-cover validation // Ecological Informatics, 2013, 14, pp. 9-16.

[8] *Yakubailik O.E., Kadochnikov A.A., Tokarev A.V.* Web geographic information system and the hardware and software ensuring rapid assessment of air pollution // Optoelectronics, Instrumentation and Data Processing, 2018, 54, pp. 243-249.

[9] *Yakubailik O.E., A. Kadochnikov A.A., Tokarev A.V.* Applied software tools and services for rapid web GIS development // International Multidisciplinary Scientific GeoConference SGEM2015 Conference Proceedings, 2015, Book 2, 1, pp. 487-494.

[10] *Rinner C., Keßler C., Andrulis S.* The use of Web 2.0 concepts to support deliberation in spatial decision-making // Computers, Environment and Urban Systems, 2008, 32, pp. 386-395.

[11] *Shaparev N., Yakubailik O.* Usage of web mapping systems and services for information support of regional management // MATEC Web of Conferences, 2016, 79, 01081.

# ИНТЕГРИРОВАННЫЕ ГЕОИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И СИСТЕМЫ В ЗАДАЧАХ МОНИТОРИНГА

# ОНТОЛОГО-ОРИЕНТИРОВАННЫЙ ПОДХОД К ОРГАНИЗАЦИИ WEB-СРЕДЫ ПО АКТИВНОЙ СЕЙСМОЛОГИИ

*Брагинская Л.П., Григорюк А.П., Ковалевский В.В.*
Институт вычислительной математики и математической геофизики СО РАН, Новосибирск

В настоящее время активно развивается подход построения информационных систем, основанных на знаниях, в которых доступ к данным организуется на основе онтологий. Этот подход делает возможным комбинирование реляционной базы для хранения данных с онтологией, представляющей понятийную систему предметной области. В статье предложена технология организации web-среды для исследований в области активной сейсмологии, которая позволяет за счет онтологии представить пользователю содержимое источника данных на концептуальном уровне и адресовать пользовательский запрос к нескольким гетерогенным источникам данных. Высокая производительность выполнения запроса на доступ и анализ экспериментальных данных обеспечивается за счет того, что этим будет заниматься уже реляционная система управления базой данных.
*Ключевые слова: активная сейсмология, онтология, информационные системы, реляционные базы данных*

**Введение.** Активная сейсмология – направление в геофизике, в котором для изучения строения земной коры и исследования геодинамических процессов в зонах землетрясений и вулканов используются управляемые источники сейсмических волн – взрывные, гидромеханические и электромагнитные импульсные источники, мощные сейсмические вибраторы [1]. Возможность изучения глубинных исследований недр Земли с помощью мощных вибрационных источников была теоретически обоснована и практически реализована в 1970-х-80-х годах. Вибросейсмические методы исследований имеют важные преимущества по сравнению с методами пассивной сейсмологии: точно известное место и время действия источника, излучение сейсмического импульса заранее заданной формы, управление экспериментом на компьютерной основе, повсеместность применения и экологическая безопасность. Вибросейсмические исследования являются наукоемкой деятельностью, в которую входят:

- разработка аппаратных средств и программно-математического обеспечения управления техническими комплексами излучения и регистрации вибрационных сигналов;
- разработка теоретических основ метода, в том числе обратные задачи теории распространения упругих волн, прямые задачи моделирования волновых сейсмических процессов, разработка теории многодисциплинарного вибросейсмического мониторинга и др.;
- экспериментальные полевые исследования по регистрации волнового поля от мощных источников на значительных, несколько сотен километров, удалениях;
- разработка алгоритмов и программ для обработки экспериментальных данных;
- численное моделирование полных волновых полей в сложнопостроенной среде с применением суперкомпьютеров;
- разработка инженерно-сейсмологических технологий исследования зданий и крупных промышленных сооружений;

Результативность исследований в активной сейсмологии зависит не только от эффективной организации доступа к актуальным знаниям и данным, имеющим непосредственное отношение к данной предметной области, но и от интеграции разнородных данных и знаний специалистов смежных областей науки. Такое значимое направление исследований в области активной сейсмологии, как математическое моделирование полных волновых полей, связано с задачей выбора адекватных моделей земной коры, полученных различными геофизическими методами. Задачи интерпретации и экспертной оценки результатов экспериментов по активной сейсмологии требуют значительного объема справочных и иных текстовых материалов. Эффективное планирование полевых экспериментов по регистрации волнового вибросейсми-

ческого поля зависит не только от доступности данных наблюдения, если таковые уже проводились в районе предполагаемых работ, но и от учета комплекса дополнительных сведений о регионе, в этом случае особенно важны удобные пользовательские ГИС-сервисы.

При множестве неоднородных источников данных встает задача организации инфраструктуры, позволяющей не просто накапливать информацию для ее повторного использования в различных исследованиях, но и способной систематизировать знания и данные предметной области, обеспечивать содержательный доступ и предварительный анализ данных. В данной работе предложен подход к организации научной инфраструктуры, для хранения, представления, формализации и систематизации информации, а также для обеспечения высокой скорости доступа и анализа числовых данных. Предлагаемая авторами инфраструктура построена путем комбинирования онтологии предметной области для интеграции разнородных информационных ресурсов без их физического слияния и реляционной базы данных, которая обеспечивают эффективное хранение и обработку структурированных данных.

**Управление экспериментальными данными.** Экспериментальные работы по вибросейсмическому мониторингу проводятся ИВМиМГ СО РАН, начиная с 1985 г. В результате многолетних наблюдений (более 50 полевых экспериментальных работ) был накоплен уникальный массив данных (более 40 000 файлов волновых форм).

Для управления данными вибросейсмического мониторинга была разработана информационно-вычислительная система (ИВС) «Вибросейсмическое просвечивание Земли» [2], включающий архив файлов волновых форм и сопутствующей информации (тип сейсмического источника, параметры излучаемого им сигнала, параметры регистратора, географические координаты источника и регистратора и т.д.).

Ядром ИВС является реляционная база данных (БД). В БД хранится структурированная информация (рис. 1), описывающая экспериментальные данные. Таблица EXP содержит перечень экспериментов и их краткое описание. Каждый из экспериментов таблица EXP по вибросейсмическому просвечиванию Земли проводится при фиксированной расстановке задействованных в эксперименте источников сейсмических волн таблица IST и регистраторов таблица RGR. В ходе эксперимента проводятся сеансы работы виброисточников. Таблица SEANS содержит расписание (дату, время) работы источника. Таблица SIGNAL содержит набор параметров, описывающих свойства излучаемого сигнала. Каждому сеансу соответствует один кортеж отношения GENER и несколько кортежей (по числу регистраторов) отношения REGA. В соответствии с заданными отношениями таблицы GENER и GENA устанавливают однозначное соответствие конкретного сеанса с источником и параметрами сигнала, а таблица REGA соответствие сеанса с регистратором.

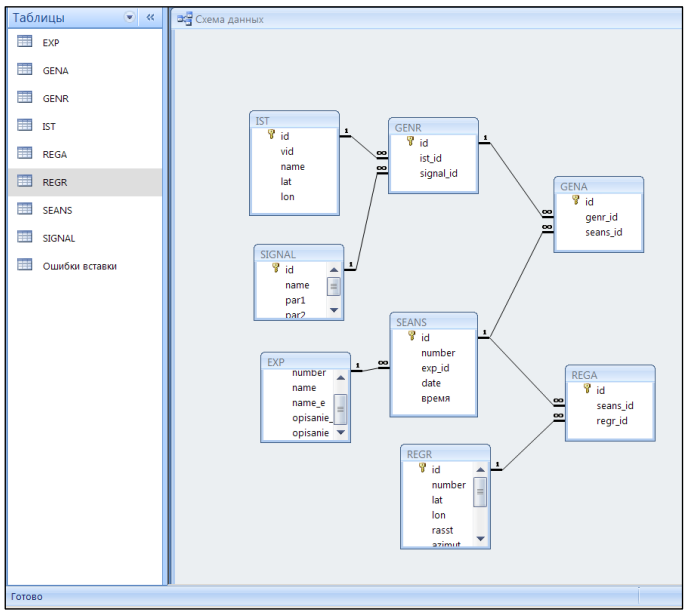



Рис. 1. Структура БД.

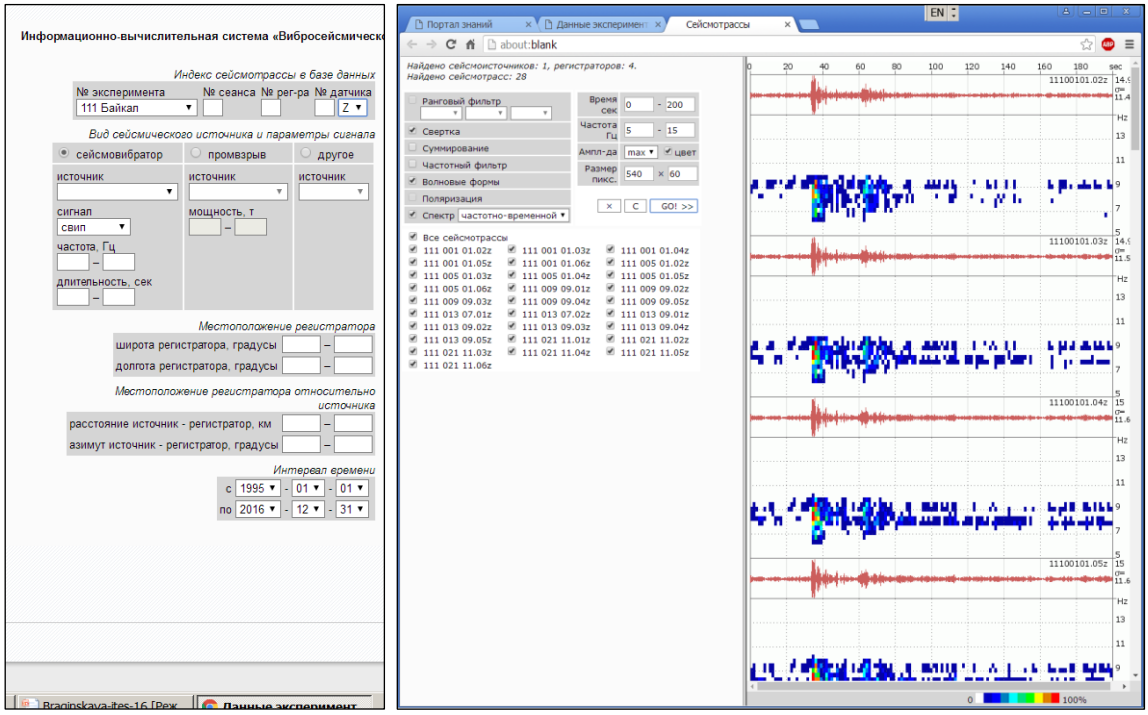



Рис. 2. Поисковая форма (слева) и результаты запроса (справа).

Каждая таблица, содержит одно ключевое поле, содержимое которого является уникальным для каждой записи данной таблицы. С помощью ключевого поля однозначно идентифицируются записи в таблице. В ответ на запрос пользователя, который может содержать любое значение интересующих параметров (до 20 параметров) ИВС формирует набор идентификаторов БД: EXP_id / SEANS_id / REGR_id.

Такая модель позволяет организовать адресацию файлового архива, имеющего следующую иерархическая структуру:

*/ номерРегистратора / номерСеанса / номерЭесперимента / <файлы волновых форм>*

Пользовательский запрос на анализ (рис. 2) должен содержать перечень процедур анализа, которые будут применены к найденным данным и параметры этих процедур.

В результате выполнения запросов данные извлекаются из файлового архива и передаются модулю анализа. Это приложение, выполняющее анализ данных в соответствии с алгоритмами, применяемыми в конкретной области экспериментальных исследований. Для обеспечения достаточного для online режима быстродействия приложение написано на языке C++ с использованием программных библиотек Intel Performance Libraries. Полученные в результате анализа числовые массивы возвращаются веб-приложению, которое «на лету» формирует графики, таблицы, текст и отправляет всё это пользователю в виде готовой веб-страницы.

**Интеграция знаний**. Интеграция знаний предметной области (ПО) «активная сейсмология» и навигации по интернет-пространству осуществляется путем построения портала знаний. Ядром информационной модели рассматриваемого портала является разработанная авторами онтология ПО «Активная сейсмология» [3].

Портал знаний [4] обеспечивает целостное представление знаний о предметной области, устанавливает взаимосвязи между относящимися к этой науке событиями, объектами, результатами и методами исследования и обеспечивает доступ к ним через интернет. Онтология портала вводит формальные описания понятий предметной области в виде классов объектов и отношений между ними, тем самым задавая структуры для представления реальных объектов и их связей. В соответствии с этим данные на портале представлены в виде семантической сети, т.е. как множество разнотипных взаимосвязанных информационных объектов. На рис. 3 представлен фрагмент онтологии активной сейсмологии, описывающий экземпляр «Эксперимент 111 “Байкал-Улан-Батор”» класса «Полевые эксперименты».

Содержательный доступ к систематизированным знаниям и информационным ресурсам обеспечивается с помощью предоставляемых порталом развитых средств навигации и поиска, функционирование которых также базируется на онтологии. На рис. 4 представлена страница Портала, соответствующая фрагменту онтологии рис. 3.

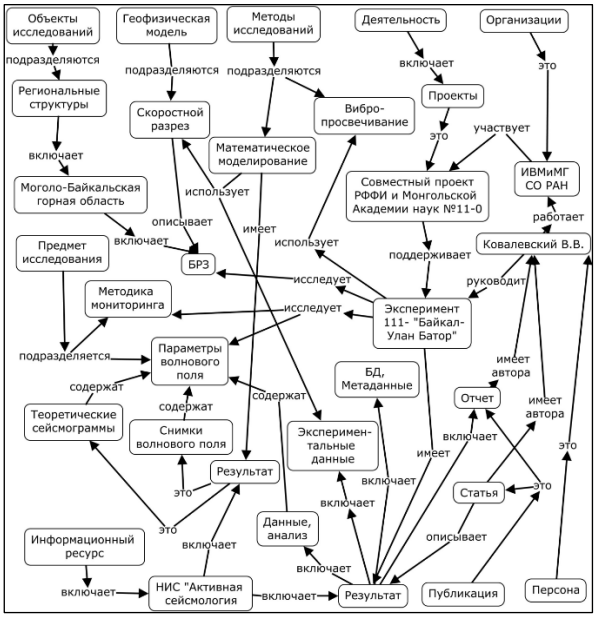
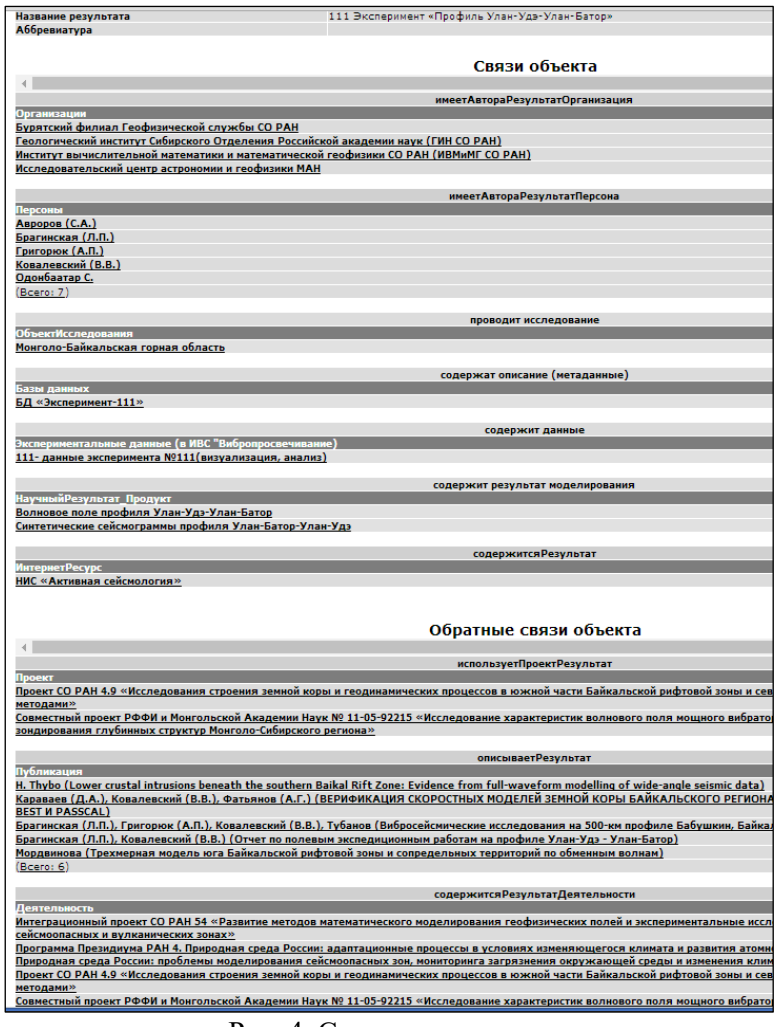


Рис. 3. Фрагмент онтологии.


Название результата 111 Эксперимент «Профиль Улан-Удэ-Улан-Батор»
Аббревиатура

Связи объекта

имеетАвтораРезультатОрганизация
Организации
Бурятский филиал Геофизической службы СО РАН
Геологический институт Сибирского Отделения Российской академии наук (ГИН СО РАН)
Институт вычислительной математики и математической геофизики СО РАН (ИВМиМГ СО РАН)
Исследовательский центр астрономии и геофизики МАН

имеетАвтораРезультатПерсона
Персоны
Авроров (С.А.)
Брагинская (Л.П.)
Григорюк (А.П.)
Ковалевский (В.В.)
Одонбаатар С.
(Всего: 7)

проводит исследование
ОбъектИсследования
Моголо-Байкальская горная область

содержат описание (метаданные)
Базы данных
БД «Эксперимент-111»

содержит данные
Экспериментальные данные (в ИВС "Вибропросвечивание")
111- данные эксперимента №111(визуализация, анализ)

содержит результат моделирования
НаучныйРезультат_Продукт
Волновое поле профиля Улан-Удэ-Улан-Батор
Синтетические сейсмограммы профиля Улан-Батор-Улан-Удэ

содержитсяРезультат
ИнтернетРесурс
НИС «Активная сейсмология»

Обратные связи объекта

используетПроектРезультат
Проект
Проект СО РАН 4.9 «Исследования строения земной коры и геодинамических процессов в южной части Байкальской рифтовой зоны и сев
методами»
Совместный проект РФФИ и Монгольской Академии Наук № 11-05-92215 «Исследование характеристик волнового поля мощного вибрато
зондирования глубинных структур Монголо-Сибирского региона»

описываетРезультат
Публикация
H. Thybo (Lower crustal intrusions beneath the southern Baikal Rift Zone: Evidence from full-waveform modelling of wide-angle seismic data)
Караваев (Д.А.), Ковалевский (В.В.), Фатьянов (А.Г.) (ВЕРИФИКАЦИЯ СКОРОСТНЫХ МОДЕЛЕЙ ЗЕМНОЙ КОРЫ БАЙКАЛЬСКОГО РЕГИОНА
BEST И PASSCAL)
Брагинская (Л.П.), Григорюк (А.П.), Ковалевский (В.В.), Тубанов (Вибросейсмические исследования на 500-км профиле Бабушкин, Байка
Брагинская (Л.П.), Ковалевский (В.В.) (Отчет по полевым экспедиционным работам на профиле Улан-Удэ - Улан-Батор)
Мордвинова (Трехмерная модель юга Байкальской рифтовой зоны и сопредельных территорий по обменным волнам)
(Всего: 6)

содержитсяРезультатДеятельности
Деятельность
Интеграционный проект СО РАН 54 «Развитие методов математического моделирования геофизических полей и экспериментальные иссл
сейсмоопасных и вулканических зонах»
Программа Президиума РАН 4. Природная среда России: адаптационные процессы в условиях изменяющегося климата и развития атомн
Природная среда России: проблемы моделирования сейсмоопасных зон, мониторинга загрязнения окружающей среды и изменения клим
Проект СО РАН 4.9 «Исследования строения земной коры и геодинамических процессов в южной части Байкальской рифтовой зоны и сев
методами»
Совместный проект РФФИ и Монгольской Академии Наук № 11-05-92215 «Исследование характеристик волнового поля мощного вибрато

Рис. 4. Страница портала.

Информационные объекты представлены на странице Портала гиперссылками. Портал знаний по гиперссылкам позволяет обратиться и к ИВС «Вибросейсмическое просвечивание Земли». Обращение к экспериментальным данным и средствам анализа производится внутри оболочки ИВС через оформление запроса (рис. 2).

**Заключение.** Разработанная web-среда по активной сейсмологии обеспечивает интеграцию тематических информационных ресурсов и содержательный доступ к результатам полевых и вычислительных экспериментов по активному вибросейсмическому мониторингу, интерактивный анализ данных, автоматическое построение интерактивных карт районов полевых работ. Взаимосвязь между деятельностью исследователей, результатами этой деятельности, персонами и организациями, осуществляющими исследования в области активной сейсмологии, обеспечивает Портал знаний, ядром которого является онтология, построенная группой экспертов, работающих в различных направлениях активной сейсмологии. На основе онтологии организуется удобная навигация по научным знаниям, а также содержательный поиск данных и средств их анализа. Интернет-ресурс доступен по адресу http://opg.sscc.ru.

## ЛИТЕРАТУРА


[1] *Alekseev A.S., Glinsky B.M., Kovalevsky V.V., Khairetdinov M.S.* Active vibromonitoring: experimental systems and fieldwork results. Handbook of Geophysical Exploration: Seismic Exploration Active geophysical monitoring, Elsevier Science, 2010, pp. 55-71

[2] *Брагинская Л.П., Григорюк А.П., Ковалевский В.В.* Научная информационная система «Активная сейсмология» для комплексных геофизических исследований // Вестник КРАУНЦ, Науки о земле, 2015, № 1, Выпуск № 25, - С. 94-98.

[3] *Боровикова О.И., Брагинская Л.П., Загорулько Ю.А., Ковалевский В.В.* Онтология предметной области «Активная сейсмология» // Материалы Всероссийской конференции с международным участием «Знания-Онтологии-Теории» (ЗОНТ-2015), Новосибирск, 2015, Т.1. С. 39-43.

[4] *Загорулько Ю.А., Боровикова О.И.* Информационная модель портала научных знаний // Информационные технологии. – 2009. – № 12. –С.2-7.

# ИНТЕРАКТИВНАЯ СРЕДА МОДЕЛИРОВАНИЯ ПРОЦЕССОВ ОБРАБОТКИ И АНАЛИЗА ДАННЫХ ДЗЗ

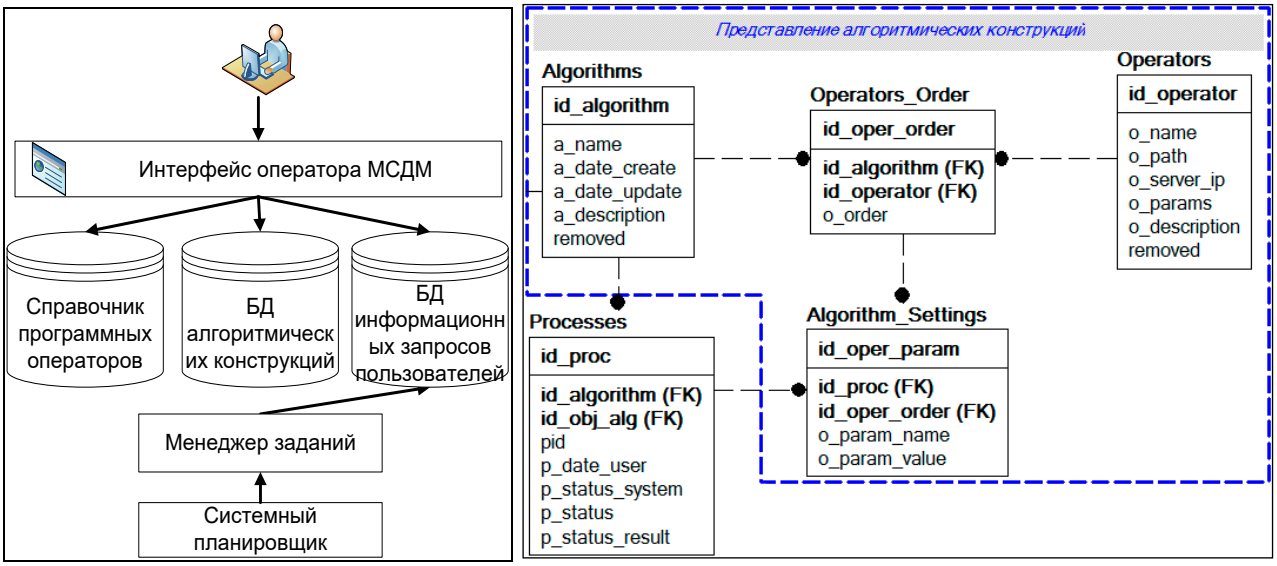
*Брежнев Р.В., Маглинец Ю.А., Раевич К.В.*
Институт космических и информационных технологий СФУ, Красноярск

В работе представлена концепция интерактивной технологии моделирования процессов обработки и анализа данных ДЗЗ, разработанная в Институте космических и информационных технологий СФУ. Технология позволяет интерпретировать модели процессов в реальные алгоритмические конструкции обработки данных в локальных и распределенных условиях, а также использовать их для автоматизации разрешения информационных запросов конечных пользователей в многоцелевой системе дистанционного мониторинга ИКИТ СФУ (МСДМ).
*Ключевые слова: дистанционное зондирование, информационные системы, обработка спутниковых данных, моделирование процессов, интерактивная среда моделирования, алгоритмическая последовательность, информационный запрос, конечный пользователь.*

**Введение.** Развитие международной группировки спутников дистанционного зондирования (ДЗ) Земли существенно расширяет спектр задач, решаемых с применением данных ДЗ. Это повышает востребованность геопространственных данных и сервисов в государственном управлении, коммерческой деятельности и среди населения. Изменяется концепция программных систем дистанционного мониторинга, которые становятся многоцелевыми. Для многоцелевых систем дистанционного мониторинга (МСДМ) характерна ориентация на конечного потребителя, предполагающая наличие развитых средств постановки и решения задач мониторинга пространственных объектов, минимизирующая роль специалистов в обработке данных и экспертов в их интерпретации, как необходимых участников процесса решения задачи. Такая возможность определяется следующими факторами: «мощностью» системы обработки пространственных данных, характеризующейся числом и разнообразием программных моделей обработки, анализа, интерпретации данных, достаточным для решения пула типовых задач в той или иной предметной области и степени ее «разумности», характеризующейся переносом экспертных знаний в базу знаний системы.

Количество и вариабельность алгоритмов обработки растровых и векторных пространственных данных, разработанных на сегодня мировым сообществом и воплощенных в соответствующих пакетах прикладных программ, таких как ENVI, Erdas Imagine, ScanEx Image Processor, QGIS и др. является достаточным для решения множества типовых задач. Дополнительно, алгоритмы и программы постоянно совершенствуются в ответ на развитие новых технических средств сбора пространственной информации, повышается точность вычислений. Однако большинство существующих пакетов программ ориентировано на использование специалистами в области обработки изображений.

Среди решений, ориентированных на конечного пользователя, следует отметить развитие web-сервисных архитектур, предоставляющих возможность решения типовых задач. Однако у конечных пользователей зачастую возникают индивидуальные требования, подразумевающие специфический порядок ввода данных, их обработки и представления результатов. Учет специфики решаемых тематических задач, как правило, ведет к необходимости построения алгоритмических последовательностей различной степени сложности на основе комбинации известных алгоритмов обработки и анализа данных. При этом решение новых задач требует не только привлечения специалиста по обработке данных для выстраивания логики процесса обработки, что вполне естественно, но и разработчика для программирования соответствующих последовательностей. Такой подход лишен гибкости, поскольку логика функционирования системы жестко зафиксирована в ее внутренней реализации и расширение функциональных возможностей требует по сути нового цикла разработки.




*а* *б*

Рис. 1. *а* – структура программного комплекса моделирования процессов; *б* – модель представления алгоритмических конструкций.

Этот факт актуализирует исследования, направленные на разработку унифицированных технологических платформ, предоставляющих специалисту в обработке данных диалоговые средства для гибкого формирования и отладки «цепочек» операторов, представляющих алгоритмические конструкции, решающих некоторую задачу [1-3] и при этом, позволяющих исключить внедрение в исходный код программных операторов обработки данных, что с позиции концепции построения технологии является неприемлемым. Примером реализации такого подхода может быть программный пакет Sentinel Application Platform (SNAP) от Европейского космического агентства (ESA) или российский продукт Image Media Center (IMC) от Центра инновационных технологий.

**Структура интерактивной среды моделирования процессов.** Основные концепции построения систем мониторинга пространственных объектов обсуждаются, в частности, в [4-8]. Рассмотренные в указанных работах системы обладают различной степенью унификации и в большинстве случаев требуют системного анализа и опыта программирования для адаптации под новые классы мониторинговых задач.

В качестве альтернативы, в настоящей статье представлена архитектура программного комплекса, обеспечивающего интерактивную среду для моделирования процессов обработки и анализа данных ДЗЗ (рис. 1,*а*). В нее входят рассмотренные ниже компоненты, представленные на концептуальном уровне.

1. ***Справочник программных операторов*** (таблица «`Operators`», рис. 1,*б*) хранит список операторов, доступных специалисту при проектировании алгоритмических конструкций. Оператор представляет собой либо скомпилированную библиотеку, либо программный модуль, который может быть реализован на любом интерпретируемом языке (Python, Perl, PHP и др.). Для того чтобы функционировать в составе алгоритмических конструкций, программные модули должны обладать следующими свойствами:
   – Автономность – одно из основных свойств функционирования программного оператора, которое характеризует способность выполнять им заданного набора функций без участия других операторов системы, что позволяет рассматривать его как самостоятельный элемент цепочки управления. При этом оператор рассматривается как программный интерфейс для взаимодействия пользователя – специалиста по настройке и объектной моделью системы, представленной совокупностью библиотек. Таким образом, автономность не приводит к дублированию функций разными

операторами. Программный интерфейс оператора может быть реализован в виде командной строки или графического интерфейса для восприятия настроек пользователя (рис. 3).

– Управляемость. Свойство подразумевает ограниченное множество входных $I_A$ и выходных $O_A$ параметров, частично означиваемых специалистом на этапе построения алгоритмической конструкции, а частично определяемых автоматически в ходе получения промежуточных результатов обработки. Входные параметры определяют сценарий поведения автономного оператора, следовательно, что сам оператор должен выстраивать логику своего функционирования в зависимости от состава параметров и от их возможных значений.

Параметры определяются по принципу «ключ – значение», где в качестве значений помимо обычных числовых или строковых данных, могут использоваться, например, SQL, SPARQL-конструкции для передачи информации о промежуточных результатах работы от предыдущего $i$-го оператора к следующему $i$+1 или $i$+$n$ ($i = \overline{1,n}$) в условиях, когда результат априори неизвестен. Так же в параметры могут учитывать поиск файлов, поиск по содержанию файлов, регулярные выражения и т.д.

– Устойчивость. Свойство, которое характеризует способность оператора корректно функционировать при неверно заданном составе или последовательности параметров. Это значит, что порядок перечисления параметров не должен быть жестко определен, а появление в списке лишнего или несуществующего параметра не должно привести к отказу работы оператора. В ситуации, при которой нужный параметр был пропущен, оператор должен сообщить об этом в диалоге с пользователем.
– Защищенность. Свойство, характеризующее способность оператора противостоять преднамеренным или неумышленным деструктивным действиям пользователя.
– Понятность характеризует наличие хорошо документированного и структурированного описания оператора, в частности, его функциональные возможности, состав и назначение входных параметров, комбинации параметров и сценариев поведения.

2. ***База данных алгоритмических конструкций.*** Каждая алгоритмическая конструкция характеризуется своим специфическим множеством параметров, обобщенно которые можно представить выражением:

$$A =< I_A, O_A, R_A, F_A >, \quad (1)$$

где $I_A$ – множество входных параметров алгоритма $A$; $O_A$ – множество выходных данных; $F_A$ – множество функций алгоритма; $R_A$ – множество отношений с другими алгоритмами, которые стоит учитывать, поскольку $F_{A_i} \cap F_{A_j}$ и на физическом уровне это свойство характеризует модульный подход к построению алгоритмов, который в свою очередь реализует принцип повторного использования программного кода или принципы объектно-ориентированного подхода.

Наиболее корректным и гибким подходом к представлению знаний системы о структуре и составе алгоритмических конструкций является реляционный подход, который позволяет связать операторы в логические последовательности, означить элементы множества $I_A$ и хранить неограниченное множество таких последовательностей. На рисунке 1, б представлен фрагмент БД, предназначенной для представления алгоритмических конструкций. Модель включает следующие таблицы:

– «`Algorithms`» – справочник созданных алгоритмических конструкций.
– Таблицы «`Operators_Settings`» и «`Operators_Order`» являются связующими сущностями для установления соответствий между алгоритмами, входящими в их состав операторами и их параметрами. Они определяют множества последовательностей операторов, их параметры и порядок в конструкции.

3. ***Интерфейс оператора МСДМ.*** Предоставляет специалисту по настройке диалоговые средства для визуального построения модели процесса и его инициализации и запуска

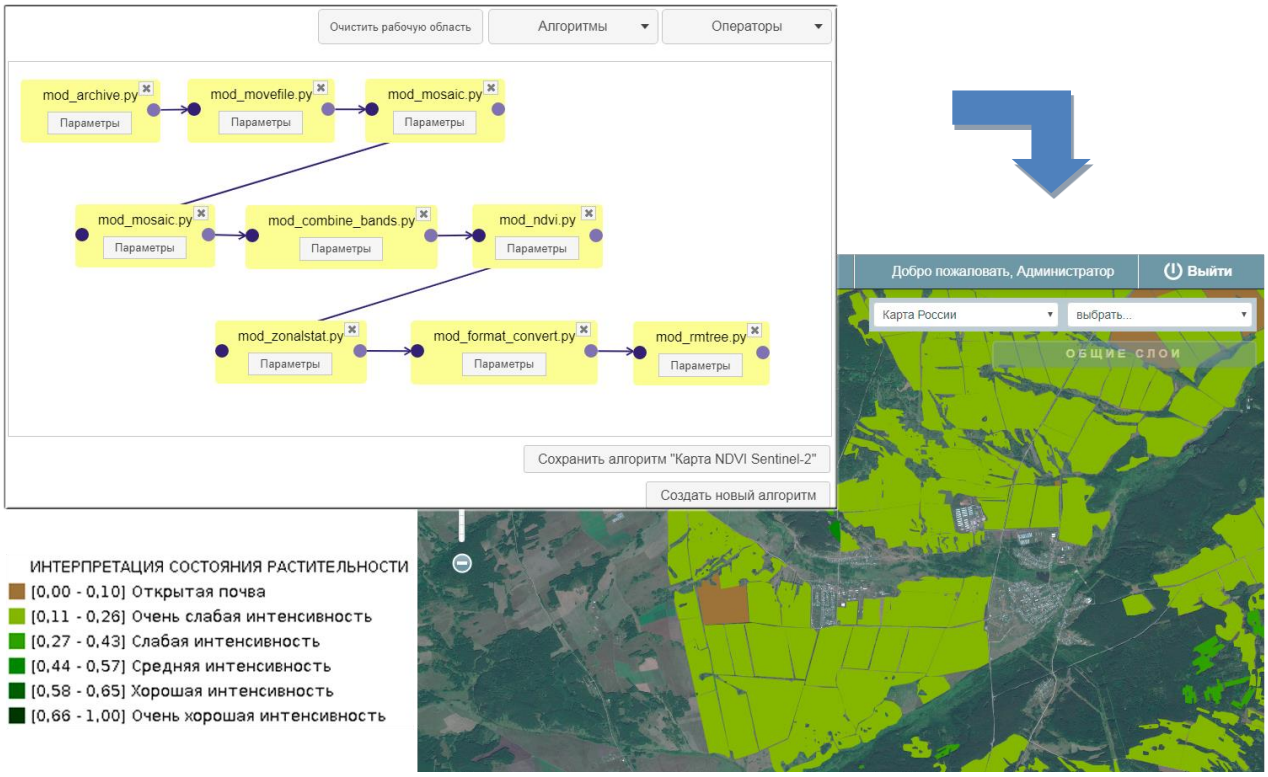



Рис. 2. Пример модели алгоритмической конструкции и результатов ее выполнения.

(рис. 2). Интерфейс так же учитывает возможности управления справочником операторов, алгоритмическими конструкциями и расписанием выполнения заданий, а именно, добавление, редактирование, удаление строк и установления связей между информационными запросами пользователей и подготовленными алгоритмическими конструкциями по разрешению запросов.

4. ***База данных информационных запросов пользователей*** представляет собой фрагмент БД (таблица «`Processes`»), предназначенный для фиксации поступающих запросов конечных пользователей на обработку данных. При этом запросы могут быть двух типов: единичные и циклические. Под единичными понимаются запросы разового расчета некоторых заданных параметров, например, расчет площади. Циклические запросы обеспечивают многократный расчет заданных параметров для заданных объектов на определенном интервале времени, что соответствует задачам мониторинга. Таким образом, модель информационного запроса пользователя позволяет представить запрос в виде плана мониторинга, описывающегося определенными статусами и связать постановщика задачи, группу объектов и соответствующую алгоритмическую конструкцию (рис. 1,*б*).
5. ***Менеджер задач*** представляет собой программный модуль, предназначенный для запуска на исполнение алгоритмических конструкций в соответствии с временным планом исполнения. Активированная алгоритмическая конструкция, с точки зрения операционной системы, представляет системный процесс, имеющий различные статусы. Менеджер позволяет интерпретировать системные статусы и изменять их. Поскольку требование многозадачности подразумевает, что операторы разных алгоритмических конструкций могут выполняться в параллельном режиме, то менеджер задач контролирует системные ресурсы, которые априори ограничены. Для этого, с одной стороны, для операторов конструкции предусмотрена технология очередей (Queue) потоков (Thread), чтобы каждый оператор выполнялся строго в соответствии с порядком, заданным специалистом. С другой стороны, подсистема управления контролирует системную память и не допускает одновременного выполнения более 5 конструкций.
6. ***Планировщик.*** Системный планировщик производит периодический старт менеджера

задач с частотой 1 раз в 1 минуту. Альтернативным подходом к решению задачи автоматического запуска программных операторов без использования планировщиков является создание службы, непрерывно просматривающей статусы заданий [9].

**Экспериментальная апробация технологии.** В качестве задачи, выбранной для экспериментальных исследований, выбрана задача мониторинга состояния яровой зерновой культуры в период вегетации. Оценка состояния основывается на периодических расчетах вегетационного индекса и формировании карт состояний, которые публикуются в web-интерфейсе МСДМ ИКИТ [10, 11].

Для разрешения данного запроса в системе подготовлена алгоритмическая конструкция, в состав которой входят следующие операторы, которые представлены в соответствующем порядке (см. рис. 2):

- *Модуль «Извлечение архива»*. В работе используются данные из открытого источника USGS, где каналы изображений распространяются в ZIP или TAR архивах. Поэтому первым действием необходимо извлечь архив в указанную ременную директорию. Модуль самостоятельно распознает тип архива. Используется пакетный режим, при котором все архивы за одну дату последовательно распаковываются в заданную директорию. Предусловием работы модуля являются заранее загруженные данные на заданную дату и область.
- *Модуль «Перемещение файлов»* используется для переноса исходных архивов в хранилище данных.
- *Модуль «Сбор мозаики»* предназначен для поканального сбора мозаик из отдельных сцен. Модуль определяет количество сцен заданного спектрального канала, если после извлечения архивов количество сцен одного канала более одного, то выполняется операция сбора мозаики, в противном случае – нет. Так, для расчета NDVI необходимо два канала: RED и NIR, поэтому модуль в конструкции используется дважды.
- *Модуль «Объединение каналов»* предназначен для объединения отдельных спектральных каналов в один многоканальный файл для оптимизации обработки.
- *Модуль «Расчет NDVI»* выполняет расчет спектрального индекса NDVI и создает канал со значениями NDVI.
- *Модуль «Зональная статистика»* реализует функции расчета статистических характеристик и их сохранения в одном из векторных форматов. Используется для расчета среднего значения NDVI ($\overline{N}$) для каждого объекта. Результаты расчетов сохраняет в файл формата GeoJSON, на основе которого создается карта состояния растительности (см. рис. 2).
- *Модуль «Преобразование форматов»* предназначен для преобразования одного растрового формата в другой. Используется для преобразования каналов спутника Sentinel из формата JPEG2000 в общепринятый формат GeoTIFF для последующей обработки другими процессами.
- *Модуль «Удаление файлов»* предназначен для рекурсивной очистки директорий. Используется для удаления временной директории с данными промежуточных результатов работы алгоритмической конструкции.

Описанный алгоритм позволяет построить карты состояния анализируемых культур по данным Sentinel-2A, B.

Моделирование процесса предусматривает всего несколько основных шагов: выбор и добавления на рабочую область программного оператора из списка и задание ему подмножества параметров (рис. 3); задание порядка выполнения операторов определяется стрелками (рис. 4).

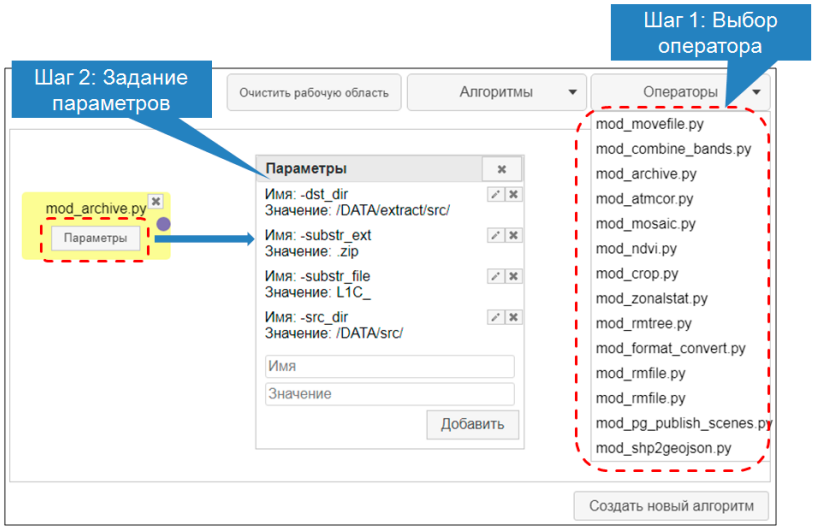



Рис. 3. Шаги 1 и 2 диалога моделирования процесса.

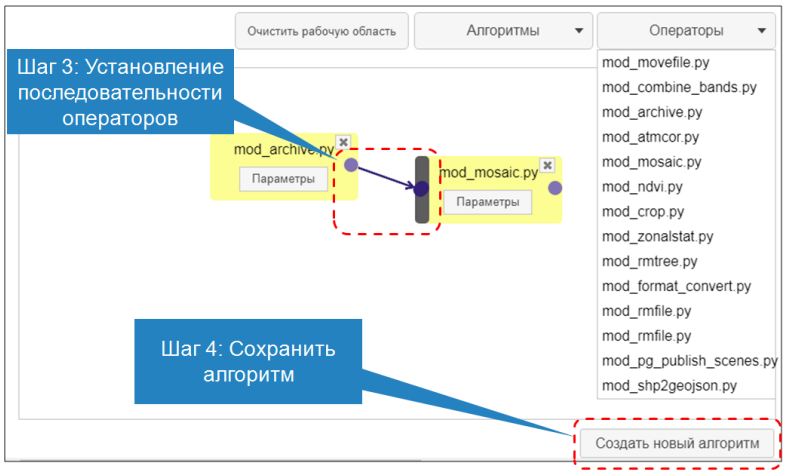



Рис. 4. Шаги 3 и 4 диалога моделирования процесса.

**Заключение.** Представленная в работе интерактивная среда моделирования процессов обработки и анализа данных ДЗЗ реализована как программно-технологический комплекс в рамках многоцелевой системы дистанционного мониторинга ИКИТ и прошла успешные тестовые испытания. На основе подготавливаемых в данной среде алгоритмических конструкций создается классификатор алгоритмов, который используется в разрабатываемом интерфейсе постановки задач [12] в виде операций доступных конечному пользователю над заданными пространственными объектами. Так же разработка позволяет организовывать вычислительные процедуры в распределенных условиях, когда, проиндексированные в едином справочнике, программные операторы физически располагаются на разных серверах.

*Исследование выполнено при финансовой поддержке РФФИ (грант № 18-47-242002-р_мк), Правительства Красноярского края и Красноярского краевого фонда науки (научный проект «Разработка технологии создания интеллектуальных информационных систем объектно-ориентированного мониторинга территорий по данным дистанционного зондирования»).*

## ЛИТЕРАТУРА

[1] *Лобзенёв В.Н.* Полный цикл обработки материалов ДЗЗ в ПК IMC // Труды Международной конференции «От снимка к карте: цифровые фотограмметрические технологии». Хайнань, Китай, 2014. С. 13-19.

[2] *Michael Foumelis, Jose Manuel Delgado Blasco, Yves-Louis Desnos and other* ESA SNAP – StaMPS Integrated Processing for Sentinel-1 Persistent Scatterer Interferometry // IGARSS. 2018. P. 1364–1367.
[3] *Брежнев Р.В., Перевалова А.А.* Технология графического построения процессов обработки и анализа данных дистанционного зондирования Земли // Труды Международной конференции «Региональные проблемы дистанционного зондирования Земли». Красноярск: Сиб. федер. ун-т, 2018. С. 40-43.
[4] *Матвеев А.А., Мамаев А.С., Прошин А.А., Флитман Е.В.* Организация контроля над функционированием распределенной системы ИСДМ Рослесхоз // Современные проблемы дистанционного зондирования Земли из космоса. Москва, 2009. Т. 6, № 2. С. 535-541.
[5] *Маглинец Ю.А., Брежнев Р.В.* Развитие средств автоматизации приёма и обработки спутниковой информации региональной системы ДЗЗ СФУ // Современные проблемы дистанционного зондирования Земли из космоса. Москва, 2011. Т. 8, № 3. С. 120-128.
[6] *Балашов И.В., Ефремов В.Ю., Мазуров-мл. А.А., Мамаев А.С., Матвеев А.М., Прошин А.А.* Особенности организации контроля и управления распределенных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. Москва, 2011. Т. 8, № 3. С. 161-166.
[7] *Лошкарев П.А., Тохиян О.О., Курлыков А.М., Кошкин К.В., Гладков А.П.* Развитие ЕТРИС ДЗЗ с применением облачных технологий // ГЕОМАТИКА. Москва, 2013. № 4. С. 22-26.
[8] *Brezhnev R.V., Maltsev E.A.* An Ontological Spatial Monitoring System for Agricultural Land Monitoring // Pattern Recognition and Image Analysis. 2015. Vol. 25, No. 2. P. 201–208.
[9] *Кашницкий А.В., Балашов И.В., Лупян Е.А., Толпин В.А., Уваров И.А.* Создание инструментов для удаленной обработки спутниковых данных в современных информационных системах // Современные проблемы дистанционного зондирования Земли из космоса. Москва, 2015. Т. 12. № 1. С. 156-170.
[10] *Brezhnev R.V., Maglinets Yu.A., Raevich K.V., Tsibulski G.M.* Modeling of Agricultural Spatial Objects with Heterogeneous Dynamically Changing Spatial Structure // CEUR Workshop Proceedings. Samara, Russia, 2018. Vol. 2210. P. 316–322.
[11] *Brezhnev R.V., Maglinets Yu.A.* The dynamic model of agricultural land structure on the space images in the precision agriculture tasks // E3S Web of Conferences. 2019. Vol. 75, 01001.
[12] *Brezhnev R.V., Maglinets Yu.A.* Information Support Technique for Solving Agricultural Land Monitoring Tasks Based on Earth Remote Sensing Data // Journal of Siberian Federal University. Engineering & Technologies. Krasnoyarsk, 2017. Vol.10. No. 6. P. 819-827.

# ПРИМЕНЕНИЕ КОМПЛЕКСА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ В ПРИКЛАДНЫХ СЕЛЕКЦИОННЫХ ЗАДАЧАХ

*Гребенникова И.Г.[(1)], Чешкова А.Ф.[(1)], Чанышев Д.И.[(1)], Стёпочкин П.И.[(1)(2)]*
[(1)] Сибирский федеральный научный центр агробиотехнологий РАН, Краснообск
[(2)] Сибирский НИИ растениеводства и селекции – филиал ИЦиГ СО РАН, Краснообск

В статье описано применение программного комплекса для информационной поддержки селекции зерновых культур. Рассмотрены основные используемые методы оценки селекционного материала: метод диаллельного анализа, метод оценки экологической пластичности сортов и линий по показателям интенсивности и устойчивости, интегральная оценка селекционной ценности методами скалярного ранжирования и статистического анализа. Приведены результаты тестирования программ на образцах яровой пшеницы и тритикале, позволившие произвести оценку коллекционных образцов и подбор родительских форм яровой тритикале для гибридизации.
*Ключевые слова: информационные технологии, компьютерная программа, зерновые культуры, селекция, диаллельный анализ, экологическая пластичность.*

**Введение.** Развитие эффективного сельскохозяйственного производства на сегодняшний день невозможно без внедрения новых технологических процессов производства и улучшения информационно-технологической базы при управлении этими процессами. Как правило, основным фактором эффективности сельскохозяйственного производства являются современные информационные технологии. Базовыми элементами новых информационных технологий являются компьютерные программы. Знания ведущих специалистов и современные методики производства сельскохозяйственной продукции отображаются в этих программах в виде математических моделей и методов обработки информации.

Создание новых высокоурожайных сортов зерновых культур требует изучения наследования признаков продуктивности растения, определения экологической приспособленности перспективных форм и линий, использования эффективных методов отбора лучших образцов по комплексу признаков. Использование компьютерных технологий в селекции зерновых культур обеспечивает информационное сопровождение селекционного процесса при решении задач классификации и прогнозирования.

**Результаты.** В СФНЦА РАН разработан комплекс компьютерных программ, предназначенный для информационно-аналитического сопровождения различных этапов селекции злаковых культур (рис. 1).




Рис. 1. Структура программно-алгоритмического комплекса.

В состав комплекса входят пакеты статистических, биометрико-генетических и селекционно-ориентированных программ, которые обеспечивают планирование полевых экспериментов, хранение полученных в результате опытов данных, оценку селекционного материала и проведение статистического анализа данных [1]. Приоритет разработки подтвержден свидетельствами о государственной регистрации программ для ЭВМ [2-5]. Тестирование комплекса осуществлялось на полученных опытным путём данных практических селекционных исследований основных хозяйственных показателей селекционных образцов. Материалом исследований служили сортообразцы яровой мягкой пшеницы сибирской селекции: Новосибирская 15, Новосибирская 31, Сибирская 12 и коллекционные формы яровой тритикале из коллекции ВИР: Укро (к-3644, Украина, Россия), Gabo (к-3722, Польша), Сокол Харьковский (к-3542, Украина), Dahbi 6/3/Ardi 1/Topo 1419 (к-3881, Мексика), Лт-F6-544-6 (к-3992, Россия), Kissa (k-3721, Мексика), а также селекционная форма F8: Сирс 57×Укро, полученная в результате диаллельных скрещиваний. Полевой опыт проводился на земельных участках СФНЦА РАН и СибНИИРС с 2009 по 2017 гг.

Исследуемые образцы существенно различаются по комплексу хозяйственно ценных признаков и принадлежат к разным группам спелости, отличаются полевой устойчивостью к листовым болезням, к пыльной и твёрдой головне, стеблевой и жёлтой ржавчине. По результатам исследований создана база данных, содержащая информацию об изучении образцов яровых тритикале и пшеницы по урожайности, качеству продукции, устойчивости к болезням, вредителям и другим неблагоприятным факторам – всего около 20 показателей.

Программный комплекс тестировался на пшенице и тритикале, однако после небольшой модификации его компоненты могут быть адаптированы для других сельскохозяйственных культур. Ниже приведено описание основных программ, входящих в комплекс, и результаты селекционной оценки исследуемых образцов яровой пшеницы и тритикале.

*Диаллельный анализ в селекции сельскохозяйственных культур «DIAS».* Программа создана на основе методических рекомендаций Р.А. Цильке и Л.П. Присяжной [6, 7] и предназначена для расчёта генетических параметров, комбинационной способности сортов зерновых культур и анализа исходного материала по количественным признакам на основе методов Гриффинга и Хеймана. Расчётные параметры позволяют оценить способности селекционного материала передавать трансгрессивное расщепление в потомстве диаллельных гибридов $F_1$, а также позволяют для конкретного селекционного материала оценить такие интегральные генетические параметры признака как среднюю степень и направление доминирования в полиморфных локусах, приблизительное число этих локусов, распределение желательных и нежелательных аллелей и др. Анализ параметров облегчает подбор оптимальной схемы селекции по ценным признакам.

Для изучения донорских способностей тритикале в 2009 г. проведена гибридизация по полной диаллельной схеме 4×4 сортов яровой тритикале Сокол Харьковский, Укро, Gabo, к–3881, а также этих сортов с озимым сортом Сирс57. С использованием программы *«DIAS»* был проведён диаллельный анализ и определены комбинационные способности исследуемых образцов тритикале по таким показателям, как число колосков в колосе, натура зерна, масса 1000 зёрен, длина колоса, число колосков в колосе [8-11]. На рис. 2 и 3 приведены результаты работы программы для признака «масса 1000 зёрен». Анализ показал, что сорт Укро может быть использован в линейной селеции для увеличения значения данного признака. Гибридную комбинацию Сокол×К-3881 целесообразно применять в селекции на гетерозис.

*Анализ экологической пластичности сельскохозяйственных культур.* В основу создания программы положена методика оценки экологической пластичности сортов и линий, разработанная Р.А. Удачиным [12]. Данная методика позволяет проводить оценку экологической пластичности образцов на разных этапах селекционного процесса при испытании их минимум на двух агрофонах методом случайных повторений и использует для этого показатели интенсивности и устойчивости индекса стабильности.

| Название сорта | Константы СКС (Sij) | | | | Эффекты ОКС (gi) | Вклад в дисперсию ОКС | Вклад в дисперсию СКС |
|---|---|---|---|---|---|---|---|
| | Габо | К-3881 | Сокол | Укро | | | |
| Габо | | | | | -3,3484 * | 8,6597 | 3,834 |
| К-3881 | 3,0135 * | | | | -0,1859 | -2,5178 | 17,6035 |
| Сокол | 2,5851 * | 48,0206 * | | | 1,1359 | -1,262 | 17,4038 |
| Укро | 5,9034 * | 1,7764 * | 1,6058 * | | 2,3984 * | 3,2002 | 3,0952 |

Рис. 2. Результаты анализа варианс общей и специфической комбинационной способности (* – оценка значима с вероятностью 95%).

| Компоненты изменчивости и их ошибки | | |
|---|---|---|
| D | 18,82 ± 9,08 | Аддитивный компонент изменчивости |
| F | 1,29 ± 23,32 | Отражает направление доминирования (в среднем по рядам в диаллел |
| H1 | 50,3 ± 26,39 | Обусловлен доминантными эффектами (положительными) |
| H2 | 45,29 ± 24,36 | Обусловлен доминантными эффектами (отрицательными) |
| h2 | 8,9 ± 16,49 | Алгебраическая сумма доминантных эффектов гетерозисных локусов |
| E | 3,69 ± 4,06 | Отражает доминантные и аддитивные эффекты для каждого ряда |

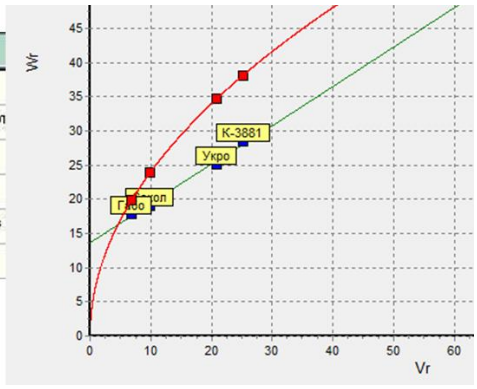



Рис. 3. Результаты расчёта генетических параметров.

| Сорт | Интенсивность,% | Классификация сортов по интенсивности | Устойчивость,% | Классификация сортов по устойчивости |
|---|---|---|---|---|
| К-3992 | 50,04 | Экстенсивный | 76,356 | Стабильный |
| Укро | 158,492 | Интенсивный | -299,977 | Стабильный |
| Kissa | 129,541 | Интенсивный | -128,584 | Стабильный |
| Сирс*Укро | 17,958 | Экстенсивный | 37,655 | Стабильный |

Рис. 4. Результаты оценки экологической пластичности.

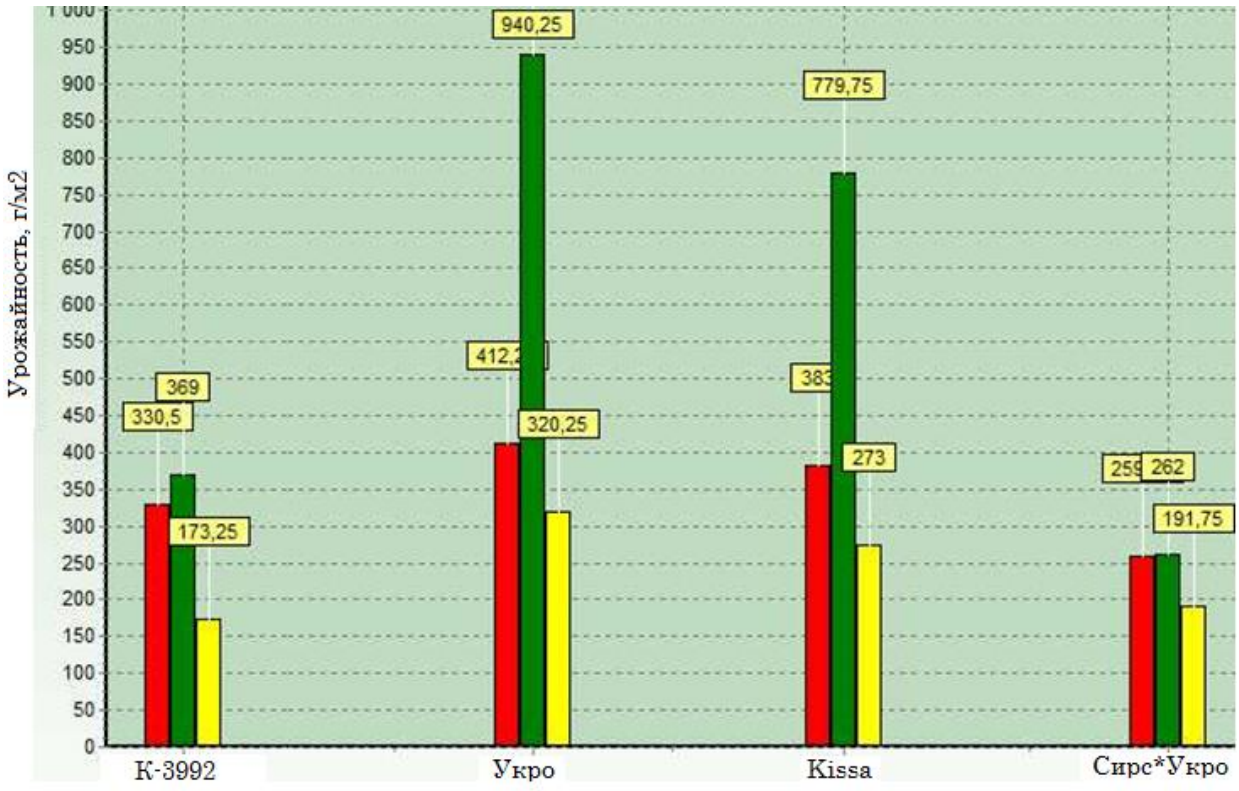



Рис. 5 Влияние агрофонов на урожайность.

Для тестирования программы было проведено экологическое испытание четырёх форм яровой тритикале на трёх агрофонах, различающихся по срокам сева и почвенным условиям. Результаты анализа (рис. 4 и 5) позволили произвести дифференцирование исследуемых образцов по их реакции на условия возделывания, а также установить критерии селекционной ценности при различном сочетании высокой урожайности со свойствами стабильность и интенсивность [13].

Все изучаемые сорта были оценены как стабильные. Наибольшие индексы стабильности отмечены у сортов Укро и Kissa, из чего следует, что они лучше приспособлены к местным условиям произрастания.

*Полевые опыты. Регистрация и оценка селекционного материала сельскохозяйственных культур.* Компьютерная программа предназначена для хранения данных, полученных в результате полевых опытов, оценки селекционной ценности культур по комплексу хозяйственно важных признаков, а также для обработки экспериментальных данных методами прикладной статистики. Программа обеспечивает выполнение следующих функций: ввод и редактирование справочной информации по изучаемым сортам и гибридам; ввод и редактирование данных, полученных в результате полевых опытов; расчёт селекционной ценности образцов на основе метода скалярного ранжирования; расчет стандартных статистических характеристик.

В связи с необходимостью сведения в один показатель ценности разнородных критериев селекционных объектов, предложен способ замены абсолютных исходных значений признаков на относительные балловые критерии. Результирующий сводный индекс представляет собой интегральную селекционную оценку исследуемого образца и количественно отображает относительную важность объекта в рассматриваемой выборке. Полученные значения ранжируются, располагая объекты в порядке степени их удовлетворения целям селекции [14]. На рис. 6 приведены результаты расчета индексов интегральной селекционной ценности по данным структурного анализа гибридов яровой тритикале $F_1$-$F_2$.

В интегральную селекционную оценку вошли 12 хозяйственно ценных признаков. На диаграммах представлены максимальный, минимальный и средний по повторностям индексы для каждого изучаемого гибрида и их родительских форм.

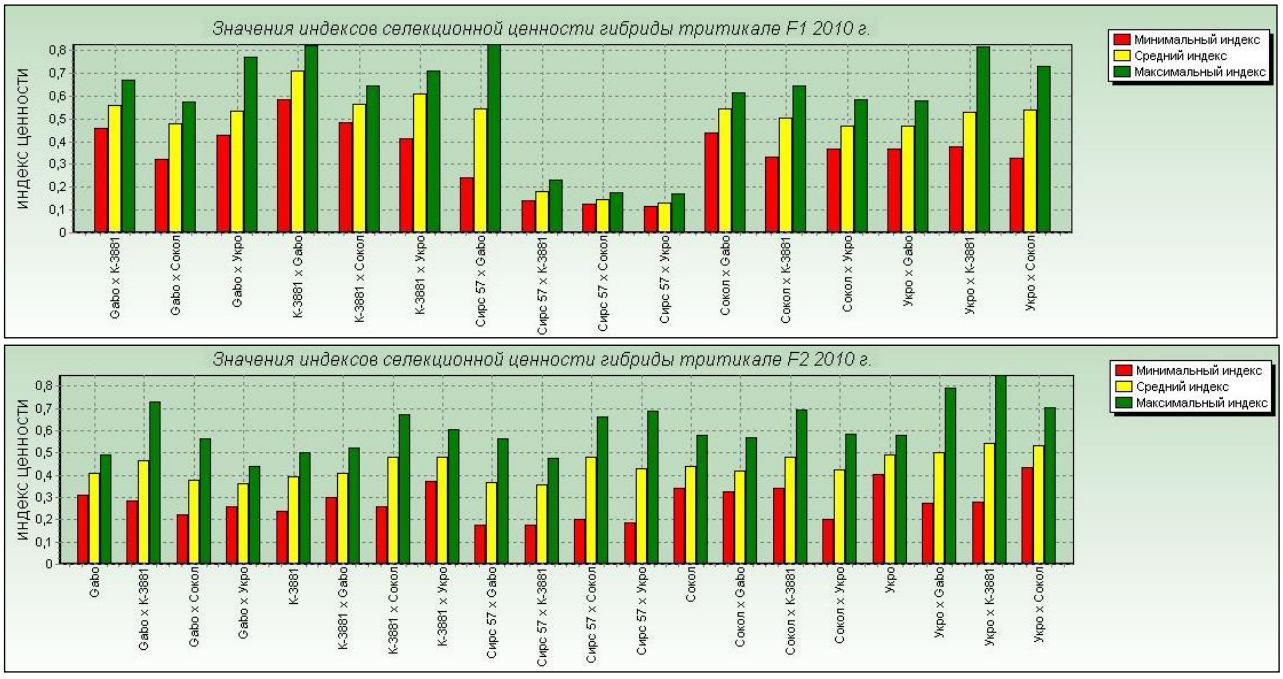
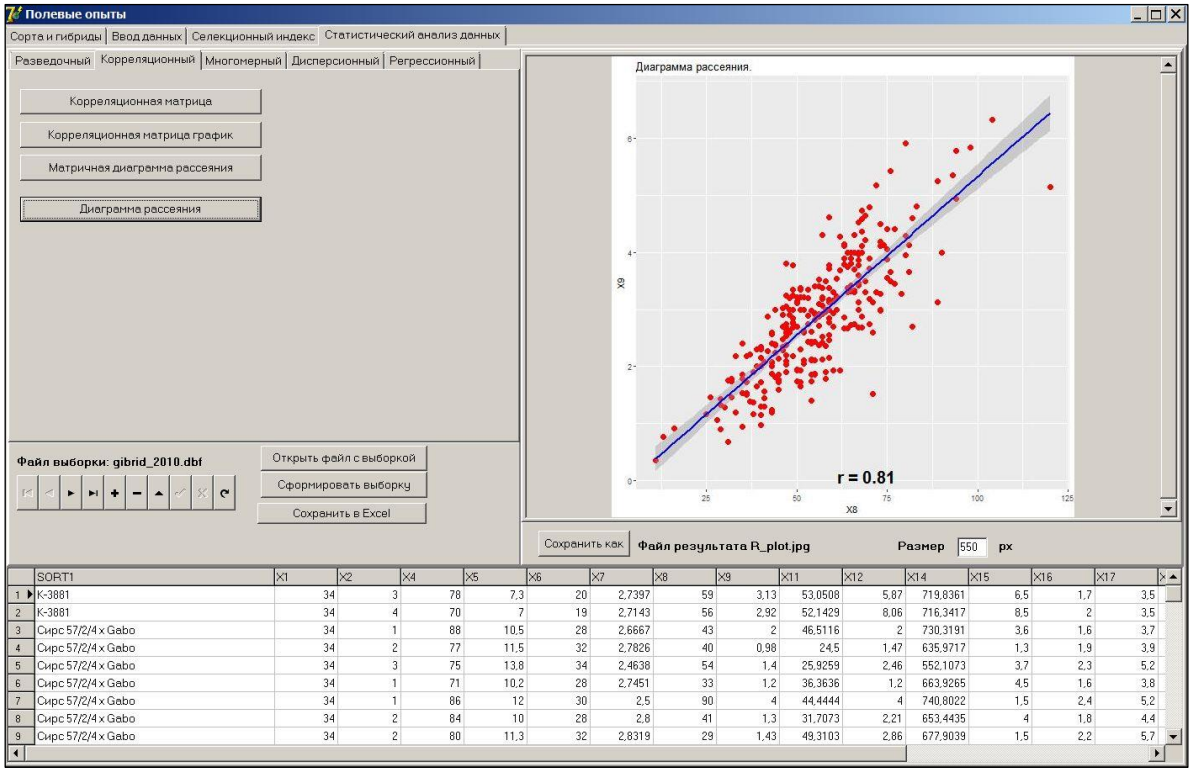



Рис. 6. Значения индексов селекционной ценности для гибридов $F_1$ (вверху) и $F_2$ (внизу).




| | SORT1 | X1 | X2 | X4 | X5 | X6 | X7 | X8 | X9 | X11 | X12 | X14 | X15 | X16 | X17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | K-3881 | 34 | 3 | 78 | 7,3 | 20 | 2,7397 | 59 | 3,13 | 53,0508 | 5,87 | 719,8361 | 6,5 | 1,7 | 3,5 |
| 2 | K-3881 | 34 | 4 | 70 | 7 | 19 | 2,7143 | 56 | 2,92 | 52,1429 | 8,06 | 716,3417 | 8,5 | 2 | 3,5 |
| 3 | Сирс 57/2/4 х Gabo | 34 | 1 | 88 | 10,5 | 28 | 2,6667 | 43 | 2 | 46,5116 | 2 | 730,3191 | 3,6 | 1,6 | 3,7 |
| 4 | Сирс 57/2/4 х Gabo | 34 | 2 | 77 | 11,5 | 32 | 2,7826 | 40 | 0,98 | 24,5 | 1,47 | 635,9717 | 1,3 | 1,9 | 3,9 |
| 5 | Сирс 57/2/4 х Gabo | 34 | 3 | 75 | 13,8 | 34 | 2,4638 | 54 | 1,4 | 25,9259 | 2,46 | 552,1073 | 3,7 | 2,3 | 5,2 |
| 6 | Сирс 57/2/4 х Gabo | 34 | 1 | 71 | 10,2 | 28 | 2,7451 | 33 | 1,2 | 36,3636 | 1,2 | 663,9265 | 4,5 | 1,6 | 3,8 |
| 7 | Сирс 57/2/4 х Gabo | 34 | 1 | 86 | 12 | 30 | 2,5 | 90 | 4 | 44,4444 | 4 | 740,8022 | 1,5 | 2,4 | 5,2 |
| 8 | Сирс 57/2/4 х Gabo | 34 | 2 | 84 | 10 | 28 | 2,8 | 41 | 1,3 | 31,7073 | 2,21 | 653,4435 | 4 | 1,8 | 4,4 |
| 9 | Сирс 57/2/4 х Gabo | 34 | 2 | 80 | 11,3 | 32 | 2,8319 | 29 | 1,43 | 49,3103 | 2,86 | 677,9039 | 1,5 | 2,2 | 5,7 |

Рис. 7. Программный интерфейс для расчета статистических параметров.

*Программа статистического анализа R.* В программе реализованы следующие процедуры: разведочный анализ данных (расчет базовых статистических показателей, проверка на нормальность распределения, график квантилей, диаграмма размахов, гистограмма, график плотности распределения); корреляционный анализ данных (корреляционная матрица, диаграмма рассеяния); метод главных компонент (матрицы и графики счетов и нагрузок, диаграмма главных компонент, проекции на компоненты); кластерный анализ данных (дендрограмма, график коэффициентов слияния, график кластеров в пространстве главных компонент, проверка устойчивости кластеров); дисперсионный анализ данных (анализ дисперсий и средних многофакторных моделей, проверка адекватности модели, критерий Тьюки для разности средних, график плотности остатков, групповая диаграмма размахов) (рис. 7).

Результаты статистических расчетов выводятся в текстовом или графическом виде.

**Выводы.** При помощи разработанного программно-алгоритмического комплекса из исходного материала выделены ценные в селекционном плане образцы яровой тритикале, адаптированные для возделывания в Западно-Сибирском регионе. Таким образом, применение программного комплекса позволяет повысить точность подбора родительских пар, конкретизировать направление селекционного процесса, выполнить дифференцирование исследуемых сортов по реакции на условия возделывания и получить оценку селекционной ценности изучаемых образцов по комплексу признаков. За счёт сокращения сроков и снижения затрат на выведения новых сортов повысится эффективность селекционного процесса зерновых культур.

*Работа поддержана СФНЦА РАН (бюджетный проект № 0778-2019-0001).*

## ЛИТЕРАТУРА

[1] *Алейников А.Ф., Чешкова А.Ф., Гребенникова И.Г., Стёпочкин П.И., Потанин В.Г., Чанышев Д.И.* Применение программно-алгоритмического комплекса для информационной поддержки селекции зерновых культур: методические указания. Новосибирск, 2017. 48 с.

[2] Диаллельный анализ селекции сельскохозяйственных культур: свидетельство об офиц. регистрации программы для ЭВМ Российская Федерация / Алейников А.Ф., Стёпочкин П.И., Гребенникова И.Г. № 2011613440; опубл. 25.04.2011.
[3] Анализ экологической пластичности сельскохозяйственных культур: свидетельство об офиц. регистрации программы для ЭВМ Российская Федерация / Алейников А.Ф., Стёпочкин П.И., Гребенникова И.Г., Чешкова А.Ф. № 2013611494; опубл. 22.01.2013.
[4] Интегральная селекционная оценка сельскохозяйственных культур: свидетельство об офиц. регистрации программы для ЭВМ Российская Федерация / Чешкова А. Ф., Алейников А.Ф., Стёпочкин П.И., Гребенникова И.Г. № 2013661141; опубл. 29.11.2013.
[5] Полевые опыты. Регистрация и оценка селекционного материала сельскохозяйственных культур. Версия 2.0: свидетельство об офиц. регистрации программы для ЭВМ Российская Федерация / Чешкова А. Ф., Алейников А.Ф., Стёпочкин П.И. № 2017613267; опубл. 29.03.2017.
[6] *Цильке Р.А., Присяжная Л.П.* Методика оценки исходного материала по комбинационной способности в диаллельных скрещиваниях. Новосибирск, 1979. 29 с.
[7] *Цильке Р.А., Присяжная Л.П.* Методика диаллельного анализа исходного материала по количественым признакам. Новосибирск, 1979. 13 с.
[8] *Гребенникова И.Г., Алейников А.Ф., Стёпочкин П.И.* Диаллельный анализ числа колосков в колосе яровой тритикале // Сибирский вестник с.-х. науки. 2011. № 7,8. С. 77-85.
[9] *Grebennikova I.G., Aleynikov A.F., Stepochkin P.I.* Diallel analysis of the number of spikelets per spike in spring triticale // Bulgarian Journal of Agricultural Science. 2011. Т. 17. № 6. С. 755-759.
[10] *Гребенникова И.Г., Алейников А.Ф., Степочкин П.И.* Диаллельный анализ длины колоса у яровой тритикале // Сибирский вестник сельскохозяйственной науки. 2010. № 12. С. 103-109.
[11] *Гребенникова И.Г., Алейников А.Ф., Стёпочкин П.И.* Компьютерная программа обеспечения селекционного процесса зерновых культур (на примере тритикале) // Ползуновский вестник. 2011. № 2/2. С. 128-133.
[12] *Удачин Р.А.* Методика оценки экологической пластичности сортов пшеницы // Селекция и семеноводство. 1990. № 5. С. 2-6.
[13] *Гребенникова И.Г., Алейников А.Ф., Степочкин П.И.* Анализ экологической пластичности тритикале // Сибирский вестник сельскохозяйственной науки. 2013. № 3. С. 101-106.
[14] *Чешкова А.Ф., Алейников А.Ф., Стёпочкин П.И., Гребенникова И.Г.* Компьютерная программа «Интегральная селекционная оценка сельскохозяйственных культур» // Достижения науки и техники АПК. 2013. № 12. С. 69-71.

# ИСПОЛЬЗОВАНИЕ МЕТОДА ПОСЛЕДОВАТЕЛЬНОГО КВАДРАТИЧНОГО ПРОГРАММИРОВАНИЯ ДЛЯ РАЗДЕЛЕНИЯ СМЕСЕЙ ОБЪЕКТОВ НА ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЯХ

*Григорьева О.В., Саидов А.Г., Чапурский Л.И.*
Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург

Рассматривается алгоритм классификации элементов ландшафта и идентификации малоразмерных объектов по гиперспектральным изображениям с помощью определения доли этих объектов в пикселе методом последовательного квадратичного программирования. Приводится пример практической реализации алгоритма в задаче определения проективного покрытия виноградников.
*Ключевые слова: гиперспектральные данные, смеси объектов, последовательное квадратичное программирование, спектральные компоненты*

**Введение.** Гиперспектральные (ГС) данные за счет высокого спектрального разрешения приобретают все большую популярность в изучении спектральных свойств объектов земной поверхности, как для их классификации и обнаружения, так и для проектирования перспективных многоспектральных средств дистанционного зондирования Земли, обеспечивая селекцию наиболее информативных спектральных каналов для решения тематических задач в различных областях исследования. Однако их невысокое пространственное разрешение не всегда однозначно позволяет идентифицировать какой-либо фрагмент изображения, поскольку пиксель содержит не один объект, а их смесь. Особенно это наблюдается в случаях, если размер объекта меньше, чем размер пикселя на местности, или при расположении пикселя на границе двух объектов. В связи с этим в этих случаях актуальной является задача разделения смеси объектов на ГС изображениях.

Зачастую допускается, что спектральная смесь объектов линейная:

$$R = M \cdot S + \varepsilon, \quad (1)$$

где $M = \{m_{ij}\}$, $i = \overline{1,l}$, $j = \overline{1,p}$ – матрица смешивания, каждый столбец которой содержит спектральный вектор конечных членов (объектов), $l$ – количество спектральных каналов, $p$ – количество объектов; $S = \{s_{ij}\}$, $i = \overline{1,p}$ и $j = \overline{1,n}$ – матрица, столбцы которой являются относительными распространенностями объектов, имеющих спектральные сигнатуры $M_j$, то есть элементы $s_{ij}$ являются вероятностью отнесения $j$-го пикселя к спектральной сигнатуре $M_j$, $n$ – количество пикселей в анализируемом фрагменте ГС изображения; $R = \{r_{ij}\}$, $i = \overline{1,l}$ и $j = \overline{1,n}$ – матрица, столбцы которой являются спектральными векторами анализируемого фрагмента ГС изображения; $\varepsilon$ – доля аддитивного шума.

При этом на коэффициенты смеси накладываются ограничения (сумма коэффициентов полагается равной единице, каждый из коэффициентов должен быть неотрицательным):

$$s_{ij} \geq 0 \text{ и } \sum_{j=1}^{p} S_j = 1. \quad (2)$$

Для того, чтобы определить доли объекта (фракции) в пикселе S, используют методы линейного спектрального разложения (англ. linear spectral unmixing): метод минимизации среднеквадратического отклонения, метод ортогональной проекции подпространства и другие, различающиеся по критериям и методам оптимизации [1, 2]. При этом часть из них не учитывает все ограничения, накладываемые на коэффициенты разложения смеси, что приводит к результатам расчета, которые трудно интерпретировать.

Возникающая задача минимизации величины среднеквадратического отклонения линейной смеси сигнатур от истинного значения пикселя ГС данных с учетом полных ограничений (Fully Constrained Linear Square, FCLS) относится к классу задач нелинейной оптимизации с ограничениями и может решаться с помощью симплекс-метода, методов проекционного градиента, обобщенного приведенного градиента, линеаризации, штрафов и лагранжевой двойственности.

В данной работе предлагается решать задачу определения коэффициентов разложения S с помощью метода последовательного квадратичного программирования (англ. sequential quadratic programming, SQP). SQP является одним из самых современных методов в области нелинейного программирования, позволяющим учесть полные ограничения, накладываемые на коэффициенты разложения смеси [3]. При этом предполагается наличие информации о спектральных сигнатурах объектов M, присутствующих на изображении, в том числе и искомого малоразмерного объекта. Для формирования массива исходной априорной спектральной информации M могут быть использованы различные алгоритмы endmember extraction: симплекс-метод минимального объема MVSA, N-FINDR, анализ независимых и зависимых компонент (ICA и DCA) и др. [4-6]. В работе [7] выделение спектральных компонент массива ГС данных также осуществляется с помощью метода SQP, для чего предварительно ставится задача оптимизация симплекса минимального объема к набору спектральных векторов R ГС изображения и используются специальные методы анализа формы спектральных сигнатур.

**Алгоритм разделения смесей объектов на основе метода SQP.** В алгоритме задача спектрального разделения смеси описывается как задача минимизации среднеквадратического отклонения $\varepsilon_i$ оцениваемой смеси фракций к исходным ГС данным. Для того, чтобы перейти в область задач квадратичного программирования, целевая функция в соответствии с уравнением (1) была переписана следующим образом:

$$f = \sum_{i=1}^{n} {\varepsilon_i}^T \varepsilon_i = \sum_{i=1}^{n} (R_i - MS_i)^T (R_i - MS_i).$$

Решение данной задачи с учетом вышеперечисленных ограничений на компоненты смеси предлагается с помощью метода SQP. Основным затруднением при реализации такого метода является необходимость формализации условий ограничений в виде линейных равенств и неравенств:

$$\mathrm{AS} \leq \mathrm{b} \text{ и } \mathrm{A}_{eq}\mathrm{S} = b_{eq}.$$

В соответствии с этим, условие равенства были преобразованы следующим образом:

$$\mathrm{A}_{eq} \cdot \mathrm{vec(S)} = b_{eq},$$

где $\mathrm{A}_{eq} = I_n \otimes (1_\mathrm{p})^\mathrm{T}$, $b_{eq} = (1_\mathrm{n})^\mathrm{T}$, vec(S) – вектор, элементы которого являются элементами столбцов матрицы *S* друг за другом; *I* – единичная матрица; ⊗ – оператор тензорного произведения Кронекера для матриц.

Условие неравенства было приведено к виду:

$$-\mathrm{A} \cdot \mathrm{vec(S)} \leq b,$$

где $\mathrm{A} = I_{p*n}$ и $\mathrm{b} = (0_{p*n})^T$.

Пиксель j идентифицировался как искомый объект *i*, если коэффициент разложения $\mathrm{s}_{ij}$ превышал определенный порог (например, 0.95).

**Заключение.** Реализация разработанного алгоритма разделения смеси апробировалась по данным космической ГС съемки, полученным космическим аппаратом Ресурс-П на территорию виноградников Крымского полуострова. На рисунке показан результат обработки этих данных в виде карты, отражающей проективное покрытие виноградников. В качестве исходных данных были заданы спектральные сигнатуры почвы и виноградников с разными индексами листовой поверхности LAI (от 1 до 2) в диапазоне 350…2500 нм, зарегистрированные в ходе натурных измерений с помощью полевого спектрорадиометра FieldSpec®4 в период съемки с космического аппарата. Полученные в результате работы алгоритма коэффициенты разложения смеси, отражающие долю почвы в пикселе $\mathrm{s}_{\mathrm{ij}}$, были использованы для расчета проективного покрытия виноградников $P_i = 1 - \mathrm{s}_{\mathrm{ij}}$. Оценка точности оценки проективного покрытия осуществлялась в ходе наземных обследований, а также по данным сверхвысокого разрешения (размер пикселя около 0,5 м), полученным цифровой цветной фотокамерой с борта самолета. В среднем точность оценки по тестовым участкам составила около 91 процентов.




*а* *б* *в*

Рис. 1. Результаты применения метода SQP для оценки проективного покрытия виноградников: *а* – исходные ГС данные космического аппарата Ресурс-П; *б* – данные цифровой цветной фотокамеры; *в* – результат обработки.

Аналогично по коэффициентам разложения смеси, извлеченным из алгоритма, осуществляется обнаружение и восстановление свойств объектов, чей размер меньше размера пикселя ГС изображения. Применение предложенного алгоритма в дальнейшем позволит более достоверно картографировать ландшафты по ГС данным невысокого пространственного разрешения с использованием алгоритмов субпиксельного картографирования, например, основанных на многоагентной системе [8].

## ЛИТЕРАТУРА

[1] *J.C. Harsanyi, C.-I. Chang* Hyperspectral Image Classification and Dimensionality Reduction: An Orthogonal Subspace Projection Approach // IEEE Trans. Geosci. Remote Sens. 32 (4), 1994. PP. 779-785.

[2] *S. Kraut, L.L. Scharf, L.T. McWhorter* Adaptive Subspace Detectors // IEEE Trans. Signal Process. 49 (1). 2001. PP. 1-16

[3] *Гладких Б.А.* Методы оптимизации и исследование операций для бакалавров информатики. Ч. II. Нелинейное и динамическое программирование: учебное пособие. Томск: Изд-во НТЛ, 2011. 264 с.

[4] *Aapo Hyvarinen* Fast and Robust Fixed-Point Algoriths for Independent Component Analysis // IEEE Trans. On Neural Networks. 10 (3). 1999. PP.626-634.

[5] *Li J., Bioucas-Dias J. M.* Minimum volume simplex analysis: a fast algorithm to unmix hyperspectral data // Proc. of IEEE International Conference on Geoscience and Remote Sensing Simposium (IGARSS). Boston, USA: IEEE. 2008. V. 3. P. 250-253.

[6] *Winter M.E.* N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data // Proc. Of SPIE Image Spectrometry V. V. 3753. Descour M., R., Shen S. S., eds. Denver, CO: SPIE. 1999. PP. 266-277.

[7] *Григорьева О.В.* Субпиксельная идентификация объектов местности по гиперспектральным данным на основе последовательного квадратичного программирования и метода анализа формы спектральных компонент // Цифровая обработка сигналов. 2018. № 3. С. 26-32.

[8] *Xiong Xu, Yanfei Zhong, Liangpei Zhang* Adaptive subpixel mapping based on a multiagent system for remote-sensing imagery // IEE Transaction on geoscience and remote sensing. 2014. Vol. 52. No. 2. PP. 787-804.

# АЛГОРИТМЫ ПОСЛЕДОВАТЕЛЬНОГО АНАЛИЗА МНОГОСПЕКТРАЛЬНЫХ АЭРОКОСМИЧЕСКИХ СНИМКОВ НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ СТРУКТУРНО-СТАТИСТИЧЕСКОГО ПОДХОДА ПРИ ДЕШИФРИРОВАНИИ ПРИРОДНЫХ ОБЪЕКТОВ

*Гук А.П.*[(1)], *Алтынцев М.А.*[(1)], *Евстратова Л.Г.*[(2)], *Алтынцева М.А.*[(1)]
[(1)] Сибирский государственный университет геосистем и технологий, Новосибирск
[(2)] Государственный университет по землеустройству, Москва

Рассматривается способ дешифрирования многоспектральных аэрокосмических снимков, основанный на непараметрическком подходе. Для анализа требуемых при распознавании объектов предлагается применять функции распределения и плотности распределения, построенных по исходным и преобразованным изображениям с помощью различных алгоритмов. Обсуждается способ повышения достоверности дешифрирования снимков путем последовательного применения алгоритмов их преобразования и использования большого числа тестовых образцов.
*Ключевые слова: дешифрирование, непараметрический подход, функция распределения, функция плотности распределения, решающее правило.*

**Введение**. Дешифрирование природных объектов по многоспектральным аэрокосмическим снимкам является основной задачей в области дистанционного зондирования Земли. Каждый канал многоспектральных аэрокосмических снимков представляет собой двумерную матрицу, в значениях которой хранится спектральная яркость элементов изображения. Спектральная яркость элементов изображения является формой записи спектральной яркости объектов – главного источника информации для распознавания различных объектов на изображениях. Распознавание объектов и определение их качественных характеристик может быть выполнено на основе классификации изображений. Но на достоверность классификации влияет большое число факторов, таких как тип и разрешение съемочной системы, ее ориентация в момент съемки, состояние атмосферы, наличие облачности, подверженность значительным изменениям спектральных коэффициентов отражения для различных объектов. То есть непосредственно определить качественные характеристики объектов невозможно. Поэтому необходимо сформировать вектор признаков, способный однозначно выделить объект и некоторые его свойства [1].

Для формирования вектора признаков чаще всего применяется такой подход, при котором создается модель, связывающая признаки и выполняемые в пространстве модели измерения. Признаки выбираются таким образом, чтобы они однозначно определяли объект. Наиболее простым видом таких моделей являются модели кластерного анализа, используемые для интерактивного дешифрирования аэрокосмических снимков. К таким моделям относятся и различные статистические модели: метод расстояния Махаланобиса, максимального правдоподобия и т.п. В большинстве таких моделях используется нормальный закон распределения. Статистические модели, соответствующие нормальному закону распределения, называют параметрическими [2].

Параметрические модели основаны на количественных признаках и применяются с целью упрощения решения задачи распознавания образов. Но если распределение оказывается отличным от нормального, достоверность распознавания объектов на изображениях с помощью применения параметрических моделей оказывается низкой. В таком случае необходимо применять непараметрические модели, основанных на качественных признаках.

**Методы анализа**. Сущность непараметрического подхода, предложенного в [3, 4], заключается в том, что для всех требуемых для распознавания классов объектов создаются эталонные признаки в виде функций плотности распределения вероятностей по выборкам достаточно большого размера. Измерения выполняются по изображениям эталонных объектов.

Функция плотности распределения вероятностей $f(x)$ является производной функций распределения *F(x)* и характеризует плотность, с которой распределяются значения случайной величины в определенной точке. А функция распределения определяет вероятность того, что в результате испытания случайная величина *X* примет значение, меньшее *x*. Значения функции распределения принадлежат интервалу [0, 1].

Прежде чем выполнять оценку функций распределения и функций плотности распределения вероятностей необходимо проверить распределение на нормальность. Проверка на нормальность осуществляется на основе критерия $\omega^2$ [3]:

$$\omega_m^2 = m\int....\int \psi^2[F(X_m,\mu)]\ [F_m(X_m) - F_m(X_m,\mu)]^2\, dF(X_m,\mu) \succ \omega_\alpha^2,$$

где $F_m(x_m)$ – эмпирическая функция распределения выборки $\{x_i\}^m$; $F_m(X_m,\mu)$ – функция нормального распределения с параметрами $\mu$ и $\Psi^2[F]$ – весовая функция.

Непараметрические модели целесообразно применять, если распределение оказывается отличным от нормального.

Для различных съемочных систем эталонные функции также различны. Поэтому при использовании непараметрического подхода следует для каждой съемочной системы создавать банк данных функций распределения и плотности распределения для всех требуемых при распознавании классов объектов. Банк данных эталонных функций получают с помощью картографических материалов. Эталонные функции строят по участкам изображения, соответствующим определенному классу объектов на карте. Затем, выполнив дешифрирование снимка на любую другую территорию методом сегментации, строят также функций распределения и плотности распределения для каждого сегмента. Построенные функции сравнивают с эталонными на основе заданного решающего правила.

В результате выполненных ранее экспериментальных исследований в качестве решающего правила при оценке функций плотности распределения вероятностей было выбрано значение коэффициента корреляции Пирсона между двумя рассчитанными функциями по изображениям, а при оценке функций распределения – предложенный авторами частный случай критерия Колмогорова.

Для сравнения функций распределения по частному случаю критерия Колмогорова в начале в каждом спектральном канале определяется наибольшее значение яркости $B_{max}$ среди всех сравниваемых участков изображения, одни из которых является дешифрируемым, а остальные – эталонными. Для соответствующего участка в каждом канале рассчитывается функция распределения в диапазоне [0, $B_{max}$]. Затем для исследуемого участка на интервале [0, 1] для значений функций распределения, кратных 0,1, определяются значения яркости *B*, на основе которых для каждого спектрального канала формируется вектор яркостей *f* размером 1x10, соответствующий значениям функции распределения, кратным 0,1. Аналогично формируются вектора $f_i$ для функций распределения эталонных участков изображения. На следующем этапе выполняется расчет расстояния *r* между вектором *f* и каждым из векторов $f_i$ по формуле:

$$r = \frac{\sum_{j=1}^{10} \left| f[j] - f_i[j] \right|}{10} \qquad (1)$$

Расстояния по функции распределения между каналами каждого эталонного участка и каналами дешифрируемого сравниваются между собой. Исследуемый участок будет принадлежать тому эталону, до которого расстояние, рассчитанное по формуле (1), будет минимальным. Также может быть рассчитано суммарное расстояние по всем каналам изображения.

Получаемые в результате такого анализа результаты могут иметь различную степень достоверности в различных спектральных каналах для определенных типов объектов. Например, водные объекты могут быть верно распознаны по различным спектральным каналам на основе сравнения как функций распределения, так и функций плотности распределения, в то время как достоверность распознавания пород леса будет существенно ниже. Может оказаться, что

определенная порода леса будет верно распознана на основе одного из видов функций, построенных по одному определенному спектральному каналу [6].

Распознавания пород леса представляет собой наибольшую трудность. Может оказаться, что порода леса не выделяется ни по одному из каналов. В этом случае в качестве исходного пространства признаков можно использовать не исходное многоспектральное изображение, а преобразованное в соответствии с априорно заданной вероятностной моделью многоспектральных измерений по одному из алгоритмов, таких как метод главных компонент, метод независимых компонент, преобразование Tasseled Cap, вегетационные индексы. Выполнив построение рассматриваемых функций по преобразованным изображениям, последовательно проанализировав результаты их построения посредством сравнения их схожести по одному из предложенных решающих правил, можно существенно повысить итоговую достоверность распознавания различных классов объектов [7, 8].

В качестве метода последовательного анализа наилучшим образов подходит алгоритм дерева решения. Дерево решений – это многоступенчатый алгоритм. Деревья решений представляют собой различные способы описания правил разделения данных в виде последовательной и иерархической структуры, где каждому объекту соответствует единственный дающий решение узел.

**Результаты**. Для исследования результатов дешифрирования природных объектов на основе применения структурно-статистического подхода и алгоритмов преобразования исходных многоспектральных изображений был выбран четырехканальный космический снимок Ikonos на территорию вблизи Академгородка г. Новосибирска. Разрешение каждого многоспектрального канала данного снимка составляет 3,2 м. По данному снимку и на основе тематической карты породного состава леса выполнялось создание эталонов (рис. 1). В качестве эталонов выбирались участки наибольшей площадью.

На рис. 2 показан пример эталонного участка, ограниченного контуром красного цвета и соответствующего сосновому лесу. Всего в качестве эталонов были выбраны следующие классы объектов: березовый лес, сосновый лес, осиновый лес, земля, водные объекты. Площадь эталонных участков составляла не менее 3 Га. Для оценки достоверности дешифрирования с применением функций распределения и плотности распределения исследуемые участки снимка выбирались также по тематической карте.

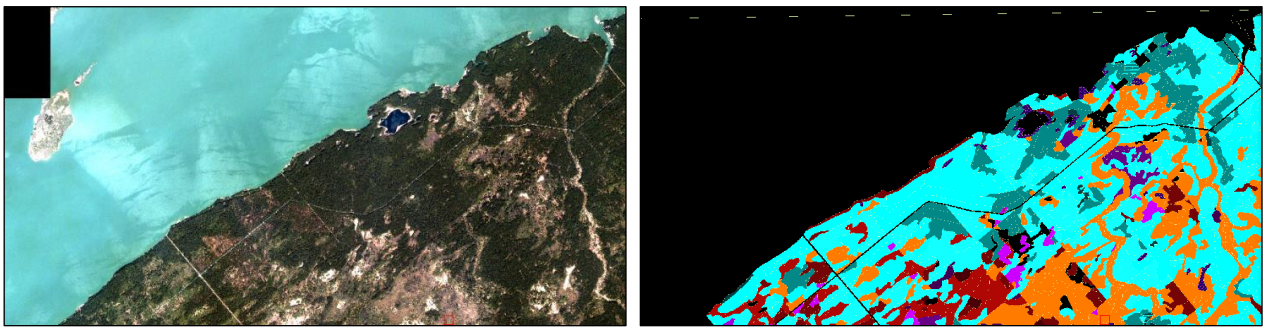


Рис. 1. Космический снимок Ikonos и тематическая карта на территорию Академгородка.



Рис. 2. Эталонный участок соснового леса.

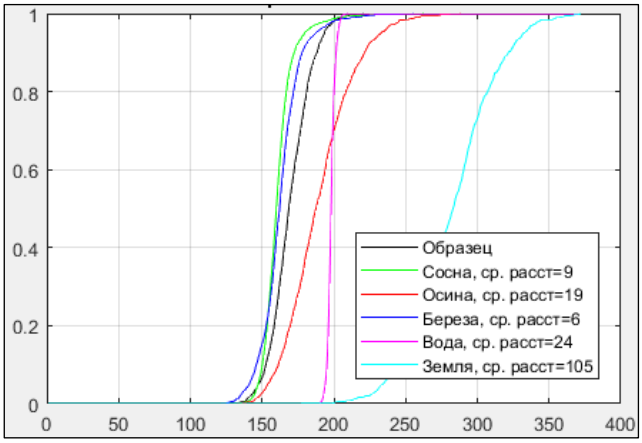


Рис. 3. Пример расчета функций распределения по красному каналу.

Функции распределения и плотности распределения рассчитывались по каждому каналу многоспектрального изображения, по преобразованным с помощью вегетационного индекса NDVI изображениям, а также по каждой компоненте преобразованного по методу главных компонент изображения. На рис. 3 показан пример построения функций распределения по красному каналу исходного изображения для эталонных участков и одного из дешифрируемых образцов. В качестве дешифрируемого образца был выбран участок березового леса. На рисунке также показано рассчитанное по формуле (1) расстояние от образца до каждого из эталонных участков. Минимальное значение расстояния было получено между функциями распределения образца березового леса и эталонного участка данного вида леса. То есть по красному каналу дешифрирование образца было выполнено верно.

На рис. 4 показан пример построения функций плотности распределения по красному каналу для тех же эталонных участков и образца.

Коэффициенты корреляции, показанные на рис. 4, рассчитывались между функцией плотности распределения дешифрируемого образца и данными функциями эталонных участков. Максимальное значение коэффициента корреляции было рассчитано между функциями плотности распределения образца березового леса и эталонного участка данного вида леса.

Таким образом, образец был верно дешифрирован по красному каналу как при использовании функции распределения, так и функции плотности распределения.

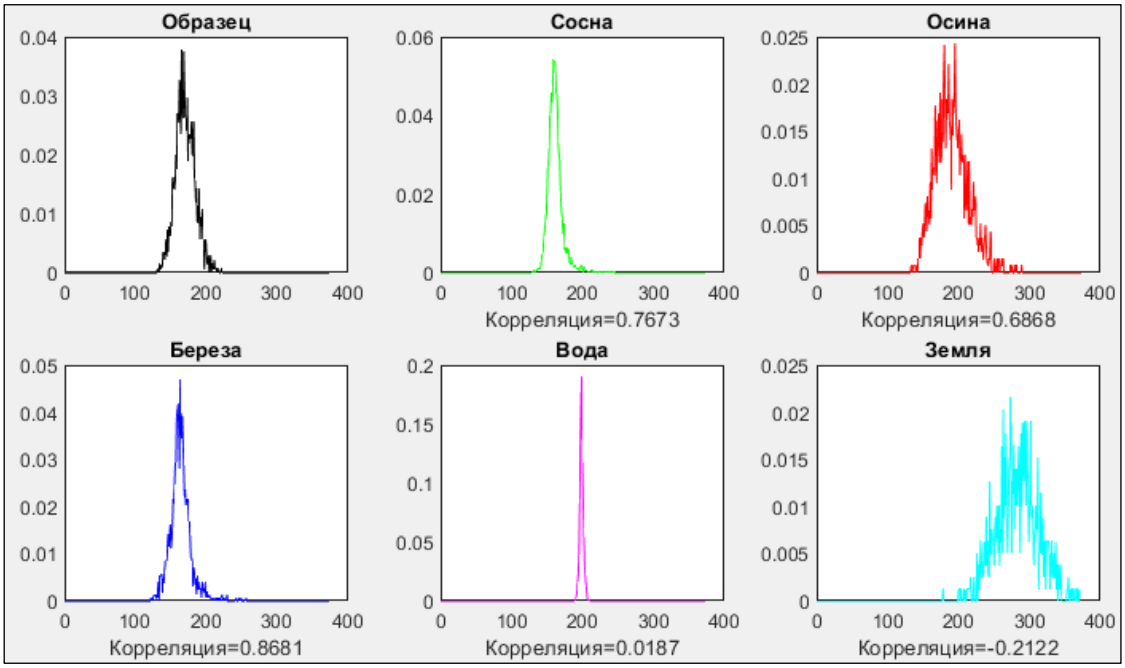


Рис. 4. Пример расчета функций плотности распределения по красному каналу.

Результаты дешифрирования образца.

| **Класс образца** | **Спектральный канал** | **Класс эталона** | **Расстояние по функции распределения от образца до эталона** | **Коэффициент корреляции по функциям плотности вероятностей между образцов и эталоном** |
|---|---|---|---|---|
| Береза | Красный | Сосна | 9.0 | 0.7673 |
| | | Осина | 19.2 | 0.6868 |
| | | Береза | **6.3** | **0.8681** |
| | | Вода | 24.0 | 0.0187 |
| | | Земля | 104.5 | -0.2122 |
| | Синий | Сосна | 2.9 | 0.8660 |
| | | Осина | 7.6 | 0.7098 |
| | | Береза | **0.1** | **0.9893** |
| | | Вода | 46.0 | -0.0401 |
| | | Земля | 39.4 | -0.0747 |
| | Зеленый | Сосна | 9.6 | 0.7393 |
| | | Осина | 11.9 | 0.8241 |
| | | Береза | **6.2** | **0.8915** |
| | | Вода | 68.5 | -0.0600 |
| | | Земля | 63.1 | 0.1426 |
| | Инфракрасный | Сосна | 23.8 | **0.7612** |
| | | Осина | **22.6** | 0.6350 |
| | | Береза | 47.0 | 0.6471 |
| | | Вода | 273.9 | -0.0821 |
| | | Земля | 16.8 | 0.6952 |
| | Четырехмерное пространство | Сосна | **27.882** | - |
| | | Осина | 33.791 | - |
| | | Береза | 47.856 | - |
| | | Вода | 288.356 | - |
| | | Земля | 130.424 | - |
| | Первая компонента | Сосна | 22.4 | **0.7547** |
| | | Осина | **21.9** | 0.6391 |
| | | Береза | 45.5 | 0.6491 |
| | | Вода | 281.6 | -0.0898 |
| | | Земля | 15.1 | 0.6772 |
| | Вторая компонента | Сосна | 11.7 | 0.7387 |
| | | Осина | 19.0 | 0.7267 |
| | | Береза | **11.1** | **0.8317** |
| | | Вода | 62.0 | -0.1005 |
| | | Земля | 111.2 | -0.4171 |
| | Третья компонента | Сосна | 3.1 | 0.8810 |
| | | Осина | 9.5 | 0.7451 |
| | | Береза | **2.0** | **0.9269** |
| | | Вода | 5.8 | 0.3535 |
| | | Земля | 57.4 | -0.1217 |
| | Четвертая компонента | Сосна | 3.1 | 0.8743 |
| | | Осина | **2.7** | **0.9244** |
| | | Береза | 4.7 | 0.8576 |
| | | Вода | 3.6 | 0.7447 |
| | | Земля | 12.0 | 0.3227 |
| | Индекс NDVI | Сосна | **0.006** | **0.8738** |
| | | Осина | 0.032 | 0.7242 |
| | | Береза | 0.038 | 0.7621 |
| | | Вода | 0.678 | -0.0967 |
| | | Земля | 0.192 | -0.0920 |

Как было отмечено выше, построение описанных функций может быть выполнено не только по исходным многоспектральным снимкам, но и по преобразованным с помощью определённого алгоритма. Преобразованные изображения могут повысить достоверность дешифрирования определённого образца. В таблице приведены результаты сравнения рассматриваемого образца с эталонами по всем спектральным каналам отдельно, по четырехмерному пространству изображения, по всем компонентам преобразованного по методу главных компонент изображения и по индексному изображению, полученному по формуле расчета вегетационного индекса NDVI. Данные результаты свидетельствуют о том, что достоверность дешифрирования существенно отличается в зависимости от того, по каким данным строится функции распределения и плотности распределения.

Чтобы объективно оценить, преобразованные по каким алгоритмам изображения обеспечивают наибольшую достоверность дешифрирования определенного класса объектов, необходимо выполнить построение функций по большому числу образцов и сравнить их с эталонами. За основу должен быть выбран тот алгоритм и та функция, обеспечивающие наибольшее различие классов. Для достижения еще большего процента достоверности также можно применить нескольких алгоритмов преобразования последовательно с помощью дерева решений.

**Заключение.** Применение непараметрического подхода при дешифрировании аэрокосмических снимков способно значительно повысить результаты распознавания различных классов объектов. Предложенные решающие правила позволяют численно оценить различия между функциями распределения и плотности распределения, построенным по исходным или преобразованным с помощью различных алгоритмов изображениям. Выполнив сравнение данных функций по большому числу образцов различных типов, можно выбрать ту функцию и подобрать те каналы исходных или преобразованных изображений, позволяющих достичь большей степени достоверности. А последовательное объединение нескольких из алгоритмов преобразования способно обеспечить достижение максимального уровня достоверности.

Дальнейшие исследования будут направлены на сбор большого числа статистической информации с целью определения устойчивых статистических характеристик распределения яркостей различных классов объектов на исходных и преобразованных многоспектральных изображениях, а также на определение последовательности применения алгоритмов преобразования и выбор характерных участков функций, определяющих наибольшее различие классов.

## ЛИТЕРАТУРА


[1] *Гук А.П., Евстратова Л.Г., Хлебникова Е.П., Арбузов С.А., Алтынцев М.А., Гордиенко А.С., Гук А.А., Симонов Д.П.* Дешифровочные признаки изображений объектов на многоспектральных космических снимках. Разработка методик автоматизированного дешифрирования аэрокосмических снимков // Геодезия и картография. 2013, № 7. С. 31-40.

[2] *Гук А.П., Евстратова Л.Г.* Исследование критериев оценки эффективности статистических непараметрических методов для дешифрирования лесных массивов // Материалы V Международной научной конференции «Региональные проблемы дистанционного зондирования Земли». Красноярск: Сиб. федер. ун-т, 2018. С. 12-15.

[3] *Гук А.П, Евстратова Л.Г.* Новый статистический подход к распознаванию лесных массивов // Материалы III Международной научной конференции «Региональные проблемы дистанционного зондирования Земли». Красноярск: Сиб. федер. ун-т, 2016. С. 14-17.

[4] *Гук А.П.* Автоматизация дешифрирования снимков. Теоретические аспекты статистического распознавания образов // Известия Вузов Геодезия и аэрофотосъемка. 2015. №5. С. 166-170.

[5] *Фукунага К.* Введение в статистическую теорию распознавания образов: Пер. с англ. / Под ред. А.А.Дорофеева. М.: Наука, 1970. 368 с.

[6] *Гук А.П., Шляхова М.М.* Исследование статистических свойств многоспектральных космических снимков леса // Материалы V Международной научной конференции «Региональные проблемы дистанционного зондирования Земли». Красноярск: Сиб. федер. ун-т, 2018. С. 105-108.

[7] *Гук А.П., Шляхова М.М.* Анализ эффективности применения метода главных компонент при использовании непараметрического статистического подхода к дешифрированию снимков // Материалы IV Международной научной конференции «Региональные проблемы дистанционного зондирования Земли». Красноярск: СФУ, 2017. С.89-94.

[8] *Гук А.П., Евстратова Л.Г.* Основные направления совершенствования автоматизации дешифрирования лесных массивов по многоспектральным аэрокосмическим снимкам // Журнал Сибирского федерального университета. Серия: Техника и технологии. 2018. Т. 11, № 8. С. 892-901.

# ОЦЕНКА ПАРАМЕТРОВ СНЕЖНОГО ПОКРОВА НА ЗЕМНОЙ ПОВЕРХНОСТИ БЕЗ РЕЛЬЕФА И С РЕЛЬЕФОМ МЕТОДОМ РАДАРНОЙ ИНТЕРФЕРОМЕТРИИ

*Дагуров П.Н.[(1)(2)], Дмитриев А.В.[(1)], Добрынин С.И.[(3)], Чимитдоржиев Т.Н.[(1)]*
[(1)] Институт физического материаловедения СО РАН, Улан-Удэ
[(2)] Бурятский государственный университет, Улан-Удэ
[(3)] Бурятский институт инфокоммуникаций, Улан-Удэ

Исследуется зондирование снежного покрова методом радарной интерферометрии. Предложена приближенная модель интерферометрического зондирования, основанная на методе малых возмущений. Проведена оценка вклада рассеяния от поверхности снега в значения амплитуды и интерферометрической фазы. Выполнен анализ влияния рельефа на оценку параметров снежного покрова. Приведены результаты численных оценок
*Ключевые слова: водный эквивалент снега, радарная интерферометрия, метод малых возмущений, рельеф.*

**Введение.** Сезонный снежный покров в регионах умеренных и северных широт является важным природным фактором. Снег оказывает большое влияние на климат, гидрологические и почвенные процессы, жизнь растений и животных, жизнедеятельность человека.

Основными характеристиками снежного покрова, определяющими его влияние на окружающую среду, являются его толщина и водный эквивалент снега (ВЭС) [1]. Водный эквивалент снега определяет содержание воды в снежном покрове В частности, в случае однородного снега с постоянной высотой водный эквивалент снега определяется как произведение глубины снежного покрова $d$ и его плотности $\rho_s$, отнесенной к плотности воды $\rho_w$, и выражается в единицах длины.

Использование метода радарной интерферометрии показало, что он является эффективным инструментом для диагностики и мониторинга различных изменений земных покровов [3-6]. Интерферометрические методы также использовались для анализа снежного покрова и оценки водного эквивалента снега [7-11]. Возможность прямых измерений SWE с помощью дифференциальной интерферометрии была впервые рассмотрена в [7]. В [8] приведены результаты экспериментов в С-диапазоне с использованием спутников дистанционного зондирования ERS-1 и ERS-2, которые показывают согласие с расчетными зависимостями. Теоретическая зависимость для интерферометрической фазы от ВЭС была также использована в [8] для сравнения с экспериментальными данными, полученными из пролетов самолетного РСА в L-диапазоне. Сравнение расчетной формулы для определения SWE сухого снега и экспериментальных результатов было проведено в [9]. В работе [10] для оценки ВЭС были использованы данные радара Sentinel-1. В статье [11] аналогичные оценки были выполнены с помощью анализа данных ALOS PALSAR 2, полученных на тестовом полигоне на побережье озера Байкал.

Ранее в работе [12] была предложена модель обратного рассеяния от снежного покрова на ровной в среднем земной поверхности. Обратное рассеяние происходит благодаря мелкомасштабным шероховатостям. В настоящей работе приводятся некоторые численные результаты для этой ситуации и рассматривается более общий случай рассеяния от снежного покрова на поверхности с рельефом.

**Обратное рассеяние от снежного покрова в отсутствие рельефа.** При построении модели обратного рассеяния микроволн от сухого снежного покрова на почве будем полагать, что снег является сплошной однородной средой и объемное рассеяние отсутствует. Это справедливо в С- и особенно в L-диапазонах, когда размеры снежных частиц много меньше длины волны.

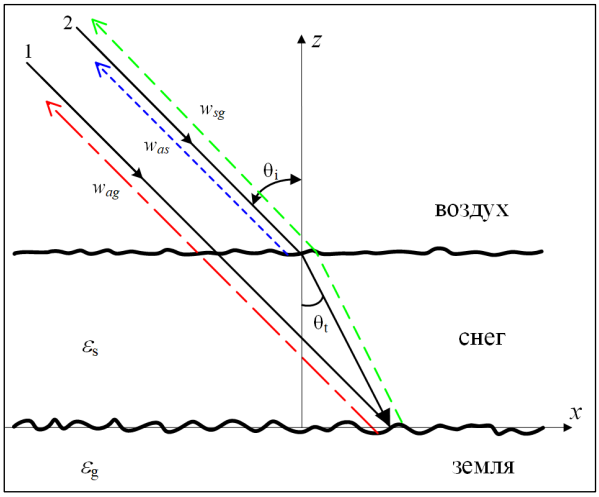


Рис. 1. Геометрия обратного рассеяния волны почвой без снежного покрова и со снежным покровом.

На рис. 1 показана геометрия задачи и траектории микроволн, по которым они распространяются в отсутствие снежного покрова и при его наличии. В отсутствие снега волна *1*, падающая из воздуха на почву под углом $\theta_i$ рассеивается обратно шероховатостями почвы в виде волны $w_{ag}$ (красная штриховая линия). Отметим, что на рис. 1 траектории падающих волн показаны сплошными линиями, а траектории рассеянных волны штриховыми линиями. В присутствии снежного покрова волна *2* при падении на снег из воздуха под углом $\theta_i$ после преломления и распространения в снежном слое падает на границу «снег – земля» под углом $\theta_t$ и рассеивается обратно теми же шероховатостями почвы в виде волны $w_{sg}$, траектория которой показана зеленой штриховой линией. В одноволновой модели [7] учитывается только эта рассеянная волна.

Рассмотрим теперь более общую модель, учитывающую волну обратного рассеяния от границы «воздух – снег». Пусть из воздуха на слой снега, покрывающий землю, падает радарная волна *2* под углом $\theta_i$ (рис.1). Радар находится в дальней зоне и падающую волну можно считать плоской волной $\vec{E}_p^i = \vec{E}_{0p} e^{jk(x\sin\theta_i - z\cos\theta_i)}$ ($k$ – волновое число в воздухе). Индекс $p$ здесь и далее описывает поляризацию излучения: $p = h$ при горизонтальной поляризации излучения и $p = v$ при горизонтальной поляризации; Однородная ground характеризуется комплексной диэлектрической проницаемостью $\varepsilon_g = \varepsilon_g' + j\varepsilon_g''$, сухой снег полагаем непоглощающей средой с диэлектрической проницаемостью $\varepsilon_s$, диэлектрическая проницаемость воздуха равна 1. Граничные поверхности «воздух – снег» и «снег – почва» являются статистически шероховатыми поверхностями со случайными неровностями, высоты которых описываются некоррелированными между собой стационарными случайными функциями $z_s(x, y)$ и $z_g(x, y)$ со средними значениями $\langle z_s \rangle = d$, и $\langle z_g \rangle = 0$ ($d$ – средняя толщина снега), среднеквадратичными отклонениями $b_s$ и $b_g$, радиусами корреляции $l_s$ и $l_g$. Полагаем, что неровности малы по сравнению с длиной волны, их наклоны невелики и выполняются условия применимости метода малых возмущений $kb_s, kb_g < 0{,}3$ и $kl_s, kl_g < 3$ [13]. Считаем, что неровности не влияют на когерентное поле (борновское приближение). Падающая волна проходит в слой снега под углом $\theta_t$, определяемым законом Снеллиуса. Рассеянное поле при выходе из снега также преломляется по закону Снеллиуса. Коэффициенты отражения и прохождения когерентных волн определяются формулами Френеля для плоской границы раздела.

Поле обратного рассеяния является когерентной суммой волн, рассеянных на неровностях границ. Это волна $w_{as}$, рассеянная границей «воздух – снег» (траектория её на рис.1 обозначена синей штриховой линией) и волна обратного рассеяния $w_{sg}$ (красная штриховая линия) от границы «снег – земля», возникающая после прохождения в слой и выхода из него (основная рассеянная волна). Кроме этих волн образуется более слабые волны обратного рассеяния,

появляющиеся вследствие отражений от границ слоя и рассеяния на неровностях, однако влиянием этих волн пренебрегаем из-за малого значения коэффициента отражения от границы «сухой снег – воздух» (см. раздел 2.2).

Вследствие малости неровностей не учитываем флуктуации фазы волны и полагаем, что в приближении метода малых возмущений значения фаз обусловлены путями, пройденными волнами в результате их взаимодействия с плоскими в среднем границами раздела. Тогда случайное электромагнитное поле, рассеянное шероховатыми границами в обратном направлении, с амплитудой $\left|E_p\right|$ и фазой $\Phi$ можно представить в виде следующей суммы

$$\left|E_p\right|e^{j\Phi} = \left|E_{sp}\right|e^{j\Phi_s} + T_{sp}T'_{sp}\left|E_{gp}\right|e^{j\Phi_g}, \qquad (1)$$

где слагаемые в правой части формулы описывают, соответственно, поля волн $w_1$, $w_2$. В формуле (1) $T_{sp}$ и $T'_{sp}$ – коэффициенты прохождения волн через границу «воздух – снег», рассматриваемую в среднем, как плоская поверхность, в прямом и обратном направлениях, соответственно; $\left|E_{sp}\right|$ – амплитуда поля, рассеянного неровностями снега на рис. 1; $\left|E_{gp}\right|$ – амплитуда поля, рассеянного неровностями ground. Интерферометрические фазы $\Phi_s$ и $\Phi_g$ отсчитываются относительно фазы волны в отсутствие снежного покрова

$$\Phi_s = \psi - \varphi_0, \quad \Phi_g = \varphi + \psi - \varphi_0, \qquad (2)$$

где $\varphi = 2k\sqrt{\varepsilon_2}d/\cos\theta_s$, $\psi = 2kd(tg\theta_i - tg\theta_t)\sin\theta_i$, $\varphi_0 = 2kd/\cos\theta_i$.

После усреднения соотношения (1) получим

$$\left\langle\left|E_p\right|\right\rangle e^{j\Phi} = \left\langle\left|E_{sp}\right|\right\rangle e^{j\Phi_s} + T_{sp}T'_{sp}\left\langle\left|E_{gp}\right|\right\rangle e^{j\Phi_g}, \qquad (3)$$

где выражения в угловых скобках представляют собой средние амплитуды рассеянных полей. Полагаем, что квадраты этих средних амплитуд пропорциональны средним квадратам амплитуд. Это справедливо, например, для случайных величин, распределенных по закону Релея. Поскольку средний квадрат амплитуды поля обратного рассеяния пропорционален коэффициенту обратного рассеяния, то из выражения (3) следует

$$\sqrt{\sigma_p^0}e^{j\Phi} = \sqrt{\sigma_{sp}^0}e^{j\Phi_s} + T_{sp}T'_{sp}\sqrt{\sigma_{gp}^0}e^{j\Phi_g}, \qquad (4)$$

где $\sigma^0$ – результирующий коэффициент обратного рассеяния, $\sigma_{sp}^0$ - коэффициент обратного рассеяния от границы «воздух – снег», $\sigma_{gp}^0$ – коэффициент обратного рассеяния от границы «снег – земля».

Величину $\sqrt{\sigma_p^0}e^{i\Phi}$ можно назвать комплексным амплитудным коэффициентом обратного рассеяния. Учитывая соотношения между коэффициентами прохождения и коэффициентом отражения $R_{sp}$ от границы «воздух – снег»: $T_{sp} = 1 + R_{sp}$, $T'_{sp} = 1 - R_{sp}$, перепишем (4) в виде

$$\sqrt{\sigma_p^o}e^{j\Phi} = \sqrt{\sigma_{sp}^0}e^{j\Phi_s} + (1 - R_{sp}^2)\sqrt{\sigma_{gp}^0}e^{j\Phi_g}. \qquad (5)$$

Далее для расчетов будем использовать известные выражения коэффициентов обратного рассеяния в приближении метода малых возмущений [12]. Диэлектрическая проницаемость сухого снега определяется выражением [14]

$$\varepsilon_s = 1 + 1.6\rho + 1.86\rho^3, \qquad (6)$$

где $\rho$ – численное значение плотности снега, выраженной в г/см$^3$. Это соотношение справедливо для частот в диапазоне 100 МГц – 10 ГГц и для плотности снега, меньшей 0,5.

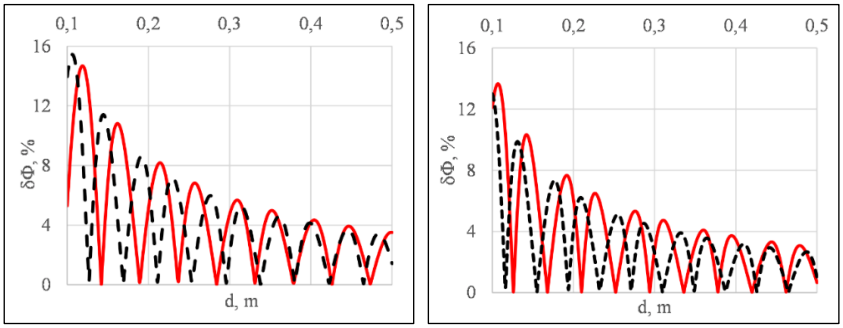



Рис. 2. Относительные вариации интерферометрической фазы в зависимости от глубины снега; сплошные красные линии – $\theta$ = 25°, штриховые синие линии – $\theta$ = 40°, а) $\rho$ = 0,2 г/см$^3$, б) $\rho$ = 0,4 г/см$^3$.

На рис. 2 представлены относительные вариации интерферометрической фазы $\delta\Phi$ в зависимости от глубины снега, полученные с п использованием (5) и (6). Представленные зависимости показывают, что относительные вариации в рассматриваемых ситуациях могут превышать 10%. Отметим, что при увеличении глубины снега относительные вариации уменьшаются. Этот результат объясняется тем, что с увеличением глубины снега значение $\Phi_g$ растет быстрее, чем $\Delta\Phi$. При глубине снега, большей 40 см, относительные вариации интерферометрической фазы не превышают 4%, т.е. рассматриваемом гипотетическом случае равенства среднеквадратичных значений и радиуса корреляции неровностей поверхности почвы и поверхности снега влияние обратного рассеяния от границы «воздух – снег» на интерферометрическую фазу является достаточно слабым.

Разность фаз основной волны обратного рассеяния, прошедшей снежный покров и волны обратного рассеяния в отсутствие снежного покрова согласно определяется соотношением [7].

$$\Phi_1 = 2kd(\sqrt{\varepsilon_s - \sin^2\theta_i} - \cos\theta_i). \qquad (7)$$

Оценки показывают, что интерферометрическая фаза основной рассеянной волны линейно связана с плотностью снега с хорошей точностью и, соответственно, она прямо пропорциональна ВЭС (SWE). Этот важный для практики результат получен впервые в [7]. Соотношения, связывающие интерферометрическую фазу с плотностью снега и ВЭС (SWE), получены в виде

$$\Phi_1 = \frac{1.5kd\rho}{\cos(\theta)}, \qquad SWE = \frac{\Phi_1\cos(\theta)}{1.5k} \qquad (8)$$

Оценка относительной погрешности определения интерферометрической фазы вследствие влияния волны, рассеянной границей «воздух - снег», и нелинейности формулы (7) получим, что результирующая погрешность формулы (8) не превышает 8% для углов падения 20° – 45° и плотности снега (0,2 – 0.3) г/см$^3$.

**Оценка влияния рельефа на интерферометрическую фазу.** Полагаем, что рельеф описывается цифровой моделью рельефа (ЦМР). Будем рассматривать взаимодействие волн с поверхностью снега и почвы в приближении касательной плоскости и геометрической оптики, полагая, что лучи взаимодействуют с плоскостями, касательными к поверхности почвы и снега в точке падения волны, как показано на рис. 3.

Таким образом, задачу можно сформулировать следующим образом. На слой снега, ограниченный граничными плоскостями «воздух – снег» и «снег – почва» падает волна под углом падения к поверхности снега $\theta_i$, который после преломления в снеге под углом $\theta_t$ и прохождения толщи снега рассеивается обратно. Плоскости «воздух – снег» и «снег – почва» параллельны друг другу и расстояние между ними по вертикали (толщина снега) равно $d$. Локальный наклон этих плоскостей, определяемый ЦМР, по отношению к горизонтальной плоскости составляет угол $\gamma$ по радиолокационной координате «дальность» (ось $y$) и $\omega$ по координате «азимут» (ось x). Угол падения волны на горизонтальную поверхность (номинальный угол падения) составляет $\theta_0$. Необходимо найти интерферометрическую фазу Ф волны в рассматриваемом случае неровного рельефа.

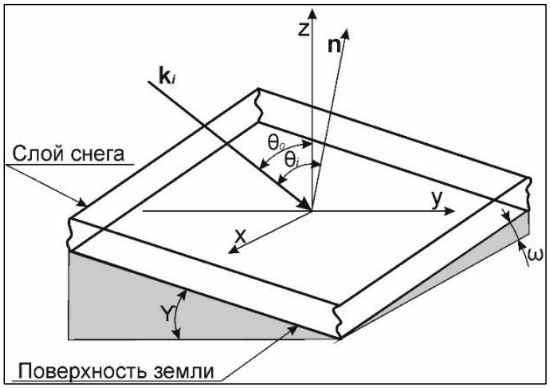


Рис. 3. Геометрия задачи о падении волны на поверхность с рельефом.

Локальный угол падения $\theta_i$, равный углу между единичной нормалью $\hat{n}$ к поверхностям снега и земли и единичным волновым вектором $\hat{k}_i$

$$\cos\theta_i = -\hat{k}_i \cdot \hat{n} = \frac{tg\gamma \sin\theta_0 + \cos\theta_0}{\sqrt{1+tg^2\gamma + tg^2\omega}}. \qquad (9)$$

Из формулы (9) можно заключить, что наибольшее влияние на локальный угол отражение оказывает наклон рельефа по координате «дальность». Далее полагаем, что $\omega = 0$.

Из геометрии задачи следует, что интерферометрическая фаза

$$\Phi = 2k\sqrt{\varepsilon_s} l_s - 2kl_a, \qquad (10)$$

где $l_s = \frac{d_1 \cos\gamma}{\cos\theta_t}$ – путь, пройденный волной в снегу, $l_a = l_s \cos(\theta_i - \theta_t)$ – путь в воздухе, пройденный в воздухе волной 2 дополнительно к пути, пройденному в воздухе волной 1.

Выражение (10) принимает вид

$$\Phi = \frac{2kd_1 \cos\gamma}{\cos\theta_t}\left[\sqrt{\varepsilon_s} - \cos(\theta_i - \theta_t)\right].$$

После упрощения получим

$$\Phi = 2kd_1 \cos\gamma\left(\sqrt{\varepsilon_s}\cos\theta_t - \cos\theta_i\right).$$

Переходя к номинальному углу падения $\theta_0$, имеем

$$\Phi = 2kd_1 \cos\gamma\left(\sqrt{\varepsilon_s - \sin^2(\theta_0 - \gamma)} - \cos(\theta_0 - \gamma)\right). \qquad (11)$$

Выражение (11) получено для случая, когда поверхность снега составляет положительный угол $\alpha$ с горизонтальной поверхностью, однако, как нетрудно видеть, она справедлива и для отрицательных значений $\alpha$.

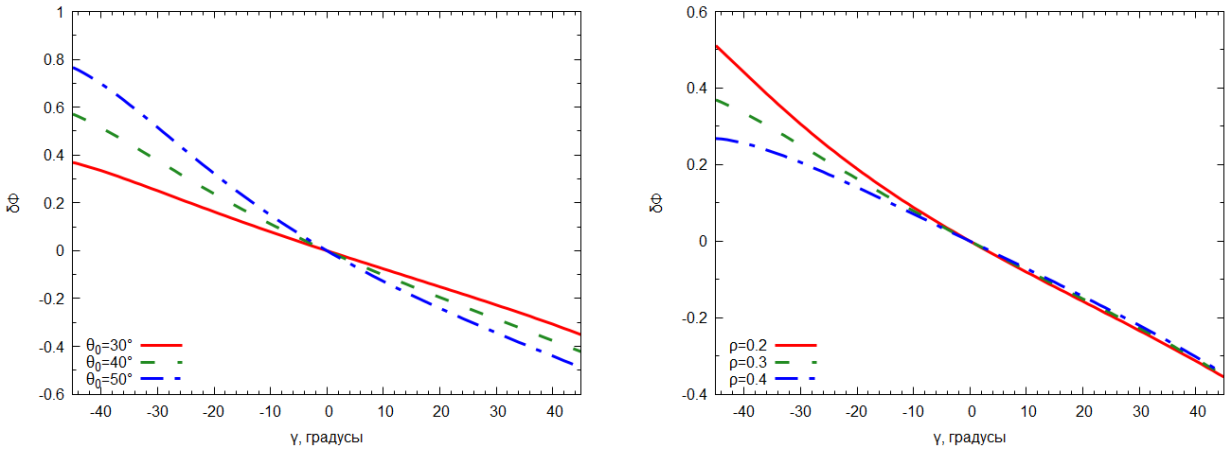
Разность интерферометрических фаз для случаев неровного и гладкого рельефа имеет значение

$$\Delta\Phi = \Phi - \Phi_0 = 2kd_1\left[\cos\gamma\left(\sqrt{\varepsilon_s - \sin^2(\theta_0 - \gamma)} - \cos(\theta_0 - \gamma)\right) - \left(\sqrt{\varepsilon_s - \sin^2\theta_0} - \cos\theta_0\right)\right]$$

и относительная разность фаз

$$\delta\Phi = \frac{\Delta\Phi}{\Phi_0} = \left[\frac{\cos\gamma\left(\sqrt{\varepsilon_s - \sin^2(\theta_0 - \gamma)} - \cos(\theta_0 - \gamma)\right)}{\left(\sqrt{\varepsilon_s - \sin^2\theta_0} - \cos\theta_0\right)} - 1\right]. \qquad (12)$$

На рис. 4 приведены результаты расчетов по формуле (12) в зависимости от угла $\gamma$ при различных значениях $\theta_0$ и $\rho$. Эти результаты показывают, что влияние наклонов рельефа на относительную фазу может быть достаточно значительным. Для крутых склонов с значениями $\gamma$ порядка 45° относительные изменения фазы достигают 40%. Однако при сравнительно пологих склонах при углах наклона, меньших 10° эти изменения не превышают 10%.




*а* *б*

Рис. 4. Зависимости относительного изменений интерферометрической фазы радарной волны от угла наклона рельефа вдоль координаты «дальность»: *а* – при различных $\theta_0$, *б* – при различных $\rho$.

**Заключение.** В работе построена приближенная модель для определения интерферометрической фазы, представляющей собой разность фаз радарных сигналов в отсутствие снега и после выпадения снега. Модель основана на методе малых возмущений. Выполнен анализ влияния волны обратного рассеяния от шероховатостей снежного покрова на фазу радарного сигнала. Выполнено обобщение модели на общий случай обратного рассеяния от снежного покрова на земной поверхности с рельефом и проведена оценка влияния углов наклона рельефа на интерферометрическую фазу.

*Работа выполнена при финансовой поддержке РФФИ (грант № 18-05-01051-а) и частично в рамках выполнения госзадания.*

ЛИТЕРАТУРА

[1] *Rees W.G.* Remote sensing of snow and ice. CRC Press, Taylor & Francis Group, 2006. 302 p.

[2] *A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. Papathanassiou.* A tutorial on synthetic aperture radar // IEEE Geosci. Remote Sensing Mag. 2013. V. 1. № 1. P. 6-43.

[3] *Hanssen R.F.* Radar Interferometry: Data Interpretation and Error Analysis. Dordrecht, Kluwer Academic Publishers. 2001. 308 p.

[4] *Zakharov A.I., Epov M.I., Mironov V.L., Chymitdorzhiev T.N., Bykov M.E., Seleznev V.S., Emanov A.F., Cherepenin V.A.* Earth surface subsidence in the Kuznetsk coal basin caused by manmade and natural seismic activity according to ALOS PALSAR interferometry. // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2013. V. 6. №. 3. P. 1578-1583.

[5] *Beck I., Ludwig R., Bernier M., Strozzi T., Boike, J.* Vertical movements of frost mounds in sub-Arctic permafrost regions analyzed using geodetic survey and satellite interferometry. // Earth Surface Dynamics. 2015. V 3. P. 409-421.

[6] *Chimitdorzhiev T.N., Dagurov P.N., Bykov M.E., Dmitriev A.V., Kirbizhekova I.I.* Comparison of ALOS PALSAR interferometry and field geodetic leveling for marshy soil thaw/freeze monitoring, case study from the Baikal lake region, Russia // Journal of Applied Remote Sensing. 2016. V. 10. № 1. P. 016006.

[7] *Guneriussen T., Hogda K. A., Johnsen H., Lauknes I.* InSAR for estimation of changes in snow water equivalent of dry snow. // IEEE Transactions on Geoscience and Remote Sensing, 2001. V. 39. № 10. P. 2101-2108.

[8] *Rott H., Nagler T., Scheiber R.* Snow mass retrieval by means of SAR interferometry. // In 3rd FRINGE workshop. European Space Agency: Earth Observation, 2003. Available at: **https://earth.esa.int/fringe03/proceedings/papers/46_rott.pdf**.

[9] *Leinss S., Wiesmann A., Lemmetyinen J., Hajnsek I.* Snow water equivalent of dry snow measured by differential interferometry. // IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. 2015. V. 8. №. 8. P. 3773-3790.

[10] *V. Conde, G. Nico, P. Mateus , J. Catalão1 , A. Kontu, M. Gritsevich* On the estimation of temporal changes of snow water equivalent by spaceborne SAR interferometry: a new application for the Sentinel-1 mission // J. Hydrol. Hydromech. 2019. V. 67. № 1. P. 93-100 DOI: 10.2478/johh-2018-0003.

[11] *Дагуров П.Н., Чимитдоржиев Т.Н., Дмитриев А.В., Добрынин С.И., Захаров А.И., Балтухаев А.К., Быков М.Е., Кирбижекова И.И.* Радиолокационная дифференциальная интерферометрия L-диапазона для определения параметров снежного покрова // Журнал радиоэлектроники. 2017. № 5. С. 14. **http://jre.cplire.ru/jre/may17/1/text.pdf**.

[12] *Дагуров П.Н., Балтухаев А.К., Дмитриев А.В., Добрынин С.И., Чимитдоржиев Т.Н.* Интерферометрическая модель радарного зондирования снежного покрова В сб.: Региональные проблемы дистанционного зондирования Земли Материалы V Межд. научн. конференции. СФУ, Институт космических и информационных технологий. 2018. С. 109-112.

[13] *Ulaby, F.T., R.K. Moore, and A.K. Fung* Microwave Remote Sensing. 1982. Chapter 12. V. 2, Dedham, MA: Artech House.

[14] *Mätzler C.* Microwave permittivity of dry snow // IEEE Trans. Geosci. Remote Sens. 1996. V. 34. №. 2. P. 573-581.

# ИНФОРМАТИВНОСТЬ ТЕКСТУРНЫХ ПРИЗНАКОВ В ЗАДАЧЕ КЛАССИФИКАЦИИ ПОЧВЕННО-РАСТИТЕЛЬНОГО ПОКРОВА

*Дмитриев Е.В.[(1)], Козодеров В.В.[(2)], Соколов А.А.[(3)]*
[(1)] Институт вычислительной математики им. Г.И. Марчука РАН, Москва
[(2)] Московский государственный университет им. М.В. Ломоносова, Москва
[(3)] Laboratoire de PhysicoChimie de l'Atmosphère Université du Littoral Cote d'Opale, Dunkerque, France

Проводится анализ информативности текстурных признаков в задаче обучаемой классификации объектов почвенно-растительного покрова на основе панхроматических изображений WorldView-2. Рассматривается 19 наиболее часто используемых признаков Харалика для различных направлений смежности пикселей. На основе корреляционного анализа исследованы взаимные зависимости признаков, а также выделены признаки не чувствительные к выбору направления смежности. Выделены наиболее информативные признаки, которые позволяют достичь достаточно высокой точности тематической обработки (ошибка классификации менее 1%).
*дистанционное зондирование, распознавание образов, текстурный анализ, изображения сверхвысокого разрешения, почвенно-растительный покров.*

**Введение.** С развитием аэрокосмических оптикоэлектронных систем мониторинга поверхности Земли в видимом и ближнем инфракрасном (ВБИК) диапазоне введены в эксплуатацию новые устройства, имеющие сверхвысокое пространственное разрешение. Ряд коммерческих спутниковых систем, таких как WorldView-2, 3, 4, GeoEye-1 и Pleades имеют пространственное разрешение 1.24-2 м в мультиспектральных каналах и 0.31-0.5 м в панхроматическом канале. Использование такого рода данных открывает новые возможности при решении различных задач дистанционного исследования почвенно-растительного покрова.

При разработке методов тематической обработки изображений лесных территорий, сверхвысокое пространственное разрешение позволяет учитывать распределение освещенности элементов лесного полога, рассматривать более широкий набор текстурных признаков, использовать результаты сегментации крон отдельных деревьев. Использование спутниковых изображений сверхвысокого разрешения (ИСВР), в конечном итоге, способствует созданию технологии точной дистанционной инвентаризации лесов, имеющей высокую актуальность для Российской Федерации и ряда других стран.

Особенности использования ИСВР обсуждаются в ряде научных публикаций последних лет. Большое внимание уделяется возможности восстановления параметров, определяющих структуру древостоя, таких как размер и плотность кроны, высота дерева, диаметр ствола и характерное расстояние между деревьями. Так, например, в работе [1] была предложена методика тематической обработки спутниковых ИСВР Quickbird и Pleades, основанная на поиске линейных зависимостей структурных параметров сосновых древостоев от спектральных признаков и текстурных признаков Харалика. Методика позволила достигнуть вполне приемлемых точностей: средняя погрешность определения диаметров крон составила 1.1 м, расстояния межу деревьями – 0.9 м, высоты – 3, диаметра стволов – 0,06 м. В работе [2] текстурные признаки Фурье, полученные при обработке фото ИСВР, были использованы для восстановления надземной биомассы древостоев северо-восточной части Китая. Сравнение с лидарными данными показало, что точность предлагаемой методики составила около 78%. Подобная задача рассматривалась также в работе [3] для тропических лесов Камбоджи. Авторами использовались текстурные признаки Харалика, Фурье и Габора, применительно к изображениям, предоставляемым системой Google Earth.

Во многих работах подобного рода возникает проблема оптимизации признакового пространства. Избыточность используемых признаков, влечет за собой возникновение проблемы проклятия размерности при обучении классификаторов и регрессионных моделей. В данной

работе рассматривается задача определения эффективной размерности признакового пространства и выбора наиболее информативного набора признаков при обработке панхроматических спутниковых ИСВР с целью классификации почвенно-растительного покрова.

**Методика текстурной классификации.** Описываемая здесь методика текстурного анализа была впервые предложена в работе [4]. Методика предназначена в первую очередь для обработки изображений в градациях серого. Схема обработки представлена на рис. 1.

На первом этапе необходимо оценить правильный размер окна – прямоугольного контура, который выделяет анализируемую часть изображения. Размер окна определяется характерным масштабом анализируемых текстур. Если размер окна выбран слишком малым, то результат текстурной классификации будет напоминать классификацию на основе яркостей отдельных пикселей. При слишком большом размере окна возрастает время расчетов и происходит чрезмерное сглаживание распознаваемых объектов. Таким образом, окно должно иметь минимально возможный размер, при котором анализируемые текстуры хорошо различимы.

Панхроматическое изображение расширяется на половинный размер окна. Центр окна пробегает все точки панхроматического изображения. При совместной обработке панхроматических и мультиспектральных (или гиперспектральных) изображений, для сокращения объема вычислений достаточно пробегать только пиксели, координаты которых соответствуют центрам пикселей многоспектрального изображения.

Для каждого положения окна производится вычисление матрицы совместной встречаемости уровней серого (GLCM – Gray-Level Co-Occurrence Matrix). Элементы GLCM представляют собой частоты встречаемости градиентов яркости по заданному направлению. Пример построения такой матрицы по горизонтальному направлению слева на право представлен на рис. 1. В данной работе рассматривается симметричный способ построения GLCM, когда наряду с заданным направлением рассматривается также и противоположное. На основе GLCM рассчитывается матрица

$$p(i,j) = \frac{GLCM(i,j)}{\sum_{i,j=1}^{N} GLCM(i,j)},$$

которая по сути является функцией вероятности распределения совместной встречаемости заданного числа градаций серого $N$ ($i, j$ – индексы элементов GLCM).

На основе значений $p(i,j)$ рассчитываются статистики, которые известны как текстурные признаки Харалика. В данной работе используются наиболее известные 19 статистик. Соответствующие формулы расчета представлены в табл. 1.

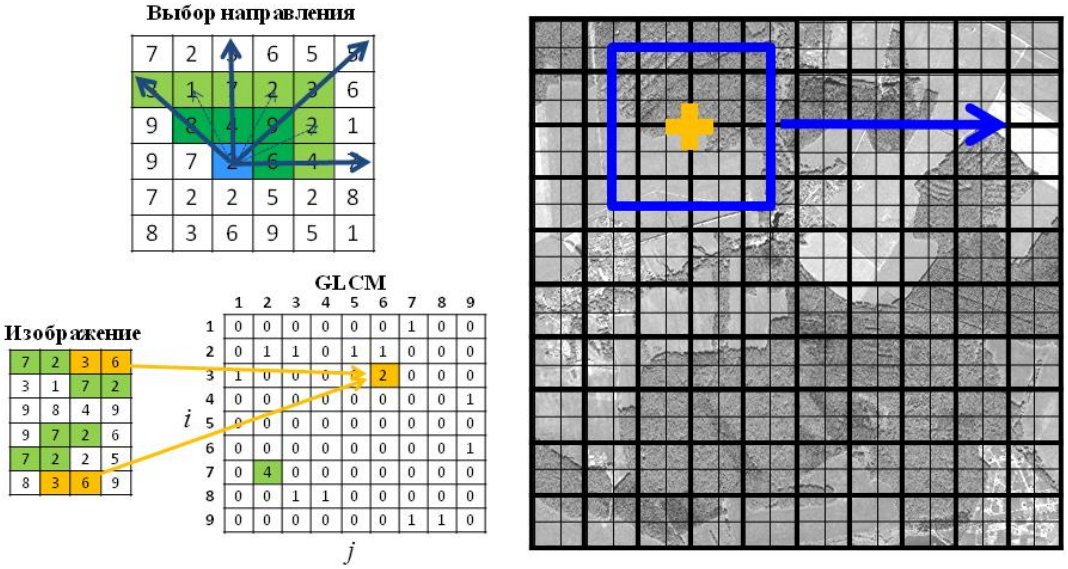



Рис. 1. Схема расчета текстурных признаков с использованием панхроматического ИСВР.

Табл. 1. Текстурные признаки Харалика, отсортированные по английским наименованиям.

| Наименование признака | Формула |
|---|---|
| Автокорреляция (Autocorrelation) | $\sum_{i=1}^{N}\sum_{j=1}^{N} i \cdot j \cdot p(i,j)$ |
| Островершинность (Cluster Prominence) | $\sum_{i=1}^{N}\sum_{j=1}^{N} (i+j-\mu_i-\mu_j)^4 \cdot p(i,j)$ |
| Асимметрия (Cluster Shade) | $\sum_{i=1}^{N}\sum_{j=1}^{N} (i+j-\mu_i-\mu_j)^3 \cdot p(i,j)$ |
| Контраст (Contrast) | $\sum_{i=1}^{N}\sum_{j=1}^{N} (i-j)^2 \cdot p(i,j)$ |
| Корреляция (Correlation) | $\sum_{i=1}^{N}\sum_{j=1}^{N} (i-\mu_i)\cdot(j-\mu_j)\cdot p(i,j) / (\sigma_i \cdot \sigma_j)$ |
| Энтропия разности (Diffrence Entropy) | $-\sum_{k=0}^{N-1} p_{i-j}(k) \cdot \ln p_{i-j}(k)$ |
| Дисперсия разности (Diffrence Variance) | $\sum_{k=0}^{N-1} (k-\mu_{i-j})^2 \cdot p_{i-j}(k)$ |
| Неоднородность (Dissimilarity) | $\sum_{i=1}^{N}\sum_{j=1}^{N} \lvert i-j \rvert \cdot p(i,j)$ |
| Энергия (Energy) | $\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j)^2$ |
| Энтропия (Entropy) | $-\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j) \cdot \ln p(i,j)$ |
| Однородность (Homogeneity) | $\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j) / (1+\lvert i-j \rvert)$ |
| Однородность2 (Homogeneity2) | $\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j) / (1+(i-j)^2)$ |
| Первая информационная мера корреляции (Information Measure of Correlation 1) | $(HXY - HXY1) / \max(HX, HY)$ |
| Вторая информационная мера корреляции (Information Measure of Correlation 2) | $\sqrt{1-\exp(-2(HXY2-HXY))}$ |
| Максимум вероятности (Maximum Probability) | $\max_{i,j} p(i,j)$ |
| Среднее суммы (Sum Average) | $\sum_{k=2}^{2N} k \cdot p_{i+j}(k)$ |
| Энтропия суммы (Sum Entropy) | $-\sum_{k=2}^{2N} p_{i+j}(k) \cdot \ln p_{i+j}(k)$ |
| Сумма квадратов (Sum Squares) | $\sum_{i=1}^{N}\sum_{j=1}^{N} (i-\mu_i)^2 \cdot p(i,j)$ |
| Дисперсия суммы (Sum Variance) | $\sum_{k=2}^{2N} (k-\mu_{i+j})^2 \cdot p_{i+j}(k)$ |

При вычислении статистик использовались следующие параметры:

1) средний индекс $\mu_i = \sum_{i=1}^{N}\sum_{j=1}^{N} i \cdot p(i,j)$, $\mu_j = \sum_{i=1}^{N}\sum_{j=1}^{N} j \cdot p(i,j)$;
2) среднеквадратичное отклонение (СКО) индекса $\sigma_i = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} (i-\mu_i)^2 \cdot p(i,j)}$;
3) вероятность разности индексов $p_{i-j}(k) = \sum_{|i-j|=k} p(i,j)$;
4) вероятность суммы индексов $p_{i+j}(k) = \sum_{i+j=k} p(i,j)$;
5) энтропии $HX = -\sum_{i=1}^{N} p_x(i) \cdot \ln p_x(i)$, $HY = -\sum_{j=1}^{N} p_y(j) \cdot \ln p_y(j)$,
$HXY = -\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j) \cdot \ln p(i,j)$, $HXY1 = -\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j) \cdot \ln\left(p_x(i) \cdot p_y(j)\right)$,
$HXY2 = -\sum_{i=1}^{N}\sum_{j=1}^{N} p_x(i) \cdot p_y(j) \cdot \ln\left(p_x(i) \cdot p_y(j)\right)$, где $p_x(i) = \sum_{j=1}^{N} p(i,j)$, $p_y(j) = \sum_{i=1}^{N} p(i,j)$.

Для проведения классификации на основе описанных выше текстурных признаков были рассмотрены три стандартных метода: нормальный байесовский классификатор, метод

k-ближайших соседей (КБС) и многоклассовый метод опорных векторов с гауссовским ядром [5, 6]. Указанные методы имеют различную постановку задачи, точность и вычислительную сложность.

Была проведена серия экспериментов, описание которых выходит за рамки данной статьи, в которых сравнивалась эффективность указанных классификаторов для решения рассматриваемой задачи. В результате была выбрана эффективная модификация метода КБС, в которой для повышения скорости вычислений был использован оптимизированный поиск с использованием kd-деревьев. Выбранное число соседей 49 обеспечивает баланс между точностью классификации и устойчивостью обучения.

**Численные эксперименты.** Для проведения расчетов использовались панхроматические изображения WorldView-2 территории Бронницкого лесничества (Московская область). Рассмотрены два тестовых участка, содержащие различные группы объектов. Участок Отра находится вблизи водоема (Татаринцевский пруд) и содержит 4 основных типа объектов, отличающихся по текстуре: водная поверхность, поле, естественный смешанный древостой с преобладанием березы и лесокультуры ели. Участок Лубнинка находится вблизи соответствующего населенного пункта. Он содержит естественные леса с преобладанием дуба и березы, а также часть территории экспериментального полигона, на котором выращивается лиственница. Отличительной особенностью лиственного древостоя является строгая упорядоченность, деревья расположены вдоль прямых линий на равных расстояниях друг от друга и имеют практически одинаковые размеры крон. При проведении текстурного анализа особый интерес составляет возможность классификации естественных и культурных насаждений.

На основе панхроматических изображений тестовых территорий были рассчитаны текстурные признаки, представленные в табл. 1 (19 параметров) для 4 направлений смежности 0, 45, 90 и 135 градусов. Таким образом, исходное признаковое пространство имеет размерность 76. Большинство признаков оказались существенным образом зависимы. На рис. 2 представлены корреляционные матрицы для 19 признаков по совокупности направлений. Оценки корреляций между признаками различаются для рассматриваемых тестовых участков, однако можно видеть, что они имеют схожую структуру.

Анализ корреляций по пороговым значениям показал следующее. 35% рассматриваемых признаков имеют взаимные корреляции более 0.8 для обоих участков. Связь между такими переменными в первую очередь объясняется способом их построения. Относительно небольшая часть признаков имеет слабо выраженную взаимную зависимость. Корреляцию менее 0.5 имеют 30% признаков для участка Отра и 25% для участка Лубнинка, а корреляцию менее 0.3 имеют 16% и 8 % признаков, соответственно. Таким образом, связи между данными признаками существенно зависят от выбора сцены.

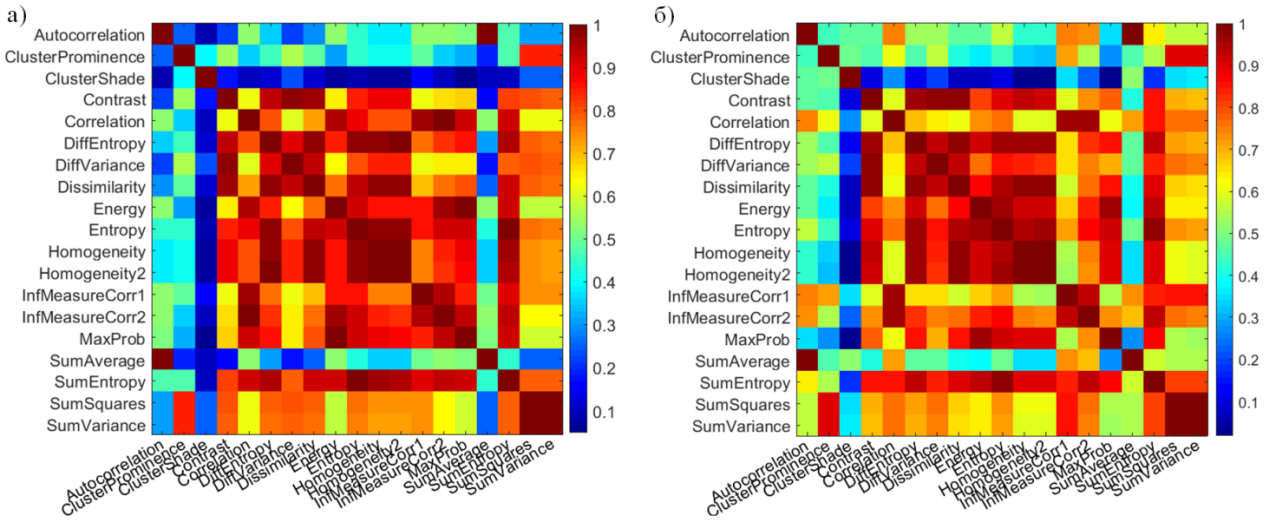



Рис. 2. Матрицы модулей корреляции текстурных признаков для тестовых участков: *а* – Отра; *б* – Лубнинка.

Табл. 2. Минимальная и максимальная корреляция текстурных признаков Харалика по направлениям смежности пикселей для участков Отра и Лубинка.

| Признак | Отра | | | | Лубинка | | | |
|---|---|---|---|---|---|---|---|---|
| | $\rho_{min}$ | $\alpha(\rho_{min})$ | $\rho_{max}$ | $\alpha(\rho_{max})$ | $\rho_{min}$ | $\alpha(\rho_{min})$ | $\rho_{max}$ | $\alpha(\rho_{max})$ |
| Autocorrelation | 1 | 90-0 | 1 | 135-45 | 1 | 90-0 | 1 | 135-45 |
| ClusterProminence | 0.99 | 90-0 | 1 | 135-90 | 1 | 90-0 | 1 | 135-45 |
| ClusterShade | 0.99 | 135-45 | 0.99 | 135-90 | 0.99 | 90-0 | 1 | 135-45 |
| Contrast | 0.81 | 135-45 | 0.95 | 135-90 | 0.83 | 135-45 | 0.94 | 90-45 |
| Correlation | 0.96 | 135-45 | 0.98 | 135-90 | 0.78 | 135-45 | 0.94 | 135-0 |
| DiffEntropy | 0.96 | 90-0 | 0.99 | 135-90 | 0.92 | 135-45 | 0.97 | 90-45 |
| DiffVariance | 0.72 | 135-45 | 0.94 | 135-90 | 0.79 | 135-45 | 0.92 | 90-45 |
| Dissimilarity | 0.93 | 90-0 | 0.97 | 135-90 | 0.9 | 135-45 | 0.96 | 90-45 |
| Energy | 1 | 90-0 | 1 | 135-45 | 0.99 | 135-45 | 1 | 135-0 |
| Entropy | 1 | 90-0 | 1 | 135-90 | 0.99 | 135-45 | 1 | 135-0 |
| Homogeneity | 0.96 | 90-0 | 0.99 | 135-90 | 0.92 | 135-45 | 0.96 | 90-45 |
| Homogeneity2 | 0.96 | 90-0 | 0.99 | 135-90 | 0.92 | 135-45 | 0.96 | 90-45 |
| InfMeasureCorr1 | 0.94 | 90-0 | 0.97 | 135-90 | 0.94 | 135-45 | 0.96 | 135-0 |
| InfMeasureCorr2 | 0.98 | 90-0 | 0.99 | 135-45 | 0.95 | 135-45 | 0.99 | 135-0 |
| MaxProb | 0.99 | 90-0 | 1 | 90-45 | 0.95 | 90-0 | 0.96 | 90-45 |
| SumAverage | 1 | 90-0 | 1 | 135-45 | 1 | 90-0 | 1 | 135-45 |
| SumEntropy | 1 | 90-0 | 1 | 90-45 | 1 | 135-45 | 1 | 90-0 |
| SumSquares | 1 | 90-0 | 1 | 135-45 | 1 | 90-0 | 1 | 135-45 |
| SumVariance | 1 | 90-0 | 1 | 135-90 | 1 | 90-0 | 1 | 135-45 |

Результаты корреляционного анализа признаков для 4 выбранных направлений представлены в табл. 2. Можно видеть, что такие признаки как Autocorrelation, Energy, Entropy, SumAverage, SumEntropy, SumSquares и SumVariance не имеют зависимости от выбора направления. Наибольшую чувствительность к выбору направления имеют признаки Contrast и DiffVariance. Следует отметить, что означенные выводы можно сделать для обоих тестовых участков.

Для эффективного сокращения признакового пространства использовался регуляризованный метод последовательного отбора [7]. Проблема стандартного метода последовательного отбора состоит в том, что получаемая последовательность наиболее информативных признаков имеет высокую чувствительность к малым изменениям обучающего множества. Регуляризованный метод позволяет получить более устойчивый результат. Возможные колебания в результатах отбора обычно соответствуют наименее информативным членам последовательности признаков. Устойчивость отбора повышается с увеличением числа повторных расчетов локально-оптимальных последовательностей признаков.

При обработке данных по участку Отра была выделена следующая последовательность признаков (в скобках указывается направление смежности): Contrast(45), Autocorrelation(0), DiffEntropy(135), Correlation(90), Homogeneity2(135), Dissimilarity(0) и Correlation(0). Полученные результаты согласуются с данными, представленными на рис. 3. Первые 3 наиболее информативных признаков имеют наибольшую вероятность вхождения в ансамбль локально-оптимальных последовательностей.

Результаты тематической обработки тестовых участков Отра и Лубинка представлены на рис. 4. Можно видеть, что целевые объекты были классифицированы достаточно точно. Черным цветом на рис. 4,*б* и *г* обозначены прочие объекты, признаки которых находятся на достаточно большом расстоянии от обучающего множества. Области прочих объектов соответствуют в основном границам между целевыми объектами, дорожной сети, прибрежному мелководью (видимость дна изменяет текстуру водной поверхности).

Для оценки ошибок распознавания использовались методы кросс-валидации k-fold [6], переклассификации (обучающий и тестовый ансамбли совпадают) и независимая валидация

(тестовый и обучающий ансамбли полностью различны). Для этих ошибок введены обозначения CV, Resub и Indep, соответственно. Общими характеристиками качества обучаемой классификации являются полная вероятность ошибки TE, средняя ошибка пропуска цели TOE, средняя ошибка ложной тревоги TCE и каппа Коэна (kappa) [6]. Значения этих ошибок представлены в таблице 3. Близость ошибок переклассификации и кросс-валидации свидетельствует об устойчивости обучения. Независимые оценки ошибок значимо превышают ошибки кросс-валидации. Таким образом, можно заключить о наличии систематических изменений значений текстурных признаков на изображении. В целом можно говорить о достаточно высокой точности классификации, для обоих тестовых участков оценки ошибок не превышают 1%, а высокие значения каппа Коэна говорят об отличном соответствии между результатами классификации и экспертными данными.

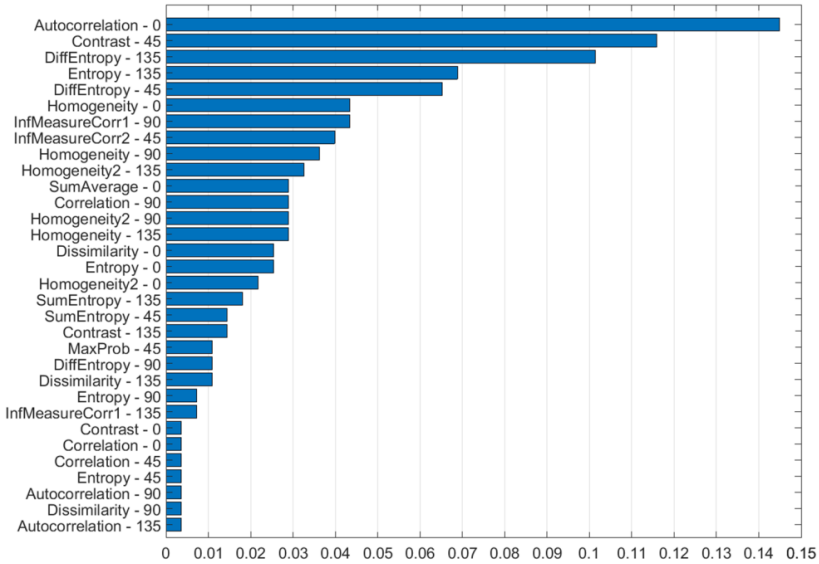
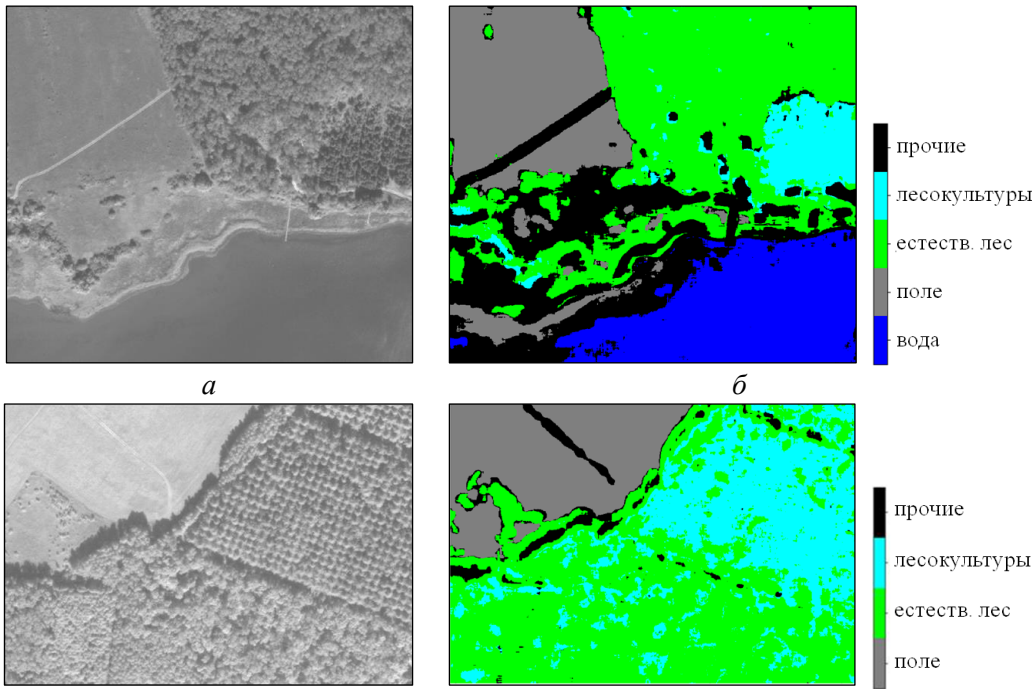



Рис. 3. Вероятности вхождения признаков в ансамбль локально-оптимальных последовательностей признаков.




*а* *б*

*в* *г*

Рис. 4. Распознавание целевых классов по текстурным признакам:
*а* и *б* – панхроматическое изображение и тематическая карта участка Отра;
*в* и *г* – панхроматическое изображение и тематическая карта участка Лубинка.

Табл. 3. Общие характеристики качества классификации.

| | Отра | | | Лубинка | | |
|---|---|---|---|---|---|---|
| | **CV** | **Resub** | **Indep** | **CV** | **Resub** | **Indep** |
| **TE** | 0.001 | 0.001 | 0.005 | 0.006 | 0.005 | 0.072 |
| **TOE** | 0.002 | 0.001 | 0.009 | 0.005 | 0.004 | 0.070 |
| **TCE** | 0.002 | 0.004 | 0.002 | 0.007 | 0.067 | 0.006 |
| **kappa** | 0.998 | 0.999 | 0.994 | 0.990 | 0.992 | 0.890 |

Табл. 4. Поклассовые характеристики качества классификации.

| | | **вода** | **поле** | **лес** | **посадки** |
|---|---|---|---|---|---|
| **Отра** | **OE** | 0.000 | 0.000 | 0.002 | 0.035 |
| | **CE** | 0.000 | 0.000 | 0.011 | 0.006 |
| **Лубинка** | **OE** | - | 0.002 | 0.070 | 0.138 |
| | **CE** | - | 0.000 | 0.091 | 0.109 |

Независимые ошибки пропуска цели OE и ложной тревоги CE приведены в таблице 4 для каждого рассматриваемого класса. Для обоих тестовых участков наименьшая точность достигается при распознавании лесокультур. Для тестового участка Лубинка ошибки распознавания территории экспериментального полигона довольно высоки, по всей видимости это связано с соответствием среднего размера крон естественных древостоев и культурных насаждений лиственницы.

*Работа выполнена при финансовой поддержке РФФИ (грант № 19-01-00215).*

## ЛИТЕРАТУРА

[1] *Beguet B., Guyon D., Boukir S., Chehata N.* Automated retrieval of forest structure variables based on multi-scale texture analysis of VHR satellite imagery // ISPRS J. Photogramm. Remote Sens. 2014. V. 96. P. 164–178.

[2] *Meng S., Pang Y., Zhang Z., Jia W., Li Z.* Mapping Aboveground Biomass using Texture Indices from Aerial Photos in a Temperate Forest of Northeastern China // Remote Sensing. 2016 V. 8. P. 230.

[3] *Singh M., Evans D., Friess, D., Tan B., Nin C.* Mapping Above-Ground Biomass in a Tropical Forest in Cambodia Using Canopy Textures Derived from Google Earth // Remote Sensing. 2015. V. 7. P. 5057-5076.

[4] Haralick R.M., Shanmugam K., Dinstein I. Textural Features for Image Classification // IEEE Transactions on Systems, Man, and Cybernetics, SMC-3. 1973. N 6. P. 610-621.

[5] *Duda R., Hart P., Stork D.* Pattern Classification, Second Edition. New York, NY, Uand Sons. 743 p.

[6] *Hastie T., Tibshirani R., Friedman J.* The Elements of Statistical Learning. New York: Springer, 2001. 745 p.

[7] *Дмитриев Е.В.* Классификация лесного покрова тверской области на основе гиперспектральных аэроизображений // Исследование Земли из космоса. 2013. №3. С. 22-32.

# ИНТЕГРИРОВАННАЯ ГЕОИНФОРМАЦИОННАЯ СИСТЕМА МОНИТОРИНГА СОСТОЯНИЯ ВНУТРИКОНТИНЕНТАЛЬНЫХ ВОДНЫХ ОБЪЕКТОВ

*Донцов А.А.[(1)], Суторихин И.А.[(1)(2)], Коломейцев А.А.[(1)]*
[(1)] Институт водных и экологических проблем СО РАН, Барнаул
[(2)] Институт вычислительных технологий СО РАН, Новосибирск

Представлено описание модульной специализированной геоинформационной системы (ГИС) мониторинга внутриконтинентальных водоёмов. Информационной основой ГИС являются данные дистанционного зондирования Земли, а также данные наземных измерительных комплексов и экспедиционных работ. Показана архитектура ГИС, основные модули и компоненты. Приводятся результаты использования этой системы в задаче сопоставления экспедиционных данных измерения концентрации хлорофилла «а» в поверхностном слое водоемов.
*Ключевые слова: дистанционное зондирование, ГИС, водохранилище, озеро, спутниковые данные, Django, MapServer, Posgresql, Posgis.*

**Введение.** Мониторинг параметров озер и водохранилищ является важной и актуальной задачей природопользования. Проблема сохранения водных ресурсов и повышения качества воды возрастает с каждым годом. Вместе с тем, увеличивается и число задач управления водными ресурсами. Эффективное управление в этой области требует учета и анализа большого числа факторов, а также доступности результатов анализа широкому кругу заинтересованных организаций и лиц, начиная с государственных структур и заканчивая общественными организациями. В последние десятилетия под влиянием глобальных и региональных изменений в экологической системе и воздействия антропогенных факторов на территории Российской Федерации наблюдаются значительные изменения состояния гидрологических и гидробиологических параметров внутриконтинентальных водных объектов [1]. В связи с этим, разрабатываются различные системы и подходы мониторинга, как правило, они предназначены для относительно больших водных объектов, таких как озеро Байкал [2, 3]. На малых озерах практически отсутствует наземная система измерительных приборов, в связи с этим основным источником информации об их состоянии являются данные дистанционного зондирования Земли из космоса и результаты экспедиционных исследований. Однако современный уровень изучения водных объектов предполагает проведение регулярных систематических измерений, что затруднительно в плане проведения экспедиционных работ, величин, характеризующих состояние водных объектов так и метеорологических параметров, приводящих к изменению этих состояний. Изучение состояния водных объектов типа озер и водохранилищ трудно представить без хорошо структурированной базы данных площадных, гидрологических и гидрохимических параметров [1, 4, 5].

Целью работы является представление результатов разработки геоинформационной веб-системы регистрации гидрологических и гидробиологических параметров озер и водохранилищ, разрабатываемой в Институте водных и экологических проблем СО РАН. Основной особенностью данной ГИС является то, что позволяет использовать при решении гидрологических и гидробиологических задач спутниковые снимки, данные наземных измерительных комплексов и результаты экспедиционных измерений. Такой подход обеспечивает возможность предоставления разноплановой информации о состоянии водных объектов пользователям ГИС.

**Описание ГИС.** Компоненты ГИС объединяет между собой веб-платформа Django, которая позволяет реализовывать модульные приложения на языке программирования Python. Для работы с системой был разработан веб-интерфейс, WMS/WPS-интерфейс и RESTfull web-api для интеграции с наземными измерительными комплексами. На рис. 1 показана общая схема работы ГИС, которая взаимодействует с удаленными источниками данных посредством HTTP запросов с целью получения данных о выбранном водном объекте. Главным потребителем результатов работы выступает пользователь, на основе запросов которого осуществляется

поиск, обработка и отображение результатов вычислений. Исходя из этого, ГИС содержит ряд сервисов: «Каталог данных», «Сервис обработки данных» и «Средства экспорта и визуализации данных».

Пользователь посредством веб-интерфейса формирует запрос на получение данных по выбранному водному объекту, модуль ГИС проверяет, есть ли запрашиваемая информация в каталоге ГИС, если её нет, то в случае со спутниковыми снимками формируется запрос к архивам ДЗЗ, информация наземных измерительных комплексов поступает в каталог через программный интерфейс или загружается администратором. После поступления в каталог данных происходит их обработка и представление в виде web-карты или таблиц. Более подробно архитектура и компоненты ГИС показаны в наших предыдущих работах [3]. ГИС можно разделить на два блока – это работа со спутниковыми данными и работа с данными наземных измерительных комплексов и натурных наблюдений. Интеграцию с наземными измерительными комплексами обеспечивает RESTfull web-api, который основывается на расширении Django REST framework (DRF). Данные API передаются в формате JSON и после валидации при помощи функционала Django-форм записываются в базу данных ГИС. Результаты экспедиционных работ также могут быть добавлены в ГИС посредством API или web-интерфейса с формой добавления и импорта данных [6].

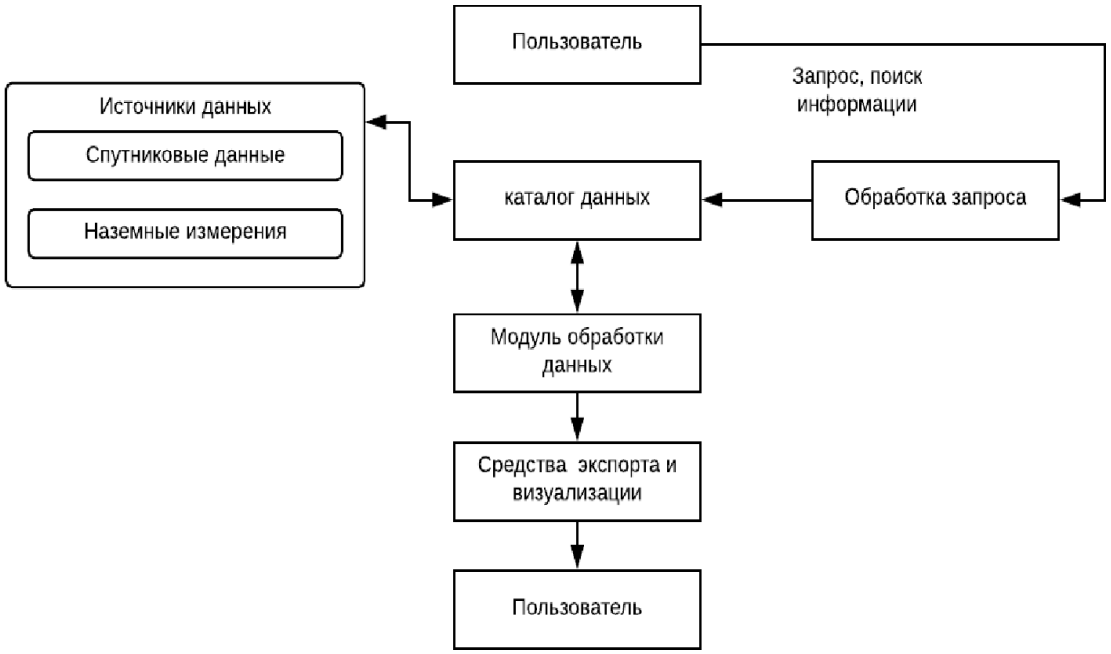



Рис. 1. Общая модель работы ГИС.




Рис. 2. Этапы создания каталога пространственных данных.

Одним из основных компонентов ГИС является каталог данных, который реализован в виде базы данных под управление СУБД Postgresql. Главным критерием создания каталога является специфика хранимых данных, которые имеют географическую привязку, и результаты обработки спутниковых снимков записываются в каталог в виде векторных полигонов, результаты наземных измерений выражены в виде точек. Согласно специфике и организации поступления данных в каталог предложена следующая схема его реализации. *Сбор данных* для хранилища данных водной тематики представляет собой низовой уровень аккумуляции в одном месте «сырой» информации. В качестве источников информации выступают традиционные карты разных масштабов, ДДЗ, данные мониторинга и натурных наблюдений. *Обработка* предполагает согласование данных для обеспечения их целостности и непротиворечивости, т.е. приведение к единому формату представления, системе координат и т.п. В данном случае, при создании каталога главным требованием для загрузки цифровых карт и ДДЗ являлось наличие единой проекции. Поэтому на этом этапе осуществляется перевод материалов в цифровую форму (векторизация карт, обработка ДДЗ и т.п.), привязка их к конкретной системе координат. В итоге формируются базы данных векторных материалов и ДДЗ, а также базы атрибутивных данных.

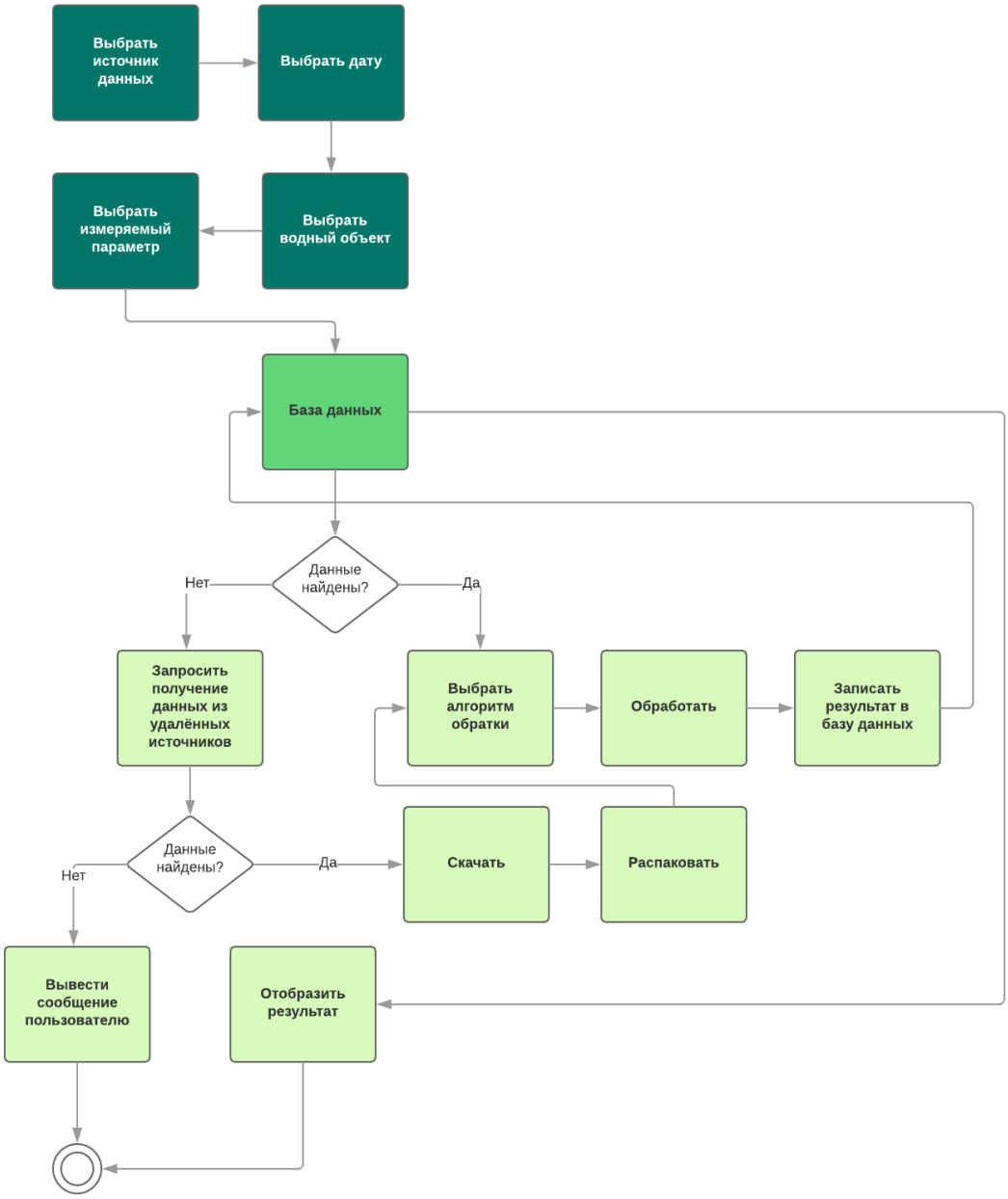



Рис. 3. Последовательность операций обработки гидрологических данных в информационной системе.

На этапе *загрузки в систему* подготовленных пространственных данных происходит отслеживание повторяемости материалов, связка таблиц данных по первичным и вторичным ключам. Работа с пространственными объектами в базе данных ГИС основывается на расширении PostGIS. PostGIS позволяет хранить различные типы пространственных данных и выполнять к ним SQL запросы. Взаимодействие базы данных с другими компонентами системы осуществляется через ORM (Object-Relational Mapping, объектно-реляционное отображение). Визуализация данных осуществляется утилитой MapServer который выступает в роли WMS-сервиса. Помимо визуализации присутствует функция экспорта файлов в виде векторных полигонов Shape формата. На рис. 3 показана схема операций обработки данных и ГИС.

**Оценка содержания хлорофилла «а» в поверхностном слое водоемов.** В целом процедуру обработки спутниковых данных для решения рассматриваемой задачи можно разделить на следующие этапы:

1. Атмосферная коррекция;
2. Вычисление спектрального индекса NDCI;
3. Вычисление концентрации содержания хлорофилла «а» в поверхностном слое;
4. Географическая привязка и экспорт, полученных результатов в формат Geotiff.

Для реализации указанных этапов был использован вычислительный пакет ACOLITE, предназначенный для обработки данных Landsat (5/7/8) и Sentinel-2 (A/B). ACOLITE позволяет производить атмосферную коррекцию по алгоритму «Dark spectrum fitting» и содержит набор алгоритмов для определения концентрации хлорофилла «а» в поверхностном слое внутриконтинентальных водных объектов [7]. Посредством программных интерфейсов вычислительный комплекс может работать в виде набора библиотек алгоритмов. На основе этого был разработан независимый модуль геоинформационной системы мониторинга параметров внутриконтинентальных водных объектов. Его работу можно разделить на следующие этапы. Вычисление спектрального индекса NDCI (Normalized Difference Chlorophyll Index), который учитывает различия в отражении в инфракрасной и зеленой областях электромагнитного спектра. Индекс NDCI позволяет с относительно небольшими временными и вычислительными ресурсами оценить состояние поверхностного слоя водоемов, данная процедура проста в технологической, программной реализации. Для вычисления концентрации хлорофилла бы использован модуль chl_o2, реализующий алгоритм отношения синего и зелёного спектральных каналов [7]. В табл. 1 и 2 показаны данные для озера Красиловского и озера Иткуль (Алтайский край).

В таблицах 1 и 2 показаны средние значения измерений по области водоемов, где были произведены измерения. Некоторая расходимость между данными спутниковых измерений и экспедиционными может быть обусловлена спецификой водных объектов. Необходимо отметить, что точность оценки концентрации хлорофилла «а» в поверхностном слое зависит от атмосферных процессов в момент измерений, алгоритмов обработки данных и времени суток, в которое были произведены спутниковые измерения [8].

Табл. 1. Концентрация хлорофилла «а» в поверхностном слое озера Иткуль по экспедиционным и спутниковым данным.

| **Дата** | **Экспедиционные измерения ($\frac{mg}{m^3}$)** | **Дата** | **Sentinel-2 ($\frac{mg}{m^3}$)** |
|---|---|---|---|
| 10.05.2018 | 1,8 | 11.05.2018 | 2.4 |
| 02.08.2018 | 6,1 | 03.08.2018 | 6.7 |
| 12.10.2018 | 7,1 | 12.10.2018 | 8.2 |

Табл. 2. Концентрация хлорофилла «а» в поверхностном слое озера Красиловское по экспедиционным и спутниковым данным.

| **Дата** | **Экспедиционные измерения ($\frac{mg}{m^3}$)** | **Дата** | **Sentinel-2 ($\frac{mg}{m^3}$)** |
|---|---|---|---|
| 11.05.2018 | 3,0 | 11.05.2018 | 2.8 |
| 02.08.2018 | 10,0 | 03.08.2018 | 10,5 |
| 12.10.2018 | 35,9 | 12.10.2018 | 37,0 |

**Заключение.** Предложенная геоинформационная система позволяет производить регулярный мониторинг параметров внутриконтинентальных водных объектов по данным спутниковой съемки, систем наземного мониторинга и результатов экспедиционных работ. Указанная ГИС может быть использована для решения широкого спектра прикладных и фундаментальных задач гидрологии внутриконтинентальных водных ресурсов.

## ЛИТЕРАТУРА


[1] Решение VII Всероссийского гидрологического съезда от 13.01.2014 г. [Электронный ресурс]. URL: **http://www.meteorf.ru/press/news/6572** (дата обращения 10.06.2019).

[2] *Бычков И.В., Васильев С.Н., Кузьмин В.А., Ступин Г.В.* Проект интегрированной геоинформационной системы ИНЦ СО РАН для поддержки фундаментальных исследований // Вычислительные технологии. 1998. Т. 3, № 5. С 11-17.

[3] *Бешенцев А.Н., Тулохонов А.К., Бычков И.В., Ружников Г.М., Батуев А.Р.* Создание геоинформационных ресурсов для междисциплинарных исследований дельтовых территорий побережья озера Байкал // Материалы Международной конференции «ИнтерКарто/ИнтерГИС». 2013. 19 (1), С. 15-18.

[4] *Сухоруков Б.Л., Ковалева Т.Е., Новиков И.В.* Оценка трофности водных объектов по данным дистанционной спектрометрии высокого разрешения видимого диапазона // Водные ресурсы. 2017. Т. 44. № 1. С. 79-90.

[5] *Boronina L., Sadchikov P., Tazhieva S. et al.* Effectiveness of the automation selection of water treatment technology in a particular water source // Advanced Materials Research. 2014. Vols. 1073-1076. Pp. 1039-1042.

[6] *Донцов А.А., Суторихин И.А.* Специализированная геоинформационная система автоматизированного мониторинга рек и водоемов // Вычислительные технологии. 2017. Т. 22, № 5. С. 39-46.

[7] ACOLITE [Electronic resource]. URL: **http://www.odnature.naturalsciences.be**

[8] *Левин И.М., Копелевич О.В.* Корреляционные соотношения между первичными гидрооптическими характеристиками в спектральном диапазоне около 550 нм. // Океанология. 2007. Т. 47. № 3. С. 374-379.

# ОБРАБОТКА СИГНАЛОВ В ПРИЕМНОЙ СИСТЕМЕ С ПРОСТРАНСТВЕННО-РАСПРЕДЕЛЕННЫМИ ПРИЕМНЫМИ ЭЛЕМЕНТАМИ

*Зима Д.Н., Соколова Д.О., Спектор А.А.*
Новосибирский государственный технический университет, Новосибирск

Рассмотрены вопросы обработки сигналов и подавления помех в системах приема с пространственно-распределенными элементами. Алгоритмы обработки строятся на теории марковских процессов. Обработка учитывает широкополосный характер исследуемых пространственно-временных сигналов. В качестве полезного сигнала взят сигнал с линейной частотной модуляцией.
*Ключевые слова: пространственно-распределенные приемные элементы, широкополосный сигнал, ЛЧМ-сигнал, марковский случайный процесс.*

**Введение.** Задача обнаружения широкополосного (в пространственно-временном смысле) сигнала пространственно-распределенными элементами представляется трудоемкой, поскольку только в некоторых случаях можно разделить временную и пространственную фильтрации [1]. Также, применение оптимального байесовского критерия обнаружения сигнала требует обращения корреляционных матриц помехи порядка, соответствующего квадрату произведения числа приемных пространственно-распределенных элементов $k_a$ и числа моментов времени $k_t$. Это требует трудоемких расчетов, приводит к появлению ошибок, а при малых углах прихода помехи корреляционная матрица близка к вырожденной матрице. Данные проблемы можно решить, если использовать модель марковского случайного процесса для описания помехи на пространственно-распределенных антенных элементах.

**Модель пространственно-временного процесса.** Поставленная задача обнаружения широкополосного сигнала в пространственно-временном смысле на фоне помехи рассмотрена для случая линейки пространственно-распределенных антенных элементов. Часто используется в радиолокации в качестве полезного сигнала сигнал с линейной частотной модуляцией (1)

$$s_i(t) = S_0 \cos\left[\omega_0\left(t-(i-1)\tau_0 + b\left(t-(i-1)\tau_0-\frac{\tau_и}{2}\right)^2+\varphi\right)\right], \quad (1)$$

где $\omega_0$ – несущая частота радиосигнала, $b$ – скорость изменения частоты при ЛЧМ, $\tau_и$ – длительность импульсного сигнала, $\tau_0$ – относительные задержки между соседними элементами пространственно-распределенной приемной системы, зависящие от угла прихода сигнала $\Theta_С$ и расстояния между антенными элементами, $0 \le t \le \tau_и$. Данный сигнал заложен в согласованный фильтр временной части пространственно-временного фильтра.

Пространственно-временной процесс при фиксации $t$ преобразуется в пространственное колебание, имеющее гармонический характер. На рис. 1 представлена картина формирования данного колебания. Считаем, что источник помехи расположен на значительном удалении от приемных антенных элементов. Имеем плоский фазовый фронт волны и наблюдаем одно и то же колебание на всех линейно распределенных антенных элементах, задержанное на время $\tau_0$ между соседними элементами.

Считаем, что помеха в пространственно-временном представлении имеет вид, представленный ниже.

$$u_i(t) = \Xi(t-(i-1)\tau_0)\cos[\omega_0(i-1)\tau_0 + \Psi(t-(i-1)\tau_0)],$$

Данная помеха при фиксированных $t$ и $\Theta_П$ является колебательной случайной функцией номера антенного элемента, а параметр в формуле (2) определяет среднюю нормированную частоту пространственных колебаний:

$$\omega_н = 2\pi\frac{d\sin\Theta_П}{c}. \quad (2)$$

Таким образом, на рис. 2 представлена картина пространственных колебаний на элементах линейной антенны без учета флюктуаций, вызванных модуляциями $\Xi(t-(i-1)\tau_0)$ и $\Psi(t-(i-1)\tau_0)$.

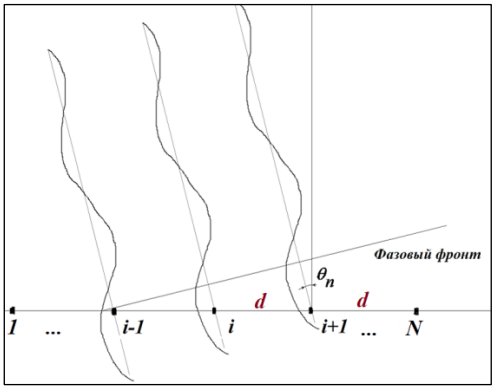



Рис. 1. Формирование пространственно-временного колебания.

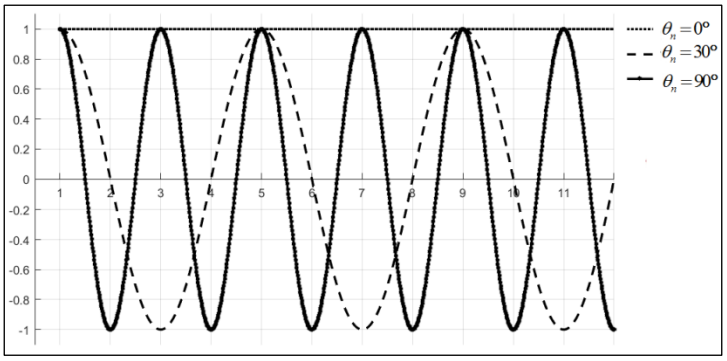



Рис. 2. Картина пространственных колебаний на линейке антенных элементов.

При этом при последовательном попадании на линейку пространственно-распределенных элементов различных участков широкополосного сигнала в дискретном спектре пространственного сигнала видно только одну частоту, не зависимо от наличия модуляции. Данное свойство удобно при наличии нескольких сигналов помехи с разными направлениями прихода. Данный характер сохраняется при рассмотрении пространственно-временного колебания. На рис. 4 представлен пространственно-временной спектр одного помехового колебания после его прохождения через согласованный фильтр с КЧХ соответствующей (1).

Следует отметить, что сигнал широкополосный во времени оказывается узкополосным в пространстве, т.е. база пространственного сигнала близка к 1.

**Марковская модель помехи.** Применялась модель на основе модели авторегрессии, т.е. линейного предсказания [2]:

$$u_i = \sum_{k=1}^{K} a_k u_{i-k} + \nu_i, \qquad (3)$$

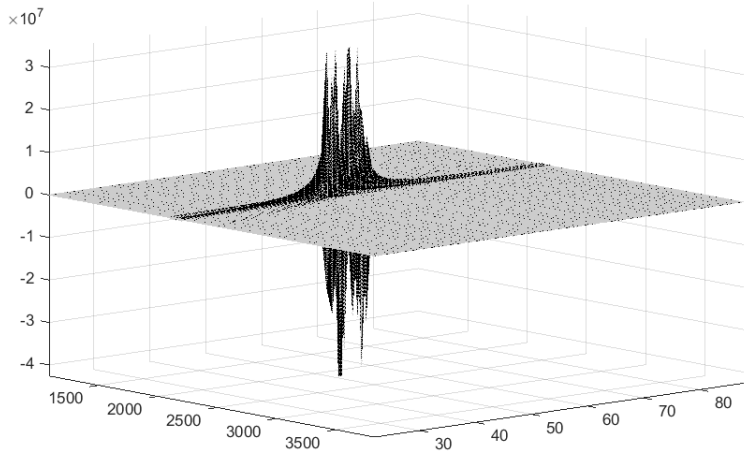



Рис. 3. Пространственно-временной спектр помехи на линейке антенных элементов.

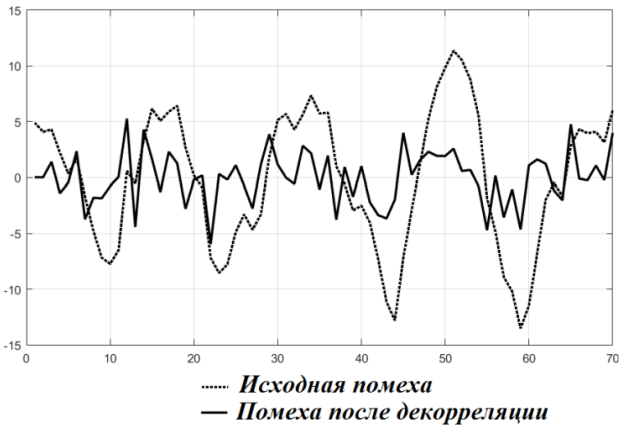



Рис. 4. Временная реализация помехи.

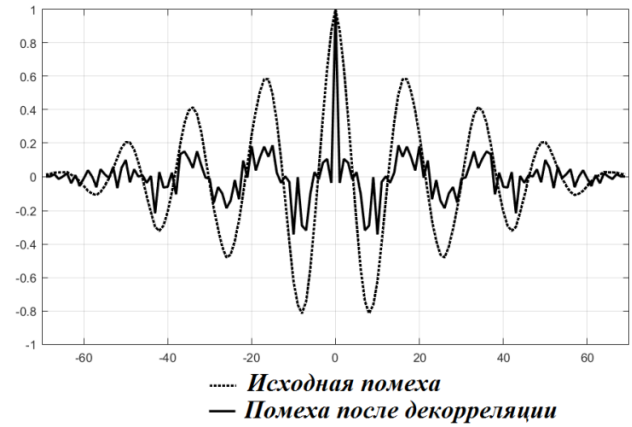



Рис. 5. АКФ помехи.

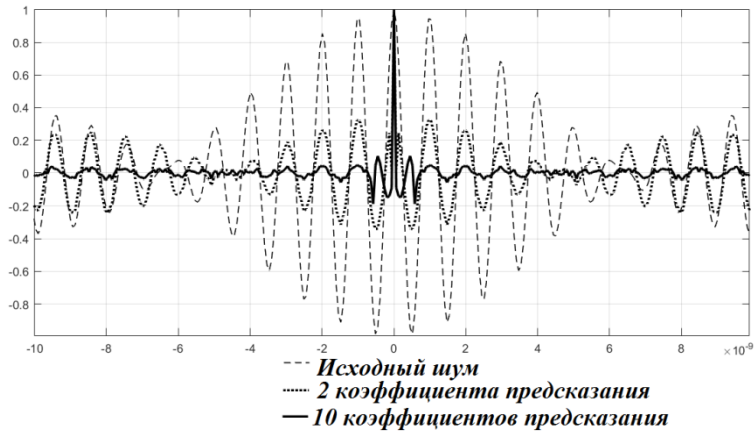



Рис. 6. АКФ шума при разном количестве коэффициентов предсказания.

Поскольку необходимо описывать квазигармонические процессы, минимальным порядком марковской модели для формирования пространственного колебания помехи является второй порядок

$$u_i = a_1 u_{i-1} + a_2 u_{i-2} + \nu_i, \qquad (4)$$

где $\nu_i$ – шум предсказания, который представляет собой информационный эквивалент исходной помехи, $a_i$ – коэффициенты предсказания. Для определения этих коэффициентов предсказания используется метод моментов, который основан на использовании соотношений, связывающих искомые параметры с моментами наблюдаемых процессов [4].

В качестве показателя соответствия помехи марковской модели на рис. 4 приведена временная реализация помехи до и после проведения процедуры декорреляции (4). После применения обработки из рис. 4 видно, что частота пересечения сигналом нулевого уровня увеличилась, что может служить признаком уменьшения корреляции между соседними отсчетами. На рис. 5 приведены нормированные АКФ помехового сигнала до и после проведения декорреляции, где видно, что АКФ результирующего информационного шума ближе к АКФ белого гауссовского шума (стремиться к дельтаобразному виду).

На рис. 6 представлены АКФ исходной помехи, помехи после декорреляции согласно (4) при использовании двух коэффициентов предсказания, и помехи после декорреляции согласно (3) при использовании десяти коэффициентов предсказания. Из графика видно, что при увеличении количества коэффициентов предсказания и объема выборки помехи качество обработки увеличивается.

Если сигнал принимается с известного направления на фоне белого шума, то его обработка может быть всегда разделена на пространственную и временную [1,5]. Тогда в данном случае совместное распределение отсчетов помехи на всех пространственно-распределенных элементах для модели с двумя коэффициентами предсказания:

$$W_0(U) = W_0(u_1, u_2, \dots, u_N) = W_0(u_1, u_2) \prod_{i=3}^{N} \pi(u_i | u_{i-1}, u_{i-2}), \quad (5)$$

Решающая статистика при (4) и (5), учитывая [3], будет иметь вид :

$$Z(U) = {U_{22}}^T {R_2}^{-1} S_{22} + \sum_{i=3}^{N} \left[ {U_{3i}}^T {R_3}^{-1} S_{3i} - {U_{2i}}^T {R_2}^{-1} S_{2i} \right], \quad (6)$$

где $U_{2i}$, $U_{3i}$, $S_{2i}$, $S_{3i}$ – укороченные векторы из отсчетов сигнального и помехового колебаний при фиксированных $t$ и $\Theta_{\Pi}$, $R_2$ и $R_3$ – корреляционные матрицы укороченных векторов. При теоретическом исследовании критерием качества предложенного метода являлось улучшение отношения сигнал/помеха (ОСП) после пространственной обработки к входному отношению сигнал/помеха. На рис. 7 отображено улучшение ОСП в зависимости от направления прихода одного помехового колебания и при фиксированном направлении на полезный сигнал. На рис. 8 отображено улучшение ОСП в зависимости от количества антенных элементов в линейке пространственно-распределенных элементов при наличии одного помехового колебания и фиксированном направлении на полезный сигнал.

При условии независимости совокупности отсчетов в данный момент времени $t$ перейдем к пространственно-временному алгоритму обнаружения, т.е. имеем вектор:

$$U^{\langle t \rangle} = \left\| u_1{}^{\langle t \rangle}, u_2{}^{\langle t \rangle}, \dots, u_N{}^{\langle t \rangle} \right\|^T. \quad (7)$$

Решающая статистика согласно (6) и (7) имеет вид:

$$Z(U) = \sum_{t=1}^{M} Z^{\langle t \rangle}\left(U^{\langle t \rangle}\right),$$

$$Z(U) = \sum_{t=1}^{M} \left\{ {U_{22}}^{\langle t \rangle^T} {R_2}^{-1} {S_{22}}^{\langle t \rangle} + \sum_{i=3}^{N} \left[ {U_{3i}}^{\langle t \rangle^T} {R_3}^{-1} {S_{3i}}^{\langle t \rangle} - {U_{2i}}^{\langle t \rangle^T} {R_2}^{-1} {S_{2i}}^{\langle t \rangle} \right] \right\}.$$

Таким образом, происходит временное накопление выше описанной пространственной обработки.

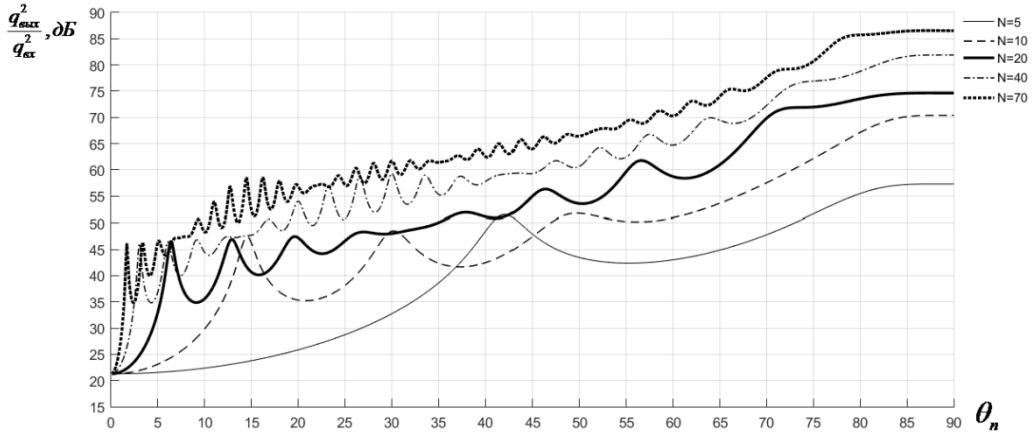



Рис. 7. Зависимость улучшения ОСП от направления прихода помехи.

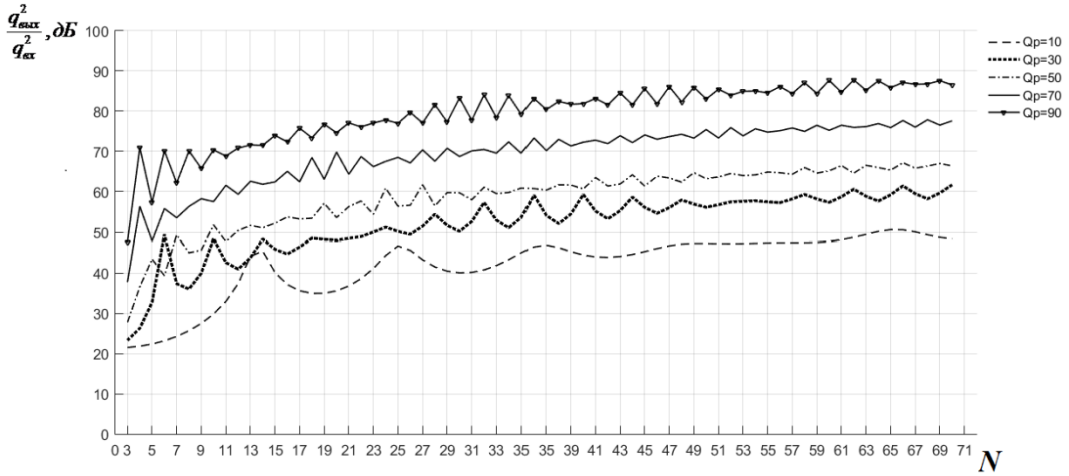



Рис. 8. Зависимость улучшения ОСП от количества антенных элементов.

Обработка сигнала при наличии нескольких помех требует большего количества коэффициентов предсказания в (3).

**Заключение.** В работе рассмотрены вопросы обработки сигналов на фоне помех в системах приема с пространственно-распределенными антенными элементами. Алгоритмы обработки построены на модели марковских случайных процессов, что дает возможность факторизации пространственно-временной обработки, а, следовательно, приводит к упрощению реализации алгоритма. Использование пространственного спектра позволяет определить количество и направление прихода активных помех.

## ЛИТЕРАТУРА


[1] *Далматов А.Д., Елисеев А.А., Лукошкин А.П., Оводенко А.А., Устинов Б.В.* Обработка сигналов в радиотехнических системах. Ленинград: Ленинградский университет, 1987. 400 с.

[2] *Казаков В.А.* Введение в теорию марковских процессов и некоторые радиотехнические задачи. Москва: «Советское радио», 1973. 232 с.

[3] *Левин Б.Р.* Теоретические основы статистической радиотехники. Книга первая. Москва: «Советское радио», 1969. 752 с.

[4] *Соколова Д.О.* Непараметрическое обнаружение и классификация в сейсмических системах охраны: Дис. канд. техн. наук, Новосибирск, НГТУ, 2013. 147 с.

[5] *Якубов В.П.* Статистическая радиофизика. Томск: НТЛ, 2006. 132 с.

# ГЕОИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ОПЕРАТИВНОЙ ОЦЕНКИ ЗАГРЯЗНЕНИЯ АТМОСФЕРЫ

*Кадочников А.А., Токарев А.В., Якубайлик О.Э.*
Институт вычислительного моделирования СО РАН, Красноярск
Федеральный исследовательский центр «Красноярский научный центр СО РАН», Красноярск

Рассматриваются особенности разработки системы экологического мониторинга на основе интеграции данных из различных источников. Реализован оперативный сбор данных с автоматизированных постов наблюдений за состоянием атмосферного воздуха и метеорологической информацией. Используется технологическая платформа геопортала, предназначенная для сбора, обработки и представления данных различных наблюдений.

*Ключевые слова: веб-картография, ГИС, окружающая среда, пространственные данные, система мониторинга.*

**Введение.** Качество атмосферного воздуха – важнейший экологический фактор, определяющий здоровье населения и состояние экосистем. Системы мониторинга окружающей среды и атмосферного воздуха в частности создаются и внедряются во многих городах мира как за рубежом, так и в России [1,2]. В г. Красноярске ряд организаций федерального и региональных уровней на основе собственных методик и технологий обработки данных выполняет мониторинг состояния загрязнения атмосферного воздуха. Разнообразие используемых решений, межведомственная и организационная разобщенность приводят к тому, что комплексный анализ и оперативная оценка всего массива регистрируемой информации представляются в настоящее время технически затруднительными и практически не проводятся. Собираемые данные часто публикуются с запозданием и не понятны для широкой общественности.

Использование международных стандартов значительно упростило бы сбор и анализ данных наблюдений из различных источников. Однако сегодня ситуация в Красноярске и России в целом далека от идеальной. В большинстве случаев доступ к данным оперативных наблюдений ограничен и нет специализированных инструментов для получения таких данных.

Необходимо отметить, что сегодня краевые власти, совместно с федеральными службами предпринимают попытку создания единой системы мониторинга качества атмосферного воздуха. Пример, «Единая информационная система мониторинга загрязнения атмосферного воздуха г. Красноярска» (http://www.feerc.ru/uisem/krasnoyarsk). Однако с помощью такой системы нельзя анализировать оперативные данные, т.к. система предоставляет только среднесуточные данные по загрязнению атмосферного воздуха.

**Методы и инструменты исследования.** В мировом сообществе инициатива по созданию сети датчиков Sensor Web Enablement (SWE) от организации Open Geospatial Consortium (OGC) давно приобрела особое значение благодаря своей зрелости и широкой поддержке со стороны ученых и промышленности. SWE стремится предоставить открытые стандарты и протоколы для улучшения взаимодействия между различными платформами и поставщиками данных. Инициатива SWE стремится облегчить поиск, повысить доступность и обеспечить контроль различных датчиков в реальном времени через Интернет [3-5]. В настоящее время семейство SWE состоит из семи стандартов: Sensor Model Language (SensorML), Observations & Measurements (O&M), Transducer Model Language (TML), Sensor Observation Service (SOS), Sensor Planning Service (SPS), Sensor Alert Service (SAS), Web Notification Service (WNS).

Наиболее интересным и полезным инструментом для анализа экологической ситуации в городе Красноярск был бы инструмент на основе стандарта Sensor Observation Service (SOS). Этот стандарт описывает сервис, обеспечивающий доступ к результатам измерений от датчика или сенсорной сети, обеспечивает программное обеспечение информацией о возможностях датчика и качестве измерений, также обеспечивает доступ к параметрам датчика. Другими

словами, SOS группирует коллекцию возможных гетерогенных датчиков и предоставляет их измерения через стандартизированный сервисный интерфейс. Спецификация SOS определяет операции, предлагаемые определенным датчиком, при этом минимальный набор методов включает GetCapabilites, DescribeSensor и GetObservation, которые возвращают информацию о наблюдениях и измерениях, поддерживаемых SOS. Типы данных, предоставляемых датчиком, и сами типы датчиков могут быть получены из реестра датчиков.

**Технология сбора данных.** В результате анализа текущей ситуации в Красноярске сформирован список из нескольких источников оперативной информации по состоянию атмосферного воздуха, которую можно получать и анализировать. Однако эти данные приходится извлекать специальными инструментами и в большинстве случаев сбор данных из внешних источников можно разделить не несколько групп:

1. Данные представлены в виде веб-сервисов, которые передают информацию в структурированном виде, например, в форматах csv, json, xml. Такие данные легко обрабатывать и проверять на наличие ошибок. К этому блоку относятся данные предоставляемые в рамках стандарта OGC SOS.
2. Специальных веб-сервисов для получения данных нет, но есть служебные сервисы ресурса необходимые для его работы. Например, сервисы для построения графиков или вывода табличных данных в веб-интерфейсе, инструменты экспорта данных и др. Такие сервисы тоже можно использовать для получения данных, но при этом необходимо разработать ряд инструментов для анализа данных, включающих проверку изменения структуры данных, изменение внутренних идентификаторов, формата вывода данных, размерность данных и т.д.
3. Данные представлены в виде html страниц Интернет-ресурса. Для получения таких данных требуется провести анализ страницы с целью поиска некоторых опорных текстовых блоков, с помощью которых можно всегда найти необходимую информацию на странице. В таком случае из страницы извлекается определенный текстовый блок и выполняется его дальнейшая обработка. В некоторых случаях с такими фрагментами страницы можно работать как с xml, что упрощает извлечение информации.

В ряде случав может использоваться комбинация из таких групп, т.к. для работы сервисов может понадобиться служебная информация, содержащаяся в самой странице. Отдельное внимание необходимо уделить структурным изменениям в полученных данных, разрабатывая серию блоков проверки, включающих изменение количества датчиков и станций наблюдения, целостность и полноту переданных данных и т.п. Для оперативной доработки программных модулей сбора должны быть созданы средства логгирования текстовой информации и оповещения, например, по электронной почте. Такие средства позволять получать информацию об изменениях на удаленном ресурсе и его состоянии постоянно. Оперативно внося исправления в программный код можно обеспечить целостность и корректность полученной информации.

В мире существует ряд систем для доступа к данным наблюдений. Например, платформа Smart Emission (https://data.smartemission.nl/data), проект OpenSensorHub (https://opensensorhub.org), проект 52°North (https://52north.org), проект istSOS OSGeo (http://istsos.org) и др.

**Результаты.** На базе геопортала Института вычислительного моделирования СО РАН разработан блок научно-исследовательского мониторинга, предназначенный для сбора, обработки и представления данных различных наблюдений [6]. Показатели с внешних источников данных собираются на сервере сбора SensorCollector (http://gis.krasn.ru/sc). Источниками данных могут быть как отдельные датчики, так и внешние базы данных и информационные системы через дополнительные адаптеры. На основе собранных данных выполняется автоматическая агрегация с разными временными интервалами, поддерживается вычисление производных показателей. Реализован программный интерфейс (API) для наполнения и извлечения данных и базовый веб-интерфейс для управления метаданными и навигации по хранилищу.

Поддерживается разграничение доступа к объектам системы в зависимости от прав пользователя. Сервис реализован на языке сценариев PHP 5 с использованием фреймворка Yii, для хранения данных применяется СУБД PostgreSQL с модулем расширения PostGIS.

Организация доступа к данным наблюдений осуществляется различными приложениями и сервисами, включающими просмотр табличных данных, экспорт, просмотр данных на картах с возможностью выбора временных интервалов и доступ с помощью общепринятых стандартов (OGC WMS, SOS). Общая схема сбора данных представлена на рис. 1.

На данный момент на портале реализован сбор оперативных данных с нескольких источников и созданы следующие разделы:

- Система мониторинга состояния атмосферного воздуха в г. Красноярске.
- Подсистема мониторинга атмосферного воздуха КВИАС КГБУ «ЦРМПиООС».
- Данные системы мониторинга ФГБУ «Среднесибирское УГМС».
- Данные гидрологических наблюдений в Красноярском крае.
- и др.

Осуществляется сбор данных с приборов Федерального исследовательского центра «КНЦ СО РАН». Данные собираются с интервалом в одну минуту с 24 постов с помощью API для доступа к данным устройств измерений автоматизированной информационной системы мониторинга качества воздуха. Приборы разработаны компанией «CityAir» (https://cityair.io) и позволяют получать данные измерений содержания взвешенных веществ в атмосферном воздухе ($PM_{2.5}$, $PM_{10}$) и основных метеорологических параметров: температуры, влажности и давления. В комплект поставки входили восемь дополнительных модулей для стационарных постов с возможностью сбора данных по некоторым газам. Четыре модуля расширения G1, осуществляющие измерение концентрации газов $NO_2$, $O_3$, CO в воздухе, и четыре модуля G2 осуществляющие, измерение концентрации $H_2S$, $SO_2$ в воздухе. Эти модули устанавливаются в качестве расширения к основным приборам и передают данные вместе с ними. Накоплен архив данных по концентрации газов в воздухе г. Красноярск за три первых месяца 2019 года. Однако в результате эксплуатации выяснилось, что для получения корректных данных модули нужно постоянно калибровать и сделать это без специалистов компании «CityAir» нельзя. В данный момент модули отключены и дорабатываются.

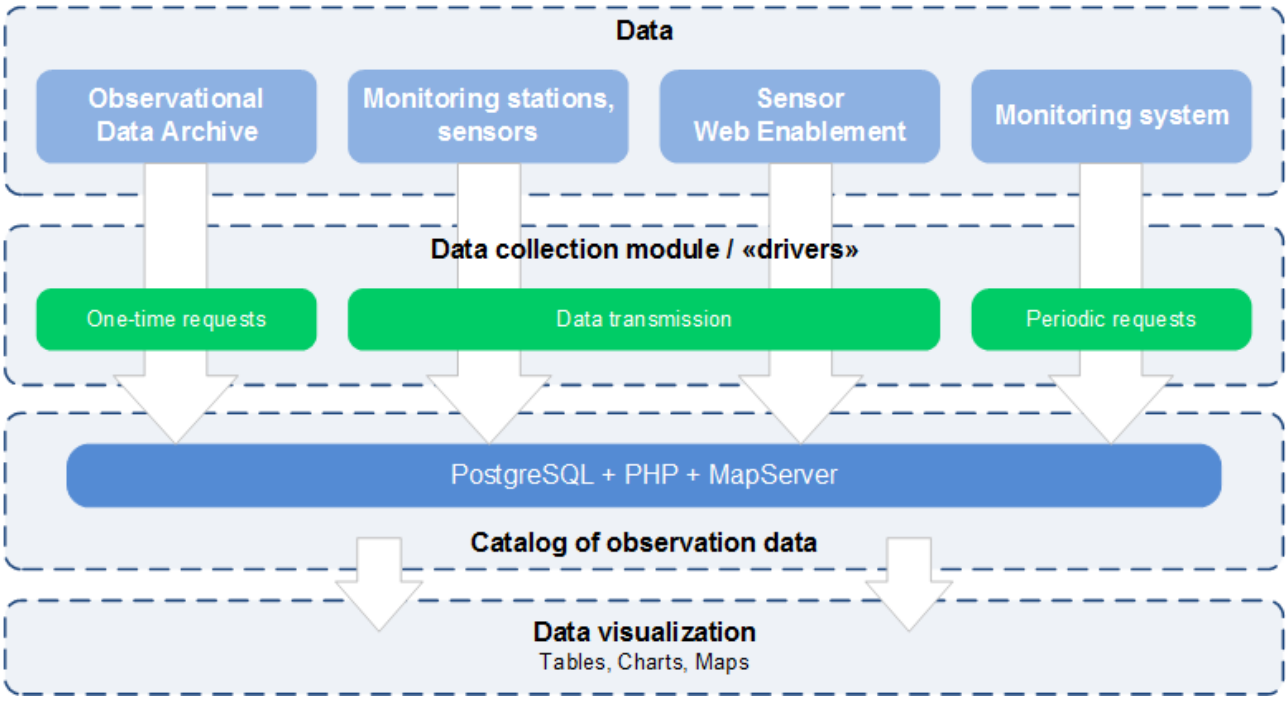



Рис. 1. Схема сбора данных наблюдений.

Настроен поток данных об атмосферном воздухе из Краевой ведомственной информационно-аналитической системы о состоянии окружающей среды Красноярского края (КВИАС) (http://krasecology.ru/Air). Наблюдения за качеством атмосферного воздуха проводятся на 11 автоматизированных постах наблюдений, 9 из которых расположены в Красноярске. Для загрузки данных с сайта КВИАС подготовлен программный модуль с соответствующим «драйвером» для обработки и преобразования входных данных, обеспечивающий периодическую загрузку данных наблюдений от этой системы через веб-сервис. В данный момент большинство загружаемых данных по стационарным постам наблюдения на удаленном сервере формируется каждые 20 минут. Данные с передвижных лабораторий появляются с частотой раз в сутки, после обработки их сотрудниками.

В автоматическом режиме загружаются метеоданные с сайта ФГБУ Среднесибирское управление по гидрометеорологии и мониторингу окружающей среды (http://meteo.krasnoyarsk.ru). Метеоданные, в реальном времени поступающие от автоматических станций, публикуются на сайте с частотой 1-3 часа. Данные собираются по следующим показателям: атмосферное давление, влажность, направление ветра, скорость ветра, температура воздуха. Архив ведется с начала 2013 года.

Собираются данные проекта Nebo community (https://nebo.live), созданного активистами из Красноярска. Проект осуществляет сбор с 22 приборов данных измерений содержания взвешенных веществ в атмосферном воздухе $PM_{2.5}$ и основных метеорологических параметров: температуру и влажность. В конце прошлого года, владельцы сервиса стали выкладывать архивы данных за прошедший период. Однако эти архивы имеют пробелы во времени, но иногда позволяют заполнить пропущенные сервисом импорта данные.

Пример пользовательского приложения, построенного на основе компонентов и API геопортала, показан на рис. 2. На карте отображены данные о концентрации взвешенных частиц $PM_{2.5}$ в атмосфере города, полученные с различных систем наблюдения. Форма пиктограмм на карте характеризует принадлежность к источнику данных, а цвет отражает концентрацию примеси.

**Заключение.** Для оперативного мониторинга и анализа экологической ситуации в Красноярске организован сбор данных из разных источников в единую систему «Данные оперативного мониторинга» Института вычислительного моделирования СО РАН.

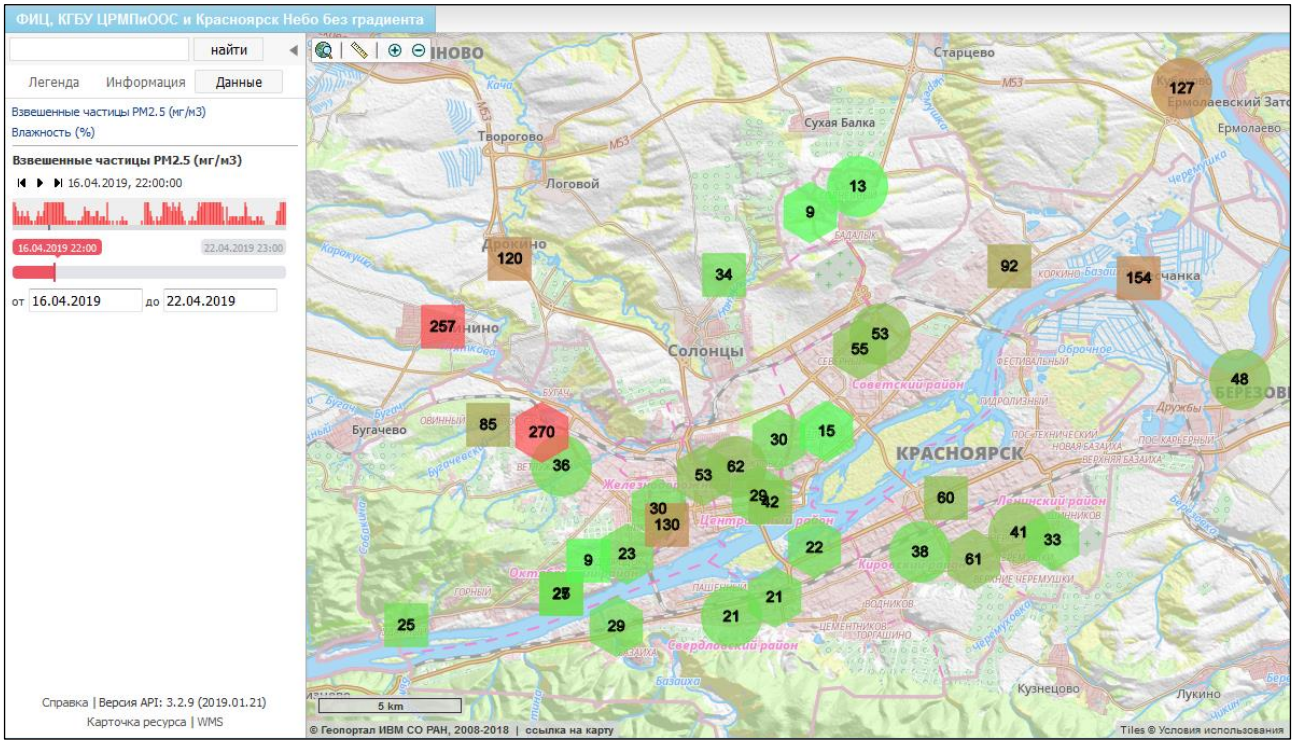



Рис. 2. Карта с данными наблюдений по $PM_{2.5}$ в г. Красноярск на 16.04.2019.

Разработанное программное обеспечение для импорта данных наблюдений с различных ресурсов позволяет автоматически выполнять загрузку данных, независимо от того, как организован доступ к этим данным. При любых изменениях и ошибках импорта отправляются сообщения по электронной почте, позволяющие оперативно внести изменения в программу импорта или в данные о датчиках и их составе. В будущем рассматривается возможность организовать доступ к данным наблюдений в городе Красноярск с помощью международного стандарта OGC SOS. К сожалению, в России этот стандарт практически не развит у поставщиков данных мониторинга.

## ЛИТЕРАТУРА


[1] *Manuel A. Regueiro, José R.R. Viqueira, Christoph Stasch, José A. Taboada.* Semantic mediation of observation datasets through Sensor Observation Services // Future Generation Computer Systems. 2017. Vol. 67, P. 47-56.

[2] *Azemov D.* Saint-Petersburg air pollution monitoring system // Environment of Saint-Petersburg. 2016. Vol. 2 (2). P. 8-14.

[3] *Botts M., Percivall G., Reed C., Davidson J.* 2008 OGC® Sensor Web Enablement: Overview and High Level Architecture. In: Nittel S, Labrinidis A, Stefanidis A (eds) GeoSensor Networks. GSN 2006. Lecture Notes in Computer Science. Vol. 4540. Springer, Berlin, Heidelberg.

[4] *Boulos K. et al.* Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples // International Journal of Health Geographics. 2011. Vol. 10. P. 67.

[5] *Schmitt R.H., Voigtmann C.* Sensor information as a service – component of networked production // J. Sens. Sens. Syst. 2018. Vol. 7. P. 389-402.

[6] *Yakubailik O.E., Kadochnikov A.A., Tokarev A.V.* Web geographic information system and the hardware and software ensuring rapid assessment of air pollution // Optoelectronics, Instrumentation and Data Processing. 2018. Vol. 54, N 3. P. 243-249.

# ПРОГНОЗ ИЗМЕНЕНИЯ КЛИМАТА АЛТАЕ-САЯНСКОЙ ГОРНОЙ СТРАНЫ ДО 2030 ГОДА

*Кирста Ю.Б., Ловцкая О.В., Пузанов А.В.*
Институт водных и экологических проблем СО РАН, Барнаул

Разработан метод пространственного обобщения температур приземного слоя воздуха и осадков с возможностью приложения в ГИС и реанализе. На примере Алтае-Саянской горной страны показано, что их относительные изменения в процентах от среднемноголетних значений имеют единообразную динамику по ее территории. Дан прогноз изменения климата по всей горной стране к 2030 году. Температуры в январе уменьшатся (~20%), в марте и апреле увеличатся (>20%), а в остальные месяцы останутся прежними. Прогнозируемые изменения осадков имеют разнонаправленный характер по месяцам года.
*Ключевые слова: температура воздуха, осадки, пространственное обобщение, горная территория, Алтай, Саяны.*

**Введение.** Для гор характерна сложная вертикальная и горизонтальная дифференциация метеорологических факторов. В связи с этим, анализ многолетних изменений последних трудоемок и не всегда возможен из-за отсутствия необходимых наблюдений. Разработанный метод пространственного обобщения среднемесячных температур приземного слоя воздуха и месячных осадков обеспечивает адекватную оценку их помесячной и многолетней динамики при опосредованном учете пространственной дифференциации [1]. Такая оценка может быть сделана для произвольных участков горной территории, в том числе тех, где метеорологические наблюдения вообще не проводились.

Выбранная в качестве тестового региона Алтае-Саянская горная страна состоит из системы хребтов, веерообразно раскрывающейся в северо-западном направлении. Это самая высокогорная область Сибири, высота которой меняется от 300 до 4500 м над уровнем моря. Важной особенностью страны является разнообразие ее ландшафтов: гляциально-нивальных, тундровых, альпийских и субальпийских лугов, лесных, степных и полупустынных. Она служит горным водосборным регионом для таких крупных рек Сибири как Обь и Иртыш.

Основной задачей работы являлся многолетний прогноз изменений температур воздуха и осадков на рассматриваемой территории. Также предполагалась возможность выполнения реанализа при малом количестве исходных данных о их помесячной и межгодовой динамике.

**Исходные данные и методика исследования.** Для оценки климатической обстановки в Алтае-Саянской горной стране использовались 11 опорных метеостанций с наблюдениями 1951-2017 гг., продолжающимися в настоящее время (табл. 1). Поскольку климат страны в значительной степени формируется общими циркуляционными процессами атмосферы [2], предполагалось подчинение среднемесячных температур воздуха и месячных осадков одним и тем же масштабным климатообразующим процессам. Среднемноголетние значения выбранных факторов в целом по рассматриваемой территории приведены в табл. 2.

Исходя из принятых требований, наблюдаемые значения температур и осадков за отдельные месяцы каждого года пересчитывались в проценты относительно их среднемноголетней величины за определенный месяц [1]. Окончательный выбор последнего определялся наименьшим среднеквадратичным (стандартным) отклонением, характеризующим разброс полученных значений в целом по всей совокупности данных. После этого выполнялся расчет временной динамики факторов (выраженной в процентах) по отдельным метеостанциям. Затем проводилось усреднение величин по всем 11 станциям для каждого месяца каждого года. Тем самым определялась многолетняя динамика температур воздуха и осадков, отражающая метеорологическую ситуацию одновременно по всей Алтае-Саянской горной стране за период наблюдений. С помощью программного обеспечения Microsoft Office Excel 2003 рассчитывались многолетние тренды обоих факторов по каждому месяцу года для осуществления по этим трендам многолетнего прогноза до 2030 г.

Табл. 1. Географическое расположение опорных метеостанций Алтае-Саянской горной страны.

| № | Метеостанция | Индекс ВМО | Широта | Долгота | Высота над у.м., м |
|---|---|---|---|---|---|
| 1 | Бийск-Зональная[1] | 29939 | 52° 41' | 84° 56' | 222 |
| 2 | Змеиногорск[2] | 36038 | 51° 09' | 82° 10' | 354 |
| 3 | Камень-на-Оби[1] | 29822 | 53° 49' | 81° 16' | 127 |
| 4 | Кара-Тюрек[2] | 36442 | 50° 02' | 86° 27' | 2601 |
| 5 | Кузедеево[3] | 29849 | 53° 20' | 87° 11' | 293 |
| 6 | Кызыл-Озек[2] | 36055 | 51° 54' | 86° 00' | 324 |
| 7 | Ребриха[1] | 29923 | 53° 05' | 82° 20' | 218 |
| 8 | Славгород[1] | 29915 | 52° 58' | 78° 39' | 125 |
| 9 | Солонешное[2] | 36045 | 51° 38' | 84° 20' | 409 |
| 10 | Усть-Кокса[2] | 36229 | 50° 16' | 85° 37' | 977 |
| 11 | Яйлю[2] | 36064 | 51° 46' | 87° 36' | 482 |

[1] – прилегающие к Горному Алтаю равнины; [2] – Горный Алтай; [3] – Кузнецкая межгорная котловина.

Табл. 2. Среднемноголетние значения температур воздуха и осадков для исследуемой территории.

| Климатическая характеристика | Месяцы | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Температуры, °С | –16,0 | –14,6 | –8,0 | 2,2 | 10,5 | 15,7 | 17,8 | 15,3 | 9,6 | 2,1 | –7,3 | –13,3 |
| Осадки, мм | 21,9 | 19,5 | 22,3 | 40,7 | 58,3 | 71,1 | 81,9 | 73,0 | 50,9 | 49,4 | 39,5 | 30,6 |

**Результаты и обсуждение**. В соответствии со среднемноголетними значениями среднемесячных температур воздуха (табл. 2) можно выделить холодный и теплый периоды года. С учетом характера температурных инверсий зимой и летнего циркуляционного механизма атмосферы, определяющих зимний и летний «эффекты высоты», к холодному периоду нами отнесены 10-12, 1-4, а к теплому – 5-9 месяцы года [1]. Выделение таких периодов обеспечило наилучшую точность расчетов с возможностью прогноза помесячной и межгодовой динамики относительных значений выбранных факторов и лучше всего отразило их пространственное единообразие по территории горной страны за счет циркуляционных атмосферных процессов. В качестве реперных месяцев, дающих наилучшую «метеорологическую» корреляцию со всеми остальными месяцами для холодного и теплого периодов года, взяты январь и июль соответственно. Эти месяцы обеспечили наименьшую дисперсию относительных значений температур воздуха по всем метеостанциям.

Как установлено ранее, на территории России сформирован вековой климатический цикл, имеющий три фазы (1918-1950, 1951-1983, 1984-2016 гг.) с определенными статистическими закономерностями многолетних изменений температур и осадков [3]. Под влиянием антропогенной деятельности мезо- и макромасштабные процессы влаго- и теплопереноса в атмосфере становятся все менее устойчивыми. Поэтому при прогнозе нам целесообразно основываться на третьей 33-летней фазе цикла, наиболее близкой к прогнозируемому периоду 2019-2030 гг. На рисунке представлены тренды температур воздуха и осадков, полученные для данной фазы в только закончившемся вековом климатическом цикле.

Традиционное описание многолетних изменений температур воздуха и осадков через их увеличение или уменьшение на некоторую величину в условиях гор является весьма приближенным. Нами дано более адекватное описание климатических трендов, когда изменения среднемесячных температур воздуха для холодного и теплого периодов года, а также осадков, характеризуются в процентах от их среднемноголетних месячных величин in situ [4-7]. Именно такие тренды относительных изменений факторов приведены на рисунке, и мы можем использовать их для долгосрочного прогноза. Наибольшие изменения температур будут наблюдаться в январе (похолодание на ~ 20% к 2030 г.) и марте-апреле (потепление >20%). Осадки заметно изменятся в большинстве месяцев года, хотя и разнонаправлено.




Тренды относительных температур воздуха (T) и осадков (P) по 12 месяцам 1984-2016 гг.

Метод пространственного обобщения метеорологических факторов позволяет решить и вторую часть задачи – возможность осуществления их реанализа для гор. Последний легко выполняется для среднемесячных температур и месячных осадков через их расчет в процентах от среднемесячных значений in situ. Для перехода к общепринятым единицам их измерения (°C, мм) достаточно знать среднемноголетние январские и июльские значения температур и осадков для характеризуемого участка. Данным методом также восстанавливаются пропуски в многолетних рядах метеорологических наблюдений, причем с лучшей точностью, чем при замене недостающих данных соответствующими среднемноголетними значениями [1].

**Заключение.** Предложенный метод пространственного обобщения климатических факторов пригоден не только для равнинных [4-7], но и горных территорий. С его помощью выполнен прогноз относительных изменений температур приземного слоя воздуха и осадков Алтае-Саянской горной страны до 2030 г. Особенностью метода является их расчет в процентах от среднемноголетних величин за реперные месяцы года (январь и июль). Определяемая таким способом многолетняя динамика климатических факторов единообразна по всей анализируемой территории, независимо от ее орографической и климатической неоднородности.

*Работа выполнена при финансовой поддержке РФФИ (грант № 18-45-220019-р_а).*

## ЛИТЕРАТУРА

[1] *Кирста Ю.Б.* Пространственное обобщение климатических характеристик для горных территорий // Мир науки, культуры, образования. 2011. № 3 (28). С. 330-337.

[2] *Модина Т.Д., Сухова М.Г.* Климат и агроклиматические ресурсы Алтая. Новосибирск: Универсальное книжное издательство, 2007. 178 с.
[3] *Кирста Ю.Б., Кирста Б.Ю.* Информационно-физический закон построения эволюционных систем. Системно-аналитическое моделирование экосистем. Барнаул: Изд-во АГУ, 2014. 283 с.
[4] *Кирста Ю.Б., Зинченко Г.С., Ловцкая О.В., Белоторонина Т.А.* Оценка и прогноз агроклиматического потенциала Алтайского края в условиях современных изменений климата // Известия Алтайского гос. ун-та. 2002. № 3. С. 61-68.
[5] *Kirsta Yu.B.* System-analytical modelling – Part II: Wheat biotime run and yield formation. Agroclimatic potential, Le Chatelier principle, changes in agroclimatic potential and climate in Russia and the U.S. // Ecol. Modelling. 2006. Vol. 191. P. 331-345.
[6] *Кирста Ю.Б., Ловцкая О.В.* Прогноз климатических изменений в зернопроизводящих зонах Сибири и России // Мир науки, культуры, образования. 2009. № 7 (19). С. 9-13.
[7] *Кирста Ю.Б., Курепина Н.Ю., Ловцкая О.В.* Прогноз климата и агроклиматического потенциала Алтайского края до 2020 года // Вестник Алтайского гос. аграрного ун-та. 2013. №1. С. 27-32.

# МЕТОДЫ ОБРАБОТКИ ПРОСТРАНСТВЕННЫХ ДАННЫХ В ЗАДАЧАХ МОНИТОРИНГА ПРИРОДНЫХ ЯВЛЕНИЙ, ОСНОВАННЫЕ НА БАЙЕСОВСКОМ ПОДХОДЕ

*Климова Е.Г.*
Институт вычислительных технологий СО РАН, Новосибирск

Одной из важных задач мониторинга природных явления является задача получения значений параметров окружающей среды на некоторой регулярной сетке. В настоящее время принято решать такие задачи с привлечением всех имеющихся данных наблюдений, а также математической модели интересующего нас процесса. Математическая постановка задачи входит в набор задач так называемого обратного моделирования. В случае, если рассматривается вероятностная постановка задачи, применяется байесовский подход. На таком подходе основаны популярные алгоритмы, такие как ансамблевый фильтр Калмана, ансамблевое сглаживание Калмана, метод частиц. В докладе приводится краткий обзор современных методов, а также подходы к их практической реализации.
*Ключевые слова: усвоение данных, ансамблевый фильтр Калмана, спутниковые данные.*

**Введение.** Одной из важных задач мониторинга природных явления является задача получения значений параметров окружающей среды на некоторой регулярной сетке. В настоящее время принято решать такие задачи с привлечением всех имеющихся данных наблюдений, а также математической модели интересующего нас процесса. Математическая постановка задачи входит в набор задач так называемого обратного моделирования. Решение задачи обратного моделирования для заданной модели процесса и набора данных наблюдений циклически по времени является задачей усвоения данных. В проблему обратного моделирования включена также задача оценки параметров модели.

В случае если рассматривается вероятностная постановка задачи, применяется байесовский подход. На таком подходе основаны популярные алгоритмы, такие как ансамблевый фильтр Калмана, ансамблевое сглаживание Калмана, метод частиц. В случае, если рассматриваемые случайные величины являются гауссовскими, а модели прогноза и наблюдений линейными, такая постановка задачи эквивалентна вариационной постановке задачи усвоения данных (4DVAR) [1].

В докладе приводится краткий обзор современных методов усвоения данных, используемых при моделировании процессов в окружающей среде, а также подходы к их практической реализации.

**Задача обратного моделирования.** Задачей обратного моделирования называют оценку неизвестного вектора параметров большой размерности $\mathbf{x} \in R^M$ по вектору данных наблюдений $\mathbf{y} \in R^P$ [2]. Предположим, что

$$\mathbf{y} = M(\mathbf{x}) + \boldsymbol{\delta},$$

где $M$ – известный оператор прогноза – наблюдений; $\boldsymbol{\delta}$ – случайный шум с заданной функцией распределения. В случае, если рассматривается временной ряд данных наблюдений, изменение по времени оцениваемой переменной описывается с помощью математической модели. Под задачей усвоения данных принято понимать последовательное по времени оценивание неизвестной величины по данным наблюдений [2].

**Байесовский поход к задаче усвоения данных.** Предположим, что изменение по времени оцениваемой величины $\mathbf{x}^k$ описывается моделью

$$\mathbf{x}^{k+1} = f_{k+1,k}(\mathbf{x}^k) + \boldsymbol{\eta}^k,$$

где $k$ – номер шага по времени. Кроме того, известны данные наблюдений $\mathbf{y}^k$:

$$\mathbf{y}^k = h_k(\mathbf{x}^k) + \boldsymbol{\varepsilon}^k.$$

В этих формулах $\boldsymbol{\eta}^k$, $\boldsymbol{\varepsilon}^k$ – случайные ошибки прогноза и наблюдений, соответственно; $f_{k+1,k}$ – оператор модели; $h_k$ – оператор трансформации прогнозируемой переменной в наблюдаемую.

Байесовский подход состоит в применении теоремы Байеса для получения оптимальной оценки по данным наблюдений и прогнозу:

$$p(\mathbf{x}\,|\,\mathbf{y}) = \frac{p(\mathbf{y}\,|\,\mathbf{x})\,p(\mathbf{x})}{p(\mathbf{y})}\,.$$

Известны различные варианты оценки состояния по данным и прогнозу: $p(\mathbf{x}_l\,|\,\mathbf{y}_{k,1}), k > l$ – прогноз; $p(\mathbf{x}_k\,|\,\mathbf{y}_{k,1})$ – фильтрация; $p(\mathbf{x}_{k,0}\,|\,\mathbf{y}_{k,1})$ – сглаживание, где $\mathbf{x}_{k,0} = \{\mathbf{x}_k, \mathbf{x}_{k-1}, \cdots, \mathbf{x}_0\}$, $\mathbf{y}_{k,1} = \{\mathbf{y}_k, \cdots, \mathbf{y}_1\}$. Обозначения и формулировки взяты из обзора [1].

В линейном гауссовском случае решением задачи фильтрации является фильтр Калмана, решением задачи сглаживания – сглаживание Калмана. В работе [3] предложено использовать метод Монте-Карло для решения задач фильтрации и сглаживания, так называемые ансамблевые алгоритмы фильтрации и сглаживания. В ансамблевом фильтре Калмана в нелинейном случае прогноз ковариационной матрицы осуществляется с помощью нелинейной модели, при этом нарушается условие гауссовости. Также в этом случае оценка на шаге анализа является приближением оценки минимальной дисперсии (linear variance minimizing). Для решения задачи сглаживания в нелинейном случае в настоящее время рассматриваются итерационные алгоритмы сглаживания с применением ансамблей [4]. Как отмечается в [4], итерационные методы дают хорошее приближение для слабо нелинейных моделей. В нелинейном негауссовском случае используется метод частиц, который также основан на байесовском подходе.

**Ансамблевый фильтр Калмана.** Вывод формул ансамблевого фильтра Калмана впервые приведен в работах Эвенсена (Evensen G.) [3]. Запишем нелинейную динамическую систему в виде уравнения процесса

$$\mathbf{x}_k^t = f(\mathbf{x}_{k-1}^t) + \boldsymbol{\eta}_{k-1}^t$$

и уравнения наблюдений

$$\mathbf{y}_k = h(\mathbf{x}_k^t) + \boldsymbol{\varepsilon}_k^t\,,$$

где $h$ – оператор, вообще говоря, нелинейный, переводящий значения прогноза в наблюдаемую переменную; $\boldsymbol{\eta}_{k-1}^t$ – вектор «шумов модели»; $\boldsymbol{\varepsilon}_k^t$ – вектор ошибок наблюдений; $\mathbf{x}_k^t$ – вектор оцениваемых переменных в момент времени $t_k$; $\boldsymbol{\varepsilon}_k^t$ и $\boldsymbol{\eta}_{k-1}^t$ – Гауссовские случайные переменные: $E[\boldsymbol{\varepsilon}_k^t \left(\boldsymbol{\varepsilon}_k^t\right)^T] = \mathbf{R}_k^t$, $E[\boldsymbol{\eta}_{k-1}^t \left(\boldsymbol{\eta}_{k-1}^t\right)^T] = \mathbf{Q}_{k-1}^t$. Будем считать $\mathbf{x}_k^t$ «истинным» значением.

Стохастический ансамблевый фильтр Калмана состоит из ансамбля прогнозов $\{\mathbf{x}_k^{f,n}, n = 1, \cdots, N\}$

$$\mathbf{x}_k^{f,n} = f(\mathbf{x}_{k-1}^{a,n}) + \boldsymbol{\eta}_{k-1}^n \tag{1}$$

и ансамбля анализов $\{\mathbf{x}_k^{a,n}, n = 1, \cdots, N\}$

$$\mathbf{x}_k^{a,n} = \mathbf{x}_k^{f,n} + \mathbf{K}_k(\mathbf{y}_k^n + \boldsymbol{\varepsilon}_k^n - h(\mathbf{x}_k^{f,n}))\,. \tag{2}$$

Ансамбли (1) и (2) задают выборку значений «истины», при этом среднее по выборке значение будет являться оптимальной оценкой, а отклонение от среднего – ансамблем ошибок анализа и прогноза, соответственно. Для осуществления ансамблевого варианта алгоритма фильтра Калмана требуется задание ансамбля ошибок наблюдений $\{\boldsymbol{\varepsilon}_k^n, n = 1, \cdots, N\}$ а также ансамбля ошибок прогноза $\{\mathbf{dx}_k^{f,n} = \mathbf{x}_k^{f,n} - \overline{\mathbf{x}_k^{f,n}}, n = 1, \cdots, N\}$, где $\overline{\mathbf{x}_k^{f,n}} \cong \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}_k^{f,n}$ и ансамбля шумов модели $\{\boldsymbol{\eta}_{k-1}^n, n = 1, \cdots, N\}$: $E[\boldsymbol{\eta}_{k-1}^n \left(\boldsymbol{\eta}_{k-1}^n\right)^T] = \mathbf{Q}_k$. Матрица $\mathbf{K}_k$ имеет вид

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k{}^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k{}^T + \mathbf{R}_k)^{-1}\,,$$

где $\mathbf{P}_k^f$ и $\mathbf{R}_k$ – матрицы, оцениваемые по ансамблю

$$\mathbf{P}_k^f \triangleq \frac{1}{N-1}\sum_{n=1}^{N}\mathbf{dx}_k^{f,n}\left(\mathbf{dx}_k^{f,n}\right)^T,\ \mathbf{R}_k \triangleq \frac{1}{N-1}\sum_{n=1}^{N}\boldsymbol{\varepsilon}_k^n\left(\boldsymbol{\varepsilon}_k^n\right)^T,$$

$\mathbf{H}_k$ – линеаризованный оператор $h(\mathbf{x}_k^{f,n})$ относительно $\overline{\mathbf{x}_k^{f,n}}$:

$$h(\mathbf{x}_k) \cong h(\overline{\mathbf{x}_k^{f,n}}) + \mathbf{H}_k\boldsymbol{\varepsilon}_k^f \ .$$

Формулы (1)-(2) представляют собой стохастический ансамблевый фильтр Калмана.

Детерминированный ансамблевый фильтр Калмана (шаг анализа) состоит из уравнения для среднего значения

$$\overline{\mathbf{x}_k^{a,n}} = \overline{\mathbf{x}_k^{f,n}} + \mathbf{K}_k(\mathbf{y}_k^n - \overline{h(\mathbf{x}_k^{f,n})})$$

и оценки ансамбля ошибок анализа так, чтобы соответствующая ковариационная матрица удовлетворяла уравнению фильтра Калмана $\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^f$ [5].

**Задача оптимального сглаживания. Ансамблевое сглаживание (EnKS – ensemble Kaman smoother).** В классической теории оптимального оценивания наряду с задачей оптимальной фильтрации – получения оптимальной оценки в конце рассматриваемого временного интервала существует также задача оптимального сглаживания, представляющая собой задачу получения оптимальной оценки на всем рассматриваемом временном интервале. Под оптимальной оценкой понимается решение, минимизирующее след матрицы ковариаций ошибок оценивания [6].

Для задачи оптимального сглаживания также применяются ансамблевые алгоритмы. В работе [3] получен важный результат, состоящий в том, что задача ансамблевого сглаживания может быть решена последовательно, используя на каждом шаге по времени имеющиеся данные для оценки значений на всем временном интервале. При этом формулы оптимальной оценки аналогичны формулам ансамблевого фильтра Калмана. При таком варианте ансамблевого сглаживания (EnKS) не требуется проводить вычисления в обратном времени. Следует отметить, что при использовании соответствующих весовых матриц алгоритм EnKS эквивалентен алгоритму вариационного усвоения данных 4DVAR [1, 2].

**Оценка параметров в процедуре усвоения данных.** Рассмотрим уравнение процесса в виде

$$\mathbf{x}_k^t = f(\mathbf{x}_{k-1}^t, \boldsymbol{\alpha}_{k-1}^t) + \boldsymbol{\eta}_{k-1}^t,$$

данные наблюдений

$$\mathbf{y}_k = h(\mathbf{x}_k^t, \boldsymbol{\alpha}_k^t) + \boldsymbol{\varepsilon}_k^t,$$

где $\boldsymbol{\alpha}_k^t$ – вектор параметров. Будем считать, что параметр не меняется со временем: $\boldsymbol{\alpha}_{k+1}^t = \boldsymbol{\alpha}_k^t$. Рассмотрим обобщенную задачу оценивания вектора $\mathbf{z} = [\mathbf{x}, \boldsymbol{\alpha}]^T$. Опуская промежуточные выкладки, запишем сразу результат процедуры оценивания в общем виде:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{P}_{xx}h_x^T(h_x\mathbf{P}_{xx}h_x^T + \mathbf{R})^{-1}[\mathbf{y} - h(\mathbf{x}^f, \boldsymbol{\alpha}^f)] + \mathbf{P}_{x\alpha}h_\alpha^T(h_x\mathbf{P}_{xx}h_x^T + \mathbf{R})^{-1}[\mathbf{y} - h(\mathbf{x}^f, \boldsymbol{\alpha}^f)],$$
$$\boldsymbol{\alpha}^a = \boldsymbol{\alpha}^f + \mathbf{P}_{\alpha x}h_x^T(h_x\mathbf{P}_{xx}h_x^T + \mathbf{R})^{-1}[\mathbf{y} - h(\mathbf{x}^f, \boldsymbol{\alpha}^f)] + \mathbf{P}_{\alpha\alpha}h_\alpha^T(h_x\mathbf{P}_{xx}h_x^T + \mathbf{R})^{-1}[\mathbf{y} - h(\mathbf{x}^f, \boldsymbol{\alpha}^f)].$$

В этих формулах индекс «k» опущен. Рассматривается шаг анализа, индекс «a» означает анализ, индекс «f» – прогноз; $\mathbf{P}_{\alpha x}$ – кросс-ковариации ошибок $\mathbf{x}$ и $\boldsymbol{\alpha}$; $\mathbf{P}_{\alpha\alpha}$ – ковариационная матрица ошибок $\boldsymbol{\alpha}$. $h_x$ и $h_\alpha$ – линеаризованные операторы по $\mathbf{x}$ и $\boldsymbol{\alpha}$, соответственно. В случае, если $h$ от $\boldsymbol{\alpha}$ не зависит, оценка $\mathbf{x}^a$ проводится по той же формуле, что и в обычном фильтре Калмана [6]. В современных работах по усвоению данных такой подход применяется к оценке потоков парниковых газов [7].

**Вариационный подход к задаче усвоения данных**. Вариационный подход к задаче усвоения данных очень популярен и используется в ведущих прогностических центрах мира в оперативных системах усвоения данных [1, 2]. В алгоритме анализа данных, называемом в литературе 3DVAR (3-Dimensional VARiational) ищется значение вектора $\mathbf{x_a}$, доставляющее минимум функционалу

$$2\mathrm{J}(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_\mathrm{b})^\mathrm{T}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_\mathrm{b}) + [\mathbf{y}_0 - \mathrm{h}(\mathbf{x})]^\mathrm{T}\mathbf{R}^{-1}[\mathbf{y}_0 - \mathrm{h}(\mathbf{x})].$$

В этой формуле **B** и **R** – ковариационные матрицы ошибок прогноза $\mathbf{x}_b$ и наблюдений $\mathbf{y}_0$, соответственно. Оператор h, вообще говоря, нелинейный, переводит прогностические значения в точки наблюдений (и наблюдаемые переменные).

Минимум функционала есть решение уравнения

$$\nabla J(\mathbf{x}_a) = 0.$$

Если градиент функционала записывается в следующем виде:

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x}_b) - \mathbf{H}^T\mathbf{R}^{-1}\{\mathbf{y}_0 - h(\mathbf{x}_b)\},$$

тогда решение задачи имеет вид

$$\mathbf{x}_a = \mathbf{x}_b + (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{P}^{-1}\mathbf{H})\{\mathbf{y}_0 - h(\mathbf{x}_b)\}.$$

где **H** – линеаризованный оператор относительно $\mathbf{x}_b$. Эта формула эквивалентна шагу анализа алгоритма фильтра Калмана [1, 2, 6].

Алгоритм 4DVAR представляет собой обобщение 3DVAR на пространственно-временной случай. 4DVAR позволяет использовать наблюдения из некоторого временного интервала $(t_n - t_0)$. Считается, что на этом интервале заданы наблюдения $\{\mathbf{y}_i^0, i = 0, \cdots, N\}$. Ищется минимум функционала

$$J[\mathbf{x}(t_0)] = \frac{1}{2}[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)]^T\mathbf{B}_0^{-1}[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)] + \frac{1}{2}\sum_{i=0}^{N}[h(\mathbf{x}_i) - \mathbf{y}_i^0]^T\mathbf{R}^{-1}[h(\mathbf{x}_i) - \mathbf{y}_i^0]$$

при условии, что

$$\mathbf{x}(t_n) = M_n[\mathbf{x}(t_0)].$$

В данной оптимизационной задаче контрольной переменной является $\mathbf{x}(t_0)$ – значение в начальный момент времени. То есть, в методе 4DVAR ищется начальное условие такое, что прогноз с этим начальным значением наилучшим образом приближает наблюдения в заданном временном интервале. Минимум функционала в 4DVAR ищется с помощью итерационных методов (например, метод квази-Ньютона). Для реализации итерационного процесса требуется оценить градиент функционала J(**x**), который, в свою очередь, вычисляется с помощью сопряженной линеаризованной модели. В настоящее время существуют схемы, основанные на 4DVAR с использованием ансамблевого подхода для оценки изменений ковариаций по времени [1, 2].

**Проблемы реализации ансамблевых алгоритмов.** Задачи усвоения данных, оценки параметров при моделировании окружающей среды являются чрезвычайно трудоемкими, требующими больших затрат машинных ресурсов. Применение ансамблевых алгоритмов частично позволяет решить эту проблему, однако все равно задача остается крайне трудоемкой. Для проведения шага анализа применяются алгоритмы трансформации ансамбля прогнозов для получения ансамбля анализов. В этом случае анализ осуществляется только для среднего по ансамблю значения, а затем вычисляется ансамбль анализов. Существуют детерминированный и стохастический варианты таких алгоритмов для реализации как ансамблевого фильтра Калмана, так и ансамблевого сглаживания [5].

Одним из таких алгоритмов является ансамблевый π-алгоритм [8-10], который представляет собой стохастический фильтр Калмана. В этом алгоритме производятся операции с матрицами порядка размерности ансамбля. Кроме того, этот алгоритм может быть использован при реализации ансамблевого сглаживания (EnKS). Важным свойством алгоритма является локальность. Алгоритм оценки искомого значения может быть реализован независимо для заданных подобластей. Это свойство может быть использовано при оценке параметров локально, в заданных регионах.

При использовании ансамблей в задачах усвоения данных возникает ряд проблем, связанных с небольшим размером выборки (ансамбля). В частности, это проблема высоких зна-

чений корреляций ошибок прогноза на больших расстояниях. В этом случае принять применять так называемую локализацию, а именно, поэлементное умножение ковариационных матриц на функцию, убывающую с расстоянием [1, 5].

Малый размер выборки также является одной из причин расходимости алгоритма. В линейных алгоритмах ансамблевого фильтра Калмана (EnKF), ансамблевого сглаживания Калмана (EnKS) используется «увеличивающий множитель» (inflation factor), при этом либо элементы ковариационной матрицы умножаются на некоторый множитель (multiplicative inflation), либо к ансамблю возмущений добавляется дополнительное возмущение (additive inflation). Следует отметить, что эти приемы используются и в ряде итерационных алгоритмов сглаживания с применением ансамблей, а также в методе частиц с применением ансамблей [1, 2, 5].

Особо выделим вопросы, связанные с усвоением спутниковой информации [11].

Спутниковые данные имеют ряд отличительных особенностей по сравнению с наземными и аэрологическими наблюдениями. К этим особенностям относятся следующие.

а. Спутник измеряет информацию об уходящем излучении, поэтому необходимо осуществлять преобразование переменных модели в наблюдаемые переменные с помощью уравнения переноса излучения. Таким образом, оператор наблюдений h нелинейный.
б. Данные поступают непрерывно по времени.
в. Ошибки наблюдений коррелируют, матрица ошибок наблюдений большой размерности и недиагональная.
г. Наличие систематической ошибки в наблюдениях (bias). Это усложняет использование формул вышеприведенных алгоритмов.

Спутниковые данные могут использоваться двумя путями: напрямую, в виде данных об уходящем излучении, либо требуется предварительно восстанавливать по спутниковым данным значения метеорологических величин, а затем использовать их в процедуре усвоения данных. Первый вариант является более предпочтительным по целому ряду причин.

а. При преобразовании данных о радиации используется информация о прогнозе, поэтому ошибки восстановленных данных и прогноза будут коррелировать.
б. Проблема оценки матрицы ковариаций ошибок наблюдений.

Спутниковые данные – это огромный массив информации. Однако, как отмечается в ряде работ, большое количество наблюдений не всегда дает наилучший результат, поскольку эти данные представляют собой массив коррелирующих между собой случайных величин и потому они мало информативны [11].

**Задача усвоения данных об окружающей среде.** Для осуществления мониторинга состояния окружающей среды в настоящее время проводится моделирование распространения в атмосфере, в частности: пассивных газовых составляющих (парниковые газы), химически активных веществ, аэрозолей.

Информация о содержании всех этих веществ измеряется с помощью как наземных измерительных приборов, так и спутников. Большое количество работ в настоящее время посвящено проблеме создания систем усвоения данных об окружающей среде, при этом используется базовый математический аппарат, наработанный в задачах усвоения метеорологических переменных [7].

Важной задачей в мониторинге состояния окружающей среды является оценка потоков газовых составляющих с поверхности Земли на основе полей, полученных с помощью системы усвоения данных [7].

В задачах моделирования состояния окружающей среды в настоящее время используются модели атмосферы с «химическим» блоком. В качестве примера можно привести популярную модель WRF-Chem [12]. Для проведения численных экспериментов используются современные базы данных метеорологических полей (ERA-interim re-analysis [13]). В ECMWF создана база данных MACC (Monitoring Atmospheric Composition and Climate), включающая в себя активные газовые составляющие, аэрозоли и парниковые газы [14]. При региональном

моделировании важной проблемой является задание значений на границах. С этой целью используются как данные реанализа, так и рассматриваются способы усвоения данных наблюдений для уточнения значений на границах области [7].

**Заключение.** В докладе представлен краткий обзор современных подходов к задаче усвоения данных наблюдений об окружающей среде. Рассматриваются основные проблемы практической реализации этой задачи.

## ЛИТЕРАТУРА


[1] *Carrassi A., Bocquet M., Bertino L., Evensen G.* Data assimilation in the geosciences: An overview of methods, issuers and perspectives // Wiley interdisciplinary reviews: Climate Change. 2018. Vol. 131, is 5. e535. DOI: 10.1002/wcc535.

[2] *Nakamura G., Potthast R.* Inverse Modeling. IOP Publishing, 2015. 509 p. DOI: 10.1088/978-0-7503-1218-9.

[3] *Evensen G.* Data assimilation. The ensemble Kalman filter. Berlin, Heideberg: Spriger-Verlag, 2009. 307 p.

[4] *Evensen G.* Analysis of iterative ensemble smoother for solving inverse problems // Computational Geosciences. 2018. Vol. 22. P. 885-908. DOI: 10.1007/s10596-018-9731-y.

[5] *Houtekamer H.L., Zhang F.* Review of the ensemble Kalman filter for atmospheric data assimilation // Monthly Weather Review. 2016. Vol. 144. P. 4489-4532.

[6] *Jazwinski A.H.* Stochastic processes and filtering theory. N.Y.: Academic Press, 1970. 376 p.

[7] *Bocquet M. et al.* Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models // Atmospheric Chemistry and Physics. 2015. Vol. 15. P. 5325-5358.

[8] *Klimova E.* A suboptimal data assimilation algorithm based on the ensemble Kalman filter // Quarterly Journal of the Royal Meteorological Society. 2012. Vol. 138. P. 2079-2085.

[9] *Klimova E.G.* Application of ensemble Kalman filter in environment data assimilation // IOP Conference Series: Earth and Environmental Science. 2018. Vol. 211. P. 012049. DOI: 10.1088/1755-1315/211/1/012049.

[10] *Климова Е.Г.* Стохастический ансамблевый фильтр Калмана с трансформацией ансамбля возмущений // Сибирский журнал вычислительной математики. 2019. Т. 22, № 1. С. 27-40.

[11] *Thepaut J.-N.* Satellite data assimilation in numerical weather prediction: an overview // Proceedings of the ECMWF Seminar on Recent development in data assimilation for atmosphere and ocean. 8-12 September 2003. P.75-96.

[12] *Grell G. et al.* Fully coupled «online» chemistry within the WRF model // Atmos. Environment. 2005. Vol. 39. P. 6957-6975.

[13] *Dee D.P. et al.* The ERA-interim reanalysis: Configuration and performance of the data assimilation system // Q. J. Roy. Meteorol. Soc. 2011. Vol. 137. P. 553-579.

[14] *Inness A. et al.* The MACC reanalysis: an 8 yr data set of atmospheric composition // Atmos. Cgem. Phys. 2013. Vol. 13. P. 4073-4109.

# КАРТОГРАФИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТЕМПЕРАТУРНЫХ ПОЛЕЙ ПОЧВ СРЕДНЕСИБИРСКОГО ТРАНСЕКТА НА ОСНОВЕ СОВМЕСТНОГО АНАЛИЗА НАЗЕМНЫХ И СПУТНИКОВЫХ ДАННЫХ

*Кудряшова С.Я.[(1)], Чумбаев А.С.[(1)], Пестунов И.А.[(2)], Синявский Ю.Н.[(2)], Чубаров Д.Л.[(2)], Безбородова А.Н.[(1)], Ермаков Н.Б.[(3)]*

[(1)] Институт почвоведения и агрохимии СО РАН, Новосибирск

[(2)] Институт вычислительных технологий СО РАН, Новосибирск

[(3)] Центральный сибирский ботанический сад СО РАН, Новосибирск

На основе совместного анализа количественных показателей температурного режима воздуха и почв, полученных с использованием наземных автоматизированных измерительных комплексов и спутниковых данных, создана серия картографических моделей, отражающих особенности температурных полей типологических единиц почвенно-биоклиматической зональности Алтае-Саянского региона.

*Ключевые слова: картографическое моделирование, температурное поле, наземный автоматизированный мониторинг, спутниковые данные.*

**Введение.** Актуальность развития геоинформационной концепции для картографирования почвенного покрова, свойств и режимов почв представлена в большом количестве приложений, разработанных с использованием геоинформационных систем и технологий. Температурный режим является одним из наиболее значимых экологических факторов, который совместно с гидрологическим режимом характеризует общий энергетический уровень формирования и функционирования почвенного покрова. Температурное поле почв представляет совокупность значений температуры в точках пространственной области, которое согласно результатам исследований, полученных как в нашей стране, так и за рубежом рассматривается в качестве ведущего фактора структурно-функциональной организации почвенного покрова. Для целей тематического картографирования используются свойства почв, которые тесно коррелируют с экологическими характеристиками окружающей среды и обусловлены действием факторов почвообразования [1]. В наиболее полном виде факторы, определяющие свойства почв в определенной точке исследуемого пространства, отражены в модели SCORPAN - почвенной пространственно предсказывающей функции, которая предполагает, что одинаковому сочетанию почвообразующих факторов – предикторов соответствуют почвы сходного генезиса, а границы почвенных структур обусловлены изменением факторов дифференциации почвенного покрова. Выделение наиболее информативных предикторов осуществляется с использованием подходов педотрансферного метода, которые позволяют рассчитывать педотрансферные функции – уравнения связи, описывающие функциональные соотношения базовых почвенных свойств и характеристик почвенного географического пространства [2]. Основное отличие данной методологии состоит в том, что по совокупности выбранных количественных почвенно-экологических показателей можно перейти к ее таксономической характеристике. Такой подход позволяет по совокупности однородных почвообразующих факторов-предикторов пространственно разделять ареалы почв и проводить между ними контурные границы. Картографические модели температурных полей, полученные в результате совместного анализа наземных и спутниковых данных, обладают достаточной информативностью в отношении энергетики почвообразовательных процессов и могут быть использованы для оценки термических условий почв малоисследованных и труднодоступных территорий Алтае-Саянского региона.

**Наземный автоматизированный мониторинг воздуха и почв** был организован с использованием специализированного регистратора температуры DS-1921G «Thermochron» с учетом показателей, отражающих генетическое единство типов климата Алтае-Саянского региона. Для фиксации температуры воздуха был установлен автономный регистратор на высоте

2 м от поверхности почвы в условиях, исключающих прямое радиационное воздействие. Динамика изменения температур на поверхности почвы и по горизонтам почвенного профиля фиксировалась в течение года с интервалом 4 часа. В результате выполненных наблюдений был получен большой объем фактических данных, который использовался в качестве основы для расчета термических ресурсов температурных полей и выявления их временных трендов [3].

**Для проведения ретроспективного анализа, статистической обработки и визуализации данных MODIS** использовался созданный в ИВТ СО РАН оригинальный программно-алгоритмический инструментарий, базирующийся на новой технологии доступа к архиву спутниковых данных, реализованной с использованием СУБД PostgreSQL с дополнительным модулем [4-7]. Этот модуль предназначен для прямого доступа к файловому архиву данных без необходимости предварительного копирования и преобразования формата данных для СУБД. Он реализует прозрачное отображение файлового архива спутниковых снимков в виртуальные таблицы базы данных. Модуль позволяет выполнять произвольные SQL-запросы к данным файлового архива, при этом планировщик запросов оптимизирует их выполнение на основе доступных метаданных, а применяемые для выполнения вычислений алгоритмы предусматривают работу с объемами информации превосходящими емкость оперативной памяти сервера СУБД. Разработанную систему извлечения данных из архива спутниковых снимков можно сопоставить с такими системами как NASA Giovanni, Google Earth Engine и европейским проектом TELEIOS. Все они нацелены на предоставление доступа к большим объемам спутниковых снимков. Созданная технология в различных аспектах превосходит описанные системы. В ней нет ограничений на вид запросов к данным, так как поддерживаются произвольные SQL-запросы. Эта система рассчитана на работу с произвольными пространственными данными без необходимости их предварительного преобразования и подготовки.

**Отработка методических положений выделения границ температурных полей** проведена на примере типологических единиц почвенного покрова Алтае-Саянского региона, включающего разнообразие степей Чулымо-Енисейского и Минусинского прогибов, Чуйской, Курайской, Турано-Уюкской, Центрально-Тувинской, Убсу-Нурской котловин и тундрово-степных комплексов плоскогорья Укок. Ключевые участки выбраны с учетом принципов ландшафтного районирования, которые позволяют распознавать, классифицировать и картировать ландшафтно-дифференцирующие факторы, компоненты ландшафта, региональную ландшафтную структуру в целом и ее динамические особенности. Метод основан на сопряженном анализе региональных структур, объективно отражающихся на космических снимках и фиксируемых на ландшафтно-типологических картах. В качестве источников информации использовались космические снимки Landsat 8/9. Тематические почвенные и климатические карты использовались как вспомогательные при дешифрировании космических снимков. Межгорные котловины Хакасии и Тувы расположены в восточном секторе Алтае-Саянской горной области в следующей последовательности с севера на юг – Северо-Минусинская, Южно-Минусинская, Турано-Уюкская, Центрально-Тувинская и Убсу-Нурская. Высотные отметки днищ котловин увеличиваются в южном направлении от 300 до 500 м н.у.м. в котловинах Хакасии и от 700 до 900 м н.у.м. в котловинах Тувы. В этом же направлении увеличивается континентальность климата – от наиболее мягких климатических условий в Северо-Минусинской, до наиболее сурового типа термического режима в Убсу-Нурской котловине. Особенности распространения и условия формирования термических условий степей левобережной части Енисея Минусинской котловины (Хакасия) связаны с их расположением между горными сооружениями Кузнецкого Алатау, Восточного и Западного Саяна (рис. 1).

На территории Тувы выделяются четко различающиеся по климатообразующим факторам и термическим условиям почвообразования Турано-Уюкская, Центрально-Тувинская и Убсу-Нурская степные котловины. В границах высокогорного плоскогорья Укок выделяются тундрово-степные комплексы с контрастным сочетанием горно-степных и горно-тундровых почв и количественных показателей почвенного климата (рис. 2).

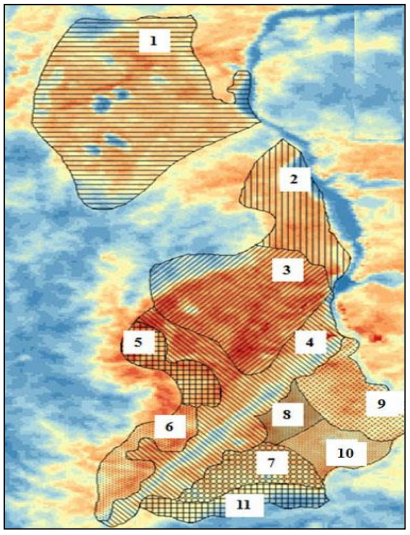



Рис. 1. Степи котловин Хакасии: 1 – Ширинская; 2 – Боградская; 3 – Уйбатская; 4 – Абаканская; 5 – Сахсарская; 6 – Аскизская; 7 – Сабинская; 8 – Сорокаозерная; 9 – Алтайская; 10 – Койбальская; 11 – Иудино-Бейская.

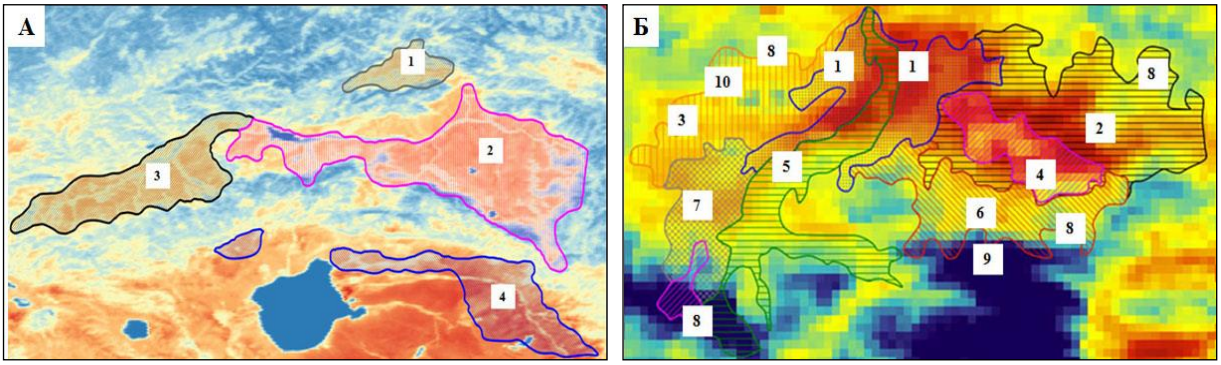



Рис. 2. *А* – степи котловин Тувы: 1 – Турано-Уюкская; 2 – Улуг-Хемская; 3 – Хемчикская; 4 – Убсу-Нурская. *Б* – тундрово-степные комплексы плоскогорья Укок: 1 – тундрово-степная равнина; 2 – степная флювиогляциальная равнина; 3 – тундрово-степная древнеозерная равнина; 4 – заболоченная луговая ледниковая равнина; 5 – лугово-степная долина; 6 – тундровое моренное плато; 7 – тундровая нагорная равнина; 8 – поверхности выравнивания и пологие склоны; 9 – крутые склоны; 10 – озера.

Контрастность климатических режимов воздуха и почв в котловинах Хакасии формируется в зависимости от выраженности направления переноса влаги, которое в восточном секторе Алтае-Саянской горной области имеет западную, сменяющуюся на северо-западную ориентацию (табл. 1).

Климатические условия, как комплекс экологических факторов, оказывают непосредственное влияние на формирование разнообразия почвенного покрова котловинных степей Хакасии. В районах Ужуро-Копьевской и Ширинской степей, в которых средняя годовая температура воздуха составляет 1,3°, продолжительность безморозного периода - 168 дней и $\Sigma t°>10°C=1573°$ преобладают черноземы южные и обыкновенные, занимающие от 14 до 34% площади котловин. В Уйбатской степи, расположенной в непосредственной близости с восточными предгорьями Кузнецкого Алатау средняя годовая температура воздуха составляет -1,1°, продолжительность безморозного периода – 170 дней и $\Sigma t°>10°C=1630°$. Поэтому здесь находится наиболее ксерофильное ядро степной растительности и формируются малоплодородные солонцовые почвы.

В котловинах Тувы четко различающихся по физико-географическим условиям среды, годовой и суточной цикличности климата выделяются температурные поля, которые могут рассматриваться как показатели различий в комплексе климатических условий, структурной организации и функционирования почвенного покрова (табл. 2).

Табл. 1. Показатели температурного режима воздуха и почв (°С) степных котловин Хакасии по глубинам почвенного профиля (см). В знаменателе указано количество дней.

| Показатели температурного режима, °С | Температура воздуха, °С | Температура по глубинам почвенного профиля (см), °С | | | | |
|---|---|---|---|---|---|---|
| | | поверхность почвы | 10 | 20 | 30 | 50 |
| Ужуро-Копьевская луговая степь с. Копьево т.1 (N54°56'19,6"; E89°52'47,1"). | | | | | | |
| **годовая** | 424,7 | 908,2 | 791,8 | 940,3 | 888,3 | 989 |
| **>10°** | 1461,5/89* | 1674,5/100 | 1427,2/90 | 1400,7/90 | 1205,3/77 | 1202,9/83 |
| **>5°** | 1755,3/131 | 1917,3/133 | 1742,8/133 | 1711,8/132 | 1571,8/123 | 1508,5/126 |
| **>0°** | 1845,6/168 | 2026,7/176 | 1827,2/173 | 1819,3/183 | 1706,3/179 | 1631,8/177 |
| **<0°** | -1420,9/165 | -1118,5/157 | -1035,3/160 | -879/150 | -817,9/154 | -642,8/156 |
| **среднегодовая** | 1,3 | 2,7 | 2,4 | 2,8 | 2,7 | 3,0 |
| Ширинская озерно-котловинная степь, оз. Тус, т.2 (N54°45'17,2"; E89°57'17,9"). | | | | | | |
| **годовая** | 430,5 | 1076,5 | 1146,2 | 872,3 | 1121,8 | 915,3 |
| **>10°** | 1558,8/90 | 1875,1/105 | 1707,9/100 | 1428,1/90 | 1359,9/87 | 1075,6/76 |
| **>5°** | 1859,6/130 | 2161/144 | 2020,7/141 | 1760,9/136 | 1721/135 | 1422,8/124 |
| **>0°** | 1974,5/169 | 2244,2/181 | 2127,2/185 | 1846/172 | 1842,5/183 | 1554,3/177 |
| **<0°** | -1544/164 | -1168,2/152 | -980/148 | -973,7/161 | -720,8/150 | -639/156 |
| **среднегодовая** | 1,3 | 3,2 | 3,4 | 2,6 | 3,4 | 2,7 |
| Уйбатская равнинно-холмистая солонцеватая степь, оз. Улуг-Холь, т.4 (N53°47'30,1"; E90°38'39,8"). | | | | | | |
| **годовая** | 371,7 | 322 | 283 | 390,6 | 389,4 | 531,3 |
| **>10°** | 1630,8/95 | 1785,7/105 | 1457,8/94 | 1363,7/90 | 1168,5/80 | 1091/79 |
| **>5°** | 1927,1/135 | 2059,4/142 | 1732,8/130 | 1674,9/132 | 1523/127 | 1407,8/121 |
| **>0°** | 2010,1/170 | 2118,6/164 | 1829,3/170 | 1767,4/169 | 1634,9/166 | 1553,8/174 |
| **<0°** | -1638,4/163 | -1796,6/169 | -1546/163 | -1367,8/164 | -1245,5/167 | -1022,5/159 |
| **среднегодовая** | 1,1 | 1,0 | 0,9 | 1,2 | 1,2 | 1,6 |

Турано-Уюкской котловина, расположенная на высоте 700-900 м н.у.м. по основным показателям климатического режима – средней годовой температуре воздуха (-2,3°), продолжительности безморозного периода (160 дней) и $\Sigma t°>10°C=1573°$ приближается к условиям зоны засушливых степей. Термические условия Улуг-Хемской и Хемчикской котловин, днища которых имеют отметки 600-800 м н.у.м. характеризуются более высокой средней годовой температурой воздуха (-0,4°), более длительной продолжительностью безморозного периода (180 дней) и $\Sigma t°>10°C=2043°$. Убсу-Нурская котловина выделяется резкой континентальностью климата – низким атмосферным увлажнением и высокими тепловыми ресурсами лета: средняя годовая температура воздуха (-1,9°), продолжительность безморозного периода – 169 дней и $\Sigma t°>10°C=2034°$, которые создают режим полупустынной зоны с каштановыми почвами и псаммоземами.

**Картографическое моделирование неоднородности температурных полей почв.** Выявление возможности картографического моделирования структурной организации и функционирования почвенного покрова выполнено на примере степных котловин Среднесибирского трансекта и тундрово-степных комплексов высокогорий Алтае-Саянского региона. Интерпретация выделения типологических единиц структурной и функциональной организации почвенного покрова степных котловин Хакасии представлена на основе статистического анализа совмещения границ и информационного содержания контуров почвенного покрова и температурных полей, полученных на основе анализа спутниковых данных дневных, ночных температур и разницы температур, полученных путем вычитания двух скомпилированных серий (рис. 3).

Полученные картографические модели дают отчетливое представление о границах контуров температурных полей типологических единиц почвенного покрова Хакасии, функцио-

нирование которых осуществляется в широком диапазоне, как положительных, так и отрицательных температур. В котловинах Тувы, отличающихся высокой степенью разнообразия характеристик почвенного географического пространства и аридизацией климата, полученные картографические модели температурных полей, могут быть использованы для выявления пространственных и временных градиентов термических ресурсов в масштабе крупных и локальных единиц почвенного покрова (рис. 4).

Картографические модели температурных полей также являются информативными для выявления условий теплообеспеченности формирования и функционирования тундрово-степных комплексов высокогорного плоскогорья Укок, для почвенного покрова которых типичными являются сочетания горно-степных и горно-тундровых почв (рис. 5). Модели температурных полей могут использоваться для получения дополнительной информации о природных комплексах, формирование и функционирование которых осуществляется в условиях ультравысоких или ультранизких значений температур.

В целом, картографические модели температурных полей, полученные в результате совместного анализа наземных и спутниковых данных, обладают достаточной информативностью в отношении энергетики почвообразовательных процессов и могут быть использованы для оценки термических условий почв малоисследованных и труднодоступных территорий Алтае-Саянского региона.

Табл. 2. Показатели температурного режима воздуха и почв (°С) степных котловин Тувы по глубинам почвенного профиля (см). В знаменателе указано количество дней.

| **Показатели температурного режима, °С** | **Температура воздуха, °С** | **Температура по глубинам почвенного профиля (см), °С** | | | |
|---|---|---|---|---|---|
| | | **поверхность почвы** | **10** | **20** | **50** |
| Турано-Уюкская котловина, Туран, т.5 (N52°08'18,8"; E93°49'25"). | | | | | |
| **годовая** | -752,3 | -85,7 | -25,5 | -33,5 | 42,6 |
| **>10°** | 1572,8/95* | 1754,1/96 | 1470,3/88 | 1284,9/82 | 908,9/70 |
| **>5°** | 1768,5/121 | 1951,4/124 | 1726/124 | 1536,5/116 | 1210,3/112 |
| **>0°** | 1874/160 | 2026,5/152 | 1822,8/161 | 1660/163 | 1312,7/156 |
| **<0°** | -2626,3/169 | -2112,3/177 | -1843,8/168 | -1693,5/166 | -1270,3/171 |
| **среднегодовая** | -2,3 | -0,3 | -0,1 | -0,1 | 0,1 |
| Улуг-Хемская котловина, Кызыл, т.6 (N51°44'36,5"; E94°19'34,4") | | | | | |
| **годовая** | 121,5 | 787,3 | 598,3 | 1126,7 | 923 |
| **>10°** | 2043,3/107 | 2327,3/120 | 1995,1/112 | 2151,3/121 | 1730,1/107 |
| **>5°** | 2349,9/149 | 2603,5/156 | 2268,9/148 | 2458,9/163 | 2083,8/156 |
| **>0°** | 2440,7/180 | 2672,4/183 | 2350,5/179 | 2515,4/192 | 2157/183 |
| **<0°** | -2319,2/149 | -1882,5/146 | -1752,3/150 | -1388,8/137 | -1234/126 |
| **среднегодовая** | 0,4 | 2,4 | 1,8 | 3,4 | 2,8 |
| Хемчикская котловина, Ак-Довурак, т.7 (N51°13'0,2"; E90°31'38") | | | | | |
| **годовая** | -203,3 | 567,2351,8 | 351,8 | 185,1 | 596,3 |
| **>10°** | 2034,2/110 | 2484,9/125 | 2238,1/121 | 2155,7/118 | 2117,8/117 |
| **>5°** | 2320,7/148 | 278,4/154 | 2450,5/149 | 2399,7/150 | 2376,7/151 |
| **>0°** | 2383,6/174 | 2767/177 | 2511,3/175 | 2459,7/173 | 2440,3/170 |
| **<0°** | -2586,9/155 | -2199,8/152 | -2159,5/154 | 1994,9/156 | -1844/159 |
| **среднегодовая** | -0,6 | 1,7 | 1,1 | 0,6 | 1,8 |
| Убсу-Нурская котловина, оз. Тере-Холь, т.8 (N50°15'23,4"; E95°0,2'32,3") | | | | | |
| **годовая** | -653,1 | 794 | 579,5 | 658,6 | 882,8 |
| **>10°** | 2207/125 | 2575/132 | 2400/133 | 2320/131 | 2226/128 |
| **>5°** | 2407/151 | 2755/155 | 2522/150 | 2480/154 | 2448/158 |
| **>0°** | 2461/169 | 2830/182 | 2603/179 | 2542/179 | 2496/179 |
| **<0°** | -3114/172 | -2035/159 | -2024/162 | -1884/162 | -1613/162 |
| **среднегодовая** | -1,9 | 2,3 | 1,7 | 1,9 | 2,6 |

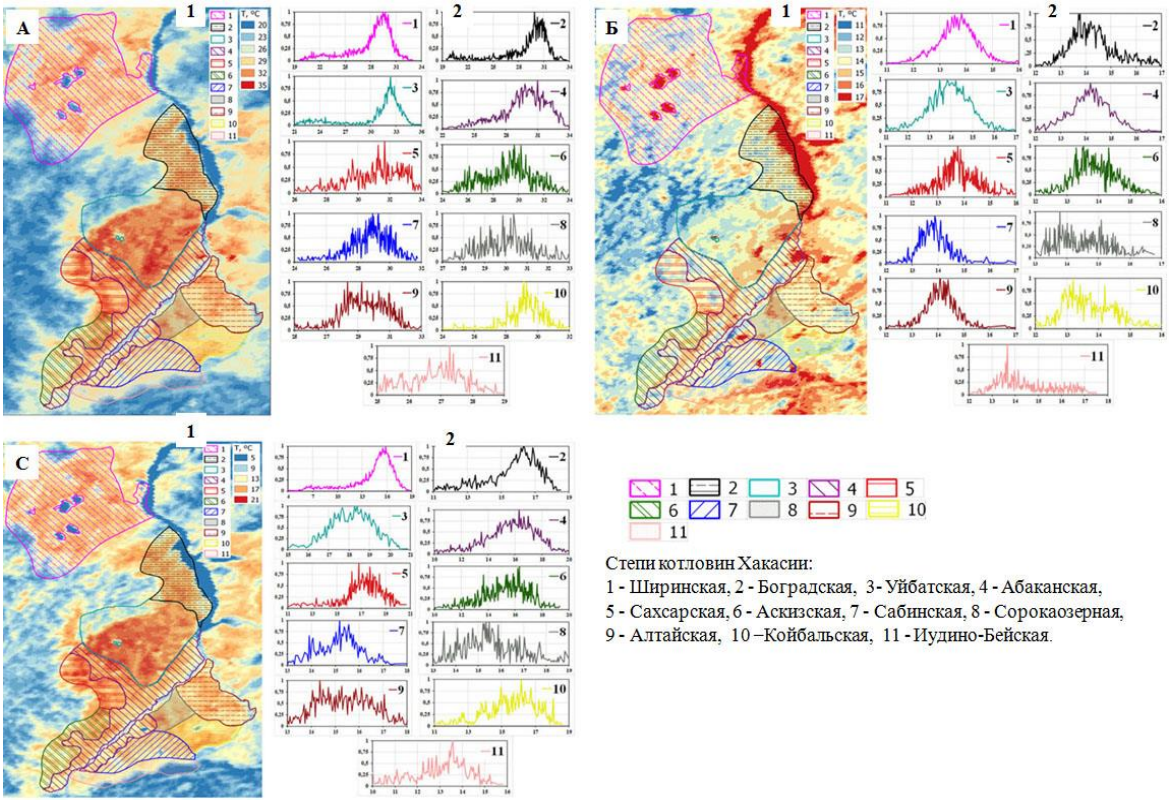



Рис. 3. Картографические модели температурных полей степей Хакасии, созданные на основе анализа спутниковых данных: *a* – дневных; *b* – ночных; *c* – разницы температур, полученных путем вычитания двух скомпилированных серий. *1* – контуры температурных полей; *2* – статистический анализ совмещения границ почвенных контуров и температурных полей.

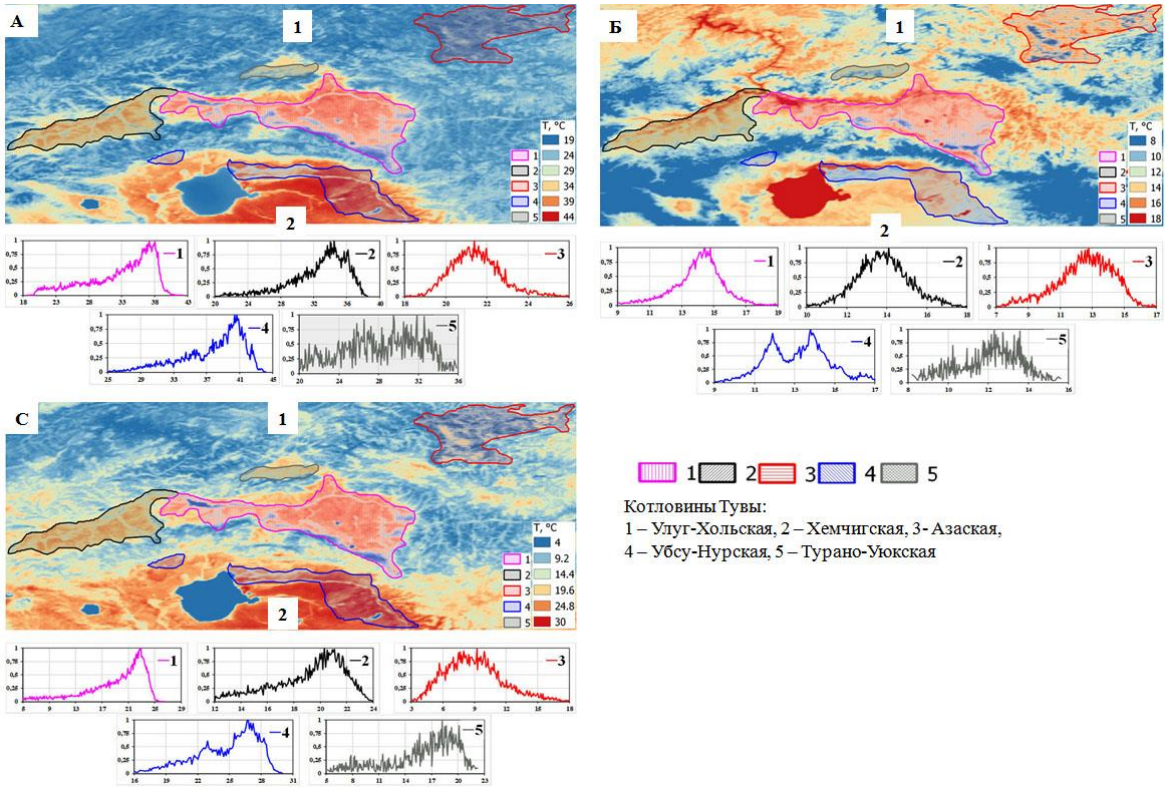



Рис. 4. Картографические модели температурных полей котловин Тувы, созданные на основе анализа спутниковых данных: *a* – дневных; *b* – ночных; *c* – разницы температур, полученных путем вычитания двух скомпилированных серий. *1* – контуры температурных полей; *2* – статистический анализ совмещения границ почвенных контуров и температурных полей.

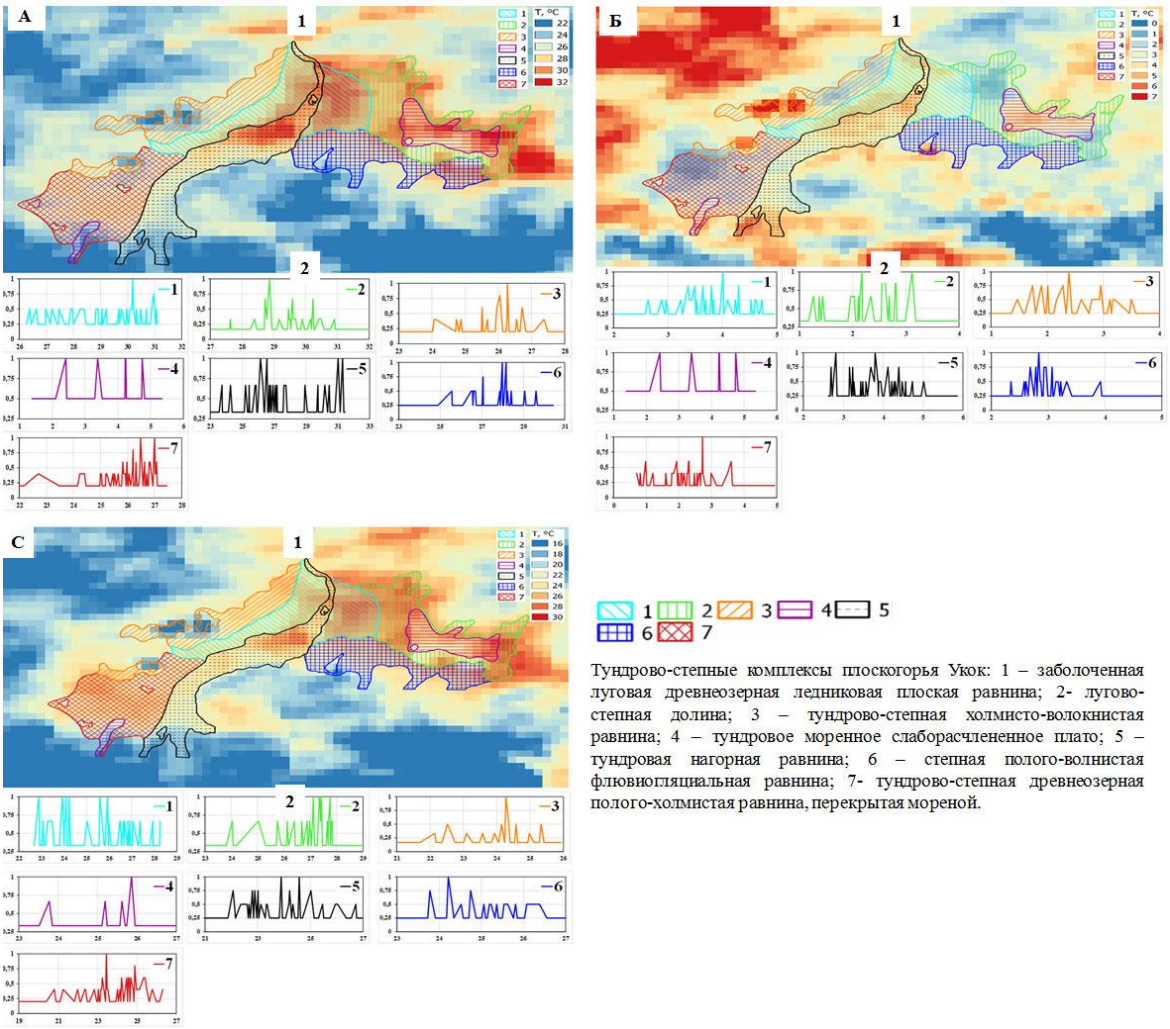



Рис. 5. Картографические модели температурных полей тундрово-степных комплексов плоскогорья Укок, созданные на основе анализа спутниковых данных: *a* – дневных; *b* – ночных; *c* – разницы температур, полученных путем вычитания двух скомпилированных серий. *1* – контуры температурных полей; *2* – статистический анализ совмещения границ почвенных контуров и температурных полей.

Новизна подхода заключается в том, что картографические модели температурных полей почв, созданные на основе количественных показателей температурного режима, обладают достаточной информативностью, позволяющей устанавливать взаимосвязи с другими характеристиками объектов природной среды и подойти к решению проблемы выделения границ температурных полей типологических единиц почвенного покрова.

*Работа выполнена при финансовой поддержке РФФИ (грант № 18-04-0063-а).*

## ЛИТЕРАТУРА

[1] *Minasny B., McBratney A.B.* Digital Soil mapping: A brief history and some lessons // Geoderma. 2015. Vol. 184. P. 201-211.

[2] *Paradian J., Morris J., Minasny B., McBratney A.B.* Pedotransfer Function and Soil Inference System // Pedometrics. 2018. P. 195-220

[3] *Кудряшова С.Я., Чумбаев А.С., Пестунов И.А., Безбородова А.Н., Курбатская С.С., Рылов С.А., Синявский Ю.Н.* Применение данных дистанционного зондирования и наземного автоматизированного мониторинга температурных полей почв для картографического моделирования почвенного покрова // Новые методы и результаты исследований ландшафтов в Европе, Центральной Азии и Сибири (в пяти томах) М.: Изд-во ФГБНУ «ВНИИ агрохимии», 2018. С. 238-242.

[4] *Кихтенко В.А., Чубаров Д.Л.* Технологии параллельной обработки данных в задачах анализа больших объемов спутниковых снимков // Сборник тезисов докладов Четырнадцатой Всероссийской открытой конференции «Современные проблемы дистанционного зондирования Земли из космоса»: Электронный сборник тезисов докладов. 2016. С. 88.

[5] *Синявский Ю.Н., Пестунов И.А., Дубровская О.А., Рылов С.А., Мельников П.В., Ермаков Н.Б., Полякова М.А.* Методы и технология сегментации мультиспектральных изображений высокого разрешения для исследования природных и антропогенных объектов // Вычислительные технологии. 2016. Т. 21, № 1. С. 127-140.

[6] *Синявский Ю.Н., Рылов С.А.* Экспериментальное сравнение непараметрических алгоритмов кластеризации для сегментации мультиспектральных изображений // Интерэкспо Гео-Сибирь. 2018. Т. 1, № 4. С. 109-114.

[7] *Добротворский Д.И., Куликова Е.А., Пестунов И.А., Синявский Ю.Н.* Веб-сервисы для непараметрической классификации спутниковых данных // Интерэкспо Гео-Сибирь. 2010. Т. 1, № 2. С. 171-175.

# РАДИОЛОКАЦИОННОЕ ОБНАРУЖЕНИЕ СОСРЕДОТОЧЕННЫХ ОБЪЕКТОВ ПРИ ИСПОЛЬЗОВАНИИ ШИРОКОПОЛОСНОГО ИЗЛУЧЕНИЯ

*Лозовский И.Ф.[(1)], Мурасев А.А.[(1)(2)], Спектор А.А.[(2)]*
[(1)] АО «НПО НИИИП-НЗиК», Новосибирск
[(2)] Новосибирский государственный технический университет, Новосибирск

В докладе изложен алгоритм обнаружителя сосредоточенных объектов при использовании широкополосного излучения. Получены характеристики разработанного обнаружителя. Приведен сравнительный анализ с некогерентным обнаружителем.
*Ключевые слова: iирокополосное излучение, радиолокационное обнаружение.*

**Введение.** Одним из интенсивно развиваемых направлений в импульсной радиолокации является применение широкополосных радиосигналов [1] с большой базой $B = \Delta f \cdot \tau_{\text{И}}$, где $\Delta f$ – ширина спектра сигнала, $\tau_{\text{И}}$ – длительность импульса. Преимущество таких сигналов заключается в возможности раздельного управления разрешающей способностью по дальности $\delta R = c/(2\Delta f)$, зависящей от ширины полосы сигнала, и энергией $E = P\tau_{\text{И}}$, определяющей дальность действия, $P$ – средняя мощность импульсного сигнала.

Классическая модель отраженного сигнала опирается на представление цели в виде совокупности одного вторичного излучателя (блестящей точки), дающего стабильный отраженный сигнал, и множества распределенных отражателей, формирующих случайную компоненту с гауссовским распределением. Такое представление цели оправдано в задачах, когда физический размер цели по дальности $l_R$ меньше разрешаемого размера $l_R < \delta R$. Увеличение разрешающей способности приводит к нарушению этого условия, возникают противоположные ситуации $l_R > \delta R$, причем, превышение размера цели над разрешением может достигать нескольких десятков раз.

В этих условиях математическую модель сигнала цели предлагается описывать в виде некоторого числа блестящих точек, число и размещение которых на ее поверхности заранее неизвестно, может изменяться в процессе наблюдения. Высокое разрешение приводит к тому, что после радиоприема отраженного сигнала, включающего в себя согласованную фильтрацию, сигнал цели имеет вид последовательности неизвестного числа разрешаемых импульсов с неизвестными заранее интервалами между ними. Вполне очевидно, что обнаружение цели следует осуществлять с учетом этих особенностей сигнала.

В данной работе изложен разработанный метод обнаружения сигнала цели в условиях высокого временного разрешения, при котором обеспечиваются обычные радиолокационные требования стабилизации вероятности ложной тревоги и достижения наибольшей вероятности обнаружения.

**Структура и рабочий алгоритм обнаружителя.** Принятый сигнал цели на выходе согласованного фильтра представляет собой последовательность импульсных сигналов, число и взаимное временное положение которых для обнаружителя следует рассматривать как случайные параметры. Для получения максимального эффекта обнаружения цели интервал времени $T_a$, используемый для анализа принятого сигнала, целесообразно согласовывать с ожидаемым размером цели, т.е. выбирать по правилу:

$$T_a = l_R/\mathrm{c},$$

где - $\mathrm{c} = 3 \cdot 10^8$ скорость распространения электромагнитных волн.

Пусть сигнал на входе обнаружителя описывается выражением

$$u(t) = \begin{cases} s(t) + n(t), & \text{при наличии цели} \\ n(t), & \text{при отсутствии цели} \end{cases},$$

где $s(t)$ и $n(t)$ – полезный сигнал и шум на выходе согласованного фильтра соответственно.

В качестве рабочего сигнала используется радиоимпульс с прямоугольной огибающей и линейной частотной модуляцией

$$\mathrm{s_1(t)} = \mathrm{S_0 cos[\omega_0 t + bt^2 + \varphi_0]},\ 0 \le \mathrm{t} \le \tau_{\text{и}},$$

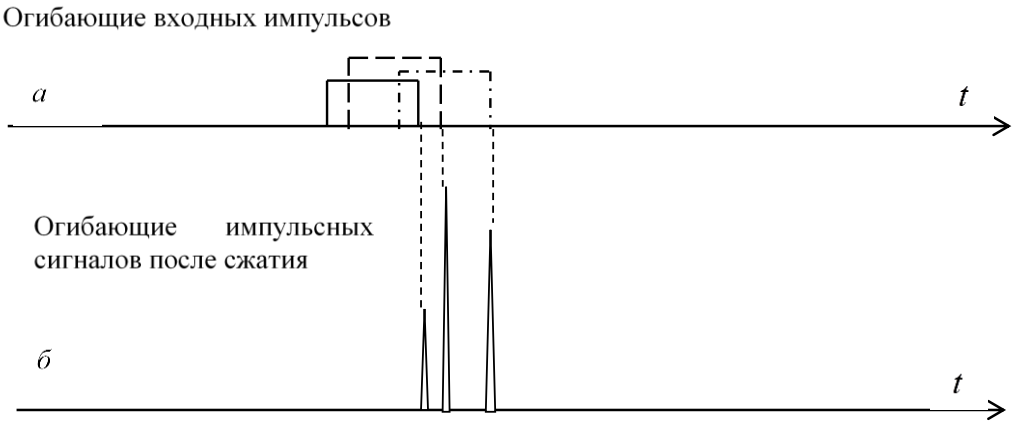



Рис. 1. Вид огибающих сигнала цели на входе (*а*) и выходе (*б*) согласованного фильтра при наличии трех блестящих точек.

где $b = \frac{\pi \Delta f_Д}{\tau_и}$ – скорость изменения частоты, $\Delta f_Д$ – девиация частоты, $\varphi_0$ – начальная фаза. Тогда сигнал, отраженный от цели с $n_{БТ}$ блестящими точками и имеющей дальность $R_Ц$, будет определяться выражением

$$s_Ц(t) = \sum_{i=1}^{n_{БТ}} s_1(t - t_Ц - t_i),$$

где $t_Ц = 2R_Ц/c$ – общее время задержки отраженного сигнала; $t_i$ – время задержки между отраженными сигналами, определяемое расстояниями между блестящими точками наблюдаемой цели. Сигналы, образуемые при отражении различными блестящими точками, накладываются друг на друга, что схематически показано на рис. 1,*а*. При большой базе сигнала $B \gg 1$ при согласованной фильтрации происходит сжатие отдельных входных импульсов в B раз, и картина выходного сигнала имеет вид рис. 1,*б*.

При таком поведении сигнала значительная часть интервала анализа будет содержать только шумовую составляющую, и, следовательно, шум будет давать существенный вклад в решающую статистику, определяемую по правилу классического некогерентного накопления (алгоритм А1) [2]

$$\sum_i {u_i}^2 \underset{<"0"}{>"1"} z_0.$$

Таким образом, возникает идея двухпороговой обработки, где в решающую статистику входят только те отсчеты сигнала, которые превысили первый порог. Структура обнаружителя представлена на рис.2. Принимаемый сигнал оцифровывается, подвергается согласованной фильтрации, поступает на амплитудный детектор и далее на обнаружитель. Сигнал на выходе $ПУ_1$ представляет собой последовательность редких выбросов, которые формируют решающую статистику по правилу

$$L = \sum_{i=1}^{n} (u_{1i} - z_1),$$

где $u_{1i}$ – отсчеты $u_i$ превысившие порог $z_1$. Решение обнаружителя определяется по правилу (алгоритм А2)

$$L \underset{<"0"}{>"1"} z_2.$$

**Характеристики обнаружения.** Определим основные характеристики обнаружителя – вероятности ложной тревоги и правильного обнаружения.

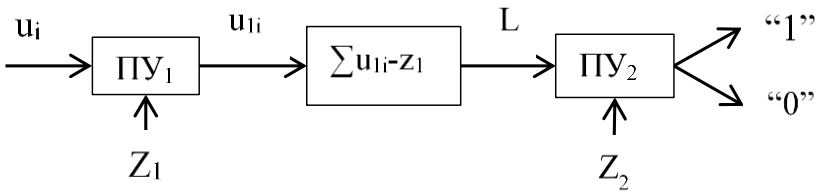



Рис. 2. Структура обнаружителя.

Число выбросов помехи, превысивших порог $z_1$, подчиняется пуассоновскому закону распределения вероятностей [3]

$$p(\mathrm{n}) = \frac{\lambda T_a}{n!} e^{-\lambda T_a},\ n = 0,1,2 \dots,$$

где λ – среднее число превышений порога $z_1$ в единицу времени

$$\lambda(\mathrm{z}_1) = \frac{\omega_1}{2\pi} \exp(-\frac{z_1^2}{2\sigma_n^2}),$$

$$\omega_1 = \sqrt{\frac{\int_0^\infty \omega^2 g(\omega) d\omega}{\int_0^\infty g(\omega) d\omega}},$$

$g(\omega)$ – спектральная плотность мощности шума на выходе согласованного фильтра.

Полагаем, что в пределах выброса имеется единственный отсчет помехи. Условная плотность распределения вероятностей (ПРВ) отсчета помехи, превысившего порог $z_1$

$$q_0(u_{1i}) = \frac{1}{\int_{z_1}^\infty p_0(u_i) du_i} p_0(u_i),$$

где $p_0(u_i)$ – безусловное распределение отсчета $u_i$

$$p_0(u_i) = \frac{u_i}{\sigma^2} \exp(-\frac{u_i^2}{2\sigma^2}),$$

$\sigma^2$ – дисперсия шума на входе обнаружителя.

Условная вероятность ложной тревоги, при условии, что произошло n превышений порога $z_1$

$$F_n = \int_{z_2}^\infty q_\mathrm{n}(L_n) dL_n,$$

где $q_\mathrm{n}(L_n)$ – условное распределение решающей статистики.

Т.к. статистика $L_n$ является суммой случайных величин $u_i$, то ее ПРВ можно определить с помощью характеристических функций [3] $Q_n$

$$q_\mathrm{n}(L_n) = F^{-1}\{Q_n\},$$
$$Q_n = F\{\mathrm{q}_0^n(u_{1i})\},$$

где $F\{\cdot\}$ и $F^{-1}\{\cdot\}$ - прямое и обратное преобразование Фурье.

Тогда полная вероятность ложной тревоги будет определяться выражением

$$F = \sum_{n=1}^\infty p_n F_n.$$

Число выбросов полезного сигнала определяется количеством $n_\mathrm{БТ}$ блестящих точек. Положим, что выбросы помехи отсутствуют на интервале анализа $T_\mathrm{a}$.

ПРВ числа превышений сигналом порога $z_1$

$w_n = C_{n_\mathrm{БТ}}^n p_1^n (1-p_1)^{n_\mathrm{БТ}-n},\ n = 0 \dots n_\mathrm{БТ},$

где $p_1$ – вероятность однократного превышения порога $z_1$

$$p_1 = \int_{z_1}^\infty p_1(u_i) du_i,$$

$p_1(u_i)$ – ПРВ отсчета $u_i$ при наличии полезного сигнала

$$p_1(u_i) = \frac{u_i}{\sigma^2} \exp(-\frac{u_i^2 - E_s^2}{2\sigma^2}) I_0(\frac{u_i E_s}{\sigma^2}),$$

$E_s$ – энергия ЛЧМ-импульса $s_1(t)$.

Условная ПРВ отсчета сигнала, превысившего порог $z_1$

$$v_0(u_{1i}) = \frac{1}{\int_{z_1}^\infty p_1(u_i) du_i} p_1(u_i),$$

Условная вероятность правильного обнаружения, при условии, что произошло n превышений порога $z_1$

$$D_n = \int_{z_2}^\infty v_\mathrm{n}(L_n) dL_n,$$

где $v_\mathrm{n}(L_n)$ – условное распределение решающей статистики

$$v_\mathrm{n}(L_n) = F^{-1}\{V_n\},$$
$$V_n = F\{v_0^n(u_{1i})\}.$$

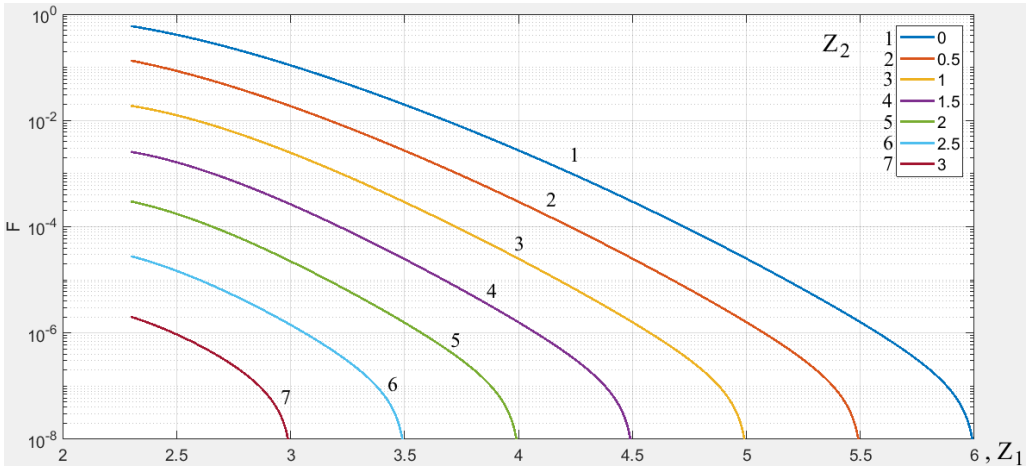


Рис. 3. Зависимость вероятности $F$ ложной тревоги от порогов $z_1$ и $z_2$.

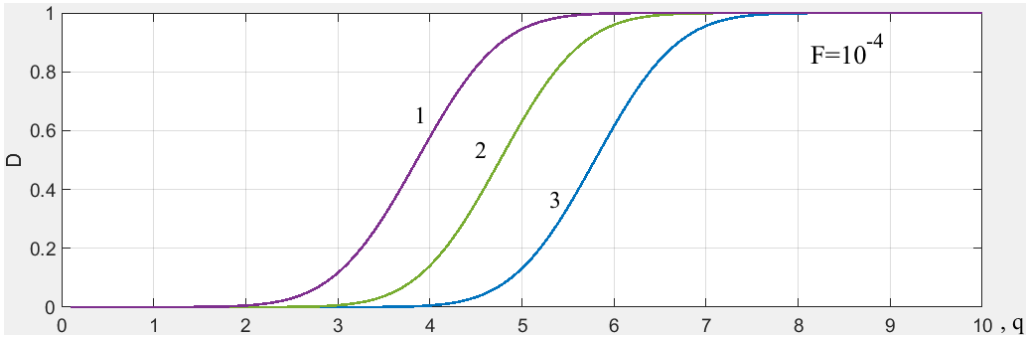


Рис. 4. Зависимость вероятности $D$ правильного обнаружения от ОСШ $q$.

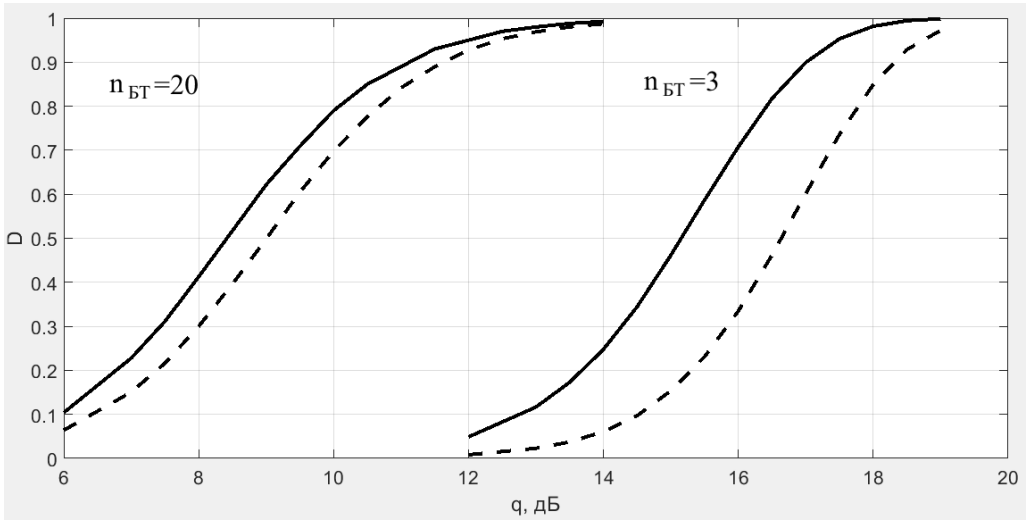


Рис. 5. Зависимость вероятности $D$ от ОСШ $q$(дБ) для двух алгоритмов обнаружения: А2 – сплошная, А1 – штриховая.

Полная вероятность правильного обнаружения

$$D = \sum_{n=1}^{\infty} w_n D_n.$$

На рис. 3 Представлена зависимость вероятности $F$ ложной тревоги от порогов $z_1$ и $z_2$. Как видно из графика, требуемая вероятность $F$ может быть получена при разных комбинациях порогов. Таким образом, подбирая пороги, можно максимизировать вероятность $D$ правильного обнаружения. Зависимость вероятности $D$ от отношения сигнал/шум (ОСШ) $q = E_s/\sigma$ при разных комбинациях порогов и $F = 10^{-4}$ изображена на рис. 4. Видно, что выбор порогов оказывает значительное влияние на качество работы обнаружителя. Выигрыш по

ОСШ для оптимального случая рис. 4 (кривая 1) относительно худшего рис. 4 (кривая 3) составляет 3 дБ по уровню $D = 0.9$. Характеристики получены при параметрах сигнала: $\tau_и = 1$ мкс, $\Delta f_Д = 150$ МГц, $f_0 = 100$ МГц.

Сравнение алгоритмов А1 и А2 , приведено на рис. 5. Результаты получены для размера цели по дальности $l_R = 20$ м, числа блестящих точек $n_{БТ} = 3$ и $n_{БТ} = 20$, вероятности ложной тревоги $F = 10^{-6}$, $\tau_и = 1$ мкс, $\Delta f_Д = 100$ МГц, $f_0 = 75$ МГц.. Как видно из графика, алгоритм А2 при $n_{БТ} = 3$ дает выигрыш по ОСШ в 1.3 дБ. При увеличении $n_{БТ}$ характеристики алгоритмов сближаются, т.к. область, занимаемая полезным сигналом на интервале анализа увеличивается.

**Заключение.** В данной работе рассмотрен алгоритм обнаружения сосредоточенных объектов при использовании широкополосного излучения. Проведен сравнительный анализ разработанного алгоритма с классическим алгоритмом некогерентного накопления. Полученные результаты подтверждают эффективность разработанного алгоритма.

## ЛИТЕРАТУРА


[1] *Лозовский И.Ф.* Цифровая обработка сигналов в РЛС обзора. Новосибирск: НГУ, 2016. 270 с.

[2] *Казаринов Ю.М.* Радиотехнические системы. М.: Изд-во «Академия», 2008. 592 с.

[3] *Левин Б.Р.* Теоретические основы статистической радиотехники. Книга первая. М.: Изд-во «Советское радио», 1969. 752 с.

# НЕПАРАМЕТРИЧЕСКОЕ ОБНАРУЖЕНИЕ СЕЙСМИЧЕСКОГО СИГНАЛА ПО ОПОРНОМУ И РАБОЧЕМУ УЧАСТКУ АМПЛИТУДНОГО СПЕКТРА

*Морозов Ю.В., Райфельд М.А., Спектор А.А.*
Новосибирский государственный технический университет, Новосибирск

В настоящей работе исследован непараметрический метод обнаружения сигнала в системе пассивной сейсмической локации. Решение об обнаружении сейсмического сигнала принимается на основе сравнения соответствующих частотных составляющих рабочего и опорного участка амплитудного спектра. Приведены статистические характеристики обнаружения, полученные путем компьютерного моделирования.
*Ключевые слова: пассивная сейсмическая локация, сейсмический сигнал, обнаружение, непараметрический метод, быстрое преобразование Фурье.*

**Введение.** Системы пассивной сейсмической локации служат для обнаружения оказывающих механическое воздействие на грунт объектов, а также определения их местоположения и классификации [1]. Решение упомянутых задач невозможно без применения алгоритмического и программного обеспечения [2].

Основное внимание в настоящей работе будет уделено обнаружению объектов с импульсным воздействием на грунт типа «человек». В любых видах локации обнаружение требует обеспечение заданной вероятности ложной тревоги при максимизации вероятности заданного обнаружения [3].

В [4] предложен алгоритм обнаружения человека на основе спектров акустических и сейсмических сигналов, где определяется апостериорная вероятность наличия человека в зоне наблюдения системы пассивной сейсмической локации с использованием байесовского подхода. Авторы указывают, что предложенный в данной работе подход обеспечивает вероятность правильного обнаружения около 70%, причем вероятность ложной тревоги ими не рассматривается.

В [5] предложен алгоритм обнаружения сейсмического сигнала человека на основе динамической фильтрации символов (SDF), полученных на основе вейвлет-преобразования сигнала во временной области. Вероятность обнаружения объекта формируется на основании вероятностей смены состояний конечного вероятностного автомата. Для ее повышения авторы предлагают дополнять геофоны инфракрасными датчиками, что влечет усложнение и удорожание локационной системы.

В [6] предлагается строить алгоритм обнаружения сейсмического сигнала на основе автокорреляционных функций во временной области. Недостатком данного подхода является, что время вычислений напрямую зависит от длительности интервала, на котором вычисляется автокорреляционная функция.

В [7] предложена процедура обнаружения сейсмического сигнала с помощью корреляционного анализа коэффициентов вэйвлет-спектра, который позволяет отделить коэффициенты, относящиеся к полезному сигналу от коэффициентов, относящихся к шуму. В данной работе не анализируются конкретные вероятностные характеристики обнаружения. Достоинством вейвлет-преобразования является возможность исследования свойств нестационарных сигналов. Однако практическое применение вейвлет-преобразования затруднено чрезмерной вычислительной трудоемкостью.

В [8,9] предложено для обнаружения и распознавания сейсмических сигналов применять нейронные сети. Предложенная в [8] нейронная сеть на основе перцептронов, которая работает со спектральными составляющими сейсмического сигнала, способна обеспечить вероятность ложной тревоги около 1%, что является слишком большой величиной для используемых на практике систем пассивной сейсмической локации. В [9] предложена нейронная сеть обратного распространения в комплексе со счетчиком пересечения нулей и отношением энергий, интегрированных по верхним и по нижним частотам.

Нейронные сети обладают существенными недостатками, к которым относятся сложность построения архитектуры сети для конкретной задачи и сложность интерпретации результатов обучения [10]. Значение параметров нейронной сети почти невозможно объяснить в терминах решаемой задачи, с связи с чем она остается «черным ящиком» и для исследователей, и для пользователей. Применение оптимизационных методов минимизации среднеквадратической ошибки при обучении сети приводит к ее переобучению [11]. Чувствительность сети к шуму сильно зависит от ее архитектуры. Обеспечение вероятности правильного обнаружения выше 90% требует использования громоздкой иерархической архитектуры, где в первом приближении вектор признаков обрабатывается более грубой сетью, а затем решение корректируется более точной и более медленной сетью. Для обучения нейронных сетей требуются большие объемы экспериментального материала, полученные при тех или иных условиях окружающей среды. Это возможно только в регионах, где длительное время стоит солнечная погода без больших температурных колебаний. В регионах, где возможны резкие изменения температуры и влажности, качественное обучение нейронной сети, обнаруживающей, принимающей решение о наличии сейсмического сигнала, весьма проблематично.

В упомянутых выше работах не уделяется достаточное внимание конечной задаче обнаружения сейсмического сигнала с учетом критерия Неймана-Пирсона.

Непараметрическое обнаружение, основанное на счетчике пересечений нуля, рассматривалось в статье [12]. Данный подход имеет значительно меньшую трудоемкость, чем ранее упомянутые подходы. Количество пересечений снижается при появлении в наблюдаемом сигнале составляющей сейсмически активного объекта. Обнаружение осуществляется по сигналу, получаемому после выбеливания фона. Полезный сигнал, также подвергающийся данной процедуре, остается коррелированным, из-за чего число пересечений нуля оказывается статистически меньшим, чем в присутствии только сейсмического фона. Недостатком этого обнаружения является высокие требования к качеству выбеливания, в частности, требование мгновенной адаптации при изменении характеристик сейсмического фона.

В настоящей работе предложена процедура непараметрического обнаружения сейсмического сигнала в частотной области.

**Постановка задачи**. Целью исследований является построение и исследование процедуры обнаружения сейсмических сигналов в системе пассивной сейсмической локации. Данная процедура должна соответствовать критерию Неймана-Пирсона. При фиксированной допустимой вероятности ложной тревоги вероятность правильного обнаружения должны достигать максимума.

**Теория**. Система пассивной сейсмической локации обычно включает в себе группу из датчиков сейсмических сигналов, расположенных, например, как показано в [1]. Расстояние между соседними датчиками обычно составляет около 10 м. При появлении сейсмоактивного объекта, оказывающего воздействие на грунт, порожденные им сигналы воспринимаются одним или несколькими датчиками. Эти датчики образуют группу активных датчиков, сигналы которых подвергаются предварительной обработке для последующего обнаружения. Примеры сигналов, получаемых с датчиков для объектов упомянутых классов, приведены в виде нескольких циклов по 1000 отсчетов в [1].

Исходные сигналы дискретизируются с частотой 600 Гц. Для дальнейшей обработки дискретные отсчеты объединяются в циклы по 1000 отсчетов, что соответствует интервалу времени 1.67 с. Декорреляция или выбеливание осуществляется для подавления сейсмического фона.

Для сигнала каждого цикла выполняется преобразование Фурье и вычисляется амплитудный спектр. На каждом n-м цикле сигнал $Sx_n$, $n=\overline{1,N}$ представляет собой *J*-элементный вектор. Для каждого из сигналов $Sx_n, n=\overline{1,N}$ выполняется преобразование Фурье и вычисляются амплитудные спектры, в результате чего образуются векторы $X^{(n)}=\left\|x_j^{(n)},\quad j=\overline{1,J}\right\|,\qquad n=\overline{1,N}$. Каждый амплитудный цикл разбивается на опорный участок

с номерами гармоник $Y^{(n)} = \left\| x_j^{(n)}, \quad j = \overline{j_1, j_2} \right\|, \qquad n = \overline{1, N}$ и рабочий участок, смещенный относительно опорного на $\Delta j$ гармоник вниз $Z^{(n)} = \left\| x_j^{(n)}, \quad j = \overline{j_1 - \Delta j, j_2 - \Delta j} \right\|, \qquad n = \overline{1, N}$. Решающая статистика $U(Y, Z)$ формируется в соответствии с соотношением:

$$U(Y,Z) = \frac{1}{N} \sum_{n=1}^{N} U^{(n)}(Y^{(n)}, Z^{(n)}), \tag{1}$$

а решение о наличии или отсутствии сейсмически активного объекта (САО) выносится по правилу:

$$U(X,Y) \begin{cases} \geq U_0 & \Rightarrow \quad "CAO \quad присутствует" \\ < U_0 & \Rightarrow \quad "CAO \quad отсутствует" \end{cases} \tag{2}$$

Статистика $U(Y, Z)$ формируется из локальных статистик отдельных циклов $U^{(n)}(Y^{(n)}, Z^{(n)})$, которые, в свою очередь, определяются по правилу:

$$U^{(n)}(Y^{(n)}, Z^{(n)}) = \sum_{j=j_1}^{j_2} u(x_{j-\Delta j}^{(n)} - x_j^{(n)}), \tag{3}$$

где $u(\zeta) = \begin{cases} 1 & \text{при } \zeta \geq 0 \\ 0 & \text{при } \zeta < 0 \end{cases}$. Значения $\zeta_j^{(n)} = u(x_{j+\Delta j}^{(n)} - x_j^{(n)}), \qquad j = \overline{1, J/2}, \quad n = \overline{1, N}$ образуют независимую двоичную последовательность нулей и единиц с равномерным распределением вероятностей $p_0 = p_1 = 0,5$, если выборки $Y$ и $Z$ образованы сейсмическим фоном и являются однородными. Независимость вытекает из известного свойства некоррелированности спектральных отсчетов стационарных случайных процессов [13], а также из принадлежности сигналов сейсмического фона гауссовскому распределению. В [14] было предложено использовать опорный и рабочий участок в разных циклах. Однако, в этом случае требуется достаточно трудоемкая корректирующая процедура, чтобы обеспечить независимость спектральных отсчетов.

На рис. 1 приведен пример амплитудного спектра сейсмического сигнала для объекта типа «человек». На частотах с 50-й по 150-ю наблюдается характерный подъем, на остальных частотах спектр напоминает спектр сейсмического фона. Следовательно, участок с 50-й по 150-ю гармонику принимается за рабочий, а участок, смещенный, например, на 200 гармоник вверх (с 250-й по 350-ю гармонику), принимается за опорный.

**Вычислительные эксперименты**. Путем компьютерного имитационного моделирования были получены серии по M=2000 циклов сигналов сейсмического фона и человека. Для каждой серии подсчитывалось количество циклов K, для которых принималось решение о наличии САО согласно правилу (2). Далее определялось экспериментальное значение вероятности обнаружения САО

$$P_э = K / M \ . \tag{4}$$

По этим данным были построена зависимость вероятности ложной тревоги от порога обнаружения и зависимость вероятности правильного обнаружения от отношения сигнал-шум (ОСШ) при разных вероятностях ложной тревоги.

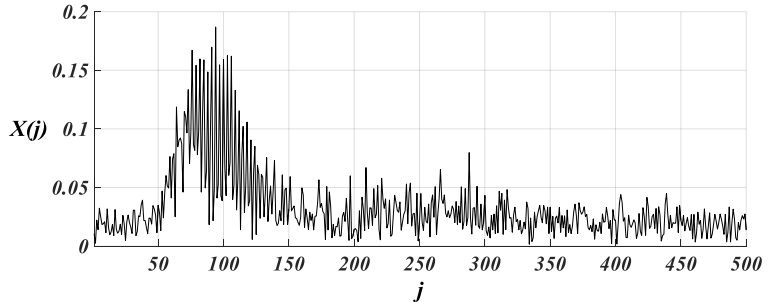



Рис. 1. Амплиутдный спектр сейсмического сигнала человека.

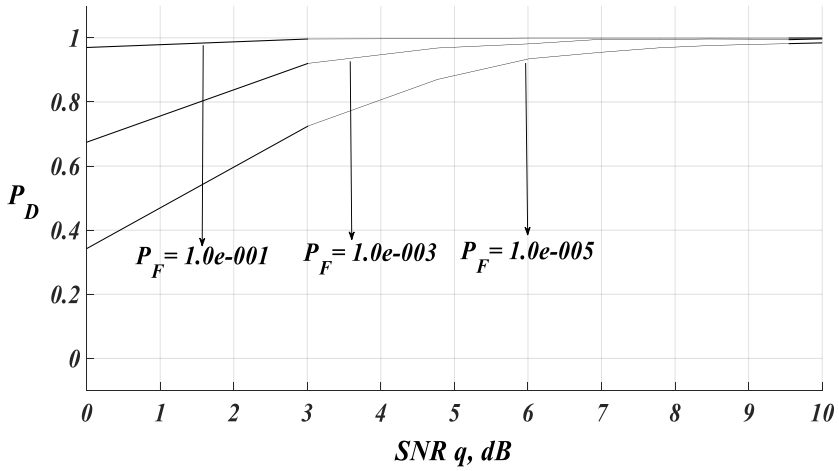



Рис. 2. Зависимость вероятности правильного обнаружения от ОСШ при $N$=1.

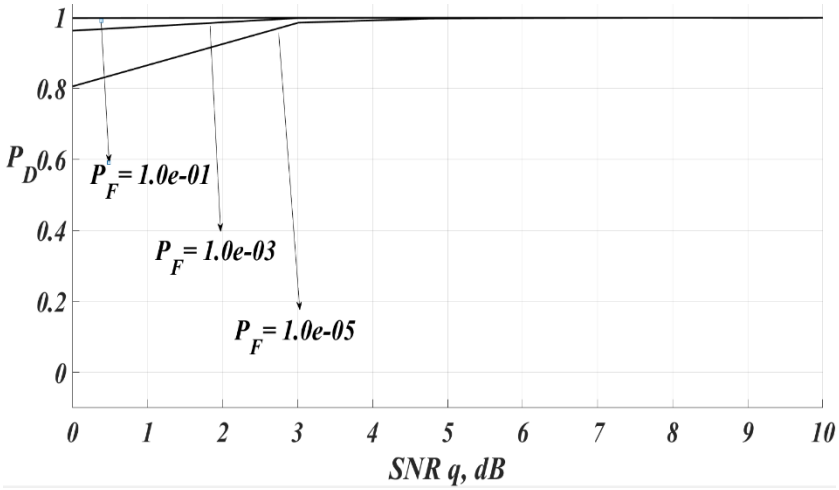



Рис. 3. Зависимость вероятности правильного обнаружения от ОСШ.

Вероятность правильного обнаружения определялась для различных значений ОСШ (рис. 2) при N=1. ОСШ определяется отношением средних квадратичных значений флюктуаций полезного сигнала и сейсмического фона.

Естественно, при любом значении вероятности ложной тревоги наблюдается монотонное увеличение вероятности правильного обнаружения с ростом отношения сигнал-шум.

При N=2 вероятность правильного обнаружения при тех же вероятностях ложной тревоги существенно повышается. Однако, в этом случае увеличивается время выполнения необходимых вычислений.

**Заключение.** Предложенный в работе метод непараметрического обнаружения сейсмических сигналов на основе подсчета количества превышений гармоник в рабочем участке амплитудного спектра над соответствующими гармониками опорного участка обеспечивает достаточно высокое значение вероятности обнаружения при стабилизации уровня ложной тревоги, что достигается за счет непараметрического свойства предложенного алгоритма. Так, например, при вероятности ложной тревоги 0.001 и ОСШ 5дБ вероятность правильного обнаружения превышает значение 0.9. При усреднении упомянутого количества превышений по двум измерительным циклам вероятности правильного обнаружения при тех же самых ОСШ и вероятностях ложной тревоги значительно возрастают. Предложенный метод обнаружения сейсмических сигналов полезен в условиях априорной неопределенности, когда имеется недостаточная информация о статистических свойствах сигналов. В дальнейшем планируется исследовать предложенный метод сейсмических сигналов от других объектов с импульсным воздействием на грунт.

## ЛИТЕРАТУРА


[1] *Morozov Y.V., Rajfeld M.A., Spector A.A.* Analysis of seismic signals preliminary processing influence on classification results // 12 International forum on strategic technology (IFOST 2017) : proc., Korea, Ulsan, 31 May - 2 June 2017., Vol. 1, P. 138-142.

[2] *Красовский А.А.* Цифровая обработка в ZETLAB при идентификации параметров сейсмического сигнала // Цифровая обработка сигналов. 2010. №3. С. 70-76.

[3] *Акимов П.С., Бакут П.А., Богданович В.А. и др.* Теория обнаружения сигналов. М.: Радио и связь, 1984. 440 с.

[4] *Damarla T., Mehmood A., Sabatier J.* Detection of People and Animals Using Non-imaging Sensors // Proceedings of 14th International Conference on Information Fusion, Chicago, Illinois, USA, July 5-8, 2011, pp. 429-436.

[5] *Jin X., Sarkar S., Ray A., Gupte S., Damarla T.* Target Detection and Classification Using Seismic and PIR Sensors // IEEE Sensors Journal, 2012, Vol. 12, No. 6, pp. 1709-1718.

[6] *Alyamkin S.A., Eremenko S.I.* Pedestrian Detection Algorithms Based on an Analysis of the Autocorrelation Function of a Seismic Signal // Optoelectronics, Instrumentation and Data Processing, 2011, Vol. 47, No. 2, pp. 124-129.

[7] *Xing H.F., Li F., Liu Yu.L.* Wavelet Denoising and Feature Extraction of Seismic Signal for Footstep Detection // Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China, 2-4 Nov, 2007, pp. 218-223.

[8] *Faghfouri A.E., Frish M.B.* Robust discrimination of human footsteps using seismic signals // Proceedings of SPIE – The International Society for Optical Engineering, May 2011. **https://www.researchgate.net/publication/241347660**.

[9] *Peng Z., Cao C., Huang J., Liu Q.* Seismic Signal Recognition Using Improved BP Neural Network and Confirmed Feature Extraction Method // J. Cent. South Univ., 2014, Vol. 21, pp. 1898-1906.

[10] *Корнеев Д.С.* Использование аппарата нейронных сетей для создания модели оценки и управления рисками предприятия // Управление большими системами: сборник трудов, 2007, Вып. 14, с. 81-102.

[11] *Ивашко А.В., Потапенко А.И.* Алогритмы выделения объектов на изображении // Вестник Нац. техн. ун-та «ХПИ»: сб. науч. тр. Темат. вып.: Автоматика и приборостроение. Харьков: НТУ «ХПИ». 2011. № 57. С. 103-111.

[12] *Соколова Д.О., Спектор А.А.* Непараметрическое обнаружение стохастических сигналов, основанное на пересечениях с «нулем» // Вестник ТГУ. Управление, вычислительная техника и информатика. 2013. №1 (22). С. 138-146.

[13] *Рытов С.М.* Введение в статистическую радиофизику. М.: Наука, 1976. Ч. 1. 494 с.

[14] *Morozov Y.V., Rajfeld M.A., Spector A.A.* Non-parametric target detection in a passive seismic locator based on spectral data // Journal of Theoretical and Applied Information Technology. 2018. Vol. 96, No. 11. pp. 3567-3577.

# НЕЙРОСЕТЕВАЯ КЛАССИФИКАЦИЯ КРУПНОРАЗМЕРНЫХ МНОГОСЕГМЕНТНЫХ ПОЛИГОНОВ С ФОРМИРОВАНИЕМ ПРИЗНАКОВ ПУТЕМ ПРЕОБРАЗОВАНИЯ ГИЛЬБЕРТА – ХУАНГА ГИПЕРСПЕКТРАЛЬНЫХ ДАННЫХ

*Нежевенко Е.С.*
Институт автоматики и электрометрии СО РАН, Новосибирск

Описан метод классификации гиперспектральных изображений участков местности, включающий предобработку спектральных данных в виде преобразования к главным компонентам, пространственную обработку, состоящую из нахождения эмпирических мод главных компонент (преобразование Гильберта – Хуанга) и собственно классификацию путем использования нейронных RBF сетей. Эксперименты проводились на крупноформатном (580×580) гиперспектральном изображении, причем трудноразличимые классы не объединялись.
*Ключевые слова: гиперспектральные изображения, главные компоненты, преобразование Гильберта – Хуанга, нейронные сети.*

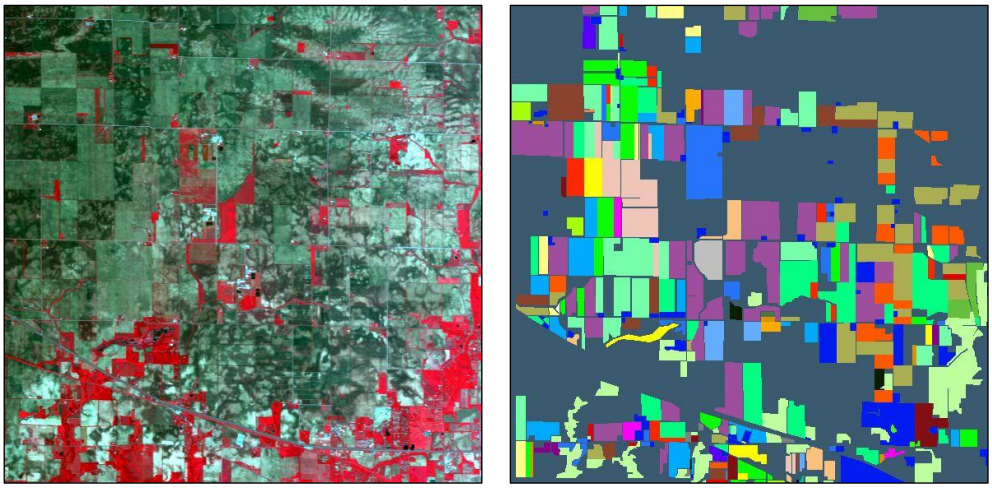
**Введение** Классификации участков подстилающей поверхности Земли по гиперспектральным данным в последнее время, с развитием средств наблюдения, становится важнейшим направлением науки о дистанционном зондировании Земли. Она позволяет решать важные задачи землепользования в сельском хозяйстве, лесоводстве, экологии, минимизации ущерба от природных катаклизмов и т.д. Этой проблеме посвящена обширная литература, сделан хороший обзор [1]. В нем представлено большое число методов классификации по гиперспектральным данным, в заключение авторы приводят конкретные результаты, полученные ими лично на крупноформатном гиперспектральном изображении (ГСИ). С сожалением следует отметить, что одно из направлений обработки в обзоре [1] упомянуто лишь мимоходом, а оно дало очень хорошие результаты на предыдущих среднеформатных как мультиспектральных, так и гиперспектральных изображениях [2, 3]. В настоящей работе мы хотим показать, что и на крупноформатном ГСИ наш подход дает такие результаты, которые можно поставить в ряду лучших, полученных другими методами. Однако сначала о самом методе. Он был описан нами подробно в [3], поэтому здесь мы изложим его очень кратко.

**Объект классификации.** Как понятно из введения, мы хотим сравнить наш метод с другими, описанными в литературе. Для того, чтобы это сделать, необходимо проверить его на подобном (а еще лучше одинаковом) объекте. В качестве такового мы взяли крупноформатное гиперспектральное изображение, полученного в рамках программы AVIRIS (Airborne Visible Infrared Imaging Spectrometer) на опытном полигоне Индиан Пайнс (штат Индиана, США) [4]. Размер изображения 614×2677 пикселей, разрешение – 20 м/пиксел. Изображение размечено на 58 классов. Кроме магистрали, железной дороги, леса, жилых построек на изображении имеются сельскохозяйственные культуры (в том числе 15 классов кукурузы и 18 сои, отличающиеся методиками возделывания почвы). Число спектральных каналов – 220 в диапазоне 0.4-2.5 мкм., при этом 20 каналов, содержащих высокий уровень шума, удалены из рассмотрения. Из этого изображения выделен информативный фрагмент 580×580 пикселов$^2$ (в псевдоцветах изображен на рис. 1,*а*). Все классы, попавшие в выделенный фрагмент, показаны в псевдоцветах на рис. 1,*б*. Для эксперимента выбрано 33 класса, в которых имеется значимое число пикселов. Из выбранных классов присутствуют 9 классов кукурузы и 12 классов сои.

**Метод классификации.** Как известно из теории распознавания, важнейшую роль при классификации играет предобработка распознаваемого объекта. Объект у нас спектрально-пространственный, соответственно, и предобработка будет вестись в спектрально-пространственной области. Как уже говорилось, каждый пиксел характеризуется 200 спектральными признаками. Анализ показывает, что они сильно коррелированы между собой, поэтому логично перейти к некоррелированным признакам. Это можно сделать разными способами, но самый эффективный – преобразование к главным компонентам (ГК). Существенный вопрос –

сколько компонент следует оставить, чтобы потерять минимум информации. Для этого применяются критерии: Кайзера, каменистой осыпи. После преобразования к главным компонентам проводится пространственное преобразование. Во многих публикациях [5-7] показано, что оно дает существенное увеличение результирующей точности классификации (overall accuracy), – процента отношения количества верно классифицированных пикселей к их общему количеству в выборке. Однако в этих публикациях, как и в большинстве других, пространственное преобразование сводится к изменению значения пиксела в зависимости от его окрестности. При этом совершенно не учитывается структура классифицируемых участков фрагмента, поэтому в случае трудноразличимых участков наблюдаются большие ошибки классификации. Из-за этого, в частности, в работах [6-8] все участки сои, отличающиеся способом обработки земли, объединены в один класс. То же относится и к кукурузе. В нашем методе использован другой прием. Каждая главная компонента была разложена на внутренние колебания (т.н. эмпирические моды, intrinsic mode function – IMF). Очень коротко скажем о них. В отличие от преобразования Фурье и вейвлет-преобразования эмпирические моды не заданы аналитически и определяются исключительно самой анализируемой последовательностью. При этом базисные функции преобразования формируются адаптивно, непосредственно из входных данных.

В основе алгоритма разложения на моды лежит построение гладких огибающих по максимумам и минимумам последовательности и дальнейшее вычитание среднего этих огибающих из исходной последовательности. Для этого производится поиск максимумов и минимумов и строятся сплайны, проходящие через эти точки. Эти сплайны и образуют верхнюю и нижнюю огибающие. Процесс построения огибающих показан на рис. 2.



*а* *б*

Рис. 1. Фрагмент гиперспектрального изображения (*а*) и его разметка на классы (*б*).

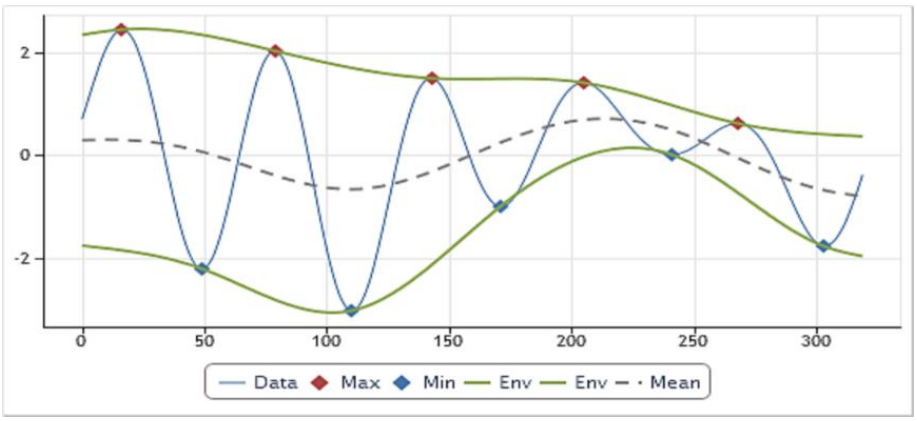



Рис. 2. Этап построения эмпирической моды одномерной функции.

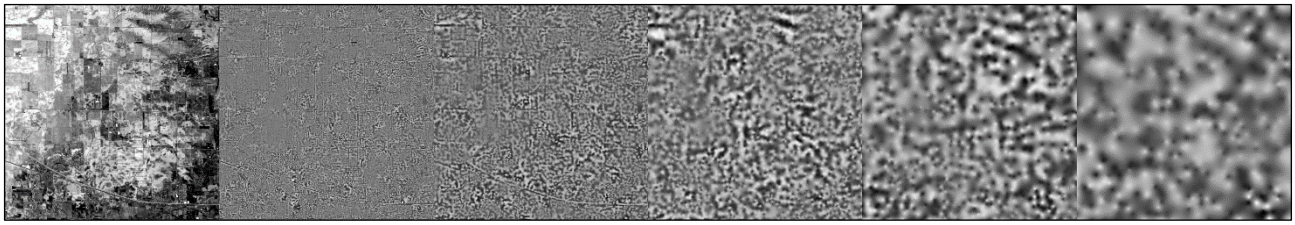


Рис. 3. Первая ГК и 5 ее эмпирических мод.

Анализируемая последовательность представлена на рис. 2 тонкой синей линией. Максимумы и минимумы этой последовательности отмечены красным и синим цветом соответственно. Огибающие показаны зеленым цветом. По двум огибающим вычисляется среднее значение. Это среднее показано на рис. 2 штриховой линией. Найденное таким образом среднее значение далее вычитается из исходной последовательности.

В результате перечисленных шагов в первом приближении находится первое приближение искомой IMF. Для полноценного выделения IMF необходимо вновь найти максимумы и минимумы у этой оценки IMF, и повторить изложенные ранее действия. Этот процесс (он называется отсеивание) продолжается до тех пор, пока не будет достигнут заданный критерий его останова. Если процесс просеивания удачно завершен, то в результате мы получаем первую найденную эмпирическую функцию. Далее, для того чтобы найти следующую IMF, необходимо из исходного сигнала вычесть уже найденную IMF и снова повторить описанную процедуру. Так продолжается до тех пор, пока не будут найдены все IMF. Алгоритм разложения двумерного сигнала на эмпирические моды формально не отличается от разложения одномерного сигнала, хотя, конечно, имеет некоторые специфические особенности, связанные с поиском экстремумов и интерполяцией [9, 10]. После нахождения $m$ эмпирических мод $n$ главных компонент каждый пиксел ГСИ описывается $m \times n$ мерным вектором признаков. Все вектора ГСИ вместе с информацией о классах поступают в нейронную сеть. Пикселы каждого класса случайным образом делятся на три выборки: обучающая (ОВ), контрольная (КВ) и тестовая (ТВ). Обучение останавливается по результатам проверки на контрольной выборке. Точность классификации проверяется на всех выборках, но главное – на тестовой, поскольку именно она определяет способность нейронной сети к обобщению.

**Результаты эксперимента.** ГСИ, представленное на рис. 1,*а*, было разложено на главные компоненты, из которых по критерию «каменистой осыпи» было отобрано 4 (на них приходится 99.42% дисперсии данных). Затем каждая из ГК была разложена на 5 эмпирических мод. Первая ГК и 5 ее эмпирических мод представлены на рис. 3.

Таким образом, после всех этих преобразований ГСИ описывается массивом 580×580×20, размеченным на 33 класса. Количество пикселов в каждом классе делится в соотношении: ОВ:КВ:ТВ= 50%:25%:25%. Обучение проводится в нейронной RBF сети. Время обучения несколько часов. Рассмотрим результаты классификации после обучения. Сначала покажем перечень всех классов с нумерацией, которая будет фигурировать в таблицах.

Суммарные результаты классификации представлены на табл. 2. Архитектура сети в левом окне таблицы обозначает: 20 входных признаков (число нейронов в первом слое, 851 нейронов (RBF функций) в скрытом слое и 33 нейрона в выходном слое (число классов). Мы видим, что точности классификации (ТК) не такие высокие, как на экспериментах, проведенных на меньших фрагментах и с меньшим количеством классов [2, 3], но еще раз напомним: здесь существуют трудноразличимые классы кукурузы и сои, не объединенные в два класса.

Как видно, процент ТК изменяется от 42% (класс 3 «здания») до 99,3% (класс 58 «лес»). Отметим, что в основном большие вероятности ТК дают классы с большим количеством пикселов. Это следует из того, что средний процент правильной классификации по классам – 83.63%, а средний процент правильной классификации по пикселам – 87.31%.

Посмотрим более подробные результаты классификации тестовой выборки с указанием процентов правильной классификации по классам (табл. 3).

Табл. 1. Перечень классов с нумерацией.

| | | | |
|---|---|---|---|
| 1 | фон | 31 | Овес |
| 2 | Голая почва | 33 | фруктовый сад |
| 3 | здания | 35 | пруд |
| 4 | Бетон / Асфальт | 36 | Соевые бобы |
| 7 | Кукуруза -запад-восток | 38 | Соевые бобы -север=юг |
| 8 | Кукуруза -север-юг | 41 | Соевые бобы обычная вспашка -запад-восток |
| 9 | Кукуруза обычная вспашка | 42 | Соевые бобы- обычная вспашка -север-юг |
| 12 | Кукуруза - обычная вспашка-орошаемая-север-юг | 43 | Соевые бобы обычная вспашка -посадка в бороздуе |
| 14 | Кукуруза-малоразрушающая вспашка | 44 | Соевые бобы- заросшие сорняками- обычная вспашка |
| 15 | Кукуруза- малоразрушающая вспашка- запад-восток | 45 | Соевые бобы посадка в борозду |
| 16 | Кукуруза- малоразрушающая вспашка-север-юг | 46 | Соевые бобы малоразрушающая вспашка |
| 17 | Кукуруза без вспашки | 47 | Соевые бобы- малоразрушающая вспашка -запад-восток |
| 18 | Кукуруза без вспашки -запад-восток | 51 | Соевые бобы - без вспашки -запад-восток |
| 25 | Травяная взлетно-посадочная полоса | 52 | Соевые бобы - без вспашки -север-юг |
| 27 | сено? | 53 | Соевые бобы - без вспашки -посадка в борозду |
| 29 | Озеро | 57 | пшеница |
| | | 58 | Лес |

Табл. 2 Суммарные точности классификации фрагмента.

| Файл «Все участки 33 класса» | | | |
|---|---|---|---|
| Архитектура сети | Точность классификации обучающая выборка % | Точность классификации контрольная выборка % | Точность классификации тестовая выборка % |
| RBF 20-851-33 | 89,35 | 87,36 | 87,31 |

Табл. 3. Результаты классификации фрагмента по классам.

| № классов | 1 | 2 | 3 | 4 | 7 | 8 | 9 | 12 | 14 | 15 | 16 | 17 | 18 | 25 | 27 | 29 | 31 | 33 | 35 | 36 | 38 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 51 | 52 | 53 | 57 | 58 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % ошибочных | 6.2 | 19.0 | 57.1 | 12.3 | 8.0 | 12.1 | 19.8 | 6.1 | 11.1 | 9.5 | 15.8 | 11.4 | 9.1 | 32.6 | 28.2 | 26.6 | 27.4 | 10.8 | 15.4 | 1.8 | 22.4 | 18.7 | 29.6 | 8.8 | 11.0 | 34.3 | 20.2 | 30.0 | 10.1 | 9.0 | 3.8 | 3.7 | 0.7 |
| % правильных | 93.8 | 81.0 | 42.9 | 87.7 | 92.0 | 87.9 | 80.2 | 93.9 | 88.9 | 90.5 | 84.2 | 88.6 | 90.9 | 67.4 | 71.8 | 73.4 | 72.6 | 89.2 | 84.6 | 98.2 | 77.6 | 81.3 | 70.4 | 91.2 | 89.0 | 65.7 | 79.8 | 70.0 | 89.9 | 91.0 | 96.2 | 96.3 | 99.3 |
| Ошибочно | 552 | 374 | 109 | 501 | 88 | 623 | 206 | 41 | 324 | 395 | 163 | 102 | 493 | 71 | 66 | 79 | 34 | 243 | 276 | 7 | 431 | 141 | 243 | 48 | 569 | 135 | 68 | 334 | 52 | 621 | 22 | 123 | 1 |
| Правильно | 8363 | 1593 | 82 | 3582 | 1013 | 4547 | 837 | 635 | 2592 | 3752 | 871 | 796 | 4949 | 147 | 168 | 218 | 90 | 2006 | 1514 | 372 | 1489 | 614 | 577 | 495 | 4615 | 259 | 268 | 781 | 464 | 6258 | 561 | 3234 | 143 |
| Всего | 8915 | 1967 | 191 | 4083 | 1101 | 5170 | 1043 | 676 | 2916 | 4147 | 1034 | 898 | 5442 | 218 | 234 | 297 | 124 | 2249 | 1790 | 379 | 1920 | 755 | 820 | 543 | 5184 | 394 | 336 | 1115 | 516 | 6879 | 583 | 3357 | 144 |

Табл. 4. Матрица ошибок классификации для фрагмента.

| | 1 | 2 | 3 | 4 | 7 | 8 | 9 | 12 | 14 | 15 | 16 | 17 | 18 | 25 | 27 | 29 | 31 | 33 | 35 | 36 | 38 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 51 | 52 | 53 | 57 | 58 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 100 | 0.71 | 0.23 | 1.30 | 0.66 | 2.06 | 0.60 | 0.35 | 0.86 | 1.11 | 0.01 | 0.07 | 1.49 | 0.01 | 0.22 | 0.00 | 0.04 | 0.04 | 0.19 | 0.01 | 1.21 | 1.03 | 1.09 | 0.51 | 2.18 | 1.04 | 0.29 | 1.20 | 0.00 | 2.88 | 0.02 | 0.29 | 0.00 |
| 2 | 1.44 | 100 | 1.57 | 0.88 | 0.00 | 0.38 | 0.06 | 0.00 | 0.31 | 0.13 | 0.00 | 0.13 | 1.76 | 1.32 | 0.25 | 0.00 | 0.13 | 6.28 | 0.06 | 0.00 | 0.75 | 0.00 | 0.82 | 0.25 | 0.75 | 0.00 | 0.00 | 0.06 | 0.00 | 0.44 | 0.38 | 2.32 | 0.00 |
| 3 | 1.22 | 1.22 | 100 | 0.00 | 0.00 | 6.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.22 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 4 | 0.47 | 0.75 | 0.00 | 100 | 0.00 | 0.64 | 1.17 | 0.00 | 0.45 | 0.00 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.08 | 0.00 | 0.03 | 2.74 | 0.00 | 0.42 | 0.08 | 0.06 | 0.00 | 1.54 | 0.00 | 0.00 | 0.45 | 0.56 | 0.39 | 0.00 | 0.11 | 0.00 |
| 7 | 2.86 | 0.10 | 0.00 | 0.00 | 100 | 0.00 | 0.10 | 0.00 | 0.39 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.97 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.10 | 0.00 |
| 8 | 2.24 | 0.53 | 1.10 | 2.27 | 0.00 | 100 | 0.33 | 0.00 | 0.24 | 0.53 | 0.00 | 0.37 | 0.90 | 0.04 | 0.00 | 0.09 | 0.04 | 0.22 | 0.02 | 0.07 | 1.14 | 0.09 | 1.67 | 0.00 | 2.60 | 0.18 | 0.31 | 0.37 | 0.00 | 1.08 | 0.00 | 0.07 | 0.00 |
| 9 | 0.96 | 0.00 | 0.00 | 4.18 | 0.48 | 0.00 | 100 | 0.00 | 0.24 | 0.96 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.67 | 0.00 | 0.00 | 0.00 | 0.12 | 0.36 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.24 | 0.00 | 0.72 | 0.00 | 0.00 | 0.12 |
| 12 | 0.31 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.47 | 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 14 | 0.54 | 0.08 | 0.00 | 0.00 | 0.12 | 0.04 | 0.19 | 0.00 | 100 | 0.46 | 0.00 | 0.04 | 0.04 | 0.04 | 0.00 | 0.04 | 0.00 | 0.08 | 0.12 | 0.00 | 0.08 | 0.00 | 0.77 | 0.00 | 0.54 | 0.00 | 0.23 | 0.08 | 0.00 | 1.16 | 0.00 | 0.04 | 0.00 |
| 15 | 1.65 | 0.24 | 0.00 | 0.03 | 0.40 | 0.56 | 0.19 | 0.00 | 0.85 | 100 | 0.00 | 0.21 | 0.91 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.03 | 1.36 | 0.72 | 0.00 | 0.00 | 1.07 | 0.11 | 0.03 | 2.11 | 0.59 | 1.95 | 0.00 | 0.11 | 0.00 |
| 16 | 0.23 | 0.00 | 0.00 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.00 | 0.11 | 100 | 0.00 | 0.92 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.95 | 0.00 | 0.92 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 2.87 | 0.00 | 0.00 | 0.00 |
| 17 | 0.38 | 0.13 | 0.00 | 0.13 | 0.00 | 7.16 | 0.00 | 0.00 | 0.63 | 1.51 | 0.00 | 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.88 | 0.00 | 0.13 | 0.00 | 0.00 | 0.00 | 2.64 | 0.00 | 0.00 | 0.50 | 0.00 | 0.63 | 0.00 | 0.00 | 0.00 |
| 18 | 0.61 | 0.10 | 0.04 | 0.12 | 0.00 | 0.53 | 0.22 | 0.00 | 0.61 | 1.29 | 0.81 | 0.00 | 100 | 0.12 | 0.00 | 0.14 | 0.00 | 0.14 | 0.18 | 0.00 | 0.30 | 0.12 | 0.73 | 0.00 | 0.89 | 0.14 | 0.06 | 0.44 | 0.00 | 2.28 | 0.00 | 0.06 | 0.00 |
| 25 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.68 | 0.00 |
| 27 | 2.38 | 3.57 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.00 | 0.60 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.19 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.60 | 0.00 |
| 29 | 0.00 | 0.00 | 0.00 | 0.00 | 0.46 | 0.00 | 0.00 | 0.00 | 2.75 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.92 | 0.92 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 31 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.00 | 2.22 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 33 | 0.40 | 6.18 | 0.00 | 0.00 | 0.00 | 0.15 | 0.00 | 0.00 | 0.00 | 0.15 | 0.00 | 0.00 | 0.80 | 0.85 | 1.79 | 0.00 | 0.00 | 100 | 0.55 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.50 | 0.00 | 0.60 | 0.00 | 0.10 | 0.25 | 0.70 | 0.00 |
| 35 | 0.59 | 0.00 | 0.00 | 4.82 | 0.00 | 2.18 | 0.00 | 0.00 | 0.00 | 0.07 | 3.70 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.00 | 0.00 | 0.40 | 0.00 | 0.00 | 0.26 | 0.00 | 0.00 | 0.00 | 0.00 | 0.13 | 0.00 | 0.00 | 0.00 |
| 36 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.61 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 38 | 1.68 | 0.87 | 0.07 | 2.75 | 0.40 | 1.21 | 1.48 | 0.00 | 0.54 | 0.94 | 0.00 | 0.47 | 0.07 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 | 0.00 | 100 | 0.00 | 0.00 | 0.00 | 1.41 | 0.00 | 0.07 | 0.20 | 0.00 | 0.20 | 0.00 | 0.00 | 0.00 |
| 41 | 2.93 | 0.00 | 0.00 | 0.00 | 0.00 | 0.81 | 0.98 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.14 | 100 | 0.16 | 0.16 | 0.00 | 0.00 | 0.65 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 42 | 0.00 | 0.52 | 0.00 | 0.00 | 0.00 | 0.69 | 0.00 | 0.00 | 0.35 | 0.00 | 0.00 | 0.00 | 0.69 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.39 | 0.00 | 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 43 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.20 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.20 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.00 | 0.00 | 0.20 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 44 | 1.67 | 0.30 | 0.17 | 1.47 | 0.04 | 2.15 | 0.59 | 0.07 | 1.82 | 0.95 | 0.67 | 0.46 | 1.91 | 0.30 | 0.00 | 0.69 | 0.02 | 0.15 | 0.30 | 0.02 | 1.73 | 0.11 | 0.02 | 0.00 | 100 | 0.13 | 0.26 | 1.39 | 0.00 | 0.85 | 0.00 | 0.07 | 0.00 |
| 45 | 3.09 | 3.09 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.39 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 46 | 1.87 | 0.00 | 0.00 | 1.12 | 0.00 | 0.75 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.12 | 0.00 | 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 47 | 0.26 | 0.00 | 0.26 | 0.51 | 0.00 | 0.00 | 0.38 | 0.00 | 0.51 | 0.51 | 0.00 | 0.13 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.26 | 0.00 | 0.00 | 0.00 | 2.18 | 0.00 | 0.00 | 100 | 0.00 | 0.90 | 0.00 | 0.38 | 0.00 |
| 51 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 6.90 | 0.00 | 0.00 | 0.22 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 1.08 | 0.00 | 0.00 | 0.00 |
| 52 | 1.52 | 0.35 | 0.02 | 0.50 | 0.03 | 2.19 | 0.19 | 0.14 | 0.69 | 1.28 | 0.56 | 0.62 | 2.27 | 0.11 | 0.00 | 0.21 | 0.24 | 0.48 | 1.01 | 0.00 | 1.15 | 0.06 | 0.03 | 0.00 | 0.56 | 0.21 | 0.03 | 0.03 | 0.16 | 100 | 0.00 | 0.21 | 0.00 |
| 53 | 0.89 | 0.18 | 0.18 | 0.00 | 0.00 | 1.43 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.53 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.18 | 0.00 | 0.00 | 100 | 1.96 | 0.00 |
| 57 | 0.09 | 1.67 | 0.00 | 0.37 | 0.00 | 0.06 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 | 0.25 | 0.12 | 0.25 | 2.35 | 0.00 | 0.00 | 0.03 | 0.00 | 0.03 | 0.00 | 0.03 | 0.00 | 0.00 | 0.09 | 0.00 | 0.00 | 0.28 | 100 | 0.00 |
| 58 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 |

Посмотрим матрицу ошибок (табл. 4), которая показывает, к каким классам относятся неправильно классифицированные данные. Каждая строчка матрицы нормирована на число правильно классифицированных в соответствующих классах пикселов.

Серыми зонами отмечены классы одной культуры: левый – все классы кукурузы, правый – все классы сои. Из сравнения этих зон с другими областями можно сделать вывод, что в области сои суммарная перекрестная ошибка в несколько раз меньше, чем в других областях. Это говорит о том, что участки сои, отличающиеся видом обработки или другими характеристиками, действительно трудноразличимы. В меньшей степени это относится к участкам кукурузы. Вероятно, такое различие связано со спецификой обработки участков кукурузы и сои.

**Заключение.** Таким образом, эксперименты по классификации крупноформатного гиперспектрального изображения подтвердили высокую эффективность метода, включающего предварительную обработку в виде преобразования к главным компонентам спектральных составляющих, пространственное преобразование путем разложения главных компонент на эмпирические моды и классификацию в нейронной сети. Метод дает высокую вероятность правильной классификации близких по спектральному составу гиперспектральных изображений и эффективно работает в случае трудноразличимых обычными методами участков местности.

## ЛИТЕРАТУРА


[1] *Борзов С.М., Потатуркин О.И.* Спектрально-пространственные методы классификации гиперспектральных изображений, обзор // Автометрия, 2018. 54, № 6, с. 64-86

[2] *Нежевенко Е.С., Феоктистов А.С.* Исследование эффективности нейросетевой классификации гиперспектральных изображений с использованием преобразования Гильберта – Хуанга // Интерэкспо Гео-Сибирь. 2016. Т. 1. С. 60-64.

[3] *Нежевенко Е.С., Феоктистов А.С., Дашевский О.Ю.* Нейросетевая классификация гиперспектральных изображений на основе преобразования Гильберта – Хуанга // Автометрия. 2017. Т. 53, № 2. С. 79-85.

[4] *Baumgardner M.F., Biehl L.L., Landgrebe D.A.* 220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3. Purdue University Research Repository. 2015. DOI: 10.4231/R7RX991C.

[5] *Кузнецов А.В., Мясников В.В.* Сравнение алгоритмов управляемой поэлементной классификации гиперспектральных изображений // Компьютерная оптика. 2014. Т. 38, № 4. С. 494-502.

[6] *Борзов С.М., Потатуркин А.О., Потатуркин О.И., Федотов А.М.* Исследование эффективности классификации гиперспектральных спутниковых изображений природных и антропогенных территорий // Автометрия. 2016. Т. 52, № 1. С. 3-14.

[7] *Борзов С.М., Потатуркин О.И.* Исследование эффективности спектрально-пространственной классификации данных гиперспектральных наблюдений // Автометрия. 2017. Т. 53, № 1. С. 32-42.

[8] *Борзов С.М., Потатуркин О.И.* Классификация гиперспектральных изображений при различных способах формирования обучающих выборок // Автометрия. 2018. Т. 54, № 1. С. 89-97.

[9] *Vincent L.* Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms // IEEE Trans. Image Process. 1993. Vol. 2, N 2. P. 176-201.

[10] *Carr J.C., Fright W.R., Beatson R.K.* Surface interpolation with radial basis functions for medical imaging // Proc. ACM SIGGRAPH. Los Angeles, USA, 12-17 Aug. 2001. P. 67-76.

# ВИРТУАЛЬНАЯ ИССЛЕДОВАТЕЛЬСКАЯ СРЕДА В ФУНДАМЕНТАЛЬНЫХ И ПРИКЛАДНЫХ ЗАДАЧАХ КЛИМАТОЛОГИИ

*Гордов Е.П.[(1)(2)], Окладников И.Г.[(1)(2)], Титов А.Г.[(1)(2)], Рязанова А.А.[(1)]*
[(1)] Институт мониторинга климатических и экологических систем СО РАН, Томск
[(2)] Томский филиал ИВТ СО РАН, Томск

Представлено описание и новые функциональные возможности разрабатываемой виртуальной вычислительно-информационной среды для анализа, оценки и прогноза последствий глобальных климатических изменений для окружающей среды и климата в выбранном регионе.
*Ключевые слова: виртуальная исследовательская среда, большие наборы данных об окружающей среде, изменения климата.*

**Введение.** Недавние исследования показали, что наука становится все более глобальной, многополярной и распределенной [1]. Это приводит к необходимости разработки современных, открытых и универсальных исследовательских программных продуктов, обеспечивающих удаленный и совместный доступ к различных архивам данных, а также их «облачную» обработку и анализ распределенными междисциплинарными группами ученых. Такие продукты в настоящее время обычно называют виртуальными исследовательскими средами [2], хотя у них есть много других названий: научные шлюзы (scientific gateways), киберинфраструктура (cyberinfrastructure) и т. д. Согласно работе Л. Канделы [3], термин виртуальная исследовательская среда (ВИС) определяет программный комплекс со следующими основными характеристиками: (i) это рабочая среда, доступная через Интернет, (ii) предназначенная для удовлетворения потребностей целевого сообщества, (iii) предоставляющая ему продукты и инструменты, необходимые для достижения поставленных целей, и (iv) способствующая обмену результатами исследований.

В науках о Земле такой подход к решению научных проблем уже стал необходимостью: соответствующие задачи по сути являются междисциплинарными, решаются географически распределенными исследовательскими коллективами, а объемы наборов данных достигают десятков петабайт (http://newsletter.copernicus.eu/article/data-volume). Продолжающиеся изменения климата, особенно их экстремальные проявления, оказывают все большее влияние на экономические, политические и социальные процессы [4-6]. Надежный анализ этих процессов имеет важное значение для разработки адекватных стратегий адаптации и смягчения их негативных последствий (например, для сельского хозяйства, лесного хозяйства или планируемой инфраструктуры).

Надежный анализ изменений климата и окружающей среды и реакции общества на них требуют навыков работы с громоздкими наборами метеорологических данных, умения взаимодействовать с мощными вычислительными ресурсами и сложными числовыми моделями, знания современных методов статистического анализа для работы с большими архивами моделей. выходы и использование языков программирования высокого уровня. Эти навыки не характерны для специалистов в области экономических, политических и социальных наук, имеющих дело с явлениями и процессами, на которые сильно влияет изменение климата. К сожалению, такие навыки совершенно нехарактерны для лиц, принимающих решения, в том числе ответственных за разработку мер по адаптации. Поэтому разработка ВИС должна быть направлена на то, чтобы предоставить специалистам и лицам, принимающим решения, надежные инструменты для изучения экономических, политических и социальных последствий изменения климата.

В статье представлена новая функциональность тематической ВИС для поддержки климатических исследований. В основе ВИС лежит специализированная веб-ГИС «Климат» (http://sclimate.scert.ru), представляющая собой кроссплатформенный клиент-серверный про-

граммный комплекс с открытым исходным кодом, сочетающий веб- и ГИС-функциональность, поддерживающий «облачную» обработку и анализ геопространственных климатических данных, и предоставляющий весь спектр функциональных возможностей в окне любого современного интернет-браузера на рабочем месте пользователя. Разрабатываемая ВИС обеспечит междисциплинарные распределенные исследовательские группы, не являющиеся экспертами в области информационных технологий (климатологи, экологи, биологи и лица, принимающие решения), доступными и надежными онлайн-инструментами для быстрого анализа и визуализации многомерных разнородных климатических наборов данных, полученных из различных источников.

**Тематическая виртуальная исследовательская среда.** Разрабатываемая ВИС предназначена для обработки, анализа и визуализации геопространственных наборов данных в задачах наук о Земле. Она создается на основе веб-ГИС «Климат», разработанной в ИМКЭС СО РАН [7-9], которая состоит из трех ключевых компонентов: вычислительного ядра на стороне сервера; промежуточного программного обеспечения на стороне сервера, представленного геопорталом и набором PHP-контроллеров; и веб-клиента, основанного на технологии AJAX и разработанного с использованием специализированной JavaScript-библиотеки, содержащей типовые виджеты графического интерфейса пользователя. Для обработки наборов геопространственных данных используется набор проверенных программных модулей вычислительного ядра. Результаты обработки геопространственных данных представляются на карте выбранной местности перекрывающимися растровыми и векторными картографическими слоями, сопровождаемыми соответствующими метаданными. Подробно архитектура и функционирование ВИС представлены в работе А. Барта [10].

Функциональные возможности ВИС включают в себя базовый и комплексный статистический анализ данных, а элементы геоинформационной системы (ГИС) дают пользователю возможность комбинировать и отображать результаты с географической привязкой на выбранной картографической основе. ВИС предоставляет специалистам и пользователям без навыков программирования надежные и удобные онлайн-инструменты для комплексного исследования изменений климата и экосистем в едином веб-интерфейсе пользователя. Некоторые примеры успешного применения разрабатываемой ВИС в исследованиях наблюдаемых изменений климата Сибири и их последствий обсуждаются в ряде работ [11-13].

**Климатические исследования на региональном уровне.** Для описания экстремальных климатических явлений требуется использование специальных статистических методов. Корректное статистическое описание экстремальных осадков и температуры может быть получено с использованием статистики экстремальных значений (extreme value statistics – EVS) [14]. Программная реализация EVS на языке R (пакет «extRemes») в виде специализированных программных модулей позволяет выполнять статистический анализ максимальных значений выбранных метеовеличин на основе нестационарного обобщенного распределения экстремальных значений. Значения, необходимые для оценки риска, могут быть рассчитаны с использованием функции распределения, описывающей вероятность того, что значение наблюдаемой переменной превысит определенный уровень. Эти уровни часто выражаются как уровни возврата $r_T$ для определенного периода возврата $T$. Величина $r_T$ определяется как уровень, который в среднем превышает каждый $T$, т.е. с вероятностью $\frac{1}{T}$.

После интеграции разработанных программных модулей в веб-ГИС «Климат», новая функциональность была апробирована на расчёте 100-летних уровней возврата максимального количества осадков в июле на основе данных ECMWF ERA Interim (рис. 1,$a$) и APHRODITE JMA (рис. 1,$b$) для Южно-Сибирского региона (52,5-60 °с.ш., 75-95 °в.д.). Хотя рисунки в целом показывают схожее поведение вычисленной характеристики, результаты JMA APHRODITE (рис. 1,$b$) демонстрируют больше деталей и более высокие значения в отдельных регионах. Это связано с разным пространственным разрешением и разными источниками анализируемых наборов данных.

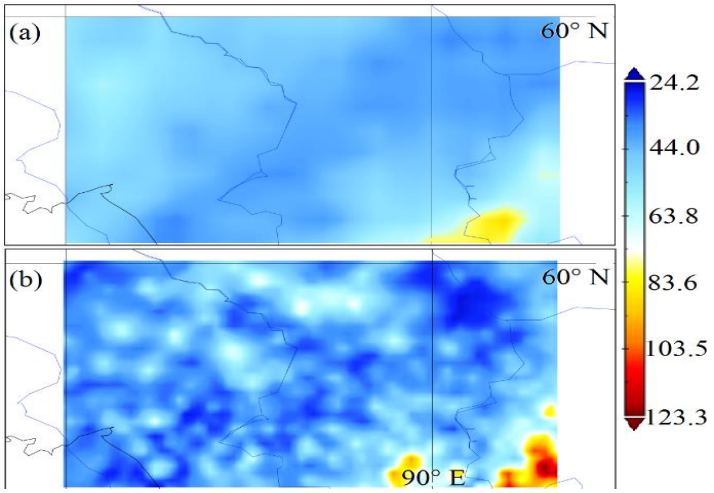


Рис.1. 100-летние уровни возврата максимальных осадков в июле для региона Южной Сибири:
*a* – данные ECMFW ERA Interim, горизонтальная сетка 0,75x0,75, 1979-2007 гг.;
*b* – данные JMA APHRODITE, горизонтальная сетка 0,25x0,25, 1979-2007 гг.

Анализ трендов в метеорологических наблюдениях является одним из наиболее распространенных подходов в исследованиях изменения климата. Квантильная регрессия обеспечивает четко определенную статистическую основу для оценки скорости изменения не только для среднего значения, как в обычной регрессии, но и в других частях функции распределения данных. Если задана случайная величина $Y$ с непрерывной кумулятивной функцией распределения $F_Y(y)$, квантильная функция $Q_Y(\tau)$ определяется из $F_Y(y)$ как $Q_Y(\tau) = F_Y^{-1}(\tau)$. Квантиль определяется как значение $Q_Y(\tau)$ такое, что $P[Y \leq Q_Y(\tau)] = \tau,\ 0 \leq \tau \leq 1$. Затем, учитывая условное распределение $Y$ с учетом $X = x$, условная квантильная функция $Q_{Y|X}(\tau; x)$ проверяет $P\left[Y \leq Q_{Y|X}(\tau; x)|X = x\right] = \tau$. В то время как обычная регрессия основана на условной функции среднего $E[Y|X = x]$ и минимизации соответствующих невязок, квантильная регрессия основана на условной функции квантиля и минимизации суммы асимметрично взвешенных абсолютных невязок $\sum_{i=1} \rho(\tau)|y_i - Q_{Y|X}(\tau; x = x_i)|$, где $\rho$ – функция абсолютного значения. Процедура расчета квантильной регрессии была реализована на языке R (пакет «quantreg») [15] и интегрирована в ВИС. Диапазон интересуемых квантильных значений задан между 0 и 1.

Для апробации новой функциональности ВИС, на основе данных реанализа ECMWF ERA 40 были рассчитаны тренды максимума температуры января для Южно-Сибирского региона (50-65 °N, 60-120 °E). Полученные результаты показывают (рис. 2), что максимум январской температуры на квантиле 0,05 изменился (как в сторону уменьшения, так и увеличения) в большей степени, чем на квантилях 0,5 и 0,95 почти на всей рассматриваемой территории. При этом на квантиле 0,95 температура изменилась в меньшей степени.

Расчет корреляции между двумя переменными основан на вычислении коэффициента корреляции Пирсона. В систему был интегрирован новый программный модуль для расчета коэффициента корреляции Пирсона для средних суточных, месячных, сезонных и годовых значений исследуемых метеорологических переменных. Предварительно рассматриваемые переменные сводятся на одну временную сетку, выбранную пользователем. Коэффициент корреляции может быть рассчитан для метеорологических переменных как из одного, так и из разных наборов данных. Если рассматриваемые переменные заданы на разных пространственных сетках, они интерполируются на сетку с более высоким пространственным разрешением. На рис. 3 показана специальная форма для расчета коэффициента корреляции между различными параметрами (температура воздуха на 2 м из наборов данных реанализа ERA Interim и метеорологических станций). В случае использования в расчете данных метеорологических станций и данных, заданных на регулярной сетке, выполняется интерполяция данных с регулярной сетки в координаты метеостанций.

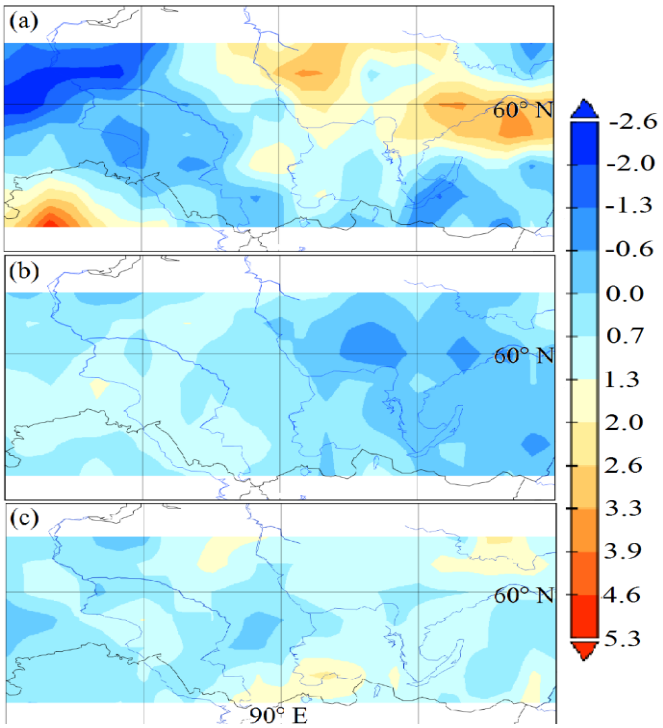



Рис. 2. Тренд максимума температуры января на основе данных реанализа ECMWF ERA 40 (горизонтальная сетка 2,5x2,5, 1961-2002 гг.): *a* – на квантиле 0,05; *b* – на квантиле 0,5; *c* – на квантиле 0,95.

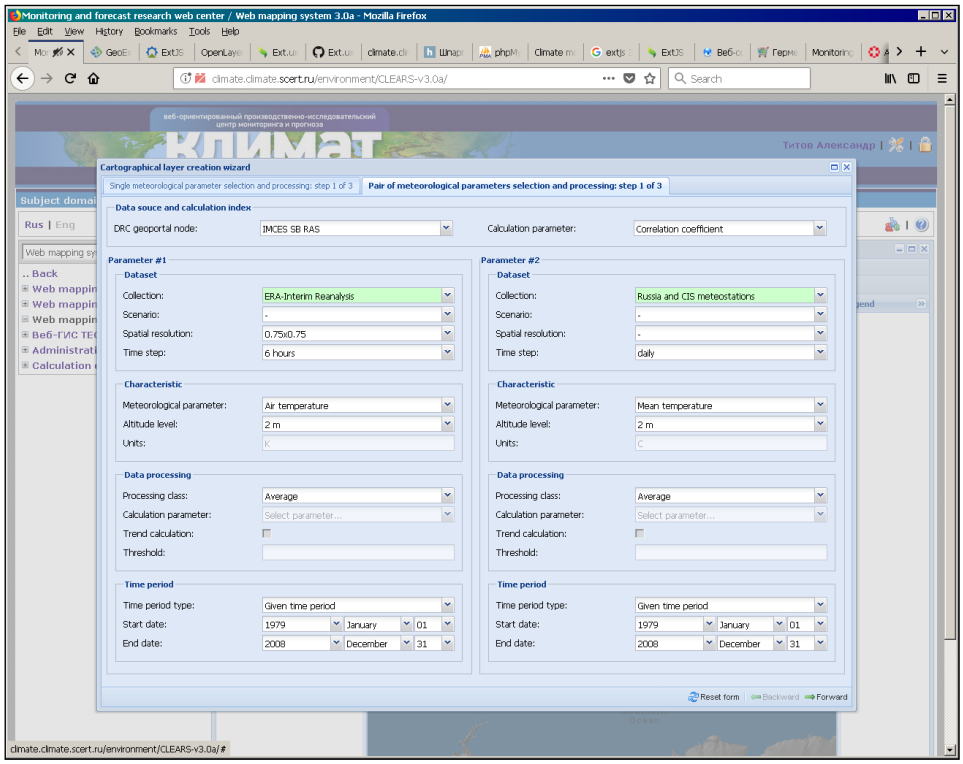



Рис. 3. Веб-ГИС клиент. Выбор пары метеопараметров и процедуры их совместного анализа.

**Расширенная функциональность веб-портала.** Среднегодовая температура на территории Российской Федерации продолжает расти. Современные оценки показывают, что экстремально высокие температуры и засухи летом, вызванные глобальным потеплением, нанесут серьезный ущерб в отдельных регионах России. В частности, они негативно повлияют на урожайность сельскохозяйственных культур и состояние лесного хозяйства. Ускорится таяние вечной мерзлоты на севере, и усилятся наводнения в некоторых регионах. Для разработки эффективной стратегии адаптации и мер по снижению негативных последствий экстремальных

климатических явления необходимы точные знания географии экстремальных климатических явлений, частоты их возникновения и интенсивности. Поскольку частота таких событий мала, для получения необходимой информации необходимо выполнять анализ этих явлений с использованием современного вероятностно-статистического аппарата и подробной метеорологической информации, накопленной за период инструментальных наблюдений на территории изучаемого региона.

Для предоставления региональным лицам, принимающим решения, информации, необходимой для их целевой деятельности, с использованием веб-ГИС «Климат» для сибирского региона был рассчитан набор соответствующих климатических характеристик. На его основе был создан открытый архив, который был интегрирован в ВИС для последующего анализа текущих изменений климата лицами, принимающими решения в регионе. Этот архив предоставляет количественные справочные материалы для оценки будущих климатических и экологических рисков и адаптации политики развития региона к этим рискам (рис. 4).

В архиве представлены различные характеристики температуры воздуха и количества осадков, рекомендованные ВМО (Всемирной метеорологической организацией) для анализа экстремальных климатических явлений. Эти характеристики предоставляют информацию о максимальных/минимальных значениях температуры и осадков, информацию о частоте и продолжительности различных экстремальных явлений, определяют количество дней, когда температура или осадки превышают некоторый порог (аномальные волна жары / холода, аномальные осадки и т.д.).

В архиве также есть ссылки для скачивания файлов в различных форматах (netCDF, GeoTIFF, WMS и WFS) с рассчитанными характеристиками. Если есть необходимость в дальнейшей работе с какой-либо характеристикой, пользователь может скачать требуемые файлы и загрузить их в стороннее программное обеспечение.

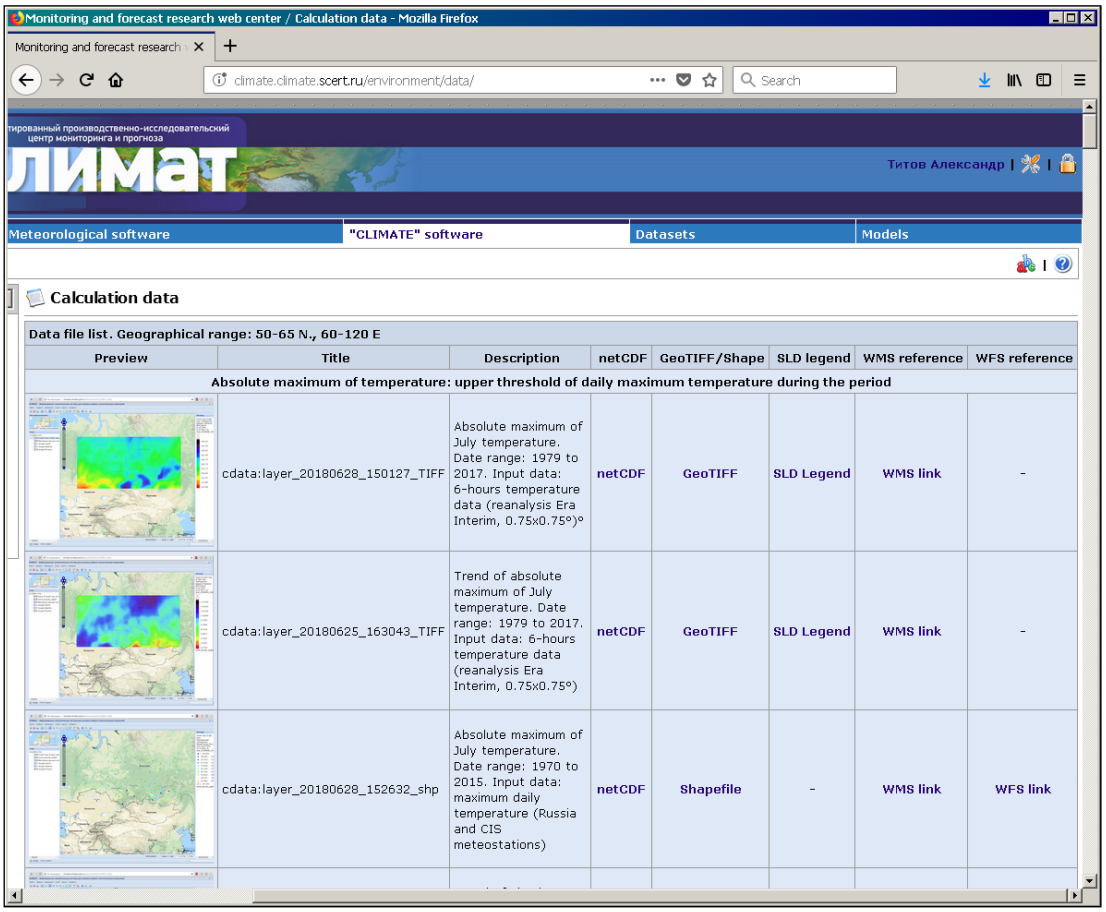



Рис.4. Веб-страница со ссылками на наборы соответствующих климатических характеристик, рассчитанных для сибирского региона.

**Заключение.** Задачи, решаемые в науках о Земле и, в частности – климатологии, как правило связаны с использованием больших архивов геопривязанных данных. Для их эффективного анализа и получения полезной информации требуется интеграция веб- и ГИС-технологий, данных, инструментов обработки и визуализации в распределенные, доступные через Интернет, тематические программные комплексы. В данной работе представлена разработка тематической ВИС для анализа метеорологических и климатических процессов. Она представляет собой бесплатный, кроссплатформенный, составной программный комплекс с открытым исходным кодом для «облачного» анализа климатических данных в окне интернет-браузера. Гибкая модульная архитектура ВИС позволяет легко добавлять новые вычислительные узлы, системы хранения данных, а также обеспечивает надежную вычислительную инфраструктуру для региональных исследований изменения климата на основе современных веб- и ГИС-технологий. Эта тематическая ВИС обеспечит междисциплинарные распределенные группы исследователей, не являющихся экспертами в области информационных технологий (климатологи, экологи, биологи и лица, принимающие решения), легкодоступными надежными онлайн-инструментами для надежного анализа и визуализации многомерных разнородных наборов климатических данных, полученных из различных источников. Возможность получения результатов анализа в форматах GeoTIFF, ESRI Shapefile, Encapsulated PostScript, CSV, XML, netCDF, float GeoTIFF открывает возможность для исследователей выполнять дальнейший анализ данных с использованием их собственного программного обеспечения.

## ЛИТЕРАТУРА


[1] *Llewellyn Smith C., Borysiewicz L., Casselton L., Conway G., Hassan M., Leach M.* Knowledge, Networks and Nations: Global Scientific Collaboration in the 21st Century. UK: The Royal Society, 2011.

[2] *Carusi A., Reimer T.* Virtual Research Environment Collaborative Landscape Study. JISC. 2010.

[3] *Candela L., Castelli D., Pagano P.* Virtual Research Environments: An Overview and a Research Agenda // Data Science Journal. 2013. V. 12. GRDI75-GRDI81. DOI: 10.2481/dsj.GRDI-013.

[4] IPCC. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge, UK: Cambridge University Press, 2012.

[5] IPCC. Fifth Assessment Report «Climate Change 2013». Cambridge, UK: Cambridge University Press, 2013.

[6] *Sillmann J., Donat M.G., Fyfe J.C., Zwiers F.W.* Observed and simulated temperature extremes during the recent warming hiatus // Environmental Research Letters. 2014. V. 9(6). P. 064023. DOI: 10.1088/1748-9326/9/6/064023.

[7] *Okladnikov I.G., Gordov E.P., Titov A.G., Shulgina T.M.* Information-computational System for Online Analysis of Georeferenced Climatological Data. // XVII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2015). Selected papers / CEUR Workshop Proceedings. 2015. Vol. 1536. P. 76-80.

[8] *Okladnikov I.G., Gordov E.P., Titov A.G.* Development of climate data storage and processing model // IOP Conference Series: Earth and Environmental Science. 2016. V. 48. P. 012030. DOI: 10.1088/1755-1315/48/1/012030.

[9] *Gordov E., Shiklomanov A., Okladnikov I., Prusevich A., Titov A.* Development of Distributed Research Center for analysis of regional climatic and environmental changes // IOP Conf. Series: Earth and Environmental Science. 2016. V. 48. P. 012033. DOI: 10.1088/1755-1315/48/1/012033.

[10] *Bart A., Fazliev A., Gordov E., Okladnikov I., Privezentsev A., Titov A.* Virtual Research Environment for Regional Climatic Processes Analysis: Ontological Approach to Spatial Data Systematization // Data Science Journal. 2018. V. 17. P. 14. DOI: 10.5334/dsj-2018-014.

[11] *Riazanova A.A., Voropay N.N., Okladnikov I.G., Gordov E.P.* Development of computational module of regional aridity for web-GIS «Climate» // IOP Conference Series: Earth and Environmental Science. 2016. V. 48. P. 012032. DOI: 10.1088/1755-1315/48/1/012032.

[12] *Ryazanova A.A., Voropay N.N.* Droughts and Excessive Moisture Events in Southern Siberia in the Late XXth – Early XXIst Centuries // IOP Conference Series: Earth and Environmental Science. 2017. V. 96. P. 012015. DOI: 10.1088/1755-1315/96/1/012015.

[13] *Shulgina T.M., Gordov E.P., Genina E.Yu.* Dynamics of climatic characteristics influencing vegetation in Siberia // Environmental Research Letters. 2011. V. 6. P. 045210.
[14] *Coles S.G.* An Introduction to Statistical Modelling of Extreme Values. London: Springer, 2001.
[15] *Koenker R., Portnoy S., Tian P., Zeileis A., Grosjean P., Ripley B.D.* Package «quantreg». The Comprehensive R Archive Network (CRAN), 2017. **https://cran.r-project.org/web/packages/quantreg/quantreg.pdf** (дата обращения 17.01.2019).

# UNISAT. ТЕХНОЛОГИЯ ПОСТРОЕНИЯ УНИФИЦИРОВАННЫХ СИСТЕМ ВЕДЕНИЯ СВЕРХБОЛЬШИХ РАСПРЕДЕЛЕННЫХ АРХИВОВ РАЗНОРОДНЫХ СПУТНИКОВЫХ ДАННЫХ

*Прошин А.А., Лупян Е.А., Кашницкий А.В., Балашов И.В.*
Институт космических исследований РАН, Москва

Ускоренное развитие систем спутникового дистанционного зондирования Земли привело к существенному возрастанию требований, предъявляемых к системам ведения архивов спутниковых данных. Статья посвящена описанию технологии UNISAT, предназначенной для построения систем ведения сверхбольших распределенных архивов разнородных спутниковых данных, обеспечивающей динамическое формирование данных по запросу пользователя и работу широкого спектра инструментов для удаленного анализа и обработки данных.
*Ключевые слова: дистанционное зондирование Земли, архивы спутниковых данных, система архивации спутниковых данных, информационная система, сверхбольшие массивы данных.*

**Введение.** Ускоренное развитие систем спутникового дистанционного зондирования Земли в последние десятилетия привело к взрывному росту объемов спутниковых данных, получаемых с большого числа различных приборов наблюдения [1]. Расширяется область применения данных дистанционного зондирования Земли, которые в настоящее время широко используются для решения самых разных исследовательских и прикладных задач, связанных с мониторингом природной среды и антропогенных объектов. Все это, в свою очередь, приводит к существенному возрастанию требований, предъявляемых к системам, обеспечивающим работу со спутниковой информацией, и в частности, к системам ведения архивов спутниковых данных, на которых базируется их реализация.

Одним из основных требований к современным системам ведения архивов спутниковых данных является поддержка работы с разнородными спутниковыми данными, получаемыми приборами наблюдения с разными техническими характеристиками (частота наблюдения, пространственное разрешение, повторяемость наблюдений и т.д.). Это приводит к необходимости унификации процедур архивации данных, выработки единой структуры баз данных, и реализации единых программных интерфейсов для доступа к разным типам спутниковой информации. Еще одним важным требованием является поддержка сверхбольших распределенных архивов спутниковых данных, которая, позволяет объединить архивы различных центров получения и хранения спутниковых данных в единый информационный ресурс. Таким образом, пользователям уже неважно, где именно физически располагаются те или иные данные.

За последние годы радикально поменялись и требования по обеспечению доступа к информации. Если раньше потребители спутниковых данных в основном довольствовались получением исходных данных для использования их в своих системах обработки и анализа, то сейчас их все больше интересуют возможности работы с готовыми информационными продуктами различного уровня обработки [2], причем число таких информационных продуктов, необходимых для решения различных задач, постоянно растет. Следует отметить, что, так как спутниковые данные, как правило, имеют значительный объем, хранение всех возможных информационных продуктов, получаемых на основе одних и тех же исходных данных, становится нецелесообразным и, во многих случаях, технически невозможным. Естественным выходом из этой ситуации является предоставление пользователям доступа к «виртуальным» информационным продуктам, т.е. продуктам, которые динамически формируются на основе исходных данных в режиме реального времени. Ключевым преимуществом такого подхода является возможность расширения списка доступных пользователям информационных продуктов без необходимости проведения массовой обработки данных в архивах.

Одним из наиболее актуальных направлений развития современных систем доступа к спутниковым данным, на наш взгляд, является создание различных инструментов для работы

со спутниковыми данными и их анализа, которые ранее были доступны только в специализированных настольных приложениях. Реализуемые посредством таких интерфейсов инструменты для анализа данных позволяют обрабатывать большие объемы имеющейся спутниковой информации, используя при этом мощности центров предоставления данных. Наиболее яркими представителями подобных систем, на наш взгляд, являются Google Earth Engine [3] (https://earthengine.google.org) и система «Вега-Science» [4], реализованная в рамках ЦКП «ИКИ-Мониторинг» [5].

Настоящая статья посвящена краткому описанию ключевых особенностей разработанной в ИКИ РАН технологии UNISAT [6], предназначенной для построения унифицированных систем ведения сверхбольших распределенных архивов разнородных спутниковых данных. Она в свою очередь базируется на использовании технологий и программного обеспечения, которые были разработаны в ИКИ РАН за последние десятилетия для построения широкого спектра различных архивов спутниковых данных [7, 8]. Технология UNISAT позволяет однотипно реализовать работу с самыми разными типами спутниковой информации, отличающейся как пространственным разрешением, так и схемой организации хранения данных. Важным преимуществом разработанной технологии является поддержка механизма «виртуальных информационных продуктов», т.е. продуктов, которые динамически формируются по запросу пользователя на основе обработки имеющихся в архивах спутниковых данных. Большое внимание при создании этой технологии было также уделено поддержке работы широкого спектра различных инструментов для удаленного анализа и обработки спутниковых данных.

**Общая архитектура.** На основе многолетнего опыта создания и эксплуатации различных систем доступа к спутниковым данным нами была разработана общая архитектура построения систем ведения архивов спутниковых данных. Ключевыми ее элементами является новая унифицированная структура базы данных, предназначенной для хранения разнородных спутниковых данных, а также структура справочной базы данных, содержащей детальную информацию о спутниках, приборах наблюдения, реализованных продуктах, а также правилах построения «виртуальных» информационных продуктов.

Общая архитектура узла распределенной системы ведения архивов UNISAT приведена на рис. 1. Данные могут поступать в архивы, как от подсистемы обработки спутниковых данных, так и от внешних поставщиков спутниковых данных. В левой части схемы изображены программные компоненты, отвечающие за архивацию спутниковых данных и обмен данными с другими информационными центрами. В средней части схемы приведена справочная БД unisat_catalog, БД unisat, содержащая метаданные имеющихся в архиве данных, и связанное с ней файловое хранилище, в котором непосредственно находятся файлы спутниковых данных. В правой части схемы изображены основные сервисы, предназначенные для обеспечения доступа к данным в архиве. Картографический WEB интерфейс является основным инструментом для предоставления доступа пользователям, как к самим спутниковым данным, так и к сервисам, предназначенным для работы с ними. На схеме пунктирными стрелками показаны запросы на получение данных или метаданных, сплошными стрелками – метаданные, а полыми стрелками – данные. Внешние компоненты, непосредственно не входящие в состав системы ведения архивов, показаны пунктирными фигурами.

На рис. 2 приведена принципиальная схема построения распределенного архива спутниковых данных, реализованного по представленной технологии. На схеме указаны основные потоки данных и метаданных, реализуемые в рамках распределенного архива. Данные могут поступать в архивы как из внешних центров данных, так и с локальных станций приема. Для каждого информационного центра может быть индивидуально определена политика экспорта данных или метаданных в остальные центры распределенного архива, но в самом простом варианте в каждом центре содержится вся информация об имеющихся в распределенном архиве данных. В каждом центре может быть реализован свой набор различных архивов спутниковых данных, но при этом обязательно содержится синхронизируемая с центрального сервера справочная БД unisat_catalog. Основным преимуществом представленной реализации распределенного архива спутниковых данных является высокая степень гибкости в определении того,

какие типы метаданных и данных должны передаваться между информационными центрами, входящими в состав распределенного архива, что практически недостижимо при использовании стандартных инструментов репликации баз данных.

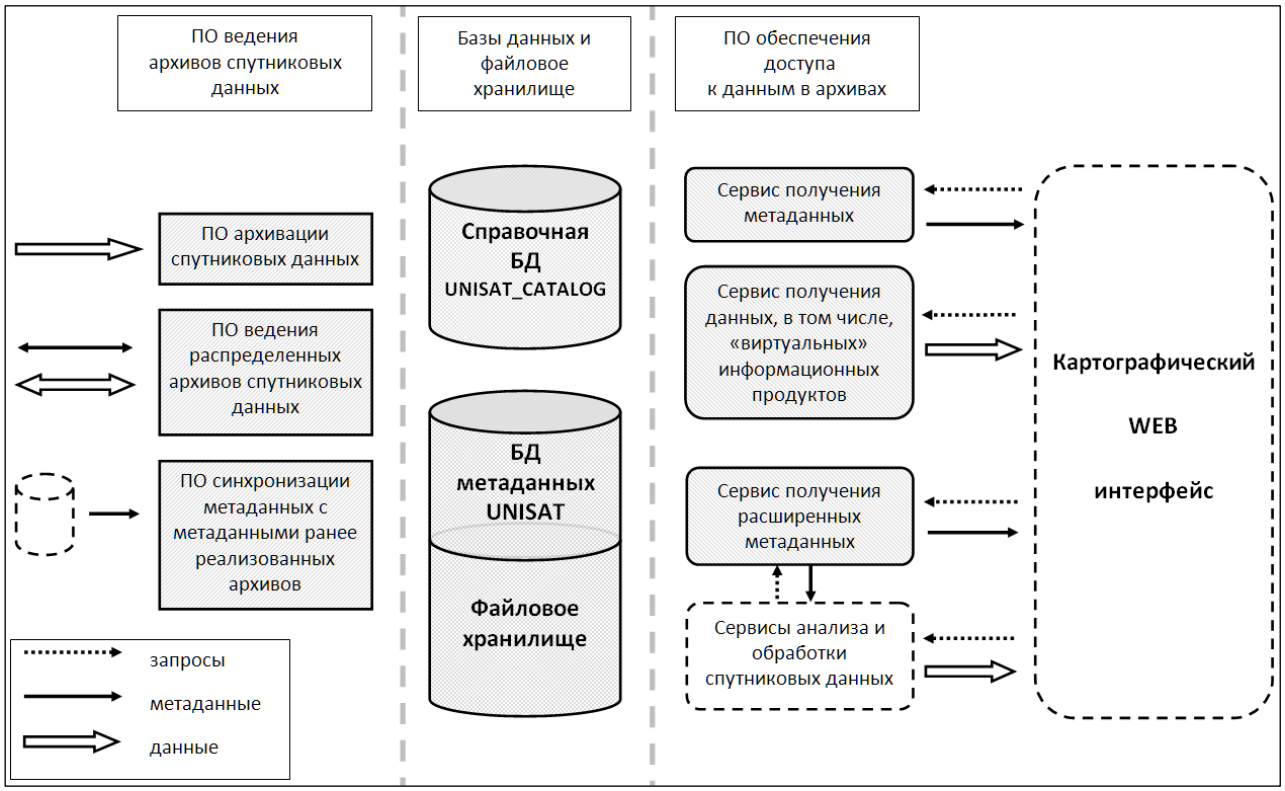



Рис. 1. Общая архитектура построения узла распределенного архива.

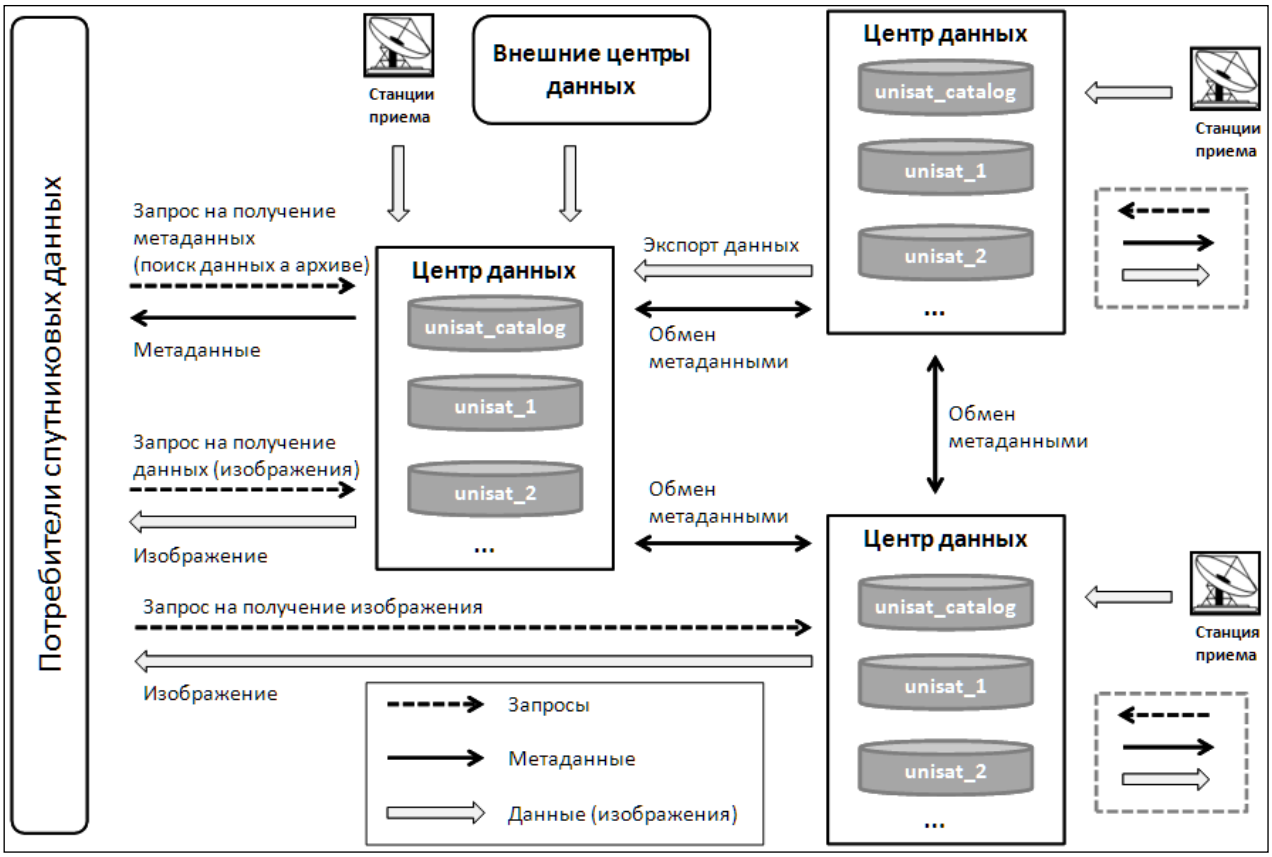



Рис. 2. Принципиальная схема построения распределенных архивов данных.

**Унифицированная структура БД UNISAT для хранения спутниковых данных.** Для описания унифицированной структуры БД введем несколько обозначений. Сеанс – это набор данных, однозначно идентифицируемый полями: дата и время (**dt**), спутник (**satellite**), станция (**station**), прибор (**device**). Фрагмент – это пространственная часть сеанса. При этом сеанс может состоять как из одного фрагмента, так и из множества фрагментов. БД **UNISAT** включает в себя две основные таблицы для хранения метаданных: таблица **fragments** предназначена для описания фрагментов, а таблица **fragment_products** – для описания информационных продуктов, относящихся к соответствующим фрагментам (см. рис. 3). На рисунке первичные ключи таблиц выделены полужирным шрифтом, а поля уникального ключа – курсивом.

В случае, когда сеанс состоит из одного фрагмента, в таблицу **fragments** прописывается одна запись с указанием типа фрагмента «**single_fragment**». Если же сеанс состоит из множества фрагментов, то каждому из них будет соответствовать отдельная запись с типом фрагмента «**fragment**» и отличающимся номером фрагмента, но при этом также автоматически генерируются фрагменты типа «**products_contour**», описывающие интегральные контура продуктов или групп продуктов по сеансу. Поле **scale_level** в таблице **fragment_products** используется для идентификации прореженных версий информационных продуктов, используемых для ускорения отображения изображений в масштабах, больших, чем разрешения самих спутниковых данных. Существенно, что одни и те же продукты могут иметь различное пространственное разбиение в разных масштабах.

Разработанная структура позволила эффективно работать с данными, представленными как в виде отдельных фрагментов, так и с продуктами, состоящими из многих тысяч отдельных фрагментов. К преимуществам описываемой структуры также относится гибкая поддержка дополнительных пространственных разбиений в прореженных масштабах, позволяющая максимально ускорить формирование требуемых изображений в заданном масштабе.

**Структура единой справочной БД unisat_catalog.** БД предназначена для ведения всей необходимой справочной информации о станциях приема, спутниках, характеристиках спутниковых приборов и др., а также содержит информацию о типах продуктов, хранящихся в архивах и правилах получения на их основе «виртуальных» информационных продуктов. Ниже в таблице приводится состав и назначение основных таблиц БД unisat_catalog с указанием типа хранимой в них справочной информации. Ключевым отличием предложенной структуры «справочной» базы данных является интеграция общей информации о спутниках, приборах и соответствующих им типах данных с информацией, необходимой для получения «виртуальных» информационных продуктов. При этом реализованные сервисы получения расширенных метаданных позволяют предоставлять информацию, необходимую для реализации инструментов для удаленного анализа и обработки спутниковых данных.

Таблица **fragments**

| fragment_id | Идентификатор фрагмента |
|---|---|
| dt | Дата и время фрагмента |
| satellite | Дата и время фрагмента |
| device | Прибор |
| station | Станция приема |
| fragment_type | Тип фрагмента:<br>single_fragment<br>fragment<br>products_contour |
| fragment_number | Номер фрагмента |
| owner | Центр-владелец продукта |
| revolution | Виток |
| min_lon100 | минимальная долгота * 100 |
| max_lon100 | максимальная долгота * 100 |
| min_lat100 | минимальная широта * 100 |
| max_lat100 | максимальная широта * 100 |
| contour | контур фрагмента |
| source_info | Информация об исходных данных |
| cloudiness | Процент облачности |
| corrected | Признак коррекции |
| daynight | Время суток |
| extra_parameters | Параметры в формате JSON |
| mtime | Время модификации записи |

Таблица **fragment_products**

| ID | FDB. Идентификатор продукта по фрагменту |
|---|---|
| fragment_id | Идентификатор фрагмента |
| product_type_id | Идентификатор типа продукта |
| scale_level | Уровень разрешения (0 - базовый) |
| product_NAME | FDB. Имя файла с изображением |
| product_SIZE | FDB. Размер файла |
| product_VOLUME | FDB. Том хранения |
| product_STORAGEFILE | FDB. Файл в хранилище |
| product_MD5 | FDB. Контрольная сумма файла |
| hist_NAME | FDB. Имя файла с «гистограммой» |
| hist_SIZE | FDB. Размер файла |
| hist_VOLUME | FDB. Том хранения |
| hist_STORAGEFILE | FDB. Файл в хранилище |
| hist_MD5 | FDB. Контрольная сумма файла |
| mtime | Время модификации записи |

Рис. 3. Структура основных таблиц БД UNISAT.

Состав и назначение таблиц БД unisat_catalog.

| **Тип информации** | **Таблица** | **Назначение таблицы** |
|---|---|---|
| Общая справочная информация | satellite | информация о спутниках |
| | satellite_device | информация о приборах, установленных на спутниках |
| | device | информация о спутниковых приборах |
| | band | информация о каналах спутниковых приборов |
| | station | информация о станциях приема спутниковых данных |
| | center | информация об информационных центрах |
| Описание типов продуктов, хранящихся в архивах | product | описание типов информационных продуктов |
| | product_cases | информация о типах продуктов, построенных на основе данных разных спутниковых приборов |
| | product_level | информация о «прореженных» масштабах продуктов |
| | channel | информация о каналах информационных продуктов |
| Правила получения «виртуальных» продуктов на основе обработки продуктов, имеющихся в архиве | vproduct | описание типов виртуальных продуктов |
| | vproduct_cases | варианты реализации виртуальных продуктов в зависимости от типов спутниковых приборов и имеющихся в архиве данных |
| | vchannel | правила получения каналов виртуальных продуктов |

**Заключение.** В настоящее время технология UNISAT успешно используется в реализации целого ряда информационных систем доступа к спутниковым данным, разрабатываемых ИКИ РАН в сотрудничестве с другими организациями. Среди них особо стоит отметить центр коллективного пользования спутниковыми данными ЦКП «ИКИ-Мониторинг» [5], обеспечивающий непосредственный доступ к спутниковым данным более чем 30 различных приборов наблюдения Земли, при этом суммарный объем доступных данных в настоящее время существенно превышает 2 петабайта. Эта же технология используется для ведения архивов единой системы работы со спутниковыми данными ФГБУ «НИЦ «Планета» [9], в которых содержатся данные как российских, так и зарубежных спутниковых систем дистанционного зондирования Земли. В перспективе планируется обеспечить эффективную поддержку более широкого набора инструментов для удаленного анализа и обработки спутниковых данных.

*Базовая функциональность созданной системы реализована при поддержке ФАНО (тема «Мониторинг», госрегистрация № 01.20.0.2.00164); элементы, связанные с поддержкой функций работы с виртуальными продуктами, реализованы при поддержке РФФИ (грант № 16-37-00427-мол_а); функции, обеспечивающие поддержку взаимодействия узлов сверхбольших архивов спутниковых данных, отрабатывались с использованием распределенных архивов спутниковых данных в центрах НИЦ «Планета» при поддержке РФФИ (грант № 15-29-07953-офи_м). С 2019 года работы по совершенствованию представленных технологий и программного обеспечения ведутся также при поддержке ФАНО (тема «Большие данные в космических исследованиях: астрофизика, солнечная система, геосфера», госрегистрация № 0024-2019-0014).*

## ЛИТЕРАТУРА

[1] *Лупян Е.А., Бурцев М.А., Прошин А.А., Кобец Д.А.* Развитие подходов к построению информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2018. Т. 15. № 3. С. 53-66. DOI: 10.21046/2070-7401-2018-15-3-53-66.

[2] *Лупян Е.А., Саворский В.П.* Базовые продукты обработки данных дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 9. № 2. С. 87-97.

[3] *Moore R.T., Hansen M.C.* Google Earth Engine: a new cloud-computing platform for global-scale earth observation data and analysis // AGU Fall Meeting Abstracts, 2011. Vol. 1. P. 2.

[4] *Толпин В.А., Лупян Е.А., Барталев С.А., Плотников Д.Е., Матвеев А.М.* Возможности анализа состояния сельскохозяйственной растительности с использованием спутникового сервиса «ВЕГА» // Оптика атмосферы и океана. 2014. Т. 27. № 7 (306). С. 581-586.

[5] *Лупян Е.А., Прошин А.А., Бурцев М.А., Балашов И.В., Барталев С.А., Ефремов В.Ю., Кашницкий А.В., Мазуров А.А., Матвеев А.М., Суднева О.А., Сычугов И.Г., Толпин В.А., Уваров И.А.* Центр коллективного пользования системами архивации, обработки и анализа спутниковых данных ИКИ РАН для решения задач изучения и мониторинга окружающей среды // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 5. С. 263-284.

[6] *Прошин А.А., Лупян Е.А., Балашов И.В., Кашницкий А.В., Бурцев М.А.* Создание унифицированной системы ведения архивов спутниковых данных, предназначенной для построения современных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13. № 3. С. 9-27. DOI: 10.21046/2070-7401-2016-13-3-9-27.

[7] *Лупян Е.А., Балашов И.В., Бурцев М.А., Ефремов В.Ю., Кашницкий А.В., Кобец Д.А., Крашенинникова Ю.С., Мазуров А.А., Назиров Р.Р., Прошин А.А., Сычугов И.Г., Толпин В.А., Уваров И.А., Флитман Е.В.* Создание технологий построения информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 5. С. 53-75.

[8] *Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В.* Технология построения автоматизированных систем хранения спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса. 2004. Выпуск 1. Т.1. С. 437-443.

[9] *Лупян Е.А., Милехин О.Е., Антонов В.Н., Крамарева Л.С., Бурцев М.А., Балашов И.В., Толпин В.А., Соловьев В.И.* Система работы с объединенными информационными ресурсами, получаемыми на основе спутниковых данных в центрах НИЦ «Планета» // Метеорология и гидрология. 2014. № 12. С. 89-97.

# ПОСТРОЕНИЕ ЗАВИСИМОСТИ РАЗМЕРА КРОНЫ СОСНЫ ОБЫКНОВЕННОЙ ОТ ДИАМЕТРА СТВОЛА

*Пятаев А.С.[(1)(2)], Вайс А.А.[(1)]*
[(1)] Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, Красноярск
[(2)] Филиал ФБУ «Рослесозащита» – «ЦЗЛ Красноярского края», Красноярск

В настоящей работе предложен способ оценки диаметров кроны по диаметрам стволов сосны обыкновенной в высокополнотных насаждения. В таких условиях методы компьютерного зрения не в состоянии четко выделить границы кроны исследуемого дерева по причине существенного взаимного проникновения крон деревьев. Для этого построены аппроксимирующие функции зависимости диаметров крон от ступеней толщины ствола дерева.
*Ключевые слова: крона дерева, аппроксимация, сосна обыкновенная.*

**Введение.** Крона дерева играет ключевую роль при определении различных характеристик лесных насаждений. Такие параметры кроны как густота и цвет характеризует санитарное состояние дерева, плотность кроны влияет на видовую конкуренцию между деревьями [1-3], густота полога оказывает влияние на потенциал возникновения и развития верховых пожаров [4], также характеристики кроны влияют на качество древесины [5]. Крона дерева формируется под воздействием различных показателей и факторов, таких как сомкнутость кроны насаждения, освещенность [6], скорость и сила ветра [7, 8] и т.д.

Одним из вариантов решения задачи автоматической оценки кроны является анализ визуальных данных, например, при определении категории санитарного состояния дерева. С помощью методов компьютерного зрения зачастую трудно отделить границы кроны конкретного дерева в высокополнотных древостоях, так как кроны деревьев существенно проникают друг в друга. На данный момент оценить крону дерева и сделать выводы можно только эмпирически с использованием человеческого ресурса, поэтому задача автоматической оценки формы кроны является актуальной. Оценку характерных диаметров проекций крон, для определенных условий мест произрастаний лесных участков, можно выполнить по статическим данным в зависимости от диаметра ствола [9]. Таким образом, оценка зависимости диаметров кроны от диаметров стволов, характерная для определенной полноты насаждения, поможет повысить достоверность определения границ кроны с помощью методов компьютерного зрения по изображениям в насаждениях с высокой полнотой.

**Построение зависимости диаметра кроны от диаметра ствола.** В работе [10] для определения категории состояния дерева предложен подход на базе нечеткой логики. Такой подход позволяет формализовать субъективные представления специалиста, на основе которых и происходит принятие решений. При этом в качестве лингвистических переменных выступили следующие: степень плотности кроны, прирост, степень усыхания ветвей, отпад коры и цвет хвои. В работе [11] проведена сплайн-аппроксимация кусочно-линейных функций принадлежности с использованием функций Гаусса. Особую трудность при работе специалиста в реальных условиях составляет определение состояния кроны, так как не всегда визуально можно определить границы кроны у конкретного исследуемого дерева из-за высокой сомкнутости полога леса (рис. 1). Кроме того, для одной категории состояния у деревьев по степени густоты кроны возможна большая вариативность внутри класса: разница в плотности кроны может составлять до 30%. Поэтому для оценки густоты и формы кроны дерева в настоящей работе предложено использовать методы оценки зависимости размеров кроны дерева от диаметра его ствола. Это позволит увеличить точность дифференциации живых деревьев по категориям состояния, тем более, что измерение диаметров стволов является одним из обязательных требований при проведении таксации лесов или лесопатологических обследований.



Рис. 1. Сосняк осочково-разнотравный.

Относительная близость произрастания деревьев влияет на взаимное проникновение крон, что в свою очередь влияет на их диаметр, он будет отличен от среднего диаметра крон отдельно стоящих деревьев. Сотрудниками кафедры лесной таксации, лесоустройства и геодезии СибГУ были проведены измерения диаметров крон и стволов в лесных насаждениях с высокой полнотой в Курагинском районе Красноярского края. Получены данные для 400 деревьев сосны обыкновенной разных диаметров, возраста и санитарного состояния.

С помощью статистического анализа данных проведенных замеров были выведены средние размеры крон, сгруппированные в зависимости от диаметров по ступеням толщины стволов. Кроме того, были рассчитаны статистические показатели, такие как стандартное отклонение и стандартная ошибка, показывающие меру различия между средними диаметрами крон и диаметрами крон выборки. В табл. 1 представлены рассчитанные статистические показатели, сгруппированные по диаметрам стволов, а также количество деревьев, участвующих при расчетах.

На рис. 2 показаны исходные и усредненные статистические данные. Функция, построенная на усредненных статистических данных и описывающая зависимость диаметра кроны от диаметра ствола, является кусочно-линейной.

Табл. 1. Распределение средних диаметров крон по ступеням толщины ствола.

| **Ступень толщины ствола, см** | **Средний диаметр кроны, м** | **Стандартное отклонение** | **Стандартная ошибка** | **Количество деревьев** |
|---|---|---|---|---|
| 4 | 1,08 | 0,24 | 0,08 | 10 |
| 8 | 1,44 | 0,48 | 0,12 | 17 |
| 12 | 1,88 | 0,64 | 0,1 | 39 |
| 16 | 2,95 | 1,48 | 0,26 | 32 |
| 20 | 3,44 | 0,9 | 0,22 | 17 |
| 24 | 4,16 | 1,58 | 0,25 | 41 |
| 28 | 4,45 | 1,78 | 0,32 | 30 |
| 32 | 4,64 | 1,96 | 0,41 | 23 |
| 36 | 4,86 | 2,04 | 0,33 | 38 |
| 40 | 6,02 | 2,13 | 0,39 | 30 |
| 44 | 5,95 | 2,28 | 0,72 | 10 |
| 56 | 6,38 | 1,03 | 0,52 | 4 |

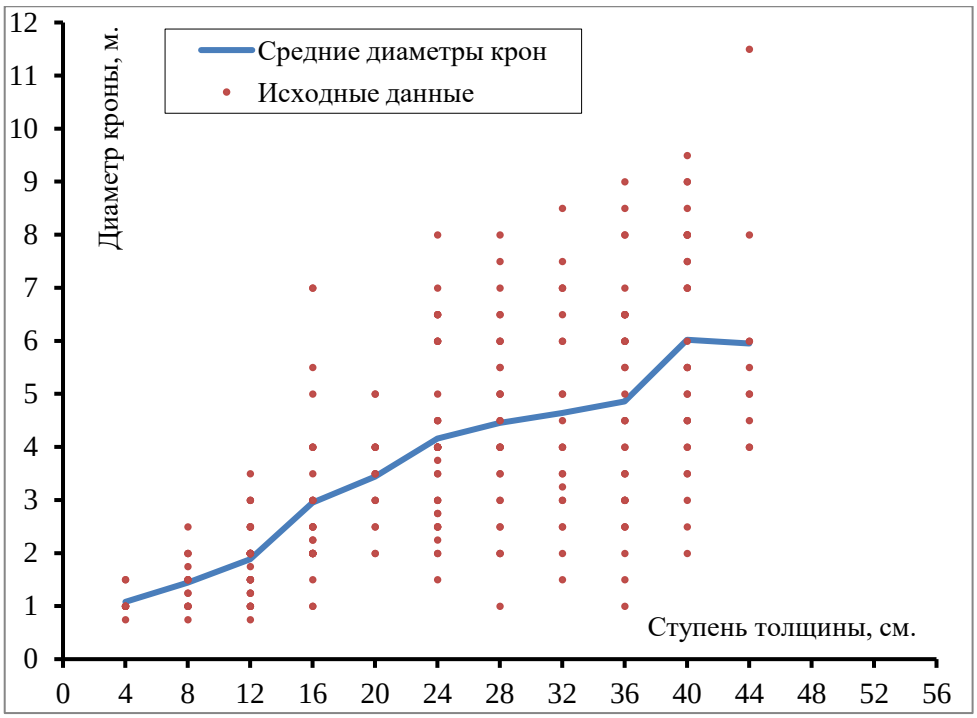


Рис. 2. Регрессионная линия связи диаметра кроны и ступени толщины.

**Аппроксимация функциями Гаусса.** Для удобства расчетов построенная кусочно-линейная функция была аппроксимирована функциями Гаусса. Было рассмотрено два подхода: аппроксимация одной функцией и аппроксимация несколькими функциями Гаусса. Первый подход представлен формулой (1).

$$g(x) = 11e^{-\frac{(x-80)^2}{2380}} + 0.5, x \leq 52 \quad (1)$$

Сравнительные графики исходной и аппроксимированной функций представлены на рис. 3. Несмотря на визуальное различие, расчеты показывают, что средняя ошибка аппроксимации составляет 13%, что укладывается в допустимые пределы.

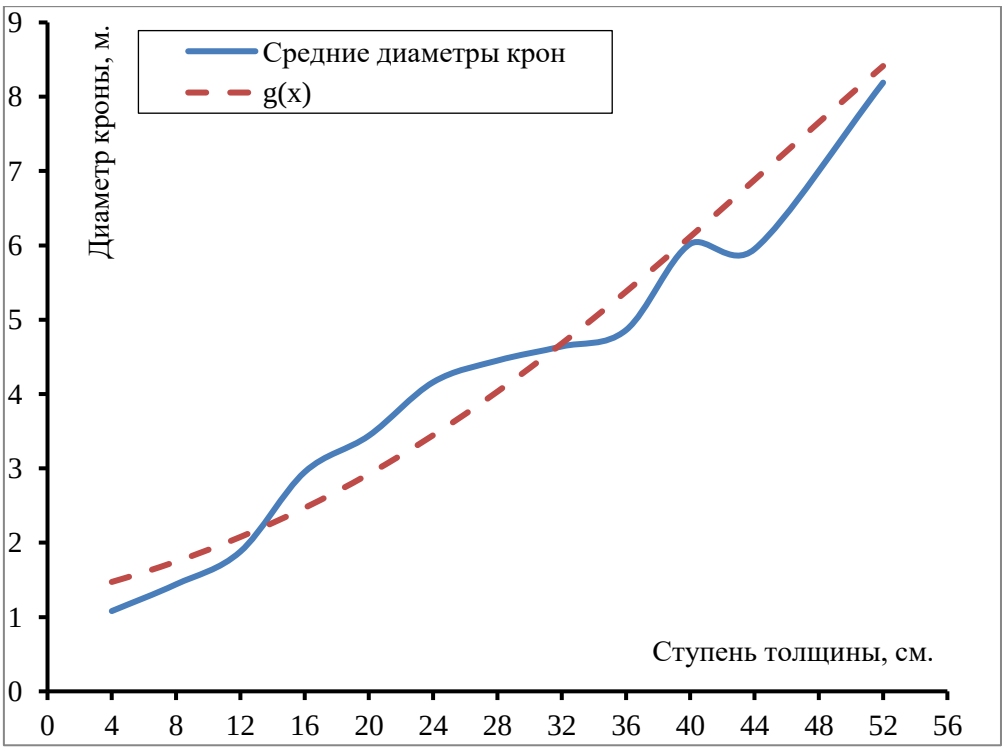


Рис. 3. Аппроксимация функции средних диаметров крон функцией Гаусса.

Табл. 2. Ошибка аппроксимации при использовании одной функции.

| Ступень толщины ствола, см | Средний диаметр, м | g(x) | Ошибка, % | Средняя ошибка, % |
|---|---|---|---|---|
| 4 | 1,08 | 1,47 | 36,29 | 13,11 |
| 8 | 1,44 | 1,75 | 21,27 | |
| 12 | 1,88 | 2,08 | 10,47 | |
| 16 | 2,95 | 2,47 | 16,32 | |
| 20 | 3,44 | 2,92 | 14,99 | |
| 24 | 4,16 | 3,45 | 17,16 | |
| 28 | 4,45 | 4,03 | 9,38 | |
| 32 | 4,64 | 4,68 | 0,84 | |
| 36 | 4,86 | 5,38 | 10,65 | |
| 40 | 6,02 | 6,12 | 1,61 | |
| 44 | 5,95 | 6,88 | 15,67 | |
| 52 | 8,19 | 8,41 | 2,73 | |

Второй вариант аппроксимации несколькими функциями Гаусса представлен формулой (2).

$$g(x) = \begin{cases} 3{,}9e^{-\frac{(x-30)^2}{327{,}68}} + 0{,}6, x \le 28 \\ 3.7e^{-\frac{(x-52)^2}{112{,}5}} + 4{,}5, x \in (28{,}52] \end{cases} \quad (2)$$

Графики этих функций представлены на рис. 4. На графиках видно, что расхождение значений исходной и аппроксимированных функций незначительно. Что подтверждается расчетами ошибок аппроксимации, которые продемонстрированы в табл. 3. Значение средней ошибки свидетельствует о хорошо подобранных аппроксимирующих функциях.

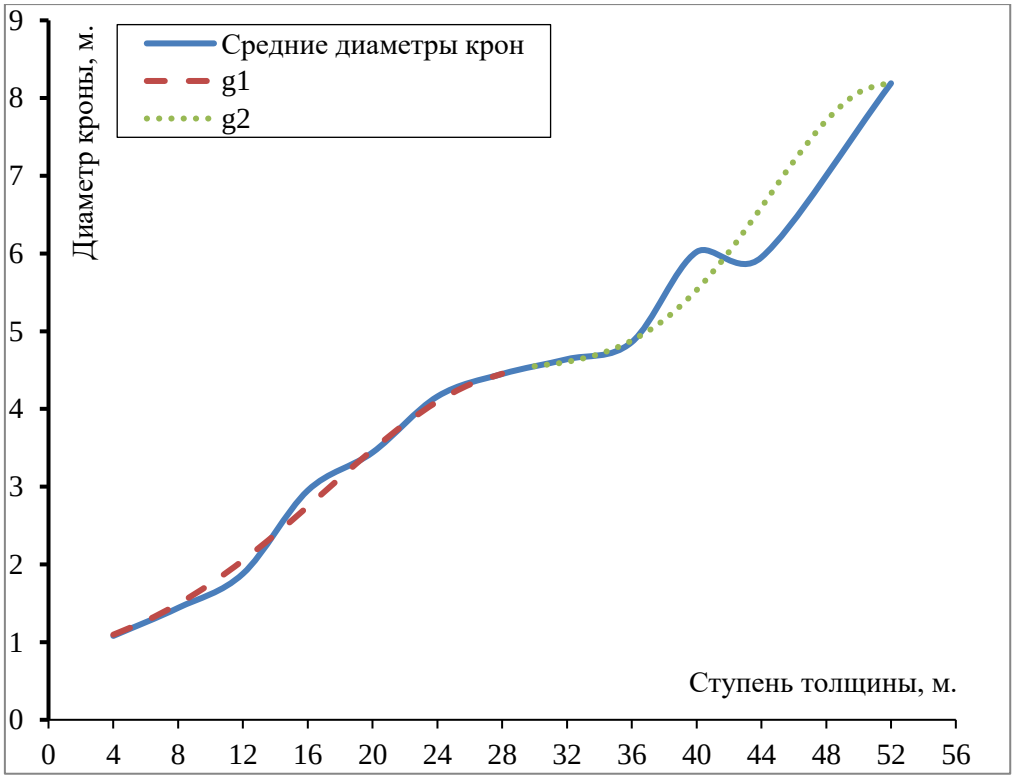


Рис. 4. Аппроксимация функции средних диаметров крон двумя функциями Гаусса.

Табл. 3. Ошибка аппроксимации при использовании двух функций.

| Ступень толщины ствола, см | Средний диаметр, м | g(x) | Ошибка, % | Средняя ошибка, % |
|---|---|---|---|---|
| 4 | 1,08 | 1,1 | 1,44 | 3,38 |
| 8 | 1,44 | 1,49 | 3,5 | |
| 12 | 1,88 | 2,05 | 9,09 | |
| 16 | 2,95 | 2,74 | 6,97 | |
| 20 | 3,44 | 3,47 | 1 | |
| 24 | 4,16 | 4,09 | 1,58 | |
| 28 | 4,45 | 4,45 | 0,06 | |
| 32 | 4,64 | 4,61 | 0,74 | 4,05 |
| 36 | 4,86 | 4,88 | 0,41 | |
| 40 | 6,02 | 5,53 | 8,16 | |
| 44 | 5,95 | 6,59 | 10,84 | |
| 52 | 8,19 | 8,2 | 0,12 | |

Таким образом, использование одной функции при аппроксимации зависимости средних диаметров крон от диаметров стволов показывает среднюю ошибку аппроксимации в 13%, а использование двух функций дает в среднем ошибку в 4%.

**Заключение.** При использовании методов компьютерного зрения при исследовании высокополнотных древостоев возникают трудности с определением границ кроны, поэтому предложенные в настоящей работе подходы могут являться корректирующим фактором параметров крон в случае спорных ситуаций.

## ЛИТЕРАТУРА


[1] *Russell D. Kramer, Stephen C. Sillett, Robert Van Pelt, Jerry F. Franklin* Neighborhood competition mediates crown development of Picea sitchensis in Olympic rainforests: Implications for restoration management // Forest Ecology and Management. 2019. Vol. 441 P. 127-143.

[2] *Stephan Getzin, Kerstin Wiegand, Jens Schumacher, Francois A.Gougeon* Scale-dependent competition at the stand level assessed from crown areas // Forest Ecology and Management. 2008. Vol. 255 P. 2478-2485.

[3] *Noriyuki Osada, Ryunosuke Tateno, Fujio Hyodo, Hiroshi Takeda* Changes in crown architecture with tree height in two deciduous tree species: developmental constraints or plastic response to the competition for light? // Forest Ecology and Management. 2004. Vol. 255 P. 337-347.

[4] *Jonathan A.Thomas, Joseph D.White, Darrel B.Murray* Tree species influence woodland canopy characteristics and crown fire potential // Forest Ecology and Management. 2016. Vol. 362 P. 169-176.

[5] *Luka Krajnc, Niall Farrelly, Annette M.Harte* The influence of crown and stem characteristics on timber quality in softwoods // Forest Ecology and Management. 2019. Vol. 435 P. 8-17.

[6] *Ethan J.Coonen, Stephen C.Sillett* Separating effects of crown structure and competition for light on trunk growth of Sequoia sempervirens // Forest Ecology and Management. 2015. Vol. 358 P. 26-40.

[7] *Shawn X.Meng, Shongming Huang, Victor J.Lieffers, Thompson Nunifu, Yuqing Yang* Wind speed and crown class influence the height–diameter relationship of lodgepole pine: Nonlinear mixed effects modeling // Forest Ecology and Management. 2008. Vol. 256 P. 570-577.

[8] *Vincent A. Webb, Mark Rudnicki, Shravan Kumar Muppa* Analysis of tree sway and crown collisions for managed Pinus resinosa in southern Maine // Forest Ecology and Management. 2013. Vol. 302 P. 193-199.

[9] *Вайс А.А.* Форма крон деревьев сосны обыкновенной (pinus silvestris l.) в чистых высокогустотных насаждениях минусинской котловины красноярского края // Хвойные бореальной зоны. 2017. Т. XXXV, № 3-4. С. 14-20.

[10] *Пятаев А.С.* Определение категорий состояний сосны на основе нечеткой логики по визуальным данным // Труды V международной научной конференции «Региональные проблемы дистанционного зондирования Земли» (РПДЗЗ-2018). Красноярск: ИКИТ СФУ, 2018. С 181-187

[11] *Пятаев А.С., Пятаева А.В., Редькин А.Ю.* Оценка категорий состояния деревьев хвойных пород бореальной зоны по глобальным признакам на основе нечеткой логики // Труды V международной конференции и молодежной школы. Под редакцией В.В. Мясникова. Том 2. Обработка изображений и дистанционные зондирование Земли. Самара: Самарский национальный исследовательский университет имени академика С.П. Королева, 2019. С. 434-438

# МЕТОД РАСПОЗНАВАНИЯ ЛИЦ ПО ВИЗУАЛЬНЫМ ДАННЫМ

*Пятаева А.В., Верхотурова М.В.*
ФГАОУ ВО «Сибирский федеральный университет», Красноярск

В работе предложен метод распознавания лиц по визуальным данным. Выделение лица на видеоизображениях реализовано с помощью метода Виолы-Джонса. При работе алгоритма выделения лиц вычисляются признаки Хаара, используется каскадная классификация. На этапе распознавания лиц использованы локальные бинарные шаблоны.
*Ключевые слова: обнаружение лиц, распознавание лиц, локальные бинарные шаблоны, алгоритм Виолы – Джонса, примитивы Хаара.*

**Введение.** Технологии распознавания лиц применяются для решения различных практических задач, таких как защита персональных данных и личной информации, обеспечение человеко-машинного взаимодействия, реализация доступа на территории объектов с повышенным уровнем безопасности, идентификация личности для обеспечения прокторинга в системах электронного обучения и многих других. Подходы к решению задачи распознавания лиц также значительно различаются. На начальном этапе развития подходов к распознаванию лиц использовались геометрические признаки для выделения характерных особенностей лица [1, 2]. В настоящее время для решения этой задачи используют технологии глубокого обучения [3, 4], эволюционные алгоритмы [5], метод роя частиц [6] и другие подходы. На эффективность распознавания могут оказывать влияние различные факторы, такие как низкий уровень освещенности и зашумленность изображений [7, 8], положение лица [9], пол и возраст и этническая принадлежность человека на изображении [10-12] и пр.

**Метод распознавания лиц.** Первым этапом распознавания лица является его обнаружение на видеопоследовательности или изображении, то есть классификация «один против всех». В таком случае объекты изображения делятся на два класса «лицо», «не лицо». Далее требуется выполнить идентификацию пользователя, то есть понять, какой именно человек снят на видеоизображении. Для решения задачи обнаружения лиц использован метод Виолы Джонса [13]. Основной областью применения метода Виолы-Джонса является задача обнаружения лиц [14, 15]. В основе работы метода Виолы-Джонса лежит выделение признаков, подобных признакам Хаара и использование каскадной модели классификации. Особенностью метода Виолы-Джонса является работа с интегральным способом представления изображения (таблицей сумм). Интегральное представление изображения – это матрица, которая совпадает по размерам с исходным изображением. Каждый ее элемент хранит сумму интенсивностей всех пикселей, находящихся левее и выше данного элемента. Элементы интегрального представления изображения для каждого пиксела исходного изображения рассчитываются по следующей формуле:

$$L(\mathrm{x},\mathrm{y}) = \sum_{i=0}^{x}\sum_{j=0}^{y} I(\mathrm{i},\mathrm{j}),$$

где $I(i,j)$ – яркость текущего пиксела исходного изображения, $(i,j)$ – координаты текущего пиксела. Таким образом, каждый элемент матрицы $L$ представляет собой сумму значений интенсивностей пикселов в прямоугольнике от пиксела (0, 0) до пиксела с координатами $(x,y)$.

Для вычисления признаков Хаара по исследуемому видеоизображению перемещается сканирующее окно, состоящее из смежных прямоугольников – примитивов Хаара. Выделяя такие характерные особенности изменения интенсивности пикселов для лица человека, становится возможным отделить его от других объектов изображения. Таким образом, путем перемещения сканирующего окна по всему изображению вычисляются признаки Хаара, дающие точечное значение перепада яркости пикселов области. Маски Хаара, использованные в работе (рис. 1), позволяют обнаружить не только лица в анфас, но и лица, обращенные к камере с углом поворота более 30 градусов [16]. Далее признаки Хаара в алгоритме Виолы-Джонса

организуются в каскадный классификатор. Результатом работы алгоритма классификации Виолы-Джонса является набор признаков для каждой области, состоящий из 200 значений перепадов яркостей, позволяющий отделить изображения, содержащие лицо, от изображений без него.

Следующим этапом решения задачи распознавания лиц является идентификация конкретной личности. Для этого в настоящем исследовании применен метод текстурного анализа на базе локальных бинарных шаблонов (Local Binary Pattern – LBP). Оператор LBP используется как составная часть многих классификаторов [17-19]. Для идентификации личности изображение, на котором обнаружено лицо, разбивается на непересекающиеся блоки. Для каждого блока вычисление бинарного кода LBP происходит следующим образом:

$$LBP_R(P) = \sum_{n=0}^{P-1} s(I_n - I_c) \cdot 2^n,$$

где $P$ – количество пикселов в окрестности, $R$ – радиус окрестности, $s(x) = 1$, если $x \geq 0$, и $s(x) = 0$ в противном случае, $I_n$ и $I_c$ – значения яркости текущего и центрального пикселов, представляющие собой значение яркости $Y$ из цветовой схемы YUV.

Формирование бинарного кода LBP выполняется следующим образом: если для текущего пиксела значение яркости $Y$ меньше центрального, то текущему разряду LBP присваивается значение «1», в противном случае значение разряда бинарного кода локальных бинарных шаблонов равно «0». Таким образом, получается P-разрядный бинарный код, который описывает окрестность пиксела. В настоящей работе для построения LBP учитывается 8 значений пикселов соседей ($R$=1), обход происходит по часовой стрелке, разрядность бинарного кода LBP равна 8. На основе вычисленных LBP по десятичному коду строится гистограмма фрагментов лица. Далее выполняется операция конкатенации гистограмм, полученных для различных фрагментов лица, в одну гистограмму. В качестве меры различия двух гистограмм использовано Евклидово расстояние, как наиболее часто рекомендуемая метрика:

$$D = \sqrt{\sum_{i=1}^{n} (hist1_i - hist2_i)^2},$$

где $hist1_i$ – столбец с номером $i$ гистограммы исследуемого изображения лица, где $hist2_i$ – столбец с номером $i$ гистограммы изображения из базы эталонов лиц, $n$ - количество столбцов гистограмм.

Применение локальных бинарных шаблонов для решения задачи распознавания лица показано на рис. 2.

Блок-схема алгоритма применения оператора локальных бинарных шаблонов для задачи распознания лица представлена на рис. 3.

Таким образом, объединенная гистограмма фрагментов лица сравнивается по порогу с каждой из эталонных гистограмм, на основе этого сравнения выполняется идентификация пользователя.

**Экспериментальные исследования.** Экспериментальные исследования для этапов обнаружения и распознавания лиц по видеоданным проведены отдельно. Для обнаружения лиц использована выборка видеоизображений, включающая 4916 примеров лиц и 8500 примеров с отсутствием лиц, взятых из хранилища данных Labeled Faces in the Wild Home [20] и набора данных Aberdeen [21]. Для проверки качества работы алгоритма распознавания лиц использован набор данных YouTubeFaces (YTF) [22], McGillFaces Database [23] и Db Fases Dataset [24]. Набор данных YouTubeFaces содержит 3425 видеоролика с изображениями 1595 различных людей, набор McGillFaces Database – 60 видеороликов с изображениями 40 различных людей, набор Db Fases Dataset – 22 видеоролика с изображениями 38 различных людей. Видеоизображения имеют различные уровни освещенности, содержат разное количество людей обоих полов. При этом на видеороликах различаются как количество людей, так и угол поворота головы людей к камере. Изображения имеют различные разрешения, минимальное разрешение изображений 160×120 пикселов, максимальное – 1280×720. На изображениях присутствуют как природные, так и human-made объекты, люди. При этом видеосъемка выполнялась как в

закрытых помещениях, так и на открытых пространствах. Кроме того, люди были совершенно свободны в своих движениях, что приводило к произвольным масштабам и выражениям лица, положению головы. Видеоролики имели формат mp4 или avi. Примеры кадров использованных видеороликов показаны в табл. 1.

Обучающая выборка составила 80 %, тестовая выборка – 20 % от общей выборки. Для оценки эффективности работы алгоритмов обнаружения и распознавания лиц использованы показатели точность обнаружения (TR), ложноположительные (FAR) и ложноотрицательные срабатывания (FRR). Лица людей, повернутые относительно камеры на угол более 55 градусов, не учитывались. Результаты обнаружения и распознавания лиц показаны в табл. 2.

Как показали результаты экспериментальных исследований, пол и возраст людей не влияют на качество работы алгоритма обнаружения и распознавания лиц. На качество работы алгоритма оказывают влияние такие факторы как, уровень освещенности сцены, разрешение видеоролика, скорость перемещения людей на сцене, угол поворота лица и степень открытости лица. Так, дополнительную погрешность в работу алгоритма обнаружения и распознавания лиц вносят надетые на лицо аксессуары, например, очки, шарфы, шапки. Негативное влияние оказывает также закрывание части лица волосами, наличие бороды или усов. Эмоциональное выражение лица в большинстве случаев не влияет на результаты работы алгоритма, однако оно способно вызвать трудности при распознавании, например, при широкой улыбке или закрытых глазах человека. Кроме того, при затенении части лица качество работы алгоритма может снижаться.

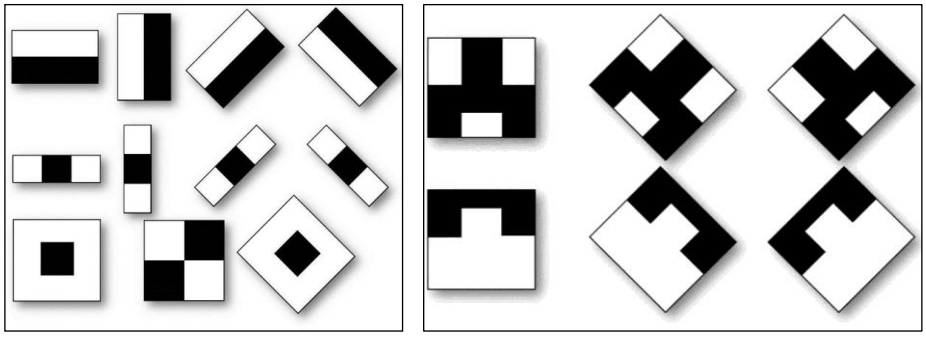


Рис. 1. Использованные примитивы Хаара – основные (слева) и дополнительные (справа).

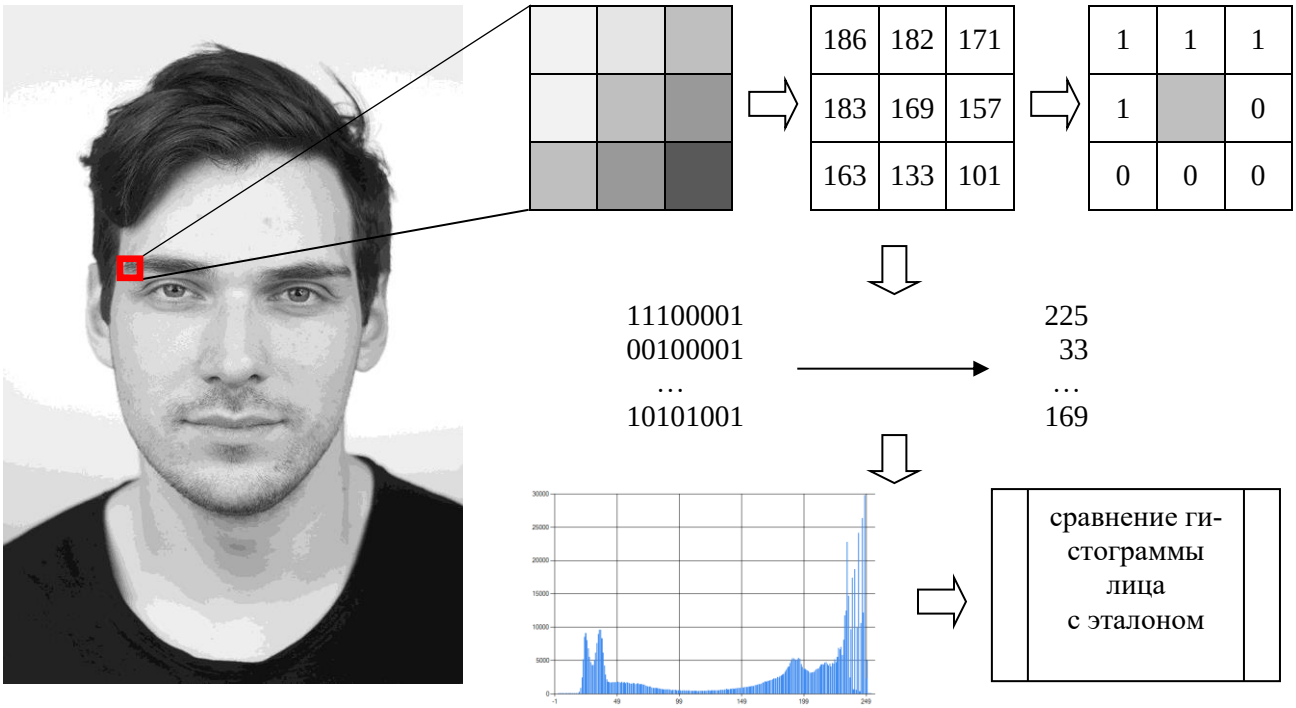



Рис. 2. Применение локальных бинарных шаблонов.

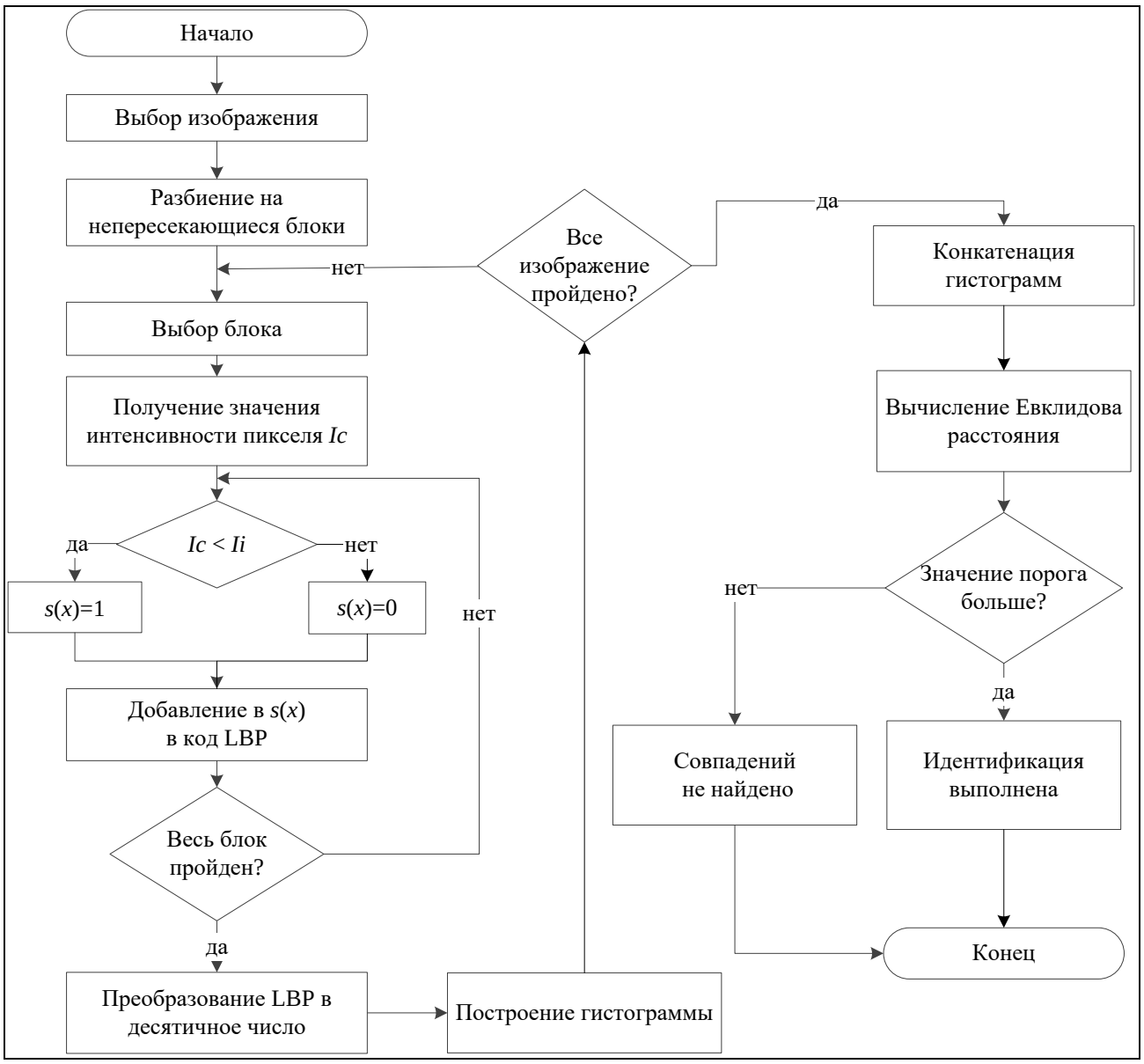



Рис. 3. Блок-схема алгоритма распознавания лица.

Обучающая выборка составила 80 %, тестовая выборка – 20 % от общей выборки. Для оценки эффективности работы алгоритмов обнаружения и распознавания лиц использованы показатели точность обнаружения (TR), ложноположительные (FAR) и ложноотрицательные срабатывания (FRR). Лица людей, повернутые относительно камеры на угол более 55 градусов, не учитывались. Результаты обнаружения и распознавания лиц показаны в табл. 2.

Как показали результаты экспериментальных исследований, пол и возраст людей не влияют на качество работы алгоритма обнаружения и распознавания лиц. На качество работы алгоритма оказывают влияние такие факторы как, уровень освещенности сцены, разрешение видеоролика, скорость перемещения людей на сцене, угол поворота лица и степень открытости лица. Так, дополнительную погрешность в работу алгоритма обнаружения и распознавания лиц вносят надетые на лицо аксессуары, например, очки, шарфы, шапки. Негативное влияние оказывает также закрывание части лица волосами, наличие бороды или усов. Эмоциональное выражение лица в большинстве случаев не влияет на результаты работы алгоритма, однако оно способно вызвать трудности при распознавании, например, при широкой улыбке или закрытых глазах человека. Кроме того, при затенении части лица качество работы алгоритма может снижаться.

Табл. 1. Примеры кадров тестовых видеопоследовательностей.

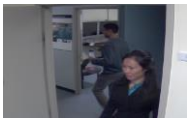
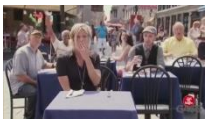
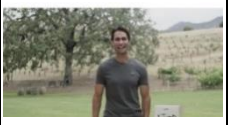
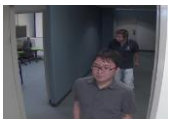
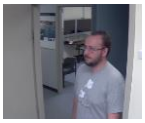
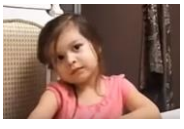
| **Описание тестовой видеопоследовательности** | **Образец кадра** | **Описание тестовой видеопоследовательности** | **Образец кадра** |
|---|---|---|---|
| YouTubeFaces\P1E_S1_C1.mp4.<br>Количество кадров: 1125.<br>Разрешение: 1280×720.<br>Различных лиц:2 женских. |  | YouTubeFaces\P1E_S2_M6.mp4.<br>Количество кадров:1025<br>Разрешение: 640×480.<br>Различных лиц:3 женских. |  |
| YouTubeFaces\P1E_S1_C4.mp4.<br>Количество кадров:1500.<br>Разрешение: 1280×720.<br>Различных лиц:1 женское, 1 мужское. |  | YouTubeFaces\P1E_S2_M1.mp4.<br>Количество кадров:875.<br>Разрешение: 640×480.<br>Различных лиц:1 женское, 1 мужское. |  |
| db\faces\crglad\00002.avi<br>Количество кадров: 300<br>Разрешение:160×120,<br>Различных лиц:1 женское. |  | YouTubeFaces\P1E_S2_M3.mp4.<br>Количество кадров: 750.<br>Разрешение: 1280×720.<br>Различных лиц:2 мужских лица. |  |
| YouTube-Faces\P1E_S2_M13.mp4.<br>Количество кадров: 850.<br>Разрешение: 640×480.<br>Различных лиц:5 женских, 1 мужское. |  | McGillFaces Database\mmdm2\video \sx372.avi.<br>Количество кадров: 400<br>Разрешение: 640×480.<br>Различных лиц:1 женское и 1 мужское лицо. |  |
| YouTubeFaces\P1E_S2_D3.mp4.<br>Количество кадров: 575.<br>Разрешение:480×320.<br>Различных лиц: 1 женское, 2 мужских. |  | YouTubeFaces\P1E_S2_D5.mp4.<br>Количество кадров: 900.<br>Разрешение: 1280×720.<br>Различных лиц:2 мужских. |  |
| McGillFaces Database\mmdm2\video\si2028.avi.<br>Количество кадров: 575.<br>Разрешение: 640×480.<br>Различных лиц:2 мужских. |  | YouTubeFaces\P1E_S1_C3.mp4.<br>Количество кадров: 1250.<br>Разрешение: 1280×720.<br>Различных лиц: 1 мужское. |  |
| McGillFaces Database\mmdm2\video\sx102.avi.<br>Количество кадров: 425.<br>Разрешение: 640×480.<br>Различных лиц: 1 мужское. |  | db\faces\crglad\000046.avi.<br>Количество кадров: 300.<br>Разрешение: 160×120.<br>Различных лиц:2 мужских. |  |
| McGillFaces Database\mmdm2\video\si1425.avi.<br>Количествокадров: 375.<br>Разрешение: 640×480.<br>Различных лиц: 1 мужское. |  | db\faces\crglad\000050.avi.<br>Количество кадров: 300.<br>Разрешение: 160×120.<br>Различных лиц:1 мужское лицо. |  |
| McGillFaces Database\mmdm2\video\sa1.avi.<br>Количество кадров: 325.<br>Разрешение: 640×480.<br>Различных лиц: 3 мужских. |  | YouTubeFaces\P1E_S2_M7.mp4.<br>Количество кадров: 700<br>Разрешение:640×480.<br>Различных лиц: 2 женских, 2 мужских. |  |
| YouTubeFaces\P1E_S1_K1.mp4.<br>Количество кадров: 300.<br>Разрешение:480×320.<br>Различных лиц: 1детское. |  | YouTubeFaces\P1E_S1_K2.mp4.<br>Количество кадров: 400.<br>Разрешение:480×320.<br>Различных лиц: 1 детское. |  |

Табл. 2. Результаты экспериментальных исследований.

| Видеоролик | Обнаружение лиц | | | Распознавание лиц | | |
|---|---|---|---|---|---|---|
| | TD,% | FRR,% | FAR,% | TR,% | FRR,% | FAR,% |
| **YouTubeFaces\P1E_S1_C1** | 100 | 0,00 | 0,00 | 99,5 | 0,50 | 0,44 |
| **YouTubeFaces\P1E_S1_C3** | 100 | 0,00 | 0,00 | 100 | 0,00 | 0,00 |
| **YouTubeFaces\P1E_S1_C4** | 100 | 0,00 | 0,00 | 99,1 | 0,01 | 0,66 |
| **YouTubeFaces\P1E_S2_M1** | 100 | 0,00 | 0,00 | 97,5 | 2,50 | 1,10 |
| **YouTubeFaces\P1E_S2_M3** | 100 | 0,00 | 0,00 | 96,2 | 4,00 | 3,80 |
| **YouTubeFaces\P1E_S2_M6** | 100 | 0,00 | 0,00 | 100 | 0,00 | 0,00 |
| **YouTubeFaces\P1E_S2_M13** | 93,9 | 6,10 | 5,80 | 87,9 | 12,1 | 11,7 |
| **YouTubeFaces\P1E_S2_M7** | 100 | 0,00 | 0,00 | 100 | 0,00 | 0,00 |
| **YouTubeFaces\P1E_S2_D3** | 100 | 0,00 | 0,00 | 100 | 0,00 | 0,00 |
| **YouTubeFaces\P1E_S2_D5** | 95,3 | 6,66 | 4,70 | 88,9 | 11,1 | 11,1 |
| **YouTubeFaces\P1E_S2_D6** | 100 | 0,00 | 0,00 | 98,2 | 1,80 | 1,78 |
| **YouTubeFaces\P1E_S2_D8** | 100 | 0,00 | 0,00 | 100 | 0,00 | 0,00 |

**Заключение.** Решение задачи распознания лиц на сегодняшний день актуально для реализации различного рода практических задач. В настоящей работе для выделения лица использован алгоритм Виолы-Джонса, для распознавания лица применены локальные бинарные шаблоны. Экспериментальные исследования, проведенные на разнородных видеоданных, подтверждают эффективность предложенных методов.

## ЛИТЕРАТУРА


[1] *Turk M., Pentland A.* Eigenfaces for recognition // J. Cognit. Neurosci. 1991. No. 3 (1). P. 71-86.

[2] *Belhumeur P.N., Hespanha J.P., Kriegman D.J.* Eigenfaces vs. fisherfaces: recognition using class specific linear projection // IEEE Trans. Pattern Anal. Mach. Intell. 1997. No. 19 (7). P. 711-720.

[3] *Taigman Y., Yang M., Ranzato M., Wolf L.* Deepface: closing the gap to human-level performance in face verification // Proc. IEEE Conf. on Comp. Vis. and Patt. Recogn. 2014. P. 1701-1708.

[4] *Parkhi O.M., Vedaldi A., Zisserman A.* Deep face recognition // Proceedings of the British Machine Vision Conference (BMVC). 2015. Vol. 1. P. 41.1-41.12.

[5] *Zhi H., Liu S.* Face recognition based on genetic algorithm // Journal of Visual Communication and Image Representation. 2019. Vol. 58. P. 495-502.

[6] *Khan S.A., Ishtiaq M., Nazir M., Shaheen M.* Face recognition under varying expressions and illumination using particle swarm optimization // Journal of Computational Science. 2018. Vol. 28. P. 94-100.

[7] *Nikan S., Ahmadi M.* A modified technique for face recognition under degraded conditions // Journal of Visual Communication and Image Representation. 2018. Vol. 55. P. 742-755.

[8] *Ding C., Tao D.* Pose-invariant face recognition with homography-based normalization // Pattern Recognition. 2017. Vol. 66. P. 144-152.

[9] *Liang Y., Zhang Y., Zeng X.X.* Pose-invariant 3D face recognition using half face // Signal Processing: Image Communication. 2017. Vol. 57. P. 84-90.

[10] *Wu S., Wang D.* Effect of subject's age and gender on face recognition results // Journal of Visual Communication and Image Representation. 2019. Vol. 60. P. 116-122.

[11] *Muikudi P.B.L., Hills P.J.* The combined influence of the own-age, -gender, and –ethnicity bases on face recognition // Acta Psychologia. 2019. Vol. 194. P. 1-6.

[12] *Segal S.C., Reyes B.N., Gobin K.C., Moulson M.C.* Children's recognition of emotion expressed by own-race versus other-race faces // J. Experim. Child Psychol. 2019. Vol. 182. P. 102-113.

[13] *Viola P., Jones M.J.* Rapid Object Detection using a Boosted Cascade of Simple Features // Proceedings IEEE Conf. on Computer Vision and Pattern Recognition. 2001. Vol. 1. P. 511-518.

[14] *Irgens P., Bader C., Lé T., Saxena D., Ababei C.* An efficient and cost effective FPGA based implementation of the Viola-Jones face detection algorithm // Hardware X. 2017. No.1. P. 68– 75.

[15] *Nguyen T., Hefenbrock D., Oberg J. et al.* A software-based dynamic-warp scheduling approach for load-balancing the viola-jones face detection algorithm on gpus // J. Paral. Distrib. Comput. 2013. N 73(5). P. 677-685.

[16] *Пятаева А.В., Верхотурова М.В.* Детектирование лиц с использованием метода Виолы-Джонса // Матер. V Междунар. науч. конф. «Региональные проблемы дистанционного зондирования Землию». Красноярск: ИКИТ СФУ, 2018. С. 188-191.
[17] *Yuan F., Shi J., Xia X., Zhang L., Li S.* Encoding pairwise Hamming distances of Local Binary Patterns for visual smoke recognition // Computer Vision and Image Understanding. 2019. Vol. 178. P. 43-53.
[18] *Xu Z., Jiang Y., Wang Y., Zhou Y., Li W., Liao Q.* Local polynomial contrast binary patterns for face recognition // Neurocomputing. 2019. Vol. 355. P. 1-12.
[19] *Hassaballah M., Alshazly H.A., Ali A.A.* Ear recognition using local binary patterns: A comparative experimental study // Expert Systems with Applications. 2019. Vol. 118. P.182-200.
[20] Labeled Faces in the Wild Home набор данных. **http://vis-www.cs.umass.edu/lfw/** (дата обращения 09.06.2019).
[21] Aberdeen набор данных. **http://pics.psych.stir.ac.uk/2D_face_sets.htm** (дата обращения 09.06.2019).
[22] YouTubeFaces набор данных. **http://www.cs.tau.ac.il/~wolf/ytfaces/index.html#download** (дата обращения 09.06.2019).
[23] McGillFaces набор данных. **https://sites.google.com/site/meltemdemirkus/mcgill-unconstrained-face-video-database** (дата обращения 09.06.2019).
[24] Db Fases набор данных. **http://www.videorecognition.com/db/video/faces/cvglab/** (дата обращения 09.06.2019).

# РАССМОТРЕНИЕ ПОДХОДОВ К ПРЕДСТАВЛЕНИЮ ЗНАНИЙ ОБ ОПЕРИРОВАНИИ ПРОСТРАНСТВЕННЫМИ ДАННЫМИ В ЗАДАЧАХ МОНИТОРИНГА

*Раевич К.В., Маглинец Ю.А., Брежнев Р.В.*
Институт космических и информационных технологий СФУ, г. Красноярск

Работа посвящена исследованию проблемы представления разнородных знаний и данных о пространственных объектах и способах оперирования ими при решении задач мониторинга под управлением многоцелевой системы дистанционного мониторинга ИКИТ (МСДМ). Представлено видение классификации знаний, которая систематизирует знания о семантике, синтаксисе, поведении объекта мониторинга и правила их интерпретации. Рассмотрен подход к представлению знаний, позволяющий формализовать постановку задачи мониторинга.
*Ключевые слова: дистанционное зондирование Земли, постановка задачи, онтологический подход, база знаний, объект мониторинга, модель знаний.*

**Введение.** Современные тенденции непрерывного роста объема, улучшения качества, доступности данных, поступающих в центры обработки от спутниковых систем дистанционного зондирования Земли, развития средств распределенного хранения и доступа, а также методов и алгоритмов обработки пространственных данных, потенциально позволяют расширять круг решаемых прикладных задач дистанционного зондирования и мониторинга наземных объектов и повысить качественные показатели их решения. Указанные факторы создают в области обработки ДДЗ ситуацию «информационного взрыва», известную в информатике с 70-х годов: присутствуют данные, имеются средства их обработки, однако в силу того, что задачи космического мониторинга во многом носят экспертный характер, лимитирующим фактором на пути их решения становится наличие экспертов в цепочке обработки данных. Это приводит к тому, что между формированием запроса конечного пользователя – постановщика задачи и получением ее решения возникает ряд операций – как автоматических, так и требующих привлечения специалистов в области алгоритмов обработки пространственных данных, что приводит к издержкам, как временным, так и технологическим.

В последнее время рядом научных коллективов достигнуты существенные достижения в области формирования технологических циклов решения задач мониторинга наземных объектов, алгоритм решения которых известен. Так, в работе [1] сформулирован ряд необходимых элементов технологии: источники и архивы данных ДЗЗ, системы первичной и тематической обработки данных для создания информационных продуктов, специализированные системные интерфейсы, средства представления данных. Развитие подходов, рассмотренных в [1, 2], указывает на смену концепций построения интерфейса «*конечный пользователь – специалист*» на тип «*конечный пользователь – решающая система*». Однако, существующие автоматизированные решения, ориентированные на конечного пользователя, как правило, не обладают должной гибкостью, так как нацелены на решение задач в узко сформулированной постановке и изменения в характеристиках задачи приводят к возникновению необходимости в привлечении экспертов. Тем самым, актуально применение методов инженерии знаний, позволяющих обособить экспертные знания, выводя эксперта их цепочки обработки данных. В статье рассмотрены основные категории знаний, проработка которых требуется в рамках разрабатываемого авторским коллективом объектно-ориентированного подхода к построению систем дистанционного зондирования и мониторинга наземных объектов.

**Основные характеристики запроса конечного пользователя к системе космического мониторинга.** Рассматривается ситуация диалогового взаимодействия между актором «Постановщик задачи», далее – Постановщик и автоматизированной системой мониторинга (далее – Системой). Основная цель Постановщика – организация дистанционного наблюдения за наземным объектом, представляющим для него интерес (далее – объект мониторинга, ОМ).

При этом Постановщик формулирует цель наблюдения и, через некоторое время, ожидает получения результата, либо совокупности результатов.

В простейшем случае, его цель – увидеть одномоментный «срез», характеризующий состояние ОМ. Это – задача разрешения информационного запроса (элементарного измерения). Данная задача важна в методическом плане, так как любая сколь угодно сложная задача наблюдения может быть решена путем выполнения множества элементарных измерений и последующей обработки их результатов.

В более общем случае, цель Постановщика – установить наблюдение за Объектом и, с определенной периодичностью, отслеживать изменения ОМ во времени. Роль Системы в этом процессе – сообщать об изменениях Постановщику. При этом необходимо определить степень существенности изменений, чтобы, с одной стороны, не пропустить важное изменение, с другой – не инициировать коммуникацию при возникновении рутинных событий. Таким образом, решается задача мониторинга.

В случае, если, помимо средств наблюдения, в арсенале Постановщика имеются средства воздействия на объект, корректирующие его состояние, Постановщик может сформулировать задачу автоматического, либо автоматизированного управления ОМ.

В свою очередь, цели Системы – определить необходимый темп измерений и источники информации; определить дешифровочные признаки аспектов объекта, подлежащих обнаружению на изображении, в контексте каждого из источников; установить соответствие между численными значениями, измеряемыми на изображении и соответствующими характеристиками объекта мониторинга, которые необходимо определить в соответствии с запросом Постановщика.

Таким образом, основные параметры постановки задачи, которые необходимо прояснить в процессе диалога между Постановщиком и Системой, это:

- Известные Постановщику наименование объекта (вида объектов) и его отличительные черты (свойства) для идентификации Системой среди многообразия объектов окружающего мира;
- Известные Постановщику структурные характеристики объекта (если они существенны для идентификации);
- Известные Постановщику отношения между ОМ с внешней средой (если они существенны для идентификации);
- Свойства объекта, которые необходимо определить;
- Структура объекта, которую необходимо определить;
- Отношения между ОМ и внешней средой, которые необходимо определить.

**Концептуальная модель знаний системы космического мониторинга.** На концептуальном уровне выделены две категории знаний: знания об объекте мониторинга; системные знания.

1. Знания об объекте мониторинга обобщенно можно выразить следующим образом:

$$O =< S_O, R, P, \Phi >, \quad (1)$$

где $O$ – объект мониторинга, $S_O$ – множество специализаций объекта; $R$ – множество бинарных отношений с другими объектами, формализованных в [5, 6]; $P$ – множество разнородных свойств объекта, при этом $S_{Oi}$ характеризуется подмножеством $\{p_{1i}, p_{2i}, p_{3i}, \dots, p_{ni}\} \in P$; $\Phi$ – множество правил интерпретации, с которыми ассоциированы методы и алгоритмы расчета признаков объекта в той или иной специализации.

1.1. *Знания о семантике ОМ.* Для согласования ментальной модели Постановщика с внутренним представлением знаний Системы предложено описание, являющееся дальнейшим развитием подхода, рассмотренного в [7]. Постановка задачи характеризуется в следующих категориях:

- Таксономии понятий предметной области. Отдельно взятой предметной области сопоставлены одна или более таксономий, каждая из которых задает иерархическую классификацию ОМ с определенной позиции. Отдельно взятый элемент (узел таксономического дерева) может иметь различные специализации.

- Свойства элементов таксономий. Отдельно взятый элемент описывается множеством свойств, характеризующих соответствующий аспект объекта. Выделяются абстрактные характеристики (классы характеристик) и конкретные характеристики, каждой из которых могут быть сопоставлены интервал значений, тип шкалы для измерения, метод измерения и возможные интерпретации.
- Отношения между элементами таксономий. Отдельно взятая пара элементов может характеризоваться множеством заданных на ней бинарных отношений.
- Агрегации предметной области. Описывают допустимые структурные комбинации исследуемого объекта, как совокупности компонент (элементов, частей).

1.2. *Знания о поведении ОМ*. Сопоставляются специализации ОМ, представленные множеством измеряемых характеристик, каждая (специализация или характеристика) из которых соотнесена с соответствующим узлом таксономии понятий и моделью плановой траектории изменения значений характеристик во времени, присущей данной специализации ОМ.

1.3. *Знания о синтаксисе ОМ*. Сводятся к совокупности ключевых особенностей (признаков) объекта мониторинга, наблюдаемых на изображении:

- структурные особенности,
- признаки формы,
- яркостные и текстурные характеристики, наблюдаемые в том или ином диапазоне, либо наборе диапазонов электромагнитного спектра,
- положение в пространстве,
- изменчивость во времени (в разрезе каждой из вышеуказанных особенностей).

1.4. *Знания о правилах интерпретации*, позволяющие на основании измерений синтаксических признаков и характеристик исходной сцены судить о значениях семантических характеристик.

2. Управляющие (системные) знания. Регламентируют управление процессами преобразования данных ДЗЗ. Большое количество различных этапов обработки данных обеспечиваются множеством методов и алгоритмов. Для построения классификатора в этом множестве важно выделить классы однотипных по функциональному назначению процедур, которые в свою очередь могут быть при необходимости разделены на атомарные действия. Системные знания можно разделить на следующие подклассы:

2.1. *Знания об алгоритмах преобразования данных* могут быть заданы декларативно посредством списков или реляционных структур [3, 4], где учтены взаимосвязи между алгоритмами, задачами и решающими правилами, что позволяет формализовать контекст задачи, способы ее решения, описание исходных данных, правила представления промежуточных и конечных результатов и прочее, в виде отдельных таксономий.

2.2. *Знания о средствах взаимодействия с конечным пользователем* описывают правила постановки задачи конечным пользователем, правила внутренней интерпретации задачи и правила представления результата разрешения задачи. Эти правила призваны описать структуру взаимодействия «*конечный пользователь – решающая система*». Указанная пользователем информация на этапе постановки задачи должна играть определяющую роль при выборе конкретного алгоритма или совокупности алгоритмов в конкретный момент времени.

2.3. *Знания о методах и алгоритмах преобразования данных*. В условиях разрешения информационного запроса пользователя уместно говорить о совокупности алгоритмов разной сложности и области действия, которые организовываются в соответствующие конструкции, объединенные логикой преобразования входных данных и требованиями к выходным данным. Построение цепочек алгоритмов является компетенцией специалиста по обработке и анализу данных. Так же специалист должен ассоциировать (задать множество $R$ для каждого алгоритма) сформированные цепочки с типовыми задачами конечных пользователей.

2.4. *Знания о преобразуемых данных*. Для успешного разрешения информационного запроса пользователя Системы должна «уметь» осуществлять поиск релевантных и актуальных данных. К таким данным относятся спутниковые снимки, которые обладают значительным

разбросом значений пространственного разрешения, а также радиометрического, спектрального и временного. Следовательно, для различных классов задач необходимо задать ограничения на возможные источники данных ДЗЗ.

**Реализация предлагаемого подхода.** Начальный этап проектирования базы знаний основан на двух основных шагах: 1) подготовка и структурирование знаний об объекте $O_i$ в выбранной спецификации $S_{O_i}$, которая включает означивание элементов множеств $R$, $P$ и $\Phi$; 2) экспорт подготовленной структуры в базу данных. Программная среда, которая обеспечивает выполнение этих шагов, включает такие основные компоненты, как редактор OWL-онтологий Protégé, формирующий RDF структуру данных, описывающую граф объектов и их отношений в формате XML; локальный веб-интерфейс, позволяющий разбирать XML конструкции структуры данных и экспортировать их в БД MySQL, структура которой представлена на рис. 1.

Логическая структура данных включает такие основные сущности, как справочник классов объектов (`Classes`), справочник экземпляров всех классов (`Individuals`), справочник возможных отношений между экземплярами классов (`Object_property`), справочник возможных свойств объектов (`Data_properties`), описание отношения класс-подкласс (`Subclasses`).

Рассмотренная концепция представления структуры знаний нашла свое практическое воплощение в созданной авторским коллективом автоматизированной системе агромониторинга [8, 9]. Данные и знания, накопленные в этой области, позволяют связать спектральные и почвенные характеристики объекта, пространственные и топологические, технологические и экономические и другие группы его свойств. Такие связи позволяют давать ответы на ряд пользовательских запросов (рис. 2), как существующих, так и перспективных.

Запросы можно разделить на единичные и циклические. Единичные запросы, как правило, представляют собой команды, требующие немедленного исполнения. Примерами таких запросов в рамках МСДМ могут быть расчеты метрических признаков, например, площади, периметра и т.д.

Циклические запросы подразумевают регулярно повторяющиеся с заданной частотой запросы на измерение заданного параметра для заданного объекта, что соответствует задачам мониторинга. Типовыми примерами может являться мониторинг структуры растительного покрова, температуры поверхности, влажности и т.д.

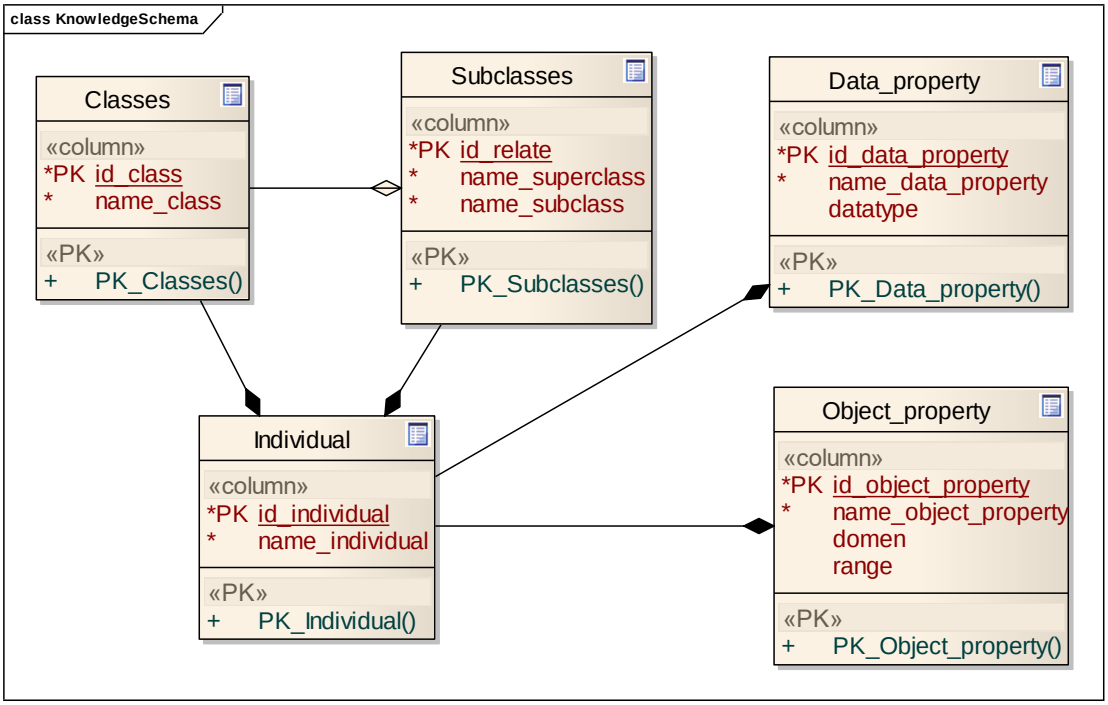



Рис. 1. Логическая схема базы знаний.




Рис. 2. Типовые запросы пользователей к решающей системе на примере оценки земель сельскохозяйственного назначения (ЗСХН).

В качестве примера рассмотрим задачу мониторинга неоднородности структуры растительного покрова в пределах сельскохозяйственного контура, модель которой можно представить в виде выражения (2):

$$\text{ОМ} =< P_{\text{П}}, NS, NP, T, N, Tm, K, E, \varphi, F > \quad (2)$$

В семантическом пространстве объект мониторинга характеризуется сельскохозяйственной культурой $k^i \in K$, характеризующейся своей моделью развития во времени $Tm$, определяемой последовательностью фенологических фаз $\varphi = \{\varphi^i\}$ и планом агротехнических мероприятий, представленных в (2) в виде совокупности событий $E = \{e^i\}$.

Множество диапазонов времени $Tm$ изменения состояний $F$ объекта характеризует поведение объекта, плановую траекторию его жизненного цикла.

Синтаксическое описание включает множество метрических признаков, описывающих форму (множество координат $P_{\text{П}}$, толщина $T$) и размеры (площадь $NS$, периметр $NP$) объекта, а также спектральный индекс NDVI.

Правила интерпретации усредненных значений $NDVI - \overline{N}$ позволяют сопоставлять состояние $f^i \in F$ каждой неоднородной области $w_j$ и объекта в целом [8, 9].

Фрагмент базы знаний, представляющий совокупность и взаимосвязь элементов таксономии при решении данной задачи, представлен на рисунке 3 в виде ориентированного графа.

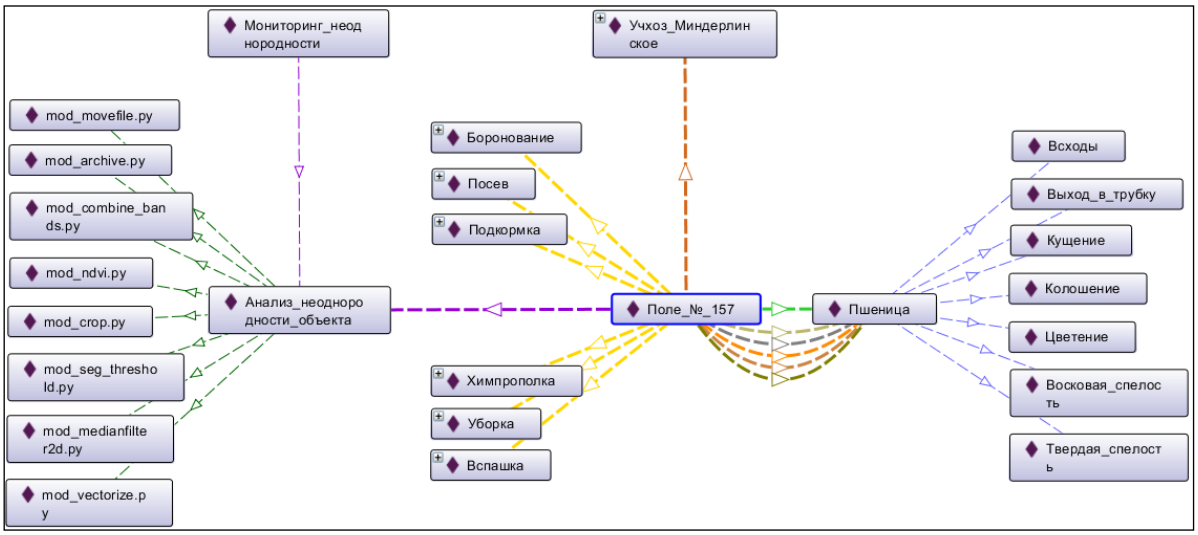



Рис. 3. Взаимосвязь объектов базы знаний МСДМ ИКИТ при решении задачи мониторинга неоднородности структуры растительности.

Центральной вершиной графа является «Поле № 157», которое принадлежит сельскохозяйственному предприятию ООО «Учхоз Миндерлинское». На поле в период с 2013 по 2018 годы сеяли пшеницу, что обозначено множеством дуг, соединяющих объект «Поле № 157» с объектом «Пшеница». На объекте «Поле» выполняются агротехнологические мероприятия, такие как боронование, посев, подкормка и другие. Экземпляры класса «Агротехнологическое мероприятие» характеризуется рекомендуемым интервалом времени выполнения.

Пшеница характеризуется набором фенологических фаз развития (всходы, кущение, выход в трубку и т.д.), что отражено на графе связями объекта «Пшеница» с экземплярами класса «Фенологическая фаза». Фенологические фазы, в свою очередь, характеризуются эталонным интервалом времени, длительностью и диапазоном значений NDVI [8].

**Заключение.** Рассмотренный в работе подход к организации знаний системы космического мониторинга предполагает использование предложенной классификации знаний, задавая наборы концептов и отношений, моделирующих интересующую пользователя предметную область, а также общесистемные представления. Созданная база знаний прошла тестовые испытания, основанные на разрешении ряда типовых SPARQL и SQL запросов, имитирующих информационные запросы пользователей. Программный модуль, отвечающий за диалоговое взаимодействие конечного пользователя с решающей системой, позволяющий формулировать задачи единичного или циклического измерения заданных параметров для оценки состояния заданного объекта в различных спецификациях, находится в стадии прототипирования.

*Исследование выполнено при финансовой поддержке РФФИ (грант № 18-47-242002-р_мк), Правительства Красноярского края, Красноярского краевого фонда науки (научный проект «Разработка технологии создания интеллектуальных информационных систем объектно-ориентированного мониторинга территорий по данным дистанционного зондирования»).*

## ЛИТЕРАТУРА

[1] *Лупян Е.А., Балашов И.В., Бурцев М.А., Ефремов В.Ю., Кашницкий А.В. и др.* Создание технологий построения информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 5. С. 53–75.

[2] *Traynor C., Williams M.G.* End users and GIS: a demonstration is worth a thousand words // Interactive Technologies. 2001. P. 115-133.

[3] *Трусова Ю.О.* Представление предметной области "Анализ изображений" в виде специализированного тезауруса: Дис. … канд. техн. наук. Москва, Учреждение Российской академии наук Вычислительный центр им. А.А. Дородницына РАН, 2009. 111 с.

[4] Патент РФ № № 2003133170/09, 10.07.2005. База знаний по обработке, анализу и распознаванию изображений // Патент России № 2256224. 2003. Бюл. № 19. / Гуревич И.Б.

[5] *Chen J. et al.* A survey of qualitative spatial representations // The Knowledge Engineering Review. 2015. Vol. 30, № 1. P. 106-136.

[6] *Papadias D., Sellis T.* Qualitative representation of spatial knowledge in two-dimensional space // The VLDB Journal. 1994. Vol. 3. No. 4. P. 479-516.

[7] *Maglinets Yu. A., Raevich K. V., Tsibulskii G. M.* Knowledge-based geoinformation technology for evaluation of agricultural lands // Procedia Engineering. 2017. Vol. 201. P. 331-340.

[8] *Brezhnev R.V., Maglinets Yu.A., Raevich K.V., Tsibulski G.M.* Modeling of Agricultural Spatial Objects with Heterogeneous Dynamically Changing Spatial Structure // CEUR Workshop Proceedings. 2018. Vol. 2210. P. 316 - 322.

[9] *Brezhnev R.V., Maglinets Yu.A.* The dynamic model of agricultural land structure on the space images in the precision agriculture tasks // E3S Web of Conferences. 2019 Vol. 75. Article Number 01001, RPERS 2018.

# ЛОКАЛИЗАЦИЯ ТОЧЕЧНЫХ ИСТОЧНИКОВ СО СЛУЧАЙНЫМ ПРОСТРАНСТВЕННЫМ РАСПОЛОЖЕНИЕМ И СЛУЧАЙНОЙ ДИСЦИПЛИНОЙ ГЕНЕРАЦИИ ИМПУЛЬСОВ

*Резник А.Л.[(1)], Тузиков А.В.[(2)], Торгов А.В.[(1)], Соловьев А.А.[(1)], Ковалев В.А.[(2)]*
[(1)] Институт автоматики и электрометрии СО РАН, Новосибирск
[(2)] Объединенный Институт проблем информатики НАН Беларуси, Минск

Обсуждаются методы и алгоритмы оптимальной по быстродействию пространственной локализации точечных импульсных источников, обнаруживающих себя генерацией в случайные моменты времени мгновенных дельта-импульсов. Предложены оптимальные поисковые процедуры, ориентированные на локализацию случайных точечно-импульсных объектов в стандартных и усложненных режимах поиска (например, при отсутствии априорной информации об интенсивности разыскиваемого импульсного источника либо когда плотность его распределения в пределах поискового интервала неизвестна).
*Ключевые слова: оптимальный поиск, случайный точечно-импульсный источник, точность локализации, приемное устройство.*

**Введение**. Исследования по оптимальному поиску случайных импульсных объектов актуальны для многих научно-технических дисциплин. Необходимость в их проведении возникает при проектировании различных электронно-оптических преобразователей и детекторов [1]; в задачах подавления импульсных помех на зашумленных и слабоконтрастных изображениях [2]; в технической диагностике при поиске неисправностей, проявляющихся в форме перемежающихся отказов [3]; в задачах обнаружения радиоактивных источников с помощью систем, состоящих из одного или нескольких сенсоров [4], в радиофизике и радиоастрономии при поиске источников гравитационных волн [5] и во многих других приложениях. В настоящем сообщении представлены методы и алгоритмы оптимального по быстродействию поиска точечных пуассоновских источников, обнаруживающих себя генерацией в случайные моменты времени мгновенных дельта-импульсов. Оптимальный алгоритм поиска должен, как правило, удовлетворять одному из двух требований: либо минимизировать суммарное поисковое усилие, необходимое для обнаружения объекта; либо максимизировать полную вероятность обнаружения при наличии ограниченного поискового усилия.

Под точечно-импульсным источником ниже будет пониматься объект пренебрежимо малых угловых размеров (математическая точка), имеющий случайное распределение на оси абсцисс с априорной плотностью вероятности $f(x)$ и излучающий бесконечно короткие импульсы ($\delta$-функции) с пуассоновской интенсивностью $\lambda$. Таким образом, временны*е* интервалы между импульсами являются случайной величиной $t$ с показательной плотностью вероятности $h(t)=\lambda exp(-\lambda t)$. Поиск объекта осуществляется с помощью регистрирующего устройства, обладающего "окном" обзора с произвольно перестраиваемой функцией времени. Импульс фиксируется, если активный объект, инициировавший импульс, находится в окне обзора регистрирующего устройства. В противном случае импульс считается пропущенным. При регистрации импульса окно сужается, в результате чего положение источника уточняется. Требуется за минимальное (в статистическом плане) время отыскать источник с точностью $\varepsilon$.

**Алгоритмы одноэтапного поиска.** Вводя в рассмотрение бинарную функцию

$$u(x,t)=\begin{cases}1, & \text{если точка } x \text{ в момент времени } t \text{ находится} \\ & \text{в окне обзора приемного устройства,} \\ 0, & \text{в противном случае,}\end{cases}$$

описывающую окно обзора в момент времени $t$, получаем соотношение для среднего времени от начала поиска до регистрации первого импульса:

$$\langle\tau\rangle = \lambda\int_0^\infty dt\int_0^\infty dx\left[tf(x)u(x,t)\exp(-\lambda\int_0^t u(x,\xi)d\xi)\right].$$

Для случайного априорного распределения импульсного источника на оси *x* построение даже одноэтапной (заканчивающейся сразу же при регистрации первого импульса) процедуры оптимального по времени поиска вызывает значительные трудности. В одношаговых алгоритмах *периодического* поиска относительная нагрузка $\varphi(x)$ на точку *x* (то есть относительное время ее пребывания в окне обзора) остается постоянной в течение всего времени поиска. При таком подходе задача состоит в том, чтобы найти функцию $\varphi(x)$, минимизирующую среднее время поиска

$$\langle\tau\rangle = \frac{1}{\lambda}\int\frac{f(x)}{\varphi(x)}dx \tag{1}$$

при соблюдении условий

$$\int\varphi(x)dx = \varepsilon, \tag{2}$$

$$0 \le \varphi(x) \le 1. \tag{3}$$

Оптимизация выражения (1) при ограничениях (2), (3) относится к задачам нелинейного программирования [7]. Для ее решения воспользуемся методом неопределенных множителей Лагранжа [8] и будем искать функцию $\varphi(x)$, минимизирующую выражение

$$\int\left[\frac{f(x)}{\varphi(x)} + \mu\varphi(x)\right]dx.$$

Дифференцируя по $\varphi$ и учитывая ограничение (2), получим

$$\varphi(x) = \varepsilon\sqrt{f(x)}\Big/\int\sqrt{f(x)}dx \quad . \tag{4}$$

Если при любых *x* условие (3) не нарушается, то функция (4) является решением сформулированной экстремальной задачи. Если же существуют такие области *x*, где решение $\varphi(x) > 1$, то внутри этих областей необходимо положить $\varphi(x) = 1$, а для оставшихся точек пересчитать неопределенный множитель $\mu$ с учетом уже изменившихся условий (2) и (3). После этого в качестве оптимальной стратегии поиска может быть выбрана любая бинарная функция $u(x,t)$, удовлетворяющая соотношениям $\int u(x,t)dx = \varepsilon; \qquad \int u(x,\xi)d\xi = \varphi(x)t.$

В общем случае построение оптимального (не обязательно периодического) алгоритма одноэтапного поиска связано с нахождением такой функции $\varphi(x,t)$ – относительной нагрузки на точку *x* в момент времени *t*, которая минимизирует среднее время локализации

$$\langle\tau\rangle = \int dt\int dx f(x)\exp(-\lambda\int_0^t\varphi(x,\xi)d\xi)$$

при условии, что

$$0 \le \varphi(x,t) \le 1 \tag{5}$$

и для любого *t*

$$\int\varphi(x,t)dx = \varepsilon. \tag{6}$$

Чтобы упростить дальнейшие расчеты, введем функцию $\alpha(x,t) = \int_0^t\varphi(x,\xi)d\xi$, соответствующую суммарному времени пребывания точки *x* в окне обзора за весь период от начала поиска до момента времени *t*. Чтобы учесть ограничения (5) и (6), снова введем неопределенный множитель Лагранжа $\mu(t)$. Тогда задача построения оптимальной стратегии поиска сведется к нахождению функции $\alpha(x,t)$, минимизирующей функционал $\int dt\int dx\left[\exp(-\lambda\alpha(x,t)f(x) + \mu(t)\alpha(x,t)\right]$ при условии, что

$$\int_{-\infty}^{\infty}\alpha(x,t)dx=\varepsilon t, \qquad (7)$$
$$0\le\alpha(x,t)\le t.$$

Решением этой вариационной задачи является функция

$$\alpha(x,t)=\begin{cases}0, \dfrac{1}{\lambda}\ln\dfrac{\lambda f(x)}{\mu(t)}<0;\\ \dfrac{1}{\lambda}\ln\dfrac{\lambda f(x)}{\mu(t)}, 0\le\dfrac{1}{\lambda}\ln\dfrac{\lambda f(x)}{\mu(t)}\le t;\\ t, t<\dfrac{1}{\lambda}\ln\dfrac{\lambda f(x)}{\mu(t)},\end{cases} \qquad (8)$$

где множитель $\mu(t)$ определяется из соотношения (7), а в качестве оптимальной стратегии поиска $u(x,t)$ может быть выбрана любая бинарная функция, удовлетворяющая условиям

$$\int_0^t u(x,\xi)d\xi=\alpha(x,t); \qquad \int u(x,t)dx=\varepsilon.$$

Использование оптимальных алгоритмов поиска на практике сталкивается с определенными трудностями. Дело в том, что в тех случаях, когда априорная функция плотности вероятности распределения разыскиваемого источника отличается от равномерной, оба предложенных алгоритма оптимального одноэтапного поиска не могут быть физически реализованы посредством перемещения односвязного (неразрывного) сканирующего окна. Поэтому в реальных процедурах поиска целесообразно выполнять одноэтапную процедуру по следующей схеме. Предварительно интервал $(0,L)$ разбивается на ряд элементов-дискретов, имеющих длину $\varepsilon$, и априорно заданная плотность $f(x)$ «ступенчато» аппроксимируется на каждом из них. Значение $\varepsilon$ считается достаточно малым (что соответствует высоким требованиям к точности локализации), чтобы можно было пренебречь вариацией функции $f(x)$ внутри одного дискрета. Поиск должен начинаться с «наблюдения» самой высокой «ступеньки», в пределах которой амплитуда функция $f(x)$ максимальна, затем по истечении времени $t_1$ окно поочередно устанавливается под двумя самыми высокими «пиками», затем по истечении времени $t_2$ проводится поочередное наблюдение трех элементов и т.д. Все моменты переключения $t_i$ определяются в точном соответствии с приведенным выше соотношением (8), являющимся основой построения оптимальной стратегии поиска.

Следует отметить, что обсуждаемый алгоритм поиска предполагает, что интенсивность источника λ заранее известна. Если же такая априорная информация отсутствует, можно рекомендовать периодическую процедуру, не зависящую от этой интенсивности. В соответствии с этой стратегией первоначально вычисляются интегралы от плотности $f(x)$ в каждом $\varepsilon$-дискрете. Если общее количество $\varepsilon$-дискретов, на которые разбивается исходный интервал поиска, равно $m$, а интегралы-«ступеньки» над каждым из них равны $P_1, P_2,\ldots, P_m$, то окно просмотра должно циклически «пробегать» все дискреты с относительной нагрузкой

$$\beta_i=\sqrt{P_i}\Big/\sum_{j=1}^{m}\sqrt{P_j}\quad,\quad (j=1,\ldots,m).$$

Эти значения $\beta_i$ легко получить, если для минимизации среднего времени поиска вновь применить метод неопределенных множителей Лагранжа:

$$\langle\tau\rangle=(1/\lambda)(P_1/\beta_1+P_2/\beta_2+\ldots+P_m/\beta_m)\Rightarrow\min, \qquad (\beta_1+\ldots+\beta_m=1).$$

**Многоэтапные алгоритмы поиска.** Поскольку при высоких требованиях к точности локализации одноэтапные алгоритмы далеки от оптимальных, при построении оптимальной поисковой процедуры нельзя ограничиваться одноэтапными алгоритмами, а следует считать поисковую процедуру состоящей из нескольких этапов. Для среднего времени $m$-этапной процедуры локализации справедливо соотношение

Табл. 1. Параметры оптимальной процедуры поиска случайного равномерно распределенного на интервале $(0, L)$ импульсного источника в зависимости от требуемой точности локализации $\varepsilon$.

| $(\varepsilon / L)$ (требуемая точность локализации) | $m_{opt}$ (оптимальное количество этапов) | $W_m, m = \overline{1, m_{opt}}$ (поисковое окно системы на каждом из $m_{opt}$ этапов оптимального поиска) | $< \tau >$ (среднее время локализации) |
|---|---|---|---|
| $\frac{1}{4} \le (\varepsilon / L) < 1$ | 1 | $W_1 = \varepsilon$ | $\frac{1}{\lambda}(\varepsilon / L)^{-1}$ |
| $\left(\frac{2}{3}\right)^6 \le (\varepsilon / L) \le \frac{1}{4}$ | 2 | $W_1 = (\varepsilon / L)^{\frac{1}{2}} \times L$<br>$W_2 = (\varepsilon / L) \times L = \varepsilon$ | $\frac{2}{\lambda}(\varepsilon / L)^{-\frac{1}{2}}$ |
| $\left(\frac{3}{4}\right)^{12} \le (\varepsilon / L) \le \left(\frac{2}{3}\right)^6$ | 3 | $W_1 = (\varepsilon / L)^{\frac{1}{3}} \times L$<br>$W_2 = (\varepsilon / L)^{\frac{2}{3}} \times L$<br>$W_3 = (\varepsilon / L) \times L = \varepsilon$ | $\frac{3}{\lambda}(\varepsilon / L)^{-\frac{1}{3}}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $\left(\frac{m}{m+1}\right)^{m(m+1)} \le (\varepsilon / L) \le \left(\frac{m-1}{m}\right)^{m(m-1)}$ | $m$ | $W_1 = (\varepsilon / L)^{\frac{1}{m}} \times L$<br>$W_2 = (\varepsilon / L)^{\frac{2}{m}} \times L$<br>$\vdots$<br>$W_i = (\varepsilon / L)^{\frac{i}{m}} \times L$<br>$\vdots$<br>$W_m = (\varepsilon / L)^{\frac{m}{m}} \times L = \varepsilon$ | $\frac{m}{\lambda}(\varepsilon / L)^{-\frac{1}{m}}$ |

$$\langle \tau \rangle = \sum_{k=1}^{m} \lambda^k \int_0^{\infty} dx f(x) \int \ldots \int t_k \times$$
$$\times \left\{ \prod_{l=1}^{k} \left[ dt_l u_l(x, \sum_{s=1}^{l} t_s, t_1, \ldots, t_{l-1}) \exp(-\lambda \int_{\sum_{s=1}^{l-1} t_s}^{\sum_{s=1}^{l} t_s} u_l(x, \xi, t_1, \ldots t_{l-1}) d\xi) \right] \right\}, \quad (9)$$
$$\int u_m(x, t, t_1, \ldots, t_{m-1}) dx = \varepsilon,$$

где $u_i(x, t, t_1, \ldots, t_{i-1})$ – функция, задающая окно обзора на $i$-м этапе при условии, что интервалы, зафиксированные между предшествующими ($i$-1) импульсами, были соответственно $t_1, t_2, \ldots, t_{i-1}$. Найти экстремали, доставляющие минимум среднему времени локализации (9), в общем случае (когда плотность вероятности $f(x)$ произвольна) не всегда возможно. Поэтому нами была разработана универсальная процедура, ориентированная на поиск источника в том случае, когда не имеется никакой априорной информации об интенсивности источника (т.е. можно считать, что источник имеет равномерное распределение на интервале поиска). Ввиду ограниченности объема настоящего сообщения ниже приводится только результирующая таблица, обобщающая параметры оптимального многоэтапного поиска случайного равномерно распределенного точечного источника для систем с одним приемным устройством, а все необходимые аналитические выкладки и численные расчеты опущены. Более подробную информацию об этом можно найти в работах [9, 10].

Принимая во внимание, что $\lim_{m \to \infty} \left(\frac{m}{m+1}\right)^m = e^{-1}$, при $(\varepsilon / L) \to 0$ получаем следующие асимптотические соотношения для систем с одним приемным устройством:

$$m_{opt} \approx \ln(L / \varepsilon); \quad W_i \approx e^{-i} \times L, i = \overline{1, m_{opt}}; \quad < \tau >_{opt} \approx \frac{e \ln(L / \varepsilon)}{\lambda}.$$

**Заключение.** Приведенные выше результаты представляют собой набор оптимальных по быстродействию алгоритмов локализации случайного точечно-импульсного источника для

систем с одним приемным устройством. Исследования проводились как для одноэтапных поисковых процедур (в случае произвольной плотности вероятности распределения случайного источника на интервале поиска), так и для многоэтапных алгоритмов локализации (для тех случаев, когда случайный точечно-импульсный объект имеет равномерное распределение на поисковом интервале).

*Работа выполнена при финансовой поддержке РФФИ (гранты №№ 19-01-128 и 18-51-00001) и Министерства науки и высшего образования (проект № AAA-A17-117052410034-6).*

ЛИТЕРАТУРА

[1] *Barnes J.H., Hieftje G.M.* Recent Advances in Detector-Array Technology for Mass Spectrometry // Int. J. Mass Spectrom. 2004. V. 238. P. 33-46.

[2] *Kirichuk V.S., Mokin K.Yu., Reznik A.L.* Algorithms for Processing of Series of Digital Aerospace Images Based on Automatic Search for the Conjugate Points // Pattern Recognition and Image Analysis, 2001, vol.11, No 1, pp. 192-194.

[3] *Gnedenko B.V., Belyayev Yu.K., Solovyev A.D.* Mathematical Methods of Reliability Theory (Nauka, Moscow, 1965*;* Academic, New York, 1969).

[4] *Mallick M., Krishnamurthy V., Vo B.* Integrated Tracking, Classification, and Sensor Management: Theory and Applications. Wiley, 2012, 736 p.

[5] *Zhu X., Wen L., Hobbs G. et al.* Detection and localization of single-source gravitational waves with pulsar timing arrays // M.N. of the Royal Astronomical Society, Oxford Academic Press, v. 449, №2, 2015, p.1650-1663.

[6] *Powell M.J.D.* A Fast Algorithm for Nonlinearly Constrained Optimization Calculations. Numerical Analysis, G.A.Watson ed., Lecture Notes in Mathematics, Springer Verlag, Vol. 630, 1978.

[7] *Bellman R.E., Glicksberg I.L., Gross O.A.* Some aspects of the mathematical theory of control processes. Santa Monica, CA: RAND Corporation, 1958, 263 p.

[8] *Bertsekas D.* Constrained optimization and Lagrange multiplier methods. New York: Academic Press. 1982.

[9] *Reznik A.L., Solov'ev A.A., Torgov A.V.* The algorithms of optimum localization of random pulsed-point source under the condition of its uniform distribution on a search interval // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications, 2018, Vol. 28, No. 2, pp.354-361.

[10] *Reznik A.L., Tuzikov A.V., Solov'ev A.A., Torgov A.V.* Time-Optimal algorithms of searching for pulsed-point sources for systems with several detectors // Optoelectronics, Instrumentation and Data Processing, 2017, Vol. 53, Issue 3, pp. 203-209.

# КАРКАСНЫЙ ПОДХОД К РЕАЛИЗАЦИИ БИБЛИОТЕК ВЫСОКОПРОИЗВОДИТЕЛЬНОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

*Русин Е.В.*
Институт вычислительной математики и математической геофизики СО РАН, Новосибирск

Предлагается подход к реализации библиотек высокопроизводительной обработки изображений на основе фреймворка. Описывается реализация макетных библиотек для выполнения обработки на многопроцессорной ЭВМ кластерного типа и графическом ускорителе.
*Ключевые слова: дистанционное зондирование, обработка изображений, высокопроизводительные вычисления, параллельные вычисления, вычислительный кластер, GPGPU, библиотека подпрограмм, фреймворк.*

**Введение.** В последние годы характерными чертами дистанционного зондирования Земли являются, с одной стороны, повышение пространственного разрешения космических снимков, а с другой – использование гиперспектральной съемки с большим числом спектральных диапазонов. Синтезированные многоспектральные данные, получаемые при совмещении изображений из нескольких спектральных каналов, могут иметь объемы в сотни мегабайт. Если для обработки данных ДЗЗ на относительно небольших областях могут быть использованы обычные настольные компьютеры, то решение крупномасштабных и глобальных задач анализа геоданных требует уже подходов, основанных на высокопроизводительных технологиях [1]: широкое использование получают здесь параллельные, кластерные и распределенные вычисления, вычисления на спецпроцессорах таких как графические ускорители или Intel Xeon Phi. На сегодняшний день существует большое количество библиотек подпрограмм обработки изображений [2-5], часть которых поддерживает использование современных средств высокопроизводительных вычислений. Все они, однако, обладают существенным с нашей точки зрения недостатком – отсутствием расширяемости, т.е. возможности добавления в библиотеку дополнительных алгоритмов: множество реализуемых типичной библиотекой алгоритмов ограничено набором, определенным разработчиком. Данное обстоятельство сильно ограничивает возможность использования существующих библиотек в исследовательских проектах, направленных на создание новых методов и алгоритмов обработки спутниковых данных. В настоящей работе обсуждается «каркасный» подход к архитектуре библиотеки обработки изображений, позволяющий избежать этого ограничения.

**Каркасный подход.** Расширяемость является одним из главных качеств программных систем, построенных на основе архитектуры программного каркаса или фреймворка: библиотека обработки изображений, организованная в виде фреймворка, состоит из набора обобщенных подпрограмм, каждая из которых реализует целое семейство алгоритмов обработки (например, семейство алгоритмов типа «пиксель-пиксель»); такая обобщенная подпрограмма содержит код, общий для всех алгоритмов данного семейства. Конкретный алгоритм данного семейства реализуется классом/подпрограммой, вызываемой обобщенной подпрограммой и выполняющей, например, вычисление значения одного пикселя результата. Параметризуя соответствующую обобщенную подпрограмму различными классами/подпрограммами конкретных алгоритмов, пользователь получает готовые реализации алгоритмов обработки данного семейства. А если интерфейс взаимодействия обобщенной подпрограммы и класса/подпрограммы конкретного алгоритма известен (опубликован), то пользователь библиотеки может самостоятельно создавать конкретные классы/подпрограммы алгоритмов, расширяя тем самым библиотеку. Важным вопросом при проектировании подобной системы является ее производительность: накладные расходы на рост уровня абстракции вычислений не должны быть значительными.

На основе данных соображений были созданы две макетных библиотеки высокопроизводительной обработки изображений.

**Библиотека ParImProLib.** Библиотека ParImProLib [6] обеспечивает выполнение операций обработки изображений на многопроцессорных системах кластерного типа. При ее создании использовались следующие методологические принципы и обусловленные ими технические решения:

1. Максимальное сокрытие использования параллелизма от пользователей библиотеки: программа, использующая библиотеку, должна выглядеть как обычная «последовательная» программа: весь код, отвечающий за выполнение операций параллельной среды (инициализация/деинициализация среды, самоидентификация узла среди всех выполняющих узлов, межузловые пересылки и синхронизация) находится в реализации библиотеки. Вместе с тем, полное сокрытие параллелизма от пользователя привело бы к потенциальной неэффективности программных систем, основанных на композиции алгоритмов (для одного алгоритма наиболее эффективен один способ распараллеливания, для другого – второй, а для их композиции – третий, возможно совпадающий с первым или вторым). Кроме того, пользователь, как правило, в состоянии указать некоторые общие соображения об эффективном распараллеливании алгоритма (например, если распараллеливаемый алгоритм – композиция трех фильтраций с ядром 5×5 пикселей, то оптимальное распараллеливание достигается при разрезании изображения на полосы с перекрытием в шесть пикселей). Поэтому библиотека оставляет пользователю возможность указать представление изображения на наборе исполняющих узлов для наиболее эффективного распараллеливания вычислительного алгоритма:
   - Полная копия изображения на каждом обрабатывающем узле.
   - Разрезание изображения на непересекающиеся горизонтальные полосы: каждый узел получает для обработки «свою» полосу.
   - Разрезание на горизонтальные полосы с перекрытием соседних полос в заданное количество пикселей.
2. Реализация в виде фреймворка, что позволяет избежать дублирования программного кода, реализующего параллельные алгоритмы, и обеспечить расширяемость библиотеки.
3. Код библиотеки должен минимизировать накладные расходы, неизбежные при росте уровня абстракции модели вычислений. Для этих целей в качестве средств реализации выбран язык C++, обеспечивающий возможность эффективного обобщения операций с помощью механизма шаблонов (templates).
4. Переносимость библиотеки на максимально широкий круг ЭВМ кластерного типа, что обеспечивается выбором в качестве средства разработки языка C++ и стандарта параллельного программирования систем с разделенной памятью MPI.

Базовыми элементами библиотеки являются следующие классы:

`Image` – класс, реализующий абстракцию изображения. Среди прочих ответственностей, прозрачно для пользователя поддерживает актуальность данных изображения на обрабатывающих узлах при распределенном представлении.

`NeighborhoodManipulator` – класс, реализующий абстракцию манипулятора окрестностью пикселя. Позволяет реализовывать алгоритмы обработки, абстрагируясь от конкретных параметров изображения (физических размеров, представления на обрабатывающих узлах и пр.), в терминах окрестности обрабатываемого пикселя. Использование данной абстракции делает легким процесс расширения библиотеки путем создания совместимых с библиотекой новых функций обработки. Следующий фрагмент кода на C++ демонстрирует класс `SobelFilter` – совместимую с библиотекой реализацию фильтра Собела:

```
1. class SobelFilter {
2. public:
3.    unsigned LeftMargin()   const { return 1; }
4.    unsigned RightMargin()  const { return 1; }
5.    unsigned TopMargin()    const { return 1; }
6.    unsigned BottomMargin() const { return 1; }
7.
8.    unsigned operator()(const NeighborhoodManipulator& n) {
9.       return abs(
```

```
10.             1 * n.PixelRelativeToCurrent(-1, -1)
11.           + 2 * n.PixelRelativeToCurrent(-1,  0)
12.           + 1 * n.PixelRelativeToCurrent(-1,  1)
13.           - 1 * n.PixelRelativeToCurrent( 1, -1)
14.           - 2 * n.PixelRelativeToCurrent( 1,  0)
15.           - 1 * n.PixelRelativeToCurrent( 1,  1))
16.          + abs(
17.             1 * n.PixelRelativeToCurrent(-1, -1)
18.           + 2 * n.PixelRelativeToCurrent( 0, -1)
19.           + 1 * n.PixelRelativeToCurrent( 1, -1)
20.           - 1 * n.PixelRelativeToCurrent(-1,  1)
21.           - 2 * n.PixelRelativeToCurrent( 0,  1)
22.           - 1 * n.PixelRelativeToCurrent( 1,  1)));
23.   }
24. };
```

- Строки 3-6 определяют границы окрестности, влияющей на результат в пикселе, что необходимо для выполнения операции в пикселях на границах изображения и на границах полос при «распределенном» представлении изображения. Для фильтра Собела все четыре границы – единицы, так как он использует только ближайших соседей.
- Строки 8-23 описывают фильтрацию в терминах обработки одного пикселя, параметром этой операции является манипулятор окрестности, предоставляющий посредством метода `PixelRelativeToCurrent(i, j)` доступ к пикселям, соседним обрабатываемому (смещенных относительно обрабатываемого пикселя на `i` пикселей по горизонтали и `j` по вертикали.

Применение данного фильтра к изображению демонстрируется следующим фрагментом кода:

```
1. Image im("image.jpg", PartitioningInfo pi(CutWithOverlap, 1));
2. im.DoNeighborhoodToPixelOperation(SobelFilter());
```

- Строка 1 считывает изображение из дискового файла и распределяет его между обрабатывающим узлами в виде полос с перекрытием в один пиксель.
- Строка 2 применяет к изображению нему обобщенную операцию обработки типа «окрестность-пиксель», параметризуя ее объектом класса, реализующего фильтр Собела. Обобщенная операция параллельно по полосам вызовет оператор обработки последовательно для каждого пикселя полосы.
- Вообще говоря, перекрытие полос, заданное в строке 1, не является строго обязательным, так как код строки 2 при необходимости подгрузит в каждом узле недостающие для выполнения обработки данные соседних полос с других узлов. Однако оно имеет смысл как оптимизация кода, которую может выполнить пользователь библиотеки, обладая априорным знанием о выполняемых в последствии операциях.

Характерным и важным моментом при использовании и расширении библиотеки является отсутствие для пользователя необходимости знать модель MPI и понимать особенности межузлового обмена в кластере. Предлагаемый подход показал свою эффективность:

1. Реализация алгоритма поиска кольцевых структур на аэрокосмоснимках [7] с помощью библиотеки оказалась медленнее реализации «с нуля» на C++ и MPI с использованием непосредственного доступ к пикселям, не более чем на 10 процентов. При этом процесс создания параллельной программы с использованием разработанного фреймворка оказывается значительно проще.
2. Эффективность реализации алгоритма поиска кольцевых структур с помощью библиотеки распараллеливания алгоритма поиска кольцевых структур составляет около 95% при исполнении на 8 узлах кластера НКС-30Т+GPU Сибирского Суперкомпьютерного Центра (ССКЦ) ИВМиМГ СО РАН [8].

**Библиотека SSCCIPGPU.** В последние годы появился большой практический интерес к использованию современных графических процессоров (Graphics Processing Unit, GPU) в ка-

честве универсального вычислителя. Вообще говоря, GPU ориентирован на эффективное решения задач машинной графики, в частности он содержит аппаратные функции, позволяющие эффективно (с производительностью в сотни ГФлопс) выполнять массовые вычисления (однотипные операции над большим объемом данных). Эти возможности позволяют использовать GPU в задачах, не связанных с визуализацией, но также опирающихся на массовые вычисления, например в задачах обработки и анализа изображений. На ряде практических задач вычисления на GPU обеспечили 70-кратное ускорение по сравнению с вычислениями на центральном процессоре, что соответствует временным результатам, получаемым на суперЭВМ. Концепция универсальных вычислений на GPU получила также поддержку производителей графических ускорителей (технология CUDA от NVIDIA), что делает доступным написание программ, использующих GPU, на языке высокого уровня без знаний архитектуры сопроцессора. Это, а также невысокая стоимость современных GPU, делает их популярным оборудованием современных суперкомпьютерных центров: так, основной на текущий момент вычислительной мощностью ССКЦ ИВМиМГ СО РАН является гибридный кластер НКС-30Т+GPU, включающий 40 узлов, каждый из которых оснащен тремя GPU NVIDIA Tesla M 2090 на архитектуре Fermi (compute capability 2.0) с 512 ядрами, 6 Гбайт памяти GDDR5. Все это делает актуальной разработку программных средств вовлечения GPU в процессы обработки и анализа ДЗЗ.

Принципы и технические решения, положенные в основу разработки библиотеки SSCCIPGPU [9] во многом аналогичны вышеприведенным принципам, на которых создавалась более ранняя библиотека ParImProLib. Добавим только, что:

1. В качестве инструментальных средств были выбраны C++ и CUDA, что сделало решение максимально переносимым, а также позволило эффективно реализовать расширяемую архитектуру фреймворка.
2. Многоуровневая иерархия памяти GPU делает желательным возможность задания размещения изображений и параметров обработки в глобальной, текстурной и константной памяти GPU.
3. Сложность модели программирования графического процессора делает желательной возможность выбора между использованием синхронных или асинхронных вызовов CUDA, а также способов распределения вычислений между несколькими GPU одного вычислительного узла (multi-GPU).

Следующий фрагмент кода демонстрирует класс `SobelFilter_GPU` – GPU-реализацию фильтра Собела:

```
1. class SobelFilter_GPU {
2. public:
3.     __host__ unsigned LeftMargin()   const { return 1; }
4.     __host__ unsigned RightMargin()  const { return 1; }
5.     __host__ unsigned TopMargin()    const { return 1; }
6.     __host__ unsigned BottomMargin() const { return 1; }
7.
8.     __host__ ParamsForDevice ExportParamsForDevice() const { return ParamsForDevice(); }
9.
10.    __device__ unsigned operator()(const NeiborhoodManupulator& n, const byte* /*params*/) {
11.       return abs(
12.                 1 * n.PixelRelativeToCurrent(-1, -1)
13.               + 2 * n.PixelRelativeToCurrent(-1,  0)
14.               + 1 * n.PixelRelativeToCurrent(-1,  1)
15.               - 1 * n.PixelRelativeToCurrent( 1, -1)
16.               - 2 * n.PixelRelativeToCurrent( 1,  0)
17.               - 1 * n.PixelRelativeToCurrent( 1,  1))
18.             + abs(
19.                 1 * n.PixelRelativeToCurrent(-1, -1)
20.               + 2 * n.PixelRelativeToCurrent( 0, -1)
21.               + 1 * n.PixelRelativeToCurrent( 1, -1)
22.               - 1 * n.PixelRelativeToCurrent(-1,  1)
23.               - 2 * n.PixelRelativeToCurrent( 0,  1)
24.               - 1 * n.PixelRelativeToCurrent( 1,  1)));
25.    }
26.};
```

- Реализация класса для GPU в основном идентична соответствующей реализации для кластера с поправкой на использование директив CUDA (`__host__` для кода, выполняющегося на главном узле, `__device__` – на GPU).
- Строка 7 определяет метод `ExportParamsForDevice`, служащий для сериализации (перевода в последовательность байтов) параметров алгоритма. Метод необходим для дальнейшего переноса фреймворком параметров алгоритма в память GPU. В случае фильтра Собела реализация метода тривиальна, так как он не имеет параметров.
- Строка 10 к описанному выше параметру манипулятора окрестностью добавляет в оператор обработки параметр `params` – последовательность байтов, перенесенная фреймворком в память GPU.
- Реализация оператора обработки (строки 10-25), как правило, дессериализует (восстанавливает) параметры алгоритма из блока памяти `params` (этот шаг отсутствует для тривиального фильтра Собела) и выполняет обработку единичного пикселя с использованием манипулятора окрестности.

Применение данного фильтра к изображению демонстрируется следующим фрагментом кода:

```
1. Image im("image.jpg");
2. im.DoNeighborhoodToPixelOperation<Global1D, Synchronous>(SobelFilter_GPU());
```

практически аналогичному коду для кластера, за исключением того, что строка 2 содержит указание для «одномерного» размещения изображения в глобальной памяти GPU, а также для использования при обработке синхронного API CUDA. Снова особо отметим здесь отсутствие необходимости знать модель программирования GPU при использовании и расширении библиотеки.

Аналогичный подход используется и для реализации обработки multi-GPU, при этом фреймворку необходимо указать стратегию распараллеливания вычисления между графическими ускорителями (использование потоков операционной системы или переключение активных устройств GPU в рамках одного потока).

Об эффективности подхода говорит следующее:

1. Реализация алгоритма поиска кольцевых структур с помощью библиотеки не уступила по быстродействию реализации «с нуля» на C++ и CUDA. При этом процесс создания высокопроизводительной программы с использованием разработанного фреймворка значительно проще.
2. Реализация алгоритма поиска кольцевых структур с помощью библиотеки при исполнении на одном GPU NVIDIA Tesla M 2090 узла кластера НКС-30Т+GPU оказалась примерно в 90 раз быстрее аналогичной реализации на его центральном процессоре Intel Xeon X5670 (2,93 ГГц).

**Заключение.** Полученные результаты свидетельствуют о перспективности каркасного подхода при реализации высокопроизводительных библиотек обработки изображений. В настоящее время созданные макетные библиотеки используются при разработке экспериментального фреймворка высокопроизводительных технологий обработки изображений SSCCIP [10], являющимся одним из элементов экспериментального облачного фреймворка [11]. В дальнейших планах авторов – распространение подхода на обработку изображений использованием процессоров типа Intel Xeon Phi, которыми оснащен новый кластер ССКЦ ИВМиМГ СО РАН НКС-1П.

*Работа выполнена в рамках государственного задания ИВМиМГ СО РАН (проект 0315-2016-0003) с использованием ресурсов ЦКП Сибирский Суперкомпьютерный Центр ИВМиМГ СО РАН.*

ЛИТЕРАТУРА


[1] *Buchnev A., Pyatkin V., Rusin E.V.* Software Technologies for Processing of Earth Remote Sensing Data // Pattern Recognition and Image Analysis. 2013. Vol. 23. No. 4. P. 474-480.

[2] OpenCV. **https://opencv.org** (дата обращения 09.06.2019).

[3] TensorFlow. **https://www.tensorflow.org** (дата обращения 09.06.2019)

[4] Computer Vision Toolbox Documentation. **https://www.mathworks.com/help/vision** (дата обращения 09.06.2019).

[5] Intel® Integrated Performance Primitives (Intel® IPP) | Intel® Software. **https://software.intel.com/intel-ipp** (дата обращения 09.06.2019).

[6] *Rusin E.V.* Object-Oriented Parallel Image Processing Library // Parallel Computing Technologies. PaCT 2009. Lecture Notes in Computer Science. 2009. Vol. 5698. P. 344-349.

[7] *Alekseev A.S, Pyatkin V.P., Salov G.I.* Crater Detection in Aero-space Imagery Using Simple Nonparametric Statistical Tests // Lecture Notes in Computer Science. 1993. Vol 179. P. 793-799.

[8] ЦКП ССКЦ ИВМиМГ СО РАН. **http://www.sscc.icmmg.nsc.ru/hardware.html** (дата обращения 09.06.2019).

[9] *Русин Е.В.* Технологии обработки данных дистанционного зондирования Земли на гибридном кластере НКС-30Т+GPU // Интерэкспо Гео-Сибирь. 2016. Т. 4. № 1. С. 46-49.

[10] *Rusin E.V.* SSCCIP – A Framework for Building Distributed High-Performance Image Processing Technologies // Parallel Computing Technologies. PaCT 2011. Lecture Notes in Computer Science. 2011. Vol. 6873. P. 467-472.

[11] *Buchnev A., Kim P., Pyatkin V., Pyatkin F., Rusin E.* Framework of cloud web services for processing remote sensing data // E3S Web of Conferences. 2019. Vol. 75. Paper 03001.

# АЛГОРИТМ ВОССТАНОВЛЕНИЯ КОЭФФИЦИЕНТОВ ОТРАЖЕНИЯ ЗЕМНОЙ ПОВЕРХНОСТИ ПО ДАННЫМ MODIS С УЧЕТОМ НЕОДНОРОДНОСТИ ЗЕМНОЙ ПОВЕРХНОСТИ

*Тарасенков М.В., Белов В.В., Зимовая А.В., Энгель М.В.*
Институт оптики атмосферы им. В.Е. Зуева СО РАН, Томск

Рассматривается алгоритм атмосферной коррекции для восстановления коэффициентов отражения земной поверхности с учетом ее неоднородности, бокового подсвета, дополнительной освещенности отраженным излучением и поляризации излучения.
*Ключевые слова: атмосферная коррекция, коэффициент отражения земной поверхности, поляризация, метод Монте – Карло.*

**Введение.** Спутниковая информация о коэффициентах отражения земной поверхности имеет широкий круг применений. Вместе с тем коэффициенты отражения представляют собой результат косвенных измерений и могут содержать существенные погрешности. Чем меньше создаваемая алгоритмом восстановления погрешность, тем шире область применения спутниковых данных. Как итог во многих ситуациях необходима атмосферная коррекция спутниковых данных. Одним из направлений выполнения атмосферной коррекции является использование RTM-алгоритмов. Подход состоит в том, что строится математическая модель переноса солнечного излучения в системе атмосфера-земная поверхность и формирования принимаемого сигнала. Далее путем решения обратной задачи определяется искомое распределение коэффициентов отражения по земной поверхности.

В общем случае интенсивность принимаемого сигнала состоит из: а) интенсивности нерассеянного излучения от наблюдаемого участка земной поверхности $I_0$, б) интенсивности излучения, рассеянного в атмосфере и не взаимодействовавшее с земной поверхностью $I_{sun}$, в) интенсивности излучения, отраженного соседними участками земной поверхности и далее рассеянного в атмосфере в направлении приемника (боковой подсвет) $I_{surf}$. Кроме того, земная поверхность освещается не только прямым и диффузным излучением Солнца, но и кратно (в общем случае) отраженным от земной поверхности излучением после рассеяния в атмосфере $E_1$,$E_2$ и т.д.

Однако в зависимости от рассматриваемых ситуаций часть из указанных величин оказывает слабое влияние на результат решения обратной задачи. Поэтому для ускорения получения решения в различных алгоритмах атмосферной коррекции строятся упрощенные модели переноса излучения в системе атмосфера-земная поверхность. Так в одной из первых работ [1] перенос излучения учитывался в приближении однократного рассеяния излучения и однородной земной поверхности. В более поздних работах (наприме, [2]) использована более адекватная модель переноса оптического излучения, но атмосферная коррекция выполнялась в приближении однородной земной поверхности. В работах подобных [3] обратная задача выполняется в два этапа: сначала определяется коэффициент отражения в приближении однородной поверхности, а затем приближенно учитывается влияние бокового подсвета. При этом предполагается, что дополнительная освещенность формируется однородной поверхностью. В работе [4] строго учитывается влияние бокового подсвета, а дополнительная освещенность отраженным излучением определяется в приближении однородной поверхности. В наших работах [5, 6] мы предлагаем алгоритм, учитывающий неоднородность земной поверхности при формировании бокового подсвета и дополнительной ее освещенности. Это позволяет расширить возможности алгоритма коррекции на ситуации с высокой мутностью атмосферы, сложным неоднородным распределением коэффициента отражения по земной поверхности.

Помимо этого, существенными факторами при атмосферной коррекции могут выступать рельеф поверхности, ее неламбертовость, поляризация излучения и влияние облачных полей на изображение безоблачных участков. Задача оценки влияния рельефа на погрешность вос-

становления отражательных свойств земной поверхности решалась в [7, 8]. Учет неламбертовости рассматривался в [4, 9, 10]. Эти факторы мы планируем рассмотреть в дальнейших работах. Исследования влияния облачных полей на безоблачные участки начато нами в работе [11]. Влияние поляризации оценивалось авторами работ [12-14] для решения прямой задачи переноса солнечного излучения в системе атмосфера-земная поверхность. По оценкам этих авторов поляризация может вносить погрешность до 10% в интенсивность принимаемого излучения. В наших работах [15, 16] показано, что при восстановлении коэффициентов отражения некоторых поверхностей неучет поляризации может приводить к абсолютным погрешностям, превышающим значение восстанавливаемого коэффициента отражения. Следовательно, для участков, например, покрытых слабоотражающей растительностью, требуется учет влияния поляризации излучения. Рассмотрим далее модифицированный алгоритм восстановления коэффициентов отражения с учетом поляризации излучения и его апробацию на примере снимков прибора MODIS.

**Алгоритм атмосферной коррекции с учетом поляризации излучения.** Разработанный алгоритм коррекции создавался при следующих допущениях: атмосфера сферическая, разделенная на однородные слои (32 слоя); атмосфера является рассеивающей и поглощающей средой; земная поверхность ламбертовская; рельеф не учитывается; атмосфера при рассеянии излучения частично поляризует излучение.

Разработанный алгоритм стал основой комплекса программ, блок-схема которого приведена на рис. 1.

Пошаговая процедура атмосферной коррекции сводится к следующему.

1) Определяется прямой коэффициент пропускания $T_i$ трассы от наблюдаемого пикселя до приемной системы.

2) Методом Монте-Карло выполняется расчет интенсивности излучения $I_{sun}$, не взаимодействовавшего с земной поверхностью, с учетом поляризации излучения для 30 узловых направлений. Исходя из этих результатов, определяются приближенные значения $I_{sun,i}$ по формуле:

$$I_{sun,i} = -\frac{B_i + \sqrt{B_i^2 - 4A_i C_{13}}}{2A_i \mu_{d,i}} \qquad (1)$$

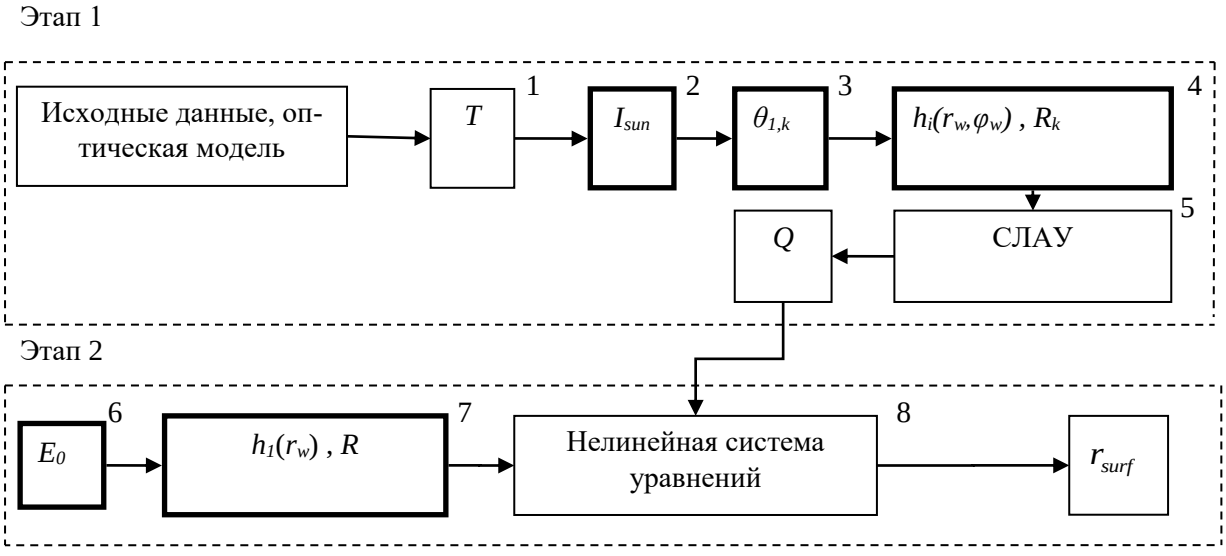



Рис. 1. Блок-схема комплекса программ. Блоки, в которых используются алгоритмы метода Монте-Карло, обозначены жирной линией.

$$A_i = \begin{cases} C_{11}\mu_{d,i}^2 + C_{21}\left(\sqrt{1-\mu_{d,i}^2}cos\varphi_i\right)^2 + C_{22}\mu_{d,i}\sqrt{1-\mu_{d,i}^2}cos\varphi_i - \left(\sqrt{1-\mu_{d,i}^2}sin\varphi_i\right)^2, \varphi_i \le 90^0 \\ C_{11}\mu_{d,i}^2 + C_{31}\left(\sqrt{1-\mu_{d,i}^2}cos\varphi_i\right)^2 + C_{32}\mu_{d,i}\sqrt{1-\mu_{d,i}^2}cos\varphi_i - \left(\sqrt{1-\mu_{d,i}^2}sin\varphi_i\right)^2, \varphi_i > 90^0 \end{cases} \quad (2)$$

$$B_i = \begin{cases} C_{12}\mu_{d,i} + C_{23}\sqrt{1-\mu_{d,i}^2}cos\varphi_i, \varphi_i \le 90^0 \\ C_{12}\mu_{d,i} + C_{33}\sqrt{1-\mu_{d,i}^2}cos\varphi_i, \varphi_i > 90^0 \end{cases} \quad (3)$$

где $\mu_{d,i}$ – косинус зенитного угла приемника при наблюдении $i$-го пикселя; $\varphi_i$ – азимут между направлениями на Солнце и на приемник в $i$-м пикселе; $C_{11}$, $C_{12}$, $C_{13}$, $C_{21}$, $C_{22}$, $C_{23}$, $C_{31}$, $C_{32}$, $C_{33}$ – константы аппроксимации, определяемые методом наименьших квадратов (МНК) по узловым значениям $I_{sun}$ для фиксированного $\theta_{sun}$.

Алгоритм расчета $I_{sun}$ с учетом поляризации и его тестирование описаны в [15,16].

3) Методом Монте-Карло вычисляются узловые значения интеграла функции размытия точки (ФРТ) канала формирования бокового подсвета $I_{dif}$. Исходя из этих результатов, определяются границы изопланарных зон $\theta_{1,k}$, (области на земной поверхности, в пределах которых можно с заданной погрешностью $\delta$ использовать одну и ту же ФРТ), по критерию:

$$\begin{cases} I_{dif}(\mu_k) = I_{dif}(1) - A(1-\mu_k)^N \\ \mu_{k+1} = 1 - \left[\frac{1}{A}\left(I_{dif}(1) - \frac{I_{dif}(\mu_k)}{1+\delta}\right)\right]^{1/N} \end{cases} \quad (4)$$

$$\delta \equiv \frac{I_{dif}(\mu_k) - I_{dif}(\mu_{k+1})}{I_{dif}(\mu_{k+1})} \quad (5)$$

$$I_{dif}(\mu) \equiv \int_S h(\mu, r_w, \varphi_w) dS \quad (6)$$

где $\mu_k = cos\theta_{1,k}$ задает границу между $k$-й и $k$+1-й изопланарными зонами; $h(\mu, r_w, \varphi_w)$ – ФРТ канала формирования бокового подсвета; $r_w$ – поверхностное расстояние от центра наблюдаемого пикселя на земной поверхности до точки на земной поверхности; $\varphi_w$ – азимутальный угол на земной поверхности между направлением на проекцию приемной системы на земную поверхность и направлением на рассматриваемую точку из наблюдаемого пикселя; $S$ – вся площадь земной поверхности; $\delta$ – предельно допустимый уровень погрешности при использовании ФРТ, соответствующей $\mu_k$ вместо ФРТ, соответствующей $\mu_{k+1}$ (в расчетах использовалась $\delta$=0.05 ); $A$, $N$ – константы аппроксимации, определяемые МНК.

4) Определяется радиус области бокового подсвета $R_k$, вне которого с заданной погрешностью $\delta_1$ можно считать влияние бокового подсвета нулевым. Можно показать, что для выполнения для произвольной неоднородной поверхности условия:

$$1 \ge \min_i \frac{Q_i}{\widetilde{Q}_i} \ge \delta_1 \quad (7)$$

где $Q_i \equiv r_{surf,i}E_{sum,i}$ – некоторое точное значение светимости земной поверхности в $i$-м пикселе; $\widetilde{Q}_i$ – приближенное значение светимости, которое получится, если считать, что вне $R_k$ боковой подсвет отсутствует; $\delta_1$ – величина, характеризующая максимальную погрешность, создаваемую использованием радиуса бокового подсвета (в расчетах $\delta_1 = 0.95$), достаточно выполнения условия:

$$f_1(R_k) \equiv \frac{\iint_{S(R_k)} h(\mu_k, r_w, \varphi_w) ds}{\iint_S h(\mu_k, r_w, \varphi_w) ds} \ge \delta_1 + (\delta_1 - 1)\frac{\frac{1}{\pi}T_k}{\iint_S h(\mu_k, r_w, \varphi_w) ds} \quad (8)$$

где $R_k$ – радиус бокового подсвета для $k$-й ФРТ, рассчитываемой для углов (4); $T_k$ – прямое пропускание, соответствующее границе $k$-й изопланарной зоны; $S(R_k)$ – площадь на земной поверхности в пределах радиуса $R_k$.

В [6] для рассматриваемых далее каналов MODIS для $0.1 \le AOT_{0.55} \le 5$, произвольного неизвестного распределения коэффициентов отражения по земной поверхности и разнообразных ситуаций оценивались сверху значения $R_k$, для которых выполнено условие (8). Значение $R_k$

лежит в пределах $3 \le R_k \le 40$ км в зависимости от $\lambda$, АОТ и расположения Солнца и приемной системы.

Для каждой из изопланарных зон $k$ методом Монте-Карло вычисляется ФРТ канала формирования бокового подсвета $h(\mu_k, r_w, \varphi_w)$ в пределах $R_k$.

5) Методом Зейделя решается система линейных алгебраических уравнений (СЛАУ) для определения распределения светимости земной поверхности $Q_i$:

$$I_{sum,i} = I_{sun,i} + \sum_{j=1}^{N_i} A_{i,j} Q_j + A_{out,i} \overline{Q_\iota} \quad i = \overline{1,N} \tag{9}$$

$$A_{i,j} \approx \begin{cases} \frac{1}{\pi} T_i + h\left(\mu_{k_i}, 0, 0\right) S_i \,, i = j \\ h\left(\mu_{k_i}, r_{w,j}, \varphi_{w,j}\right) S_j \,, i \neq j \end{cases} \tag{10}$$

$$A_{out,i} = I_{dif}\left(\mu_{k_i}\right) - \sum_{j=1}^{N_i} A_{i,j} \tag{11}$$

$$\overline{Q_\iota} = \frac{I_{sum,i} - I_{sun,i}}{\frac{1}{\pi} T_i + I_{dif}\left(\mu_{d,i}\right)} \tag{12}$$

где $I_{sum,i}$ – интенсивность суммарного, принимаемого спутниковой системой излучения; $N$ – количество пикселей рассматриваемого участка; $N_i$ – количество пикселей в пределах радиуса $R_k$ от $i$-го пикселя; $\mu_{k_i}$ – граница $k$-й изопланарной зоны, в которую попадает $i$-й пиксель; $\overline{Q_\iota}$ – величина, приближенно оценивающая светимость поверхности вне рассматриваемой области; $S_i$ – площадь $i$-го пикселя.

Решение системы уравнений (9) для всего рассматриваемого участка позволяет учесть влияние неоднородности земной поверхности.

6) Методом Монте-Карло вычисляется освещенность земной поверхности без учета переотражений $E_0$.

7) Вычисляется радиус области формирования дополнительной освещенности однократно отраженным излучением $R$, вне которой можно считать дополнительную освещенность равной 0. Можно показать, что для выполнения для произвольной неоднородной поверхности условия:

$$1 \ge \min_j \frac{r_{surf,j}}{\tilde{r}_{surf,j}} \ge \delta_2 \tag{13}$$

где $r_{surf,i}$ – некоторый коэффициент отражения $i$-го пикселя; $\widetilde{r_{surf,\iota}}$ – приближенное значение коэффициента отражения, получаемое при использовании радиуса области формирования дополнительной освещенности; $\delta_2$ – величина, характеризующая максимальную погрешность, создаваемую использованием радиуса формирования дополнительной освещенности (в расчетах $\delta_2 = 0.95$), достаточно выполнения условия:

$$f_2(R) \equiv \frac{\iint_{S(R)} h_1(r_w) ds}{\gamma_1} \ge \frac{\delta_2}{\gamma_1}\left((\delta_2 - 1) + \delta_2 \gamma_1\right) \tag{14}$$

$$\gamma_1 = \iint_S h_1(r_w) ds \tag{15}$$

где $h_1(r_w)$ – значение ФРТ канала формирования дополнительной освещенности.

В [6] для рассматриваемых далее каналов MODIS, для $0.1 \le AOT_{0.55} \le 5$ и разнообразных ситуаций оценивались значения $R$, для которых выполнено условие (14). Значение $R$ лежит в пределах $0 \le R \le 15$ км.

Методом Монте – Карло рассчитывается ФРТ канала формирования дополнительной освещенности земной поверхности отраженным в системе атмосфера-земная поверхность излучением $h_1(r_w)$ в пределах радиуса $R$.

8) Методом Ньютона с итерационным решением вспомогательных СЛАУ методом Зейделя решается нелинейная система уравнений относительно $r_{surf}$ вида:

$$\frac{Q_i}{E_0} = r_{surf,i}\left(1 + \sum_{j=1}^{M_i} C_{i,j} r_{surf,j} + C_{out,i} \overline{r_{surf,\iota}} + \frac{\left(\overline{r_{surf,\iota}} \gamma_1\right)^2}{1 - \overline{r_{surf,\iota}} \gamma_1}\right) \tag{16}$$

$$C_{i,j} \approx h_1\left(r_{w,ij}\right) S_j \tag{17}$$

$$C_{out,i} = \gamma_1 - \sum_{j=1}^{M_i} C_{i,j} \tag{18}$$

$$\overline{r_{surf,i}} = \frac{\overline{Q_i}/E_0}{1+\gamma_1\overline{Q_i}/E_0} \quad (19)$$

где $r_{surf,i}$ – коэффициент отражения $i$-го пикселя изображения; $\overline{r_{surf,i}}$ – коэффициент отражения $i$-го пикселя, полученный в приближении однородной земной поверхности; $M_i$ – количество пикселей, попадающих в область радиуса $R$ вокруг $i$-го пикселя.

Решение нелинейной системы (16) позволяет учесть влияние неоднородности земной поверхности на ее освещенность. Выполненный анализ показывает, что для предельной из рассмотренных ситуации $AOT_{0.55}$=5, $r_{surf}$≤0.4 погрешность использования (16) не более 3%. В то время как неучет влияния неоднородности земной поверхности при формировании дополнительной освещенности в этой ситуации приводит к погрешностям в пределах 19%. Для молекулярной атмосферы ($AOT_{0.55}$=0) и $r_{surf}$ ≤0.4 погрешность использования (16) не превышает 1%, а однородного приближения (19) – 10.6%.

Анализ условий сходимости систем уравнений (9) и (16) в работе [6] показывает, что система (9) сходится при любом размере пикселя при AOT≤1, а при разрешении в надире в 1 км при AOT≤4. Система (16) сходится при всех рассмотренных в [6] ситуациях ($0.1 \le AOT_{0.55} \le 5$).

**Апробация алгоритма.** Для апробации и оценки погрешности алгоритма восстановления коэффициентов отражения с учетом влияния поляризации были рассмотрены снимки MODIS для 4 каналов: 1 - центр канала λ=0.649 мкм, 2 - λ=0.860 мкм, 3 - λ=0.469 мкм, 4 - λ=0.555 мкм и 8 - λ=0.412 мкм, пространственное разрешение - 1000 м. Рассматривались три тестовые области: 1) участок юга Томской области ($55.95^0$ – $56.85^0$ с.ш. и $84.05^0$ – $84.95^0$ в.д.) 7 снимков с 17.06.2012 по 23.06.2012 г., 2) участок Московской области ($55.72^0$-$55.95^0$ с.ш. и $37.56^0$-$38.10^0$ в.д.) 5 снимков с 06.05.2017 по 07.05.2017 г. и 3) участок Иркутской области ($51.42^0$ -$52.67^0$ с.ш. и $103.64^0$ -$105.47^0$ в.д.) 4 снимка с 20.06.2017г по 21.06.2017г. Результаты предлагаемого алгоритма с учетом и без учета поляризации сопоставлялись с результатами алгоритма MOD09 и результатами, полученными без атмосферной коррекции.

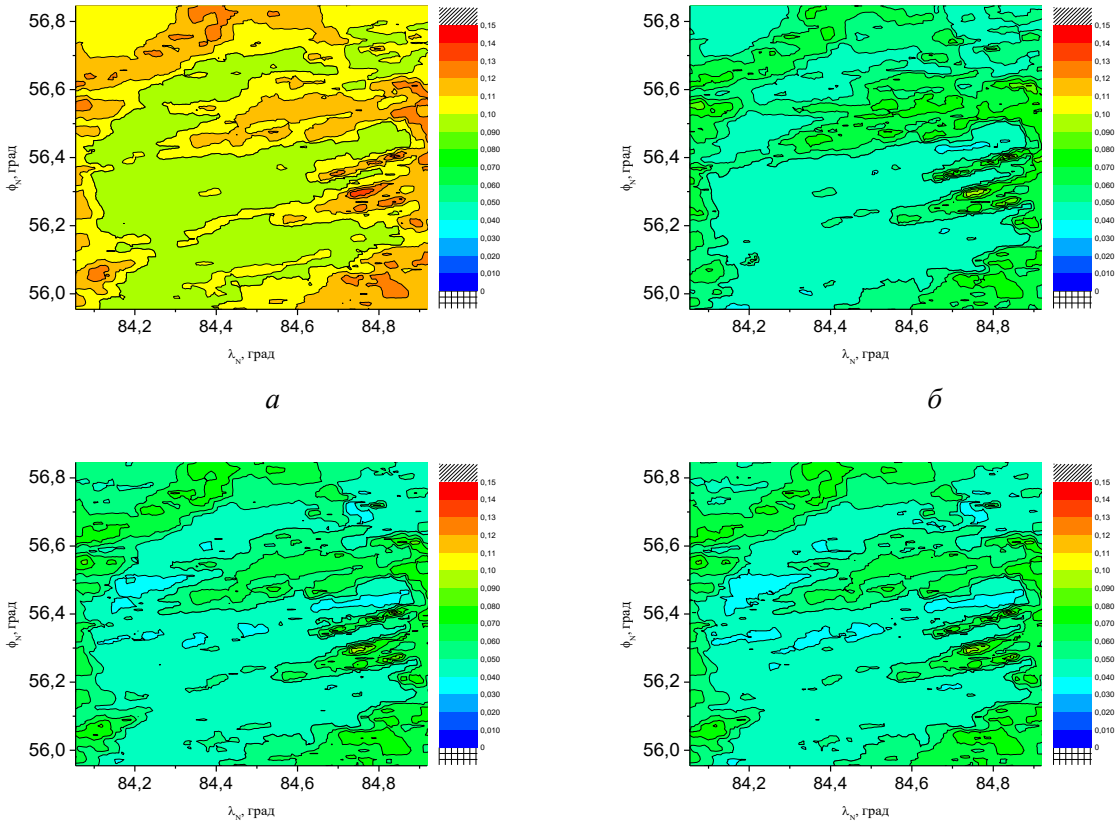
Для оценки погрешностей алгоритмов рассматривались тестовые точки в центре хвойных лесных массивов. Погрешность алгоритмов оценивалась как отличие восстановленных коэффициентов от измерений [17], значения которых приведены в табл. 1.

Как показал анализ, восстановленные коэффициенты отражения в каналах MODIS с центрами λ=0.412, 0.469, 0.555 и 0.649 мкм при отсутствии снега и малом содержании аэрозоля ($AOT_{0.55} \le 0.1$) слабо отличаются от наземных измерений из работы [17]. Следовательно, для этих каналов и участков данные [17] можно использовать как эталонные, а отличия от данных [17] приближенно оценивают погрешность алгоритмов. Коэффициенты отражения в канале MODIS с центром λ= 0.860 мкм отличается от измерений [17] заметно больше, потому что отражение в этом канале существенно зависит от состояния растительности (ее продуктивности). Поэтому в ряде ситуаций для этого канала данные [17] будут непригодны как эталонные.

В качестве исходных данных для задания модели атмосферы использовались данные Aeronet [18] по аэрозольной оптической толщине (AOT), распределению частиц по размерам и комплексному показателю преломления. Помимо этого использовались профили температуры и давления из измерений MODIS [19]. Для участка Томской области $AOT_{0.55}$ лежала в пределах от 0.1 до 1.52, для участка Московской области – от 0.04 до 0.07, а для участка Иркутской области – от 0.04 до 0.06. Пример восстановленных распределений коэффициентов отражения по земной поверхности приведен на рис. 2.

Табл. 1. Значения коэффициентов отражения молодой хвои зрелой сосны в летний период года [17].

| **λ, мкм** | 0.412 | 0.469 | 0.555 | 0.645 | 0.860 |
|---|---|---|---|---|---|
| $\mathbf{r_{surf,meas}(\lambda)}$ | 0.021 | 0.026 | 0.050 | 0.038 | 0.310 |




*а* *б*

*в* *г*

Рис. 2. Пример восстановления коэффициента отражения для участка Томской области за 22.06.2012г., время: 7:45, спутник TERRA. *а* – без коррекции; *б* – алгоритм MOD09; *в* – алгоритм без учета поляризации; *г* – алгоритм с учетом поляризации.

Табл. 2. Среднее отличие $\Delta r_{surf}$ от наземных измерений [17].

| **Номер участка** | **λ, мкм** | **$\Delta r_{surf}$, алгоритм без коррекции** | **$\Delta r_{surf}$, алгоритм MOD09** | **$\Delta r_{surf}$, алгоритм без учета поляризации** | **$\Delta r_{surf}$, алгоритм с учетом поляризации** |
|---|---|---|---|---|---|
| 1 | 0,412 | 0,193 | 0,012 | 0,035 | 0,033 |
| 1 | 0,469 | 0,116 | 0,013 | 0,013 | 0,006 |
| 1 | 0,555 | 0,053 | 0,011 | 0,010 | 0,003 |
| 1 | 0,649 | 0,030 | 0,017 | 0,007 | 3,0E-5 |
| 1 | 0,860 | 0,031 | 0,021 | 0,034 | 0,027 |
| 2 | 0,412 | 0,164 | 0,017 | 0,018 | 0,018 |
| 2 | 0,469 | 0,094 | 0,009 | 0,009 | 0,008 |
| 2 | 0,555 | 0,035 | 0,012 | 0,011 | 0,011 |
| 2 | 0,649 | 0,027 | 0,010 | 0,010 | 0,010 |
| 2 | 0,860 | 0,155 | 0,157 | 0,155 | 0,156 |
| 3 | 0,412 | 0,173 | 0,014 | 0,008 | 0,011 |
| 3 | 0,469 | 0,098 | 0,010 | 0,009 | 0,008 |
| 3 | 0,555 | 0,047 | 0,012 | 0,010 | 0,010 |
| 3 | 0,649 | 0,023 | 0,011 | 0,008 | 0,010 |
| 3 | 0,860 | 0,035 | 0,036 | 0,036 | 0,030 |

Оценка погрешности рассматриваемых алгоритмов выполнялась для 3 точек (по одной для каждого участка): 1) точка на территории Томского государственного природного заказника $56.2^0$ с.ш., $84.3^0$ в.д.; 2) точка в Национальном парке «Лосиный остров» $58.85^0$ с.ш., $37.83^0$ в.д.; 3) точка в государственном природном заказнике «Красный Яр» $52.52^0$ с.ш. $105.06^0$ в.д.

Для каждого канала и рассматриваемой точки определялось среднее по рассматриваемым снимкам отличие от эталонного значения $\Delta r_{surf}$. Полученные средние значения $\Delta r_{surf}$ приведены в табл. 2.

Сравнение $\Delta r_{surf}$ предлагаемого алгоритма с $\Delta r_{surf}$ алгоритмом MOD09 для точки в Томской области показывает, что предлагаемый алгоритм с учетом поляризации для рассматриваемых тестовых снимков имеет заметно меньшие отличия от эталонных значений для каналов с центрами λ=0.649, 0.469 и 0.555 мкм и почти одинаковые отличия для λ=0.860 мкм по сравнению с алгоритмом MOD09 NASA. Однако на длине волны 0.412 мкм значения, полученные алгоритмом MOD09, ближе к данным, приведенным в [17].

Для тестовой точки Московской области отличие результатов восстановления коэффициентов отражения поверхности Земли от эталонных значений, полученных рассмотренными алгоритмами (за исключением алгоритма без атмосферной коррекции), практически не отличаются друг от друга. В отличие от первой точки, для данной точки при λ=0.860 мкм наблюдается существенно большее отличие от [17]. Причина этого состоит, вероятно, в том, что значения из [17] существенно отличаются от реального значения коэффициента отражения в этом канале и не подходят для рассматриваемой ситуации как эталон.

Сравнение результатов для тестовой точки в Иркутской области показывает, что предлагаемый алгоритм с учетом поляризации для рассматриваемых снимков имеет несколько меньшие отличия от эталонных значений для длин волн 0.469, 0.555 и 0.412 мкм, чем MOD09, и почти одинаковые отличия для λ=0.860 и 0.649 мкм.

**Заключение.** Сравнение с алгоритмом MOD09 показывает, что этот алгоритм дает погрешности восстановления $r_{surf}$ существенно ниже, чем алгоритм MOD09 NASA при λ= 0.469, 0.555 и 0.649 мкм, погрешность одного порядка при λ= 0.860 мкм. При λ= 0.412 мкм для одних ситуаций погрешности восстановления $r_{surf}$ с помощью алгоритма MOD09 меньше, для других – предпочтительнее предложенный алгоритм, а для третьих алгоритмы приводят к погрешностям одного порядка.

ЛИТЕРАТУРА


[1] *Otterman J., Fraser R.S.* Adjacency effects on imaging by surface reflection and atmospheric scattering: cross radiance to zenith // Applied optics. 1979. V. 18, No. 16. P. 2852-2860.

[2] *Vermote E.F., El Saleous N., Justice C.O., Kaufman Y.J., Privette J.L., Remer L., Roger J.C. and Tanre D.* Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation // Journal of geophysical research. 1997. V. 102, No. D14. P. 17,131-17,141.

[3] *Vermote E.F., Vermeulen A.* Atmospheric correction algorithm: spectral reflectances (MOD09). Algorithm Theoretical Background document, version 4.0. 1999. [Электронный ресурс]: **http://modis.gsfc.nasa.gov/atbd/atbd_mod08.pdf**.

[4] *Lyapustin A., Martonchik J., Wang Y., Laszlo I., Korkin S*. Multiangle implementation of atmospheric correction (MAIAC): 3. Atmospheric correction // Remote Sensing of Environment. 2012. V. 127. P. 385-393.

[5] *Тарасенков М.В., Белов В.В.* Комплекс программ восстановления отражательных свойств земной поверхности в видимом и УФ-диапазонах // Оптика атмосферы и океана. 2014. Т. 27. № 07. С. 622-627.

[6] *Tarasenkov M.V., Belov V.V., Engel M.V.* Algorithm for reconstruction of the Earth surface reflectance from Modis satellite measurements in a turbid atmosphere // Proceedings of SPIE. 2018. V. 10833. CID: 10833 16 [10833-58].

[7] *Gaofei Yin, Ainong Li, Shengbiao Wu, Weiliang Fan, Yelu Zeng, Kai Yan, Baodong Xu, Jing Li, Qinhuo Liu*. PLC: A simple and semi-physical topographic correction method for vegetation canopies based on path length correction // Remote Sensing of Environment. 2018. V. 215. P. 184-198.

[8] *Егоров В.А., Барталев С.А.* Метод радиометрической коррекции искажений отражательных характеристик земного покрова в данных спутниковых измерений, вызванных влиянием рельефа местности // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13. № 5. С. 192-201.
[9] *Breon F.M., Vermote E.* Correction of MODIS surface reflectance time series for BRDF effects // Remote Sensing of Environment. 2012. V. 125. P. 1-9.
[10] *Lee T.Y., Kaufman Y.J.* Non-Lambertian Effects on Remote-Sensing of Surface Reflectance and Vegetation Index // IEEE Transactions on Geoscience and Remote Sensing. 1986. V. 24, No. 5. P. 699-708.
[11] *Тарасенков М.В., Кирнос И.В., Белов В.В*. Наблюдение земной поверхности из космоса через просвет в облачном поле // Оптика атмосферы и океана. 2016. Т. 29. № 09. С. 767-771.
[12] *Гермогенова Т.А.* О влиянии поляризации на распределение интенсивности рассеянного излучения // Изв. АН СССР. Сер.геофиз. 1962. № 6. С. 854-856.
[13] *Сушкевич Т.А*. Математические модели переноса излучения. М.: БИНОН. Лаборатория знаний, 2005. 661 с.
[14] *Назаралиев М.А.* Статистическое моделирование радиационных процессов в атмосфере. М.: Наука, 1990. 227 с.
[15] *Зимовая А.В., Тарасенков М.В., Белов В.В*. Влияние поляризации излучения на восстановление коэффициента отражения земной поверхности по спутниковым данным в видимом диапазоне длин волн // Оптика атмосферы и океана. 2017. Т. 30. № 11. С. 927-932.
[16] *Zimovaya A.V., Tarasenkov M.V., Belov V.V*. Effect of radiation polarization on reconstruction of the earth's surface reflection coefficient from satellite data in the visible wavelength range // Proceedings of SPIE. 2017. V. 10466. CID: 10466 10. [10466-54].
[17] *Кринов Е.Л.* Спектральная отражательная способность природных образований. Ленинград. Издательство Академии Наук СССР, 1947. 271 с.
[18] Aerosol Robotic Network (AERONET). **http://aeronet.gsfc.nasa.gov**.
[19] **https://lpdaac.usgs.gov/data_access**

# ВОССТАНОВЛЕНИЕ КОЭФФИЦИЕНТОВ ОТРАЖЕНИЯ ЗЕМНОЙ ПОВЕРХНОСТИ ПО ДАННЫМ СПУТНИКОВ РЕСУРС-П И МЕТЕОР-М

*Тарасенков М.В., Белов В.В., Энгель М.В., Гриднев Ю.В., Зимовая А.В.*
Институт оптики атмосферы им. В.Е. Зуева СО РАН, Томск

Рассматриваются результаты восстановления коэффициентов отражения земной поверхности с учетом атмосферных искажений по спутниковым измерениям приборов Сангур и ГСА спутника Ресурс-П и прибора МСУ-100 спутника Метеор-М.
*Ключевые слова: атмосферная коррекция, коэффициент отражения земной поверхности, метод Монте-Карло.*

**Введение.** В настоящий момент на орбите Земли функционирует ряд отечественных спутников дистанционного зондирования земной поверхности. Это Ресурс-П, Метеор-М, Канопус-В и Электро-Л. В своем составе эти спутники имеют разнообразную аппаратуру [1]. В частности, на спутнике Ресурс-П размещены гиперспектральная аппаратура (ГСА), оптико-электронная аппаратура высокого разрешения (Геотон-Л1) и широкозахватная мультиспектральная аппаратура (ШМСА). На спутнике Канопус-В установлена панхроматическая съемочная и многозональная съёмочная системы. Спутник Метеор-М содержит комплексы аппаратуры многозональной спутниковой съемки (МСУ-50 и МСУ-100) и многозональное сканирующее устройство малого разрешения (МСУ-МР). На спутниковой системе Электро-Л установлено многозональное сканирующее устройство (МСУ-ГС).

Одной из величин, характеризующей состояние и тип земной поверхности, является коэффициент отражения земной поверхности. Значение коэффициента отражения можно восстановить из спутниковых измерений путем решения обратной задачи. При этом во многих ситуациях решению обратной задачи должна предшествовать атмосферная коррекция – устранение из измерений искажающего влияния атмосферы. Решение этой задачи в комплексе состоит в построении модели атмосферы, решении уравнения переноса излучения относительно интенсивности излучения, формирующего принимаемый сигнал, и решении обратной задачи. По этому принципу построены активно развивающиеся и постоянно обновляющиеся алгоритмы атмосферной коррекции для спутниковых приборов MODIS [2], Polder [3], MISR [4] и др. Разработанный нами алгоритм [5, 6] для снимков MODIS, учитывающий влияние неоднородности земной поверхности, также основан на решении обратной задачи.

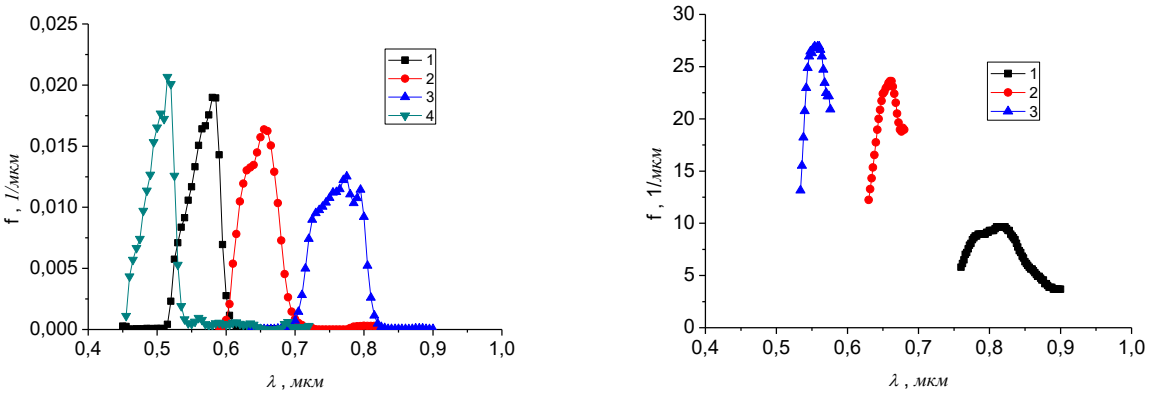
В настоящий момент для отечественных спутниковых приборов нет стандартных продуктов, содержащих восстановленные коэффициенты отражения земной поверхности после атмосферной коррекции. В рамках настоящей статьи будут кратко рассмотрены результаты нашей работы по созданию алгоритмов и программ атмосферной коррекции данных отечественных спутниковых приборов Геотон-Л1 и ГСА спутника Ресурс-П и МСУ-10 спутника Метеор-М.

**Исходные данные алгоритма.** В качестве исходных данных использовались откалиброванные интенсивности излучения, регистрируемого от наблюдаемых пикселей, координаты пикселей, а также информация о взаимном расположении Солнца, приемника и наблюдаемого участка. Эти данные для приборов спутников Ресурс-П, Канопус-В и Метеор-М можно заказать через геопортал Роскосмоса [7]. Калибровочные константы и информация о взаимном расположении Солнца, приемника и наблюдаемого участка хранятся в файлах метаданных соответствующего снимка. При работе с метаданными рассматриваемых приборов есть своя специфика. Так, для каждого прибора файл метаданных имеет свой формат и содержит свой набор данных, что требует создания процедур считывания метаданных отдельно для каждого прибора. Вторая особенность – метаданные снимков приборов Геотон-Л1 и ГСА содержат информацию для центра снимка, что допустимо в силу небольших пространственных размеров снимков. Третья особенность – взаимное положение приемника и наблюдаемого участка зада-

ется углом крена и азимутом направления движения спутника, а положение Солнца – зенитным углом Солнца и азимутом от направления на север. В отличие от этого, в стандартных продуктах MODIS, например, положение Солнца и спутника относительно каждого пикселя задается зенитным углом и азимутом от направления на север. Четвертой особенностью ряда каналов рассматриваемых приборов является то, что для этих спектральных интервалов существенно поглощение излучения атмосферой. На рис. 1 приведены аппаратные функции для приборов Геотон и МСУ-100. На рис. 2 для сопоставления приведено молекулярное пропускание атмосферы из [8]. Каналы ГСА существенно уже, чем каналы Геотона и МСУ-100 (табл. 1), но и среди них есть каналы, содержащие диапазоны длин волн с существенным молекулярным поглощением.

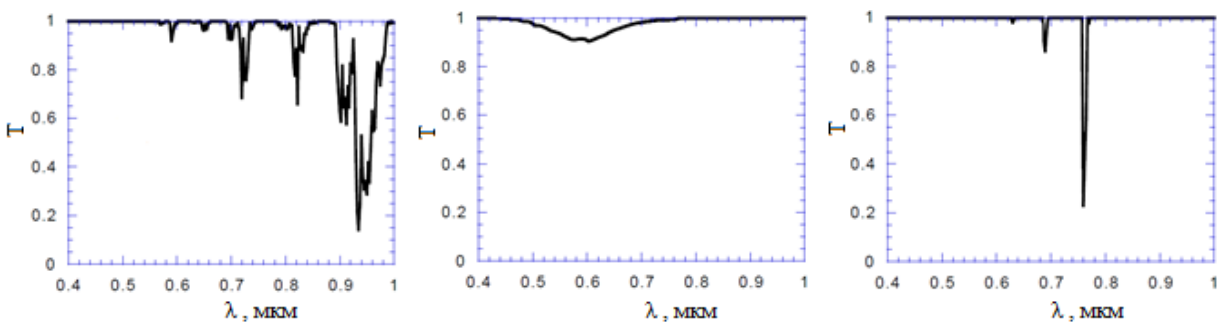
Помимо этого, при формировании модели атмосферы, используемой в алгоритме восстановления, использовались спутниковые данные MODIS об аэрозольной оптической толщине атмосферы (АОТ) $\tau_A$ на длинах волн $\lambda$=0.47, 0.55 и 0.66 мкм, профили температур $T$ и давления $P$, значение общего содержания водяного пара $W_{H20}$ и общего содержания озона $W_{O3}$ для участков и времени измерения, максимально близкими к рассматриваемым. При задании профилей молекулярного поглощения использовались также сечения поглощения $O_2$, $O_3$ и $H_2O$ для широкого набора температур и давлений [9,10]. Для задания профилей молекулярного рассеяния использовались данные из [11] о коэффициентах молекулярного рассеяния при нормальных условиях. При задании профилей аэрозольных коэффициентов использовались модели аэрозоля LOWTRAN-7 [12]. Значения солнечных постоянных выбирались из работы [13].

**Алгоритм восстановления коэффициентов отражения.** Алгоритм восстановления коэффициентов отражения состоит из двух этапов: 1) построение модели атмосферы, 2) атмосферная коррекция.




*а* *б*

Рис. 1. Аппаратные функции каналов прибора Геотон-Л1 (*а*) и МСУ-100 (*б*).




*а* *б* *в*

Рис. 2. Молекулярное пропускание T(λ) для $H_2O$ (*а*), $O_3$ (*б*) и $O_2$(*в*) из [8].

Табл. 1. Значения центра $\lambda_{0,i}$ и ширин $\widetilde{\Delta\lambda_i}$ спектральных диапазонов каналов прибора ГСА.

| № канала | $\lambda_{0,i}$ , нм | $\widetilde{\Delta\lambda_i}$ , нм | № канала | $\lambda_{0,i}$ , нм | $\widetilde{\Delta\lambda_i}$ , нм | № канала | $\lambda_{0,i}$ , нм | $\widetilde{\Delta\lambda_i}$ , нм |
|---|---|---|---|---|---|---|---|---|
| 7 | 431,63 | 3,94 | 46 | 575,82 | 4,09 | 88 | 771,61 | 5,79 |
| 8 | 435,62 | 4,09 | 47 | 579,91 | 4,19 | 89 | 777,45 | 5,99 |
| 9 | 439,75 | 4,23 | 48 | 584,1 | 4,31 | 90 | 783,49 | 6,19 |
| 10 | 444,02 | 4,36 | 49 | 588,4 | 4,42 | 91 | 789,73 | 6,41 |
| 11 | 448,43 | 4,5 | 50 | 592,82 | 4,54 | 92 | 796,2 | 6,64 |
| 12 | 452,98 | 4,65 | 51 | 597,36 | 4,66 | 93 | 802,9 | 6,88 |
| 13 | 457,68 | 4,8 | 52 | 602,02 | 4,78 | 94 | 809,84 | 7,14 |
| 14 | 462,54 | 4,97 | 53 | 606,81 | 4,91 | 95 | 817,05 | 7,41 |
| 15 | 466,71 | 3,42 | 54 | 611,72 | 5,05 | 96 | 822,63 | 3,81 |
| 16 | 470,15 | 3,51 | 55 | 616,77 | 5,18 | 97 | 826,44 | 3,89 |
| 17 | 473,68 | 3,6 | 56 | 621,95 | 5,32 | 98 | 830,33 | 3,96 |
| 18 | 477,31 | 3,71 | 57 | 627,27 | 5,46 | 99 | 834,3 | 4,04 |
| 19 | 481,05 | 3,82 | 59 | 632,73 | 5,61 | 100 | 838,34 | 4,12 |
| 20 | 484,9 | 3,95 | 61 | 638,34 | 5,76 | 101 | 842,46 | 4,2 |
| 21 | 488,89 | 4,09 | 63 | 644,1 | 5,92 | 102 | 846,66 | 4,29 |
| 22 | 493,01 | 4,24 | 64 | 649,22 | 6,28 | 103 | 850,95 | 4,37 |
| 23 | 497,29 | 4,4 | 65 | 654,77 | 4,88 | 104 | 855,32 | 4,46 |
| 24 | 501,73 | 4,57 | 66 | 659,7 | 5,01 | 105 | 859,78 | 4,55 |
| 25 | 506,35 | 4,76 | 67 | 664,75 | 5,14 | 106 | 864,34 | 4,64 |
| 26 | 511,16 | 4,97 | 68 | 669,93 | 5,27 | 107 | 868,98 | 4,74 |
| 27 | 514,9 | 2,57 | 69 | 675,24 | 5,39 | 108 | 873,72 | 4,84 |
| 28 | 517,46 | 2,62 | 70 | 680,67 | 5,52 | 109 | 878,56 | 4,94 |
| 29 | 520,09 | 2,68 | 71 | 686,23 | 5,65 | 110 | 883,5 | 5,04 |
| 30 | 522,77 | 2,75 | 72 | 691,92 | 5,78 | 111 | 888,53 | 5,14 |
| 31 | 525,52 | 2,81 | 73 | 697,75 | 5,93 | 112 | 893,68 | 5,25 |
| 32 | 528,33 | 2,88 | 74 | 703,74 | 6,09 | 113 | 898,93 | 5,36 |
| 33 | 531,2 | 2,95 | 75 | 708,85 | 4,16 | 114 | 904,28 | 5,47 |
| 34 | 534,15 | 3,02 | 76 | 713,02 | 4,24 | 115 | 909,75 | 5,58 |
| 35 | 537,16 | 3,09 | 77 | 717,28 | 4,33 | 116 | 915,34 | 5,7 |
| 36 | 540,26 | 3,17 | 78 | 721,63 | 4,42 | 117 | 921,04 | 5,82 |
| 37 | 543,42 | 3,25 | 79 | 726,07 | 4,52 | 118 | 926,86 | 5,94 |
| 38 | 546,67 | 3,33 | 80 | 730,62 | 4,63 | 119 | 932,8 | 6,07 |
| 39 | 550 | 3,41 | 81 | 735,27 | 4,74 | 120 | 938,87 | 6,19 |
| 40 | 553,41 | 3,5 | 82 | 740,04 | 4,86 | 121 | 945,06 | 6,32 |
| 41 | 556,91 | 3,59 | 83 | 744,94 | 4,99 | 122 | 951,38 | 6,46 |
| 42 | 560,5 | 3,68 | 84 | 749,97 | 5,13 | 123 | 957,84 | 6,59 |
| 43 | 564,18 | 3,78 | 85 | 755,14 | 5,28 | 124 | 964,43 | 6,73 |
| 44 | 567,96 | 3,88 | 86 | 760,46 | 5,44 | | | |
| 45 | 571,84 | 3,98 | 87 | 765,95 | 5,61 | | | |

Процедура построения модели атмосферы сводится к следующему:

1) По спутниковым измерениям АОТ при λ=0.47, 0.55 и 0.66 мкм с помощью метода наименьших квадратов определяется параметр Ангстрема А. Тогда АОТ на длине волны $\lambda$ приближенно оценивается как:

$$\tau_A(\lambda) = \tau_A(0.55)\left(\frac{0.55}{\lambda}\right)^A, \quad (1)$$

где $\tau_A(0.55)$ –АОТ при $\lambda$=0.55 мкм.

2) Среди моделей LOWTRAN-7 подбирается модель наиболее близкая по $\tau_A(0.55)$. Исходя из этой модели, находятся профили индикатрис аэрозольного рассеяния и коэффициентов аэрозольного ослабления и рассеяния:

$$\sigma_{t,a}(i,l) = \frac{\sum_{j=1}^{N} \tilde{\sigma}_{t,a}(\lambda_{ij},l) f_{ij} \Delta\lambda_j}{\sum_{j=1}^{N} f_{ij} \Delta\lambda_j}, \quad (2)$$

$$\sigma_{s,a}(i,l) = \frac{\sum_{j=1}^{N} \tilde{\sigma}_{s,a}(\lambda_{ij},l) f_{ij} \Delta\lambda_j}{\sum_{j=1}^{N} f_{ij} \Delta\lambda_j}, \quad (3)$$

$$g_a(i,l,cos\gamma_k) = \frac{\sum_{j=1}^{N} \tilde{g}_a(\lambda_{ij},l,cos\gamma_k) f_{ij} \Delta\lambda_j}{\sum_{j=1}^{N} f_{ij} \Delta\lambda_j}, \quad (4)$$

где $\sigma_{t,a}(i,l)$, $\sigma_{s,a}(i,l)$, $g_a(i,l,cos\gamma_k)$ – значение коэффициентов аэрозольного ослабления и рассеяния и индикатрисы аэрозольного рассеяния в $i$-м канале прибора в $l$-м слое атмосферы; $\gamma_k$ – $k$-й угол рассеяния; $\tilde{\sigma}_{t,a}(\lambda_{ij},l)$, $\tilde{\sigma}_{s,a}(\lambda_{ij},l)$, $\tilde{g}_a(\lambda_{ij},l,cos\gamma_k)$ – значения аэрозольных параметров из модели LOWTRAN-7; $\Delta\lambda_j$ – шаг по длине волны, с которым заданы аппаратные функции; $f_{ij}$ – значение аппаратной функции в $i$-м канале на $j$-й длине волны.

3) Исходя из профилей $T$ и $P$ и аппаратной функции канала, определяется профиль коэффициента молекулярного рассеяния:

$$\sigma_{s,m}(i,l) = \frac{\sum_{j=1}^{N} \tilde{\sigma}_{s,m}(\lambda_{ij},l) f_{ij} \Delta\lambda_j}{\sum_{j=1}^{N} f_{ij} \Delta\lambda_j}, \quad (5)$$

$$\tilde{\sigma}_{s,m}(\lambda_{ij},l) = \sigma_{s,m0}(\lambda_{ij}) \frac{P(l)}{P_0} \frac{T_0}{T(l)}, \quad (6)$$

где $\sigma_{s,m}(i,l)$ – коэффициент молекулярного рассеяния для $i$-го канала прибора в $l$-м слое атмосферы; $\tilde{\sigma}_{s,m}(\lambda_{ij},l)$ – коэффициент молекулярного рассеяния на длине волны $\lambda_{ij}$; $\sigma_{s,m0}(\lambda_{ij})$ – коэффициент молекулярного рассеяния на длине волны $\lambda_{ij}$ при нормальных условиях ($P_0 =$ 1013 ГПа , $T_0$=273.1 К); $P(l)$ – давление в $l$-м слое; $T(l)$ – температура в $l$-м слое.

4) Исходя из величины $W_{O3}$ и профиля счетной концентрации молекул озона из моделей LOWTRAN-7, создается профиль счетной концентрации молекул озона:

$$n_{O3,l} = \tilde{n}_{O3,l} \frac{N_{O3}}{\tilde{N}_{O3}}, \quad (7)$$

$$N_{O3} = \frac{W_{O3} \cdot N_A}{V_m}, \quad (8)$$

где $n_{O3,l}$ – счетная концентрация $O_3$ в $l$-м слое атмосферы; $\tilde{n}_{O3,l}$ – счетная концентрация $O_3$ в $l$-м слое атмосферы для модели LOWTRAN-7; $\tilde{N}_{O3}$ - интегральное по слою атмосферы количество молекул для модели LOWTRAN-7; $N_A$ – число Авогадро, $V_m$ – молярный объем.

5) Создается профиль счетной концентрации молекул кислорода:

$$n_{O2,l} = \frac{P_l V_{O2,l} T_0 N_A}{P_0 T_l V_m}, \quad (9)$$

где $P_l$, $T_l$ – давление и температура в $l$-м слое; $V_{O2,l}$ – объемная концентрация $O_2$; $T_0$ и $P_0$ – температура и давление при нормальных условиях.

6) По спутниковым измерениям $W_{H2O}$ создается профиль счетной концентрации водяного пара:

$$n_{H2O,l} = \begin{cases} \tilde{n}_{H2O,l} \cdot \frac{N_{H2O} - \tilde{N}_{H2O,2}}{\tilde{N}_{H2O}}, z \le 15 \text{ км} \\ \tilde{n}_{H2O,l}, z > 15 \text{ км} \end{cases}, \quad (10)$$

$$N_{H2O} = \frac{W_{H2O} N_A}{M_{H2O}}, \quad (11)$$

где $\tilde{n}_{H2O,l}$ – счетная концентрация $H_2O$ в $l$-м слое атмосферы для модели LOWTRAN-7; $\tilde{N}_{H2O}$ - интегральное по слою атмосферы количество молекул для модели LOWTRAN-7; $\tilde{N}_{H2O,2}$ – интегральное по слою атмосферы выше 15 км количество молекул для модели LOWTRAN-7; $N_{H2O}$ – интегральное по слою атмосферы количество молекул; $M_{H2O}$ – молярная масса молекулы воды.

7) Задается профиль молекулярного поглощения с учетом аппаратной функции канала:

$$\sigma_{a,m}(i,l) = \left( n_{O3,l} \frac{\sum_{j=1}^{N1} k_{O3}(P_l,T_l) f_{ij} \Delta\lambda_{ij}}{\sum_{j=1}^{N} f_{ij} \Delta\lambda_{ij}} + n_{O2,l} \frac{\sum_{j=1}^{N2} k_{O2}(P_l,T_l) f_{ij} \Delta\lambda_{ij}}{\sum_{j=1}^{N} f_{ij} \Delta\lambda_{ij}} + \right.$$
$$\left. + n_{H2O,l} \frac{\sum_{j=1}^{N3} k_{H2O}(P_l,T_l) f_{ij} \Delta\lambda_{ij}}{\sum_{j=1}^{N} f_{ij} \Delta\lambda_{ij}} \right), \quad (12)$$

где $k_{O3}(P_l, T_l)$, $k_{O2}(P_l, T_l)$, $k_{H2O}(P_l, T_l)$ – сечения поглощения $O_3$, $O_2$ и $H_2O$ при температуре $T_l$ и давлении $P_l$ соответственно; $N_1$, $N_2$, $N_3$ – количество длин волн сечений $O_3$, $O_2$ и $H_2O$, для которых аппаратная функция $i$-го канала не равна 0.

8) Задается солнечная постоянная в канале прибора с учетом аппаратной функции:

$$\pi S_{\lambda,i} = \frac{\sum_{j=1}^{N} \pi \tilde{S}_\lambda(\lambda_{ij}) f_{ij} \Delta\lambda_j}{\sum_{j=1}^{N} f_{ij} \Delta\lambda_j}, \quad (13)$$

где $\pi S_{\lambda,i}$ – солнечная постоянная $i$-го канала прибора; $N$ – количество длин волн, для которых аппаратная функция $i$-го канала не равна 0; $\pi \tilde{S}_\lambda$ – солнечная постоянная на длине волны $\lambda_{ij}$; $\Delta\lambda_j$ шаг по длине волны в данных о солнечной постоянной.

Что касается алгоритма атмосферной коррекции, то снимки приборов Геотон, Сангур и МСУ-100 содержат на порядок больше пикселей, чем снимки MODIS. Поэтому выполнить расчет, с помощью алгоритма из [5, 6] для снимков этих приборов в настоящем виде не удалось. В качестве альтернативы была использована формула в приближении однородной поверхности вида:

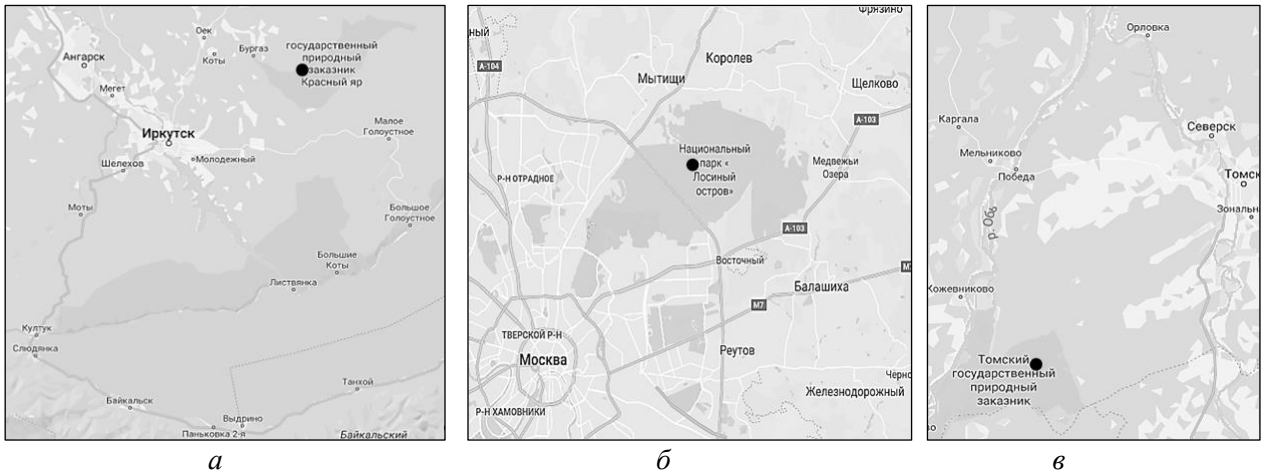
$$r_{surf,n} = \frac{Q_n / E_0}{1 + \gamma_1 Q_n / E_0} \quad (14)$$

$$Q_n = \frac{I_{sum,n} - I_{sun,n}}{\frac{1}{\pi} T_n + I_{dif,n}} \quad (15)$$

где $r_{surf,n}$ – коэффициент отражения $n$-го наблюдаемого пикселя; $E_0$ – освещенность земной поверхности без учета вклада отраженного излучения; $\gamma_1$ – сферическое альбедо атмосферы; $I_{sum,n}$ – интенсивность суммарного принимаемого излучения; $I_{sun,n}$ – интенсивность излучения, не взаимодействовавшего с земной поверхностью; $T_n$ – прямой коэффициент пропускания от наблюдаемого пикселя до приемной системы; $I_{dif,n}$ – интенсивность рассеянного излучения от поверхности при ее единичной светимости.

Величины $E_0$, $\gamma_1$, $I_{sun,n}$, $I_{dif,n}$ определяются разработанными программами метода Монте-Карло для сферической системы атмосфера-земная поверхность. Основные принципы метода Монте-Карло изложены в таких работах как [14].

Пример апробации алгоритма. В качестве примера рассмотрим результаты восстановления коэффициентов отражения земной поверхности для прибора МСУ-100. Для оценки возможностей разработанного комплекса программ рассматривалось 10 снимков, содержащих три тестовых участка, покрытых хвойными лесами: 1) Иркутская область $52.52^0$с.ш., $105.06^0$в.д., даты измерений 07.05.2017, 12.06.2017, 22.06.2017 г. (рис. 3,*а*); 2) Московская область $55.85^0$с.ш., $37.83^0$в.д., даты измерений 03.05.2017, 14.05.2017, 25.07.2016, 30.07.2016 г. (рис. 3,*б*); 3) Томская область $56.5^0$с.ш., $84.3^0$в.д., даты измерений 02.08.2017, 13.06.2017, 27.08.2017 г. (рис 3,*в*). На рис. 4 приведен пример полученного распределения коэффициента отражения по земной поверхности.




*а* *б* *в*

Рис. 3. Расположения тестовых участков для снимков прибора МСУ-100 спутника Метеор-М: *а* – Иркутская область; *б* – Московская область; *в* – Томская область.

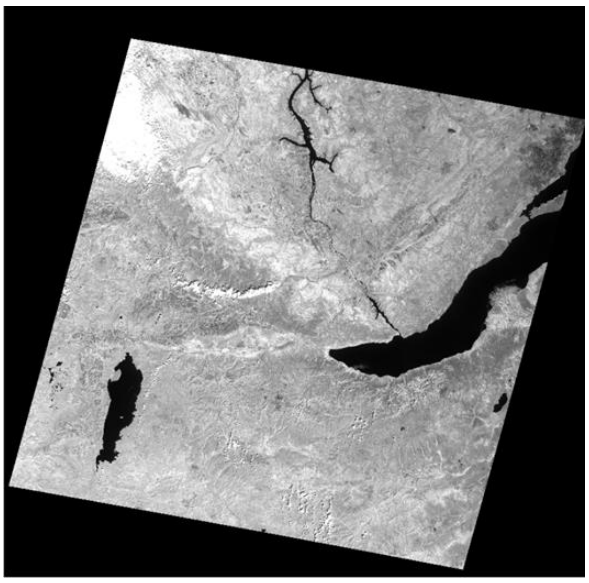


Рис. 4. Восстановленные по данным первого канала МСУ-100 коэффициенты отражения земной поверхности для участка Иркутской области дата измерений 07.05.2017 г.

Таблица 2. Значения коэффициентов отражения $r_{surf}$
для трех каналов МСУ-100, которые использовались как эталонные.

| **№ канала** | $r_{surf}$ |
|---|---|
| 1 | 0.30487 |
| 2 | 0.04838 |
| 3 | 0.04742 |

Для оценки погрешности восстановления коэффициентов отражения на рассматриваемых участках были выбраны 3 тестовые точки (по одной на каждом участке) в центре областей, покрытых хвойными лесами: 1) точка на территории Томского государственного природного заказника $56.2^0$ с.ш., $84.3^0$ в.д., 2) точка в Национальном парке «Лосиный остров» $58.85^0$ с.ш., $37.83^0$ в.д., 3) точка в государственном природном заказнике «Красный Яр» $52.52^0$ с.ш. $105.06^0$ в.д. Погрешность оценивалась как отличие от эталонных наземных измерений. В качестве эталона были взяты измерения из [15]. Значения коэффициентов отражения земной поверхности, с которыми выполнялось сравнение, приведены в табл. 2.

В табл. 3 приведены оценки средних по рассматриваемым снимкам погрешностей восстановления коэффициентов отражения для снимков МСУ-100.

Анализ результатов показывает, что полученные коэффициенты отражения МСУ-100 правдоподобны и могут использоваться для широкого круга задач, где полученная погрешность допустима. Что касается снимков приборов ГСА и Геотон-Л, выполненные оценки также показывают их пригодность для решения широкого круга задач.

Табл. 3. Средние погрешности восстановления коэффициентов отражения земной поверхности $\overline{\Delta r_{surf}}$
относительно эталонных значений [15] для трех тестовых участков для данных прибора МСУ-100.

| **Тестовый участок** | **№ канала** | $\overline{\Delta r_{surf}}$ |
|---|---|---|
| Иркутская область | 1 | 0,055 |
| Иркутская область | 2 | 0,012 |
| Иркутская область | 3 | 0,034 |
| г. Москва | 1 | 0,072 |
| г. Москва | 2 | 0,018 |
| г. Москва | 3 | 0,005 |
| Томская область | 1 | 0,100 |
| Томская область | 2 | 0,024 |
| Томская область | 3 | 0,020 |

**Заключение.** Разработанный комплекс программ позволяет в рамках приближения однородности земной поверхности по измерениям спутниковых приборов ГСА, Геотон-Л и МСУ-100 с дополнительным привлечением данных MODIS о параметрах атмосферы восстанавливать коэффициенты отражение земной поверхности с учетом атмосферных искажений. Оценка погрешностей показывает приемлемость (в смысле достоверности) результатов обработки с помощью созданного нами программного комплекса данных отечественных спутниковых ДЗЗ и целесообразность его использования для решения практических задач.

## ЛИТЕРАТУРА


[1] **www.ntsomz.ru**.

[2] *Breon FM., Vermote E.* Correction of MODIS surface reflectance time series for BRDF effects // Remote Sensing of Environment. 2012. V. 125. P. 1-9.

[3] *Leroy M., Deuzé J., Bréon F., Hautecoeur O., Herman M., Buriez J., Tanré D., Bouffies S., Chazette P., Roujean J.* Retrieval of atmospheric properties and surface bidirectional reflectances over land from POLDER/ADEOS // *J. Geophys. Res.* 1997. V. 102. P. 17023-17037.

[4] *Diner D.J., Martonchik J.V., Borel C., Gerstl S.A.W., Gordon H.R., Knyazikhin Y., Myneni R., Pinty B., Michel V.M.* Multi-Angle Imaging Spectro-Radiometer Level 2 Surface Retrieval Algorithm Theoretical Basis; Jet Propulsion Laboratory: La Cañada Flintridge, CA, USA, 2008.

[5] *Тарасенков М.В., Белов В.В.* Комплекс программ восстановления отражательных свойств земной поверхности в видимом и УФ-диапазонах // Оптика атмосферы и океана. 2014. Т. 27. № 07. С. 622-627.

[6] *Tarasenkov M.V., Belov V.V., Engel M.V.* Algorithm for reconstruction of the Earth surface reflectance from Modis satellite measurements in a turbid atmosphere // Proceedings of SPIE. 2018. Vol. 10833. CID: 10833 16 [10833-58].

[7] Геопортал Роскосмоса: **https://gptl.ru**.

[8] *Vermote E.F., Vermeulen A.* Atmospheric correction algorithm: spectral reflectances (MOD09). Algorithm Theoretical Background document, version 4.0. 1999. **http://modis.gsfc.nasa.gov/atbd/atbd_mod08.pdf**.

[9] **hitran.iao.ru**.

[10] **http://www.iup.uni-bremen.de/gruppen/molspec/databases/referencespectra/o3spectra2011 /index.html**.

[11] *Bucholtz A.* Rayleigh-scattering calculations for the terrestrial atmosphere // Applied optics. 1995. V. 34. № 15. P. 2765-2773.

[12] *Kneizys F.X., Shettle E.P., Anderson G.P., Abreu L.W., Chetwynd J.H., Selby J.E.A. Clough S.A., Gallery W.O.* User guide to LOWTRAN-7. ARGL-TR-86-0177.ERP 2010 / Hansom AFB. MA 01731. 137 p.

[13] *Gueymard C.A.* The sun's total and spectral irradiance for solar energy applications and solar radiation models // Solar Energy. 2004. No.76. P. 423-453.

[14] *Марчук Г.И., Михайлов Г.А., Назаралиев М.А., Дарбинян Р.А., Каргин Б.А., Елепов Б.С.* Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука, Сибирское отделение. 1976. 284 с.

[15] *Кринов Е.Л.* Спектральная отражательная способность природных образований. Ленинград: Издательство Академии Наук СССР. 1947. 271 с.

# МНОГОМЕРНОЕ ПРЕДСТАВЛЕНИЕ И АНАЛИЗ ДАННЫХ СТАЦИОНАРНЫХ НАЗЕМНЫХ НАБЛЮДЕНИЙ ЗА КОНЦЕНТРАЦИЯМИ АТМОСФЕРНЫХ ПРИМЕСЕЙ ДЛЯ ВЕРИФИКАЦИИ ГИДРОДИНАМИЧЕСКИХ МОДЕЛЕЙ ПРОГНОЗА

*Токарев В.М., Здерева М.Я.*
ФГБУ «Сибирский региональный научно-исследовательский гидрометеорологический институт»

В статье описана возможность оценки влияния погодных условий на концентрацию антропогенных примесей с помощью параметров гидродинамических моделей атмосферы. В качестве выборки использованы официальные данные наблюдений за одиннадцатью примесями на десяти постах города Новосибирска за четыре года и синхронные параметры, полученные из выходных полей глобальных гидродинамических моделей атмосферы. Предложен алгоритм факторного анализа для статистически неоднородных выборок. Продемонстрирована возможность использования бинарных деревьев для оценки зависимостей примесей от параметров атмосферы.
*Ключевые слова: динамика концентраций атмосферных примесей, факторный анализ, разложение дисперсий, бинарные деревья*

Динамика концентраций антропогенных атмосферных примесей, несомненно, зависит от погодных условий. Однако эмиссии таких примесей, за редким исключением, имеют мало общего с атмосферными процессами. Таким образом, задача моделирования и прогнозирования аэрозольного и газового состава атмосферного воздуха сильно осложняется. На входе модели надо иметь не только пространственные данные метеорологических наблюдений, но и оценки эмиссий от множества разнородных источников. К сожалению, системы мониторинга качества воздуха пока не могут предоставить данные наблюдений достаточного для моделирования и верификации объема и детальности. В случае же отсутствия подробной карты эмиссий, задача представляется невыполнимой. Попробуем все-таки извлечь полезную информацию из доступной с помощью не совсем традиционных подходов и алгоритмов.

Для исследования были использованы архивные данные наблюдений за концентрациями атмосферных примесей Новосибирского ЦМС (4 года, 11 контролируемых примесей на 10 постах) и синхронные архивные данные глобальных прогностических моделей Exeter и ECMWF.

На первом этапе из данных наблюдений примесей формировался многомерный массив: 11 примесей * 10 постов * 3 суточных срока * 12 месяцев * 4 года. Таким образом, далее можно было оценивать влияние таких факторов, как местоположение на территории мегаполиса, суточный и сезонный ход, и межгодовой тренд. Однако от оценок суточного хода пришлось отказаться из-за отсутствия ночных наблюдений (в выходные также были пропуски).

Качественный характер динамики примесей оценивался по перекрестным сравнительным графикам с разным временным масштабом (от суток до годов). И первые выводы были неутешительные: в рядах присутствует очевидная сильная нестационарность в масштабах от часов до нескольких суток, а тренды и сезонность трудноразличимы и немонотонны. Это означало, что придется иметь дело с выделением очень слабого полезного сигнала (влияния погодных условий) на фоне преобладающего шума от неизвестной динамики эмиссий и, возможно, технологических особенностей мониторинга.

Следующим шагом было получение количественных оценок факторов. Для этого использовались условные статистические оценки из известного разложения дисперсии [1]:

$$DY = DM(Y|x) + MD(Y|x), \quad (1)$$

где $DM(Y|x)$ определяет долю дисперсии Y за счет влияния X; $MD(Y|x)$ – неопределенная (остаточная) дисперсия Y. Тогда $DM(Y|x)/DY$ – дисперсионное отношение, мера (нелинейной) связи.

Дополнительную полезную информацию можно получить, рассматривая отдельно выборки из минимальных и максимальных по городу наблюдений. При этом минимум можно

условно интерпретировать, как «фон», т.е. пункт, достаточно удаленный от максимума, т.е. источника «выброса».

Все расчетные характеристики сведены в табл. 1.

Полученная таблица позволяет выполнить довольно тонкий статистический анализ. Так, средний максимум по городу для диоксида серы превышает средний минимум более, чем в 10 раз (больше всех таблице). То есть, на первый взгляд, для этой примеси местоположение поста является ключевым фактором. Однако, дисперсионное отношение оставляет этому фактору лишь 6% всей дисперсии, а сезонной компоненте – от 13% («фон») до 23% («выброс»). Так что все весьма относительно. И преобладание сезонных колебаний выбросов $SO_2$ за счет сжигания угля котельными и частным сектором вполне логично и объяснимо, впрочем, этим же объясняются такие различия в местоположении (частный сектор в Новосибирске занимает отдельные компактные территории).

Самый большой трендовый вклад – у оксидов азота и формальдегида, и это никак не радует, поскольку по графикам – это почти монотонный рост по годам. И, если оксиды азота объясняются ростом парка автотранспорта, то тренд формальдегида, скорее всего, связан с неконтролируемым использованием пластиков и эпоксидных смол.

Табл. 1. Компоненты дисперсионной идентификации концентрации примесей по пространтсвенно-временным осям матрицы данных.

| | по городу | | | | посты | | | тренд | | сезонность | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | макс | мин | сред | дисп | | DM(Y/X) | DM(Y/X)/DY | DM(Y/X)/DY | | DM(Y/X)/DY | |
| | MY | MY | MY | DY | N | | | выброс | фон | выброс | фон |
| | | | | | | | % | % | % | % | % |
| Пыль | 0.55405 | 0.12934 | 0.30519 | 0.04998 | 10 | 0.00293 | 5.86 | 2.71 | 2.24 | 6.95 | 13.96 |
| SO2 | 0.01860 | 0.00136 | 0.00782 | 0.00026 | 6 | 0.00001 | 6.18 | 8.42 | 5.07 | 22.53 | 13.32 |
| CO | 4.78524 | 1.28747 | 2.58200 | 4.07565 | 10 | 0.27580 | 6.76 | 11.38 | 0.25 | 2.46 | 19.45 |
| NO2 | 0.16659 | 0.03972 | 0.09193 | 0.00533 | 10 | 0.00030 | 5.66 | 12.62 | 53.07 | 2.62 | 0.61 |
| NO | 0.08238 | 0.02959 | 0.05714 | 0.00274 | 7 | 0.00066 | 24.32 | 37.35 | 51.67 | 1.80 | 0.85 |
| H2S | 0.00046 | 0.00046 | 0.00046 | 0.00000 | 1 | 0.00000 | 0.00 | 5.05 | 5.05 | 0.46 | 0.46 |
| Фенол | 0.00740 | 0.00124 | 0.00415 | 0.00004 | 3 | 0.00000 | 0.47 | 11.80 | 10.21 | 0.42 | 0.37 |
| Сажа | 0.09714 | 0.00164 | 0.03330 | 0.00527 | 10 | 0.00038 | 7.38 | 0.67 | 4.26 | 7.49 | 9.87 |
| HF | 0.00763 | 0.00763 | 0.00760 | 0.00008 | 1 | 0.00000 | 0.00 | 1.49 | 1.49 | 3.21 | 3.21 |
| Аммиак | 0.13945 | 0.02303 | 0.07739 | 0.01305 | 3 | 0.00067 | 5.17 | 5.27 | 1.27 | 2.20 | 1.90 |
| Форм | 0.02357 | 0.00734 | 0.01525 | 0.00018 | 6 | 0.00001 | 10.65 | 21.87 | 52.55 | 0.47 | 0.86 |

Табл. 2. Прямые и обратные связи превышения ПДК примесей и параметров атмосферы.

| | Примеси | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Параметры атмосферы | Пыль | CO | NO2 | фенол2 | фенол1 | сажа | аммиак2 | аммиак1 | форм |
| r0 | - | - | 0 | - | 0 | - | 0 | + | - |
| VV | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P0 | 0 | - | 0 | + | 0 | 0 | 0 | 0 | 0 |
| gradP0 | 0 | 0 | + | + | 0 | 0 | 0 | 0 | 0 |
| lapP0(+цикл,-анти) | 0 | 0 | 0 | 0 | + | 0 | 0 | 0 | 0 |
| dP0(t-24) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | | | | | | | | |
| Tz-T850 | - | 0 | - | 0 | - | - | + | 0 | 0 |
| lapT850(+холод,-тепло) | 0 | 0 | 0 | 0 | 0 | 0 | - | 0 | - |
| gradT850 | 0 | 0 | - | 0 | 0 | 0 | 0 | 0 | 0 |
| dT850(t-24) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | | | | | | | | |
| H500 | + | + | - | 0 | 0 | 0 | + | 0 | 0 |
| lapH500(+цикл,-анти) | 0 | - | 0 | - | + | 0 | 0 | 0 | 0 |
| gradH500 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | + | 0 |
| dH500(t-24) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | + | 0 |

Подобный анализ можно продолжить, но нам важно, что примененный алгоритм позволяет извлекать полезную количественную информацию из очень статистически непростых для обработки данных.

Остается вопрос, а где же здесь погода? Да где-то в очень значительной остаточной дисперсии (для фенола это почти 90%), как трудноразличимый фактор на фоне высокоамплитудных аварийных или «планово-незаконных» выбросов, которые хорошо видны на графиках. Анализируемую выборку сложно разбить на отрезки с «хорошей» и «плохой» погодой для концентраций всех примесей. Нет такого одного модельного параметра атмосферы. А чтобы это проверить и оценить используем один из алгоритмов распознавания – построение логического бинарного дерева решений. Собственно, он может сразу строить прогностические решающие правила по обучающей выборке, что и делали авторы в другой работе.

Логически алгоритм [2] очень прост. Он последовательно разделяет исходную выборку на две с разнесением вероятностей бинарного предиктанта в каждой подвыборке. Так получается дерево с двоящимися ветвями, в узлах которых фиксируется выбранный лучший параметр разделения (предиктор) и значение этого порога. Таким образом, выходное дерево содержит список выбранных лучших параметров с лучшими пороговыми значениями распознавания двух образов предиктанта («да» или «нет») и со знаковой информацией о связи предиктора и предиктанта. Эта связь определяется тем, совпадают ли условия деления по порогу предиктора (больше/меньше порога) с увеличением/уменьшением соответствующей вероятности образа предиктанта («да»/«нет»). Совпадение означает прямую связь: чем больше значение предиктора, тем больше вероятность 1-го образа предиктанта («да») и наоборот. Не вдаваясь в подробности, по каким критериям выбираются лучшие варианты разделения ветвей и правило останова (они важнее для целей прогноза), перейдем к получению статистических связей для нашей задачи. Сведение концентраций примесей к бинарной форме выполнялось по индивидуальному пороговому условию: больше или меньше ПДК. Матрица предикторов формировалась из списка различных выходных параметров атмосферы глобальных моделей Exeter и ECMWF в ближайших узлах сетки ГРИБ у земли, на верхней границе пограничного слоя и в свободной атмосфере.

Полученные результаты обучения сведены в таблицу (табл. 2) Для фенола и аммиака к обычной обучающей выборке фактов превышения ПДК – «хотя бы на одном посту» добавлено обучение для двух и более постов.

Первое, что можно отметить, это немногочисленность найденных связей, но это ожидаемо для таких зашумленных данных. Также ожидаема отрицательная связь относительной влажности у земли (r0), но для аммиака это не так, причем, и другие связи у аммиака чисто индивидуальные. Логично и рассеяние примесей с изменением вертикального градиента температуры в сторону неустойчивости (Tz-T850). Следует также отметить вероятность такого сценария рассеяния примеси, когда слабый ветер и устойчивая стратификация способствуют накоплению концентрации непосредственно около точечного источника, а неустойчивая погода будет рассеивать примесь, донося ее до постов наблюдений.

Еще одно методическое замечание относится к выбору критического порога для разбиения предиктанта на образы. ПДК – не физическая характеристика, а биологическая, так что выбор других порогов, например, единых для всех примесей частотных квантилей, должен обеспечить более четкое разделение градаций концентраций по условиям погоды.

Полученные в работе результаты не совсем очевидны и могут помочь в настройках численных моделей на прогноз уровней загрязнения атмосферы антропогенными примесями и последующей валидации и анализа результатов.

## ЛИТЕРАТУРА


[1] Дисперсионная идентификация / Под ред. Н.С. Райбмана. М.: Наука,1981. 336 с.

[2] *Манохин А.Н.* Алгоритм DW для распознавания образов: Пакет прикладных программ ОТЭКС. Новосибирск: изд-во Новосибирского государственного университета, 1981. С. 3-30.

# АВТОМАТИЗАЦИЯ ФОРМИРОВАНИЯ БАЗЫ ДАННЫХ ЛЕДНИКОВ НА ОСНОВЕ ДЗЗ

*Фёдоров Р.К.[(1)], Китов А.Д.[(2)], Авраменко Ю.В.[(1)]*
[(1)] Институт динамики систем и теории управления СО РАН, Иркутск
[(2)] Институт географии СО РАН, Иркутск,

В статье рассматривается технология формирования базы данных ледников на основе совокупности WPS сервисов. Сервисы технологии применяются для выделения границ видимой части ледников по космоснимкам с использованием обучаемой нейронной сети. По результатам выделения ледника рассчитываются его характеристики и формируется база данных по заданным правилам.
*Ключевые слова: комоснимки, ледники, нейронные сети, обучение распознаванию, базы данных*.

**Введение**. В горах Южной Сибири сохранилось множество малых современных ледников. Эти ледники являются хорошими индикаторами изменения климата. Ледники и их динамику отражают такие характеристики как площадь, длина, экспозиция, верхняя и нижняя высотные отметки, высота фирновой линии и др. Одной из характерных характеристик является площадь ледника. Выделяют общую площадь и площадь открытой части. По известной площади и типу ледника можно также оценить его мощность (объем накопленного льда). Наиболее показательна для анализа динамики ледника и его реакции на изменение климата – открытая часть (не покрытый поверхностными моренами участок, как правило верхний, ледника со снегом и льдом). В то же время, ледники представляют трудно доступные и опасные объекты. Поэтому использование данных дистанционного зондирования Земли (ДЗЗ), космоснимков, для измерения характеристик ледников приобрело особое значение. Именно открытая часть ледника хорошо выделяется на данных ДЗЗ. Кроме того, наиболее фотогеничны ледники в узкий промежуток времени, как правило в Сибири с середины июля до середины августа, а то и в меньший промежуток – с последней декады июля по первую декаду августа.

Векторизация границ ледника достаточно трудоёмкий процесс, который требует автоматизации и разработки новых методов компьютерного анализа. В то же время «ручное» дешифрирование трудно стандартизовать. Алгоритмический подход позволяет единообразно, на выбранных критериях, выделять границы ледников. Но учитывая разную экспозиция, характер съёмки объективное выделение границ может оказаться не столь простой задачей. Нужны обучающиеся алгоритмы. Одним из таких эффективных подходов является распознавание с использованием нейронных сетей. Применение современных методов ДЗЗ является достаточно сложным процессом, требующим некоторой квалификации в информационных технологиях. Исследователю требуется выбрать из большого количества методов ДЗЗ наиболее подходящий, установить, правильно его настроить и применить. Кроме того, на результат обработки нейронными сетями значительно влияет размер и представительность обучающей выборки. Формирование обучающей выборки достаточно трудоемкий процесс.

Поэтому является необходимым разработка технологии, которая позволит:

1) упростить применение современных методов обработки ДЗЗ;
2) обеспечит возможность расширения набора методов обработки данных ДЗЗ;
3) объединит усилия специалистов в части формирования обучающих выборок данных;
4) предоставит доступ к результатам обработки ДЗЗ и обеспечит каталогизацию и поиск.

В соответствии с современным развитием сети Интернет и программного обеспечения данная технология должна базироваться на сервис-ориентированном подходе, который позволяет удаленно применять различные методы без необходимости установки и настройки их на локальной машине.

**Материалы и методы**. В качестве характерного объекта были выбраны ледники горного массива Мунку-Сардык – ледник Перетолчина (наибольший ледник этой территории) и ледник Радде, имеющие историю наблюдения более 100 лет [1, 2].

Авторами статьи разработана технология (рис. 1) в виде совокупности WPS сервисов [3], каждый из которых может применять отдельно. Разработанная технология состоит из следующих этапов:

1) получение снимков, на этом этапе пользователь производит выбор подходящего снимка, на котором имеется ледник;
2) формирование обучающей выборки, на этом этапе пользователь указывает границы ледников. Данный этап является опциональным;
3) создание классификатора, на текущий момент применяются пока только нейронные сети. Данный этап является опциональным;
4) классификация снимков, пользователь может использовать готовые классификаторы или создать свои с помощью предыдущих этапов;
5) включение результатов классификации в каталог, который позволяет быстро найти нужный ледник и сохранить в одном из форматов.

Рассмотрим эти этапы подробнее.

**Формирование обучающей выборки.** Часто существующие классификаторы могут не корректно выделять ледники из-за наличия специфики исследуемой территории. В этом случае необходимо сформировать обучающую выборку, включающую специфичные случаи, и создать новый классификатор. В рамках геопортала [4] создана таблица (рис. 2), в которой пользователь указывает границы ледников на различных снимках.

Таблица состоит из следующих атрибутов:

1) name – название ледника;
2) boundaries – границы ледника в виде полигонального объекта;
3) date – дата снимка;
4) imagename – путь к файлу снимка в директории системы хранения данных геопортала;
5) imprecisedate – неточная дата снимка.

В каждой записи таблицы пользователь определяет границы одного ледника на конкретном снимке. Вся область снимка делится на два класса, область, покрытая полигонами из таблицы для этого снимка, считается ледником. Вся оставшаяся область изображения противоположным классом. Соответственно в процессе составления обучающей выборки необходимо полностью покрыть ледники полигонами.




Рис. 1. Основные этапы технологии выделения ледников.

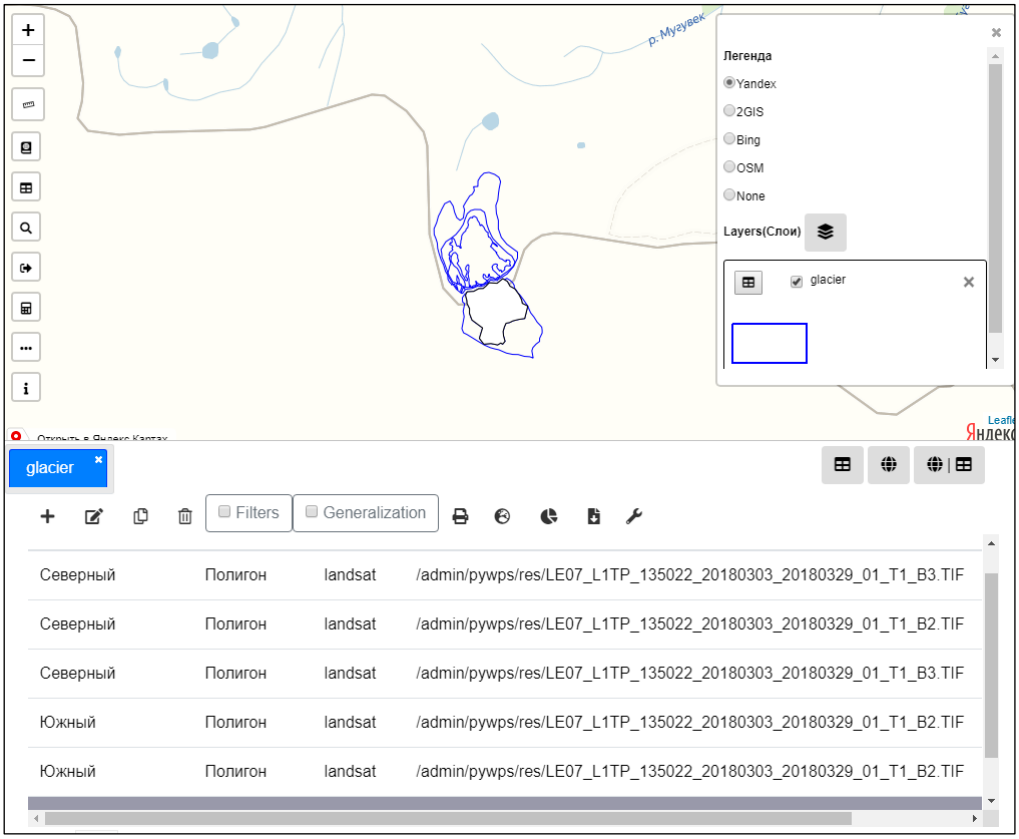



Рис. 2. Таблица обучающей выборки.

Космоснимки, полученные со спутника Landsat, представлены в виде набора файлов в формате GEOTIFF, в каждом из которых находится отдельный канал (монохромное изображение). Для отображения на карте геопортала необходимо совокупность файлов преобразовать в RGB представление. Для этого используется сервис Bands_to_RGB, который принимает на вход три монохромных изображения, соответствующих каналам модели RGB и строит из них цветное. В основе сервиса используется модуль gdal_merge_simple (https://github.com/gina-alaska/dans-gdal-scripts#gdal_merge_simple).

**Создание классификатора.** Для создания классификатора на основе обучающей выборки разработан набор сервисов, на вход которых подается таблица и на выходе получаем файл классификатора. В таблице можно отфильтровать записи по всем атрибутам, в том числе указать область обработки. Например, задать диапазон дат или определенный ледник. Для каждого изображения в обучающей выборке формируется набор записей. Затем используется сервис Rasterize, который принимает на вход векторный файл в формате SHAPE, определяющий положения прецедентов, исходное изображение в формате GEOTIFF и область прямоугольную область интереса. На рис. 3 представлен пример запуска сервиса Rasterize.

Результатом работы сервиса является растровое изображение – бинарная маска (рис. 4), содержащая положения прецедентов по заданной области интереса. В основе сервиса используется инструмент gdal_rasterize (https://gdal.org/programs/gdal_rasterize.html). Сервис позволяет задать размер выходного изображения в пикселях.

После сервиса Rasterize используется сервис Learning, который непосредственно производит обучение нейронной сети. На вход сервису подается растровое изображение модели RGB и бинарная маска, с указанием положений ледников. На выходе формируется файл модели. В основе сервиса используется проект Segmentation models (https://github.com/qubvel/segmentation_models). Данный проект позволяет использовать одну из предобученных нейросетей, распространенных архитектур Unet, FPN, Linknet или PSPNet. Принцип работы сервиса следующий. Из изображения и маски методом скользящего окна

формируется обучающая выборка. После чего вызывается метод обучения модели. Затем полученная модель возвращается пользователю в качестве результата. Так как процесс обучения достаточно долгий, то данный сервис вызывается в асинхронном режиме. Пользователь может следить за статусом его выполнения.

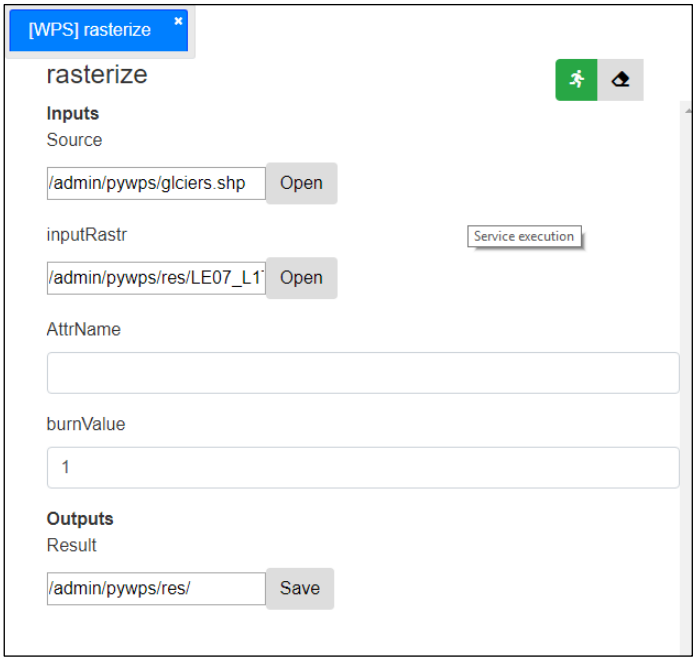
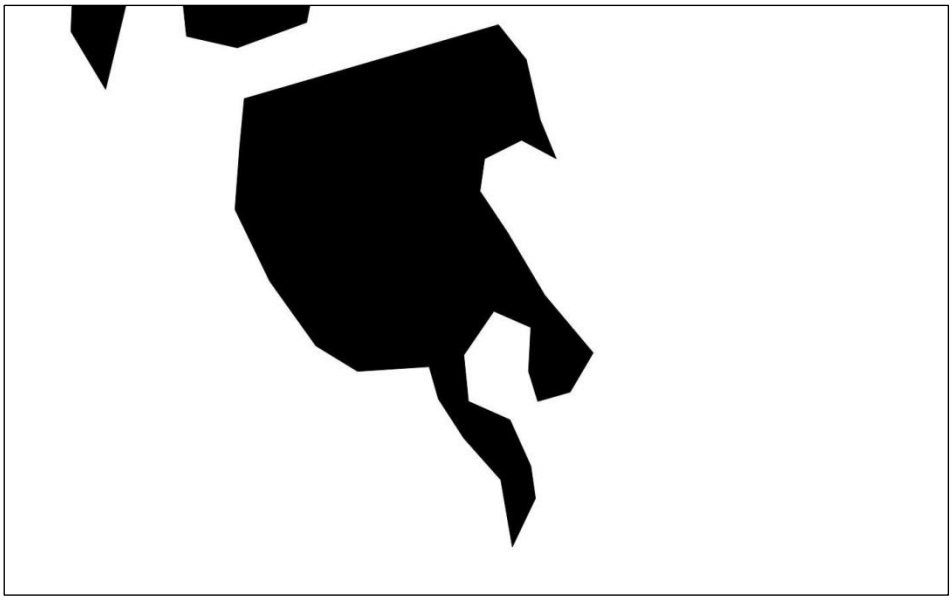
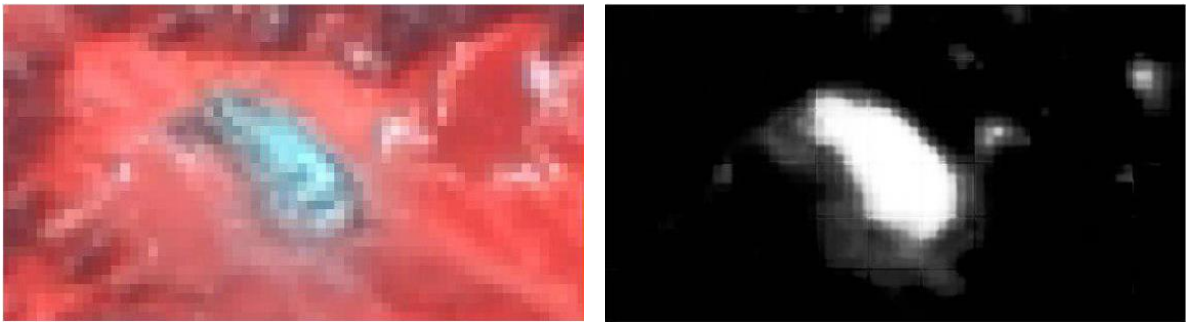


Рис. 3. Запуск сервиса Rasterize.


Рис. 4. Маска ледника, полученная на основе таблицы.


Рис. 5. Снимок Landsat и результат классификации.

**Классификация снимков.** Классификация осуществляется с помощью сервиса Segmentation. На вход сервису подается растровое изображение модели RGB. На выходе получается бинарная маска, с отмеченными положениями объектов. В основе сервиса используется проект Segmentation models (https://github.com/qubvel/segmentation_models). Сервис получает на вход изображение, после чего пробегается по нему скользящим окном. В результате чего изображение делится на строки и столбцы. Из каждой строки формируется набор данных и передается на вход нейронной сети. Затем результат работы записывается на соответствующие позиции копии исходного изображения и так далее пока не будет обработано все изображение целиком.

**Заключение**. Разработана технология применение нейронных сетей для анализа космоснимков, которая позволяет упростить применение современных методов обработки ДЗЗ. Применение методов в виде сервисов обеспечивает возможность расширения набора методов обработки данных ДЗЗ. Вместо нейронных сетей может быть использован любой другой метод, реализованный в виде WPS сервиса. Геопортал позволяет одновременно работать специалистам при формировании обучающих выборок данных, создании классификаторов, обработке данных ДЗЗ. Данной технологией могут воспользоваться для решения подобных задач на других территориях. Результаты анализа предоставляются в общий доступ.

*Работа выполнена при финансовой поддержке РФФИ (грант № 17-05-0040) и Сибирского отделения РАН (интеграционный проект № 37).*

ЛИТЕРАТУРА

[1] *Перетолчин С.П.* Ледники хребта Мунку-Сардык. Изв. Томск.техн. ин-та. Т. 9. Томск: Типолитография Сиб. т-ва печатного дела, 1908. 60 с

[2] *Kitov A.D., Kovalenko S.N., Plyusnin V.M.* The results of 100-year-long observations of the glacial geosystem dynamics in the Munku-Sardyk massif. Geography and natural resources, 2009. V. 30. No 3. P. 272-278. DOI: 10.1016/j.gnr.2009.09.012

[3] Geospatial and location standards [Электронный ресурс] // Open Geospatial Consortium. URL: **http://www.opengeospatial.org** (дата обращения: 04.09.2014).

[4] *Paramonov V., Fedorov R., Ruzhnikov G., Shumilov A.* Web-Based Analytical Information System for Spatial Data Processing // Communications in Computer and Information Science. 2013. Vol. 403. P. 93-101.

# ОЦЕНКА ГАЗООБМЕНА В ИРКУТСКОМ ВОДОХРАНИЛИЩЕ

*Хмельнов А.Е.[(1)], Гаченко А.С.[(1)], Михайлов А.А.[(1)], Минаев В.В.[(2)], Минаев А.В.[(2)]*
[(1)] Федеральное государственное бюджетное учреждение науки Институт динамики систем и теории управления имени В.М. Матросова СО РАН, Иркутск
[(2)] Федеральное государственное бюджетное учреждение науки Лимнологический институт Сибирского отделения Российской академии наук СО РАН, Иркутск

В статье рассмотрена задача анализа газообмена растворённых в воде парниковых газов ($CO_2$ и $CH_4$) через водную поверхность с атмосферой и выполнена оценка сезонных изменений такого газообмена для Иркутского водохранилища. Предложен оригинальный подход к решению поставленной задачи, в котором использованы современные ГИС технологии и технологии анализа данных.
*Ключевые слова: парниковые газы, газообмен, водохранилище, кластерный анализ, метан, углекислый газ.*

**Введение**. Возрастающий глобальный спрос на энергию и обеспокоенность в связи с изменениями в окружающей среде привели к стремлению оценить влияние ГЭС и образованных водохранилищ с точки зрения их оценки в качестве источников выбросов этих газов в атмосферу Земли. Обеспокоенность в связи с изменениями в окружающей среде привели к необходимости получения количественных оценок влияния ГЭС и образованных водохранилищ на газообмен парниковых газов ($CO_2$ и $CH_4$) в системе «вода-атмосфера». В первых научных работах [1, 2] основное внимание уделялось водохранилищам, расположенным в Канаде, Южной Америке, позже в Финляндии, Швеции и Швейцарии. В первом глобальном обзоре [4] отмечено, что такие факторы, как возраст, температура воды и количество затопленного органического углерода, регулируют выбросы парниковых газов водохранилищами. Из-за более значительного влияния $CH_4$ на глобальное потепление по сравнению с $CO_2$, поток метана в атмосферу может превышать такой благоприятный эффект, как фиксация углекислого газа водорослями.

В исследованиях, выполненных с начала 2000 года показано, что величина потока $CO_2$ и $CH_4$ из водохранилищ зависит от возраста водоема и географической широты расположения водоема [5]. Повышенные выбросы парниковых газов наблюдаются от молодых (менее 10 лет) водохранилищ [6-10] и обусловлены они быстрым разложением наиболее лабильного наземного органического вещества, и со временем водохранилища по выбросам парниковых газов могут характеризоваться как естественные озера, но некоторые водохранилища могут попрежнему иметь повышенные выбросы парниковых газов, по меньшей мере и через 20 лет после затопления [11].

В соответствии с распоряжением от 14 апреля 2016 г. № 670-р Российская Федерация подписала Парижское соглашение 21-й конференции сторон рамочной конвенции ООН по изменению климата, и рекомендации к применению разработанного под эгидой ЮНЕСКО «Руководства по Расчёту парниковых газов в пресноводных водохранилищах», которое представляет из себя стандартные процедуры расчёта основных выбросов парниковых газов из-за создания водохранилищ в бассейнах рек. В связи с планирующимся внесением изменений в действующее законодательство Российской Федерации в части разработки порядка представления отчетности об объеме выбросов парниковых газов возникла насущная потребность в понимании процессов выбросов - поглощений парниковых газов и объёмах выбросов-поглощений парниковых газов в климатических условиях Восточной Сибири.

Во всех климатических сценариях и исследованиях глобального баланса углерода Сибирь рассматривается, как один из наиболее значимых регуляторов углеродосодержащих газов в Северном полушарии [1-3].

В тоже время ряд исследователей отмечает, что выбросы от этих водоемов значительно выше, чем принято считать [12-15]. Поскольку в настоящее время отсутствует единое мнение

относительно выбросов парниковых газов пресноводными водохранилищами и, соответственно, углеродной нейтральности водохранилищ крупных ГЭС, а также относительно оценки их поглощающей способности, требуется проведение измерений выбросов парниковых газов и поглощающей способности пресноводных водохранилищ.

В данной работе рассматривается подход, использованный для обработки сделанных на водохранилище Иркутской ГЭС измерений в целях получения количественной оценки потоков парниковых газов и полученные с его использованием результаты.

**Исходные данные.** Были проведены несколько экспедиций и получены результаты измерений параметров газообмена и других характеристик окружающей среды в различных точках водохранилища в различные моменты времени. Пример покрытия водохранилища измерениями показан на рис. 1.

**Кластерный анализ результатов измерений.** Для оценки суммарных характеристик газообмена необходимо выполнить пространственную и временную интерполяцию сделанных измерений. При этом необходимо учитывать, что измерения в отдельной точке могут определяться различными факторами, поэтому в первую очередь были проанализированы зависимости между полученными показателями. Для анализа зависимостей между измеренными показателями в различных точках использовался кластерный анализ, позволяющий сгруппировать точки, обладающие сходными свойствами.

При кластерном анализе полученных проб был использован следующий набор параметров: широта, долгота, парциальное давление $CO_2$ в воздухе, парциальное давление $CO_2$ в воде (uatm), парциальное давление $CH_4$ в воздухе (uatm), парциальное давление $CH_4$ в воде (uatm), содержание $CO_2$ в воде (mg/l), содержание $CH_4$ в воде (mg/l), атмосферное давление (mmHg), скорость ветра (m/s), температура воздуха (°C), солнечная радиация (r.u.), влажность воздуха (%), температура воды (°C). В пространстве этих координат был выполнен кластерный анализ методом К-средних для результатов каждой из проведённых экспедиций. Посредством обобщения границ найденных кластеров было получено пространственное разбиение водохранилища на зоны (Рис. 2). Точки выполнения измерений, относящиеся к каждой из зон, обладают сходными характеристиками (относятся к одному кластеру и компактно расположены в пространстве параметров).

**Пространственная интерполяция.** По измерениям, сделанным в отдельных точках, необходимо оценить газообмен на всей площади водохранилища. Для этого можно было использовать различные методы пространственной интерполяции, такие как: метод естественных соседей, кригинг и т.д. Однако необходимо учитывать, что измерения в разных точках были выполнены ещё и в разное время, поэтому необходимо усреднять результаты ещё и по времени. Кроме того, при неудачном стечении обстоятельств методы пространственной интерполяции могут распространить особенности отдельной точки (которая может быть и выбросом) на слишком большой фрагмент акватории.

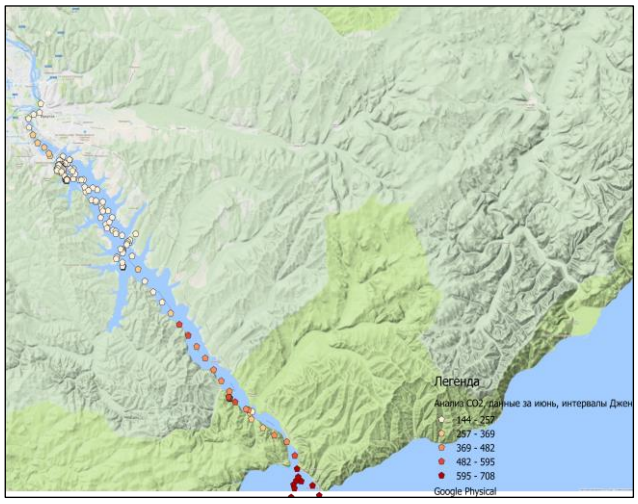
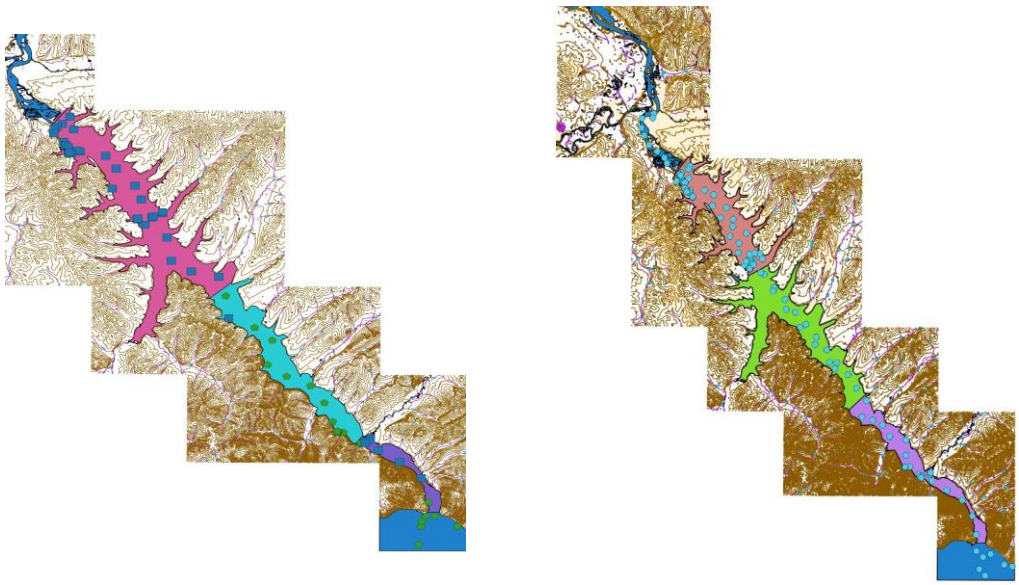



Рис. 1. Точки взятия проб в июне 2017 г.



*а* *б*

Рис. 2. Результат разбиения на зоны Иркутского водохранилища на зоны по поглощению $CO_2$ (*а*) и по выделению $CH_4$ (*б*). С северо-запада на юго-восток: 1 – зона влияния плотины; 2 – промежуточная зона; 3 – зона активного влияния озера Байкал.

В результате был сделан выбор в пользу применения простого усреднения сделанных измерений в рамках зон разбиения. Заметим, что выбор разбиения на зоны будет влиять на окончательный результат усреднения, поскольку при этом вес точки определяется площадью зоны, к которой она отнесена. Основными показателями, которые используются в дальнейших расчётах, являются разности парциальных давлений рассматриваемых в данной работе газов ($CO_2$ и $CH_4$). Также учитывается температура воды и скорость ветра, поскольку эти показатели влияют на скорость газообмена.

**Оценка объёмов газообмена.** Поток $F$ свободного $CO_2$ через единичную площадку поверхности озера пропорционален разности парциальных давлений в воздухе $pCO_{2a}$ и в воде $pCO_2w$ на контакте между ними [1]:

$$F = \alpha(pCO_{2w} - pCO_{2a}) \quad (1)$$

где $\alpha$ – коэффициент скорости растворения $CO_2$, который зависит от ряда факторов: температуры, скорости ветра, солёности воды и т.д. При задании парциального давления в долях атмосферы $\alpha$ имеет размерность мгм$^{-2}$ атм$^{-1}$ сут$^{-1}$. В работе [16] положительный поток направлен в воду, но, вслед за работами [17, 18], мы будем использовать противоположное направление (отрицательный поток направлен в воду).

В работах [17, 18] приводится тот факт, что коэффициент скорости растворения $CO_2$ очень слабо зависит от температуры, так, в диапазоне температур 5-25° он изменяется лишь на 7%. Таким образом, при пренебрежении температурной зависимостью точность полученного результата изменится очень слабо.

В работе [16] с использованием расчёта баланса углерода выводится оценка среднего значения коэффициента скорости растворения $CO_2$ для Байкала

$$\alpha = 6.3 \times 10^5 \text{ мг/(м}^2 \cdot \text{атм} \cdot \text{сут)} = 630\text{г/(м}^2 \cdot \text{атм} \cdot \text{сут)} \quad (2)$$

Разница парциальных давлений традиционно измеряется в миллионных долях атмосферы $\mu$атм = $10^{-6}$атм. При измерении площади в км$^2$ = $10^6$м$^2$ эти коэффициенты ($10^{-6}$ и $10^6$) взаимно сокращаются, поэтому также

$$\alpha = 630\text{г/(км}^2 \cdot \mu\text{атм} \cdot \text{сут)} \quad (3)$$

**Формулы расчёта коэффициента скорости растворения.** Для получения более точных оценок, а также для определения коэффициента скорости растворения метана воспользуемся формулами из [17, 18]. Далее изложим некоторую информацию из статьи [18] в таком виде, который поможет сэкономить время, потраченное на её понимание.

Поток газа из воды $F$ определяется разницей концентраций газа в воде $C_w$ и в воздухе $C_a$.

$$F = k(C_w - C_a) \tag{4}$$

Поток газа $F$ измеряется, как (масса площадь$^{-1}$ время$^{-1}$); $k$ – скорость перехода газа (длина время$^{-1}$); $C_w$ и $C_a$ – концентрации (масса объём$^{-1}$) газа в воде и атмосфере (точнее, в слое воды, непосредственно соприкасающемся с атмосферой). Скорость перехода газа зависит от температуры, скорости ветра и солёности воды.

Поток $F$ газа $G$ также может быть выражен через разность парциальных давлений

$$F = \alpha(pG_w - pG_a) = kK_0(pG_w - pG_a) \tag{5}$$

Здесь $K_0$ – растворимость газа в воде (масса объём$^{-1}$ давление$^{-1}$), а $pG_w$ и $pG_a$ – парциальные давления газа в равновесном состоянии с поверхностью воды и в приповерхностном слое воздуха соответственно (давление). Нам требуется именно эта формула, поскольку все летние измерения содержат парциальные давления, а концентрации во многих точках отсутствуют.

В работах [17, 18] предлагается аппроксимировать зависимость скорости перехода газа от скорости ветра квадратичной функцией. Эта аппроксимация действительна на интервале скоростей ветра 3-15 м/с. При этом отмечается, что ветра более 15м/с дуют лишь 3-5% времени. Примерно также редко случаются ветра менее 3 м/с. Глобальная средняя скорость ветра составляет 7.4 м/с.

Зависимость скорости перехода газа от скорости ветра описывается формулой

$$k = 0.251 \frac{U^2}{\sqrt{Sc/660}} \tag{6}$$

При этом принято измерять $k$ в см/ч. $U$ - средняя скорость ветра в 10 м над водой (м/с), $Sc$ - безразмерное число Шмидта. Единицы измерения коэффициента 0.251 - см/(ч(м/с)$^2$).

Таким образом

$$\alpha = 0.251 \frac{U^2}{\sqrt{Sc/660}} K_0 \tag{7}$$

Поскольку для $CO_2$ отношение $\frac{K_0}{\sqrt{Sc/660}}$ почти постоянно, выражение (7) в этом случае может быть упрощено для морской воды до

$$\alpha_{м,CO_2} = 7.710^{-4}\, U^2 \tag{8}$$

при этом $\alpha_{CO_2}$ измеряется в моль/(м$^2$ год $\mu$атм).

В формуле (7) выделим коэффициент

$$\gamma = \frac{\alpha}{U^2} = 0.251 \frac{K_0}{\sqrt{Sc/660}}, \tag{9}$$

который необходимо вычислить. Для этого число Шмидта вычисляем с использованием таблиц и формул из [3] (его величина зависит от температуры). В табл. 1 приведены значения ряда величин, при разных значениях температуры и солёности воды.

По аналогии возьмём значение $9.68 \cdot 10^{-4}$ при $5^{\circ}C$ в качестве коэффициента $\gamma$ для пресной воды. Сравним полученный коэффициент с оценкой (2). Молярная масса $M_{C02} = 44.01$г/моль.

Табл. 1. Вычисление $\alpha_{CO_2}$.

| **Величина** | **Пресная вода** | | | **Солёная вода** | | |
|---|---|---|---|---|---|---|
| $t,^{\circ}C$ | 0 | 5 | 10 | 0 | 0 | 10 |
| $Sc$ | 1923.6 | 1397.5 | 1032.1 | 2116.8 | 1542.9 | 1143.1 |
| $\beta$ | 1.738 | 1.436 | 1.202 | 1.448 | 1.201 | 1.010 |
| $K_0$ | 0.0775 | 0.0641 | 0.0536 | 0.0646 | 0.0536 | 0.0451 |
| $\frac{\alpha}{U2}$ | $9.99 \cdot 10^{-4}$ | $9.68 \cdot 10^{-4}$ | $9.43 \cdot 10^{-4}$ | $7.93 \cdot 10^{-4}$ | $7.70 \cdot 10^{-4}$ | $7.53 \cdot 10^{-4}$ |

$$K_{\alpha,5°C,C0_2} = \frac{\alpha}{U^2} \cdot M_{C0_2} = 9.68 \cdot 10^{-4}(\frac{\text{моль}}{\text{м}^2\ \mu\text{атм год}\ (\text{м/с})^2}) \cdot 44.01\text{г/моль} =$$
$$9.68 \cdot 10^2 \cdot 44.01/365(\frac{\text{г}}{\text{км}^2\ \mu\text{атм сут}\ (\text{м/с})^2}) = 116.717(\frac{\text{г}}{\text{км}^2\ \mu\text{атм сут}\ (\text{м/с})^2}) \quad (10)$$

Отсюда с использованием глобальной средней скорости ветра 7.4м/с получим оценку

$$\alpha_{CO_2}{}^1 = 6391\text{г/(км}^2 \cdot \mu\text{атм} \cdot \text{сут}) \quad (11)$$

которая в 10 раз превышает предыдущую.

С другой стороны, если учесть, что для Иркутска средняя скорость ветра летом и осенью составляет 2.0 м/с [19], то получим оценку

$$\alpha_{CO_2}{}^2 = 467\ \text{г/(км}^2 \cdot \mu\text{атм} \cdot \text{сут}) \quad (12)$$

которая гораздо ближе к предыдущей. Далее будем использовать эту же среднюю скорость ветра и для метана.

Заметим, что при средней скорости ветра 2.0 м/с мы используем формулы из работы [3] за пределами интервала скоростей ветра 3 – 15 м/с, для которого получена аппроксимация. В частности, при полном штиле получим значение $\alpha_{CO_2}{}^2 = 0$, что явно не соответствует действительности. Истинный график функции, описывающей данную зависимость, должен быть монотонной функцией, принимающей при $U = 0$ положительное значение, соответствующее коэффициенту скорости растворения при штиле. Таким образом, истинный график должен постепенно смещаться выше, чем используемый нами, при изменении скорости ветра от 3 до 0 м/с, поэтому можно считать, что в результате мы получаем оценку снизу для коэффициента скорости растворения, что хорошо соответствует целям данного проекта.

Теперь аналогичным образом вычислим $\alpha_{CH_4}{}^2$ сразу для пресной воды. Заметим, что скорость изменения величины $\frac{\alpha}{U^2}$ для метана не слишком отличается от скорости изменения аналогичной величины для $CO_2$, поэтому также можно в первом приближении пренебречь её зависимостью от температуры.

Молярная масса метана $M_{CH_4} = 16.04$г/моль.

$$K_{\alpha,5°C,CH_4} = \frac{\alpha}{U^2} \cdot M_{CH_4} = 3.34 \cdot 10^{-5}(\frac{\text{моль}}{\text{м}^2\ \mu\text{атм год}\ (\text{м/с})^2}) \cdot 16.04\text{г/моль} =$$
$$33.4 \cdot 16.04/365(\frac{\text{г}}{\text{км}^2\ \mu\text{атм сут}\ (\text{м/с})^2}) = 1.468(\frac{\text{г}}{\text{км}^2\ \mu\text{атм сут}\ (\text{м/с})^2}) \quad (15)$$

Для средней скорости ветра 2.0 м/с получим оценку

$$\alpha_{CH}{}^2{}_4 = 5.87\text{г/(км}^2 \cdot \mu\text{атм} \cdot \text{сут}) \quad (16)$$

**Интерполяция по времени.** Далее между моментами проведения экспедиций выполняется линейная интерполяция по времени вычисленных средних и суммарных значений для каждой из зон. В результате получаются оценки, характеризующие значения используемых показателей в каждый из моментов времени в каждой точке водохранилища. На Рис 3 и Рис 4 показаны полученные графики для лета 2017 г.

Табл. 2. Вычисление $\alpha_{CH_4}$.

| ***t,°C*** | **0** | **5** | **10** | **15** |
|---|---|---|---|---|
| *Sc* | 1909.4 | 1399.7 | 1043.1 | 794.0 |
| $\beta$ | 0.0574 | 0.0495 | 0.0434 | 0.0385 |
| $K_0$ | 0.00256 | 0.00221 | 0.00194 | 0.00172 |
| $\frac{\alpha}{U2}$ | $3.31 \cdot 10^{-5}$ | $3.34 \cdot 10^{-5}$ | $3.39 \cdot 10^{-5}$ | $3.45 \cdot 10^{-5}$ |

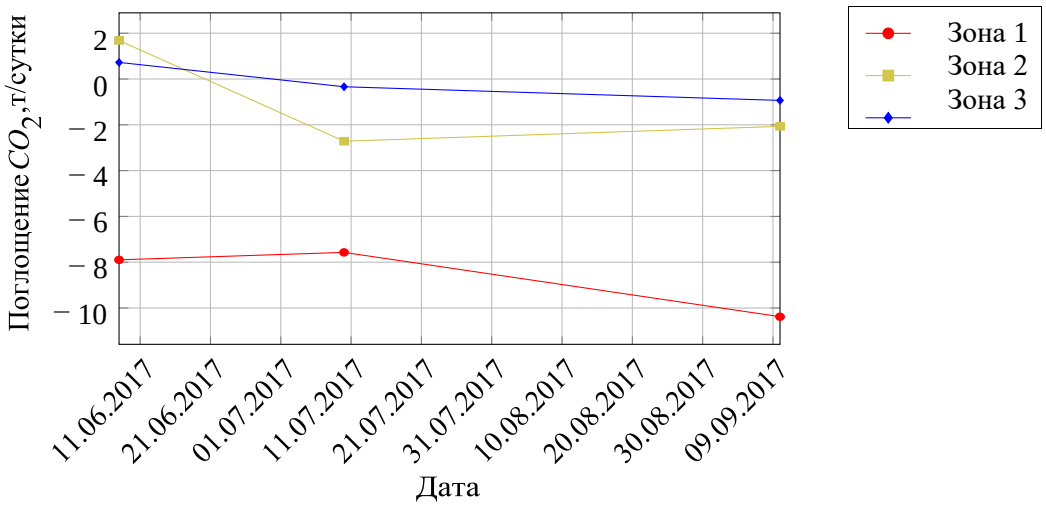



Рис. 3. График поглощения $CO_2$ Иркутским водохранилищем.

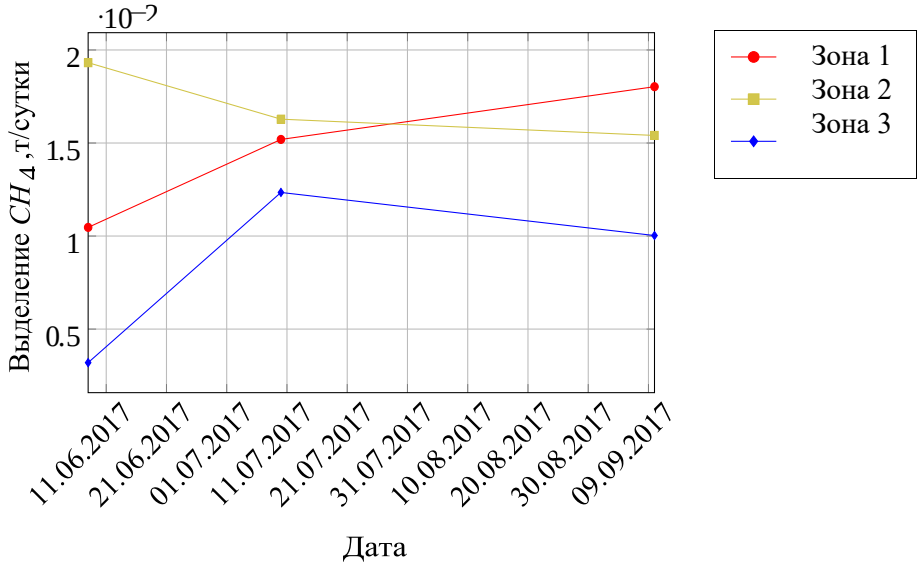



Рис. 4. График выделения $CH_4$ Иркутским водохранилищем.

В результате интегрирования полученных интерполированных оценок получаем суммарные оценки объёмов газообмена за период наблюдений. При этом сначала получены более грубые интегральные оценки, использующие средние значения скорости ветра для рассматриваемой территории. Далее с использованием метеоданных ближайших к водохранилищу метеостанций эти оценки были уточнены. Заметим, что оценки, получаемые с использованием конкретных метеоданных, больше отражают ситуацию за конкретный сезон, а оценки, получаемые с помощью среднемноголетних показателей для территории должны быть несколько ближе к средним оценкам газообмена, хотя, конечно, для получения более точных средних оценок необходимо регулярно выполнять подобные измерения.

**Заключение.** В результате проведенных измерений по Иркутскому водохранилищу и их анализа были получены следующие результаты.

Согласно оценке, полученной с использованием средней скорости ветра, за период наблюдений, продолжительность которого составила 94 суток, через поверхность Иркутского водохранилища было поглощено 1002 т $CO_2$. Выделено было 3.116 т $CH_4$.

При использовании информации о скорости ветра на Иркутской метеостанции эти оценки незначительно изменяются до поглощения 936 т $CO_2$ (−7%) и выделения 3.679 т $CH_4$ (+18%).

Однако, если использовать информацию о скорости ветра за время наблюдений на Иркутском водохранилище, то первоначальные оценки придётся увеличить в 1.5 раза до поглощения 1500 т $CO_2$ и выделения 4.674 т $CH_4$.

Таким образом, предложенный метод обработки измерений позволяет оценить фактические объёмы газообмена через поверхность водохранилища. Несмотря на достаточно большую площадь водохранилища, эти объёмы оказываются достаточно небольшими. Они оказываются на порядки меньше объёма растворённых газов, поступающих в водохранилище из Байкала и выходящих из него в нижний бьеф Иркутской ГЭС.

## ЛИТЕРАТУРА


[1] *Louis VL St., Kelly CA, Duchemin É, Rudd JW, Rosenberg DM.* Reservoir surfaces as sources of greenhouse gases to the atmosphere: A global estimate // BioScience. 2000. Vol. 50, is 9. P. 766-775.

[2] *Rudd JWM, Hamilton RD.* Methane cycling in a eutrohic shield lake and its effects on whole lake metabolism // Limnology and Oceanography. 1978. Vol. 23. P. 337-348.

[3] *Barros N., Cole J.J., Tranvik L.J., Prairie Y.T., Bastviken D., del Giorgio P.A,. Roland F., Huszar V.L.M.* Carbon emission from hydroelectric reservoirs linked to reservoir age and latitude // Nature Geoscience. 2011. Vol. 4. P. 593-596.

[4] UNESCO–IHA International Hydropower Association, United Nations Educational, Scientific, and Cultural Organization GHG Risk Assessment Tool (Beta Version). 2012. User Manual IHA (23 August 2016).

[5] *Abril G., Guérin F., Richard S. et al.* Carbon dioxide and methane emissions and the carbon budget of a 10-year old tropical reservoir (Petit Saut, French Guiana) // Global Biogeochemical Cycles. 2005. P. 19.

[6] *Bastien J., Demarty M., Tremblay A.* $CO_2$ and $CH_4$ diffusive and degassing fluxes from 2003 to 2009 at Eastmain 1 reservoir, Québec, Canada // Inland Waters. 2011. vol. 1. p. 113-123.

[7] *Teodoru C.R. et al.* The net carbon footprint of a newly created boreal hydroelectric reservoir // Global Biogeochemical Cycles. 2012. p. 26 (art. GB2016).

[8] *Kemenes A., Forsberg B.R., Melack J.M.* $CO_2$ emissions from a tropical hydroelectric reservoir (Balbina, Brazil) // Journal of Geophysical Research: Biogeosciences. 2011. vol. 116. (art. G03004).

[9] *Brothers S.M., Giorgio P.A., Teodoru C.R. et al.* Landscape heterogeneity influences carbon dioxide production in a young boreal reservoir // Canad. J. Fisheries and Aquatic Sci. 2012. Vol. 69. P. 447-456.

[10] *Hendzel L.L., Matthews C.J.D., Venkiteswaran J.J. et al.* Nitrous oxide fluxes in three experimental boreal forest reservoirs // Env. Sci. and Tech. 2005. Vol. 39. P. 4353-4360.

[11] *Matthews C.J., Joyce E.M., Louis V.L.S., Schiff S.L., Venkiteswaran J.J., Hall B.D., Bodaly R.A., Beaty K.G.* Carbon dioxide and methane production in small reservoirs flooding upland boreal forest // Ecosystems. 2005. vol. 8. p. 267-285.

[12] *DelSontro T., McGinnis D.F., Wehrli B., Ostrovsky I.* Size does matter: Importance of large bubbles and small-scale hot spots for methane transport // Environmental Science and Technology. 2015. vol. 49. p. 1268-1276.

[13] *Bridget R. Deemer, John A. Harrison, Siyue Li, Jake J. Beaulieu, Tonya DelSontro, Nathan Barros, José F. Bezerra-Neto, Stephen M. Powers, Marco A. dos Santos, J. Arie Vonk* Greenhouse Gas Emissions from Reservoir Water Surfaces: A New Global Synthesis // BioScience. 2016. Vol. 66. Issue 11. p. 949-964. DOI: 10.1093/biosci/biw117nk.

[14] *Fearnside P.M., Pueyo S.* Greenhouse-gas emissions from tropical dams // Nat. Clim. Change. 2012. № 2. p. 382-384.

[15] *Luke Gibson, Elspeth N. Wilman, William F. Laurance.* How Green is 'Green' Energy? // Trends in Ecology & Evolution. 2017. Vol 32. Issue 12. p. 922-935, DOI: 10.1016/j.tree.2017.09.007.

[16] *Мизандронцев И.Б., Мизандронцева К.Н.* Парциальное давление и обмен свободной $CO_2$ в системе вода-атмосфера на Байкале, Водные ресурсы. 1994, том 21. №1. с. 72-79.

[17] *Wanninkhof R.* Relationship between wind speed and gas exchange over the ocean // J. Geophys. Res. 1992. Vol. 97(C5). P. 7373-7382. DOI: 10.1029/92JC00188.

[18] *Wanninkhof Rik* Relationship between wind speed and gas exchange over the ocean revisited // Limnol. Oceanogr. Methods. 2014. Vol. 12. DOI: 10.4319/lom.2014.12.351.

[19] Ветра в Иркутской области **http://energywind.ru/recomendacii/karta-rossii/sibir/irkutskaya-oblast** (дата обращения 15.07.2019).

# КОМБИНАЦИЯ НЕЙРОННОЙ СЕТИ И ЛИНЕЙНОЙ ФИЛЬТРАЦИИ ДЛЯ ОБНАРУЖЕНИЯ ОБЪЕКТОВ

*Шакенов А.К.*
Институт автоматики и электрометрии СО РАН, Новосибирск

Рассмотрено несколько подходов к применению нейронных сетей для обнаружения объектов на пространственно неоднородных фонах. Реализован метод построения классификатора для обнаружения объектов непосредственно по наблюдаемым фрагментам. Предложен подход, заключающийся в комбинации метода оптимальной линейной фильтрации и сверточных нейронных сетей. Показано, что применённый подход позволяет уменьшить вероятность ложной тревоги при сохранении вероятности обнаружения объекта.
*Ключевые слова: обнаружение и распознавание объектов, сверточные нейронные сети, машинное обучение, малоразмерные объекты.*

**Введение.** Задача обнаружения малоразмерных слабоконтрастных объектов активно исследуется в течение последних десятилетий. Из-за широкого разброса условий регистрации данных и сопутствующих помех [1] исследования в этой области продолжают оставаться актуальными, что подтверждается большим количеством работ по этой тематике, опубликованным в последние годы. При обнаружении малоразмерных слабоконтрастных объектов их форма и размер определяются аппаратной функцией системы, и не содержат достаточно информации для надежного обнаружения. Важной особенностью алгоритмов обнаружения таких объектов является необходимость оценить и исключить из рассмотрения подстилающий фон. Наиболее эффективным подходом для этого является пространственно-временная фильтрация серии последовательных изображений [2]. Однако в ряде случаев из-за особенностей геометрии съемки или вычислительных ограничений системы обработки данных необходимо производить оценку и подавление фона по одному изображению. Известен подход к решению рассматриваемой задачи, позволяющий получить оптимальный линейный фильтр для случая стационарного фона с известной ковариационной матрицей [3]. Активно разрабатываются и применяются различные алгоритмы оценки и фильтрации фона по наблюдаемой локальной окрестности, например, билатеральная [4], медианная [5] фильтрация, оптимальный линейный прогноз [6, 7], другие эвристические методы [8-11]. Метод оптимальной линейной фильтрации разработан в предположении, что статистические свойства фона одинаковы по всему полю кадра. Это предположение может не выполняться для широкого круга наблюдаемых фонов. Этим обусловлена актуальность поиска новых подходов к решению задачи обнаружения малоразмерных слабоконтрастных объектов на пространственно неоднородных фонах.

В последнее время активно развиваются методы распознавания и обнаружения объектов с использованием обучаемых нейронных сетей [12]. Примеры использования этого подхода к обнаружению малоразмерных объектов можно найти в работах [13-16]. Поскольку нейронные сети используют достаточно большое количество промежуточных признаков в процессе обработки фрагмента, можно ожидать, что их применение позволит улучшить результаты линейной фильтрации именно на пространственно неоднородных фонах. Кроме того, возможность обучать сеть непосредственно по наблюдаемым данным, дает возможность легко адаптировать данный подход к большому количеству фонов и наблюдаемых объектов.

**Постановка задачи.** Необходимо разработать алгоритм обнаружения объектов на неоднородных фонах, улучшающий характеристики обнаружения по сравнению с алгоритмом оптимальной линейной фильтрации за счет применения обучаемых нейронных сетей.

**Обнаружение объектов с обучением по наблюдаемым фрагментам.** Одним из способов применения нейронных сетей для обнаружения объектов является обучение классификатора, который характеризует каждый фрагмент наблюдаемого изображения как содержащий объект, либо только фон. При этом размер обрабатываемого фрагмента выбирается равным размеру изображения объекта. Процедура обнаружения при таком подходе заключается в последовательном переборе всех фрагментов изображения и проверки их на наличие объекта с

помощью обученного классификатора. Для обнаружения использовалась трехслойная сверточная нейронная сеть, схематически изображенная на рис. 1.

Первый слой сети выполняет свертку с 32 различными фильтрами размера 9x9 значений и уменьшение размера полученных в результате свертки изображений в два раза. Уменьшение осуществляется методом выбора наибольшего элемента из окрестности размера 2x2 пиксела. Второй слой сети аналогичным образом выполнял свертку с двумя фильтрами размера 9x9 значений и уменьшение размера выходных массивов в два раза. Третий слой переводит полученный массив данных в один вектор признаков, содержащий 1024 элементов. Полученный вектор признаков затем характеризуется как содержащий, либо не содержащий объект.

**Комбинация оптимальной линейной фильтрации и нейронной сети.** Детектирование объектов описанным выше способом достаточно трудоемко с вычислительной точки зрения, поскольку каждый фрагмент изображения необходимо обрабатывать нейронной сетью, которая содержит каскад значительного числа фильтров. Метод оптимальной линейной фильтрации дает достаточно хорошие результаты для широкого круга реальных фонов, при этом, если ковариационная матрица фона оценена заранее, то вычисление состоит в фильтрации одним линейным фильтром. Таким образом, возникает идея на первом этапе обработки использовать оптимальный линейный фильтр, а затем применить обученную нейронную сеть. Зарегистрированное изображение можно представить в векторной форме следующим образом:

$$f_{src} = f_o + n,$$

где $f_o$ – вектор объекта, $n$ – вектор коррелированного шума (фона). Если $K$ – матрица ковариации шума, то оптимальный в смысле увеличения отношения сигнал/шум линейный фильтр m, имеет вид [2]:

$$m = K^{-1} f_o.$$

На практике матрица K, как правило, не известна. В данной работе использовалась численная оценка матрицы K, полученная непосредственно по входным изображениям фона. Рассчитав таким образом линейный фильтр, дальнейшую обработку можно провести по схеме, приведенной на рис. 2.

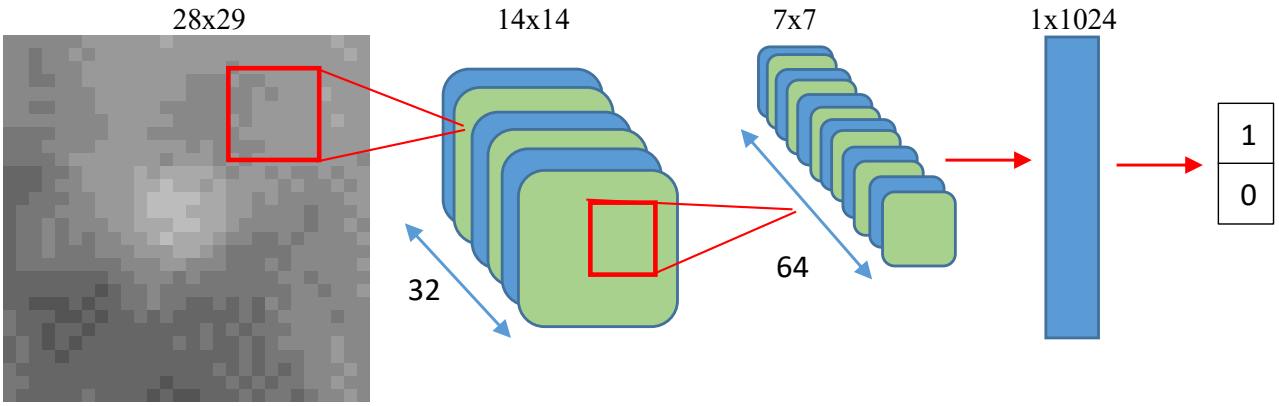



Рис. 1. Схема использованной нейронной сети.

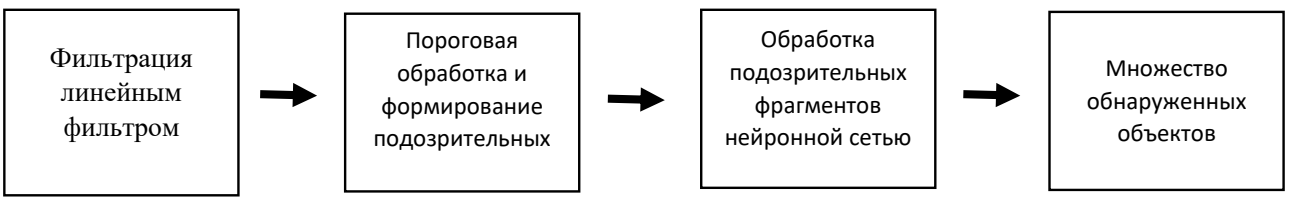



Рис. 2. Комбинация оптимальной линейной фильтрации и нейронной сети.

**Исходные данные и обучение сети.** Для экспериментов использовались изображения Земли со спутника Электро Л-1 доступные в открытом доступе в сети Интернет [17]. В работе рассмотрены точечные объекты, размеры и форма изображения которых определяются аппаратной функцией системы. Форма объекта моделировалась с использованием функции Гаусса, применялся аддитивный метод нанесения объекта. Для обучения сети распознаванию фрагментов после процедуры фильтрации данные получались следующим образом. На исходное изображение наносилось значительное количество объектов на расстоянии в несколько раз превышающем размер объектов. Фрагменты изображения, содержащие объекты, сохранялись и использовались для обучения сети. Строился согласованный фильтр и проводилась фильтрация изображения с нанесенными объектами, а также исходного изображения без объекта. По обработанному изображению выбирался порог яркости, определяющий вероятность обнаружения объекта и ложной тревоги. На исходном изображении, не содержащем объектов выбирались фрагменты, отклик на фильтр в которых превышал пороговое значение (ложно обнаруженные фрагменты). Эти фрагменты в дальнейшем использовались при обучении нейронной сети, как примеры фона, не содержащего объекты.

**Экспериментальные результаты.** Для сравнения эффективности оптимальной линейной фильтрации и нейронной сети проводился следующий эксперимент. Одно случайно выбранное изображение фона использовалось для обучения классификатора, дальнейшее сравнение алгоритмов проводилось на других фоновых изображениях. В качестве объекта была использована гауссова функция с максимальной интенсивностью равной одному СКО фона и параметром $\sigma$ равным 3. Для получения тестового множества фрагментов на изображения фонов наносилось 14000 объектов. Фрагменты с нанесенными объектами использовались для оценки вероятности обнаружения объектов. Для оценки вероятности ложной тревоги из исходного изображения в произвольных местах вырезалось 14000 фрагментов такого же размера. Полученные множества обрабатывались обученным классификатором. Результаты обработки нескольких фоновых изображений приведены в табл. 1.

В столбце «Сеть» таблицы 1 приведены значения ложной тревоги при использовании алгоритма обнаружения на основе обученного классификатора, а столбце «ОЛФ» – при использовании оптимальной линейной фильтрации. Из таблицы видно, что в проведенных экспериментах нейронная сеть показала результаты хуже, чем оптимальная линейная фильтрация. Скорее всего это связано с тем, что в отсутствие четкой методологии достаточно сложно подобрать обучающее множество для получения результата близкого к оптимальному. Приведенные данные показывают, что оба алгоритма дают стабильные результаты при обработке различных входных изображений.

Табл. 1. Сравнение нейронной сети и оптимальной линейной фильтрации.

| **Текстура** | **Вероятность обнаружения** | **Ложная тревога** | |
|---|---|---|---|
| | | Сеть | ОЛФ |
| Фон 1 | 0,9993 | 0,0057 | 0,0011 |
| Фон 2 | 0,9998 | 0,0061 | 0,0007 |
| Фон 3 | 0,9996 | 0,0085 | 0,0004 |
| Фон 4 | 0,9997 | 0,0055 | 0,0005 |

Табл. 2. Комбинация нейронной сети и оптимальной линейной фильтрации.

| ОЛФ 1 | | Нейросеть | | **ОЛФ+Сеть** | | **ОЛФ** |
|---|---|---|---|---|---|---|
| $\alpha_1$ | $\beta_1$ | $\alpha_N$ | $\beta_N$ | $\boldsymbol{\alpha_C}$ | $\boldsymbol{\beta_C}$ | $\boldsymbol{\beta_0}$ |
| 0,999 | $1{,}17*10^{-3}$ | 0,968 | 0,097 | **0,967** | $\mathbf{1{,}11*10^{-4}}$ | $\mathbf{1{,}59*10^{-4}}$ |
| 0,98 | $1{,}93*10^{-4}$ | 0,969 | 0,404 | **0,95** | $\mathbf{7{,}56*10^{-5}}$ | $\mathbf{1{,}27*10^{-4}}$ |
| 0,96 | $1{,}38*10^{-4}$ | 0,971 | 0,480 | **0,932** | $\mathbf{6{,}62*10^{-5}}$ | $\mathbf{1{,}07*10^{-4}}$ |
| 0,94 | $1{,}16*10^{-4}$ | 0,976 | 0,5 | **0,918** | $\mathbf{5{,}61*10^{-5}}$ | $\mathbf{9{,}92*10^{-5}}$ |
| 0,92 | $1{,}0*10^{-4}$ | 0,979 | 0,545 | **0,901** | $\mathbf{5{,}29*10^{-5}}$ | $\mathbf{9{,}11*10^{-5}}$ |
| 0,90 | $9{,}03*10^{-5}$ | 0,980 | 0,564 | **0,882** | $\mathbf{4{,}94*10^{-5}}$ | $\mathbf{8{,}13*10^{-5}}$ |

В табл. 2 приведены результаты работы алгоритма, комбинирующего оптимальную линейную фильтрацию и нейронную сеть. В предыдущем эксперименте на вход нейронной сети подавались все фрагменты обрабатываемого изображения, в данном эксперименте – только фрагменты подозрительные на наличие объекта по результатам линейной фильтрации. В данном случае использовалось изображение «Фон 4» с 900 нанесенными объектами. Эксперименты с другими фонами дали схожие результаты. На первом этапе использовался алгоритм оптимальной линейной фильтрации, при этом подбирались значения порога, дающие вероятности обнаружения, указанные в таблице в столбце $\alpha_1$. Каждое значение порога определяет два множества: $A_1$ – множество правильно обнаруженных объектов, и $B_1$ – множество ложно обнаруженных фрагментов, не содержащих объекты. Обозначим $N_{obj}$ – количество всех нанесенных объектов, $N_{pix}$ – количество всех пикселов изображения. Используя множества $A_1$ и $B_1$ были оценены вероятности обнаружения и ложной тревоги при использовании алгоритма оптимальной линейной фильтрации, приведенные в столбце «ОЛФ1». Вероятность обнаружения оценивалась как $\alpha_1 = |A_1|/N_{obj}$, вероятность ложной тревоги $\beta_1 = |B_1|/N_{pix}$. Затем множества $A_1$ и $B_1$ подавались на вход нейронной сети. В результате получались два множества: $A_2$ – множество объектов правильно классифицированных нейросетью и $B_2$ – множество ложных фрагментов неверно классифицированных нейросетью как объекты. Эффективность работы нейросети при обработке множеств $A_1$ и $B_1$ показана в столбце «Нейросеть». В нем указаны значения $\alpha_N = |A_2|/|A_1|$ и $\beta_N = |B_2|/|B_1|$. Итоговая вероятности обнаружения и ложной тревоги при рассмотренном методе обнаружения приведена в столбце «ОЛФ+Сеть». В нем указаны значения $\alpha_C = |A_2|/N_{obj}$ и $\beta_C = |B_2|/N_{pix}$. Для сравнения предложенного подхода были измерены вероятности ложной тревоги при применении алгоритма оптимальной линейной фильтрации для вероятностей обнаружения, указанных в столбце $\alpha_C$. Эти значения приведены в столбце «ОЛФ» и обозначены как $\beta_0$.

При сравнении значений $\beta_C$ и $\beta_0$ видно, что предложенный подход позволил на 40-60 процентов уменьшить вероятность ложной тревоги, при той же вероятности обнаружения. Характер изменений значений $\alpha_N$ и $\beta_N$ показывает, что результаты обнаружения с помощью нейронной сети имеют корреляцию с результатами обнаружения алгоритмом оптимальной линейной фильтрации. Для получения меньших значений $\alpha_1$ необходимо использовать более высокий порог на этапе пороговой обработки, что дает множества $A_1$ и $B_1$, содержащие фрагменты с большей интенсивностью отклика на фильтр. Поскольку основным признаком наличия объекта является дополнительная зарегистрированная интенсивность, то можно предположить, что множество истинных объектов с большей интенсивностью $A_1$ становится проще для правильного распознавания, а множество ложных фрагментов с большей интенсивностью $B_1$ – сложнее. Этим можно объяснить характер изменения величин $\alpha_N$ и $\beta_N$. Уменьшение $\beta_C$ по отношению к $\beta_0$ объясняется скорее всего тем, что нейросеть использует дополнительные к величине отклика на фильтр признаки, применение которых позволяет улучшить итоговые результаты.

**Заключение.** В проведенных экспериментах прямое применение нейронной сети для классификации фрагментов в рассмотренном диапазоне вероятностей обнаружений не позволило улучшить результаты, полученные методом оптимальной линейной фильтрации. При этом показана возможность эффективно использовать комбинацию оптимальной линейной фильтрации и нейронной сети. В результате применения предложенного подхода была повышена эффективность детектирования объектов, вероятность ложной тревоги была снижена на 40-60 процентов при той же вероятности обнаружения объекта. Дальнейшие исследования могут быть направлены на более точную настройку параметров сети и на использование больших объемов данных в процессе обучения.

## ЛИТЕРАТУРА


[1] *Gong Cheng, Junwei Han A* survey on object detection in optical remote sensing images // ISPRS Journal of Photogrammetry and Remote Sensing. 2016. 117. P. 11-28

[2] *Киричук В.С., Косых И.П., Попов С.А., Синельщиков В.В.* Подавление квазистационарного фона в последовательности изображений посредством межкадровой обработки // Автометрия, 2014, т. 50, № 2. С. 3-13.

[3] *Pratt W.K.* Digital image processing: PIKS Scientific Inside. PixelSoft, Inc. Los Altos, California p. 662.

[4] *Tae-Wuk Bae, Kyu-Ik Sohng* Small Target Detection Using Bilateral Filter Based on Edge Component // Springer, J Infrared Milli Terahz Waves 2010 Vol. 31, p. 735-743.

[5] *Deshpande S.D., Er M.H., Ronda V., Chan Ph.* Max-mean and max-median filters for detection of small-targets // Proc. SPIE 3809 (1999): p. 74-83

[6] *Soni T., Zeidler R., Ku W.H.* Performance evaluation of 2D adaptive prediction filters for detection of small object in image data // IEEE Transactions on Image Processing 1993 2 (3), p. 327-340.

[7] *Ffrench P. A., Zeidler J. R., Ku W. H.* Enhanced detectability of small objects in correlated clutter using an improved 2-D adaptive lattice algorithm // IEEE Transactions on Image Processing. 1997.

[8] *Hong P., Wang C., Zhang Z.* Weak point target detection in the complicated infrared background // Proc. SPIE 2011, Vol. 8200, International Conference on Optical Instruments and Technology: Optoelectronic Imaging and Processing Technology,

[9] *DONG Yu-xing, LI Yan, ZHANG* Hai-bo Research on Infrared Dim-point Target detection and Tracking under Sea-Sky-Line Complex Background // International Symposium on Photoelectronic Detection and Imaging 2011: Advances in Infrared Imaging and Applications. Proc. of SPIE Vol. 8193 (2011).

[10] *Иванов В.А., Киричук В.С., Косых В.П., Синельщиков В.В.* Особенности обнаружения точечных объектов в изображениях, формируемых матричным приемником // Автометрия, 2016, т. 52, № 2, с. 10-19.

[11] *Шакенов А.К.* Алгоритмы подавления фона в задаче обнаружения точечных объектов по изображениям // Автометрия, 2014, т. 50, № 4, с. 81-87.

[12] *Weibo Liua, Zidong Wanga, Xiaohui Liua, Nianyin Zengb, Yurong Liucd, Fuad E.Alsaadi* A survey of deep neural network architectures and their applications // Neurocomputing. 2017. 234. P. 11-26.

[13] *Shangnan Zhao, Yong Song, Yufei Zhao, Yun Li, Xu Li, Yurong Jiang, Lin Li* Infrared dim small target segmentation method based on ALI-PCNN model // Proc. SPIE. 2017. 10459. P 104590A-1-104590A-9.

[14] *Junhwan Ryu, Sungho Kim* Small infrared target detection by data-driven proposal and deep learning-based classification // Proc. SPIE. 2018. 10624. P. 106241J.

[15] *Zunlin Fan, Duyan Bi, Lei Xiong, Shiping Ma, Linyuan He, Wenshan Ding* Dim infrared image enhancement based on convolutional neural network // Neurocomputing. 2018. 272. P. 396-404.

[16] *Peng Zhang Jianxun Li* Neural-network-based single-frame detection of dim spot target in infrared images // Optical Engineering. 2007. 46. P. 076401.

[17] Электро-Л / Земля из космоса. **http://electro.ntsomz.ru** (дата обращения 01.02.2019).

# ПРОГРАММНЫЙ КОМПЛЕКС АТМОСФЕРНОЙ КОРРЕКЦИИ ДАННЫХ РОССИЙСКИХ СПУТНИКОВЫХ ПРИБОРОВ

*Белов В.В., Гриднев Ю.В., Зимовая А.В., Тарасенков М.В., Энгель М.В.*
Институт оптики атмосферы им. В.Е. Зуева СО РАН, Томск

Представлены результаты работы по созданию программного обеспечения для тематической обработки данных российских спутниковых систем, созданного на основе алгоритма атмосферной коррекции, разработанного в ИОА СО РАН. Программный комплекс позволяет восстанавливать коэффициенты отражения земной поверхности в видимом и ближнем ИК-диапазоне в приближении однородной поверхности. В состав программного обеспечения входят конверторы для чтения входных форматов спутниковых данных, модуль атмосферной коррекции, модуль для создания модели атмосферы.
*Ключевые слова: спутниковые данные, статистическое моделирование, атмосферная коррекция.*

**Введение.** В 2018 году в ИОА СО РАН выполнена работа по созданию программного и информационного обеспечения данных некоторых приборов дистанционного зондирования Земли (ДЗЗ) российской спутниковой группировки. Программный комплекс предназначен для восстановления коэффициентов отражения земной поверхности в видимом и ближнем ИК-диапазонах на основе принимаемого спутниковыми приборами оптического излучения.

Одним из способов решения задачи атмосферной коррекции спутниковых изображений земной поверхности является решение уравнения переноса излучения (УПИ) через многокомпонентную среду, где необходимо учитывать механизмы поглощения и рассеяния излучения молекулами и атмосферными частицами (такими, как аэрозоль, облака, осадки, атмосферные газы, и т.д.).

Для расчета интенсивности и потоков, формирующих принимаемый спутниковым прибором сигнал, обычно используют два подхода, основанные на применении численных методов. Одним является метод дискретных ординат (DISORT), реализованный в таких известных программных кодах как MODTRAN, 6S, Streamer. Метод DISORT используется, главным образом, для спектральных областей с преобладанием рассеяния [[1], [2]]. Для областей с сильным атмосферным поглощением используется более точный алгоритм k-корреляции [[3]].

Представленное программное обеспечение создано с использованием методических средств проведения атмосферной коррекции спутниковых изображений земной поверхности, разработанных в ИОА СО РАН [[4], [5]]. Предложенные алгоритмы основаны на методе Монте-Карло как на наиболее универсальном методе решения УПИ, практически не имеющем ограничений на геометрию задачи и оптические параметры атмосферы.

Однако для реализации методов, основанных на решении УПИ, необходимо наличие данных об оптическом и метеорологическом состоянии атмосферы в момент проведения спутниковых измерений. При создании программного обеспечения для проведения атмосферной коррекции спутниковых изображений также решался вопрос определения и получения оптимального набора данных об атмосфере, позволяющего произвести расчет искажающих характеристик атмосферы с приемлемой точностью; и создания на основе этих данных оптической модели атмосферы.

**Алгоритмическая основа.** В ИОА СО РАН разработан алгоритм атмосферной коррекции для измерений в видимом и ближнем ИК диапазонах волнового спектра, более качественно учитывающий боковой подсвет и дополнительную освещенность земной поверхности отраженным излучением. В рамках данного алгоритма задача восстановления распределения коэффициентов отражения по земной поверхности решается в следующей постановке (рис. 1). Рассматривается сферическая система атмосфера – земная поверхность. Атмосфера считается разделенной на сферические однородные слои, оптические параметры которых задаются генератором оптических моделей LOWTRAN-7 [[6]]. Земная поверхность представляет собой

ламбертовскую поверхность с неизвестным распределением коэффициента отражения по поверхности. На высоте $h_d$ от земной поверхности располагается пассивная спутниковая система, ориентированная в направлении $\omega_d$ и ведущая наблюдение за участком земной поверхности. Пусть пространственное разрешение оптического приемника, формирующего изображение, постоянно в пределах наблюдаемой области. На верхнюю границу атмосферы падает параллельный поток солнечного излучения в направлении $\omega_{sun}$. Принимаемое спутниковой системой излучение состоит из: солнечной дымки $I_{sun}$ – излучения Солнца рассеянного в атмосфере и не взаимодействовавшего с земной поверхностью, нерассеянного излучения, отраженного от наблюдаемого участка земной поверхности $I_0$, а также поверхностной дымки $I_{surf}$ – рассеянного излучения, отраженного от земной поверхности.

Требуется, зная оптические параметры атмосферы и значения интенсивности, измеренные спутниковой системой, восстановить распределение коэффициента отражения по наблюдаемой области земной поверхности.

Для решения задачи используется теория линейных систем в приложении к оптике дисперсных сред [[7]]. Реализация алгоритма требует существенных затрат машинного времени, поскольку расчет производится для каждого пикселя спутникового изображения. Для ускорения расчетов разработчиками метода предложены следующие приемы:

Разбиение изображения на изопланарные зоны, в пределах каждой из которых функцию размытия точки (ФРТ) можно считать постоянной, что позволяет рассчитывать ФРТ один раз для каждого пикселя изопланарной зоны.

1. Задание радиуса бокового подсвета.
2. Задание радиуса области формирования дополнительной освещенности.
3. Расчет интенсивности солнечной дымки в узловых точках вместо расчета для каждого направления визирования спутниковой системы.

Алгоритм атмосферной коррекции спутниковых изображений земной поверхности в видимом и ближнем ИК диапазонах длин волн позволяет корректно учитывать следующие факторы:

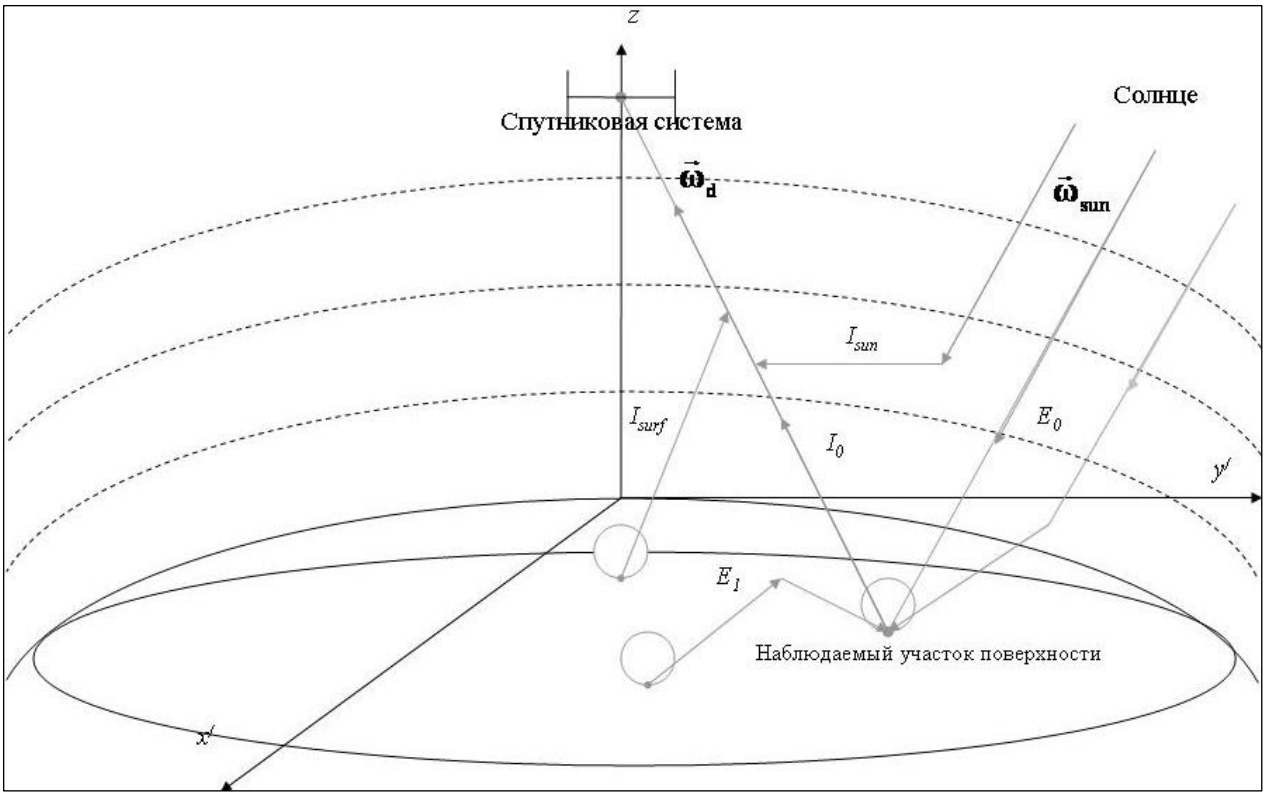



Рис. 1. Геометрическая схема постановки задачи.

а) интенсивность солнечной дымки;
б) боковой подсвет в области изопланарности изображений;
в) дополнительный подсвет поверхности, обусловленный процессом отражения излучения от нее;
г) сферичность земной поверхности.

Однако в данной версии программного обеспечения задача восстановления коэффициентов отражения решена в однократном приближении с использованием формул (1), (2).

$$r_{surf,i} = \frac{Q_i/E_0}{1+\gamma_1 Q_i/E_0} \quad (1)$$

$$\overline{Q_i} = \frac{I_{sum,i}-I_{sun,i}}{\frac{1}{\pi}T_i+I_{dif,i}} \quad (2)$$

где $r_{surf,i}$ – коэффициент отражения земной поверхности; $E_0$ – освещенность земной поверхности без учета вклада отраженного излучения; $\gamma_1$ – сферическое альбедо атмосферы; $I_{sum,i}$ – интенсивность суммарного принимаемого излучения; $I_{sun,i}$ – интенсивность солнечной дымки; $T_i$ – прямой коэффициент пропускания от наблюдаемого пикселя до приемной системы; $I_{dif,i}$ – интенсивность поверхностной дымки при единичной светимости земной поверхности.

**Модель атмосферы.** В качестве информации об оптическом и метеорологическом состоянии атмосферы в момент проведения спутниковых измерений могут быть использованы модельные данные, данные ДЗЗ, прогностические данные, данные радиозондов, данные сети AERONET и др. Среди перечисленных типов только данные ДЗЗ позволяют снять вопрос об оперативности их получения и обеспечить относительно высокое пространственное разрешение, сравнимое с пространственным разрешением спутниковых радиометрических каналов.

В результате исследований, выполненных в ИОА СО РАН, была предложена и обоснована методика использования данных о состоянии атмосферы по данным измерений спектрорадиометра MODIS при проведении атмосферной коррекции спутниковых ИК-изображений [8, 9]. Проведенная комплексная валидация спутниковых данных MODIS подтвердила возможность их использования для проведения атмосферной коррекции [10-13].

В рамках разработанного программного обеспечения данные об оптических свойствах атмосферы записываются в оптическую модель атмосферы, для генерации которой разработана отдельная программа. Модель атмосферы формируется для данных каждого канала спутникового прибора и включает в себя следующие величины: солнечную постоянную, профили коэффициентов аэрозольного и молекулярного ослабления и рассеяния и индикатрисы рассеяния для 32 слоев атмосферы в диапазоне высот от 0 до 100 км. Кроме того, модель включает аппаратные характеристики приборов.

В качестве источников информации об атмосферных параметрах используются данные измерений спектрорадиометра MODIS. В качестве альтернативы могут быть привлечены результаты наземных измерений. Исходными характеристиками для построения оптической модели являются длина волны излучения центра $\lambda$, ширина $\Delta\lambda$ и аппаратная функция канала регистрирующей аппаратуры, аэрозольная оптическая толщина атмосферы при $\lambda = 0.47$, 0.55 и 0.66 мкм, общее содержание озона, приземное давление, общее содержание водяного пара, вертикальный профиль температуры.

Для формирования моделей оптических свойств атмосферы в момент наблюдения (или близкий к нему) в программе используются данные о солнечных постоянных, спектрах поглощения $O_2$, $O_3$, $H_2O$, аэрозольные модели LOWTRAN-7, общее содержание водяного пара, общее содержание озона и данные о коэффициентах молекулярного рассеяния воздуха при нормальных условиях.

**Архитектура программного комплекса.** Программный комплекс атмосферной коррекции позволяет выполнять следующие функции:
1) расшифровка данных спутниковых измерений и сопутствующих метаданных, подготовка данных к расчетам, запись результатов расчетов в стандартный формат;
2) формирование оптической модели атмосферы;

3) расчет потоков излучения, формирующих принимаемое спутниковым прибором излучение;
4) проведение атмосферной коррекции.

Функциональность комплекса реализована через создание набора модулей.

Модуль подготовки данных включает конвертор для чтения и записи исходных данных и результатов расчетов в формат GeoTIFF [14], в настоящее время преимущественно используемый для хранения данных измерений российских спутниковых приборов и метаданных. Этот формат обеспечивает возможность использования стандартных программных средств для просмотра и анализа результатов расчетов. Конвертор разработан на основе пакета GDAL [15] и имеет интерфейс для использования совместно с модулями, созданными на языках Python, C и Fortran. Планируется расширение функционала модуля за счет добавления программ для конвертации данных, записанных в других стандартных форматах.

Модуль расчета корректирующих атмосферных поправок представляет собой набор отдельных программ, реализованных на языке Fortran. Расчеты осуществляются методом Монте – Карло с помощью алгоритмов, учитывающих сферичность атмосферы [16, 17]. В данной версии программного обеспечения выполняется статистическое моделирование освещенности земной поверхности; осуществляются расчеты сферического альбедо атмосферы, интенсивности излучения, не взаимодействовавшего с земной поверхностью, и интенсивности отраженного поверхностью излучения. Атмосферная коррекция выполняется с учетом рассчитанных поправок.

Последовательность обработки данных и вызов конверторов и расчетных модулей задаются при работе управляющего модуля, реализованного на языке Python. Представленное программное обеспечение предназначено для работы в операционной системе Linux. Разработка и тестирование работы программного комплекса производилась в среде операционной системы openSUSE leap 42.3. Схема работы программного комплекса приведена на рис. 2.

В качестве инструментов разработки использованы свободно распространяемые программные средства.

**Заключение.** Представленное программное обеспечение предназначено для восстановления коэффициентов отражения (альбедо) земной поверхности по спутниковым измерениям световых потоков отечественными приборами Сангур и ГСА спутника Ресурс-П и КМСС спутника Метеор-М. В первой версии данного программного комплекса для каждого из приборов разработан отдельный управляющий модуль, что связано с недостаточной стандартизацией форматов спутниковых данных и метаданных.

Дальнейшее развитие представленного программного комплекса тематической обработки данных российских спутниковых приборов предполагается вести в следующих направлениях:

1. Включение дополнительных расчетных модулей, обеспечивающих расчеты корректирующих атмосферных поправок.

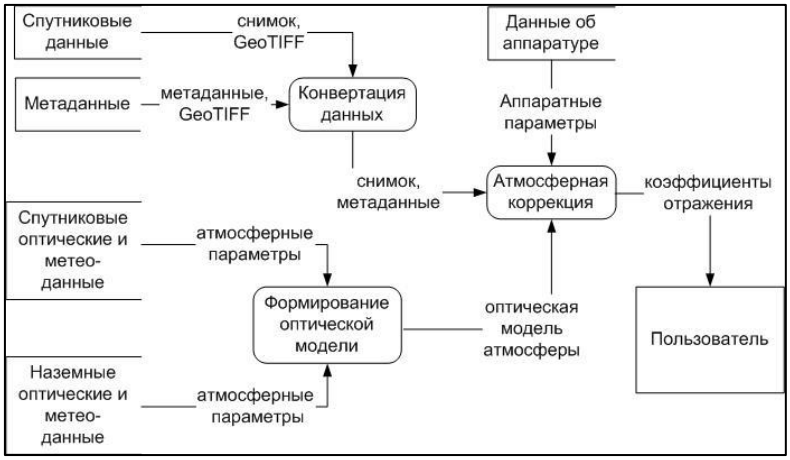



Рис. 2. Схема работы программного комплекса.

2. Увеличение быстродействия расчетов за счет использования технологий параллельного программирования.
3. Включение генератора оптических моделей атмосферы в состав комплекса. Использование в составе комплекса сетевых источников данных, необходимых для формирования оптических моделей атмосферы.
4. Развитие интерфейсной части комплекса. Создание средств визуализации и анализа спутниковых изображений.
5. Увеличение номенклатуры спутниковых приборов, данные которых можно обрабатывать с помощью предлагаемого программного обеспечения.

## ЛИТЕРАТУРА


[1] *Stamnes K., Tsay S.C., Wiscombe W., Jayaweera K.* Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media // Appl. Opt. 1988. V.27. N12. P. 2502-2509.

[2] *Stamnes K., Tsay S.C., Wiscombe W., Laszlo I.* A General-Purpose Numerically Stable Computer Code for Discrete-Ordinate-Method Radiative Transfer in Scattering and Emitting Layered Media, DISORT Report v1.1. 2000.

[3] *Berk A., Bernstein L.S., Anderson G.P. et al.* MODTRAN cloud and multiple scattering upgrades with application to AVIRIS // Remote Sensing of Environment. 1998. Vol. 65. P. 367-375.

[4] *Белов В.В., Тарасенков М.В.* О точности и быстродействии RTM-алгоритмов атмосферной коррекции спутниковых изображений в видимом и УФ-диапазонах // Оптика атмосферы и океана. 2013. Т. 26, № 07. С. 564-571.

[5] *Тарасенков М.В., Белов В.В.* Комплекс программ восстановления отражательных свойств земной поверхности в видимом и УФ-диапазонах // Оптика атмосферы и океана. 2014. Т. 27, № 07. С. 622-627.

[6] Kneizys F.X., Shettle E.P., Anderson G.P. et al. User Guide to LOWTRAN-7. ARGL-TR-86-0177. ERP 1010. Gallery. Hansom AFB. MA 01731. 1988. 137 p.

[7] *Зуев В.Е., Белов В.В., Веретенников В.В.* Теория систем в оптике дисперсных сред. Томск: Изд-во «Спектр» Института оптики атмосферы СО РАН, 1997. 402 с.

[8] *Афонин С.В., Соломатов Д.В.* Методика учета оптико-метеорологического состояния атмосферы для решения задач атмосферной коррекции спутниковых ИК-измерений // Оптика атмосферы и океана. 2008. Т. 21, № 2. С. 147-153.

[9] *Афонин С.В.* К вопросу о применимости восстановленных из космоса метеоданных MODIS для атмосферной коррекции спутниковых ИК измерений // Оптика атмосферы и океана. 2010. Т. 23, № 8. С. 684-688.

[10] *Афонин С.В., Белов В.В., Энгель М.В., Кох А.М.* Валидация региональных спутниковых данных MODIS Aerosol Product (MOD04) // Вторая Всерос. конф. «Дистанционное зондирование земных покровов и атмосферы космическими средствами»: Тез. докл. С.-Пб.: Изд-во РГГМУ, 2004. Т.2. С. 27-31.

[11] *Afonin S.V., Belov V.V., Engel' M.V.* Statistical analysis of the MODIS Atmosphere Products for the Tomsk Region // Proc. SPIE. 2005. V. 5979. P. 164-172.

[12] *Афонин С.В., Белов В.В., Энгель М.В.* Сравнительный анализ спутниковых аэрозольных данных типа MODIS Aerosol Products // Оптика атмосф. и океана. 2008. Т.21, № 3. С. 235-239.

[13] *Афонин С.В., Белов В.В., Панченко М.В., Сакерин С.М., Энгель М.В.* Корреляционный анализ пространственных полей аэрозольной оптической толщи на основе спутниковых данных MODIS // Оптика атмосф. и океана. 2008. Т. 21, № 6. С. 510-515.

[14] GeoTIFF. Описание формата. **https://trac.osgeo.org/geotiff** (дата обращения 10.06.2019).

[15] GDAL – Geospatial Data Abstraction Library. **https://www.gdal.org** (дата обращения 10.06.2019).

[16] *Белов В.В., Тарасенков М.В., Пискунов К.П.* Параметрическая модель солнечной дымки в видимой и УФ-области спектра // Оптика атмосферы и океана. 2010. Т. 23, № 04. С. 294-297.

[17] *Белов В.В., Тарасенков М.В.* Статистическое моделирование интенсивности световых потоков, отраженных сферической земной поверхностью // Оптика атмосферы и океана. 2010. Т. 23. № 1. С. 14-20.

# ИНФОРМАЦИОННАЯ СИСТЕМА СБОРА, ХРАНЕНИЯ И ОБРАБОТКИ СПУТНИКОВЫХ ДАННЫХ ИВТ СО РАН КАК ИНСТРУМЕНТ ДЛЯ РЕШЕНИЯ ЗАДАЧ МОНИТОРИНГА ПРИРОДНЫХ И АНТРОПОГЕННЫХ ОБЪЕКТОВ

*Шокин Ю.И.[(1)], Пестунов И.А.[(1)], Чубаров Д.Л.[(1)], Кихтенко В.А.[(1)], Мамаш Е.А.[(1)], Добрецов Н.Н.[(1)(2)], Воронина П.В.[(1)], Синявский Ю.Н.[(1)]*
[(1)] Институт вычислительных технологий СО РАН, Новосибирск
[(2)] Институт геологии и минералогии СО РАН, Новосибирск

В докладе рассматривается структура и основные принципы работы развернутой в ИВТ СО РАН системы сбора, хранения и обработки спутниковых данных и обсуждается спектр возможностей системы для решения различных тематических задач.
*Ключевые слова: информационная система, спутниковый мониторинг, обработка пространственных данных.*

**Общая структура системы.** В последние несколько лет все большую популярность приобретают задачи, связанные с анализом накопленного за последние десятилетия архива спутниковых данных, позволяющего исследовать динамику и закономерности большого количества природных и антропогенных процессов. В связи с этим, становится актуальным переход от традиционной схемы работы со спутниковой информацией, основанной на обработке отдельных сцен, к концептуально новой модели работы с архивом спутниковых данных как с единым целым. Результатом такого подхода является эффективная реализация процедур статистического анализа, решение большого класса новых тематических задач, включая исследование динамики процессов в отдельном пикселе или на заданном полигоне, а также возможность совместного использования данных из разных источников.

Начиная с 2007 года, в ИВТ СО РАН функционирует информационно-вычислительная инфраструктура приема, глубокой обработки и проведения статистического анализа больших объемов спутниковых данных и другой пространственной информации [1-7]. В базе данных ИВТ СО РАН в настоящее время представлены данные как с зарубежных, так и с отечественных космических аппаратов: SPOT, (Terra+Aqua)/MODIS, Landsat, Метеор, Ресурс-П, Канопус-В.

Инфраструктура основана на ранее разработанных в ИВТ СО РАН технологиях и состоит из двух основных блоков: комплекса обработки спутниковых снимков и блока информационной аналитики (рис. 1).

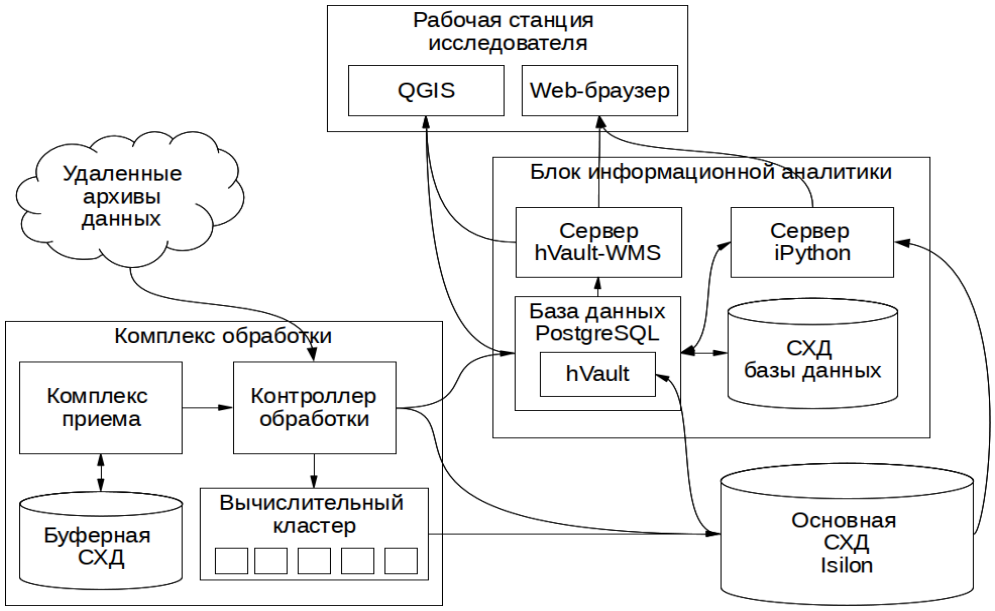



Рис. 1. Схема взаимодействия компонент информационно-вычислительной инфраструктуры.

Комплекс приема и обработки обеспечивает своевременное получение спутниковых данных и их глубокую обработку. Исходные данные принимаются как непосредственно с космических аппаратов в режиме прямой передачи данных, так и загружаются из удалённых архивов, что, с одной стороны обеспечивает оперативность предоставления обработанных данных исследователю, а с другой расширяет спектр доступной информации для ретроспективного анализа.

Система управления потоковыми процессами автоматически обеспечивает параллельное выполнение обработки данных в распределенном режиме на узлах кластера. Тем самым обеспечивается получение стандартных продуктов глубокой обработки в режиме близком к реальному времени.

Весь архив информационных продуктов, произведенных на основе данных дистанционного зондирования, находится в постоянной доступности. Хранение данных и их каталогизация внутри одного дерева директорий обеспечивается горизонтально масштабируемым сетевым хранилищем Isilon на основе технологии OneFS [8]. Реализованная конфигурация хранилища предназначена для решения задачи долговременного хранения нескольких сотен терабайт данных с обеспечением их постоянной доступности.

Блок информационной аналитики предоставляет исследователям средства виртуальной интеграции разнородных данных, а также инструментарий для проведения ретроспективного анализа, статистической обработки и визуализации. Его ключевым элементом является СУБД PostgreSQL с разработанным в ИВТ СО РАН дополнительным модулем hVault для прямого доступа к файловому архиву геопространственной информации без необходимости предварительного копирования и преобразования форматов данных [9].

Существует два концептуально разных подхода в работе с пространственными данными. Первый, традиционный для спутниковых данных, состоит в поддержании сети распределенных архивов, единицей хранения которых являются файлы сцен. В другом подходе, традиционном для реляционных баз данных, элементом хранения является отдельная запись. Недостатком подхода, основанного на выгрузке отдельных сцен, является избыточное копирование, что снижает его эффективность при работе с большими архивами, поэтому актуальной становится задача обеспечения эффективной работы с архивом как с базой данных.

Задачи информационной аналитики, требующие обращения к значительному объему записей, хранящихся в реляционной базе данных могут быть решены либо внешними средствами с выгрузкой данных в специальную структуру, оптимизированную для работы с отношениями, определенными на временных и пространственных шкалах (OLAP), либо за счет интеграции функций для работы с временными и пространственными данными в механизм выполнения реляционных запросов (in-database processing).

Технологии, основанные на многомерном представлении данных – OLAP-технологии – обеспечивают высокую скорость выполнения запросов, что позволяет работать в интерактивном режиме, но такие решения требуют выгрузки данных в новую структуру, выполнения предварительного агрегирования данных, что неосуществимо при работе с архивами, объем которых может составлять сотни терабайт. С другой стороны, технологии, основанные на выполнении вычислений в процессе исполнения запроса, особенно эффективны при параллельной обработке.

В ИВТ СО РАН предложен и реализован подход к работе с архивом пространственных данных, объем которого превышает 100 ТВ, основанный на отображении данных в единую реляционную модель [9, 10]. Модуль hVault реализует прозрачное отображение файлового архива спутниковых снимков в виртуальные таблицы базы данных. Он позволяет выполнять произвольные SQL-запросы к данным файлового архива, при этом планировщик запросов оптимизирует их выполнение на основе доступных метаданных, а применяемые для выполнения вычислений алгоритмы предусматривают работу с объёмами информации превосходящими ёмкость оперативной памяти сервера СУБД (рис. 2).

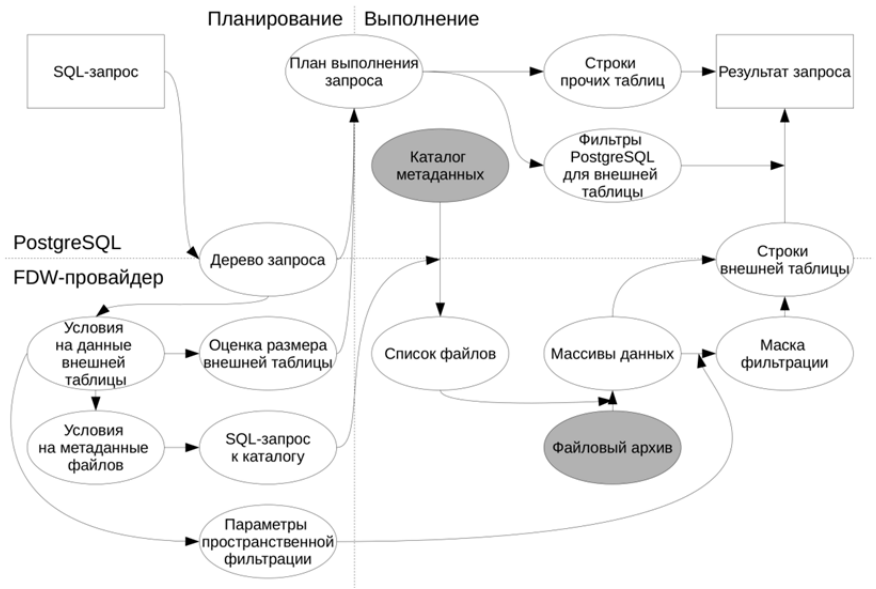



Рис. 2. Схема выполнения запроса в системе hVault.

На основе hVault построен сервер протокола WMS (Web Map Service), предоставляющий доступ ко всему архиву. Особенностью этого сервера является поддержка произвольных SQL запросов к системе, что позволяет реализовать произвольные вычисления на стороне сервера перед отображением карты.

Таким образом, исследователю предоставляется широкий ассортимент инструментов для работы с данными. Все современные ГИС-системы поддерживают работу с пространственными СУБД, в их числе такие свободно распространяемые пакеты программ, как QGIS и GRASS GIS. Также возможно использование такого популярного среди научных работников интерактивного инструмента для совместной работы с данными, как iPython Notebook Server, основанного на языке программирования Python.

Предложенная технология предоставляет следующие преимущества по сравнению с традиционной технологией работы с данными дистанционного зондирования:

- работа с данными эквивалентна работе с таблицами значений. Она не требует использования специализированного программного обеспечения, в том числе выгрузки и анализа отдельных снимков — одним запросом обрабатывается группа снимков, выбранных по заданным признакам;
- пользователю предоставляется возможность исследования многолетней динамики процессов как в отдельно взятой точке, так и на выделенном полигоне;
- математическая и статистическая обработка данных производится ядром СУБД в момент выполнения SQL запросов;
- возможность быстрой оценки основных статистических характеристик исследуемой выборки спутниковых данных и её качества.

Для повышения производительности системы была предложена структура данных для эффективной агрегации данных в приложениях в области интерактивной визуализации – интервальное дерево отрезков. Средствами Apache Spark развернута масштабируемая вычислительная инфраструктура, основанная на модели распределенных массивов. Выполнена реализация интервального дерева отрезков в этой модели. Модель распределенных массивов и, в частности, средства Apache Spark в настоящее время широко применяются для построения распределенных систем обработки данных ведущими технологическими компаниями. В частности, средства для работы с данными дистанционного зондирования в этой модели разрабатываются в качестве перспективного инструмента для доступа к данным, хранящимся в архивах NASA [11].

Разработанную систему извлечения данных из архива спутниковых снимков можно сопоставить с такими системами как NASA Giovanni [12], Google Earth Engine [13] и европейским проектом TELEIOS [14]. Все они нацелены на предоставление доступа к большим объемам спутниковых снимков.

Для поддержки исследований, выполняемых с помощью данных из распределенных архивов центра Годдарда (NASA), была разработана упомянутая выше система Giovanni. Лептух и Эккер реализовали первую версию Giovanni на основе системы GrADS [15], предназначенной для совместной обработки разнородных данных, определенных на сетках, заданных в трех пространственных, временном, а также одном дополнительном измерениях. Функциональные возможности системы hVault соответствуют уровню доступа к данным, однако, превосходят возможности GrADS в нескольких аспектах: промежуточные результаты, полученные при выполнении запроса, также могут быть сохранены в hVault и доступны для последующих запросов, кроме того, возможности пространственной СУБД позволяют оптимизировать выполнение запросов, требующих извлечения небольшого объема данных из большого числа исходных файлов.

Еще одна информационная система для совместной работы с данными в науках о Земле разрабатывается в рамках проекта компании Google (Google Earth Engine, GEE). Подход, реализованный в этом проекте, отличается от подходов, предлагаемых в рамках ESC и в настоящей работе, тем, что интеграция данных из архивов в информационную систему осуществляется посредством импорта и преобразования самих данных к единому формату. На уровне алгоритмов тематической обработки данных вся загруженная в систему информация приведена к единому типу, соответствующему растровому покрытию некоторой области пространства. Такие покрытия, допускающие единообразную обработку, организуются в коллекции. Так, каждому информационному продукту MODIS соответствует отдельная коллекция, содержащая все файлы, импортированные в систему. Отличие обусловлено тем, что перед разработчиками GEE не стоят ограничения на объем данных, а также отсутствует необходимость хранения данных в оригинальных форматах. Система GEE также поддерживает выполнение вычислений в удаленном режиме и сохранение промежуточных результатов в виде отдельных изображений или массивов векторных данных, а также их выгрузку из системы на компьютер пользователя. Как и в системе Giovanni, в GEE работа с данными без предварительного их пересчета на регулярную сетку не поддерживается. В отличие от GEE, система hVault предназначена для работы с архивами данных в исходных форматах.

Проект TELEIOS [14] нацелен на создание инфраструктуры виртуальной обсерватории для данных зондирования Земли. В части организации доступа к данным он опирается на технологию колоночной базы данных MonetDB Data Vault [16]. Эта технология позволяет обращаться к файловым архивам спутниковых снимков средствами базы данных, посредством автоматической загрузки и преобразования данных в таблицы базы данных. Ограничением этой технологии является объем доступной оперативной памяти сервера.

В отличие от описанных систем, в системе hVault нет ограничений на вид запросов к данным, так как поддерживаются произвольные SQL-запросы, в том числе, с использованием хранимых процедур, что позволяет реализовать выполнение произвольного алгоритма. Кроме того, она рассчитана на работу с пространственными данными, представленными в виде набора файлов, любого объема без необходимости их предварительного преобразования и подготовки, что упрощает ее развертывание для существующих архивов данных дистанционного зондирования и других архивов пространственных данных.

Ниже приведены примеры использования предложенной технологии при решении некоторых тематических задач.

**Анализ пространственно-временного распределения точек термальных аномалий.** Разработанные программные инструменты позволили решить ряд прикладных задач, одной из которых является исследование пространственно-временного распределения термальных аномалий (термоточек) на территории Евразии.

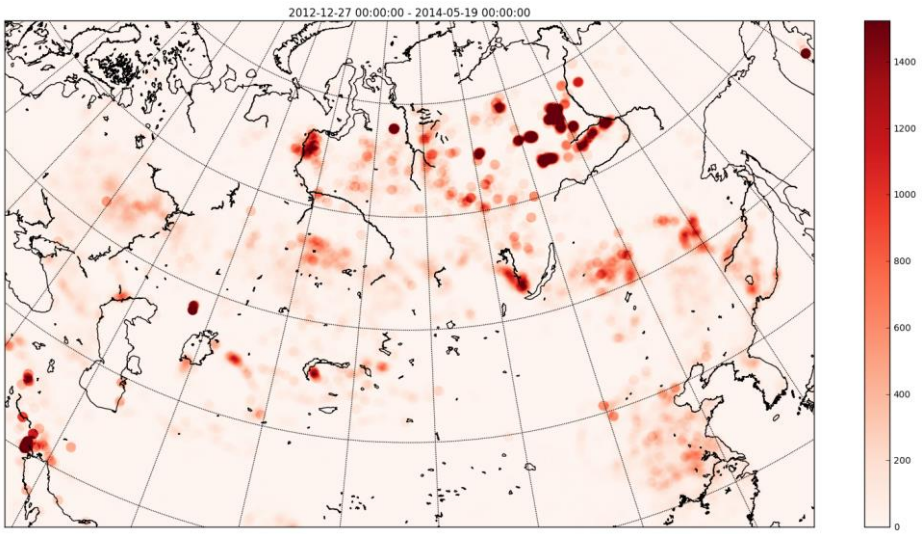



Рис. 3. Демонстрация работы приложения для интерактивной агрегации большого объема данных: плотность термоточек, зарегистрированных на территории Северной Евразии, интервал агрегирования 27.12.2012-19.05.2014.

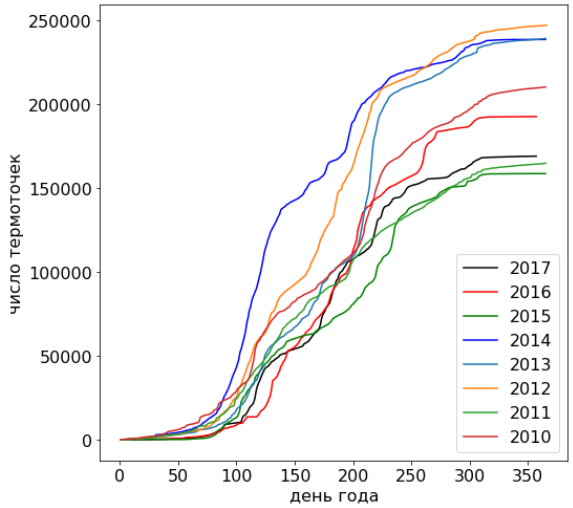



Рис. 4. Число термоточек, зарегистрированных на территории Сибири и Дальнего Востока в 2010-2017 гг.

На основе Apache Spark создан прототип системы агрегации разнородных пространственных данных, ориентированный на масштабируемые вычислительные архитектуры. Прототип реализован в рамках сервисной инфраструктуры рабочих пространств для интерактивного удаленного счета (ipython.ict.nsc.ru). Для демонстрации возможностей прототипа на его основе реализовано приложение, обеспечивающее интерактивное отображение картосхемы плотности термоточек (hotspot) по данным MODIS (MOD14) за произвольный период времени (рис. 3).

Пространственное распределение термоточек существенно различается год от года, что отражает различия пожарных сезонов. Ещё одним способом сравнения пожарных сезонов, является сравнение роста числа термоточек в течение сезона, подсчитанных с нарастающим итогом (рис. 4).

На графиках, подобных приведенному на рис. 4, построенных для различных территорий, можно выделить группы сезонов в зависимости от интенсивности появления термоточек. Ещё одна функция таких графиков возможность прогнозирования по году аналогу.

Одним из ограничений статистических наблюдений за динамикой природных пожаров, построенных на подсчете термоточек, является наличие среди них тех, что не связаны с

природными пожарами. Значительную долю составляют те термоточки, которые связаны с промышленными источниками тепла. Один из способов выявления термоточек, связанных с промышленными объектами, основан на том факте, что они, в отличие от природных пожаров, являются постоянными и неподвижными источниками теплового излучения. Таким образом, вокруг таких источников могут накапливаться кластеры. С помощью методов автоматической кластеризации на территории Сибири и Дальнего Востока было выявлено несколько сотен кластеров термоточек. С помощью данных более высокого разрешения была выполнена валидация полученных кластеров, которая подтвердила, что они связаны с различными промышленными источниками тепла. В основном, – с установками сжигания попутного газа на нефтяных месторождениях.

Дальнейшее исследование было направлено на каталогизацию найденных источников. Цель продолжения исследований состоит в изучении возможности использования постоянных источников тепла для точной географической привязки тепловых спутниковых снимков. О возможностях этого метода можно судить по тому, что положение источника на снимке может быть определено с точностью значительно превышающей разрешение кадра (рис. 5). Центр теплового пятна, зарегистрированного на снимке в канале B11 (КА Landsat, сенсор TIRS, 11 мкм), с точностью до нескольких метров совпадает с положением установки сжигания попутного газа. Представляет интерес возможность повышения точности привязки не только данных высокого и среднего, но и низкого разрешения, таких как данные MODIS. Для этого потребуется точное вычисление положения источника по небольшому числу пикселей исходного изображения. Одним из способов решения этой проблемы является интерполяция значений на сетку более крупного масштаба.

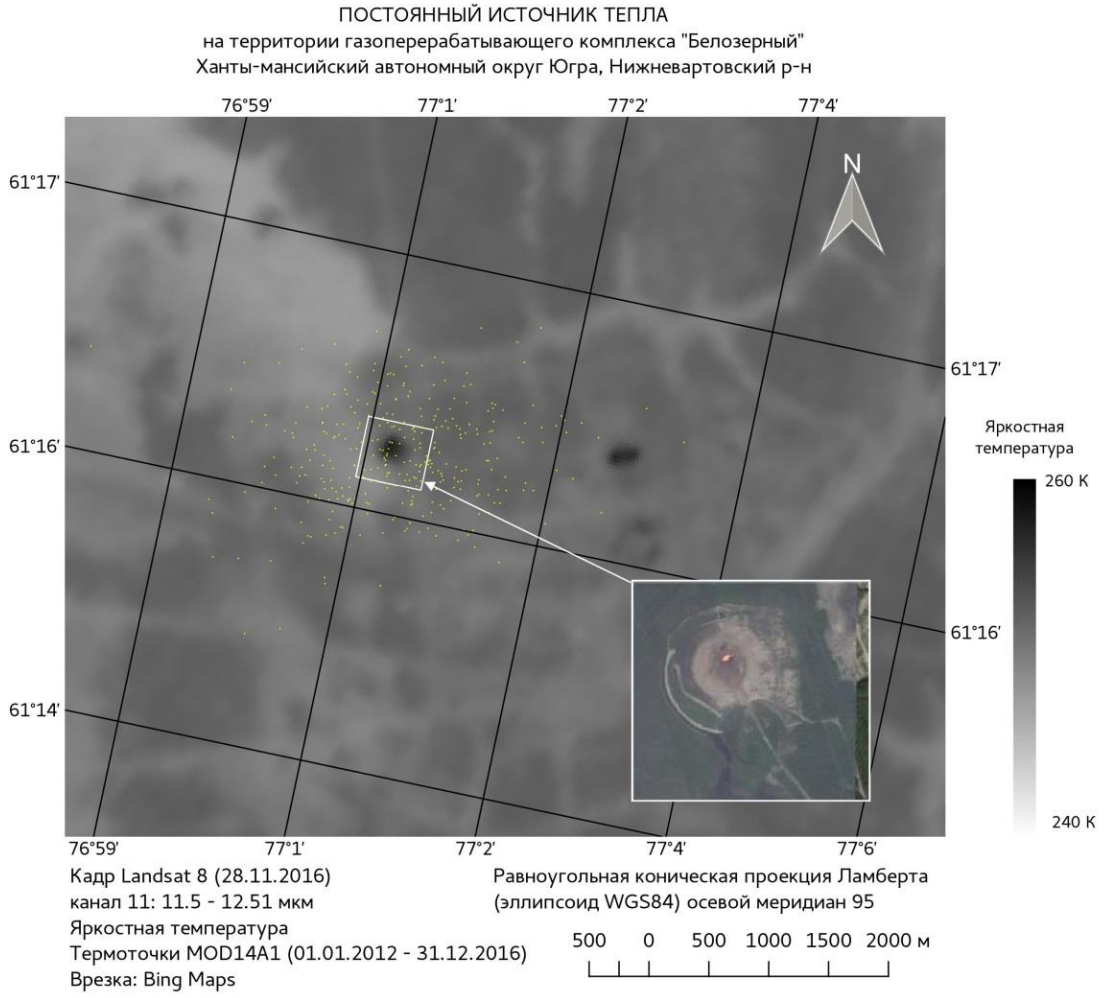



Рис. 5. Образ найденной с помощью кластеризации установки сжигания попутного газа на снимке, выполненном с разрешением 100 м.

**Оценка качества спутниковых данных на примере вычисления индекса NDVI.** Одной из важных и часто решаемых задач с помощью ДДЗ является вычисление всевозможных индексов, и отдельный интерес заключается в оценке соответствия значений индексов, полученных на основе спутниковых данных реальному состоянию исследуемых природных объектов. Это, прежде всего, зависит от качества используемой информации. Созданная система спутниковых данных позволяет решать такого рода задачи. Например, была выполнена оценка применимости спутниковых данных для исследования состояния посевных сельхозкультур. Ход развития посевных культур отражается на интенсивности фотосинтетической активности, одним из показателей которой является индекс NDVI. Данные MODIS позволяют при благоприятных условиях съемки восстанавливать значения индекса для отдельных полей. Для того, чтобы определить количество дней, когда восстановленный индекс может быть пригоден для оценки состояния растений, были извлечены данные за период с 1-го марта по 1-е сентября 2017 для одного из полей сельскохозяйственного предприятия, расположенного на территории Новосибирской области. Полученные результаты представлены на рис. 6.

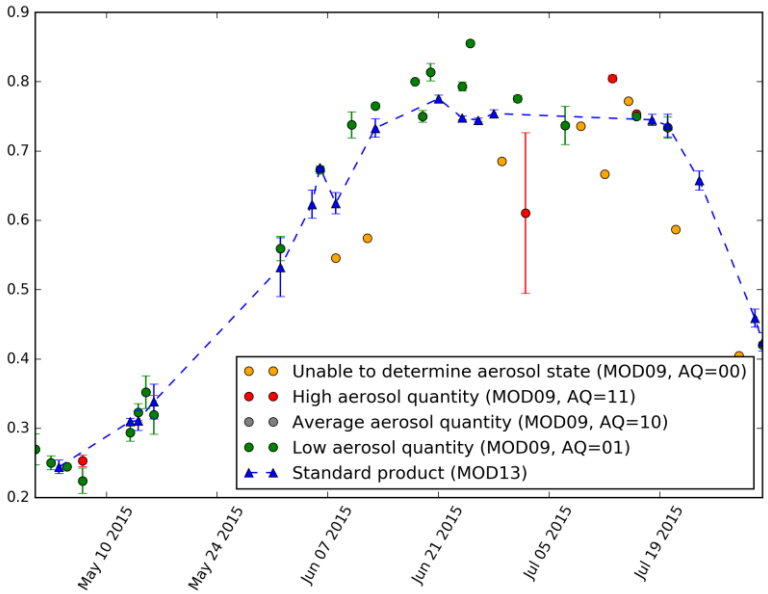



Рис. 6. Динамика NDVI для поля пшеницы по данным из различных источников.

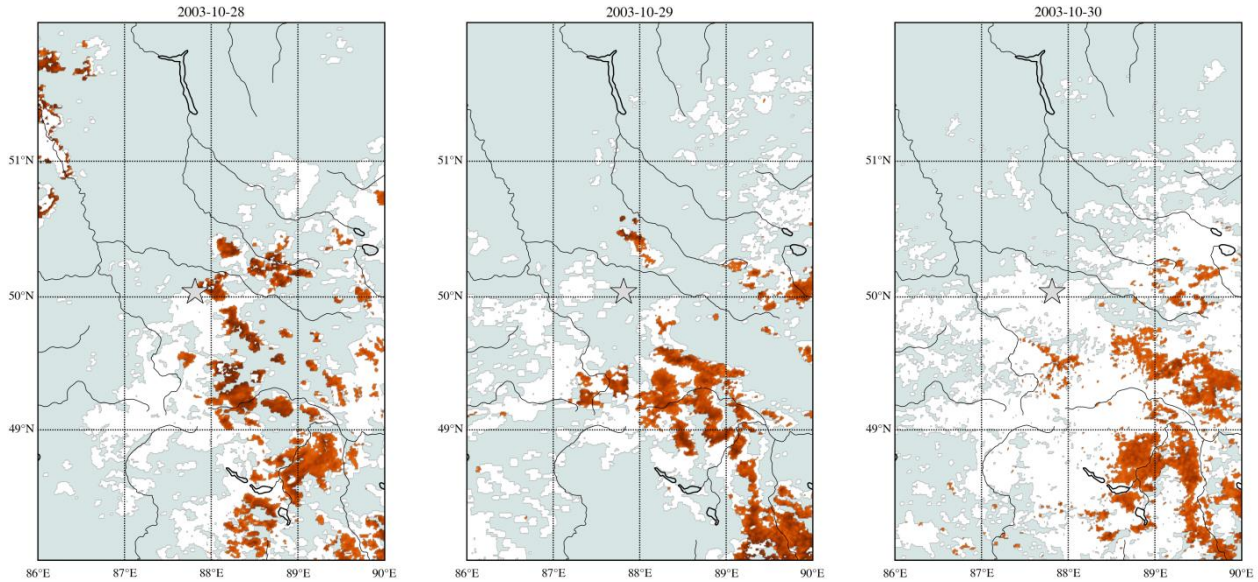



Рис. 7. Области с аномальными значениями температуры в дни после Чуйского землетрясения (27.09.2003). Коричневым цветом выделены области с наибольшими значениями температуры.

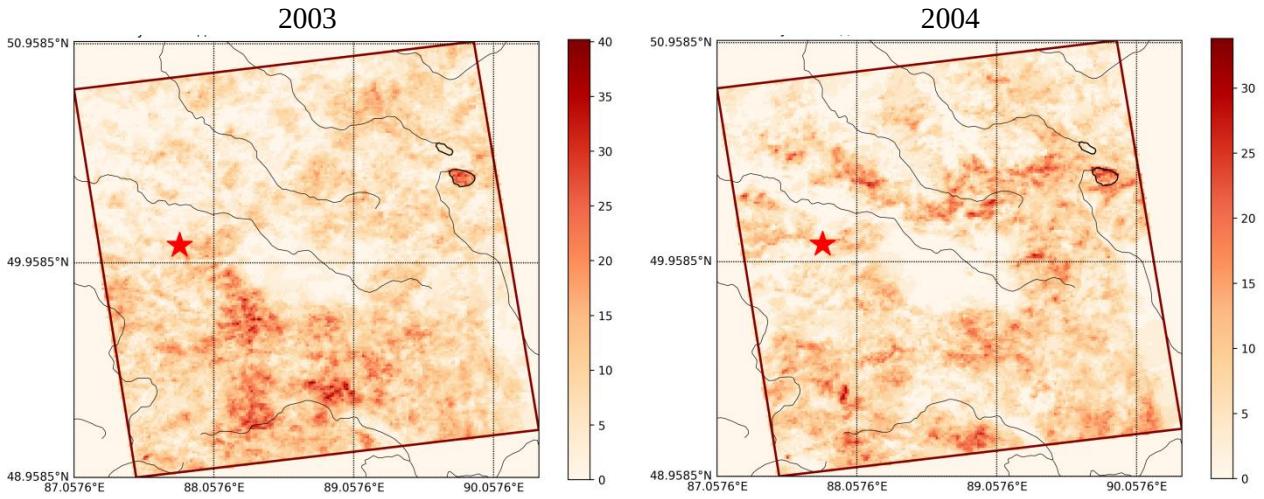


Рис. 8. Накопительная сумма аномальных значений индекса.

Поиск статистических отклонений температуры земной поверхности методом RST (Robust Satellite Technique) для территории юга Сибири. Созданная система также является эффективным инструментом статистического анализа накопленной за годы спутниковой информации. Эта эффективность позволила реализовать метод RST (Robust Satellite Techniques), в основе которого лежит статистический анализ наборов спутниковых данных температуры земной поверхности для выделенной территории [17]. По выбранной области вычисляется индекс, основным достоинством которого является то, что при выборе области и временного интервала подходящих размеров он позволяет исключить влияние вариаций температуры, вызванных климатическими процессами, неоднородностью рельефа и погодными условиями [7, 18]. На рис. 7 представлены результаты применения метода RST для территории юга Сибири.

Также была предложена модификация метода RST, которая заключается в суммировании только аномальных значений индекса в пикселе. Предложенная модификация метода RST позволяет частично отфильтровывать их случайные проявления. С использованием модифицированной методики проведен анализ полей температуры поверхности для сейсмически активной территории Горного Алтая для интервала 2001-2005 гг. Результаты для 2003 и 2004 годов приведены на рис. 8. Выявлены участки с устойчивыми по пространству и времени тепловыми аномалиями, а также участки с отсутствием аномалий. Модифицированная методика теоретически позволяет разделять аномальные проявления на две категории. Первая – это «пролонгированные аномалии», при которых аномальные точки непрерывно проявляются несколько раз подряд, формируя тем самым некое событие, продолжающееся во времени. Оно аномально по отношению к фону, но достаточно длительное по отношению к частоте измерений. И вторая – это «пульсирующие аномалии», в которых аномальные точки проявляются несколько раз в течение исследуемого интервала времени, но по отношению друг к другу возникают в разное время.

**Применение алгоритмов многомерной интерполяции для решения задачи восстановления пропусков в спутниковых данных.** Наличие пропусков в ДДЗ затрудняет использование данных для моделирования процессов, протекающих на поверхности Земли, традиционными средствами ГИС. Для решения задачи восстановления пропущенных данных средствами разработанной системы работы со спутниковыми данными были реализованы некоторые известные методы построения полных временных рядов по частично определенным данным. Была решена задача повышения точности метода RST за счет предварительного восстановления временных рядов, полученных по данным дистанционного зондирования. Использовались такие методы аппроксимации временных рядов как HANTS [19] и полиномиальные аппроксимации (рис. 9).

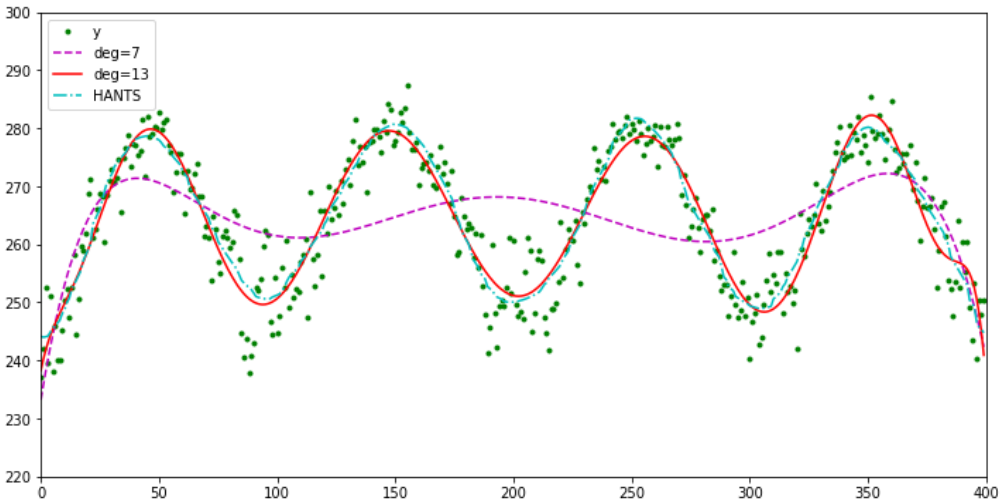



Рис. 9. Сравнение аппроксимации методом HANTS и полиномами различных степеней.

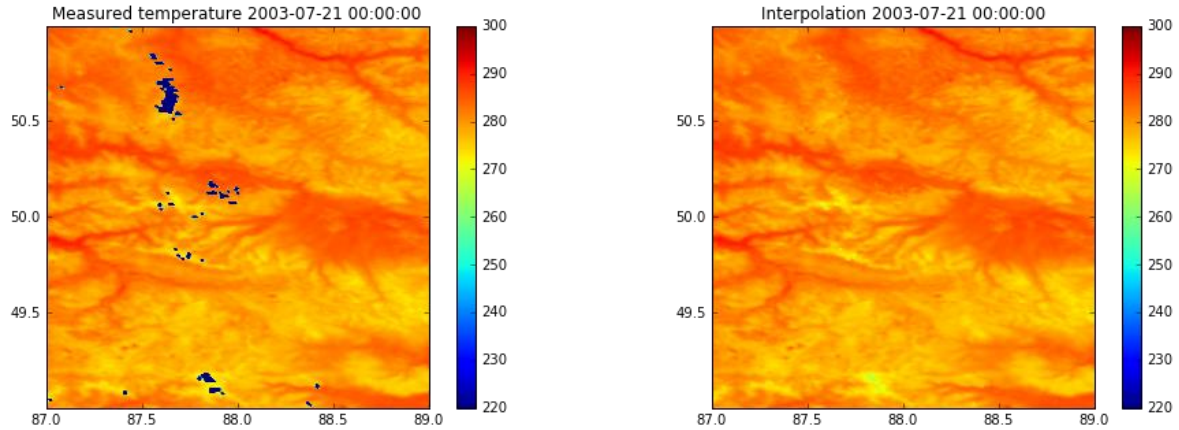



Рис. 10. Результаты восстановления значений температуры по методу HANTS.

Выполнено сравнение аппроксимации, полученной методом HANTS, с аппроксимацией полиномами различной степени.

Получены результаты восстановления рядов температуры поверхности Земли, полученных по данным дистанционного зондирования, для территории Горного Алтая, представляющей интерес с точки зрения мониторинга активных геологических разломов (рис. 10).

**Мониторинг температурных полей почв для картографического моделирования почвенного покрова.** Температурный режим является одним из наиболее значимых экологических факторов, который совместно с гидрологическим режимом характеризует общий энергетический уровень формирования и функционирования почвенного покрова. Температурное поле почв представляет совокупность значений температуры в точках пространственной области, которое согласно результатам исследований, полученных как в нашей стране, так и за рубежом, рассматривается в качестве ведущего фактора структурно-функциональной организации почвенного покрова.

С помощью системы hVault по данным сенсора MODIS (продукт MOD11), установленным на спутниках Terra/Aqua, было получено распределение средних дневных и ночных температур за 16-ти летний период (2001-2016 гг.) Среднесибирского трансекта и плоскогорья Укок (рис. 11).

На рис. 12 в качестве примера представлены результаты интерпретации выделения типологических единиц структурной и функциональной организации почвенного покрова плоскогорья Укок на основе статистического анализа совмещения границ контуров почвенного покрова и температурных полей, полученных по спутниковым данным.

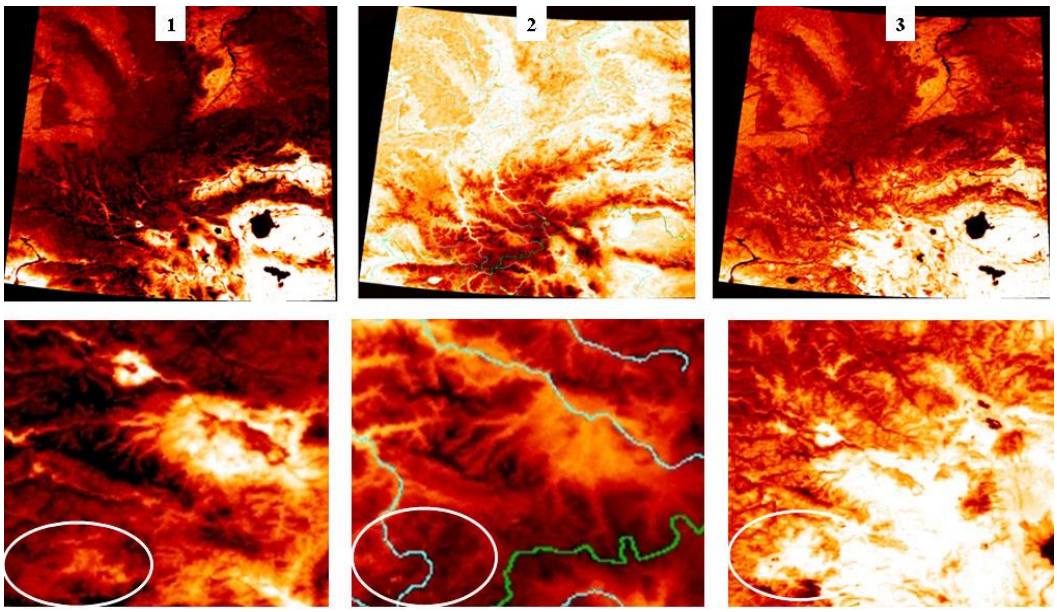



Рис. 11. Распределение температурных полей среднесибирского трансекта и плоскогорья Укок (внизу), рассчитанные за 16-ти летний период (2001-2016 гг.) по данным сенсора MODIS: 1 – дневные температуры; 2 – ночные температуры; 3 – разность дневных и ночных температур.

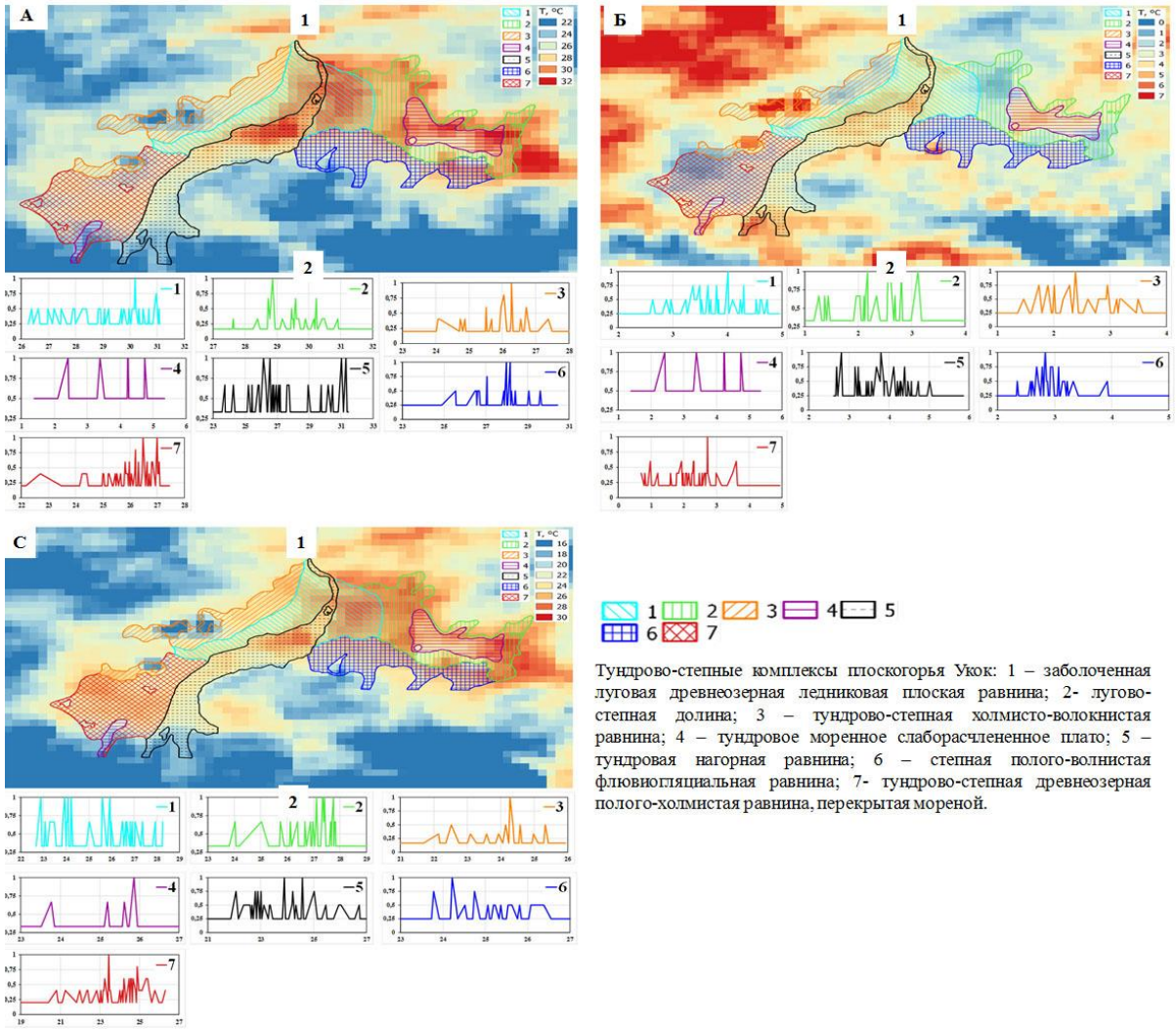



Рис. 2. Картографические модели температурных полей тундрово-степных комплексов плоскогорья Укок, созданные на основе анализа спутниковых данных: А – дневных, Б – ночных и С – разницы температур. 1 – контуры температурных полей; 2 – статистический анализ совмещения границ почвенных контуров и температурных полей

Показано [20], что картографические модели температурных полей, полученные в результате совместного анализа наземных и спутниковых данных, обладают достаточной информативностью в отношении энергетики почвообразовательных процессов и могут быть использованы для оценки термических условий почв малоисследованных и труднодоступных территорий Алтае-Саянского региона.

Наиболее интересные результаты и возможности разработанной системы представлены на сайте http://sdc.esemc.nsc.ru, где также приводится информация о структуре системы, о данных и продуктах комплекса, возможности визуализации и доступа к ним, приводятся примеры космоснимков различных наблюдаемых природных и техногенных явлений, результаты обработки пространственных данных, извлеченных из архива.

## ЛИТЕРАТУРА


[1] *Шокин Ю.И., Жижимов О.Л., Пестунов И.А., Синявский Ю.Н., Смирнов В.В.* Распределенная информационно-аналитическая система для поиска, обработки и анализа пространственных данных // Вычислительные технологии. 2007. Т. 12, № 3. С. 108-115.

[2] *Пестунов И.А., Смирнов В.В., Жижимов О.Л., Синявский Ю.Н., Скачкова А.П., Дубров И.С.* Каталог пространственных данных для решения задач регионального мониторинга // Вестник Казахского национального университета. Серия: Математика, механика, информатика. 2008. Т. 13, № 43. С. 71-76.

[3] *Шокин Ю.И., Пестунов И.А., Смирнов В.В., Синявский Ю.Н., Скачкова А.П., Дубров И.С., Левин В.А., Алексанин А.И., Алексанина М.Г., Бабяк П.В., Громов А.В., Недолужко И.В.* Распределенная информационная система сбора, хранения и обработки спутниковых данных для мониторинга территорий сибири и дальнего востока // Журнал Сибирского федерального университета. Серия: Техника и технологии. 2008. Т. 1, № 4. С. 291-314.

[4] *Шокин Ю.И., Добрецов Н.Н., Смирнов В.В., Лагутин А.А., Антонов В.Н., Калашников А.В.* Система информационной поддержки задач оперативного мониторинга на основе данных дистанционного зондирования // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8, № 3. С. 109-112.

[5] *Шокин Ю.И., Антонов В.Н., Добрецов Н.Н., Кихтенко В.А., Лагутин А.А., Смирнов В.В., Чубаров Д.Л., Чубаров Л.Б.* Распределенная система приема и обработки спутниковых данных Сибири и Дальнего Востока. Текущее состояние и перспективы развития // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 9, № 5. С. 45-54.

[6] *Шокин Ю.И., Чубаров Л.Б., Добрецов Н.Н., Кихтенко В.А., Смирнов В.В., Чубаров Д.Л.* О распределенной инфраструктуре системы оперативного спутникового мониторинга ЦКП ДДЗ СО РАН // Вычислительные технологии. 2013. Т. 18, Специальный выпуск, С. 86-94

[7] *Шокин Ю.И., Добрецов Н.Н., Мамаш Е.А., Кихтенко В.А., Воронина П.В., Смирнов В.В., Чубаров Д.Л.* Информационная система приема, обработки и доступа к спутниковым данным и ее применение для решения задач мониторинга окружающей среды // Вычислительные технологии. 2015. Т. 20, № 5. С. 157-174.

[8] *Patel N.N., Angiuli E., Gamba P., Gaughan A., Lisini G., Stevens F.R., Tatem A.J., Trianni G.* Multitemporal settlement and population mapping from Landsat using Google Earth Engine // International Journal of Applied Earth Observation and Geoinformation. 2015. V.35. P. 199-208.

[9] *Кихтенко В.А., Чубаров Д.Л., Смирнов В.В.* hVault: технология доступа к архиву спутниковых данных // XII Всероссийская открытая конференция «Современные проблемы дистанционного зондирования Земли из космоса»: Сборник тезисов докладов конференции (Москва, Россия, 10-14 ноября 2014). Москва: ИКИ РАН, 2014. С. 117.

[10] *Чубаров, Д.Л., Добрецов, Н.Н., Кихтенко, В.А.* Отображение модели данных netCDF в реляционную модель для работы с коллекциями данных дистанционного зондирования // Труды всероссийской конференции «Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов», SDM-2017. 2017. С. 324-328

[11] *Kuo K.-S., Lynnes C.S., Ramachandran R. A.* Proposed Earth Science Collaboratory for Remote Sensing Data Analysis // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2012. Vol. 5, N 6. P. 1612-1616.

[12] *Acker J. G., Leptoukh G.* Online analysis enhances use of NASA Earth science data // EOS Trans. AGU. 2007. Vol. 88, N 2. P. 14-17.
[13] *Hansen M.C. Potapov P.V., Moore R., Hancher M., Turubanova S.A., Tyukavina A., Thau D., Stehman S.V, Goetz S.J., Loveland T.R., Kommareddy A., Egorov A., Chini L., Justice C.O., Townshend J.R.G.* High-Resolution Global Maps of 21st-Century Forest Cover // Science. 2013. Vol. 342, N 6160. P. 850-853.
[14] *Koubarakis M., Datcu M., Kontoes C., Di Giammatteo U., Manegold S., Klien E.* TELEIOS: a database-powered virtual earth observatory // Proceedings of the VLDB Endowment. 2012. Vol. 5, N 12. P. 2010-2013.
[15] *Doty B.E., Kinter J.L.* Geophysical data analysis and visualization using the Grid Analysis and Display System // Visualization Techniques in Space and Atmospheric Sciences, NASA SP-519. 1995. P. 209-217.
[16] *Ivanova M., Kersten M.L., Manegold S.* Data Vaults: A Symbiosis Between Database Technology And Scientific File Repositories // Proceedings of International Conference on Scientific and Statistical Database Management. 2012. P. 485-494.
[17] *Tramutoli V., Di Bello G., Pergola N., Piscitelli S.* Robust satellite techniques for remote sensing of seismically active areas // Ann. di Geofis. Vol. 44, N 2. P. 295-312.
[18] *Воронина П.В., Чубаров Д.Л., Кихтенко В.А.* Статистические отклонения в пространственно-временных рядах спутниковых данных температуры поверхности Земли на территории юга Сибири // Интерэкспо Гео-Сибирь. 2017. Т. 4, № 2. С. 46-50.
[19] *Menenti M., Azzali S., Verhoef W., Van Swol R.* Mapping agroecological zones and time lag in vegetation growth by means of Fourier analysis of time series of NDVI images // Advances in Space Research. Vol. 13, N 5. P. 233-237.
[20] *Кудряшова С.Я., Чумбаев А.С., Пестунов И.А., Синявский Ю.Н., Чубаров Д.Л., Безбородова А.Н., Ермаков Н.Б.* Картографическое моделирование температурных полей почв Среднесибирского трансекта на основе совместного анализа наземных автоматизированных и спутниковых данных температурного мониторинга (в настоящем сборнике).

# О ГЕОГРАФИЧЕСКОЙ ПРИВЯЗКЕ КОНТЕНТА ТЕКСТОВЫХ ДОКУМЕНТОВ

*Жижимов О.Л., Леонова Ю.В.*
Институт вычислительных технологий СО РАН, Новосибирск

Извлечение географических названий из произвольных текстовых документов имеет важное значение в задачах обработки больших массивов документов и привязки их контента к определенному географическому региону. В самом простом виде модель извлечения географических названий из текста выглядит как последовательность действий с текстом, при этом на каждом этапе решается своя задача. Среди этих задач, несомненно, присутствуют: парсинг текста, анализатор элементов текста, обработка синонимов и сокращений, приведение элементов текста к нормальной форме с возможных словоформ и правил грамматики, сравнение элементов текста с элементами словарей географических названий, добавление в текст специальных меток для однозначной идентификации географических названий. В предлагаемой работе описана технология, реализующая перечисленные выше задачи на базе свободно распространяемой СУБД PostgreSQL. При этом используется стандартная конфигурация, все настройки серверной части выполнены в рамках штатных документированных процедур. В качестве авторитетной базы данных географических названий применены база данных GeoNames Gazetteer, базы данных Open Street Map (OSM), классификаторы ОКАТО и КЛАДР.
*Ключевые слова: географические названия, полнотекстовый поиск, модель извлечения названий, обработка текста, PostgreSQL, географический поиск.*

**Введение.** Цель настоящей работы – создание модели извлечения географических названий из произвольного текста с его индексирование по географическим атрибутам, например, по географическим координатам, с возможностью дальнейшей организации геометрического поиска.

Следует отметить, что существующие программные комплексы для организации доступа к текстовым информационным ресурсам не обладают необходимой функциональностью по хранению и обработке географических данных. Наделение же их требуемой функциональностью осложняется отсутствием единых стандартов на поиск и представление данных, связанных с географическим аспектом, которые сопрягались бы с существующими геоинформационными системами (ГИС), т. е. с системами, для которых географический аспект информации является основным [1]. Отсюда вытекает актуальность и перспективность создания технологии, обеспечивающей обработку географической информации в «негеографических» информационных системах общего назначения [2].

**Модель и алгоритмы.** Если очень коротко описать предлагаемую модель фиксирования географического контента в текстовом массиве данных для последующей индексации, то она будет выглядеть следующим образом.

1. Первое, что необходимо сделать при обработке произвольного текста – раскрыть все сокращения. В тексте заменяются слова-сокращения на их несокращенные значения. Эта процедура существенна для дальнейшего анализа, т.к. в текстах географические названия как правило сопровождаются сокращенными обозначениями типа географического объекта: г. – город, оз. – озеро, обл. – область и т.п. При этом необходима не только простая механическая подстановка значений в соответствии со словарем сокращений, но и анализ сопутствующего контента. В частности, сокращение «г.» может восприниматься те только как «год», но и как «город», в зависимости от окружающих слов. Формализованные правила, в соответствии с которыми происходит раскрытие сокращений, образуют специальный словарь шаблонов сокращений.
2. Полученный в результате вышеописанной процедуры текст разбивается на отдельные слова (токенизация) с фиксацией порядкового номера каждого слова в исходном тексте. При этом также происходит удаление стоп-слов, определенных в специальном словаре, и приведение остальных слов к нормальной форме в соответствии с морфологическим

словарем, который может сводить множество разных лингвистических форм слова к одной лексеме.

3. Следующий желательный, но не обязательный шаг, - раскрытие перечислений. Дело в том, что в различных текстах часто встречаются различные перечисления географических названий с групповым указанием типа объекта. Например, текст «... исследования были проведены в Новосибирской, Кемеровской и Омской областях» для однозначной фиксации географических объектов требует его преобразования к виду «... исследования были проведены в Новосибирской области, Кемеровской области и Омской области».
4. После выполнения вышеперечисленных процедур можно выполнить фиксацию географических объектов - приписать специальные метки соответствующим комбинациям слов или заменить соответствующую комбинацию слов на специальную метку. Первый вариант необходим в случае дальнейшей индексации текста как для геометрического, так и для полнотекстового поиска, а второй – для индексации географических объектов только для поиска геометрического.

Географическая фиксация текстов осуществляется путем поиска в тексте географических названий из географического словаря, в котором кроме написания этих названий указываются принадлежность к географической реальности: городу, поселку, горе, болоту, реке, морю и т.д. Текст документа сравнивается с шаблонами, содержащих общие географические термины – слова, которые определяют характер географического объекта, его род и вид (например, город, гора, озеро). Как говорилось выше, географическое название может представлять собой последовательность из нескольких слов, содержать тире, часть слов или все слова могут быть написаны с заглавной буквы, также в тексте могут идти подряд несколько названий. Задача извлечения географического названия состоит в определении границы этого названия (последовательности слов, в него входящей) в окрестности географического термина. Географические названия в тексте имеют разнородную внутреннюю структуру. Частью автоматической обработки является выявление этой структуры. Можно считать, что географические названия образуются по определенным формулам – словообразовательным моделям. Для выявления географических в тексте необходимо учитывать их правила написания в русском языке [10], используемые для задания конструкций лексико-семантических шаблонов, которые определяют входящие в конструкцию слова с учетом их морфологических характеристик.

Выявленные названия должны проверяться на правильность написания в соответствии с правилами грамматического согласования. Для определения начальной формы географического названия, чтобы избежать омонимии – совпадения в некоторых падежных формах, необходимо определить род, число и падеж анализируемых слов. Общие географические термины выполняют роль родовых слов, с которыми согласуются географические названия.

При определении принадлежности лексем к словарю рассматривается термин и его окрестность из двух слов, что позволяет находить названия, состоящие из одного или двух слов, употребляющихся в сочетании с общим географическим термином, например, город Великие Луки.

Удобным средством извлечения географических названий из теста являются лингвистические шаблоны. Лингвистический шаблон содержит формальное описание (образец) языковой конструкции, которую необходимо найти в тексте, чтобы извлечь название. Шаблон может быть записан с использованием регулярных выражений и учитывать особенности слов: регистр букв, последовательности букв, например, ([А-Я] [а-я]+) край

Специальная метка может представляет собой уникальный идентификатор географического объекта в базе данных географических названий. Формально вся процедура сводится к замене нормализованных лексем на специальные метки с идентификатором объекта или на метки с лексемами. Соответствие лексем и меток содержится в специальном географическом словаре.

5. Наконец, последний шаг – разрешение проблемы многозначности географических названий. Например, вполне определенной форме «Советский район» может быть поставлено в соответствие более 40 географических объектов (на основе данных [5]). Однако, среди всех возможных нужно выбрать тот, который максимально соответствует окружающему контексту. Здесь возможны несколько вариантов разрешения конфликта.
   - На основе иерархических связей - решение об идентификации объекта среди конкурирующих принимается на основе анализа иерархических связей соседних по тексту полностью идентифицированных объектов. Иерархические связи (административное подчинение, географическое расположение и пр.) как правило присутствуют в базах данных географических названий. Более того, идентификаторы объектов некоторых баз данных хранят эту иерархию в значении идентификационного кода, например, справочник ОКАТО [3]. В частности, для города Карасук код ОКАТО 50217501 содержит информацию о Карасукском районе (ОКАТО 50217000) и Новосибирской области (ОКАТО 50000000).
   - На основе геометрических параметров - решение об идентификации объекта среди конкурирующих принимается на основе минимизации расстояния до соседних по тексту полностью идентифицированных объектов. Расстояние вычисляется на основе координат объектов, присутствующих в базе данных географических названий. При этом возможны различные варианты критерия принятия решения.

Алгоритм фиксации географических объектов в произвольном тексте изображен на рис. 1.

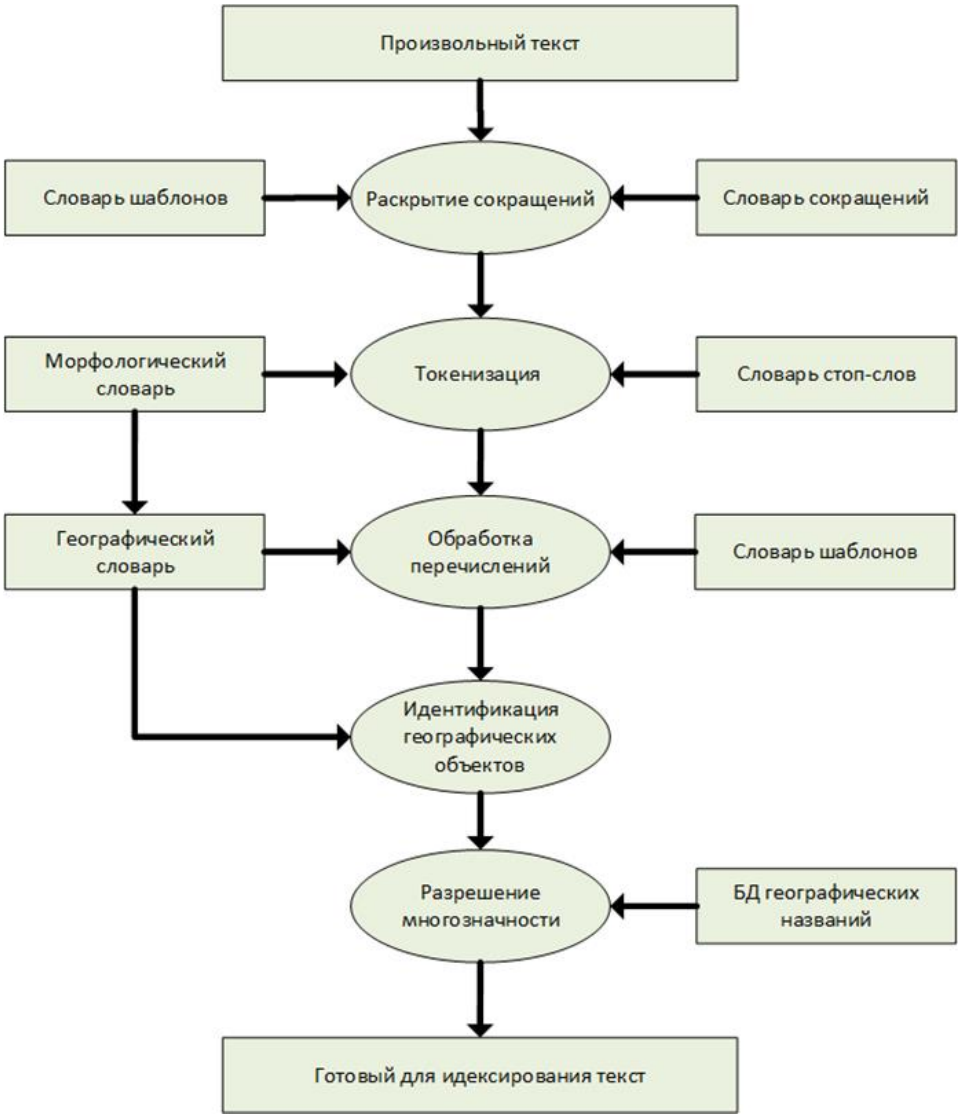



Рис. 1. Алгоритм фиксации географических объектов в произвольном тексте.

**Справочники.** Как следует из вышеописанной технологии, качество ее работы зависит от качества и полноты справочников, содержащих информацию о географических названиях. Сегодня в открытом доступе существует достаточно информации для создания собственной базы данных географических названий. Источниками информации могут быть:

- ОКАТО – общероссийский классификатор объектов административно-территориального деления [3].
- КЛАДР – класификатор адресов Российской Федерации [4].
- GeoNames – база данных, содержащая свыше 10 млн. географических названий и информацию о более 7,5 млн их уникальных характеристик. Среди характеристик: названия мест на разных языках, широта, долгота, высота над уровнем моря. Все эти характеристики разбиты по категориям, так что каждая характеристика географического объекта относится к одному из девяти классов. А каждая из этих категорий, в свою очередь, делится на подкатегории, общее количество которых более 600. Кроме наименований на различных языках, хранятся географические координаты, высота над уровнем моря, численность населения, административное деление и почтовые индексы. К сожалению, база данных содержит дубли, ошибки в наименованиях и другие неточности.
- База данных OSM (Open Street Map) [6] – открытая база данных географических объектов, включающая их геометрические и географические характеристики
- Getty тезаурус географических названий (TGN) [7] – содержит географические названия с точечными координатами, в том числе и ретроспективные. Недостаток – российские названия даны в транскрипции.
- Государственный каталог географических названий РОСРЕЕСТР [8] – содержит полный реестр официальных географических названий по регионам с точечными координатами.

Перечисленные информационные ресурсы содержат исходные данные, на основе которых формируется собственная база данных географических объектов, описанная ниже.

**Прототип стенда.** Для отработки технологии извлечения географических названий из текстов, проведения тестирования алгоритмов и сбора информации об ошибках был создан программный стенд, в котором реализованы описанные выше алгоритмы. В качестве системной основы реализации алгоритмов был выбран вариант на базе СУБД PostgreSQL, реализующей полный цикл обработки текстовой информации с возможностью расширения базовых функциональных возможностей как за счет дополнительных словарей и конфигураций, так и написанием дополнительных модулей на различных языках программирования [9].

Созданный прототип стенда включает в себя:

- Базу данных СУБД PostgreSQL 9.4 со специальной конфигурацией для полнотекстового поиска, ориентированного на географические объекты.
- Набор серверных WEB приложений (PHP скрипты), которые выполняются на стороне WEB сервера. Эти приложения обеспечивают связь с сервером баз данных PostgreSQL и клиентскими приложениями. Отдельным серверным приложением также является модуль для ZooSPACE, позволяющий анализировать текстовые данные, извлекаемые из различных библиографических баз данных.
- Набор клиентских WEB приложений (Java скрипты), которые выполняются на стороне WEB клиента. Эти приложения реализуют функции графических интерфейсов пользователя для управления работой стенда и визуализации найденных географических объектов на картах.

Для обеспечения работы стенда созданы

- Словари
  - Словарь сокращений с шаблонами на основе регулярных выражений - при помощи этого словаря раскрываются сокращения во входном тексте (шаг 1).
  - Словарь стоп-слов русского языка (russian.stop). Этот словарь входит в поставку PostgreSQL и в нашей конфигурации не менялся (шаг 2).
  - Морфологический словарь русского языка (ispell) с добавлением географических названий и орфографических правил для этих названий (ru_geo1.dict).

Фрагмент ru_geo1.dict

```
. . .
абажур/K
 . . .
Кольцово/M
Мошковский/A
Новосибирск-Южный/AEZ
. . .
```

- Географический словарь для замены лексем на комбинацию «метка + лексема». Этот словарь (geor1.ths) соответствует шаблону тезауруса (в терминах PostgreSQL тезаурус - это словарь замен: левая часть от символа «:» заменяется на правую часть, наличие символа «*» в первой позиции правой части предписывает не контролировать правую часть морфологическим словарем) и состоит из записей вида:

  Фрагмент файла geor1.ths

```
. . .
Бердск: */gn/1510350 Бердск
город Бердск: */gn/1510350 город Бердск
Советский район: */gn/490026,/gn/1491227 Советский район
. . .
```

– Конфигурация FTS (в терминах PostgreSQL), определяющую список словарей и порядок обработки текста (rugeo1):

Команды создания конфигурации FTS rugeo1

```
CREATE TEXT SEARCH DICTIONARY rugeo_ispell (TEMPLATE = ispell,
    dictfile = 'ru_geo1', afffile = 'ru', stopwords = 'russian');
CREATE TEXT SEARCH DICTIONARY tz_geo_1 (TEMPLATE = thesaurus,
    dictfile = 'geor1', dictionary = 'rugeo_ispell');
CREATE TEXT SEARCH CONFIGURATION rugeo1 (PARSER = "default");
ALTER TEXT SEARCH CONFIGURATION rugeo1 ADD MAPPING
    FOR hword WITH tz_geo_1,rugeo_ispell,russian_stem;
ALTER TEXT SEARCH CONFIGURATION rugeo1 ADD MAPPING
    FOR hword_part WITH tz_geo_1,rugeo_ispell,russian_stem;
```

Работу алгоритма фиксации географических названий можно проиллюстрировать на примере обработки фрагмента текста «В окрестностях города Новосибирска находятся: город Бердск, город Обь, поселок Краснообск и Наукоград Кольцово». В результате выполнения запроса

```
SELECT plainto_tsquery('rugeo1', 'В окрестностях города Новосибирска нахо-
дятся:  город Бердск, город Обь, поселок Краснообск и Наукоград Кольцово');
```

получим ответ – размеченный текст

```
'окрестность /gn/1496747 город новосибирск находиться
/gn/1510350 город бердск /gn/1496421 город обь /gn/1502091 поселок
краснообск /gn/1502847 наукоград кольцово'
```

Другой запрос

```
SELECT to_tsvector('rugeo1',
'В окрестностях города Новосибирска находятся: город Бердск, город Обь,
поселок Краснообск и Наукоград Кольцово');
```

вернет список лексем с указанием их позиции в тексте

```
"'/gn/1496421':10 '/gn/1496747':3 '/gn/1502091':13
'/gn/1502847':17 '/gn/1510350':7 'бердск':9 'город':4,8,11
'кольцово':19 'краснообск':15 'наукоград':18 'находиться':6
'новосибирск':5 'обь':12 'окрестность':2 'поселок':14"
```

Вид интерфейсов приложения для тестирования алгоритмов на стенде представлен на рис. 2.

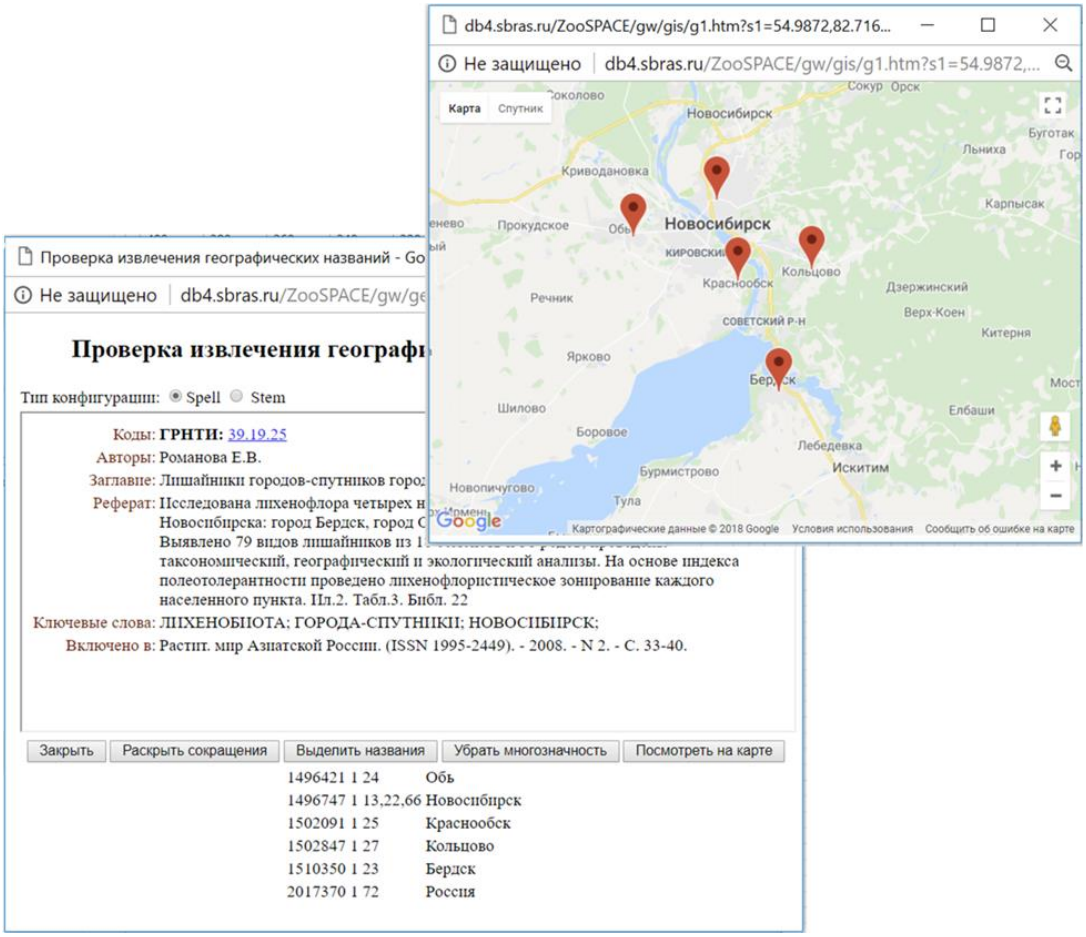



Рис. 2. Интерфейсы приложения для тестирования алгоритмов.

**Заключение.** В результате выполненных работ был создан прототип стенда для тестирования моделей и алгоритмов извлечения географических названий из неструктурированного текста для построения индексов как для текстового, так и для геометрического поиска. Предварительное тестирование показало, что предложенная технология обеспечивает высокую степень достоверности результатов при условии, что все справочники содержат информацию об определяемых географических объектах. Эффективность технологии зависит от полноты справочников. В настоящее время созданные справочники содержат информацию по географическим объектам Новосибирской области. В дальнейшем планируется расширить номенклатуру поддерживаемых регионов.

*Работа выполнена при частичной финансовой поддержке Интеграционного Проекта СО РАН (АААА18-118022190008-8), проекта по фундаментальным научным исследованиям (АААА-А17-117120670141-7), проекта РФФИ № 18-07-01457-а, а также при частичной поддержке грантового финансирования научных и (или) научно-технических исследований на 2018-2020 гг. МОН РК (№ АР 05133546).*

## ЛИТЕРАТУРА

[1] *Жижимов О.Л., Мазов Н.А.* Проблемы географической привязки цифровых объектов в электронных библиотеках // Тр. XII Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» (RCDL'2010). Казань, с. 207-214. (2010).

[2] *Барахнин В.Б., Жижимов О.Л., Куперштох А.А., Скачков Д.М., Федотов А.М.* Алгоритм извлечения из текстовых документов географических названий, отражающих содержание. Вестник Новосибирского государственного университета. Серия: Информационные технологии, Т.10, № 1, С.109-120. (2012).

[3] Общероссийский классификатор объектов административно-территориального деления (ОК 019-95). **http://protect.gost.ru/document.aspx?control=20&id= 134377**.
[4] Класификатор адресов Российской Федерации (КЛАДР). **http://kladr-rf.ru**.
[5] The GeoNames geographical database. **http://www.geonames.org**.
[6] Open Street Map. **http://wiki.openstreetmap.org**.
[7] Getty Thesaurus of Geographic Names (TGN). **http://www.getty.edu/research/tools/vocabularies/tgn/index.html**.
[8] Государственный каталог географических названий, РосРеестр. **https://rosreestr.ru/site/activity/geodeziya-i-kartografiya/naimenovaniya-geograficheskikh-obektov/gosudarstvennyy-katalog-geograficheskikh-nazvaniy**.
[9] *Бартунов О., Сигаев Ф.* Введение в полнотекстовый поиск в PostgreSQL. **http://citforum.ru/database/postgres/fts/bib.shtml**.
[10] Основные правила написания географических названий. **http://www.wikiznanie.ru/ru-wz/index.php/Основные_правила_написания_географических_названий**.

# ПРИМЕНЕНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ЗАДАЧИ РАСПОЗНАВАНИЯ ТЕСТА

*Пятаева А.В., Генза С.А.*
ФГАОУ ВО Сибирский федеральный университет, Красноярск

В работе показано применение искусственных нейронных сетей для задачи распознавания текста. Проведен сравнительный анализ эффективности использования различных архитектур нейронных сетей. Экспериментальные исследования проведены на тестовом наборе данных MNIST.
*Ключевые слова: сверточные искусственные нейронные сети, распознавание текста, тестовый набор MNIST.*

**Введение.** В настоящее время множество работ в рамках компьютерных технологий на базе методов глубокого обучения посвящается тематике распознавания текста [1-4]. Среди методов глубокого обучения свёрточные нейронные сети являются наиболее изученными [5-10]. Часто для проведения экспериментальных исследований в работах по обнаружению текста используется популярный набор данных MNIST, представляющий собой базу данных образцов рукописного написания текста. В работе [11] для классификации рукописных цифр исследованы возможности улучшения сверточных нейронных сетей для повышения точности классификации. Авторы [12] провели моделирование структуры нейронной сети обоняния мотыльков, затем созданные вычислительные модели обучаются чтению рукописных цифр. Большинство методов на основе свёрточных нейронных сетей (Convolutional Neural Network – CNN) извлекают признаки изображений на последнем уровне сети, используя единую архитектуру CNN с неограниченным количеством подходов квантования, это ограничивает использование промежуточных слоев для идентификации локальных особенностей изображения.

**Разработка нейронной сети для распознавания текста.** Сверточные нейронные сети относятся к технологии глубокого обучения, являющейся темой множества научных работ в последнее десятилетние. Глубокое обучение (англ. deep learnig) – совокупность методов машинного обучения, основанных на обучении признакам, а не на специализированных алгоритмах, применяемых для решения конкретных задач. Технология глубокого обучения активно используется в компьютерном зрении, машинном переводе, распознавании речи, при этом качество решения задач в этих областях с применением нейросетей глубокого обучения превосходит эффективность человека. Сверточная нейронная сеть (англ. convolutional neural network, CNN) – специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном [13, 14] нацеленная на эффективное распознавание изображений, входит в состав технологий глубокого обучения.

Идея CNN заключается в чередовании сверточных слоев (convolution layers) и слоев подвыборки или субдискретизирующих (subsampling layers). При этом конволюционная нейросеть может обладать разделяемыми весами, то есть часть нейронов некоторого слоя может использовать одни и те же весовые коэффициенты. Нейроны, использующие одни и те же веса, объединяются в карты признаков (feature maps) и каждый нейрон такой карты связан с частью нейронов предыдущего слоя. При вычислении выхода CNN получается, что каждый нейрон выполняет свертку – операцию конволюции – некоторой области предыдущего слоя (определяющего множество нейронов, связанных с этим нейроном).

Свертка – математическая операция, применяемая к двум функциям f и g, порождающая третью функцию, которая может быть рассмотрена как модифицированная версия одной из первоначальных. Также свертка является особым видом интегрального преобразования. Рассмотрим математическую формулировку операции свертки. Пусть f и g: Rd→R две функции, интегрируемые относительно меры Лебега на пространстве Rd. Тогда их сверткой называется функция $f \otimes g$: Rd→R, которая определяется формулой:

$$(f \otimes g)(x) \overset{def}{=} \int_{R^d} f(y)\, g(x-y)dy = \int_{R^d} f(x-y)\, g(y)dy.$$

С точки зрения анализа изображений операция свертки – это операция вычисления нового значения выбранного пиксела изображения, с учетом значений окружающих его пикселов. Для вычисления результата операции свертки используется матрица, называемая ядром свертки. Ядро свертки представляется обычно квадратной (или прямоугольной) матрицей n×n, n – нечетное число. Во время вычисления нового значения выбранного пиксела ядро свертки как бы «прикладывается» своим центром к данному пикселю. Окружающие пиксели так же накрываются ядром. Далее высчитывается сумма, где слагаемыми являются произведения значений пикселей на значения ячейки ядра, накрывшей данный пиксель. Сумма делится на сумму всех элементов ядра свертки. Полученное значение и является новым значением выбранного пикселя. Если применить свертку к каждому пикселю изображения, то в результате получится некий эффект, зависящий от выбранного ядра свертки. В настоящей работе выполнено конструирование CNN для задачи распознавания рукописного написания арабских цифр. Для распознавания рукописных цифр разработаны нейронные сети следующих архитектур.

*Архитектура 1.* Нейронная сеть с 2 полносвязными слоями:
- 1-й полносвязный слой, функция активации ReLU, 512 нейронов;
- 2-й полносвязный слой, функция активации Softmax, 10 нейронов;
- функция потерь перекрестной энтропии (cross-entropy loss).

*Архитектура 2.* Нейронная сеть с 3 полносвязными слоями:
- 1-й полносвязный слой, функция активации ReLU, 1024 нейрона;
- 2-й полносвязный слой, функция активации ReLU, 512 нейронов;
- 3-й полносвязный слой, функция активации Softmax, 10 нейронов;
- функция потерь перекрестной энтропии (cross-entropy loss).

*Архитектура 3.* Сверточная нейронная сеть с 3 слоями батч-нормализации:
- 1-й сверточный слой, функция активации ReLU, ядро 5x5, (28, 28, 32);
- 2-й сверточный слой, функция активации ReLU, ядро 5x5, (28, 28, 32);
- 1-й слой батч-нормализации;
- 1-й max-pooling слой, ядро 2x2, шаг 2x2, (14, 14, 32);
- 3-й сверточный слой, функция активации ReLU, ядро 3x3, (14, 14, 64);
- 4-й сверточный слой, функция активации ReLU, ядро 3x3, (14, 14, 64);
- 2-й слой батч-нормализации;
- 2-й max-pooling слой, ядро 2x2, шаг 2x2, (7, 7, 64);
- 1-й полносвязный слой, функция активации ReLU, 256 нейронов;
- 3-й слой батч-нормализации;
- 2-й полносвязный слой, функция активации Softmax, 10 нейронов;
- функция потерь перекрестной энтропии (cross-entropy loss).

*Архитектура 4.* Сверточная нейронная сеть с 2 слоями Dropout:
- 1-й сверточный слой, функция активации ReLU, ядро 5x5, (28, 28, 32);
- 2-й сверточный слой, функция активации ReLU, ядро 5x5, (28, 28, 32);
- 1-й max-pooling слой, ядро 2x2, шаг 2x2, (14, 14, 32);
- 1-й слой Dropout, вероятность исключения нейронов – 0.5;
- 3-й сверточный слой, функция активации ReLU, ядро 3x3, (14, 14, 64);
- 4-й сверточный слой, функция активации ReLU, ядро 3x3, (14, 14, 64);
- 2-й max-pooling слой, ядро 2x2, шаг 2x2, (7, 7, 64);
- 2-й слой Dropout, вероятность исключения нейронов – 0.5;
- 1-й полносвязный слой, функция активации ReLU, 256 нейронов;
- 2-й полносвязный слой, функция активации Softmax, 10 нейронов;
- функция потерь перекрестной энтропии (cross-entropy loss).

Перед обучением нейронной сети была проведена аугментация (augmentation) тренировочного набора данных. Каждое изображение подвергалось следующей аугментации:

- поворот на случайный угол от 0 до 10 градусов;
- случайное приближение и отдаление в пропорциях от 0.9 до 1.1 от исходного изображения;
- случайный горизонтальный сдвиг в пропорциях от -0.1 до 0.1 от ширины исходного изображения;
- случайный вертикальный сдвиг в пропорциях от -0.1 до 0.1 от высоты исходного изображения;

Для реализации предложенных архитектур нейронных сетей с использованием фреймворка Tensorflow [15] на языке программирования Python разработана программа для ЭВМ. Входное изображение размером 28x28 пикселей (для нейронных сетей без сверточных слоев изображение приводилось к вектору размера 784) в градациях серого подвергалось нормализации путем приведения интервала значений яркости пикселей к диапазону от 0 до 1, после чего подавалось на вход нейронной сети вместе с соответствующей меткой класса. Метка класса перед подачей на вход сети для корректной работы Softmax-функции и функции кросс-энтропии преобразовывалась в унитарный код (one-hot encoding). В качестве оптимизатора выбран Adam с шагом обучения 0.001, размер входного пакета (batch), для которого происходит единовременная корректировка весов – 128. Веса сети до обучения инициализируются нормально распределенными значениями с математическим ожиданием, равным 0 и стандартным отклонением, равным 0.1. Смещения (bias) слоев инициализируются нулевыми значениями. Для крайних пикселей в сверточных слоях используется автоматическое заполнение значений нулями для получения на выходе слоя той же размерности, что и на входе. Нейронная сеть обучалась 20 эпох на графическом процессоре NVIDIA GeForce GTX 650, итоговая точность выбиралась по лучшей эпохе.

**Экспериментальные исследования.** Для проведения экспериментальных исследований использован тестовый набор MNIST [16] (Modified National Institute of Standards and Technology) – база данных образцов рукописного написания цифр. Изображения из этого набора данных являются переработанными изображениями из другого набора данных – NIST. Образцы NIST имели размерность 20x20 и были взяты из бюро переписи населения США с добавлением тестовых образцов, написанных студентами американских университетов, после чего были нормализованы, прошли сглаживание и приведены к размеру 28x28 пикселей. База данных MNIST содержит 60000 тренировочных и 10000 тестовых полутоновых изображений размеров 28x28 в градациях серого. Примеры изображений приведены на рис. 1.

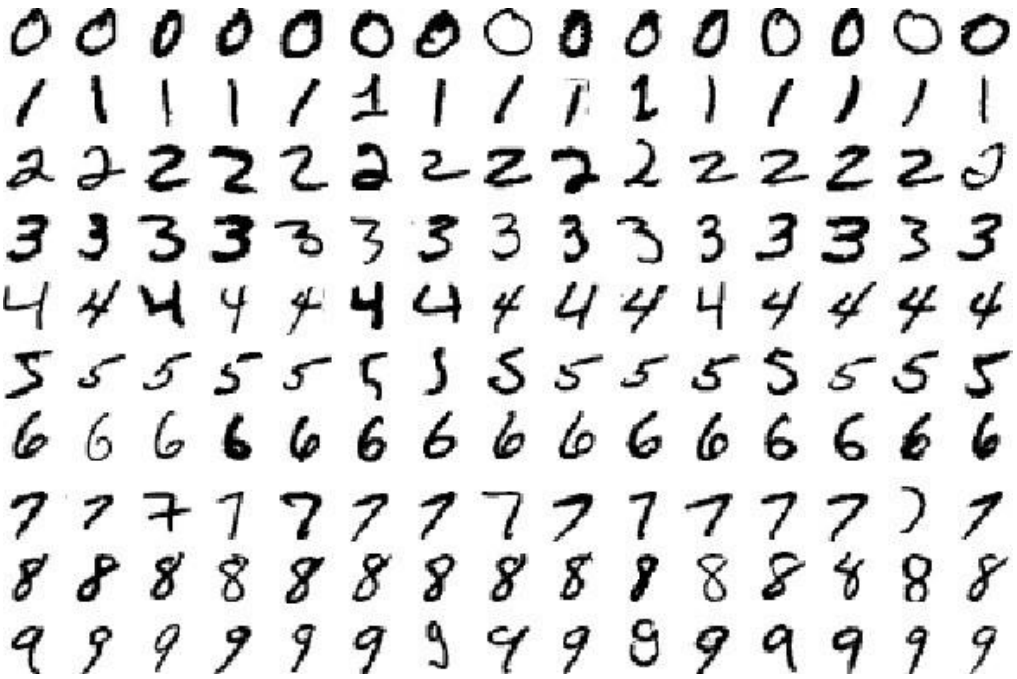


Рис. 1. Примеры изображений из набора данных MNIST.

Результаты сравнения нейронных сетей.

| **Архитектура нейронной сети** | **Количество оптимизируемых параметров** | **Точность** |
|---|---|---|
| Нейронная сеть с 2 полносвязными слоями | 407050 | 98,76 |
| Нейронная сеть с 3 полносвязными слоями | 1333770 | 98,90 |
| Сверточная сеть с Dropout | 887530 | 99,42 |
| Сверточная сеть с батч-нормализацией | 887978 | 99,58 |

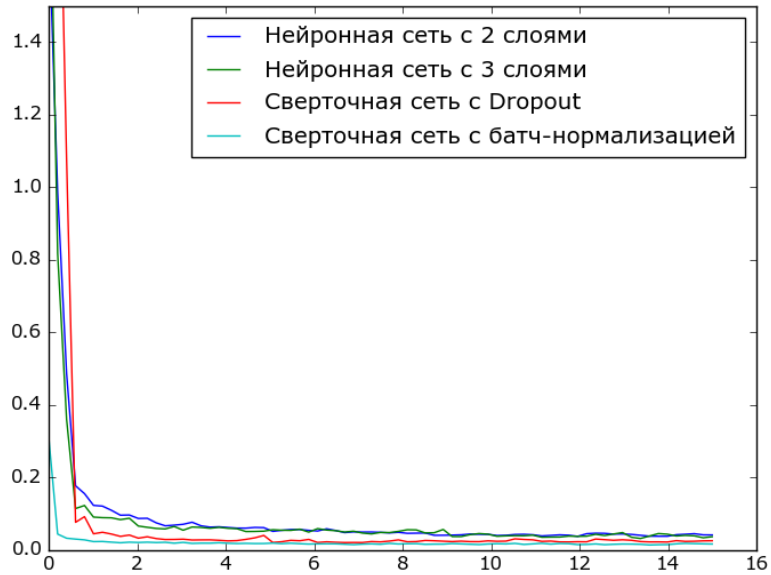



Рис. 2. Изменение ошибки для нейронных сетей различной архитектуры.

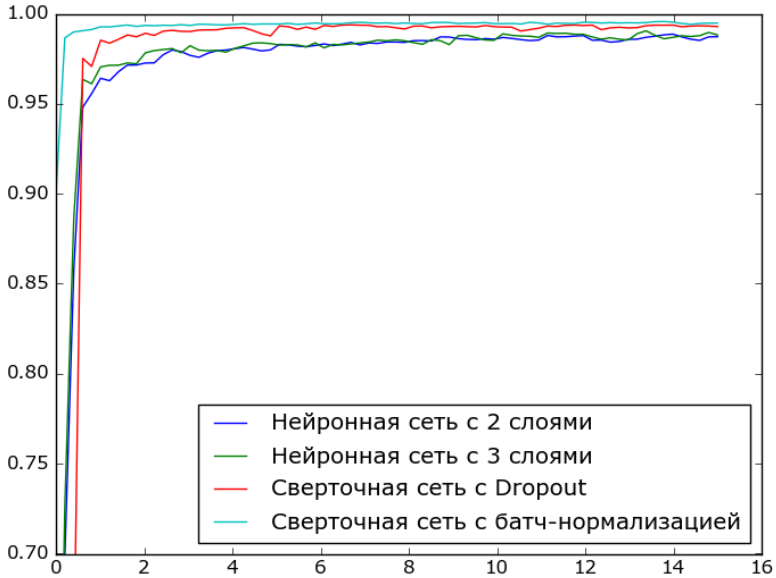



Рис. 3. Изменение точности распознавания для нейронных сетей различной архитектуры.

Сравнительные результаты экспериментальных исследований качества распознавания рукописных цифр для сверточных нейронных сетей различной архитектуры показаны в таблице. Точностью распознавания рукописных цифр рассчитывается как отношение правильно предсказанных классов к общему числу классов выборки.

График изменения ошибки на тестовых данных для каждой сети и их сравнение эффективности их работы приведено на рис. 2.

График изменения точности на тестовых данных для каждой сети и их сравнение приведено на рис. 3.

**Заключение.** По итогам эксперимента на тестовой части набора данных MNIST была получена максимальная точность для лучшей эпохи 99,58%. Такую точность показала сверточная нейронная сеть со слоями нормализации по мини-батчам, такая сеть сходится заметно быстрее других, имея при этом наибольшую итоговую точность. Дальнейшим развитием настоящей работы является исследование сочетания сверточных и рекуррентных слоев нейронной сети для задачи распознавания текста на изображениях сложных графических сцен.

## ЛИТЕРАТУРА


[1] *Jaderberg M., Simonyan K., Vedaldi A., Zisserman A.* Reading Text in the Wild with Convolutional Neural Networks // International Journal of Computer Vision. 2016. P. 1-20.

[2] *Graves A., Fernandez S., Gomez F., Schmidhuber J.* Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks // ICML '06 Proceedings of the 23rd international conference on Machine learning. 2006. P. 369-376.

[3] *Shi B., Bai X., Yao C.* An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015. P. 99.

[4] *Bissacco A., Cummins M., Netzer Y., Neven H.* Photoocr: Reading text in uncontrolled conditions // ICCV. 2013. P. 785-792.

[5] *Zheng Y., Iwana B.K., Uchida S.* Mining the displacement of max-pooling for text recognition // Pattern Recognition. 2019. Vol. 93. P. 558-596.

[6] *Banerjee I., Ling Y., Chen M.C., Hasa S.A., Langlotz C.P., Moradzadeh N., Chapman B., Amrhein T., Mong D., Rubin D.L., Farri O., Lungren M.P.* Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification // Artificial Intelligence in Medicine. 2019. Vol. 97. P. 79-88.

[7] *Gu J., Wang Z., Kuen J., Ma L., Sharoudy A., Shuai B., Liu t., Wang X., Wang G., Cai J., Chen T.* Recent Advances in Convolutional Neural Networks // Pattern Recognition. 2018. Vol. 77. Issue C. P. 354-377.

[8] *Parkhi O.M., Vedaldi A., Zisserman A.* Deep face recognition // Proceedings of the British Machine Vision Conference (BMVC). 2015. Vol. 1. P. 41.1-41.12.

[9] *Wang Haiyan, He Zhongshi, Huang Yongwen, Chen Dingding, Zhou Zexun.* Bodhisattva head images modeling style recognition of Dazu Rock Carvings based on deep convolutional network // Journal of Cultural Heritage. 2017. Vol. 27. P. 60-71.

[10] *Kinghorn P., Zhang L., Shao L.* A Hierarchical and Regional Deep Learning Architecture for Image Description Generation // Pattern Recognition Letters // Pattern Recognition Letters. 2019. Vol. 119. P. 77-85.

[11] *Alevar-Sandoval R.F., Sanco-Gomer J.I., Figueiras-Vidal A.R.* On improving CNNs performance: The case of MNIST // Information Fusion. 2019. Vol. 52. P. 106-109.

[12] *Delahunt C.B., Kutz J.N.* Putting a bug in ML: The moth olfactory network learns to read MNIST // Neural Networks. 2019. Vol. 188. P. 54-64.

[13] *LeCun Y.* Learning processes in an asymmetric threshold network. Disordered Systems and Biological Organization. NATO ASI Series (Series F: Computer and Systems Sciences). Springer, Berlin, Heidelberg. 1986. Vol 20.

[14] *LeCun Y.* Theoretical framework for back-propagation // InProceedings of the 1988 Connec-tionist Models Summer School. Morgan Kaufmann, CMU. 1988. P.21-28.

[15] An end-to-end open source machine learning platform. **https://www.tensorflow.org** (дата обращения 22.06.2019).

[16] *Qiao, Yu.* The MNIST Database of handwritten digits. **http://www.gavo.t.u-tokyo.ac.jp/~qiao/database.html** (дата обращения 22.06.2019).

# ОПЕРАТИВНЫЙ РЕГИОНАЛЬНЫЙ СПУТНИКОВЫЙ МОНИТОРИНГ ОКРУЖАЮЩЕЙ СРЕДЫ

# СЕРВИСЫ АНАЛИЗА СЕРИИ КОСМОСНИМКОВ ДЛЯ ОЦЕНКИ ДИНАМИКИ ИЗМЕНЕНИЯ ОБЪЕКТОВ

*Авраменко Ю.В.[(1)], Ружников Г.М.[(2)], Фёдоров Р.К.[(1)]*
[(1)] Институт динамики систем и теории управления СО РАН, Иркутск
[(2)] Иркутский научный центр СО РАН, Иркутск

Для повышения эффективности использования данных ДЗЗ в задачах управления территориальным развитием реализованы вспомогательные сервисы, позволяющие каталогизировать и упрощать использование методов анализа данных ДЗЗ в виде сервисов. Разработаны сервисы, применяющие методы опорных векторов и нейронные сети распространенных архитектур Unet, FPN, Linknet и PSPNet. Вспомогательных сервисы позволяют регулярно и автоматически проводить обработку данных ДЗЗ, в том числе с использованием сценариев, связывающих сервисы и формирующих единый процесс обработки.
*Ключевые слова: ДЗЗ, космоснимки, нейронные сети, обучение распознаванию*.

**Введение**. Данные дистанционного зондирования Земли (ДЗЗ) активно используются в задачах управления территориальным развитием, таких как мониторинг социально-эколого-экономического положения, оценка природно-ресурсного потенциала, проведение архитектурных и проектных работ и т.д. Большие объемы собираемых данных ДЗЗ, быстрое развитие коммуникаций, рост числа сенсоров, повышение их разрешающей способности способствуют разработке и внедрению современных методов, технологий (нейронные сети, кластерный анализ), получения и обработки данных ДЗЗ. Однако их применение не тривиально, занимает достаточно много времени из-за разрозненности источников данных и методов обработки, необходимости ручной настройки и установки программного обеспечения.

Возможность применения методов ДЗЗ для решения задачи обработки зависит от многих факторов: формата, разрешающей способности, используемого сенсора, анализируемой территории, целей обработки и т.д. Особенностью существующих методов является то, что методы, показывающие хорошие результаты на одной территории, могут оказаться неработоспособны при решении подобных задач на других территориях. Для методов обработки данных, реализующих распознавание с учителем, требуется формирование обучающих выборок, учитывающих территориальную специфику.

В настоящее время комплексный анализ данных ДЗЗ для задач управления территориальным развитием обосновывает необходимость интеграции и системного применения совокупности различных методов обработки ДЗЗ, многие из которых могут быть реализованы в виде сервисов, что упрощает их использование. Также требуется создать вспомогательные сервисы, позволяющие каталогизировать и упростить использование сервисов обработки и анализа, а также выбрать наиболее подходящие сервисы для задачи и территории, при необходимости провести обучение и настройку. Реализация сервисов анализа данных ДЗЗ и вспомогательных сервисов упростит применение данных ДЗЗ, уменьшит порог вхождения пользователей и позволит более активно использовать данные ДЗЗ в задачах управления территориальным развитием.

**Инфраструктурные функции.** В ИДСТУ СО РАН на основе геопортала [1, 2] разработаны некоторые вспомогательные функции. Общая схема приведена на рис. 1. Рассмотрим их подробнее.

**Каталог сервисов.** Информацию о сервисах предоставляет каталог сервисов. Помимо хранения информации о сервисах, каталог предоставляет интерфейс для регистрации, поиска и запуска сервисов. Для удобства и корректности заполнения информации о регистрируемых сервисах была разработана специальная форма. Каждый сервис в каталоге задается параметрами: имя, тип сервиса (WPS [3], REST [4] и т.д.), список вычислительных узлов, на которых данный сервис может выполняться. Пользователь так же определяет, может ли сервис выполняться длительное время (больше 60 минуты). Если сервис помечен как длительный, то каталог сервисов будет взаимодействовать с ним в асинхронном режиме. Например, для стандарта

WPS при вызове сервиса с длительным выполнением возвращается URL XML-файла, в котором отображается процент сделанной работы, а по завершении WPS-сервиса – результат. При указании сетевого адреса вычислительного узла система автоматически опрашивает, какие сервисы данный узел предоставляет (получение информации о поддерживаемых сервисах осуществляется с помощью специальных команд соответствующих стандартов, например GetCapabilities у стандарта WPS). Пользователь выбирает один из предоставленных сервисов, и система загружает описание параметров сервиса и отображает как входные, так и выходные параметры. Для каждого входного параметра пользователь выбирает элемент управления (система элементов управления является частью геопортала), с помощью которого будут вводиться параметры для данного сервиса. После сохранения указанной информации в каталоге, сервис готов к использованию.

В каталоге имеется модуль выполнения сервисов. При выполнении сервисов ДЗЗ требуется учитывать размер космоснимков, особенно высокого и сверхвысокого разрешения, который может быть достаточно большим. Обработка снимков может занять много времени. Некоторые методы полностью загружают космоснимки в оперативную память. Соответственно на больших изображениях корректно не работают или используют файл подкачки, что приводит значительному увеличению времени обработки. С целью ускорения работы веб-сервиса по времени, уменьшения требований к объему оперативной памяти, предлагается входные данные разделить на пространственные ячейки, каждую ячейку обработать на отдельном вычислительном узле с установленным сервисом, а затем объединить результаты (по аналогии с программной моделью MapReduce [5]). Схема [6] параллельного выполнения веб-сервиса приведена на рис. 2. Для входных параметров, соответствующих обрабатываемым изображениям, задается спецификация, по которой производится деление на ячейки (минимальный и максимальный размер области обработки, размер наложения ячеек и т.д.). Для выходных параметров определяется спецификация объединения данных (обработчик, решение конфликтных ситуаций и т.д.). Модуль геопортала принимает на вход параметры, передаваемые на вход сервису. Далее параметры, для которых определена спецификация, разделяются. Затем происходит запуск копий сервиса на отдельных вычислительных узлах с передачей соответствующих ячеек изображений. По мере завершения работы копий сервисов данные собираются, вызывается обработчик для их объединения, и полученный результат возвращается пользователю.

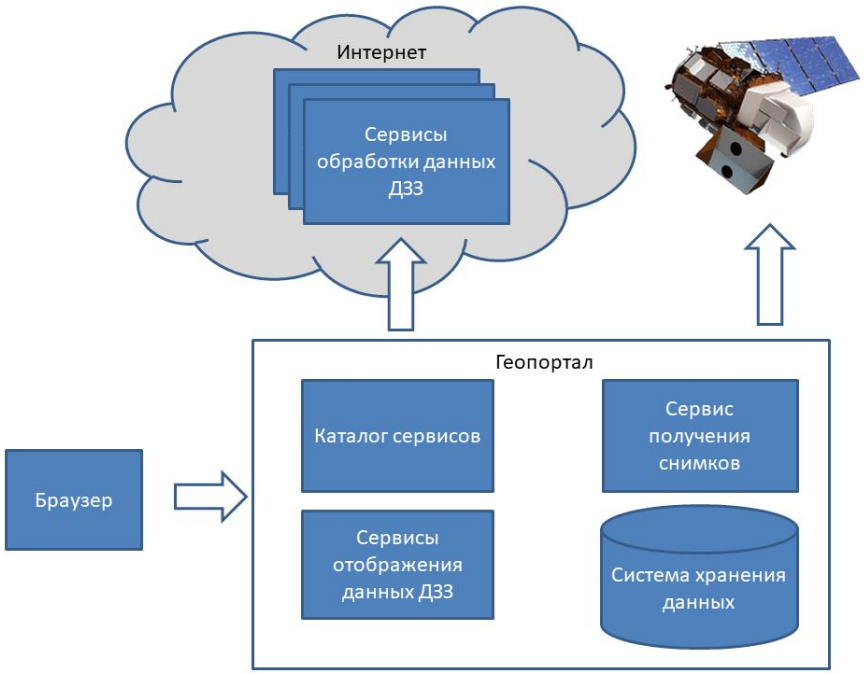



Рис. 1. Инфраструктура обработки данных ДЗЗ.

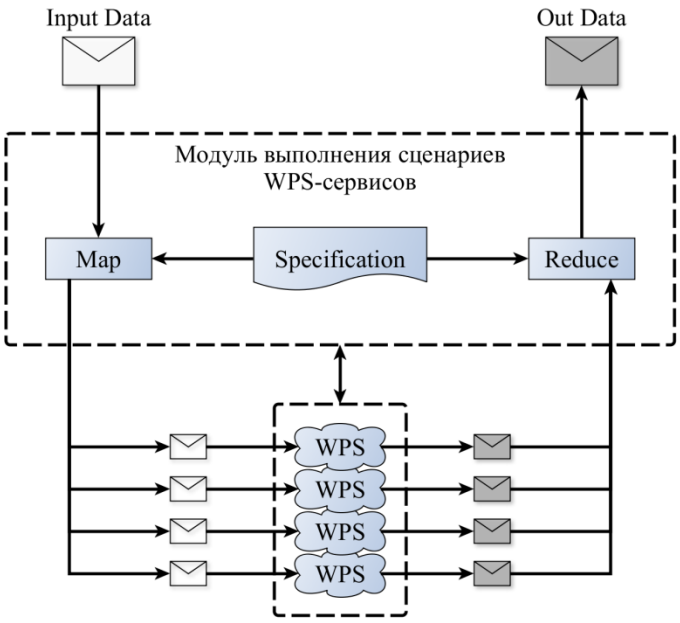


Рис. 2. Схема распараллеливания выполнения веб-сервиса.

Распараллеливание выполнения веб-сервиса позволяет уменьшить общее время выполнения веб-сервиса, а также задействовать доступные аппаратные мощности для решения задачи.

**Система хранения данных.** Система хранения данных (СХД) предоставляет пользователю веб-интерфейс для работы с файлами (космоснимки, результаты обработки, классификаторы, и т.д.). При работе с данными ДЗЗ система хранения данных является критическим местом из-за их объема. Требуется постоянное расширение объема системы.

**Сервис получения снимков.** Ряд задач требует регулярной обработки космоснимков без участия человека, поэтому необходимо автоматизировать получение снимков из различных источников, предоставляющих API. Для этого разработан сервис, который получает снимки с сенсоров Landsat 5,7,8. Входными параметрами сервиса являются: период съемки, задаваемый диапазоном дат начала и конца; тип сенсора; область интереса. Результатом работы сервиса является набор сцен. В основе сервиса лежит проект Landsat578 (https://github.com/dgketchum/Landsat578). Данный проект позволяет снимки в соответствии с системами разграфки WRS-1 и WRS-2. Поэтому для удобства пользователей было реализовано автоматическое определение номера сцен (path, row), соответствующих области интереса. Сервис последовательно получает слои каждой сцены, сохраняет их в архив и возвращает пользователю.

**Сервисы отображения данных ДЗЗ.** Разработаны сервисы для отображения данных ДЗЗ на карте геопортала. Первый сервис реализует публикацию растровых данных на основе стандарта WMS [2], что позволяет отображать картографическую информацию в рамках геопортала и на других ресурсах и системах. На вход сервису подается два параметра: растровый файл в формате GEOTIFF; стили отображения в формате JSON. Отображение можно задать следующими способами: градиентное изменение цвета в зависимости от значений; для каждого диапазона значений задается свой цвет; RGB отображение при наличии соответствующих каналов. В рамках геопортала разработан редактор для задания отображения. Сервис разработан на основе открытой программной системы Mapserver [7]. Для ускорения отображения данных за счет кэширования карт в виде тайлов применяется Mapcache [8]. Часто данные ДЗЗ предоставляются в виде набора файлов, соответствующих спектральным каналам. Информативным отображением для пользователя является цветное, в котором комбинируется каналы. Второй сервис Bands_to_RGB строит RGB файл в формате GEOTIFF из трех монохромных изображений, соответствующих каналам Landsat. В основе сервиса используется модуль gdal_merge_simple (https://github.com/gina-alaska/dans-gdal-scripts#gdal_merge_simple).

**Сервисы обработки данных ДЗЗ.** Перечислим некоторые из реализованных в виде сервисов методов обработки ДЗЗ.

Сервис обработки SRTM данных. Создает карты уклонов и экспозиций. Перед применением необходимо задать входные параметры: extent – размер входного растра, elevation – имя выходного растра, slope – имя выходного растра уклонов, aspect – имя выходного растра направлений уклонов, cell_size – размер ячейки (масштаб) результирующего растра. Размер входного растра определяется автоматически, через взаимодействие с интерактивной картой, пользователь с помощью инструмента Rectangle (рамка) выделяет интересующую область. Сервис находит соответствующие области файлы исходных данных. Исходные SRTM данные хранятся в системе хранения данных. Сервис производит склеивание и обрезание исходных файлов (формирование входного растра), перерасчет масштаба в соответствии с нужным пользователю, обработка входного растра инструментами уклон и экспозиция. На выходе сервиса пользователь получает три выходных файла в формате GEOTIFF в системе координат WGS84: выходной растр рельефа, растр уклонов, растр направлений уклонов. Эти файлы сохраняются на системе хранения данных по пути указанному пользователем.

Сервис вегетационного индекса NDVI. Пользователь указывает расположение исходного файла и путь, по которому будет сохранен результирующий файл. Выходной файл содержит в себе значения NDVI в диапазоне [-1,1], по желанию можно изменить диапазон выходных значений на диапазон [0,255] что удобно для представления в градациях серого или диапазон [0,200], что подходит для формирования карты цветов.

Сервис сегментации космоснимков. При помощи сегментации осуществляется поиск местоположений объектов на растровых изображениях. Сервис использует библиотеку Segmentation models (https://github.com/qubvel/segmentation_models), основанную на фреймворке Keras [9] (TensorFlow [10]), предоставляющую высокоуровневый API для использования нейронных сетей известных архитектур (Unet, FPN, Linknet, PSPNet) с целью двоичной и мультиклассовой сегментации. С помощью данной библиотеки, возможно, проводить обучение нейронной сети либо осуществить загрузку весов предобученных моделей (VGG, ResNet, SE-ResNet, ResNeXt, SE-ResNeXt, SENet154, DenseNet, Inception, MobileNet, EfficientNet) для более быстрого обучения. Текущая реализация сервиса работает на модели U-net и весах resnet34. Сервис получает на вход растровое изображение модели RGB, делит его на ячейки и производит классификацию каждой из них. Из обработанных ячеек формируется выходной растр – бинарная маска, с отмеченными положениями объектов. Размер ячейки (32x32 пикселя) соответствует искомым объектам и подобран таким образом, чтобы нейронная сеть могла верно, классифицировать как весь объект, так и его часть. На рис. 3 показан результат применения сервиса для поиска местоположений ледников.

Сервис SVM. Метод опорных векторов с одной стороны позволяет по малой обучающей выборке построить работающий классификатор, что полезно для тестирования и получения оперативных результатов; с другой стороны, он обладает ограничением на количество распознаваемых классов, не более 2-х. Особенность реализации сервиса заключается в том, что изображение представляется в виде вектора признаков. Вектор признаков – характеристики частот вхождения в область изображения раз личных элементов. Элементами выступают текстурные (метод Local Binary Pattern [11]) и спектральные (значения яркости пикселей цветовой модели RGB) признаки. Сочетание текстурных и спектральных признаков позволяет получить более точный результат. При реализации использовался программный интерфейс передачи информации MPI [12] библиотека SVM-light [13]. На рисунке 4 показан результат выделения растительного покрова.

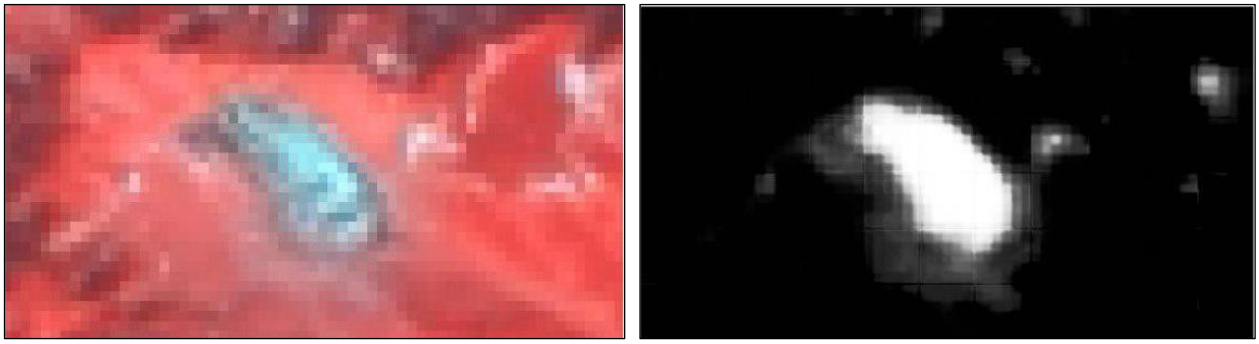


Рис. 3. Снимок Landsat и результат классификации.

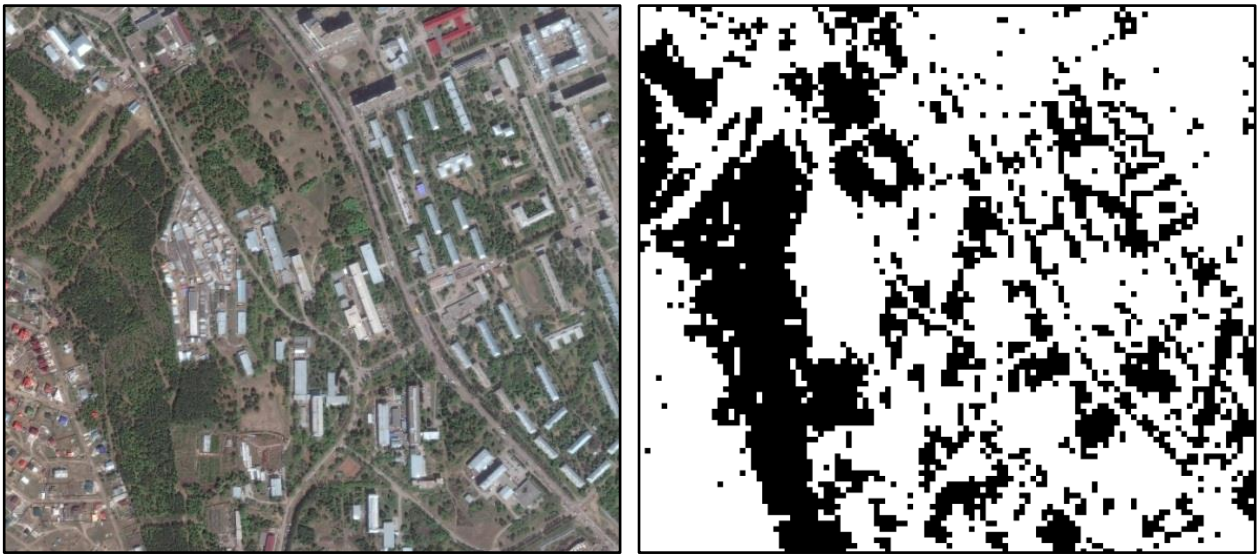


Рис. 4. Исходное изображение слева, результат классификации методом SVM справа.

**Заключение**. На основе разработанных вспомогательных сервисов будет формироваться набор сервисов обработки данных ДЗЗ, ориентированных на решение практических задач управления территориальным развитием. На текущий момент реализованы сервисы, применяющие метод опорных векторов, нейронный сети на основе распространенных архитектур Unet, FPN, Linknet или PSPNet. Реализованные сервисы позволяют регулярное и автоматическое применение сервисов, в том числе с использованием сценариев, связывающих сервисы и формирующих единый процесс обработки.

*Работа выполнена при частичной финансовой поддержке РФФИ (гранты №№ 18-07-00758-а, 17-57-44006-монг_а, 17-47-380007-р), интеграционной программы СО РАН (интеграционного проекта №37) и ЦКП ИИВС ИРНОК.*

## ЛИТЕРАТУРА

[1] *Бычков И.В.* Инфраструктура информационных ресурсов и технологии создания информационно-аналитических систем территориального управления / И.В. Бычков, Г.М. Ружников, Р.К. Фёдоров, Ю.В. Авраменко [и др.]. Новосибирск: Издательство СО РАН, 2016. 242 с.

[2] *Paramonov V., Fedorov R., Ruzhnikov G., Shumilov A.* Web-Based Analytical Information System for Spatial Data Processing // Communications in Computer and Information Science. 2013. Vol. 403. P. 93-101.

[3] Geospatial and location standards [Электронный ресурс] // Open Geospatial Consortium. URL: **http://www.opengeospatial.org** (дата обращения: 04.06.2019).

[4] *Erik Wilde, Cesare Pautasso.* REST: From Research to Practice. Springer Science & Business Media, 2011. 528 p.

[5] *Созыкин А.В., Гольдштейн М.Л.* Система обработки изображений с автоматическим распараллеливанием на основе MapReduce // Вестник Южно-Уральского государственного университета. 2012. № 27 (286). С. 109-118.
[6] *Шумилов А.С., Авраменко Ю.В.* Спецификация распараллеливания обработки векторных данных в модели MapReduce // Информационные и математические технологии в науке и управлении. 2017. Т. 4. С. 71-79.
[7] Mapserver [Электронный ресурс] // Mapserver Consortium. URL: **http://mapserver.org** (дата обращения: 04.06.2019).
[8] MapCache [Электронный ресурс] // MapCache. URL: **https://mapserver.org/mapcache** (дата обращения: 04.06.2019).
[9] Keras. URL: **https://keras.io**.
[10] TensorFlow. URL: **https://www.tensorflow.org**.
[11] *Ojala T.* A comparative study of texture measures with classification based on featured distributions / T. Ojala, M. Pietikäinen, D. Harwood // Pattern Recognition. 1996. Vol. 29, № 1. P.51-59.
[12] MPI: A Message-Passing Interface Standard. URL: **http://www.mpi-forum.org/docs/mpi-3.0/mpi30-report.pdf**.
[13] SVMlight Support Vector Machine. URL: **http://svmlight.joachims.org**.

# СОПОСТАВЛЕНИЕ ДВУХ МЕТОДИК ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ОБЛАЧНОГО ПОКРОВА ПО ДАННЫМ СКАНЕРА МСУ-МР (КА МЕТЕОР-М №2) И AVHRR (КА NOAA-18, 19, МЕТОР-В)

*Волкова Е.В.[(1)], Косторная А.А.[(2)], Голомолзин В.В.[(2)], Кузьминых С.А.[(2)], Амикишиева Р.А.[(2)]*
[(1)] ФГБУ «НИЦ «Планета», Москва
[(2)] СЦ ФГБУ «НИЦ «Планета», Новосибирск

Обсуждаются результаты сравнения расчетов двух методик дешифрирования облачности и определения ее параметров – комплексной пороговой методики (КПМ), разработанной в ФГБУ «НИЦ «Планета», и методики дешифрирования облачного покрова, разработанной в СЦ ФГБУ «НИЦ «Планета». Для территории Западной Сибири для теплого периода (апрель-август) 2019 г. определена точность восстановления по данным радиометров МСУ-МР и AVHRR следующих параметров облачности: тип облачности, высота и температура верхней границы облачности, оптическая плотность облачного слоя, фазовое состояние и эффективный радиус облачных частиц.
*Ключевые слова: параметры облачного покрова, МСУ-МР, Метеор-М №2, Западная Сибирь, AVHRR.*

**Введение.** Использование специализированной спутниковой информации об облачности в оперативно-прогностической работе метеослужб дает возможность более точного анализа синоптической ситуации. Облачный покров – самый наглядный индикатор погодообразующих процессов. На сегодняшний день существует множество методик автоматического дешифрирования облачности и определения ее параметров, основанных на разных подходах. В работе проводится сравнение полученной информационной продукции по двум методикам дешифрирования облачности и определения ее параметров: комплексной пороговой методики (КПМ) для обработки данных МСУ-МР, созданной в ФГБУ «НИЦ «Планета», и методики дешифрирования облачного покрова, разработанной в СЦ ФГБУ «НИЦ «Планета» для данных AVHRR.

**Комплексная пороговая методика.** Созданный в ФГБУ «НИЦ «Планета», Специализированный Программный Комплекс (СПК) «Западная Сибирь» позволяет автоматически в круглосуточном режиме определять параметры облачного покрова, осадков, опасных явлений погоды (ОЯП), подстилающей поверхности и приземного слоя воздуха для территории Западной Сибири (49-66° с.ш. и 60-95° в.д.) по данным сканера МСУ-МР, расположенного на полярно-орбитальном метеоспутнике «Метеор-М» №2. Кроме того, СПК проводит сравнение результатов расчетов с наземными метеонаблюдениями, климатическими оценками, аналогичными расчетами по данным других спутниковых приборов и выдает автоматическую оценку качества классификации. В основе СПК лежат Комплексная пороговая методика (КПМ) автоматической классификации спутниковых данных с целью определения параметров облачного покрова, осадков и ОЯП [1, 2] и методика расчёта параметров подстилающей поверхности и приземного слоя воздуха [3].

В качестве предикторов КПМ использует измерения МСУ-МР в каналах 1-6 ($\lambda$=0,6, 0,85, 1,7, 3,8, 11 и 12 мкм), а также их разности. Дополнительно используются прогностические поля о вертикальном распределении температуры воздуха на стандартных барических уровнях атмосферы и у поверхности земли и атмосферном давлении на уровне моря, а также цифровая модель рельефа и параметры облачности и осадков, полученные на начальных этапах классификации. Пороговые значения предикторов рассчитываются для каждого пиксела спутникового изображения как функции высоты места над уровнем моря, высоты солнца, номера календарного дня, приземной температуры воздуха и др.

Информационными продуктами, полученными в СПК (с разрешением 1’ по широте и 1,5‘ по долготе), являются оценки параметров облачного покрова (облачная маска, типы облачности, максимальная и суммарная водность облачного слоя, температура и высота верхней границы облачности (ВГО), высота нижней границы облачности, фазовое состояние воды в

облачных частицах в слое вблизи ВГО, оптические плотность и толщина облачного слоя, эффективный радиус облачных частиц), осадков (зоны осадков разной интенсивности, тип осадков у поверхности земли, суточные, месячные и годовые суммы осадков) и ОЯП (зоны гроз, града и обледенения разной интенсивности, высота верхней и нижней границ каждого слоя обледенения), а также параметров подстилающей поверхности (температура почвы, эффективная температура подстилающей поверхности, излучательная способность поверхности почвы, вегетационный индекс, листовой индекс и проективное покрытие растительностью) и приземного воздуха (температура воздуха у поверхности растительного покрова).

Проведена первичная валидация информационных продуктов СПК с данными наземных наблюдений на метеостанциях, климатическими оценками и аналогичными оценками по данным AVHRR с КА NOAA [4] на примере архива синхронных спутниковых и наземных наблюдений за 2015-2018 гг. Результаты валидации показывают хорошее качество получаемых спутниковых оценок для территории Западной Сибири, сопоставимое с аналогичными оценками для ЕТР по данным МСУ-МР [5] и AVHRR и зарубежными аналогами (например, полученных в центрах EUMETSAT SAF CM (http://www.cmsaf.eu)). Восстановленные в СПК «Западная Сибирь» характеристики облачности, в основном, удовлетворяют предъявляемым к ним пользователями требованиям и могут быть рекомендованы к использованию в качестве дополнения и равноценной замены наземным метеонаблюдениям для мезомасштабного мониторинга облачного покрова и подстилающей поверхности в оперативном режиме и для климатических исследований.

**Методика дешифрирования облачности.** Разработанная в СЦ ФГБУ «НИЦ «Планета» методика автоматического дешифрирования облачного покрова и расчета его характеристик в качестве исходных данных использует данные радиометра AVHRR (Level-1b) космических аппаратов серии NOAA и MetOp-B. Итоговыми информационными продуктами являются карты классификации облачности по типам, высоте и температуре ВГО. В основе методики лежат расчеты алгоритмов AWG, после выполнения которых производятся поэтапные оценки количественных характеристик облачности. На заключительном этапе производится подробная морфологическая классификация облачности с помощью порогового метода [6]. Для определения степени достоверности расчетов методики были проведены оценки качества полученных информационных продуктов по данным наземной наблюдательной сети, данным доплеровского метеолокатора и продуктам КА CALIPSO [7]. Получены следующие средние значения достоверности восстановленных параметров: для типов облачности ~97 %, для высоты ВГО – 80 %; вероятность случаев ложного детектирования облачности – 1,6 %. Обнаружены месяцы (апрель-сентябрь) с наивысшим качеством восстановления параметров облачности. Выявлены синоптические ситуации, при которых отмечается рост случаев ложного детектирования облачности [8].

**Условия сравнения.** Сопоставление двух методик проводилось для территории, ограниченной 49-66° с.ш. и 60-95° в.д (рис. 1), с учетом временно́го и пространственного совмещения. Для сопоставления выбирались пиксели на классифицированных изображениях МСУ-МР и AVHRR, для которых рассчитанное в заданном окне значение корреляции больше порогового. С учетом времени развития облачности в летний период разница во времени съемок задавалась не более 10 минут. Сравнение проводилось для весеннего и летнего периода года (апрель – август 2019 г.) по шести видам продуктов: тип облачности, фазовое состояние и эффективный радиус облачных частиц, оптическая плотность облачного слоя, температура и высота ВГО. Минимальное покрытие рассматриваемой территории расчетными данными составляло не менее 30%.

**Результаты.** При проведении сравнения двух методик рассчитывались: вероятность совпадения оценок параметров облачности обеими методиками относительно КПМ и относительно методики дешифрирования облачности, а также среднее отклонение, среднее абсолютное отклонение, среднее квадратичное отклонение и стандартное отклонение, рассчитанные как по каждому классу облачности, так и суммарно по всей облачности.

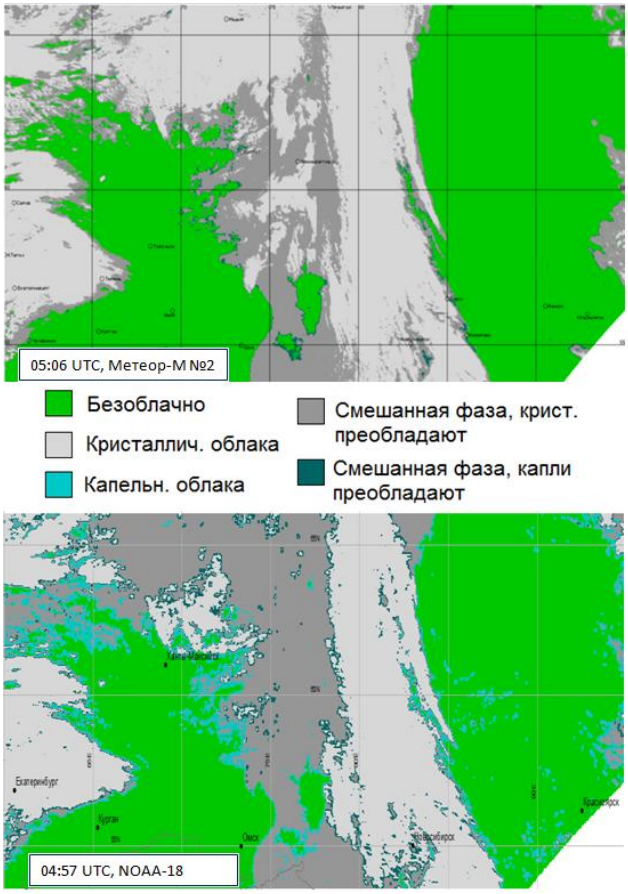



Рис. 1. Фазовое состояние облачности, рассчитанное КПМ (сверху)
и методикой дешифрирования облачности (снизу), 27.03.2019 г.

**Заключение.** Проводимое сопоставление направлено на повышение достоверности восстановления характеристик облачности в Сибири. В дальнейшем, исследование будет проведено для осеннего и зимнего периодов года.

## ЛИТЕРАТУРА

[1] *Волкова Е.В.* Определение параметров облачного покрова и осадков по данным МСУ-МР с полярно-орбитального метеоспутника «Метеор-М» №2 для Европейской территории России // Современные проблемы дистанционного зондирования Земли из космоса. 2017. Т. 14. С. 300-320.

[2] *Волкова Е.В., Гиричев Д.Б.* Определение параметров облачного покрова и осадков по данным МСУ-МР с полярно-орбитального метеоспутника «Метеор-М» №2 для территории Западной Сибири // Сборник тезисов Шестнадцатой Всероссийской открытой конференции «Современные проблемы дистанционного зондирования Земли из космоса». Москва, ИКИ РАН, 12-16 ноября 2018 г. С. 166. ISBN 978-5-00015-018-4. DOI: 10.21046/2070-16DZZconf-2018a.

[3] *Волкова Е.В., Музылев Е.Л., Старцева З.П.* Определение температуры подстилающей поверхности по данным МСУ-МР/Метеор-М №2 на примере Центрально-Черноземного региона ЕТР // Сборник тезисов Международного Симпозиума «Атмосферная радиация и динамика» (МСАРД-2017), 27-30 июня 2017 г. Санкт-Петербург. С. 47-49.

[4] *Волкова Е.В.* Сравнение двух методик определения параметров облачного покрова и осадков по данным наблюдений с полярно-орбитальных метеоспутников для Европейской территории России // Сборник тезисов Шестнадцатой Всероссийской открытой конференции «Современные проблемы дистанционного зондирования Земли из космоса». Москва, ИКИ РАН, 12-16 ноября 2018 г. С. 165. ISBN 978-5-00015-018-4. DOI: 10.21046/2070-16DZZconf-2018a.

[5] *Волкова Е.В.* Оценки суточных сумм осадков по данным МСУ-МР с полярно-орбитального КА Метеор-М №2 для Центрально-Черноземного региона Европейской территории России» // Сборник тезисов Международного Симпозиума «Атмосферная радиация и динамика» (МСАРД-2017), 27-30 июня 2017 г. Санкт-Петербург. С.43-45.

[6] *Косторная А.А., Захватов М.Г., Токарева Ю.В.* Результаты испытания методики дешифрирования облачного покрова по спутниковым данным // Информационный сборник № 44. Результаты испытания новых и усовершенствованных технологий, моделей и методов гидрометеорологических прогнозов / под ред. канд. геогр. наук А.А. Алексеевой. М.; Обнинск: ИГ-СОЦИН, 2017. С. 158-170.

[7] *Косторная А.А., Рублев И.В., Беляев Д.Ю., Соловьева И.А.* Методика попиксельной автоматической классификации облачного покрова и восстановления его количественных характеристик по данным измерений радиометра AVHRR // Сборник докладов Всероссийской молодежной научно-практической конференции «Орбита молодежи» и перспективы развития Российской космонавтики». 2017. ISBN 978-5-4387-0776-9. С. 27-28.

[8] *Косторная А.А., Сапрыкин Е.И., Захватов М.Г., Токарева Ю.В.* Метод дешифрирования облачного покрова по спутниковым данным // Метеорология и гидрология. 2017. № 12. С. 16-24.

# ОСОБЕННОСТИ ВАРИАЦИЙ ТЕМПЕРАТУРЫ И ОЗОНА В АТМОСФЕРЕ ВО ВРЕМЯ СЕЙСМОСОБЫТИЙ НА БЛИЖНЕМ ВОСТОКЕ

*Кашкин В.Б., Рублева Т.В., Одинцов Р.В.*
Сибирский федеральный университет, Красноярск

На основе спутниковых данных были изучены атмосферные возмущения над афтершоковыми полями сильных землетрясений с M=7.3 и M=6.3. Обнаружена отрицательная корреляция рядов температуры в тропосфере и нижней стратосфере в ноябре-декабре при подготовке сейсмособытий в Ираке и Иране. В ноябре 2017 г. при потеплении тропосферы ОСО уменьшалось, а при похолодании нижней стратосферы – ОСО возрастало. В ноябре 2018 г. рост температуры в тропосфере и ОСО в нижней стратосфере происходит синхронно.
*Ключевые слова: спутниковые данные, сейсмические процессы, землетрясения, геоиндикаторы, атмосфера, аномалии, температура, озон*

**Введение.** При сейсмических процессах регистрируются различные геофизические проявления в атмосфере Земли над сейсмически опасными районами [1, 2]. В частности, во время сейсмических событий в атмосферных слоях происходит изменение метеорологических характеристик и резкое увеличение содержания газов, например, приземного озона [3]. Сейсмические события сопровождаются подвижками земной коры, которые возбуждают внутренние гравитационные волны в атмосфере [4]. Атмосфера является своеобразной «антенной», регистрирующей эти волны [5]. Изменяется температурное поле в тропосфере и нижней стратосфере, что может быть зарегистрировано с помощью сенсоров космического мониторинга. Реакцию атмосферы в виде различных аномалий можно рассматривать в качестве геоиндикаторов сейсмической активности.

Данное исследование посвящено изучению атмосферных возмущений над сейсмически опасными приграничными районами Ирака-Ирана, возникших при сильных коровых землетрясениях в конце 2017 и 2018 годов. Наиболее сильные из них были зарегистрированы 12 ноября 2017 г. в 18:18:17 UTC магнитудой M=7.3 и 25 ноября 2018 г. в 16:37:32 UTC (M=6.3) на северо-восточной границе Ирака-Ирана. Эпицентр первого землетрясения с координатами 34.911° с.ш. и 45.959° в.д. находился в 29 км к югу от Иракского г. Халабаджана. Эпицентр второго землетрясения был расположен в 15 км от Иранского г. Саре-Роле-Захаб, его координаты – 34.361° с.ш. и 45.744° в.д.

По данным [6], в ноябре-декабре 2017 г. в очаговой области Иракского землетрясения было зарегистрировано 56 сейсмособытий с M≥4.0 и глубиной возникновения 7,6-19 км. В ноябре-декабре 2018 г. в очаге Иранского землетрясения произошло 18 сейсмособытий с M≥4.2, гипоцентры которых находилось на глубинах 10-18 км. Их афтершоковые поля пространственно накладывались друг на друга с временным сдвигом в 1 год и находились в зоне столкновения двух плит Аравийской и Евразийской.

**Исходные спутниковые данные.** Космические аппараты NOAA серии POES (Polar Orbital Environmental Satellites) с помощью аппаратуры ATOVS (Advanced TIROS Operational Vertical Sounder) позволяют получать вертикальные профили температуры и влажности в атмосфере [7]. В спектрометре OMI/AURA (Ozone Monitoring Instrument, KA AURA, USA) измеряется интенсивность уходящего солнечного УФ излучения и определяется общее содержание озона (ОСО) [8]. ОСО – это толщина слоя озона в вертикальном столбе атмосферы, выделенном при давлении 1013,25 гПа и температуре 288,15 К [3].

Спутниковые данные ATOVS и OMI имеют пространственное разрешение 1x1º. Информация по вертикальным профилям атмосферы представлена в [7], а по общему содержанию озона в [9]. Сформирован архив спутниковой информации, содержащий температурные профили на 18-ти изобарических уровнях от 1000 до 100 гПа и данные об ОСО.

**Методика исследования.** Наша методика анализа температурных профилей и вариаций озона по спутниковым данным обстоятельно описана в [3, 5]. В данной работе проводился

корреляционный анализ температурных рядов в возмущенной тропосфере и нижней стратосфере во время сейсмических процессов в приграничных районах Ирака и Ирана в исследуемый период.

По модулю максимального коэффициента корреляции были выбраны наиболее информативные изобарические поверхности. В ноябре-декабре 2017 г. максимальный коэффициент корреляции между температурными рядами $T_{300}$ на изобарическом уровне 300 гПа и $T_{100}$ на уровне 100 гПа равен $R_1$ = -0.68. Для исследования возмущенной атмосферы над афтершоковым полем землетрясения с М=7.3 выбраны изобарические поверхности в верхней тропосфере 300 гПа (высота h ~ 9 км) и в нижней стратосфере 100 гПа (h ~16 км). В целом, за исследуемый период с 1 ноября по 31 декабря в 2017 году, обнаружена отрицательная корреляция температуры в верхней тропосфере и нижней стратосфере. Это означает, что в верхней тропосфере и нижней стратосфере температура изменялась противофазно. Данный результат согласуется с выводами работы [10].

В ноябре-декабре 2018 г. максимальный коэффициент корреляции между рядами температур $T_{450}$ на уровне 450 гПа и $T_{200}$ на уровне 200 гПа составил $R_2$ =-0.62. В качестве изобарических поверхностей над афтершоковым полем сейсмособытия с М=6.3 выбраны уровни 450 гПа (h~6 км, тропосфера) и 200 гПа (h ~12 км, тропопауза).

Для выявления аномалий, возникших в тропосфере и нижней стратосфере при сильных землетрясениях над исследуемыми территориями, нами создавались цифровые карты температурных полей для возмущенных и фоновых атмосферных условий. По этим данным были построены разностные карты и отклонений изотерм в виде меридиональных разрезов. Фоновые температурные поля выбирались в тех же координатах и аналогичных изобарических уровнях, что и афтершоковые поля будущих сильных землетрясений с магнитудами 7.3 и 6.3. Для Иракского сейсмособытия использовалось поле температур за 12.11.2016 г. Для Иранского – за 25.11.2016 г.

**Вариации температур в тропосфере, тропопаузе и нижней стратосфере во время коровых землетрясений.** Для возмущенной атмосферы над афтершоковым полем Иракского землетрясения графики температурных рядов $T_{300}$ и $T_{100}$ показаны на рис. 1 с 1 ноября по 31 декабря 2017 г. Стрелками на рис. 1 отмечены некоторые даты произошеших сейсмособытий и их магнитуды. Среди них толчки 20 ноября с М=4.9, 26 ноября с М=4.6, 6 декабря с М=4.8, 11 декабря с М=5.4, 20 декабря с М=4.7 и 26 декабря с М=4.5 [6]. Обнаружено, что во время подготовки данных событий происходило понижение температуры на изобарическом уровне 300 гПа и ее повышение на 100 гПа.

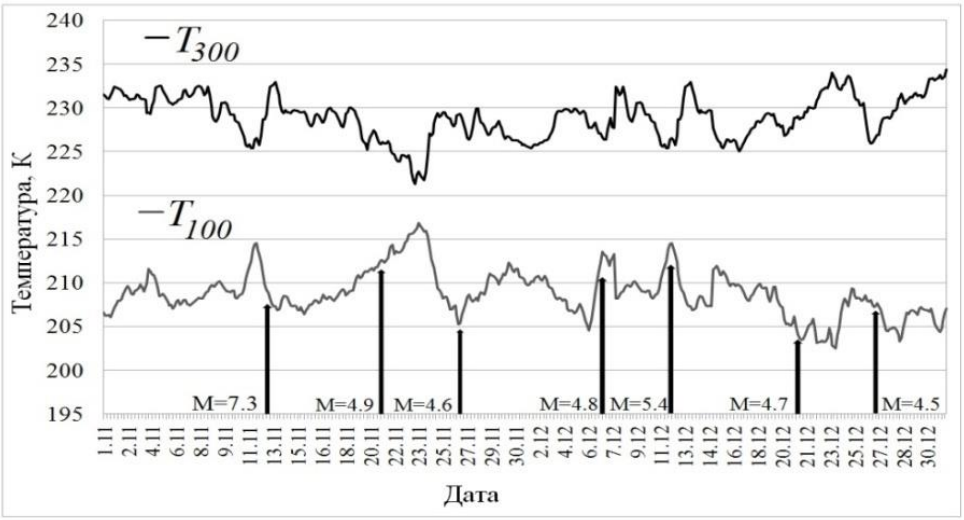



Рис. 1. Температурные профили на изобарических уровнях 100 и 300 гПа в ноябре-декабре 2017 г. над эпицентральной областью землетрясения М=7.3.

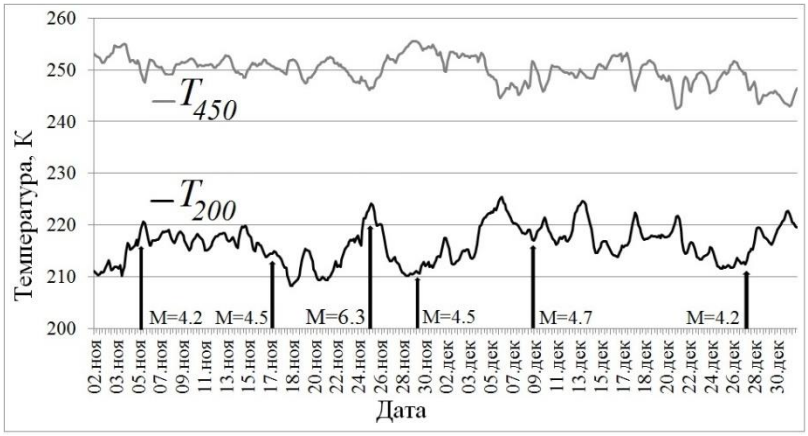



Рис. 2. Температурные профили на изобарических уровнях 450 и 200 гПа в ноябре-декабре 2018 г. над эпицентральной областью землетрясения M=6.3

Графики температурных рядов $T_{450}$ и $T_{200}$ для ноября-декабря 2018 г. в условиях возмущенной атмосферы над афтершоковым полем Иранского землетрясения приведены на рис. 2. Стрелками здесь обозначены отдельные даты сейсмособытий и их магнитуды, а именно 5 ноября с M=4.2, 16 ноября с M=4.5, 8 декабря с M=4.7 и 27 декабря с M=4.2. Во время подготовки указанных сейсмособытий обнаружено понижение температуры на изобарическом уровне 450 гПа и повышение $T$ на 200 гПа, т.е. при сейсмических процессах в конце 2018 г. над афтершоковым полем Иранского землетрясения в тропосфере температура возрастала, а в этот период в тропопаузе она уменьшалась.

**Температурные аномалии в тропосфере, тропопаузе и нижней стратосфере**. Получены цифровые карты разности температурных полей и соответствующие ей изотермы. На цифровых картах, приведенных на рис. 3 и 4, можно увидеть существующие в этот момент температурные аномалии. Шкала отклонений температур $\Delta T$ на всех рисунках показана справой стороны. Эпицентры сильных землетрясений показаны жирными точками, их очаговые области – сплошной линией, а зоны подготовки – пунктирной линией. Температурное поле на рис. 3 показано в границах 25-45° с.ш. и 34-60° в.д. Температурное поле на рис. 4 имеет границы 20-44° с.ш. и 35-55° в.д.

На рис. 3, *а* шкала отклонений температур обозначена как $\Delta T_{1\text{-}2}$, где $\Delta T_{1\text{-}2}=T_1 - T_2$. Здесь $T_1$ – значения температуры в возмущенных атмосферных условиях за 12.11.2017 г. на изобарическом уровне 300 гПа , а $T_2$ – температура на этом же уровне для фонового режима за 12.11.2016 г. Аномальная область с отрицательными значениями температур $\Delta T_{1\text{-}2}$ показана на рис. 3,*а*. Она ориентирована в широтном направлении на запад-восток. Аномалия имеет размеры 1140x2553 км.

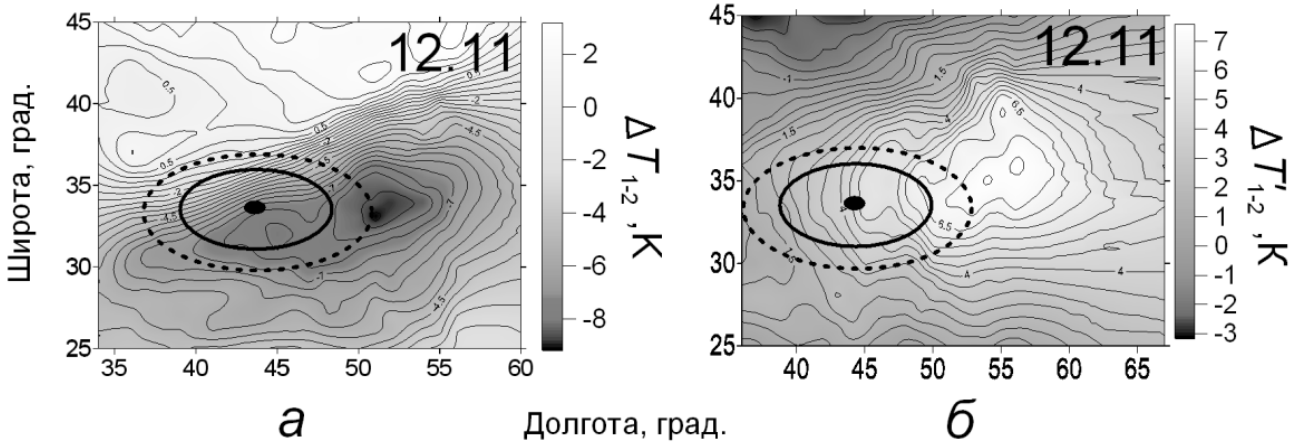



Рис. 3. Карты разности температурных полей и соответствующие ей изотермы за 12.11.2017 г. на уровнях 300 гПа (*а*) и 100 гПа (*б*).

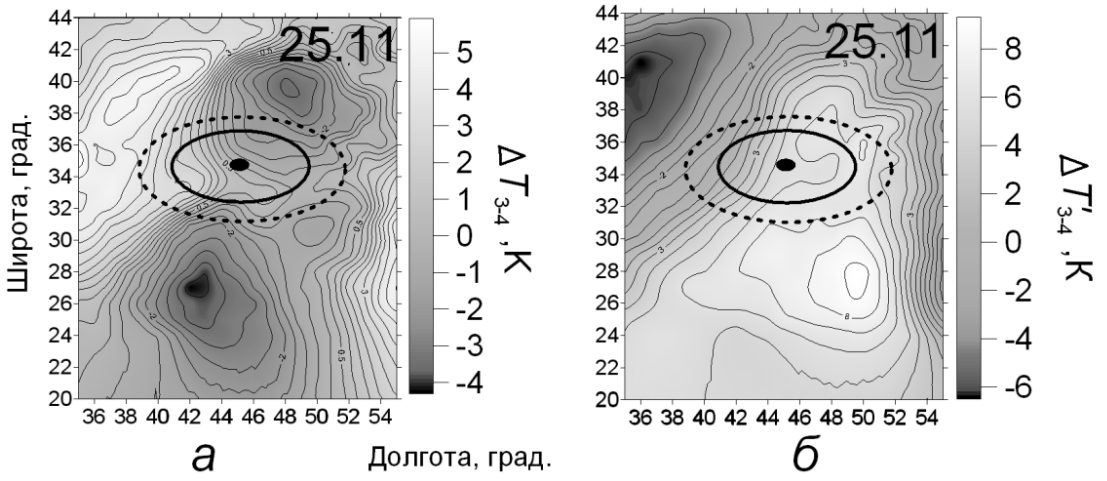



Рис. 4. Карты разности температурных полей и соответствующие ей изотермы за 25.11.2018 г. на уровнях 450 гПа (*а*) и 200 гПа (*б*).

На рис. 4,*а* видно, что эпицентр Иранского землетрясения, его очаг и область подготовки находятся между двумя тропосферными аномалиями с отрицательными значениями температур $\Delta T_{3\text{-}4}$. Ориентация этих аномалий меридиональная в направлении ЮЗ-СВ. По оценкам, размеры первой аномалии, расположенной в северной части цифровой карты, составляют 512х1110 км. Вторая аномалия находится в южной части карты в границах 1200х1110 км.

На рис. 4,*б* шкала отклонений температур $\Delta T'_{3\text{-}4}$ задавалась как разность $T'_3 - T'_4$, где $T'_3$ – значения температуры в возмущенных атмосферных условиях за 25.11.2018 г. на изобарическом уровне 200 гПа, а $T'_4$ – температура на этом же уровне для фонового режима за 25.11.2016 г. На рис. 4, *б* видно, что эпицентр Иранского землетрясения, его очаг и область подготовки в зоне наибольших искажений изолиний температурного поля. Аномальная область на уровне тропопаузы ориентирована также по меридиану в направлении с юго-запада на северо-восток. Значения $\Delta T'_{3\text{-}4}$ в аномалии положительные (рис. 4,*б*). Ее размеры составляют 1900х1332 км.

**Вариации озона и температуры во время сильных землетрясений.** Из температурного ряда $T_{300}$ были выбраны значения температуры для момента времени $18^h$, т.е. за 18 минут до основного толчка M=7.3, для периода с 5 по 27 ноября 2017 г. Для этого периода, по данным OMI/AURA, были получены значения общего содержания озона над эпицентральной областью. Графики рядов ОСО и температуры $T'_{300}$ на уровне 300 гПа приведены на рис. 5. Коэффициент корреляции между рядами ОСО и $T'_{300}$ равен -0.74. В ноябре 2017 г. в период подготовки сейсмособытий афтершокового поля Иракского землетрясения обнаружено, что при повышении температуры в верхней тропосфере происходит уменьшение озона в нижней стратосфере.

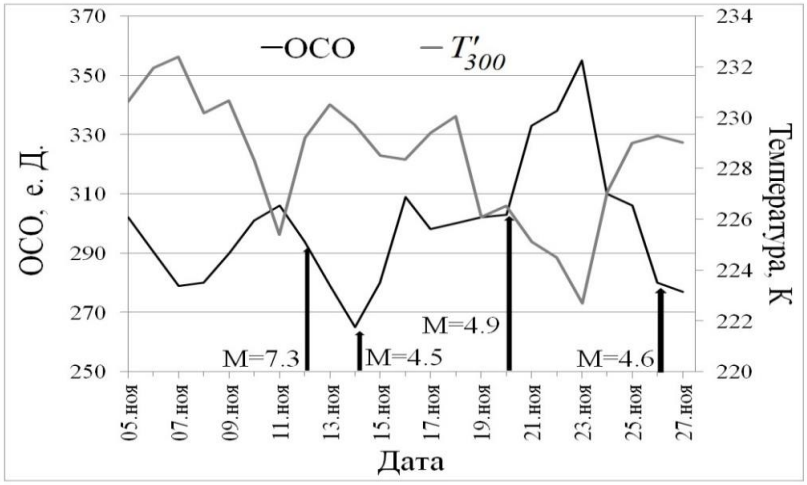



Рис. 5. Ряды температуры $T'_{300}$ на уровне 300 гПа и ОСО в ноябре 2017 г.

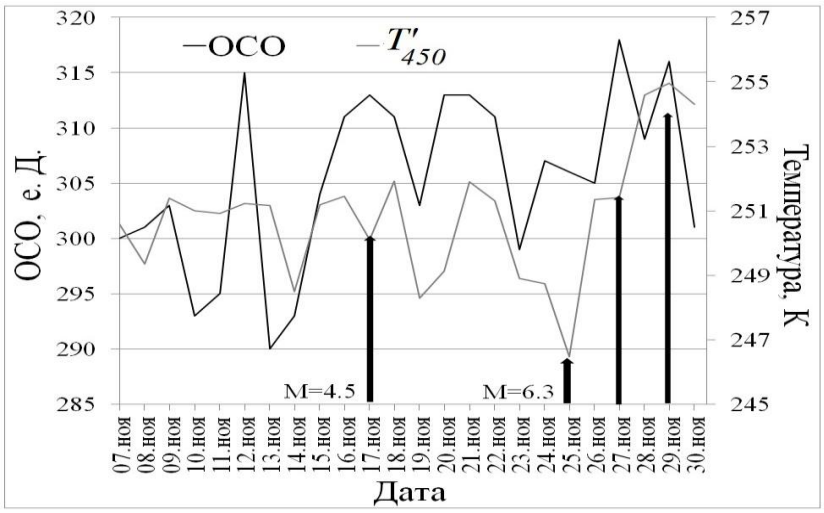



Рис. 6. Вариации ОСО и температуры $T'_{450}$ на уровне 450 гПа в ноябре 2018 г.

Аналогичным образом из температурного ряда $T_{450}$ были выбраны значения температуры для $18^h$ в ноябре 2018 г. По озоновым данным OMI/AURA построены вариации ОСО за данный период (рис. 6). Здесь же показан график температурных профилей $T'_{450}$ на изобарическом уровне 450 гПа. Коэффициент корреляции между рядами ОСО и $T'_{450}$ положительный и равен 0.26. Следовательно, в ноябре 2018 г. при подготовке землетрясений сейсмособытий афтершокового поля Иранского землетрясения рост температуры в тропосфере и общего содержание озона в нижней стратосфере происходит синхронно.

Из рядов температуры $T_{100}$ на уровне 100 гПа и $T_{200}$ на уровне 200 гПа также были выбраны значения температуры для $18^h$. Графики рядов $T'_{100}$ для ноября 2017 г. и $T'_{200}$ для ноября 2018 приведены на рис. 7 и 8, соответственно. Здесь показаны вариации значений ОСО в ноябре 2017 г. (рис. 7) и 2018 г. (рис. 8).

Коэффициент корреляции между рядами ОСО и $T'_{100}$ равен 0.84. С увеличением температуры в нижней стратосфере возрастет и общее содержание озона во время подготовки сейсмособытий в афтершоковом поле Иракского землетрясения.

Коэффициент корреляции между рядами ОСО и $T'_{200}$ составляет -0.56. При подготовке сейсмособытий в афтершоковом поле Иранского землетрясения в ноябре 2018 г. температура на уровне тропопаузы и общее содержание озона в нижней стратосфере изменяются противофазно.

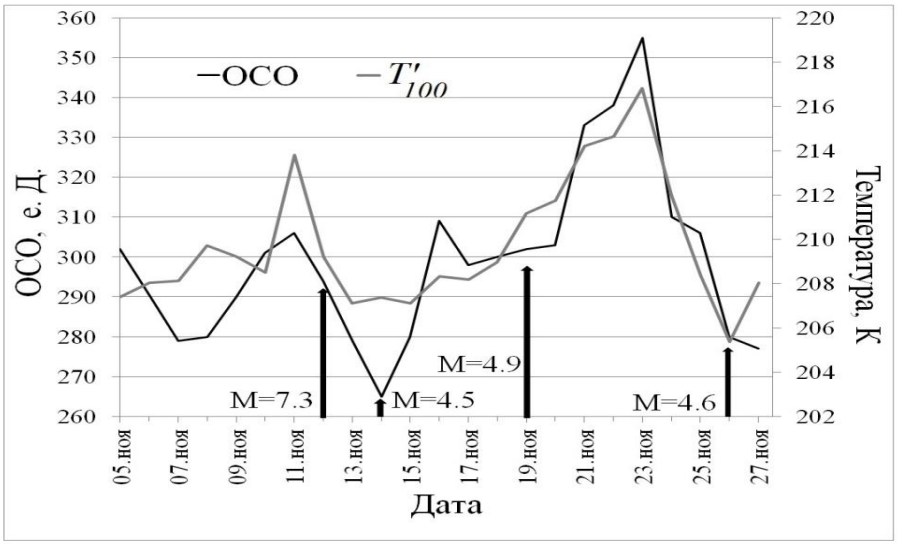



Рис. 7. Ряды температуры $T'_{100}$ на уровне 100 гПа и ОСО в ноябре 2017 г.

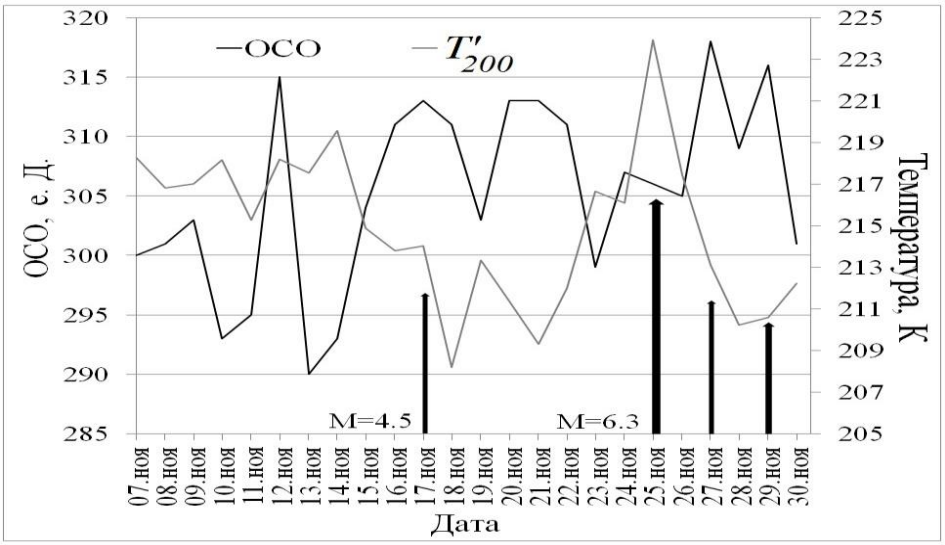



Рис. 8. Вариации общего содержания озона и температуры на уровне 200 гПа в ноябре 2018 г.

**Заключение.** Изучены атмосферные возмущения на основе спутниковых данных над афтершоковыми полями сильных землетрясений с магнитудами 7.3 и 6.3. Сформирован архив спутниковой информации, содержащий значения температур для 18-ти изобарических уровней от 1000 до 100 гПа и данные ОСО. Исследованы возмущения температурных профилей в тропосфере, на уровне тропопаузы и нижней стратосфере над сейсмоактивными районами Ирака-Ирана, где были зарегистрированы коровые землетрясения. Обнаружена отрицательная корреляция рядов температуры в тропосфере и нижней стратосфере в ноябре-декабре 2017 и 2018 годов над афтершоковыми полями Иракского и Иранского землетрясений. Выявлено, что в ноябре 2017 г. потепление тропосферы сопровождалось уменьшением общего содержания озона, а похолодание нижней стратосферы - увеличением общего содержания озона. В ноябре 2018 г. рост температуры в тропосфере и ОСО в нижней стратосфере происходит синхронно. Однако, температура на уровне тропопаузы и общее содержание озона в этот период изменяются противофазно. Вероятно, это связано с нестационарностью исследуемых процессов.

Вариации температуры и общего содержания озона могут быть геоиндикаторами готовящегося сильного землетрясения.

## ЛИТЕРАТУРА


[1] *Гуфельд И.Л., Матвеева М.И., Новоселов О.Н.* Почему мы не можем прогнозировать сильные коровые землетрясения? // Геодинамика и тектонофизика. 2011. Т. 2, № 4. С. 378-415. DOI: 10.5800/GT-2011-2-4-005.

[2] *Пулинец С.А., Узунов Д.П., Карелин А.В., Боярчук К.А., Тертышников А.В., Юдин И.А.* Единая концепция обнаружения признаков подготовки сильного землетрясения в комплексной системе литосфера-атмосфера-ионосфера-магнитосфера // Гелиофизические исследования. 2013. Т. 6. С. 81-90.

[3] *Кашкин В.Б., Рублева Т.В., Хлебопрос Р.Г.* Стратосферный озон: вид с космической орбиты. Красноярск: Сиб. федер. ун-т, 2015. 184 с.

[4] *Кашкин* В.Б. Внутренние гравитационные волны в тропосфере // Оптика атмосферы и океана. 2013. Т. 26. № 10. С. 908-916.

[5] *Kashkin V. B., Rubleva T.V., Odintsov R.V.* Abnormal Geophysical Events in the Northern Hemisphere in 2010 and 2011 // Journal of Siberian Federal University. Engineering and Technologies. 2018. Vol. 11, N 8. P. 982-988.

[6] United States Geological Survey (USGS). [Электронный ресурс]: URL: **http://www.usgs.gov** (дата обращения 25.01.2019).

[7] NOAA. [Электронный ресурс]: URL: **http://www.arl.noaa.gov/ready/cmet.html** (дата обращения 11.01.2019).

[8] Ozone Monitoring Instrument (OMI) Data User's Guide. [Электронный ресурс]: URL: **https://docserver.gesdisc.eosdis.nasa.gov/public/project/OMI/README.OMI_DUG.pdf** (дата обращения 15.02.2019)
[9] NASA. [Электронный ресурс]: URL: **ftp://toms.gsfc.nasa.gov** (дата обращения 28.02.2019).
[10] *Зуев В.Е., Комаров В.С.* Статистические модели температуры и газовых компонент атмосферы Ленинград: Гидрометеоиздат, 1986. Т. 1. 264 с.

# ОБНАРУЖЕНИЕ МАЛОРАЗМЕРНЫХ ОБЪЕКТОВ В ДВУХЦВЕТНЫХ ИЗОБРАЖЕНИЯХ С ПРОСТРАНСТВЕННО-НЕСТАЦИОНАРНЫМ ФОНОМ

*Косых В.П.[(1)(2)], Громилин Г.И.[(1)], Яковенко Н.С.[(1)]*
[(1)] Институт автоматики и электрометрии СО РАН, Новосибирск
[(2)] Новосибирский государственный университет, Новосибирск

Рассматривается способ повышения надежности обнаружения слабоконтрастных малоразмерных объектов в двухцветных изображениях посредством предварительного подавления в них пространственно-нестационарного фона. Подавление осуществляется за счет построения локально-стационарной модели фона с помощью оптимального линейного прогноза. Показано, как совместная обработка пары изображений, полученных в разных спектральных диапазонах, сказывается на вероятности обнаружения при заданной вероятности ложной тревоги.
*Ключевые слова: Обнаружение слабоконтрастных малоразмерных объектов, пространственно-нестационарный фон, двухцветные изображения, оптимальный линейный прогноз.*

**Введение.** Одним из часто формулируемых требований в задаче раннего обнаружения потенциально опасных объектов и процессов на поверхности Земли по изображениям, получаемым аппаратурой, размещенной на различных носителях, является необходимость выделения в изображениях регистрируемого параметра (яркости, интенсивности цвета, длины волны и пр.) слабых аномалий, сопоставимых по размеру с размером функции рассеяния точки (ФРТ) регистрирующей оптической системы. Если фоновая составляющая изображения пространственно-стационарна, задача поиска таких аномалий может быть решена посредством согласованной фильтрации изображений, учитывающей корреляционные свойства фона и форму пятна рассеяния оптической системы. Однако если фоновая компонента изображения не подчиняется пространственно-стационарной модели, согласованная фильтрация не обеспечит эффективного выделения аномалий. В этом случае при обнаружении малоразмерных слабоконтрастных аномалий применяется предварительное подавление пространственно-нестационарной составляющей фона. Если справедливо представление анализируемого изображения в виде

$$D(i,j) = F(i,j) + A \cdot O(i - i_0, j - j_0) + \xi(i,j), \quad (1)$$

где $F(i,j)$ – фоновая составляющая изображения, $O(i - i_0, j - j_0)$ – изображение объекта с центром, расположенным в точке $(i_0, j_0)$ и с амплитудой $A$, $\xi(i,j)$ – случайный некоррелированный шум регистрации, то при наличии модели фона $\hat{F}(i,j)$ поиск объектов целесообразно проводить в модифицированном изображении

$$\tilde{D}(i,j) = F(i,j) + A \cdot O(i - i_0, j - j_0) + \xi(i,j) - \hat{F}(i,j) = A \cdot O(i - i_0, j - j_0) + \theta(i,j), \quad (2)$$

где $\theta(i,j) = F(i,j) - \hat{F}(i,j) + \xi(i,j)$ – помеха, которую составляет остаточная часть фона с существенно ослабленной при удачно подобранной модели пространственной корреляцией и шумовая составляющая исходного изображения.

Одним из весьма эффективных средств подавления является межкадровая обработка, когда модель фоновой составляющей анализируемого кадра оценивается по ранее полученным изображениям того же участка местности [1-3]. Если же предшествующие изображения недоступны, подавление основывается на предсказании модели фона по фрагменту текущего изображения в окрестности предполагаемого присутствия аномалии. Исследованию возможности построения модели с применением оптимального линейного прогноза (ОЛП) [4] и применению этого подхода при анализе пары изображений одной сцены, полученных в разных спектральных диапазонах, посвящена данная работа.

**Построение модели фона с помощью ОЛП.** Основанной на оптимальном линейном прогнозе локально-стационарной моделью фона в точке $(i', j')$ является линейная комбинация отсчетов изображения в окрестности $W$ этой точки

$$F(i',j') = \sum_{i,j\in W} D(i-i', j-j')\, h(i,j)\,. \quad (3)$$

Здесь $h(i,j)$ – набор весовых коэффициентов, которые определяются как

$$\hat{\mathbf{h}} = argmin\left\{\sum_{i',j'\in\Omega}\left[D(i',j') - \mathbf{h}^T\mathbf{d}_{i',j'}\right]^2\right\}, \quad (4)$$

где $\mathbf{d}_{i',j'}$ – вектор, состоящий из лексикографически упорядоченных отсчетов окрестности $W$ точки $(i',j')$, $\mathbf{h}$ – вектор аналогичным образом упорядоченных весовых коэффициентов, $\Omega$ – область стационарности фона. Для исключения влияния объекта, который может оказаться в точке $(i',j')$, сама эта точка и ее ближайшие соседи в окрестность $W$ не входят. Достаточно очевидно, что наихудшее подавление фона будет в зонах изображения, где нарушается пространственная стационарность фона, в частности, в области резких перепадов яркости. Поэтому область стационарности $\Omega$ не может быть слишком большой.

**Совместная обработка изображений.** Предлагаемый метод обнаружения малоразмерных объектов в двухцветных изображениях заключается в совместном анализе изображений двух спектральных диапазонов, предварительно модифицированных согласно выражению (2). При этом предполагается, что спектр объектов перекрывает анализируемые диапазоны. Подавление основных структурных особенностей изображений в каждом из каналов своими локальными моделями приводит к снижению межканальной корреляции фоновой составляющей, сохраняя при этом корреляцию между изображениями объектов. При совместной обработке за счет этого можно ожидать снижения вероятности ложных тревог, или, при заданной вероятности ложной тревоги – повышения вероятности обнаружения. Наиболее эффективным способом совместной обработки было бы построение общей локально-стационарной модели фона, аналогичной (4), в которой $\mathbf{d}_{i',j'}$ содержит соответствующим образом упорядоченные точки двух изображений. Проблема заключается в том, что при двукратном увеличении вектора весовых коэффициентов $\mathbf{h}$ многократно возрастают вычислительные затраты на его оценивание. Поэтому для каждого цвета строиласьь независимо своя модель фона, выделлись локальные максимумы, превышающие порог, обеспечивающий заданную вероятность ложной тревоги, а затем эти максимумы объединялись. Верхний предел вероятности ложной тревоги при таком методе обработки будет равен

$$P_{\text{лт}} = P_{1\text{лт}} + P_{2\text{лт}} - 2P_{1\text{лт}}P_{2\text{лт}} \approx P_{1\text{лт}} + P_{2\text{лт}}, \quad (5)$$

(приближенное равенство справедливо при $P_{1\text{лт}}, P_{2\text{лт}} \ll 1$, что обычно выполняется) и достигаться он будет при отсутствии корреляции между фоновыми составляющими изображений. В противном случае вероятность ложной тревоги будет ниже.

**Численный эксперимент.** Повышение надежности обнаружения подтверждается численным экспериментом, выполненным с мультиспектральными изображениями (размер 1024×1024, спектральные диапазоны: 1 – 0.45-0.515; 2 – 0.525-0.605; 3 – 0.63-0.69; 4 – 0.775-0.90; 5 – 1.55-1.75 мкм), полученными спутником Landsat 7 [5]. На пару изображений разных спектральных диапазонов в соответствии с выражением (1) дополнительно наносились малоразмерные объекты со средней амплитудой, равной 1.5 СКО фона, и некоррелированный шум с СКО, равным 0.1 СКО фона. Обнаружение объектов выполнялось после фильтрации изображений адаптивными фильтрами, реализуемыми согласно (3) и (4), пороговой обработкой, обеспечивающей заданную вероятность ложной тревоги Рлт. Первый вариант обнаружения состоял в выборе порога отдельно для каждого изображения и объединении превысивших порог точек в одно изображение. Второй заключался в выборе порога и выделении точек, превысивших порог, в сумме фильтрованных изображений. Далее оценивались вероятности обнаружения в каждом изображении (**Р1** и **Р2**), в суммарном изображении **РΣ**, и вероятность обнаружения **Рили**, полученная в результате объединения точек, выделенных в каждом изображении.

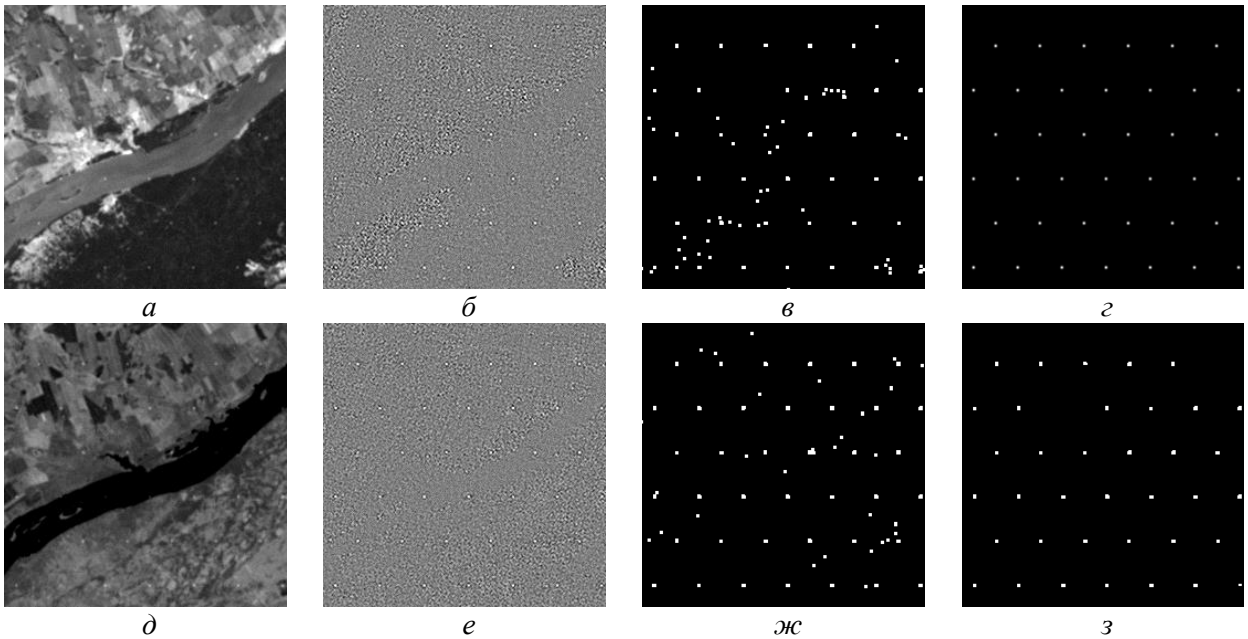
Результаты обработки для заданной вероятности ложной тревоги Рлт = 0.005 иллюстрирует приведенная ниже таблица. В колонке **R** таблицы указаны коэффициенты корреляции между изображениями в выбранных каналах, а в колонках **Кш** – отношение СКО остаточной помехи (фон + шум после фильтрации) к СКО шума в исходном изображении. Эти числа по-

казывают, что применение локально-стационарной модели позволяет снизить средний уровень фона до величины, сопоставимой с уровнем шума в исходных изображениях. В скобках в колонках **P1** и **P2** приведены вероятности обнаружения для $P_{1лт} = P_{2лт} = P_{лт}/2$, что, согласно (5), при совместной обработке обеспечивает верхний предел вероятности ложной тревоги, равный $P_{лт}$.

Из таблицы следует, во-первых, что фильтрация изображений, основанная на формировании локальной модели фона, уменьшает СКО фоновой составляющей до величины, сопоставимой с СКО шума. Во-вторых, совместная обработка изображений разных спектральных диапазонов приводит к снижению вероятности ложной тревоги. В-третьих, вероятность ложной тревоги тем меньше, чем меньше взаимная корреляция между фоновой составляющей изображений. В-четвертых, при фиксированной вероятности ложной тревоги совместная обработка увеличивает вероятность обнаружения объектов. Рисунок иллюстрирует процесс и результат совместной обработки двухцветного изображения (приведены фрагменты изображений размером 256×256 пикселей) при Рлт=0.005.

Результаты совместной обработки пар изображений разных спектральных диапазонов.

| **Каналы** | **R** | **Рлт** | **P1** | **Кш1** | **P2** | **Кш2** | **Рили** | **PΣ** | **КшΣ** |
|---|---|---|---|---|---|---|---|---|---|
| **1-2** | 0.91 | 0.005 | 0.95(0.88) | 1.69 | 0.89(0.80) | 1.81 | 0.95 | 0.96 | 1.48 |
| **1-3** | 0.92 | 0.005 | 0.95(0.88) | 1.69 | 0.92(0.85) | 1.76 | 0.95 | 0.98 | 1.48 |
| **1-4** | -0.49 | 0.005 | 0.95(0.88) | 1.69 | 0.99(0.97) | 1.56 | 1.00 | 1.00 | 1.05 |
| **1-5** | 0.45 | 0.005 | 0.95(0.88) | 1.69 | 1.00(0.99) | 1.46 | 1.00 | 0.99 | 1.28 |
| **2-3** | 0.85 | 0.005 | 0.89(0.80) | 1.81 | 0.92(0.85) | 1.76 | 0.93 | 0.94 | 1.51 |
| **2-4** | -0.22 | 0.005 | 0.89(0.80) | 1.81 | 0.99(0.97) | 1.56 | 1.00 | 1.00 | 1.16 |
| **2-5** | 0.47 | 0.005 | 0.89(0.80) | 1.81 | 1.00(0.99) | 1.46 | 0.99 | 0.99 | 1.33 |
| **3-4** | -0.42 | 0.005 | 0.92(0.85) | 1.76 | 0.99(0.97) | 1.56 | 1.00 | 1.00 | 1.06 |
| **3-5** | 0.65 | 0.005 | 0.92(0.85) | 1.76 | 1.00(0.99) | 1.46 | 0.99 | 0.99 | 1.32 |
| **4-5** | 0.25 | 0.005 | 0.99(0.97) | 1.56 | 1.00(0.99) | 1.46 | 1.00 | 1.00 | 1.04 |




Совместная обработка изображений в 3 и 4 каналах: *а,д* – фрагменты исходных изображений; *б,е* – изображения с подавленным фоном; *в,ж* – изображения точек, превысивших порог, обеспечивающий заданную вероятность ложной тревоги; *г* – нанесенные объекты; *з* – объекты, обнаруженные после совместной обработки.

**Заключение.** Предложенный способ аппроксимации пространственно-нестационарного фона локально-стационарной моделью, основанной на оптимальном линейном прогнозе, снижает уровень фона в изображениях, содержащих малоразмерные слабоконтрастные объекты, до величины, сопоставимой с уровнем шума. Последующая совместная обработка сигналов, обнаруженных в разных каналах двухцветного изображения, позволяет при фиксированной вероятности ложной тревоги повысить вероятность обнаружения объектов.

*Работа выполнена при финансовой поддержке Минобрнауки России (проект АААА-А18-118051890038-8).*

## ЛИТЕРАТУРА

[1] *Киричук В.С., Пустовских А.И.* Применение статистических методов в задаче оценивания стационарной части фона по серии изображений // Автометрия. 1988. № 3. С. 74-78.

[2] *Tartakovsky A.G., Brown A.P., Brown J.* Nonstationary EO/IR Clutter Suppression and Dim Object Tracking // Proceedings of the 2010 Advanced Maui Optical and Space Surveillance Technologies (AMOS) Conference, Maui, Hawaii, September 14-17, 2010.

[3] *Громилин Г.И., Косых В.П., Попов С.А., Стрельцов В.А.* Подавление фона с резкими перепадами яркости в последовательности изображений динамических малоразмерных объектов // Автометрия. 2019. Т. 55, № 3. С. 3-12.

[4] *Андерсен Т.* Статистический анализ временных рядов. М.: Мир, 1976. 756 с.

[5] Global Land Cover Facility. **http://glcf.umiacf.umd.edu** (дата обращения 23.09.2017).

# ОЦЕНКА ЗАГРЯЗНЕНИЯ АТМОСФЕРЫ ПО СПУТНИКОВЫМ ДАННЫМ РАЗЛИЧНОГО ПРОСТРАНСТВЕННОГО РАЗРЕШЕНИЯ

*Краснощеков К.В.[(1)(2)], Якубайлик О.Э.[(1)(2)(3)]*
[(1)] Федеральный исследовательский центр КНЦ СО РАН, Красноярск
[(2)] Институт вычислительного моделирования СО РАН, Красноярск
[(3)] Сибирский Федеральный Университет, Красноярск

На основании измерений спектрометра MODIS, установленного на спутниках TERRA и AQUA, формируются данные об аэрозольная оптическая толщина (АОТ) с различным пространственным разрешением: 10, 3, 1 км. Исследовалась взаимосвязь между значениями АОТ измеренными при помощи дистанционного зондирования и $PM_{2.5}$, измеренными на автоматизированных постах наблюдения (АПН). Показано, что данные с пространственным разрешением в 1 км позволяют увидеть зоны пылевых загрязнений внутри города. Эта информация может быть использована в качестве объективной оценки экологической обстановки.
*Ключевые слова: твердые частицы, аэрозольная оптическая толщина, MODIS, алгоритм MAIAC, дистанционное зондирование, АПН.*

**Введение.** Аэрозоли или взвешенные в воздухе твердые частицы (PM) естественного или антропогенного происхождения оказывают существенное влияние на климат, окружающую среду и здоровье человека [1]. Многочисленные эпидемиологические исследования показали, что существует связь между концентрациями PM и различными неблагоприятными последствиями для здоровья человека. [2]. Следовательно, оценка качества воздуха, особенно с точки зрения $PM_{10}$ и $PM_{2.5}$ (PM с диаметрами менее 10 и 2,5 мкм соответственно) является актуальной проблемой на данный момент. Наземные наблюдения с автоматизированных постов наблюдений (АПН) показывают важную пространственную и временную информацию о концентрации PM в атмосфере.

Мониторинг PM основан в основном на наземных измерениях. Хотя сети станций существуют в крупных городах, точечные измерения не дают информации о пространственных характеристиках и распределении PM по городским районам, представляющим интерес. Временной охват измерений PM на месте также сильно варьируется в зависимости от периода работы прибора и его функциональности. Эти причины привели к постоянным усилиям по оценке PM с использованием методов спутникового дистанционного зондирования.

Аэрозольная оптическая толщина (АОТ) является параметром, полученным со спутника, который чаще всего используется в качестве основы для оценки PM [3]. АОТ – интегрированное атмосферное рассеивание излучения аэрозолями в вертикальном столбе атмосферы. Этот параметр пропорционален количеству частиц в воздухе и, зависит от их массовой концентрации. АОТ обычно используется в качестве основы для оценки PM. Несколько методов были использованы для корреляции АОТ с дистанционным зондированием с измеренным на поверхности PM. К ним относятся линейные отношения [4], статистические и химические транспортные модели [5], множественный регрессионный анализ [6] и нейронные сети [7].

Существует несколько факторов, ограничивающих корреляцию АОТ-$PM_{2.5}$: влияние вертикального профиля аэрозоля, которое отвечает за разницу между измерениями в столбе атмосферы (АОТ) и приповерхностными ($PM_{2.5}$); сложная роль относительной влажности; скорость ветра; Распределение частиц по размерам; а также состав частиц и т.д. [7]. В работе [8] исследовалась корреляция между измерениями общего столбца АОТ и приповерхностного $PM_{2.5}$ и этими переменными. Эти исследования показали широкий диапазон корреляций между АОТ и массой $PM_{2.5}$.

В этой статье мы используем данные АОТ с разрешением 1 и 10 км, полученные для г. Красноярск, чтобы определить, становится ли связь между концентрациями $PM_{2.5}$,

измеренными на земле, и значениями АОТ более сильными, когда пространственное разрешение АОТ увеличивается.

В данном исследовании использовался линейный метод для оценки $PM_{2.5}$ над г. Красноярском, Россия, на основе спутникового АОТ. Данный параметр доступен со спектрорадиометра (MODIS) на борту спутников NASA Terra и Aqua.

**Материалы и методы.** На момент июля 2018 года в г. Красноярск осуществляло мониторинг $PM_{2.5}$ 3 автоматизированных поста наблюдения (АПН) региональной экологической системы. На рис. 1 показано расположение АПН, данные с которых были использованы в нашем исследовании.

Для измерения концентрации $PM_{2.5}$ на АПН используется радиоизотопный принцип действия, который является общепринятым во всем мире. Он основан на поглощении β-излучения частицами пыли, осажденными на фильтрующую ленту. В качестве источника β-излучения используется изотоп С14. Пыль осаждается на фильтрующей ленте в результате прокачивания пробы воздуха насосом. Измерение величины поглощения излучения осуществляется с помощью встроенного в анализатор детектора – счетчика Гейгера – Мюллера. Для оценки количества загрязнения воздуха мы использовали среднесуточные концентрации.

Алгоритм MAIAC [9] был разработан для обработки данных MODIS. MAIAC извлекает параметры аэрозоля над землей с разрешением 1 км. Продукт MCD19A2 (MAIAC) содержит данные спекторофотометра MODIS установленного на спутниках Terra и Aqua. Этот продукт был опубликован 30 мая 2018 г, и содержит данные по АОТ с 1 февраля 2000 г. [10]. Параметры аэрозоля включают оптическую глубину на длине волны от 0,47 до 0,67 мкм и тип аэрозоля, включая модели фона, дыма и пыли [11].

В нашем исследовании мы использовали данные АОТ на длине волны 0.47 мкм.

В работе [12] была проведена корреляция между наземными измерениями $PM_{2.5}$ и спутниковыми измерениями АОТ на различных длинах волн. Его исследование показало, что корреляция между $PM_{2.5}$ больше для длины волны 0.47 мкм.

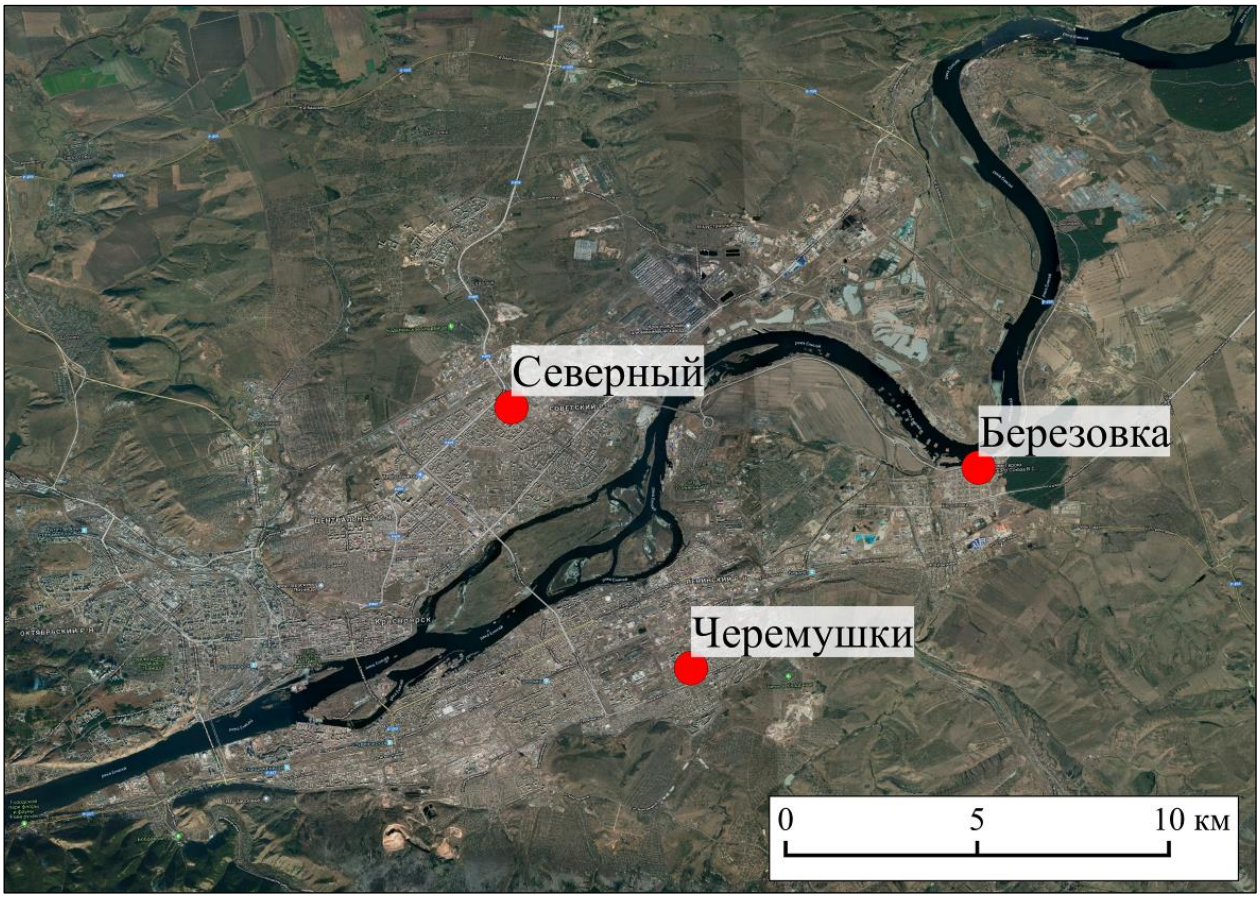



Рис. 1. Расположение используемых АПН в г. Красноярск.

Повышенная точность MAIAC является результатом использования метода явной характеристики поверхности в отличие от эмпирического подхода к параметризации поверхности, который используется в алгоритмах MOD04 и MYD04. Кроме того, MAIAC включает алгоритм облачной маски, основанный на пространственно-временном анализе, который дополняет традиционные методы обнаружения облаков на уровне пикселей [13]. MAIAC обеспечивает равномерное сеточное разрешение 1 км в выбранной проекции независимо от угла сканирования.

В дополнение к данным MAIAC мы использовали ежедневные аэрозольные данные MODIS Level 2, Collection 6.1 со спутников Aqua и Terra, которые получены с пространственным разрешением 10×10 км$^2$ (в надире). Аэрозольные продукты MYD04 и MOD04 получены на основе спектрального излучения, измеренного с помощью MODIS с использованием семи спектральных каналов в диапазоне длин волн от 470 до 2130 нм [14].

Дополнительные длины волн в других частях спектра используются для идентификации и маскировки облаков, снега и взвешенных речных отложений [15].

В нашем исследовании мы использовали измерения $PM_{2.5}$ с АПН и измерения АОТ за июль 2018 г. Мы изучали связь между измерениями АОТ и $PM_{2.5}$ в масштабе г. Красноярск. Частота измерений АПН составляет 1 измерение в 20 минут. Мы использовали усреднённые значения $PM_{2.5}$ за сутки. Для корреляции на уровне города между данными АОТ и $PM_{2.5}$ было доступно 10 дней, всего 30 пар. Дни выбирались с учётом отсутствия облачности.

**Результаты и обсуждения.** На рис. 2 показано высокое разрешение MAIAC 1 км (слева) и низкое разрешение, равное 10 км (справа). Данные высокого пространственного разрешения показывают пространственную изменчивость АОТ как при умеренных, так и при низких уровнях загрязнения, которую не могут дать данные более низкого пространственного разрешения.

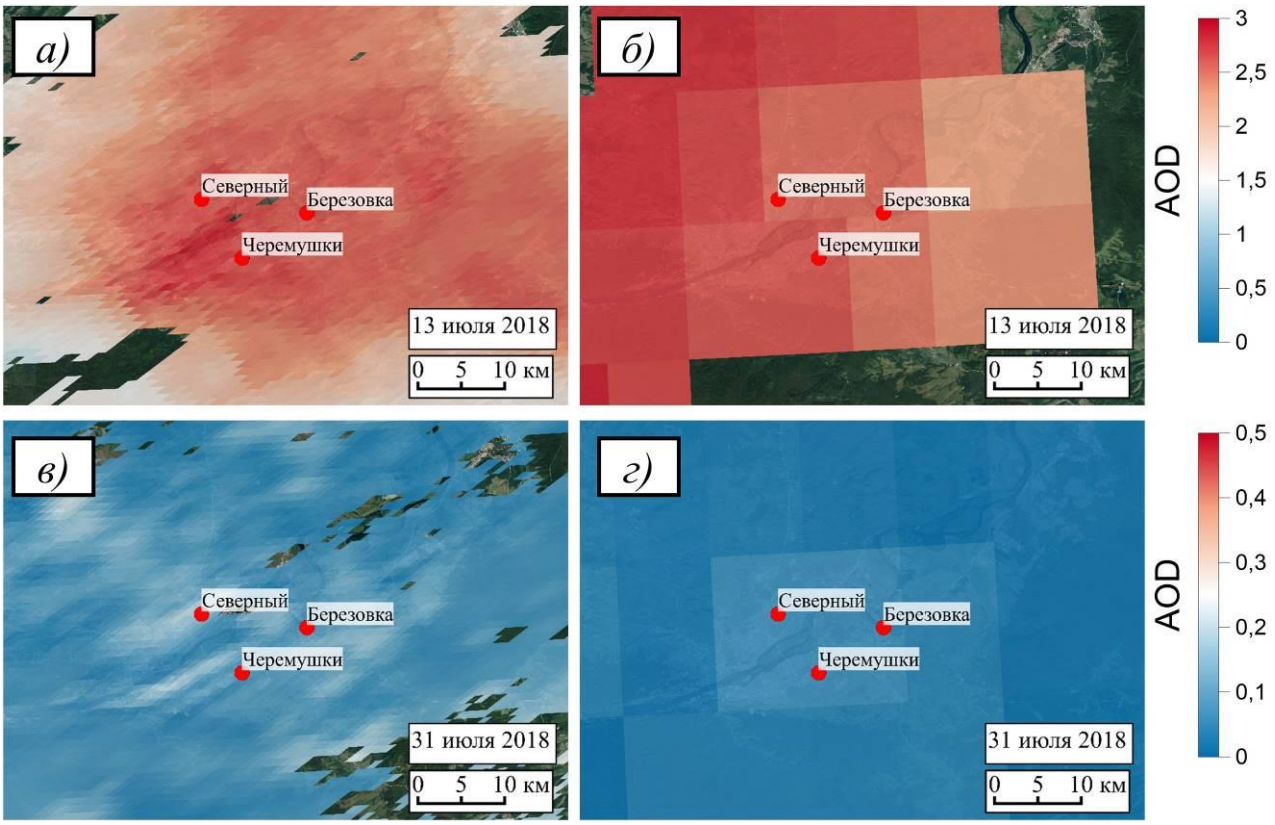



Рис. 2. Значение АОТ над г. Красноярск. Данные MIAIAC 1 км под (*а* и *в*), данные 10 км пространственного разрешения (*б* и *г*), на которых показаны дни с низким (*а* и *б*) и умеренным (*в* и *г*) уровнем загрязнения.

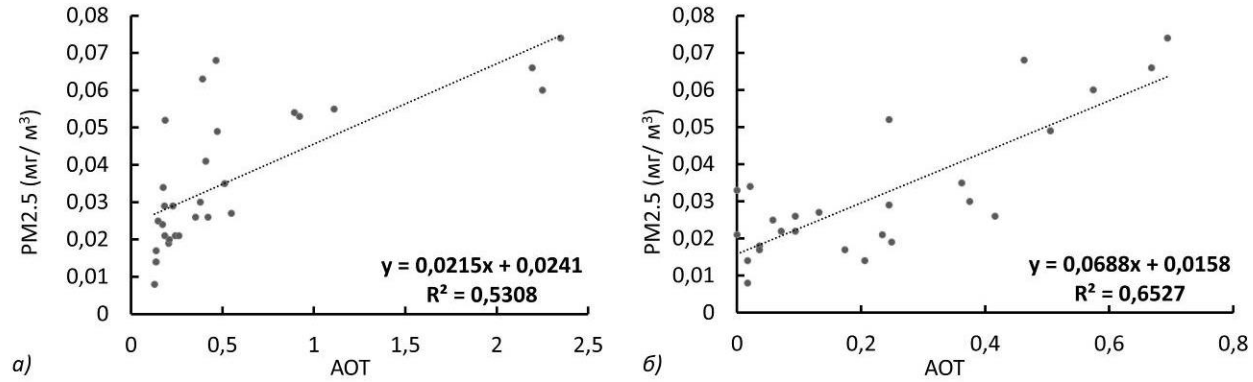



Рис. 3. Корреляция данных AOT с 1 км разрешением (*а*) и 10 км разрешением (*б*) с наземными измерениями $PM_{2.5}$.

На рис. 3 показана взаимосвязь между результатами определение AOT в один из дней июля 2018 года по алгоритму MAIAC с пространственным разрешением 1 км и среднесуточным значением $PM_{2.5}$ измеренным на 3 АПН в г. Красноярск. Коэффициент детерминации в этом случае составляет 0,53 ($R^2$=0.53). Следует отметить, что для разных спутниковых изображений в течение периода исследования коэффициент детерминации варьировался в диапазоне от 0,5 до 0,8, что указывает на хорошую связь с наземными измерениями PM2,5. Значения корреляции для алгоритма разрешением в 10 км были сопоставимы.

**Заключение.** В этой статье мы использовали новый алгоритм MAIAC для оценки AOD по данным MODIS с пространственным разрешением 1 км, сравнивая его с классическим алгоритмом с более грубым пространственным разрешением 10 км. Наш анализ показывает, что корреляция между $PM_{2.5}$ и AOT с пространственным разрешением в 1 и 10 км приблизительно схожи. Однако используя более высокое пространственное разрешение можно выделить области пылевых загрязнений в городе. Это позволит более качественно определять экологически неблагоприятные районы города. Используя данные АПН, в дополнении со спутниковыми данными высокого пространственного разрешениями (MAIAC), можно создать информационную основу для современной системы мониторинга окружающей среды в региональном масштабе и способствовать улучшению экологической ситуации в городе.

Важным недостатком рассматриваемой методики является невозможность получения данных AOT в зимний период и когда облака не позволяют видеть поверхность Земли со спутника.

## ЛИТЕРАТУРА


[1] *Kaufman Y. J., Tanré D., Boucher O.* A satellite view of aerosols in the climate system // Nature. 2002. Т. 419, №. 6903. С. 215-223

[2] *Boldo E. et al*. Apheis: Health impact assessment of long-term exposure to PM 2.5 in 23 European cities // European journal of epidemiology. 2006. Т. 21. №. 6. С. 449-458.

[3] *Van de Kassteele J. et al*. Statistical mapping of PM10 concentrations over Western Europe using secondary information from dispersion modeling and MODIS satellite observations // Stochastic environmental research and risk assessment. 2006. Т. 21. №. 2. С. 183194.

[4] *Yap X.Q., Hashim M*. A robust calibration approach for PM 10 prediction from MODIS aerosol optical depth // Atmospheric Chemistry & Physics Discussions. 2012. Т. 12. №. 12. С. 3148331505.

[5] *Kloog I. et al*. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM2. 5 exposures in the MidAtlantic states // Environmental science & technology. 2012. Т. 46. №. 21. С. 1191311921.

[6] *Gupta P., Christopher S.A.* Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach // Journal of Geophysical Research: Atmospheres. 2009. Т. 114. №. D14.

[7] *Wang J., Christopher S.A.* Intercomparison between satellite-derived aerosol optical thickness and PM2. 5 mass: Implications for air quality studies //Geophysical research letters. 2003. Т. 30. №. 21.
[8] *Hoff R. M., Christopher S. A.* Remote sensing of particulate pollution from space: have we reached the promised land? // Journal of the Air & Waste Management Association. 2009. Т. 59. №. 6. С. 645675.
[9] *Lyapustin A. et al.* MODIS Collection 6 MAIAC algorithm // Atmospheric Measurement Techniques. 2018. Т. 11. №. 10. С. 57415765.
[10] *Lyapustin A., Wang Y.* Release of MODIS Version 6 MAIAC Data Products. **https://lpdaac.usgs.gov/news/releaseofmodisversion6maiacdataproducts** (дата обращения 1.06.2019).
[11] *Lyapustin A. et al*. Corrigendum to «Discrimination of biomass burning smoke and clouds in MAIAC algorithm» published in Atmos. Chem. Phys., 12, 9679-9686, 2012 // Atmospheric Chemistry and Physics. 2012. Т. 12. №. 21. С. 1063110631.
[12] *Tian J., Chen D.* Spectral, spatial, and temporal sensitivity of correlating MODIS aerosol optical depth with groundbased fine particulate matter (PM2. 5) across southern Ontario // Canadian Journal of Remote Sensing. 2010. Т. 36. №. 2. С. 119128.
[13] *Lyapustin A., Wang Y., Frey R.* An automatic cloud mask algorithm based on time series of MODIS measurements // Journal of Geophysical Research: Atmospheres. 2008. Т. 113. №. D16.
[14] *Remer L.A. et al*. The MODIS aerosol algorithm, products, and validation // Journal of the atmospheric sciences. 2005. Т. 62. №. 4. С. 947973.

# АСТРОКЛИМАТИЧЕСКИЕ УСЛОВИЯ РЕСПУБЛИКИ АЛТАЙ ПО ДАННЫМ СПУТНИКОВЫХ НАБЛЮДЕНИЙ

*Мордвин Е.Ю.[(1)], Лагутин А.А.[(1)(2)], Волков Н.В.[(1)(2)], Райкин Р.И.[(1)], Ревякин А.И.[(1)], Клёпов М.А.[(1)]*

[(1)] Алтайский государственный университет, Барнаул

[(2)] Институт вычислительных технологий СО РАН, Новосибирск

Представлены результаты исследований астроклиматических условий для территории Республики Алтай для проведения наземных наблюдений инициируемых космическими лучами сверхвысоких энергий широких атмосферных ливней с использованием комплекса атмосферных черенковских телескопов. Информация о доле и пространственном распределении безоблачных событий для исследуемых районов получена по ночным данным радиометра VIIRS спутниковой платформы Suomi NPP для периода октябрь 2017 – март 2018 г. Установлена наиболее подходящая зона для размещения оптических черенковских телескопов.

*Ключевые слова: астроклимат, астрофизические наблюдения, космические лучи, Республика Алтай, Курайская степь, Чуйская степь, радиометр VIIRS/SNPP.*

**Введение.** Результаты многолетних астрофизических наблюдений, проводимых комплексом телескопов гамма-обсерватории TAIGA [1], позволили сформулировать новый подход по регистрации гамма-излучения. Восстановление энергии, направления и положения оси широких атмосферных ливней (ШАЛ) производится по данным атмосферных черенковских телескопов TAIGA-HiSCORE. Для определения типа частицы, породившей ШАЛ, используется полученная с помощью установки TAIGA-IACT информация о форме черенковского изображения ШАЛ. Совместная работа телескопов этих двух типов установок, с одной стороны, снимает ограничения на расстояние между установками комплекса, оно может достигать ~1 км, с другой стороны, предъявляются большие требования к площади всего комплекса. Первая очередь проекта TAIGA, реализующая гибридный подход в измерениях черенковского света, к 2020 г. практически достигнет предела чувствительности для регистрации гамма-излучения с энергией выше 100 ТэВ. Для достижения большей чувствительности площадь гамма-обсерватории должна быть порядка 10 км$^2$ и более. В настоящее время обсуждается проект второй очереди обсерватории.

Для проведения ночных оптических астрофизических наблюдений к территории расположения обсерватории, также предъявляется ряд серьезных требований по содержанию водяного пара и аэрозолей в атмосфере, частоте появления облачности, уровню светового загрязнения. Важным элементом является также наличие транспортной и телекоммуникационной инфраструктуры.

При проектировании второй очереди обсерватории TAIGA с учетом приведенных выше требований важное место могут играть данные современных спутниковых систем. Такой прибор, как VIIRS (Visible Infrared Imaging Radiometer Suite) [2], расположенный на спутниковой платформе Suomi NPP (Suomi National Polar-orbiting Partnership) [3], предоставляет достаточное количество информации об основных метеорологических характеристиках региона, а методы современных геоинформационных систем (ГИС) позволяют использовать данные из гетерогенных источников и значительно сократить время, затрачиваемое на обработку и анализ результата.

Целью работы является изучение астроклиматических условий для территории Республики Алтай с использованием данных спутниковых наблюдений, а также установление зон, подходящих для проведения астрофизических наблюдений.

**Техническая и информационная база.** Информационной основой проводимого в работе исследования являются данные радиометра VIIRS спутника Suomi NPP, получаемые

станцией УниСкан-24 центра космического мониторинга Алтайского госуниверситета в режиме прямого вещания. Схема приема и обработки спутниковой информации радиометра VIIRS реализованная и используемая в Центре, состоит из следующих основных этапов:

- распаковка сырого потока спутниковых данных;
- геолоцирование и калибровка данных;
- построение маски облачности;
- определение измеряемых величин (геофизических параметров атмосферы и подстилающей поверхности) по реализованным в центре алгоритмам, интеграция этих данных с ГИС.

Основным источником информации, используемым для решения поставленной цели, являлся продукт VIIRS Cloud Mask (маска облачности) алгоритма обработки данных радиометра VIIRS платформы Suomi NPP версии 1.5.08.04 [4].

Для определения параметров рельефа подстилающей поверхности (ПП) использовались данные цифровой модели местности SRTM (Shuttle Radar Topographic Mission) с пространственным разрешением 90x90 м [5].

Статистическая обработка и постобработка данных выполнялись с использованием ГИС GRASS (Geographic Resources Analysis Support System) версии 7.2 [6].

Зоны территории Республики Алтай, наилучшим образом подходящие для выполнения астрофизических наблюдений, отбирались на основе следующих требований:

- максимальная доля безоблачных событий в период активной фазы наблюдений с октября по март;
- высота полигона над уровнем моря от 1500м;
- угол наклона местности в пределах полигона не более 5$^{o}$;
- развитая транспортная и телекоммуникационная инфраструктура.

**Результаты.** На первом этапе установлена доля безоблачных событий для территории Республики Алтай с использованием ночных данных радиометра VIIRS/SNPP для периода октябрь 2017 – март 2018 г. На рис. 1 представлена доля событий с безоблачной атмосферой, полученная по данным VIIRS/SNPP для исследуемого периода. Видно, что районы с наибольшей долей безоблачных ночей находятся в основном в высокогорной части республики Алтай, преимущественно в труднодоступных горных массивах.

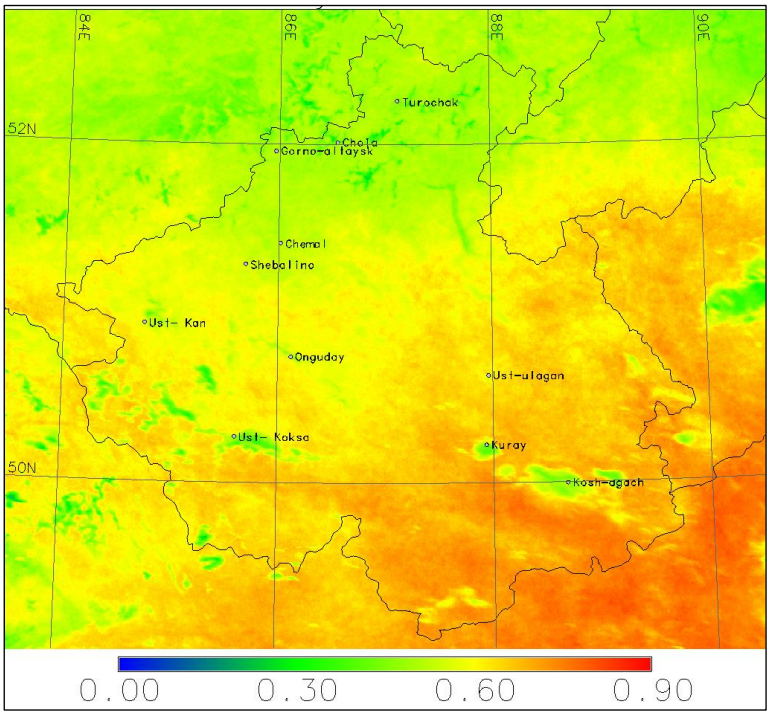



Рис. 1. Доля событий с безоблачной атмосферой по данным VIIRS/SNPP для территории Республики Алтай за период октябрь 2017 – март 2018 г.

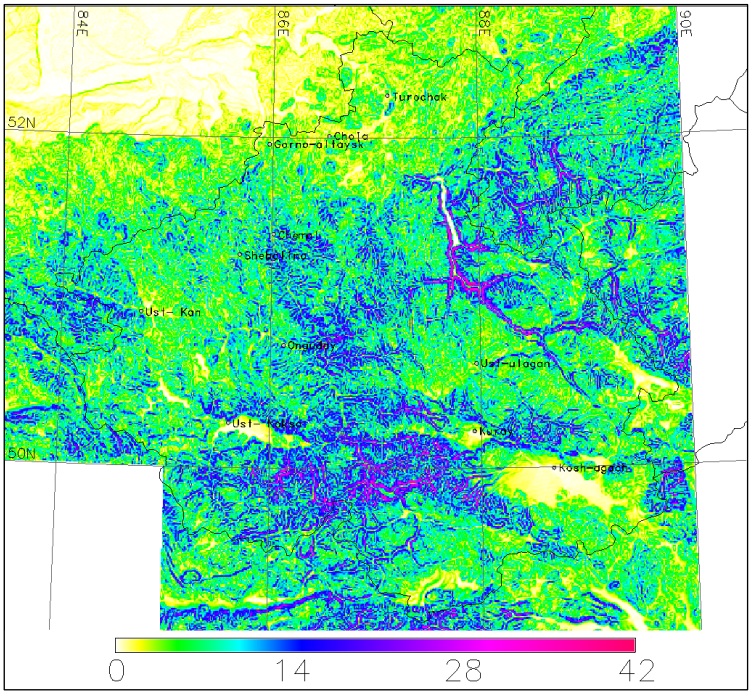



Рис. 2. Уклоны участков (угловые градусы) ПП для территории Республики Алтай, полученные по данным цифровой модели местности SRTM.

На рис. 2 показаны результаты обработки данных цифровой модели местности SRTM. Показаны углы уклонов участков ПП для территории Республики Алтай.

Совместный анализ данных, показанных на рис. 1 и 2, позволил идентифицировать районы Республики Алтай, для которых высота участков ПП составляет 1000-2000 м над уровнем моря и угол уклона местности не превышает 3-5º. На рис. 3 показаны доли безоблачных событий в районах, удовлетворяющих этим условиям. Наиболее подходящим для размещения гамма-астрономического эксперимента районом является западная часть Чуйской котловины с долей безоблачных ночей, достигающей 0,75.

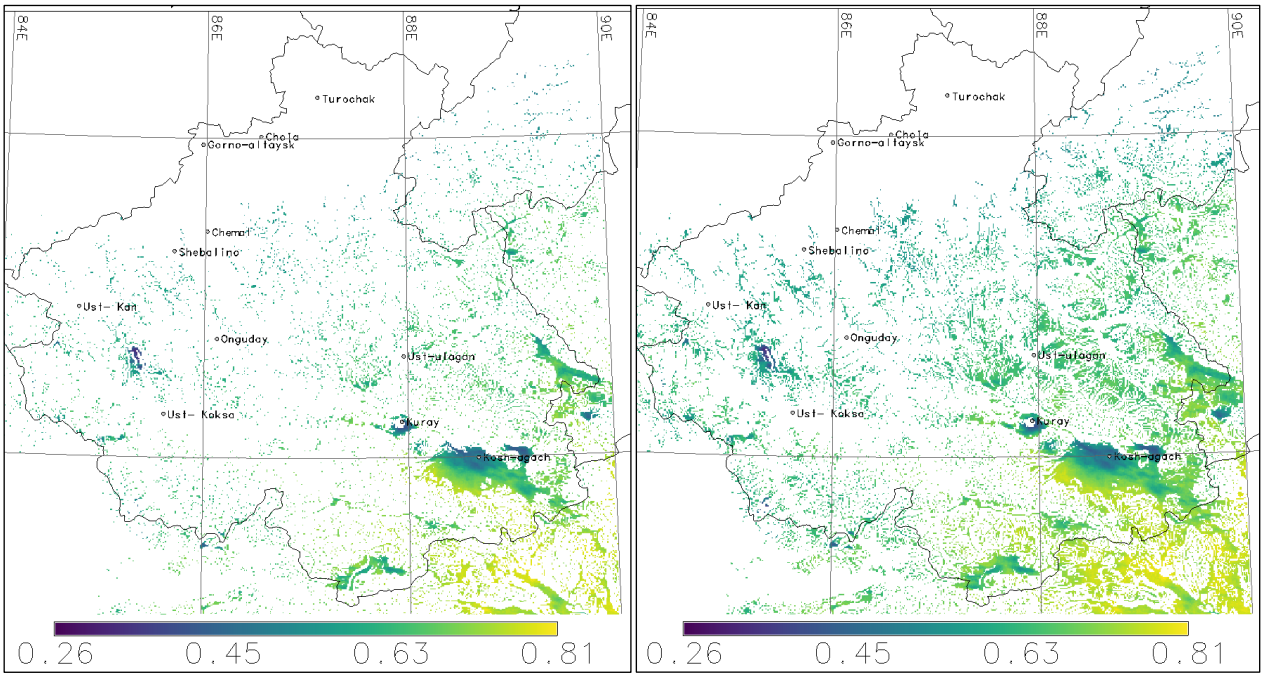



Рис. 3. Доля событий с безоблачной атмосферой по данным VIIRS/SNPP для территории Республики Алтай за период октябрь 2017 — март 2018 г. на высотах >1500 м над уровнем моря и уклоном не более 3º (слева) и 5º (справа).

**Заключение.** Проведено исследование астроклиматических условий для выполнения ночных астрофизических наблюдений на территории республики Алтай. Получено пространственное распределение безоблачных событий по ночным данным радиометра VIIRS спутниковой платформы Suomi NPP для периода октябрь 2017 – март 2018 г.

В качестве возможных регионов для размещения второй очереди обсерватории TAIGA по топографическим и инфраструктурным критериям наиболее подходят зоны Курайской и Чуйской котловин. Значительная часть этих зон расположена на высотах ~1000-2000 м над уровнем моря, что соответствует требованиям, предъявляемым к наземным гамма-астрономическим наблюдениям. Удаленное расположение этих регионов от крупных агломераций и промышленных центров приводит к незначительному содержанию аэрозоля в атмосфере и низкому уровню светового загрязнения.

Вместе с тем, установлено, что доля безоблачных событий в Курайской долине (средняя высота ~1000 м) в исследуемый период ночью не превышает 35%. Для западной части Чуйской степи (средняя высота ~2000 м) эта доля достигает ~75%, что делает этот район наиболее подходящим местом для проведения гамма-астрономических наблюдений с использованием оптических черенковских телескопов.

## ЛИТЕРАТУРА


[1] *Budnev N., Astapov I., Bezyazeekov P. et al.* // JINST. 2017. V. 12. C08018.

[2] *Hillger D., Kopp T., Lee T.* First-Light Imagery from Suomi NPP VIIRS // BAMS. 2013. V. 94. P. 1019-1029.

[3] *Murphy R.E.* The NPOESS Preparatory Project // J.J. Qu, W. Gao, M. Kafatos et al. Earth Science Satellite Remote Sensing. Springer, Berlin, Heidelberg, 2006.

[4] NASA Goddard Space Flight Center's (GSFC) Direct Readout Laboratory (DRL). VIIRS Cloud Mask Science Processing Algorithm. User's Guide, Greenbelt, Maryland, 2018, 11 p.

[5] *Reuter H.I, Nelson A., Jarvis A.* An evaluation of void filling interpolation methods for SRTM data // International Journal of Geographic Information Science. 2007. Vol. 21, N 9. P. 983-1008.

[6] Geographic Resources Analysis Support System. **https://grass.osgeo.org** (дата обращения 06.06.2019).

# НАБЛЮДЕНИЕ КОЛЬЦЕВЫХ СТРУКТУР НА ЛЬДУ БАЙКАЛА С ПОМОЩЬЮ СПУТНИКОВ SENTINEL 1 И SENTINEL 2 ВЕСНОЙ 2018-2019

*Родионова Н.В.*
Фрязинский филиал Института радиотехники и электроники им. В.А. Котельникова РАН, Фрязино

Проведен анализ изображений кольцевых структур на льду Байкала, полученных с помощью радарных данных Sentinel 1 С- диапазона и мультиспектральных данных Sentinel 2 с пространственным разрешением 10 метров за апрель 2018-2019 годов. Кольцевые структуры на изображениях обнаружены в районе мыса Нижнее Изголовье полуострова Святой Нос, мыса Улан-Нур, залива Посольский Сор в 2018 году и в районе мыса Н. Изголовье, в северной части Малого моря и в южной части Байкала в районе Слюдянки в 2019 году.
*Ключевые слова: Sentinel 1 радарные данные С-диапазона, Sentinel 2 мультиспектральные данные, кольцевые структуры на льду Байкала, коэффициент обратного рассеяния.*

**Введение.** На спутниковых снимках Байкала на весеннем льду можно увидеть тёмные кольца диаметром 5-7 км и шириной около 1 км. Чаще всего такие кольца появляются в апреле или в начале мая, иногда в январе. Средняя продолжительность существования этих явлений – 7-10 дней. Лёд в центре и вне кольца более толстый и белый, а в самом кольце – тёмный и более тонкий. Кольца возникают достаточно непредсказуемо в разных местах и в разные годы [1]. Исследованием кольцевых структур спутниковыми (LANDSAT, MODIS, SPOT, Sentinel) и наземными методами и построением гипотез и моделей возникновения данных структур занимались многие авторы [2-7]. Как отмечал Кураев [7], ледовые кольца не новый феномен. Они были обнаружены на архивных снимках американского спутника LANDSAT, датированных ещё началом 1970-х годов. Кольца на льду Байкала возникают на прежних местах не каждый год, возможно их появление на новых местах, поэтому мониторинг колец методами ДЗЗ необходим как с научной точки зрения, так и с точки зрения безопасности передвижения по льду людей и автотранспорта. В работе Кураева и др. [7] приведен наиболее полный перечень обнаруженных кольцевых структур на льду Байкала.

В данной работе приводятся изображения кольцевых структур на льду Байкала, обнаруженных в весенний период 2018, 2019 годов по радарным данным Sentinel-1 (S1) и оптическим данным Sentinel-2 (S2).

Преимуществом радарной съемки является получение изображения поверхности независимо от погоды и времени суток. Для получения качественных оптических снимков серьезным препятствием являются облачность и ограниченное время съемки.

Видимость кольцевых структур на льду в оптическом диапазоне совершенно другая, чем на радиолокационном изображении (РЛИ) из-за разных физических механизмов формирования изображений. В видимом диапазоне (в отраженном солнечном свете) кольцо на льду состоит из белого ледяного центра, где лед толще, и правильного темного кольца, где лед тоньше. В то же время на РЛИ кольцо можно обнаружить из-за шероховатости поверхности льда, а именно, по сети видимых на РЛИ периферийных трещин, которые могут указывать на таяние льда, определенно связанное с кольцом. Увеличение шероховатости приводит к увеличению коэффициента обратного рассеяния (КОР) и более ярким областям на РЛИ. Таяние льда иногда приводит к образованию полыньи, что на РЛИ обнаруживается по темному тону.

**Исходные спутниковые данные.** Спутник ESA Sentinel 1A запущен в апреле 2014 года, Sentinel 1B в апреле 2016 года. На борту S1 установлен РСА (радар с синтезом апертуры), который обеспечивает всепогодную поставку космических снимков в открытом доступе. Съемка выполняется в С-диапазоне (длина волны 5.4 см) с разным пространственным разрешением в зависимости от режима съемки. В данной работе использовался режим IW (interferometric wide swath) моды с поляризациями VV и VH и пространственным разрешением 10 м. Спутники находятся на одной орбите, осуществляя съемку одной и той же территории с

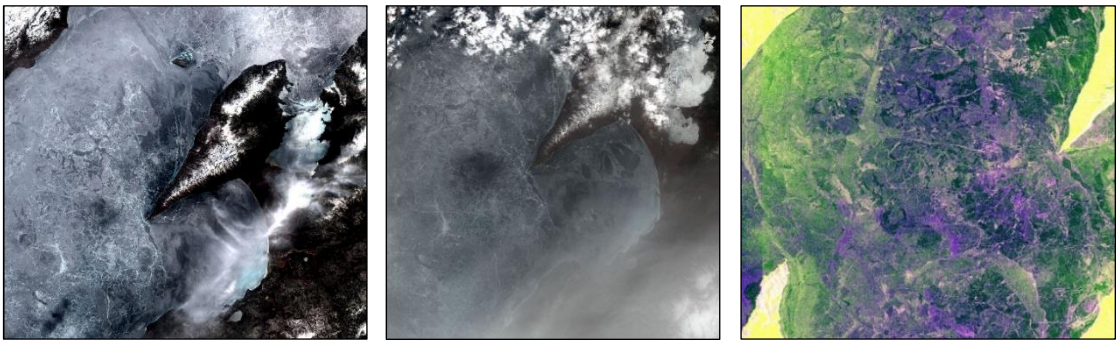
периодичностью 12 дней, при работе двух спутников 6 дней. Работа с изображениями S1 осуществлялась с помощью программы S1Toolbox и позднее SNAP [8]. Предварительная обработка радарных данных включала в себя выделение фрагмента с исследуемой областью и радиометрическую калибровку.

Спутник ESA Sentinel 2A запущен в июне 2015 года, второй Sentinel 2B в марте 2017 года. Периодичность мультиспектральной съемки каждым спутником составляет 10 дней, а при работе двух аппаратов– 5 дней. В мультиспектральной камере 13 каналов с разным пространственным разрешением (ПР) от 10 до 60 метров. Для каналов B2 (490 нм), B3 (560 нм), B4 (665 нм) и B8 (842 нм) ПР=10 м. Для получения изображений в естественных цветах в RGB кодировке используется комбинация каналов 4 3 2. Обработка изображений осуществлялась программой SNAP.

**Кольцевые структуры 2018-2019 годов по данным S1 и S2.** *А. П-ов Святой Нос, мыс Нижнее Изголовье, апрель 2018 и 2019.* В апреле 2018 года кольцевые структуры на льду появились у мыса Н. Изголовье на том же месте, что и в предыдущие годы. На рис. 1,*а* и *б* показаны снимки S2 за 23 и 28 апреля 2018 года в комбинации каналов 4 3 2, и на рис. 1,*в* приведено поляриметрическое изображение за 26 апреля 2018 года в комбинации каналов VV и VH (RGB кодировка: красный цвет – VV, зеленый – VH, синий – VV/VH). В данном случае на радарном изображении проблематично найти кольцевые структуры без «оптической подсказки».

Особенность кольцевых структур в районе мыса Н. Изголовье в 2019 году состоит в появлении 2-х колец (рис. 2), одного кольца на обычном месте (слева от мыса) с координатами центра 53,456$^0$ с.ш., 108,316$^0$ в.д. и второго справа от мыса с координатами 53,46 с.ш., 108,626 в.д. (координаты даны по снимку S1 за 3 мая 2019 года).

В данном случае не только на оптическом, но на радарном снимке визуально обнаруживаются два кольца в районе мыса Н. Изголовье.



*а* *б* *в*

Рис. 1. Снимки S2B L1C, дата съемки 23.4.2018 (*а*) и 28.4.2018 (*б*), каналы 4-3-2; поляриметрическое изображение (*в*) за 26.4.2018.

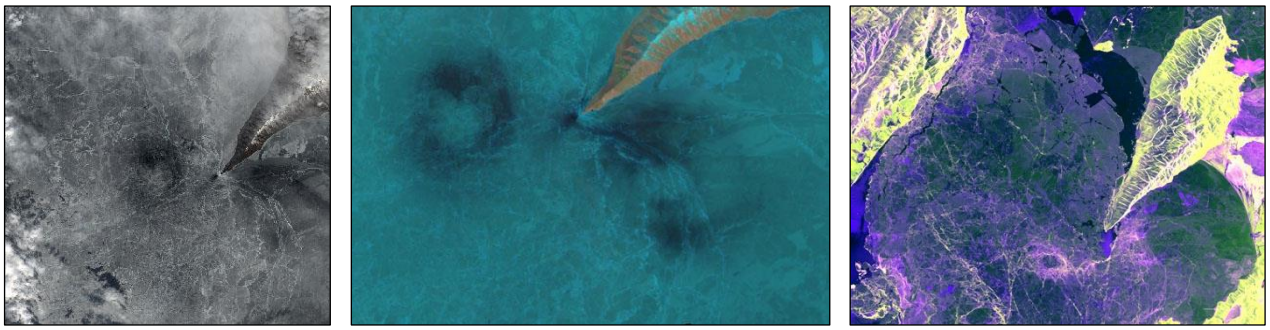


Рис. 2. Снимки S2B L1C, дата съемки 23 и 28 апреля 2019 (слева и в центре) каналы 4-3-2, поляриметрическое изображение дата съемки 3.5.2019, угол обзора 35$^0$ (справа).

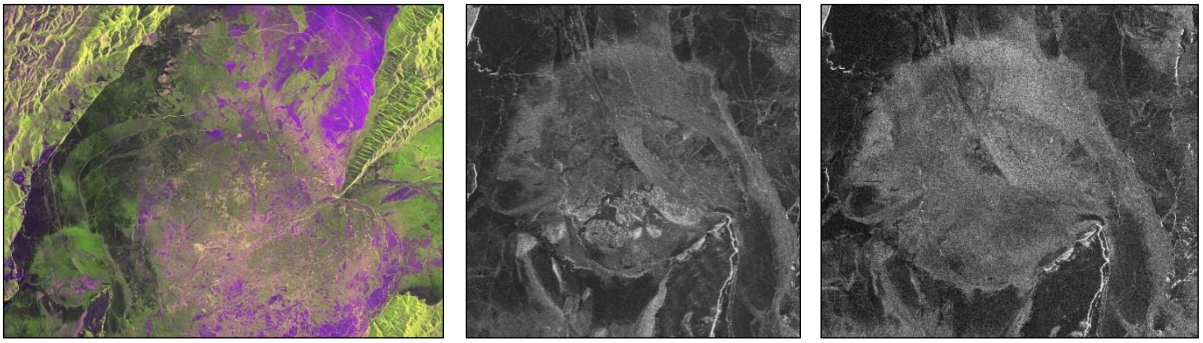


Рис. 3. Поляриметрическое изображение кольца в районе северной части Малого моря, и радарные изображения кольца поляризации VV и VH дата съемки 9 апреля 2019 года.

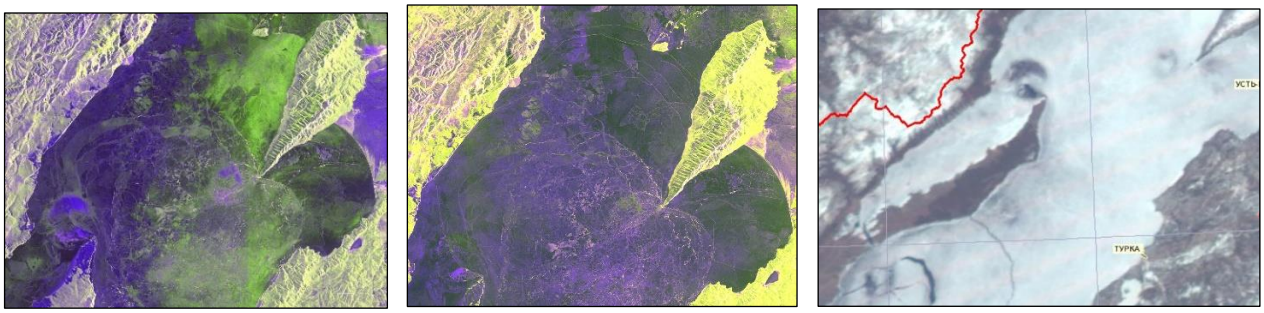
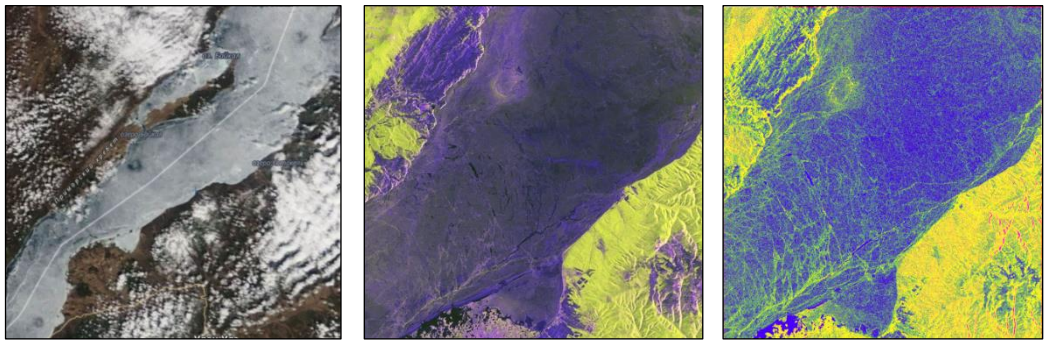



Рис. 4. S1 изображения за 16.4.2019, 21.4.2019, и Terra MODIS изображение за 24.4.2019.



*а* *б* *в*

Рис. 5. *А* – снимок MODIS, дата съемки 25.4.2018; S1 поляриметрическое (*б*) и текстурное (*в*) изображение за 26.4.2018.

*Б. Малое море север.* В апреле 2019 года в северной части пролива Малое море наблюдалась обширная кольцевая структура, визуально определяемая как на радарном, так и на оптическом изображениях. На рис. 3 (слева) показано поляриметрическое изображение за 9.4.2019, угол обзора $36^0$, координаты центра кольца 53,46 с.ш, 107,7 в.д., диаметр кольца порядка 17 км. На рис. 3 (в центре и справа) показана структура внутри кольца на радарных изображениях для поляризаций VV и VH.

На рис. 4 приведена последовательность радарных поляриметрических изображений кольца в районе северной части Малого моря за 16 и 21 апреля 2019 года.

Кольцевые структуры наблюдалась в 2019 году также с восточной стороны Ольхона. На рис. 4 (справа) дан снимок Terra MODIS за 24.4.2019 года в комбинации каналов 1 4 3, где обнаруживаются две кольцевые структуры с восточной стороны Ольхона: одна справа от пролива Ольхонские ворота и вторая в направлении на северо-восток. На этом же изображении видны два кольца по обе стороны мыса Н. Изголовье и кольцо в районе Малого моря, север.

*В. Мыс Улан-Нур, апрель 2018.* На изображении MODIS (рис. 5,*а*) дата съемки 25.4.2018 в комбинации каналов 1-4-3 (каналы 1 и 2 (0.62-0.67 мкм и 0.841-0.876 мкм, соответственно),

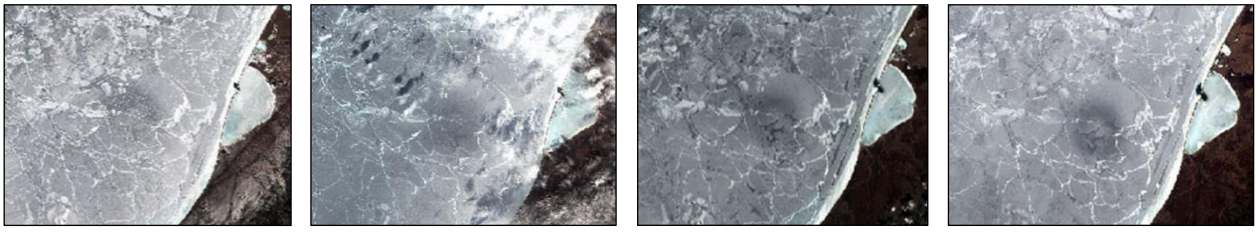
ПР=250 м; канал 4 (0.545-0.565), ПР =500м) наблюдаются три кольцевые структуры: в районе мыса Н. Изголовье, мыса Улан-Нур и залива Посольский Сор.

На S1 поляриметрическом РЛИ (RGB кодировка: красный цвет- VV поляризация, зеленый цвет – VH поляризация, синий цвет – VV/VH) за 26.4.2018, угол обзора $39.2^0$, видим кольцевую структуру с координатами центра $52.798^0$с.ш., $106.8055^0$в.д. в районе мыса Улан-Нур (рис. 5,*б*). Размеры кольца порядка 4.5 км. На текстурном изображении (рис. 5,*в*) (текстурные признаки Харалика [9]) кольцо лучше различимо в сравнении с поляриметрическим изображением, и текстурные признаки позволяют осуществить интерпретацию изображения. А, именно, контур кольца (желтый цвет) свидетельствует о повышенном значении «контраста» и «энтропии», т.е. наличии перепадов высот и хаоса на фоне основного синего цвета (гладкая поверхность льда).

*Г. Залив Посольский Сор, март, апрель 2018.* На рис. 6 показана последовательность изображений S2 с кольцевой структурой с координатами $51.9254^0$с.ш., $105.9647^0$в.д. в районе залива Посольский Сор, даты съемки 29 марта, 8, 16 и 23 апреля 2018 года. На рис. 7,*а* показано поляриметрическое изображение этого района, дата съемки 19.4.2018. Распознать кольцевую структуру (показана стрелкой и выделена окружностью) на РЛИ проблематично (несколько светлее тон). На текстурном изображении (рис. 7,*б*) в области ледового кольца отмечается более высокое значение «энтропии» (зеленый цвет) – повышенный хаос – в сравнении с окружением, и структура более различима визуально в сравнении с поляриметрическим изображением. Отметим, что в области ледового кольца у мыса Улан-Нур на текстурном изображении текстурный признак «энтропия» также имеет повышенное значение.

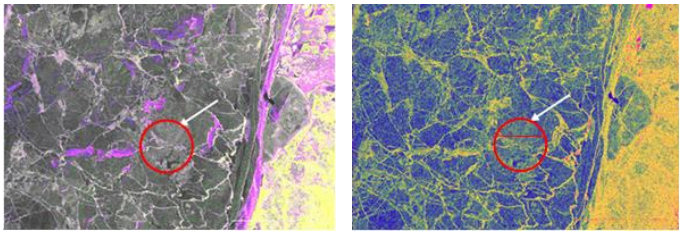
*Д. Слюдянка, апрель 2019 года.* На рис. 8 показана последовательность изображений S2 кольцевой структуры в южной части Байкала в районе Слюдянки, координаты кольца $51.685^0$ с.ш., $103.85^0$ в.д. Кольцо наблюдалось даже 1 мая 2019 года.

**Причины и механизмы образования кольцевых структур на льду. Версии. Гипотезы.** В настоящее время есть несколько гипотез о причинах и механизмах образования кольцевых ледовых явлений. Так, по мнению ряда авторов, Балханова и др. [4], Гранина и др. [3], Гунина [5], образование кругов связано с выбросами природного горючего газа (метана) из осадочной толщи дна Байкала. Механизм образования колец на ледяной поверхности, возможно, таков: газ, поднимаясь со дна Байкала, действует на восходящий водный поток, который при подъеме закручивается силами Кориолиса. В итоге в приповерхностном слое воды подо льдом формируется круговое относительно теплое течение, разрушающее ледяной покров снизу, напитывая его водой, и на поверхности льда проявляется темное кольцо [3].



29.3.2018 8.4.2018 16.4.2018 23.4.2018

Рис. 6. Кольцевая структура в районе залива Посольский Сор по снимкам S2A MSI L1C в марте и апреле 2018 года, комбинация каналов 4-3-2.



*а* *б*

Рис. 7. S1 поляриметрическое (*а*) и текстурное (*б*) изображения района залива Посольский Сор, дата съемки 19.4.2018.


Рис. 8. S2 изображения за 14.4.2019, 29.4.2019 и 1.5.2019.

По версии Бордонского и Крылова [2] кольца образуются из-за нарушения структуры льда вследствие воздействия на лёд переменных потоков тепла в условиях резко континентального климата, когда возникают значительные перепады дневных и ночных температур.

Авторы [6] считают главной причиной образования колец на льду озера Байкал подледные антициклонические вихри, которые возникают в определенных местах озера благодаря общей циркуляции и особенностям подводной топографии. Основным свидетельством в пользу вихрей является их генерация в тех же самых местах в условиях отсутствия льда, что видно время от времени на оптических IR и радарных изображениях.

Переосмыслению существующих гипотез и предложению гипотезы о линзовидных вихрях как причине появления колец на льду Байкала посвящена работа Кураева и др. [7], основу которой составили натурные измерения на более чем 250 станциях в районах уже известных или возможных появлений кольцевых структур.

Дальнейшие спутниковые наблюдения, натурные измерения, создание моделей генерации волн в замкнутом бассейне со сложной береговой структурой и подводной топографией, не без влияния придонных процессов, помогут лучше понять феномен ледовых колец и прогнозировать время и место их появления на Байкале.

## ЛИТЕРАТУРА


[1] **http://www.geol.irk.ru/baikal/baikal/krugi/baikal_ring.html**.

[2] *Бордонский Г.С., Крылов С.Д.* О природе кольцевых образований на спутниковых снимках ледяного покрова озера Байкал // ИЗК. 2014. №4. С. 27-31.

[3] *Гранин Н.Г., Козлов В.В., Цветова Е.А., Гнатовский Р.Ю.* Полевые исследования и некоторые результаты численного моделирования кольцевой структуры озера Байкал // ДАН. 2015. Т. 461. №. 1. С. 316-320.

[4] *Balkhanov V.K., Bashkuev Yu.B, Khaptanov V.B.* Formation of circular rings on the snow-covered ice field of Lake Baikal // Technical Physics. 2010. 55(9). P. 1266-1269.

[5] *Gunin V.I.* Assessment of gas-hydrothermal activity of the Baikal rift in the lake area from numerical experiment data // Geodynamics & Tectonophysics. 2014. 5(3). P. 763–775.

[6] *Ivanov A., Evtushenko N., Filimonova N., Terleeva N.* Atmospheric and oceanic phenomena in Lake Baikal visible from space on SAR and optical images // Proceedings ESA Living Planet Symposium. 2016. 9-13 May 2016, Prague, Czech Republic (ESA SP-740).

[7] *Kouraev V., Zakharova E.A., Remy Fr., Kostianoy A.G., Shimaraev M.N., Hall N.M.J., Suknev A.Y.* Giant ice rings on lakes Baikal and Hovsgol: Inventory, associated water structure and potential formation mechanism // Limnol. Oceanogr. 2016. 61. P. 1001–1014. DOI: 10.1002/lno.10268.

[8] **https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-1**.

[9] *Haralick R.M.* Statistical and structural approaches to texture // Proc. IEEE. 1979. 67 (5). P. 786-804.

# ОПРЕДЕЛЕНИЕ ТАЛОГО/МЕРЗЛОГО СОСТОЯНИЯ ПОЧВ ЯКУТИИ ПО РАДАРНЫМ ДАННЫМ SENTINEL 1

*Родионова Н.В.*
Фрязинский филиал Института радиотехники и электроники им. В.А. Котельникова РАН, Фрязино

Рассматривается вопрос об идентификации талых/мерзлых почв в приповерхностном слое в районе метеостанций Тикси, Верхоянска, Томтора (Оймякона) и Якутска с широтой места от $62^0$ до $71^0$ с.ш. по радарным данным Sentinel 1B С-диапазона с поляризациями VV и VH за период 2017-2018 годов. Определение состояния мерзлой/талой почвы осуществляется тремя способами: 1) по многовременным радарным данным на основе значительного на 3-5 дБ перепада значений коэффициента обратного рассеяния (КОР) при переходе к состоянию замерзания/оттаивания почвы, 2) по нахождению порогового значения $\sigma^0_{порог}$, при котором температура в верхнем слое почвы опускается ниже $0^0$С, 3) по текстурным признакам.
*Ключевые слова: радарные данные С-диапазона, коэффициент обратного рассеяния, температура воздуха, талое/мерзлое состояние почвы, коэффициент корреляции Спирмена.*

**Введение.** Дистанционный радарный мониторинг приповерхностного слоя мерзлых почв Арктических регионов России стал доступен на постоянной основе с запуском аппаратов Sentinel 1, осуществляющих радиолокационную съемку в С-диапазоне длин волн с периодом в 12 дней с разрешением 10 м. Исследование мерзлых почв, как правило, связано с наземными измерениями глубины протаивания. Дистанционный мониторинг важен прежде всего из-за обширности масштаба распространения вечной мерзлоты на территории России и ограниченной доступности территорий. Однако, возможности радарной съемки ограничены из-за проникновения волны в мерзлую почву лишь на несколько см. Радар различает мерзлую почву по изменению содержания в ней жидкой воды, и, соответственно, по изменению диэлектрической проницаемости (ДП) почвы [1-3]. Несколько работ посвящены использованию Sentinel 1 для идентификации мерзлых почв [4-6].

Целью данной работы является определение на основе многовременных радарных данных Sentinel 1 талых/мерзлых почв на территории Якутии, которая является самым холодным обитаемым регионом в Северном полушарии Земли.

**Исходные радарные данные Sentinel 1B.** В работе использованы находящиеся в открытом доступе Sentinel 1B (S1) радарные данные С-диапазона IW (interferometric wide swath) моды с поляризацией VV и VH и пространственным разрешением 10 м. Работа с изображениями S1 осуществлялась с помощью программы S1Toolbox и позднее SNAP [7]. Съемку исследуемых территорий России выполняли аппараты Sentinel 1A и Sentinel 1B, первый из них с апреля 2015 года по сентябрь 2016 года, второй – с конца сентября 2016 года.

Предварительная обработка данных включала в себя выделение фрагмента с исследуемой областью и радиометрическую калибровку.

В таблице приведены данные для каждого исследуемого места по числу обработанных сеансов съемки, углу обзора, временному периоду съемки и количеству точек при усреднении КОР по профилю.

**Районы исследования**. Среди 4 районов исследования в Якутии 3 города относятся к самым холодным городам России: Верхоянск, Оймякон, Якутск.

***Тикси.*** Тикси – поселок городского типа, центр Булунского района республики Саха-Якутия находится за полярным кругом, к востоку от устья Лены на берегу одноименной бухты в Море Лаптевых. Координаты метеостанции аэропорта Тикси (WMO 21824) $71^0$41.85’ с.ш., $128^0$ 54.18’ в.д.

Информация об использованных радарных данных Sentinel 1B.

| Название станции и координаты | Число обработанных сеансов съемки | Угол обзора | Период съемки | Количество точек в профиле |
|---|---|---|---|---|
| **Тикси** 71.7$^0$ с.ш., 128.9$^0$ в.д. | 39 | 40,5$^0$ | 19.2.2017–14.6.2018 | ~20 |
| **Верхоянск** 67.547$^0$с.ш., 133.427$^0$в.д. | 29 | 43$^0$ | 21.10.2017-4.10.2018 | ~15 |
| **Томтор (Оймякон)** 63.264$^0$с.ш., 143.21$^0$в.д. | 25 | 37$^0$ | 10.10.2017-25.7.2018 | ~15 |
| **Якутск** 62.013$^0$с.ш., 129.657$^0$в.д. | 39 | 35,3$^0$ | 21.2.2017-16.6.2018 | ~15 |

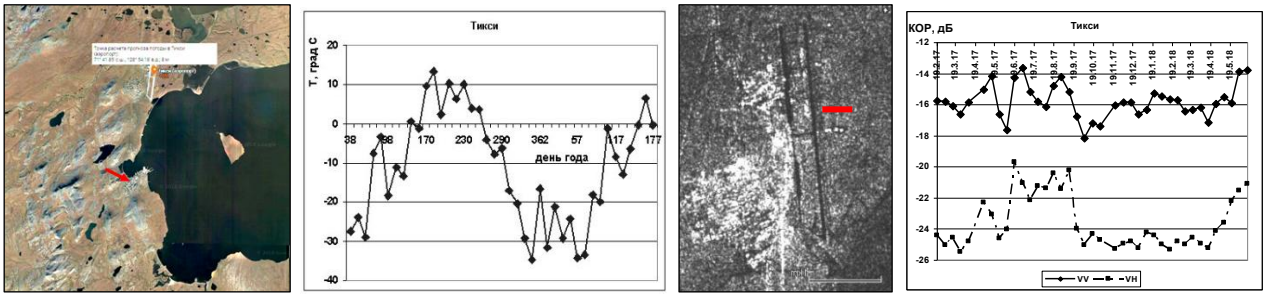



Рис. 1. Бухта с поселком Тикси и аэропортом с точкой расчета прогноза погоды, температура воздуха в Тикси, радарное изображение с профилем (красная линия), среднее значение КОР по профилю.

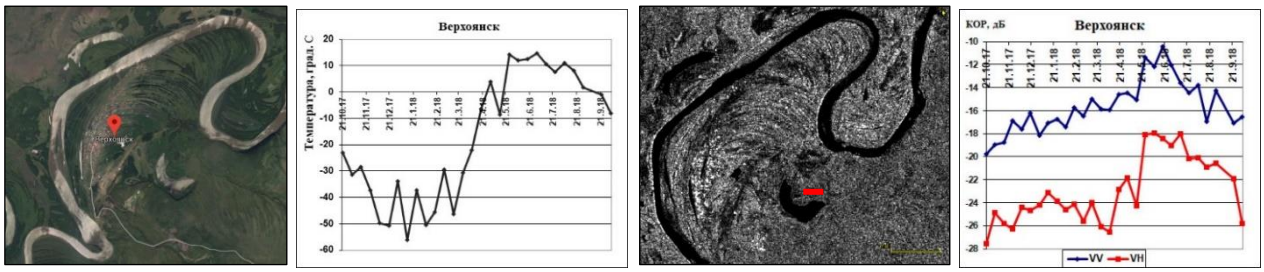



Рис. 2. Верхоянск на google map, температура воздуха, радарное изображение с профилем (красная линия), среднее значение коэффициента обратного рассеяния по профилю.

Бухта Тикси окружена невысокими щебнистыми безлесными холмами, даже кустарника здесь почти нет. Климат суровый, арктический. Морозы немного смягчает море, средняя температура января -37.3$^0$С, июля и августа – всего +7.6$^0$С и +7.7$^0$С, соответственно. Абсолютный минимум температуры в Тикси -50.5$^0$С. Самая большая возможная высота снежного покрова составляет 99 см. Апрель в Тикси имеет самую низкую среднемесячную температуру в России, составлявшую в 2006 году -27.6$^0$С. На рис. 1 показаны спутниковое изображение Тикси, температура воздуха в Тикси за период 19.2.2017-14.6.2018 по данным rp5.ru, радарное изображение района аэропорта Тикси и линия профиля, вдоль которого вычислялось среднее значение коэффициента обратного рассеяния (КОР) (координаты средней точки профиля 71.7015$^0$ с.ш. и 128.9205$^0$ в.д.) и графики изменения КОР двух поляризаций. Значения температуры воздуха взяты на момент съемки территории радаром S1.

***Верхоянск.*** Верхоя́нск – город в Верхоянском улусе Якутии, на правом берегу реки Яна, в 92 км к юго-западу от административного центра улуса – Батагая. Это самый северный город Якутии и один из самых маленьких населённых пунктов в России, имеющих статус города. Суровый климат, отсутствие осадков привели к тому, что в городе живет 1122 человека (2018 год). Это граница субарктического и умеренного поясов. Климат характеризуется коротким, сухим и теплым летом, и зимой со средней температурой -40$^0$. Верхоянск – одно из самых

холодных мест на Земле, самый холодный город в мире. Самая низкая температура -67,7 С была зарегистрирована здесь в феврале 1892 года. Верхоянск часто называют Полюсом холода северного полушария. На рис. 2 показаны спутниковое изображение Верхоянска, температура воздуха за период 21.10.2017-4.10.2018 по данным rp5.ru, положение профиля на радарном изображении для определения среднего значения КОР. Координаты средней точки на профиле $67.5468^0$ с.ш. и $133.427^0$ в.д.

***Томтор (Оймякон).*** Хотя официально статус Полюса холода отдан Верхоянску, но право о статусе самого холодного места северного полушария оспаривается Оймяконом (-65,4 °С). Численность населения села Оймякон 500 человек. Оймякон находится во впадине и со всех сторон защищен горами, задерживающими выход тяжелого холодного воздуха. Здесь резко континентальный климат. Среднемесячная температура января $-61^0$С. Село Томтор с метеостанцией (координаты $63.26382^0$ с.ш., $143.210462^0$ в.д.) находится в 30 км на юго-восток от Оймякона. В двух км от Томтора находится аэропорт. Населения здесь в 2 раза больше, чем в Оймяконе. На рис. 3 показаны спутниковое изображение Томтора, температура воздуха за период 10.10.2017-25.7.2018 по данным rp5.ru, положение профиля на радарном изображении вблизи аэропорта с координатами средней точки $63.2454^0$ с.ш., $143.1754^0$ в.д. и графики среднего значения КОР по профилю за исследуемый период.

***Якутск.*** Якутск – столица республики Саха. Самый крупный из городов, расположенных в зоне вечной мерзлоты. Город расположен в равнинной местности, в долине Туймаада на левом берегу реки Лены. Находится несколько севернее параллели 62 градуса северной широты. Климат города – резко континентальный. Среднегодовая температура – -8,8 С. Зима в Якутске исключительно сурова, средняя температура января составляет около -40 С. Самая низкая зафиксированная температура $-64^0$С. Для лета характерно небольшое количество осадков и часто сильная жара до $+40^0$. Годовая амплитуда $T^0$ Якутска – одна из наибольших на планете, примерно равна годовой амплитуде «полюсов холода» – Оймякона и Верхоянска, и превышает 100 С (102,8 С). Здесь располагается CALM площадка R42 с координатами $62.013306^0$с.ш., $129.656991^0$ в.д. [8]. Ландшафт – луг, почва сухая – супесь, песок. Глубина протаивания составляет 185-215 см. На рис. 4 показаны спутниковое изображение Якутска, температура воздуха за период 21.2.2017-16.6.2018, положение профиля на радарном изображении для определения среднего значения КОР и изменение значений КОР за исследуемый период времени.

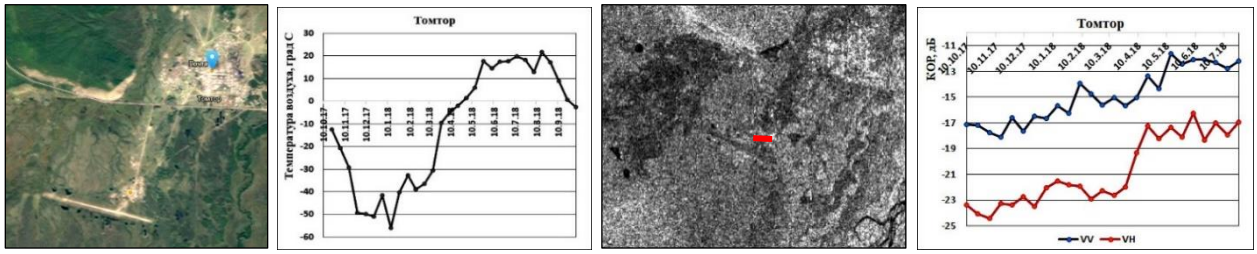


Рис. 3. Спутниковый снимок Томтора, температура воздуха по данным rp5.ru, РЛИ за 25.7.2018 VV поляризации и графики КОР за период 10.10.2017-25.7.2018.

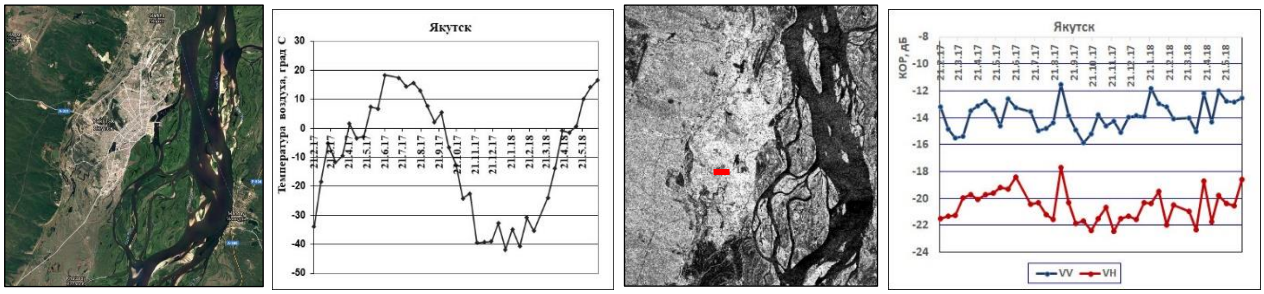


Рис. 4. Якутск, температура воздуха, РЛИ с линией профиля и графики изменения КОР.

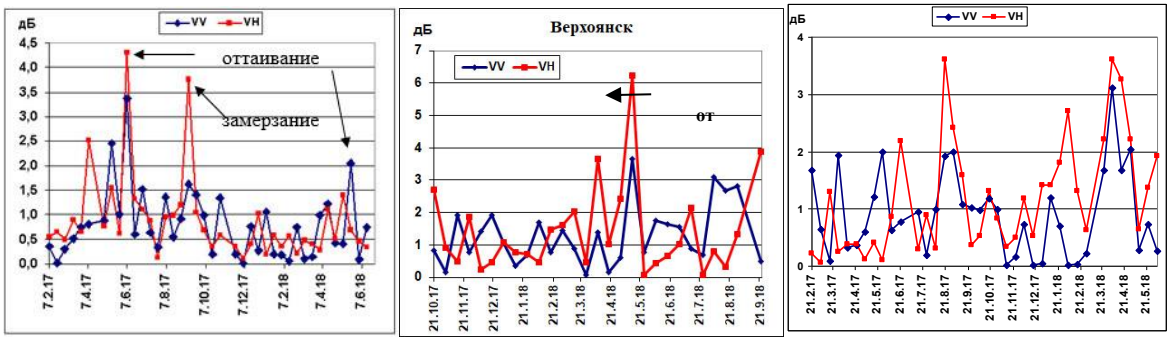
Коэффициент корреляции Спирмена между температурой воздуха и КОР для Тикси равен $\rho_S$=0.21 (p=0.1) для VV поляризации (слабая корреляции) и $\rho_S$=0.69 (p<0.01) для VH поляризации. Для Верхоянска $\rho_S$=0.68 (p<0.01) – VV и $\rho_S$=0.67 (p<0.01) – VH поляризация. Для Томтора $\rho_S$=0.77 (p<0.01) – VV и $\rho_S$=0.76 (p<0.01) – VH поляризация. Для Якутска $\rho_S$=0.198 (p=0.113) – VV и $\rho_S$=0.34 (p=0.017) – VH поляризация. На кросс-поляризации корреляция выше или близка к корреляции на согласованной поляризации.

**Определение состояния талой/мерзлой почвы по многовременным радарным данным**. Сигнал радара от поверхности преимущественно зависит от диэлектрической проницаемости почвы, которая напрямую связана с содержанием в ней воды и льда, кроме того, на КОР влияют шероховатость поверхности и растительный покров**.** Глубина проникновения э/м волны в почву определяется по формуле $\delta p = \frac{\lambda \cdot \sqrt{\varepsilon'}}{2 \cdot \pi \cdot \varepsilon''}$, где $\lambda$ - длина э/м волны, $\varepsilon = \varepsilon' + i\varepsilon''$ – диэлектрическая проницаемость почвы. Для Sentinel 1 длина волны равна 5.4 см. В случае мерзлой почвы (МП) с $\varepsilon' \approx 5.5$ и $0.1 \le \varepsilon'' \le 0.5$ получаем глубину проникновения э/м волны в почву $4 \le \delta p \le 20$ см. С помощью радара дистанционно можно определить состояние только приповерхностного слоя почвы, которое меняется при падении температуры почвы ниже $0^0$C с переходом воды в почве от жидкого к мерзлому состоянию. Данный процесс приводит к скачку значения ДП почвы и, соответственно, скачку в значениях КОР.

***Использование многовременных радарных данных.*** Используя многовременную последовательность радарных данных, построим графики разности значений КОР по абсолютной величине между соседними датами радарной съемки. Это позволит выявить скачки значений КОР, и, следовательно, даты начала процессов замерзания/оттаивания почвы. Кроме того, по значениям скачков КОР можно оценить изменения значений ДП почвы, используя формулу Dubois et al. [9], которая связывает КОР согласованной поляризации с параметрами почвы (реальной частью ДП почвы и шероховатостью поверхности). На рис. 5 показаны графики разности значений $\sigma^0$ по абсолютной величине в дБ для соседних дат съемки тестовых участков Тикси (а), Верхоянска (б) и Якутска (в), где по максимумам $|\sigma^0|$ определяем даты начала замерзания и оттаивания почвы. Следует обратить внимание, что перепады $|\sigma^0|$ для кросс-поляризации превышают значения перепадов для согласованной поляризации, что свидетельствует о большей чувствительности VH поляризации к мерзлым почвам.

***Использование фактора состояния поверхности.*** В работе [10] предложено определять состояние поверхности почвы через фактор состояния поверхности (СПФ), определяемый по радарным данным*:* $СПФ = \frac{1}{2} + \frac{\sigma^0(t) - \sigma^0_{лето}}{\sigma^0_{лето} - \sigma^0_{зима}}$, где $\sigma^0_{лето}$ *и* $\sigma^0_{зима}$ - средние значения КОР летом и зимой в дБ, $\sigma^0(t)$ – текущее значение КОР, *t* –время. Авторы [10] утверждают, что сезонные вариации СПФ сильно коррелируют с сезонными вариациями температуры почвы, принимая отрицательные значения при температуре почвы ниже $0^0$C. Тогда 5-см верхний слой почвы можно классифицировать как замерзший, если значение СПФ, измеренное радаром, становится отрицательным. На рисунке 6 показаны графики СПФ для Тикси и Верхоянска.

Назовем значение $\sigma^0(t)$, при котором СПФ=0, пороговым значением $\sigma^0_{порог}$. Тогда условие $\sigma^0 < \sigma^0_{порог}$ определяет мерзлую почву, а условие $\sigma^0 \ge \sigma^0_{порог}$ – талую почву. Для Тикси получаем $\sigma^0_{порог}$=-15,6 дБ для VV поляризации и $\sigma^0_{порог}$=-22,8 дБ для VH поляризации. Для Верхоянска $\sigma^0_{порог}$=-15,1 дБ для VV и $\sigma^0_{порог}$=-22,1 дБ для VH поляризации. В работе [5] приведена регрессия для зависимости $\sigma^0_{порог}$ от широты места. Для Тикси и Верхоянска также прослеживается такая зависимость.




*а* *б* *в*

Рис. 5. Разность значений $|\sigma^0|$ в дБ для соседних дат радарной съемки в районах исследования Тикси (*а*), Верхоянска (*б*) и Якутска (*в*).

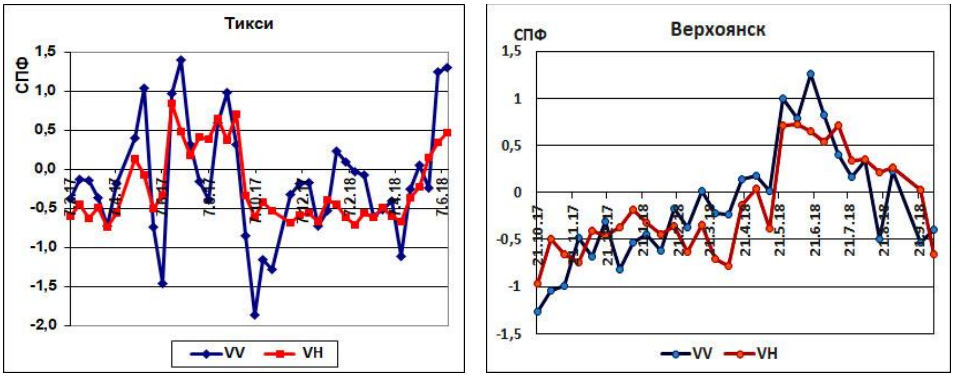
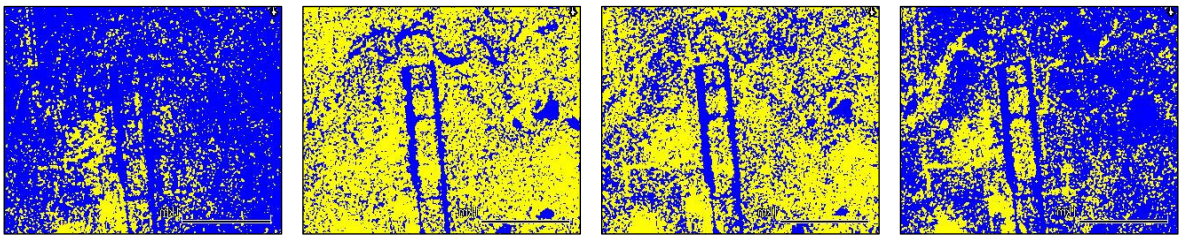



Рис. 6. Изменение СПФ для Тикси и Верхоянска.



7/6/2017 оттаивание 19/6/2017 11/9/2017 замерзание 23/9/2017

Рис. 7. Тикси, даты съемки 7.06.2017 и 19.06.2017 (период оттаивания почвы), и даты съемки 11.09.2017 и 23.09.2017 (период замерзания почвы). Синий цвет – значения $\sigma^0 < \sigma^0_{порог}$ (мерзлая почва), желтый цвет – значения $\sigma^0 > \sigma^0_{порог}$ (талая почва).

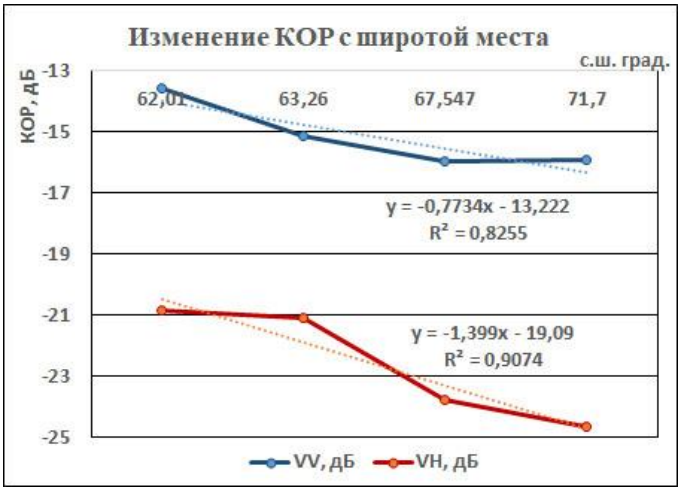



Рис. 8. Изменение КОР VV и VH поляризаций с широтой места.

Используя пороговые значения $\sigma^0_{порог}$, построим локальные карты изменения мест талой и мерзлой почвы в районе аэропорта Тикси для дат 7.6.2017 и 19.6.2017 – оттаивание, и 11.9.2017 и 23.9.2017 – замерзание. Полученные изображения показаны на рис. 7. Синий цвет – значения $\sigma^0 < \sigma^0_{порог}$ (мерзлая почва), желтый цвет – значения $\sigma^0 \geq \sigma^0_{порог}$ (талая почва). Отмечаем значительное уменьшение площади мерзлой почвы на изображении за 19.06.2017 при оттаивании и, соответственно, увеличение площади мерзлой почвы на изображении за 23.9.2017 при замерзании.

Пороговое значение $\sigma^0_{порог}$, как отмечено в [10] и получено в работе [5] для Аляски, зависит от широты места, уменьшаясь при увеличении широты. Аналогичная зависимость есть и для значений КОР. На рис. 8 приведены графики изменения среднего значения КОР VV и VH поляризаций за период с октября 2017 по июнь 2018 года для Якутска, Томтора, Верхоянска и Тикси в порядке возрастания широты места с коэффициентом детерминации регрессии 0.9 для кросс-поляризации и 0.8 для согласованной поляризации.

***Использование текстурных признаков***. Текстурные признаки Харалика [11] позволяют осуществить интерпретацию изображений, выделяя области с перепадом высот, гладкие поверхности. В работе [4] подробно описывается применение текстурных признаков для определения состояния мерзлой/талой почвы и приведены локальные карты мест талой/мерзлой почвы, где индикатором служит текстурный признак ‘обратный момент’, значение которого растет в местах с замершей почвой.

**Заключение.** В работе показаны некоторые возможности получения информации о состоянии мерзлой почвы в ряде мест Якутии по радарной съемке с помощью РСА Sentinel 1. По перепадам КОР возможно определить время начала замерзания/оттаивания почвы. По величинам перепадов КОР оценить величины перепадов значений ДП почвы в процессе замерзания/оттаивания почвы. По фактору состояния поверхности почвы можно получить пороговое значение КОР, позволяющее построить локальные карты мест талой/мерзлой почвы.

## ЛИТЕРАТУРА


[1] *Khaldoune J., Van Bochove E., Bernier M., Nolin M.C.* Mapping agricultural frozen soil on the watershed scale using remote sensing data // Appl. Environment Soil Sci. 2011. Article ID 193237. P. 1-16. DOI: 10.1155/2011/193237.

[2] *Ulaby F.T., Moore R.K., Fung A.K.* Microwave Remote Sensing: Active and Passive: Vol. 1. Fundamentals and Radiometry. Artech House, Dedham, Mass, USA, 1982.

[3] *Hallikainen M.T., Ulaby F.T., Dobson M.C., El-Rayes M.A., Wu L.K.* Microwave dielectric behavior of wet soil. part 1: empirical models and experimental observations // IEEE Trans. GRS. 1985. 23(1). P. 25-34.

[4] *Родионова Н.В.* Анализ радарных данных Sentinel 1 для идентификации талых и мерзлых почв в районе Анадыря (Чукотка) и Белой Горы (Якутия) // ИЗК. 2019. № 1. С. 29-37. DOI: 10.31857/S0205-96142019129-37.

[5] *Родионова Н.В*. Обратное рассеяние от приповерхностного слоя талых/мерзлых почв Аляски по радарным данным Sentinel 1 // РЭНСИТ. 2019. Т. 11. №1. С. 21-30.

[6] *Baghdadi N., Bazzi H., El Hajj M., Zribi M.* Detection of frozen soil using Sentinel-1 SAR data // Remote Sens. 2018. 10. 1182; DOI: 10.3390/rs10081182.

[7] **https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-1**.

[8] **https://www2.gwu.edu/~calm/data.north.html**.

[9] *Dubois P.C., van Zyl J.J., Engman T.* Measuring soil moisture with imaging radars // IEEE Trans. GRS. 1995. 33(4). P. 916-926.

[10] *Mironov V.L., Muzalevsky K.V.* Spaceborne radar monitoring of soil freezing/thawing processes in the arctic tundra // Russian Physics Journal. 2013. 55(8). (Russian Original №. 8. August, 2012).

[11] *Haralick R.M.* Statistical and structural approaches to texture // Proc. IEEE. 1979. 67(5). P. 786-804.

# ПРОГРАММНО-ТЕХНОЛОГИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ГЕОИНФОРМАЦИОННОЙ ВЕБ-СИСТЕМЫ ОПЕРАТИВНОЙ ОБРАБОТКИ И ВИЗУАЛИЗАЦИИ СПУТНИКОВЫХ ДАННЫХ

*Кадочников А.А.*
Институт вычислительного моделирования СО РАН, Красноярск
Федеральный исследовательский центр «Красноярский научный центр СО РАН», Красноярск

В работе рассматриваются некоторые особенности разработки программных инструментальных средств для загрузки, обработки и публикации данных дистанционного зондирования. Описаны механизмы пакетной обработки поступающих данных и особенности их подготовки перед публикацией в веб-приложении. Предложены методы по снижению размера архива промежуточных данных для работы этого веб-приложения, структура веб-приложения и используемые технологии и стандарты.
*Ключевые слова: данные дистанционного зондирования, веб-картография, ГИС, геопортал, геопространственные данные.*

Красноярский край обладает рядом географических и экономических особенностей, такими как протяженная территория, множество труднодоступных и слабо освоенных зон и их удаленность от мест концентрации промышленного производства и населения и др. Доступность регулярных данных дистанционного зондирования Земли (ДДЗ) высокого разрешения позволяет эффективно решать задачи во многих областях хозяйственной деятельности Красноярского края. Внедрение данных ДДЗ для края в тематических ГИС позволит решать задачи в области оперативного мониторинга чрезвычайных ситуаций и экологической обстановки, эффективности использования природных ресурсов и экономического мониторинга, а также в задачах лесного и сельского хозяйства.

Для решения задачи по обеспечению эффективного использования результатов космической деятельности в Красноярском крае в мае 2017 года был введен в эксплуатацию новый спутниковый приемный комплекс Федерального исследовательского центра «КНЦ СО РАН» (ФИЦ КНЦ СО РАН) на базе Единого регионального центра дистанционного зондирования Земли Красноярского края (ЕРЦ ДЗЗ) (http://ksc.krasn.ru/news/2017-05-12/). Приемный комплекс построен на основе приемной станции Сканекс УниСкан-3.7. В настоящее время с его помощью осуществляется прием спутниковых данных с космических аппаратов (КА) Terra, Aqua, Suomi NPP, NOAA-20 и FENG-YUN. Использование данных Suomi NPP и NOAA-20 позволяет существенно расширить возможности систем дистанционного мониторинга с привлечением новых данных повышенного пространственного разрешения и радиометрической чувствительности. Дополнительно для информационного обеспечения решаемых задач в ЕРЦ ДЗЗ началась работа по формированию локального архива актуальных снимков КА Landsat и Sentinel на территорию Красноярского края. В рамках работ, проводимых при взаимодействии с Сибирским региональным центром ДЗЗ в Железногорске, были выполнены работы по созданию архива спутниковых данных с отечественного аппарата Метеор-М2.

Для доступа к данным приемного комплекса и созданным архивам разработан каталог спутниковых данных, сформирован комплекс инструментов для обработки исходных данных и подготовки их для быстрого отображения посредством веб-браузера. Перед разработкой программного обеспечения были проанализированы существующие на сегодняшний день инструменты каталогизации спутниковых данных в мире и России [1]. Рассмотренные решения можно условно отнести к двум группам:

- простой инструмент для поиска по каталогу снимков с их просмотром в виде мелкомасштабных цветных растровых изображений;
- инструменты поиска и навигации с возможностью визуализации разных каналов снимка и их комбинаций, а также инструменты для анализа спутниковых данных.

Первая группа позволяет найти снимки в каталоге, а затем просмотреть их виде одного или нескольких обзорных изображений «quicklook» (небольшое изображение с упрощенной детализацией), либо позволяет посмотреть более детальное цветное изображение в естественных цветах на ограниченном масштабе. Вторая группа Интернет-сервисов позволяет просматривать детальные изображения, найденных снимков, а также позволяет управлять каналами снимков для создания тематических продуктов с искусственными цветами.

В результате анализа существующих продуктов было решено создать сервис доступа к каталогу спутниковых данных на территорию, охватываемую нашей приемной станцией в виде комбинации инструментов из двух рассмотренных групп. В настоящее время нашему приемному центру доступно файловое хранилище объемом порядка 100 Тб, что недостаточно для хранения всего спектра готовых продуктов. Для решения проблемы с ограниченным размером файлового хранилища, предварительно формируется архив из снимков с ограниченным числом каналов для визуализации и анализа спутниковых данных. Дополнительной задачей являлась разработка технологии для публикации в онлайн режиме данных из каталога снимков, т.к. внутренняя структура существующих решений слабо представлена в публикациях и Интернет.

Выбран подход, который позволяет экономить дисковое пространство и позволяет получить приемлемый результат. На основе исходных снимков ежедневно формируется два «базовых продукта» для данных с приборов MODIS КА Terra и Aqua, VIIRS КА Suomi NPP и NOAA-20 и данных загруженных архивов для прибора MSI КА Sentinel-2A и данные с КА Landsat-8:

- Многоканальное изображение с неполным набором каналов в отличие от исходного снимка, но достаточным списком каналов для создания основных продуктов с изображением в искусственных цветах. Для всех доступных в каталоге данных с космических аппаратов формируются такие комбинации, как цветосинтезированное изображение в «естественных цветах», комбинация каналов снимка, наиболее подходящая для обнаружения пожаров и огня и комбинация, используемая для отображения снега и льда. Дополнительно новое изображение содержит каналы для расчета индексов NDVI (нормализованный относительный индекс биомассы, Normalized Difference Vegetation Index) [2], NDWI (нормализованный разностный водный индекс, Normalized Difference Water Index) и температуры поверхности (при наличии соответствующего канала) [3].
- Одноканальное изображение для определения процента облачности. Так как в большинстве случаев снимки покрывают очень большую площадь и практически всегда имеют участки, покрытые облачностью на территории нашей страны, общая облачность всего снимка интереса не представляет. Такие изображения с грубой маской облачности позволяют осуществлять поиск снимков по каталогу по проценту облачности в пределах небольшой выбранной территории.

Разработан набор программ для создания небольших обзорных цветных изображений в естественных цветах в формате png для просмотра нескольких снимков одновременно. Такие изображения позволяют получить общую картину за сутки или несколько дней в области приема комплекса. Для дневных сцен используется определенный набор каналов для каждого прибора в диапазоне от 0 до 100% альбедо с гаммой 1,8. В ночное время используется одноканальное изображение сна основе одного канала в диапазоне от 220К до 300К с гаммой 0.5 в негативе.

Для прибора MODIS при создании обзорных изображений используется комбинация каналов 7-2-1, для многоканальных изображений используются следующие наборы каналов:

- 1-4-3 естественные цвета;
- 7-2-1 для обнаружения пожаров комбинация;
- 3-6-7 для отображения снега и льда комбинация;
- 31 температура поверхности;
- 1 и 2 для расчета индекса NDVI.

Для прибора VIIRS при создании обзорных изображений используется комбинация каналов M11-I2-I1, для многоканальных изображений используются следующие наборы каналов, наиболее близкие к прибору MODIS:

– I5-M4-M3 естественные цвета;
– M11-I2-I1 для обнаружения пожаров комбинация;
– M3-I3-M11 для отображения снега и льда комбинация;
– I5 температура поверхности;
– I1 и I2 для расчета индекса NDVI.

Для прибора MSI КА Sentinel-2A при создании обзорных изображений используется комбинация каналов 12-8-4, для многоканальных изображений используются следующие наборы каналов, наиболее близкие к прибору MODIS:

– 4-3-2 естественные цвета;
– 12-8-4 для обнаружения пожаров комбинация;
– 2-11-12 для отображения снега и льда комбинация;
– 4 и 8 для расчета индекса NDVI.

Температура поверхности для Sentinel-2A не использовалась.

В результате для быстрого просмотра обзорных изображений используются файлы в формате png, построенные на основе многоканальных изображений. В этой части разрабатываемая система работает, как представитель большинства существующих систем каталогизации спутниковых снимков. Для анализа спутниковых снимков в веб-браузере также используется многоканальное изображение, которое позволяет комбинировать в онлайн режиме разные доступные каналы. Дополнительно формируются продукты в виде одноканальных снимков, требующие расчета. Например, индекс NDVI, рассчитывающийся из данных красного (1 канал для MODIS, I1 канал для VIIRS, 4 канал для MSI) и инфракрасного каналов (2 канал для MODIS, I2 канал для VIIRS, 8 канал для MSI) по формуле (2-1)/(2+1) для прибора MODIS, (I2-I1)/(I2+I1) для прибора VIIRS, (8-4)/(8+4) для прибора MIS. Или температура, приведенная в градусы Цельсия из яркости пикселей соответствующего канала (рис. 1).

При формировании цветных изображений используется LUT (Look Up Table), это своеобразная «таблица поправок» для внесения изменений в каждый из трех каналов [4]. LUT позволяет изменять значение яркости точек изображения при преобразовании снимков с помощью линейной интерполяции и сохраняет общую контрастность для всех снимков. Дополнительно производится преобразование значение яркости пикселей из 16 бит к 8 битам.

Перед отображением снимков в веб-интерфейсе выполнялась настройка гаммы для программного обеспечения MapServer для 16 битных изображений с помощью таблиц LUT для 10 опорных точек. Без этих изменений растры выглядели намного темнее, приходилось использовать автоматический расчет гистограммы растра для каждого снимка, что в свою очередь приводило к некоторым артефактам при просмотре снимков в веб-интерфейсе и общая контрастность всех снимков отличалась, так как автоматический расчет выполнялся для каждого растра независимо.

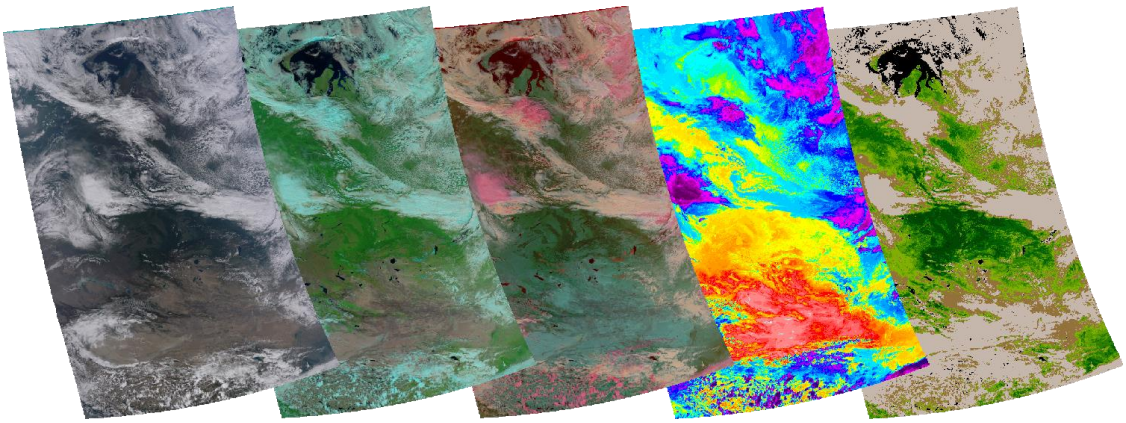


Рис. 1. Комбинации слоев (естественные цвета, пожары, снег и лед, температура, NDVI).

Обзорные изображения в формате png создаются с помощью библиотеки gdal с преобразованием формата и размера изображения. Применяется нелинейное масштабирование с помощью степенной функции для изменения яркости и контрастности снимка. Операция выполняется инструментом gdal_translate, предназначенным для преобразования растровых данных между различными форматами:

gdal_translate -of png -a_nodata 0 -co worldfile=yes -outsize xsize ysize -b b1 -b b2 -b b3 -exponent exp_val -scale 0 65535 0 255 file1 file2

Дополнительно для веб-интерфейса были созданы альтернативные карты подложки с применением проекции, более подходящей для территории Красноярского края. В частности, использовалась азимутальная равновеликая проекция Ламберта (код EPSG:3576) [5]. Эта же проекция используется для генерации «базовых продуктов», с целью исключить ресурсоемкую операцию перепроектирования при просмотре таких снимков.

В результате работы для быстрого и удобного поиска в каталоге спутниковых данных, для минимизации нагрузки на серверное программное и аппаратное обеспечение подготовлен набор серверных приложений для предварительной обработки спутниковых данных, включающих следующие этапы обработки:

1. Преобразование исходных растровых данных в формат GeoTiff с преобразованием исходной проекции в азимутальную равновеликую проекцию Ламберта.
2. Создание базового цветного изображения для разных масштабов отображения, состоящего из нескольких спектральных каналов. Такие изображения будут использоваться для детального просмотра спутникового снимка с сохранением исходного разрешения.
3. Создание растрового изображения в формате png для показа «quicklook» в веб-приложении для разных масштабов отображения, но с меньшим разрешением (до 1 км на точку). В отличие от базового цветного изображения «quicklook» в веб-приложении могут отображаться группами, в зависимости от выбора пользователя, тогда как базовое цветное изображение будет активно только для одного выбранного снимка.
4. Создание продуктов в виде растрового многоканального изображения в формате GeoTIFF с индексами NDVI и NDWI для КА Метеор-М №2 и одноканальных изображений с индексом NDVI для приборов MODIS, VIIRS и MSI.
5. Преобразование из 16-битного формата к 8-битному, более подходящему для показа спутниковых снимков в веб-приложении и требующее значительно меньших ресурсов в системе хранения данных.

В настоящее время разрабатывается набор программ для подготовки и публикации с помощью веб-каталога спутниковых данных ограниченного пользования с Российских космических аппаратов Ресурс-П и Канопус-В, получаемых от НЦ ОМЗ АО «Российские космические системы». Данные централизованно загружаются для нужд подразделений, входящих в ФИЦ КНЦ СО РАН, по общей заявке для всех подразделений. Для удобства конечных пользователей веб-каталога полученные данные будут обработаны и приведены к единому формату. Сейчас эти данные автоматически загружаются с геопортала Роскосмоса (**https://www.gptl.ru**) и предварительно обрабатываются в ручном режиме. Дополнительная полуавтоматическая обработка связана с тем, что полученные файлы исходных растровых снимков имеют различную иерархию папок, имена файлов отличаются по формату и не имеют данных о времени создания снимка (в имени снимка содержится только время его обработки оператором). Разные снимки имеют разный набор файлов растровых изображений, файлов привязки и метаданных в XML формате. XML формат метаданных отличается для разных уровней обработки полученных снимков. Имена файлов могут содержать русские символы, что вносит ряд проблем при работе с ними в Unix системах. Разрабатываемая система будет иметь интерактивные инструменты для решения проблем, связанных с ошибками подготовки снимков операторами. Т.к. бывают ситуации, когда файлы попадают в ошибочные папки на ftp сервере, имена файлов задаются некорректно и эти ошибки исправляются операторами только в следующих выгруз-

ках. Разрабатываемая система будет пытаться обрабатывать данные автоматически, но в ситуациях, когда это невозможно сделать, будет отсылать сообщение администратору системы для принятия решения и ручного внесения исправлений.

Для поиска и навигации по каталогу спутниковых данных разработано веб-приложение с применением серверного языка программирования PHP, СУБД PostgreSQL, SQLite3 и клиентского веб-интерфейса на языке программирования TypeScript с использованием фреймворка Angular 5.

Анализ существующих сегодня технологий и программного обеспечения предназначенных для работы с пространственными данными в среде Интернет, многолетний опыт коллектива авторов и разработанных ими программных средств привели к использованию открытых технологий и программного обеспечения. В качестве основы использовались разработанные программные средства для анализа пространственных данных в среде геопортала Института вычислительного моделирования СО РАН с использованием технологий, предлагаемых международной организацией OGC (Open Geospatial Consortium) и программного обеспечения MapServer и MapProxy. MapProxy используется для создания карты из фрагментов. Спутниковые снимки для веб-приложения формируются с использованием программного обеспечения MapServer и GDAL. Набор инструментов для предварительной обработки снимков разработан с использованием языка Python 3 и процессора BASH в среде Unix. Основными модулями для приложений на Python выступали библиотеки GDAL и NumPy. Пример веб-интерфейса каталога спутниковых снимков представлен на рис. 2.

Разработанная платформа для каталога спутниковых снимков обеспечивает пользователя современными инструментами для работы со спутниковыми данными и навигацией среди них в рамках приемного комплекса регионального центра в Красноярском крае и соседних регионах. Возможности создаваемого программного продукта позволяют комбинировать любые сочетания каналов, доступных в изображении без дополнительной настройки серверного программного обеспечения, что выделяет созданную систему среди подобных систем каталогизации спутниковых данных. Разработанные технологии и алгоритмы позволяют внедрять элементы системы для других архивов спутниковой информации. В ближайшем будущем разработанная платформа позволит создавать новые специализированные продукты на основе спутниковых данных в сотрудничестве с другими институтами Сибирского отделения РАН, которые могут быть использованы для анализа и мониторинга различных процессов.

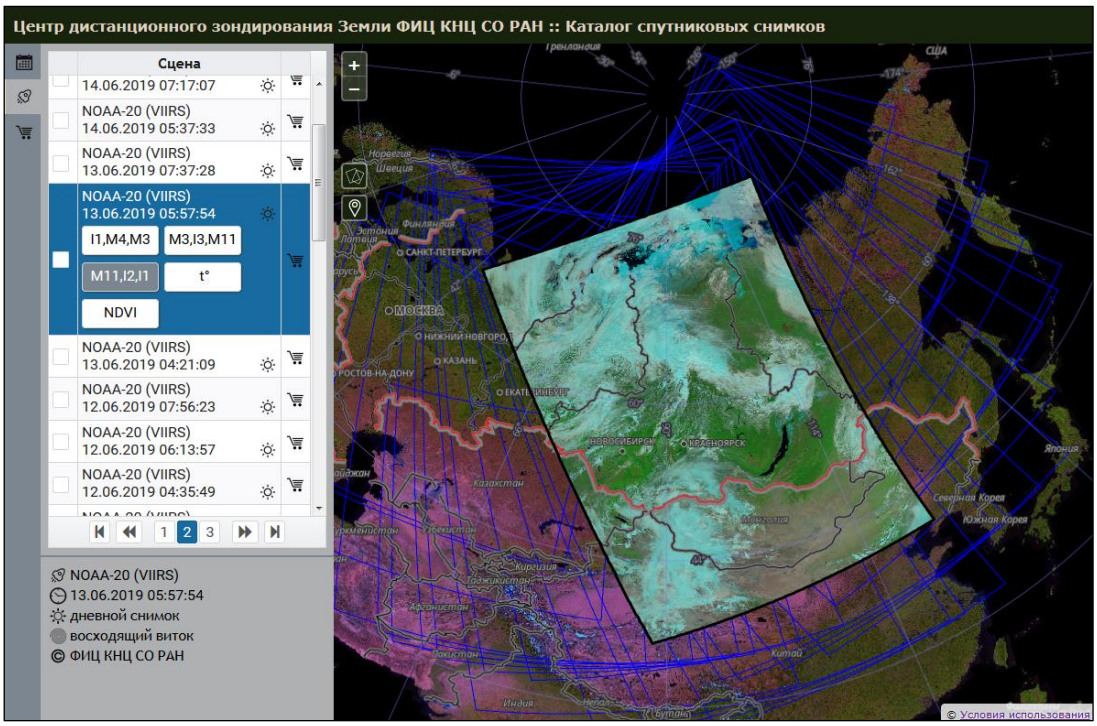


Рис. 2. Веб-интерфейс каталога спутниковых снимков.

## ЛИТЕРАТУРА


[1] *Grill S., Jedlicka J., Schneider M., Stefanova E.* Archive and catalogue system for receiving satellite data as a part of academic SDI // Imagin [e,g] Europe. Proceedings of the 29th Symposium of the European Association of Remote Sensing Laboratories. 2010. P. 150-157.

[2] *Шукилович, А.Ю., Федотова, Е.В., Маглинец, Ю.А.* Применение сенсора MODIS для оперативного мониторинга земель сельскохозяйственного назначения // Журнал Сибирского федерального университета. Серия: Техника и технологии. 2016. № 9 (7). С. 1035-1044.

[3] *Пчельников Д.В., Добрецов Н.Н.* Построение временных рядов с одновременным использованием данных дистанционного зондирования SUOMI NPP VIIRS и TERRA/AQUA MODIS // Материалы Международной конференции «ИнтерКарто/ИнтерГИС». 2017. № 23 (3). С. 46-51.

[4] *Li Gf., Li Gj., Han G.* Enhancement of low contrast images based on effective space combined with pixel learning // Information. 2017. Vol. 8. P. 135.

[5] *Кадочников А.А.* Особенности построения геопространственных веб-приложений и сервисов для систем мониторинга состояния окружающей природной среды // Журнал Сибирского федерального университета. Серия: Техника и технологии. 2015. № 8 (7). С. 908-916.

# МОДЕЛИРОВАНИЕ И МОНИТОРИНГ ЭКОЛОГИЧЕСКИХ И ТЕХНОГЕННЫХ ПРОЦЕССОВ И СИСТЕМ

# КРИТЕРИИ КАЧЕСТВА АТМОСФЕРНОГО ВОЗДУХА ГОРОДСКОЙ СРЕДЫ ДЛЯ ВЕЩЕСТВ КАНЦЕРОГЕННОГО ДЕЙСТВИЯ

*Бельская Е.Н.[(1)], Иванова У.С.[(2)], Тасейко О.В.[(1)(2)]*
[(1)] Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева, Красноярск
[(2)] Институт вычислительных технологий СО РАН, Красноярск

Проведена количественная оценка индивидуального риска здоровью населения, при воздействии канцерогенных химических веществ, с использованием их фактических концентраций; предложены расчетные значения предельно-допустимых концентраций приоритетных веществ канцерогенного действия, с учетом уровней приемлемого риска для разных категорий населения.
*Ключевые слова: загрязнение атмосферного воздуха, оценка риска, предельно-допустимая концентрация, вещества канцерогенного действия.*

Загрязнение окружающей среды – один из значимых факторов, обусловливающих ухудшение здоровья и высокую смертность населения [1-3]. В России в экологически неблагоприятных регионах проживает около 50% населения, и загрязнение атмосферного воздуха является причиной до 17,5% от общей смертности городского населения [1, 4]. В городах Европы загрязненный атмосферный воздух является причиной до 23 тысяч дополнительных смертей в год [5].

Критериями качества атмосферного воздуха в городах России на сегодняшний день служат предельно-допустимые концентрации (ПДК), являющиеся гигиеническими нормативами. В Российской Федерации утверждено более 1000 нормативов ПДК и это наиболее обширная из существующих систем нормирования качества воздушной среды [6]. Система критериев, определяющая безопасное состояние окружающей среды, в том числе ПДК, несовершенна и не позволяет определять меру воздействия на человека и ущерб, наносимый здоровью населения, в случае превышения установленных нормативов [7]. Альтернативу концепции ПДК составляет система оценки рисков. Для оценки рисков вводится ряд параметров - маркеров риска, позволяющих оценивать опасность появления симптомов токсичности при кратковременном, продолжительном либо хроническом воздействии (экспозиции) определенных уровней загрязнения (концентрации загрязняющих веществ) для разных групп населения.

Согласно «Руководства по оценке риска для здоровья населения при воздействии химических веществ, загрязняющих окружающую среду. Р 2.1.10.1920-04» (утв. Главным государственным санитарным врачом РФ 05.03.2004) [8] оценка риска для здоровья человека – это количественная и/или качественная характеристика вредных эффектов, способных развиться в результате воздействия факторов среды обитания человека на конкретную группу людей при специфических условиях экспозиции.

В настоящее время концепция оценки риска практически во всех странах мира и международных организациях рассматривается в качестве главного механизма разработки и принятия управленческих решений как на международном, государственном или региональном уровнях, так и на уровне отдельного производства или другого потенциального источника загрязнения окружающей среды [8]. Несмотря на очевидные преимущества методологии оценки риска, позволяющей учесть, в том числе, беспороговость воздействия веществ канцерогенного действия на здоровье населения, при реализации программ мониторинга качества атмосферного воздуха в населенных пунктах обойтись без использования предельно-допустимых концентраций сложно. Поэтому возникает необходимость учета рисков здоровью при определении уровней гигиенических ПДК.

В работе решены следующие задачи:

- на примере некоторых городов Сибирского федерального округа выполнен расчет нормативов качества атмосферного воздуха на основе концепции методологии оценки ингаляционного риска здоровью населения с использованием фактических среднегодовых концентраций приоритетных вредных веществ канцерогенного действия;
- предложены расчетные ПДК приоритетных веществ канцерогенного действия с учетом уровней приемлемого риска для разных категорий населения.

Использованные для расчета индивидуального канцерогенного риска фактические среднегодовые концентрации приоритетных канцерогенных веществ в городах: Красноярске, Кемерово, Омске и Черногорске, получены с постов наблюдения за загрязнением атмосферы Среднесибирского, Западно-Сибирского и Обь-Иртышского УГМС и Федеральной службы по гидрометеорологии и мониторингу окружающей среды [10] за 2008-2018 гг., и с постов наблюдения за загрязнением атмосферы Хакасским центром по гидрометеорологии и мониторингу филиала ФГБУ «Среднесибирское УГМС» за 2010-2017 гг.

На рис. 1 приведена динамика среднегодовых концентраций бенз(а)пирена и формальдегида за 2010-2017 гг для четырех городов СФО.

Наибольшее содержание бенз(а)пирена наблюдается в городах Красноярск и Черногорск (республика Хакассия), высокое содержание формальдегида в гг. Красноярск и Кемерово. При этом во всех рассмотренных городах содержание этих веществ превышает годовые нормативы качества атмосферного воздуха населенных мест.

Единичный риск, согласно [8], рассчитывается по формуле 1, с использованием значений факторов канцерогенного потенциала *SF* и стандартных значений массы тела человека (70 кг), суточного потребления воздуха (20 м$^3$/сут.)

$$UR_i = SF_i \cdot \frac{1}{MT} \cdot ДП = SF_i \cdot \frac{1}{70} \cdot 20, \quad (1)$$

Расчет индивидуального канцерогенного риска осуществляется по формуле 2, с использованием данных о величине экспозиции и значениях единичного риска. Как правило, для канцерогенных химических веществ дополнительная вероятность развития рака у индивидуума на всем протяжении жизни (*ИКР*)оценивается, при использовании величины единичного риска (*UR*), следующим образом:

$$ИКР = C \cdot UR, \quad (2)$$

где *C* – средняя концентрация вещества в исследуемом объекте окружающей среды за весь период усреднения экспозиции (воздух, мг/м$^3$); *UR* – единичный риск для воздуха (риск на 1мг/м$^3$).

В табл. 1 приведены результаты расчета *ИКР* по усредненным среднегодовым концентрациям канцерогенных химических веществ.

Индивидуальный риск развития канцерогенных эффектов считается приемлемым, если его значения не превышают $10^{-6}$ год$^{-1}$ для населения (один пострадавший на миллион населения) [8] и для профессиональных групп от $10^{-4}$ год$^{-1}$ (для функционирующих объектов) до $10^{-5}$ год$^{-1}$ (для вновь строящихся объектов) [9].

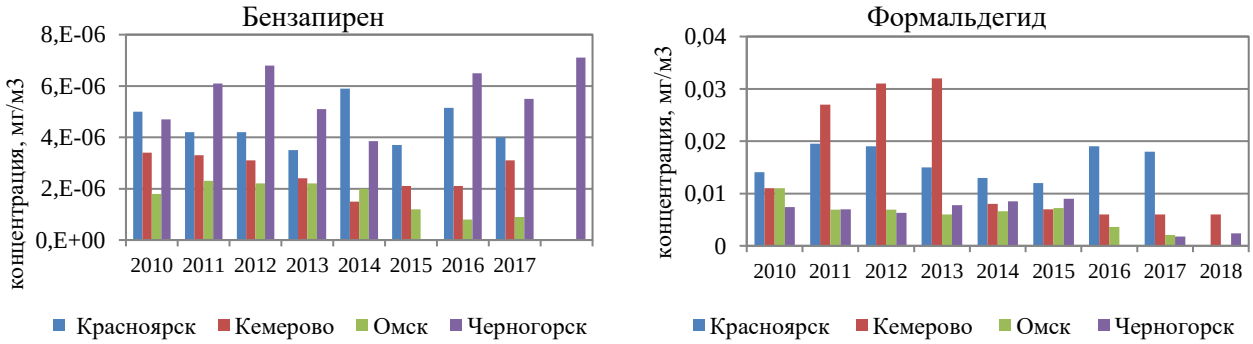



Рис. 1. Среднегодовые концентрации бенз(а)пирена и формальдегида в некоторых городах СФО.

Табл. 1. Индивидуальный канцерогенный риск.

| год/ вещество | Красноярск | | | | | Кемерово | | |
|---|---|---|---|---|---|---|---|---|
| | Бенз(а)пирен | Формальдегид | Свинец | Этилбензол | Бензол | Бенз(а)пирен | Формальдегид | Анилин |
| **2010** | **$5{,}57\cdot10^{-6}$** | $1{,}83\cdot10^{-4}$ | **$1{,}44\cdot10^{-6}$** | $1{,}80\cdot10^{-5}$ | $2{,}08\cdot10^{-4}$ | **$3{,}79\cdot10^{-6}$** | $1{,}43\cdot10^{-4}$ | $<1{,}6\cdot10^{-3}$ |
| **2011** | **$4{,}68\cdot10^{-6}$** | $2{,}54\cdot10^{-4}$ | **$2{,}04\cdot10^{-6}$** | $1{,}70\cdot10^{-5}$ | $1{,}84\cdot10^{-4}$ | **$4{,}29\cdot10^{-6}$** | $3{,}15\cdot10^{-4}$ | $<1{,}6\cdot10^{-3}$ |
| **2012** | **$4{,}68\cdot10^{-6}$** | $2{,}47\cdot10^{-4}$ | **$2{,}04\cdot10^{-6}$** | $1{,}70\cdot10^{-5}$ | $2{,}32\cdot10^{-4}$ | **$3{,}45\cdot10^{-6}$** | $4{,}03\cdot10^{-4}$ | $1{,}6\cdot10^{-3}$ |
| **2013** | **$3{,}90\cdot10^{-6}$** | $1{,}95\cdot10^{-4}$ | - | $1{,}70\cdot10^{-5}$ | $2{,}72\cdot10^{-4}$ | **$2{,}67\cdot10^{-6}$** | $4{,}16\cdot10^{-4}$ | $<1{,}6\cdot10^{-3}$ |
| **2014** | **$6{,}57\cdot10^{-6}$** | $1{,}59\cdot10^{-4}$ | - | $1{,}40\cdot10^{-5}$ | $1{,}68\cdot10^{-4}$ | **$1{,}67\cdot10^{-6}$** | $1{,}04\cdot10^{-4}$ | $1{,}6\cdot10^{-3}$ |
| **2015** | **$4{,}12\cdot10^{-6}$** | $1{,}56\cdot10^{-4}$ | - | **$7{,}00\cdot10^{-6}$** | $1{,}92\cdot10^{-4}$ | **$2{,}34\cdot10^{-6}$** | $0{,}91\cdot10^{-4}$ | $<1{,}6\cdot10^{-3}$ |
| **2016** | **$5{,}74\cdot10^{-6}$** | $2{,}47\cdot10^{-4}$ | - | **$3{,}00\cdot10^{-6}$** | $9{,}60\cdot10^{-5}$ | **$2{,}34\cdot10^{-6}$** | $0{,}78\cdot10^{-4}$ | $1{,}6\cdot10^{-3}$ |
| **2017** | **$4{,}64\cdot10^{-6}$** | $2{,}34\cdot10^{-4}$ | - | - | - | **$3{,}45\cdot10^{-6}$** | $0{,}78\cdot10^{-4}$ | $1{,}60\cdot10^{-3}$ |
| **2018** | - | - | - | - | - | **$3{,}79\cdot10^{-6}$** | $0{,}78\cdot10^{-4}$ | 0 |

| год/ вещество | Черногорск | | Омск | |
|---|---|---|---|---|
| | Бенз(а)пирен | Формальдегид | Бенз(а)пирен | Формальдегид |
| **2008** | **$5{,}24\cdot10^{-6}$** | $9{,}62\cdot10^{-5}$ | **$2{,}01\cdot10^{-6}$** | $1{,}43\cdot10^{-4}$ |
| **2009** | **$6{,}80\cdot10^{-6}$** | $9{,}10\cdot10^{-5}$ | **$2{,}56\cdot10^{-6}$** | $8{,}97\cdot10^{-5}$ |
| **2010** | **$7{,}58\cdot10^{-6}$** | $8{,}19\cdot10^{-5}$ | **$2{,}45\cdot10^{-6}$** | $8{,}97\cdot10^{-5}$ |
| **2011** | **$5{,}68\cdot10^{-6}$** | $1{,}01\cdot10^{-4}$ | **$2{,}45\cdot10^{-6}$** | $7{,}80\cdot10^{-5}$ |
| **2012** | **$4{,}29\cdot10^{-6}$** | $1{,}11\cdot10^{-4}$ | **$2{,}23\cdot10^{-6}$** | $8{,}58\cdot10^{-5}$ |
| **2013** | - | $1{,}17\cdot10^{-4}$ | **$1{,}34\cdot10^{-6}$** | $9{,}36\cdot10^{-5}$ |
| **2015** | **$6{,}50\cdot10^{-6}$** | - | - | - |
| **2016** | **$6{,}13\cdot10^{-6}$** | $2{,}34\cdot10^{-5}$ | **$8{,}91\cdot10^{-7}$** | $4{,}68\cdot10^{-5}$ |
| **2017** | **$7{,}91\cdot10^{-6}$** | $3{,}12\cdot10^{-5}$ | **$1{,}00\cdot10^{-7}$** | $2{,}73\cdot10^{-5}$ |

«-» – нет данных; выделенные значения рисков превышают уровень приемлемого риска для населения.

Риск здоровью населения от ингаляционного воздействия, превышающий приемлемый уровень, формируется под воздействием бенз(а)пирена, этилбензола и свинца во всех рассмотренных городах СФО. В целом для одних и тех же веществ уровни индивидуальных канцерогенных рисков в разных городах отличаются несущественно.

Учитывая уровни приемлемого индивидуального риска развития канцерогенных эффектов $10^{-6}$ год$^{-1}$ для населения и $10^{-4}$ год$^{-1}$ для профессиональных групп, было получено соотношение 3 для определения ПДК вышерассмотренных веществ:

$$ПДК_{расчетная} = ИКР \cdot \frac{МТ}{SF_i \cdot ДП} \qquad (3)$$

Результаты расчетов $ПДК_{расчетная}$ приведены в табл. 2.

В табл. 2 приведены основные показатели опасности развития канцерогенных эффектов – сведения, установленные различными гигиеническими нормативными документами: единичный риск для ингаляционного пути поступления, среднесуточные и максимально-разовые предельно-допустимые концентрации, референтные концентрации для вышерассмотренных веществ, также установлены пересмотренные значения ПДК для населения и профессиональных групп, с учётом приемлемости риска.

Нормативы содержания в атмосферном воздухе загрязняющих веществ с учетом их влияния на здоровье населения, определенные в соответствии с методологией оценки рисков здоровью населения, меньше ПДК, установленных для соответствующих веществ на 1-3 порядка. Наибольшее отличие действующего уровня ПДК от расчетного, основанного на непосредственном влиянии вещества на здоровье человека, выявлено для бензола. При этом, бензол относится к приоритетным веществам при изучении влияния окружающей среды на здоровье

населения. Недооценка уровней загрязнения атмосферного воздуха городов этим веществом может приводить к росту количества патологических состояний, связанных с его воздействием, таких как заболевания крови и онкологические заболевания [12].

Другое отличие предлагаемых нормативов от действующих ПДК в том, что применяемые в нашей стране предельно допустимые концентрации не всегда обоснованы по эффектам, напрямую связанными со здоровьем: 53% ПДК для водных объектов (ПДК в.в.) обоснованы по органолептическому признаку вредности и 20% – по общесанитарному показателю; 30% ПДК для атмосферного воздуха населенных мест установлена по рефлекторным реакциям человека [13]. В то время как оценка рисков, используемая в данном подходе, основывается исключительно на критериях, отражающих непосредственное влияние химических веществ на здоровье наиболее чувствительных групп населения. Кроме того, регламентированные ПДК относятся к нормативам среднесуточным, предложенные же нормативы являются стандартами среднегодовыми.

Табл. 2. Сведения о показателях опасности развития канцерогенных эффектов и пороговые значения концентраций [8, 10, 11].

| **Вещество/ показатель** | **Бенз(а)пирен** | **Формальдегид** | **Свинец** | **Этилбензол** | **Бензол** | **Анилин** |
|---|---|---|---|---|---|---|
| **МАИР** | 2A | 2A | 2A | 2B | 1 | 3 |
| **EPA** | B2 | B1 | B2 | D | A | B2 |
| **Фактор канцерог. потенциала** (ингаляционно) *SFi*, $((кг·сут)/мг)^{-1}$ | 3,9 | $4{,}6\cdot10^{-2}$ | $4{,}2\cdot10^{-2}$ | $3{,}9\cdot10^{-3}$ | $2{,}7\cdot10^{-2}$ | $5{,}7\cdot10^{-3}$ |
| **Единичный риск** (ингаляционно) *URi*, ($м^3$/мг) | 1,114 | $1{,}3\cdot10^{-2}$ | $1{,}2\cdot10^{-2}$ | $1{,}0\cdot10^{-3}$ | $8{,}0\cdot10^{-3}$ | $1{,}6\cdot10^{-2}$ |
| **Класс опасности** | 1 | 2 | 1 | 3 | 2 | 2 |
| **$ПДК_{сс}$** (среднесуточная), ($мг/м^3$) | $1{,}0\cdot10^{-6}$ | $3{,}0\cdot10^{-3}$ | $3{,}0\cdot10^{-4}$ | $2{,}0\cdot10^{-3}$ | 0,1 | $3{,}0\cdot10^{-2}$ |
| **$ПДК_{мр}$** (маскимально-разовая), ($мг/м^3$) | - | $3{,}5\cdot10^{-2}$ | $1{,}0\cdot10^{-3}$ | $4{,}0\cdot10^{-2}$ | 0,3 | $5{,}0\cdot10^{-2}$ |
| **Референтные концентрации** для хронич. ингаляцион. воздействия (мг/кг) | $1{,}0\cdot10^{-6}$ | $3{,}0\cdot10^{-3}$ | $5{,}0\cdot10^{-4}$ | 1 | $3{,}0\cdot10^{-2}$ | $1{,}0\cdot10^{-3}$ |
| **ПДК расчетная для населения** ($мг/м^3$) | $8{,}97\cdot10^{-7}$ | $7{,}61\cdot10^{-5}$ | $8{,}3\cdot10^{-5}$ | $9{,}1\cdot10^{-4}$ | $1{,}3\cdot10^{-4}$ | $6{,}1\cdot10^{-4}$ |
| **ПДК расчетная для профессиональных групп** ($мг/м^3$) | $8{,}97\cdot10^{-5}$ | $7{,}61\cdot10^{-3}$ | $8{,}3\cdot10^{-3}$ | $9{,}1\cdot10^{-2}$ | $1{,}3\cdot10^{-2}$ | $6{,}1\cdot10^{-2}$ |

МАИР – классификация Международного агентства по изучению рака (категория 1 – канцерогенные для человека; 2A – весьма вероятно канцерогенные для человека; 2B – вероятно канцерогенные для человека; 3 – невозможно классифицировать с точки зрения канцерогенности для человека); EPA – классификация степени доказанности канцерогенности для человека U.S. EPA (A – канцерогены для человека; B1 – вероятные канцерогены для человека (ограниченные доказательства для человека); B2 – вероятные канцерогены для человека (достаточные доказательства для животных и недостаточные доказательства или отсутствие данных для человека); D – не классифицируемые как канцерогены для человека); SFi – факторы канцерогенного потенциала для ингаляционного пути поступления $((кг·сут)/мг)^{-1}$; URi – единичный риск при ингаляционном воздействии на 1 $мг/м^3$.

Сравнение предложенных нормативов среднегодовых ПДК с наблюдаемым уровнем загрязнения в промышленных центрах Сибири, показывает существенное превышение содержания основных индикаторов канцерогенной опасности атмосферного воздуха для здоровья населения, таких как бенз(а)пирен, формальдегид, свинец, бензол и этилбензол.

В ходе дальнейших исследований будут определены расчетные среднегодовые нормативы ПДК с учетом реального влияния загрязнения на здоровье населения для других веществ канцерогенного и неканцерогенного действия, поступающих в организм человека не только ингаляционным, но и пероральным путем. Кроме того, предложенный подход может применяться для пересмотра производственных нормативов ПДВ и уточнения сумм ущербов, связанных с воздействием загрязнения атмосферного воздуха на здоровье населения.

*Исследование выполнено при финансовой поддержке РФФИ, Правительства Красноярского края, Красноярского краевого фонда науки в рамках научного проекта № 18-47-240006: «Методы и информационные технологии оценки рисков развития социально-природно-техногенных систем промышленного региона».*

## ЛИТЕРАТУРА

[1] *Акимов В.А., Лапин В.Л., Попов В.М. и др.* Надежность технических систем и техногенный риск. М.: ЗАО ФИД «Деловой экспресс», 2002. 368 с.

[2] *Акимов В.А., Лесных В.В., Радаев Н.Н.* Основы анализа и управления риском в природной и техногенной сферах. М.: ЗАО ФИД «Деловой экспресс», 2004. 352 с.

[3] *Алымов В.Т., Крапчатов В.П., Тарасова Н.П.* Анализ техногенного риска. М.: Круглый год, 2000. 160 с.

[4] *Балдин К.В., Воробьев С.Н.* Управление рисками. М.: ЮНИТИ-ДАНА, 2005. 511 с.

[5] Безопасность России. Правовые, социально-экономические и научно-технические аспекты. Региональные проблемы безопасности. Красноярский край. М.: Знание, 2001. 576 с.

[6] *Тасейко О.В., Михалюта С.В., Леженин А.А.* Обоснование нормативов качества атмосферного воздуха в городе // Экология и промышленность России. 2013. С. 56-61.

[7] Безопасность России. Функционирование и развитие сложных народохозяйственных, технических, энергетических, транспортных систем, систем связи и коммуникаций. М.: Знание, 1998. 448 с.

[8] Руководство по оценке риска для здоровья населения при воздействии химических веществ, загрязняющих окружающую среду. Руководство. Р 2.1.10.1920-04 (утв. Главным государственным санитарным врачом РФ 05.03.2004).

[9] Декларация Российского научного общества анализа риска «Об установлении предельно допустимого уровня риска». **http://www.sra-russia.ru** (дата обращения: 02.05.2018).

[10] Федеральная служба по гидрометеорологии и мониторингу окружающей среды. Среднегодовые концентрации бенз(а)пирена в городах на территории России. **http://voeikovmgo.ru/?option=com_content&view=article&id=683:27042015aq3-2&catid=15&Itemid=11&lang=ru** (дата обращения: 11.04.2019).

[11] ГН 2.1.6.3492-17 «Предельно допустимые концентрации (ПДК) загрязняющих веществ в атмосферном воздухе городских и сельских поселений» (с изменениями на 31 мая 2018 года). **http://docs.cntd.ru/document/556185926** (дата обращения: 02.05.2018).

[12] Руководство по комплексной профилактике экологически обусловленных заболеваний на основе оценки риска. М.: 2017 г. 68 стр. **https://static-3.rosminzdrav.ru/system/attachments/attaches/000/036/600/original/Руководство.pdf?1512992582** (дата обращения: 20.06.2019).

[13] МосМР 2.1.9.004-03 «Критерии оценки риска для здоровья населения приоритетных химических веществ, загрязняющих окружающую среду». **http://docs.cntd.ru/document/3715847** (дата обращения: 20.06.2019).

# ОБРАБОТКА ДАННЫХ НАТУРНЫХ НАБЛЮДЕНИЙ ПОВЕРХНОСТНОЙ СКОРОСТИ И ТЕМПЕРАТУРЫ В ОЗЕРЕ ШИРА

*Володько О.С.[(1)(2)], Компаниец Л.А.[(1)(2)], Толомеев А.П.[(3)], Дроботов А.В.[(3)], Гаврилова Л.В.[(2)]*
[(1)] Ииститут вычислительного моделирования СО РАН, Красноярск
[(2)] Сибирский федеральный университет, Красноярск
[(3)] Институт биофизики СО РАН, Красноярск

В 2013-2018 гг. для озера Шира были получены данные длительных натурных измерений скорости течения, скорости ветра и температуры. В 2018 г. были проведены первые исследования поверхностного течения с использованием лагранжевого дрифтера. В данной работе для выявления основных составляющих приповерхностного течения и поверхностной температуры использовался метод главных компонент. Данные, полученные с использованием дрифтера, анализировались с точки зрения их связи с ветровой картиной.
*Ключевые слова: натурные измерения, дрифтер, метод главных компонент.*

**Введение.** В настоящее время накопилось большое количество информации относительно гидрофизических характеристик озера Шира. Получены данные о температуре с использованием термоплетей, данные о скоростях с использованием ADCP, поверхностного течения с использованием дрифтера. Структура полученных данных чрезвычайно сложна и для выделения основных процессов необходимо использовать специальные методы. Анализ поверхностной температуры и приповерхностной скорости проводился с использованием метода главных компонент (эмпирических ортогональных функций). После выделения главных компонент для данных поверхностной температуры (2014, 2015, 2018 г.) подсчитывалась корреляция между первым модальным коэффициентом и температурой воздуха. Для скорости в приповерхностном слое первая мода сравнивалась с распределением скорости стационарного течения однородной жидкости и на этой основе определялся вид коэффициента вертикального турбулентного обмена.

**Анализ поверхностной скорости с использованием дрифтера.** Исследования с использованием дрифтеров позволяет выявить, в отличие от точечных измерений скорости, вихревую структуру течения.

Как известно, вихри являются механизмом как горизонтального перемешивания, приводящего к однородному горизонтальному распределения фито- и зоопланктона, так и вертикального, что приводит к переносу питательных веществ из глубоководной зоны в прибрежную. Лагранжевы дрифтеры широко используются для изучения течений в океане и в больших озерах. Например, были проведены исследования прибрежного течения в Северной Каролине [1], ветрового течения в северной части Тихого океана [2], приповерхностных циркуляций в Северных морях [3], поверхностных циркуляций в Северной Атлантике [4] и в Бискайском заливе [5], известны исследования для Великих американских озер [6, 7], озера Киннерет [8] и озера Штехлин [9].

В летний период 2018 г. Институтом Биофизики СО РАН на озере Шира (Республика Хакассия, Россия) проведены измерения скоростей течения с помощью спутникового квазилагранжевого дрифтера.

На рис. 1 сплошной линией обозначена траектория движения дрифтера во время эксперимента, точками на линии обозначены моменты времени, когда ветер менял направление. Начало эксперимента 12/07/18 в 7:22, продолжительность — 23 часа. Средняя скорость движения дрифтера 9 см/с.

Для объяснения структуры течения, обозначенной дрифтером, был проведен анализ ветровой картины в районе местоположения озера Шира. Рассматривались данные для направления и скорости ветра с нескольких метеостанций, расположенных вокруг озера Шира (рис. 2).

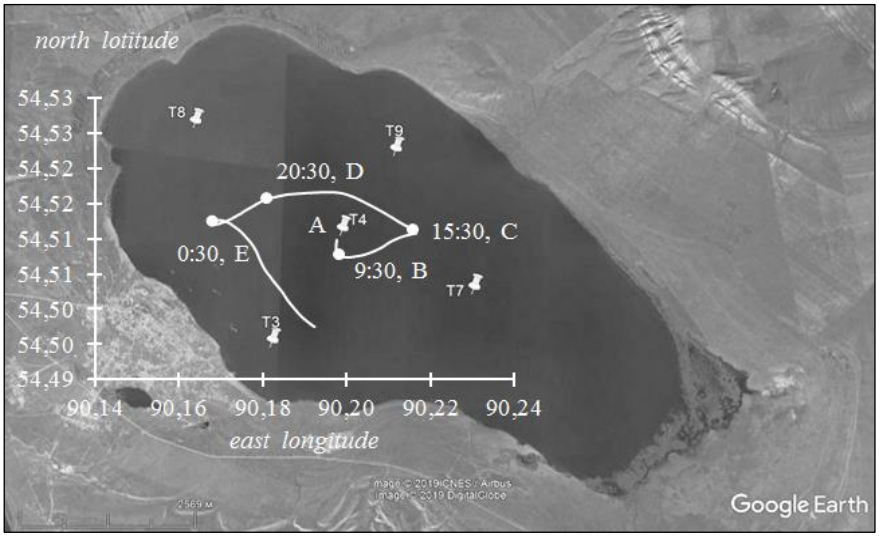



Рис. 1. Дрифтерный эксперимент 12-13 июля 2018 г. Траектория движения дрифтера.

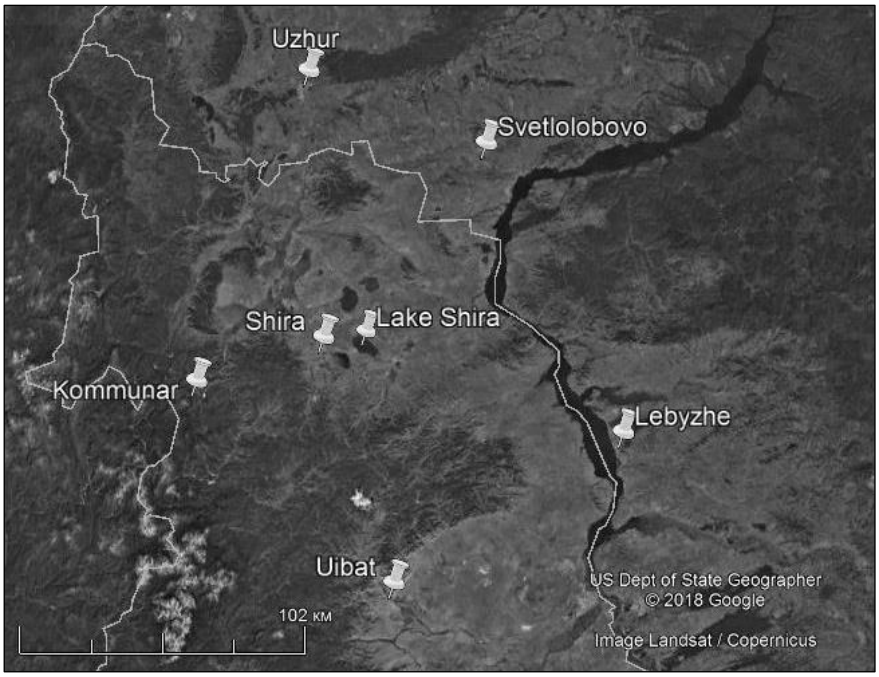



Рис. 2. Метеостанции, расположенные вокруг оз. Шира.

Для анализа были выбраны три станции, расположенные в вершинах треугольника, охватывающего озеро. Рассматривался «треугольник»: Ужур – Лебяжье – Уйбат и проводилась линейная интерполяция внутри этого треугольника.

На рис. 3 представлено поле скоростей ветра, построенное по данным трех метеостанций (1 – Ужур; 2 – Лебяжье; 3 – Уйбат) на разные моменты времени, которые соответствуют времени проведения дрифтерного эксперимента. Цифрой 4 и прямоугольником отмечено местоположение оз. Шира.

Результаты расчетов показывают, что ветер имеет циклоническую завихренность во все моменты времени, за исключением 12/07/2018 22:00, что соответствует траектории движения дрифтера.

Практически во все время измерения ветер провоцировал циклонический вихрь, что полностью соответствует рис. 1. Изменение направленности вихревого движения в ночь на 13 июня согласовано с изменением направления ветра (рис. 1 и правый верхний график на рис. 3).

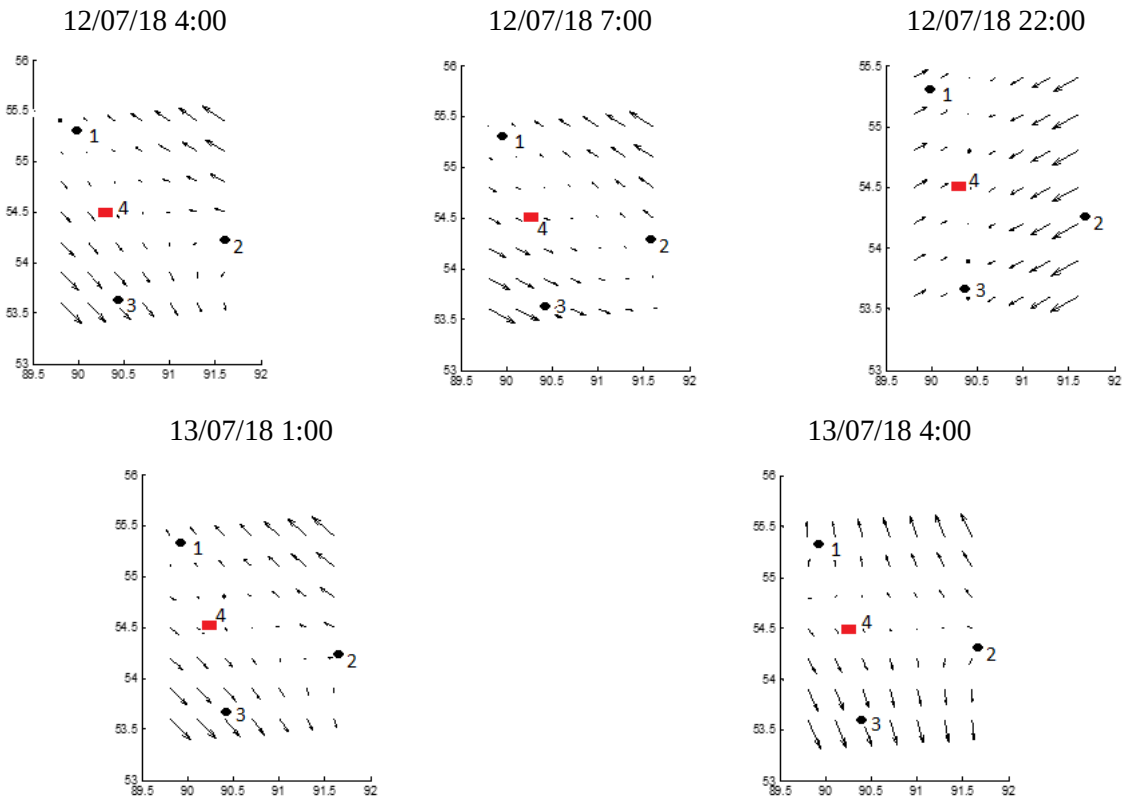



Рис. 3. Скорость ветра в районе озера Шира.

Анализ приповерхностной скорости. В 2014-2015 гг. в озере осуществлялись длительные измерения скорости, которые потом обрабатывались с использованием метода главных компонент, являющимся одним из вариантов статистической обработки данных. После использования этого метода данные натурных наблюдений заисываются в виде конечной суммы слагаемых разного масштаба, представляющих произведение функции, зависящей от пространственных переменных (моды) на зависящие от времени модальные коэффициенты.

В результате применения метода главных компонент к измерениям в летний период 2014 г. получаем первую моду – комплекснозначный вектор, не зависящий от времени, размерность которого равна числу точек измерения по пространству.

Известно, что в летний период озеро является стратифицированным по температуре и солености.

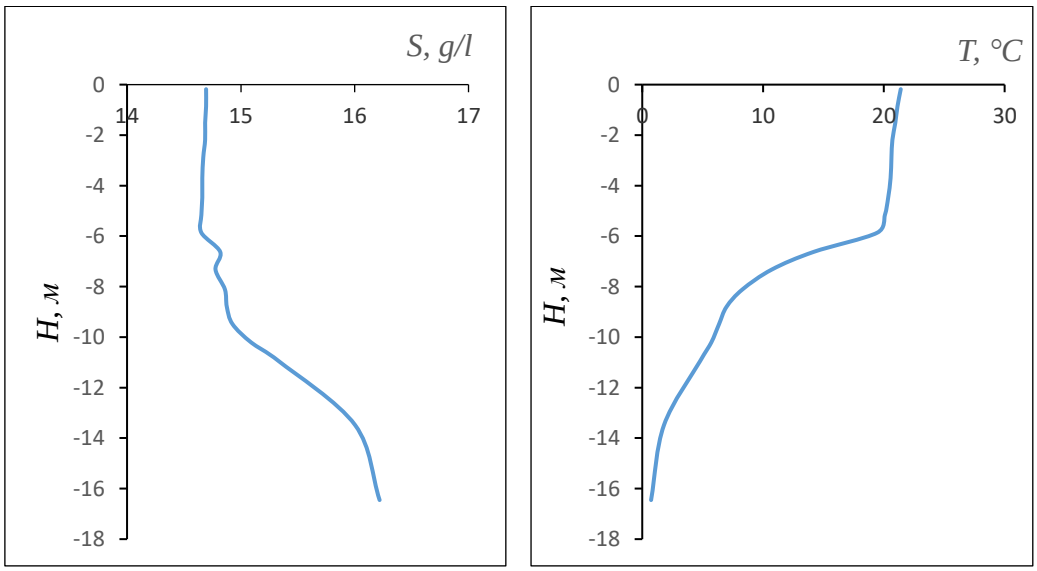



Рис. 4. Соленость и температура в озере Шира летний период 2013-2014 гг.

Верхний перемешанный слой можно рассматривать как слой, в котором жидкость однородная. Это дает основание, чтобы в этом слое на глубине до 9 м сравнить первую моду, полученную при использовании метода главных компонент, с известными решениями для стационарного течения однородной жидкости.

Для сравнения бралось аналитическое решение для дрейфовой составляющей модели Экмана [10, 11].

Решение для постоянного коэффициента вертикального турбулентного обмена $K_z$ имеет вид

$$w = u + iv = \frac{\tau^w}{\rho_0} \cdot \frac{\mathrm{ch}(\alpha(z+H)) + \dfrac{k_b}{K_z\alpha}\mathrm{sh}(\alpha(z+H))}{K_z\alpha\,\mathrm{sh}(\alpha H) + k_b\mathrm{ch}(\alpha H)}.$$

Здесь $w = u + iv$ – комплексная горизонтальная скорость; $H$ – глубина бассейна; $k_b$ – коэффициент придонного трения; $\rho_0$ – средняя плотность; $\tau^w = \tau_x^w + i\tau_y^w$, $\tau_x^w$, $\tau_y^w$ – напряжения ветра вдоль осей $Ox, Oy$ соответственно; $\alpha = \sqrt{\dfrac{il}{K_z}}$; $l$ – параметр Кориолиса.

В случае, когда коэффициент вертикального турбулентного обмена определяется формулой $K_z = \delta e^{\lambda z}$, решение выписывается через модифицированные функции Бесселя $I_1(\xi)$, $K_1(\xi)$

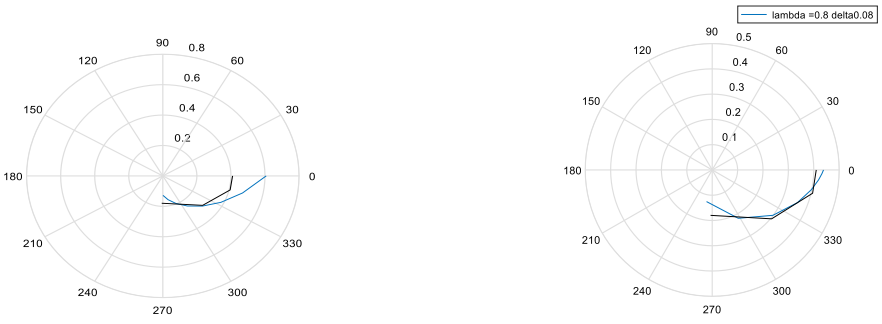
$$w = C_1\xi I_1\left(\frac{2}{\lambda}\sqrt{\frac{\alpha}{\delta}}\xi\right) + C_2\xi K_1\left(\frac{2}{\lambda}\sqrt{\frac{\alpha}{\delta}}\xi\right).$$

Произвольные константы $C_1$, $C_2$ находятся из граничных условий на дне и на поверхности.

Для сравнения выделим в аналитическом решении восточную и северную компоненты скорости и сравним их с соответствующими скоростями в первой моде для первых девяти точек измерения, считая от поверхности до глубины 9.58 м.

С этой целью определялось среднеквадратичное отклонение значений для первой моды от аналитического решения. В результате получено, что в приповерхностном слое наилучшее значение достигается для экспоненциально убывающего по глубине коэффициента вертикального турбулентного обмена.

Наилучшее приближение в смысле среднеквадратичного отклонения получено для постоянного $K_z = 0,0011$ $м^2/с$ при среднеквадратичном отклонении $M = 0,174$ (рис. 5,*а*), для экспоненциального $K_z = \lambda e^{-\delta z}$ при $\lambda = 0.8,\ \delta = 0.08$ и среднеквадратичном отклонении $M = 0,07$ (рис. 5,*б*).




*а* *б*

Рис. 5. Годограф, полученный с помощью метода главных компонент (черная линия), и годограф скорости для модели Экмана (синяя линия) с постоянным $K_z$ = 0,0011 м$^2$/с (*а*) и экспоненциальным $K_z = \delta e^{\lambda z}$; $\delta$ = 0,08; $\lambda$ = 0,8 (*б*).

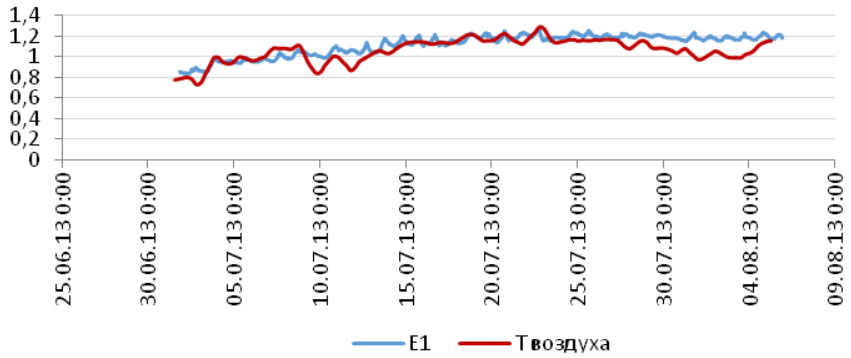



Рис. 6. 2013. $r$ = 0,74, $r$ = 0,744 (запаздывание на 1,5 часа), $r$ = 0,746 (запаздывание на 3 часа).

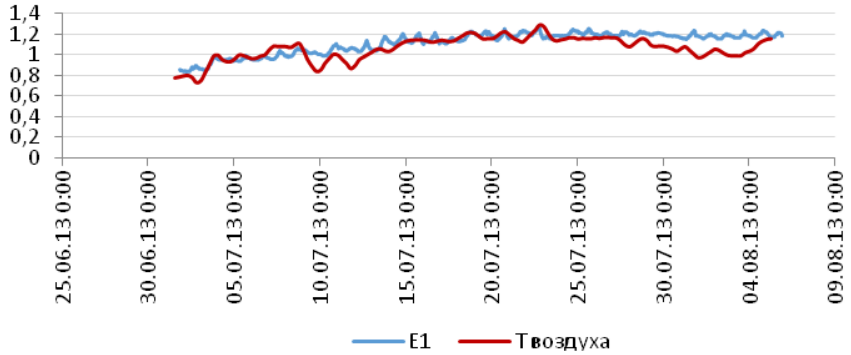



Рис. 7. 2015, $r$ = 0,7. $r$ = 0,73 (запаздывание на 3 часа).

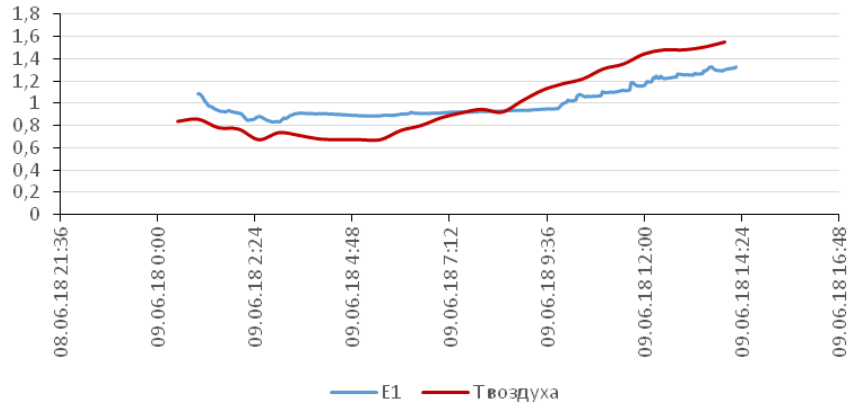



Рис. 8. 2018, $r$ = 0,91, $r$ = 0,96 (с запаздыванием на 1,5 часа), $r$ = 0,92 (с запаздыванием на 3 часа).

**Длительные измерения значений температуры.** В 2013, 2015, 2018 гг. были проведены измерения температуры по глубине и анализ поверхностной температуры методом главных компонент. В результате получены следующие оценки корреляции с температурой воздуха (рис. 6-8).

**Заключение.** Применение метода главных компонент при анализе длительных измерений скорости течения и температуры в озере Шира позволило выявить особенности течения в озере в летний период.

Так корреляция между первым модальным коэффициентом и поверхностной температурой воздуха максимальна, если ее рассчитывать со сдвигом на 1.5-2 часа.

Анализ первой моды для скорости и ее сравнение с аналитическим решением для стационарного течения однородной жидкости показывает преимущество применения экспоненциально убывающего коэффициента вертикального турбулентного обмена по сравнению с постоянным при моделировании течения в верхнем перемешанном слое.

Анализ измерения течения с использованием дрифтера показал в целом согласованность направления циркуляции ветра в районе озера Шира и поверхностного течения.

ЛИТЕРАТУРА


[1] *Davis R.E.* Drifter observations of coastal surface currents during CODE: The statistical and dynamical views // Journal of Geophysical Research: Oceans. 1985. V. 90, № C3. P. 4756-4772.

[2] *Niiler P.P., Paduan, J.D.* Wind-driven motions in the northeast Pacific as measured by Lagrangian drifters // Journal of Physical Oceanography. 1995. V. 25, № 11. P. 2819-2830.

[3] *Poulain P.M., Warn-Varnas A., Niiler P.P.* Near-surface circulation of the Nordic seas as measured by Lagrangian drifters // Journal of Geophysical Research: Oceans. 1996. V. 101, № C8. P.18237-18258.

[4] *Fratantoni D.M.* North Atlantic surface circulation during the 1990's observed with satellite-tracked drifters // Journal of Geophysical Research: Oceans. 2001. V. 106, № C10. P. 22067-22093.

[5] *Porter M., Inall M.E., Green J.A.M., Simpson J.H., Dale A.C., Miller P.I.* Drifter observations in the summer time Bay of Biscay slope current // Journal of Marine Systems. 2016. № 157. P. 65-74.

[6] *Sanderson B.* An analysis of Lagrangian kinematics in Lake Erie // Journal of Great Lakes Research. 1987. V. 13, № 4. P.559-567.

[7] *Pal B.K., Murthy R., Thomson R.E.* Lagrangian measurements in Lake Ontario // Journal of Great Lakes Research, 1998. V. 24, № 3. P.681-697.

[8] *Stocker R., Imberger J.* Horizontal transport and dispersion in the surface layer of a medium-sized lake // Limnology and Oceanography, 2003. V. 48, № 3. P.971-982.

[9] *Kirillin G., Golosov S.A.* Mesoscale vortex in a small stratified lake // Environ Fluid Mech. 2008 № 8. P. 349-366.

[10] *Welander P.* Wind action on a shallow sea: some generalizations of Ekman's theory //Tellus. 1957. V. 9, №. 1. P. 45-52.

[11] *Ekman V.W.* On the influence of the earth's rotation on ocean-currents // Arkiv, Mat. Astr. Fysik. 1905. V. 2, №. 4.

# ИССЛЕДОВАНИЕ ТЕМПЕРАТУРНОГО РЕЖИМА НОВОСИБИРСКОЙ АГЛОМЕРАЦИИ ПО ДАННЫМ СПУТНИКОВОГО ЗОНДИРОВАНИЯ

*Воронина П.В.[(1)(2)], Чубаров Д.Л.[(1)], Добрецов Н.Н.[(1)(3)]*
[(1)] Институт вычислительных технологий СО РАН, Новосибирск
[(2)] Новосибирский государственный университет, Новосибирск
[(3)] Институт геологии и минералогии им. В.С. Соболева, Новосибирск

В работе представлены результаты предварительного исследования температурного режима Новосибирской агломерации по данным спутникового зондирования.
Ключевые слова: MODIS, обработка спутниковых данных, температура поверхности, температурные аномалии, статистические выбросы, метод RST (Robust Satellite Technique).

В настоящее время в городах сосредоточено более половины населения планеты, которое с каждым годом растет. Увеличивается строительство в городах, в силу экономических причин усиливается связь между крупными населенными пунктами и окружающими их небольшими поселениями, формируя тем самым городскую агломерацию. Под городской агломерацией будем понимать компактное скопление населённых пунктов, главным образом городских, местами срастающихся, объединённых в сложную многокомпонентную динамическую систему с интенсивными производственными, транспортными и культурными связями. Развивающаяся городская агломерация оказывает мощное воздействие на природную среду: происходит загрязнение вод, воздуха, почв; меняются атмосферные процессы. Образующиеся на территории городов устойчивые положительные аномалии температуры воздуха получили название «островов тепла». Это явление впервые было описано еще в XIX веке, возникает оно вследствие поступления в атмосферный воздух различных примесей от производственных объектов, транспорта и других источников загрязнения атмосферы и снижения ее прозрачности. В результате застройки увеличивается доля поглощенной солнечной радиации по сравнению с естественными ландшафтами. За счет сокращения площадей с открытым почвенным покровом и зелеными насаждениями снижается расход тепла на испарение, что приводит к росту теплового баланса.

Помимо формирования городского «острова тепла» происходит и повышение температуры поверхности. Температурный режим этого сложного природно-антропогенного комплекса является одним из инструментов в исследовании экологического состояния территории [1-4]. Сопоставление температур поверхностей в городах с температурами пригородной местности, эколого-географическое картографирование урбанизированных территорий предложено в [5, 6].

В процессе изучения температуры поверхности городской агломерации и прилегающих территорий одним из источников информации могут выступать данные космической съёмки, проведенные в тепловом инфракрасном диапазоне [7-9]. В качестве таких данных разумно использовать данные среднего пространственного разрешения (около 1 км) (например, данные, полученные со спектрорадиометра MODIS, установленного на спутнике Terra) [4, 5, 10, 11]. Преимущество применения таких материалов заключается в высокой повторяемости съёмки. И, несмотря на невысокое пространственное разрешение, такие данные дают возможность проводить исследования не только самого городского «острова тепла», но и его влияния на окрестности, позволяют оценить общую мощность и протяжённость такого «теплового острова» [4, 10].

Настоящая работа посвящена исследованию температурного режима Новосибирской агломерации и окрестностей по данным дистанционного зондирования среднего пространственного разрешения (MODIS/Terra). Для этого предполагается проанализировать средние в пикселе температуры поверхности Земли, полученные по данным дистанционного зондирования, для различных временных интервалов. Для получения более информативной картины распре-

деления температуры поверхности внутри агломерации будут изучены пространственно-временные ряды данных о температуре с целью выявления выбросов в рядах по алгоритму Robust Satellite Technique (RST) [12, 13].

Новосибирская городская агломерация является лидером по развитию крупных городских территорий в азиатской части России, сформировалась вокруг центра Новосибирской области г. Новосибирска. Ядро агломерации образуют города Новосибирск, Бердск, Искитим, Обь, рабочий посёлок Краснообск и наукоград Кольцово. Существуют еще небольшие поселки, окруженные городскими районами, но не являющиеся административными городскими единицами (рис. 1). Новосибирская агломерация занимает территорию более 36 тыс. кв. км, на которой проживают более 2 млн. человек. Характерной особенностью Новосибирской городской агломерации является ее выраженная моноцентричность: г. Новосибирск намного больше всех остальных населенных пунктов, входящих в ее состав, вместе взятых.

В ИВТ СО РАН развивается информационная инфраструктура, обеспечивающая хранение, архивацию и пользовательский доступ к данным дистанционного зондирования Земли [14]. В отличии от традиционной выгрузки отдельных сцен из архивов и подключение к геоинформационной системе, что позволяет визуализировать данные в виде наборов тематических слоев, но затрудняет использование большого числа спутниковых снимков с применением сложных функций их обработки, предложенная в [14] технология hVault обеспечивает виртуальную интеграцию представленных в архиве данных в реляционной СУБД. В основе этой технологии лежит принцип представления данных в виде набора таблиц, содержащих данные спутниковых изображений или построенных на их основе информационных продуктов, с последующей реализацией алгоритмов анализа пространственно-временных рядов средствами СУБД.

Для решения поставленной задачи использовались продукты, получаемые в результате обработки данных приборов Terra/MODIS. Восстановленные по измерениям интенсивности инфракрасного излучения, регистрируемого в каналах 31 (длина волны – 11 мкм) и 32 (длина волны – 12 мкм), значения температуры на основании алгоритма [15] составляют информационные продукты MOD11A1. Содержащиеся в MOD11A1 данные представлены на регулярной сетке в синусоидальной проекции с размером ячейки около 1 км. При хороших атмосферных условиях наблюдения алгоритм [15] обеспечивает точность восстановления температуры в пределах 1 K.



Рис. 1. Новосибирская агломерация с пригородами.

Анализ пространственно-временных рядов температуры поверхности территории Новосибирской агломерации проведен по рассчитанным средним значениям температуры в пикселе для каждого календарного года с 2001 по 2014 (данные взяты из архива ИВТ СО РАН) для всего года; для бесснежного периода с 1 апреля по 31 октября; для зимнего периода с 1 ноября по 31 марта. В качестве примера результатов проведенного исследования рассмотрим период с 1 апреля 2006 г. по 31 марта 2007 г., который разобьем на два: «условно» бесснежный с 1 апреля по 31 октября 2006 г. и «условно» снежный с 1 ноября 2006 г. по 31 марта 2007 г. (термин «условно» появился в связи с тем, что в разные годы установление снежного покрова и его сход даже в пределах города происходят в разные дни).

Процесс изучения температурного режима Новосибирской агломерации последовательно состоял из решения следующих задач:

1. оценка пространственной структуры температуры поверхности территории с позиции ее неоднородности по средней температуре в пикселе за период;
2. получение сравнительных оценок температуры поверхности в разные сезоны по средней температуре в пикселе за период;
3. выявление структуры теплового поля в черте г. Новосибирска по средней температуре в пикселе;
4. выявление тепловых аномалий, являющихся статистическими выбросами в пространственно-временных рядах данных температуры поверхности, по методу Robust Satellite Techniques (RST);
5. проверка гипотезы наличия «острова тепла» в г. Новосибирске и поселениях, входящих в Новосибирскую городскую агломерацию.

Рассмотрим результаты по каждой поставленной задаче.

1. На рис. 2 представлено распределение средней температуры в пикселе за календарный год и с 2001 по 2014 гг. Структура теплового поля, как и ожидалось, неоднородна. Первое, что обращает на себя внимание – это «теплые» водные поверхности (Обское водохранилище, Бердский залив и р. Обь). Минимальная средняя температура в области исследования составила –6,3°С, максимальная: +1,7°С. Далее отметим более высокую по сравнению с окрестностями среднюю температуру поверхности внутри г. Новосибирска (более детально о температурном распределении внутри города поговорим позже). Средняя температура внутри городской черты превышает среднюю температуру удаленных окрестностей на 4-6 градусов. Также видно, что средняя температура поверхности внутри небольших городов и поселений, входящих в Новосибирскую агломерацию (Бердск, Горный, Колывань, Коченево, Чик) выше на 2-4 градуса по сравнению со средней температурой близлежащих территорий, но ниже, чем в г. Новосибирске (примерно такая же, как на окраинах г. Новосибирска). А вот в г. Искитим средняя температура близка к средней температуре г. Новосибирска, отличие составляет 1-2 градуса.

2. С 1 апреля по 31 октября 2006 г. в пределах анализируемого полигона, охватывающего г. Новосибирск и его окрестности (см. рис. 1), средняя температура поверхности изменялась от +0,5°С до +12°С. На большей части территории средняя температура в пикселе оказывается ниже +6,3°С. На территории г. Новосибирска средняя температура распределена неравномерно. Большая часть территории города имеет температуру поверхности от +8°С до +11°С (рис. 3,*а*). В городах Бердск и Искитим средняя температура поверхности составляет +8-9°С, что сравнимо со средней температурой г. Новосибирска. Тогда как средняя температура в пикселе за указанный бесснежный период в г. Обь и небольших поселках (Горный, Колывань, Кольцово, Коченево, Краснообск, Мошково, Чик) не отличается от температуры окрестностей с отсутствующей городской застройкой.

Рассмотрим период с 1 ноября 2006 г. по 31 марта 2007 г., когда в исследуемой территории практически ежегодно устанавливается снежный покров и начинается зима. Средняя температура поверхности в 2006-2007 гг. изменялась от –21,5°С до –10,7°С. На большей части территории средняя температура в пикселе оказывается ниже –18°С. В городах областного

подчинения (Бердск, Искитим, Обь) и небольших поселках (Горный, Колывань, Кольцово, Коченево, Краснообск, Мошково, Чик) средняя температура в пикселе выше: около –16°С. На территории г. Новосибирска средняя температура распределена неравномерно. Большая часть территории города имеет температуру поверхности от –12,5 до –13,5°С (рис. 3,*б*).

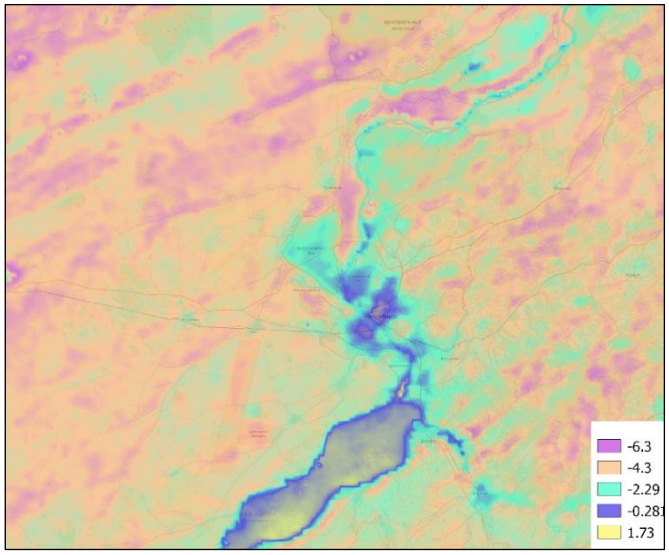
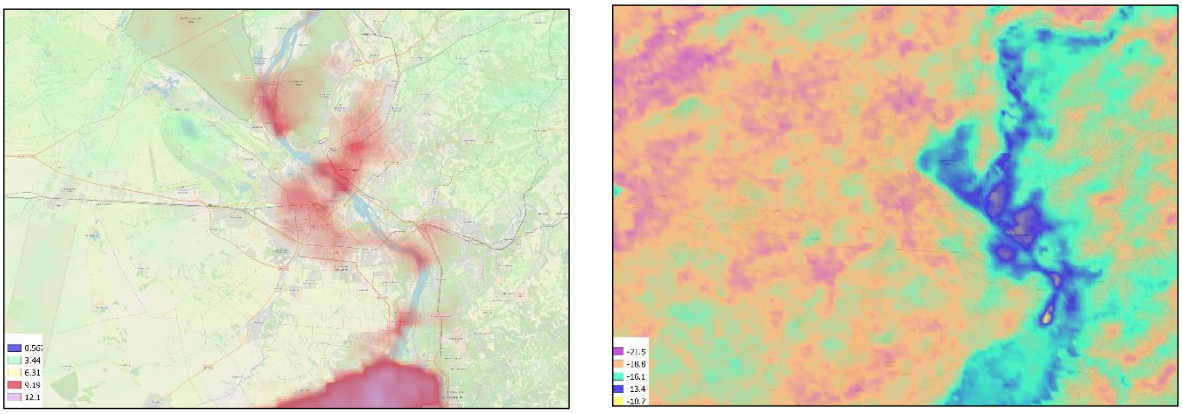



Рис. 2. Распределение средней температуры в пикселе
с 1 января 2001 г. по 31 декабря 2014 г. (температура указана в градусах Цельсия).



*а* *б*

Рис. 3. Распределение средней температуры в пикселе: *а* – с 1 апреля по 31 октября 2006 г.;
*б* – с 1 ноября 2006 г. по 31 марта 2007 г. (температура указана в градусах Цельсия).




Рис. 4. Распределение средней температуры в пикселе с 1 ноября 2006 г. по 31 марта 2007 г.
в г. Новосибирске (температура указана в градусах Цельсия).

3. На рис. 4 представлен фрагмент карты г. Новосибирска с представлением только двух интервалов с самой высокой температурой для зимнего периода (выше –13°С), что позволяет выявить структуру теплового поля в городе. На территории основного субъекта агломерации – г. Новосибирска – выделяются несколько районов с более высокой средней температурой по отношению к остальной территории: это Заельцовский парк, часть города, ограниченная улицами Дуси Ковальчук, Владимировская, Фабричная, Никитина, Волочаевская, Учительская, Георгия Колонды. В левобережной части города выделяется район, ограниченный улицами Немировича-Данченко, Станционной, с более высокой по отношению ко всей части города средней температурой.

4. Для полноты изучения распределения теплового поля городской агломерации был выполнен анализ пространственно-временных рядов температуры поверхности территории Новосибирской агломерации по методу RST (Robust Satellite Techniques) [12] с целью выявления статистических выбросов. В основе метода лежит статистический анализ наборов спутниковых данных температуры земной поверхности для выделенной территории. Для исключения влияния сезонного хода температур и неоднородности рельефа происходит преобразование данных. Сначала фиксируется интересующий исследователя временной интервал, извлекаются температуры в эти дни года за несколько лет. Далее по выбранной области вычисляется индекс RST. Основным достоинством индекса является то, что при выборе области и временного интервала подходящих размеров он позволяет исключить влияние вариаций температуры, вызванных климатическими процессами, неоднородностью рельефа и погодными условиями [13, 16]. В работе [16] авторами предложена модификация метода RST путем вычисления накопительной суммы значений индекса, признанных аномальными, т.е. суммируются значения индекса, которые превышают так называемое пороговое значение, в нашем случае оно равно 2. Одновременно считаются и количество дней с аномальным значением индекса. Модифицированная методика позволяет выделять аномальные проявления, случающиеся несколько раз в течение исследуемого временного интервала, но проявляющиеся в разные дни.

Метод RST был применен для каждого года, начиная с 2001 по 2014, и для каждого периода: с 1 апреля по 31 октября и с 1 ноября по 31 марта.

Интересен результат 2006 г. с 1 апреля по 31 октября. В этом году в рассматриваемый временной интервал выявлены тепловые аномалии, проявившиеся в разных районах г. Новосибирска: на территории Кудряшовского бора, Криводановке, с различной степенью интенсивности аномальной оказалась практически вся территория города. Аномальными оказались и г. Обь, р.п. Чик, Коченево, Колывань, Горный, г. Искитим, чуть-чуть г. Бердск.

5. По результатам проведенных исследований подтверждена гипотеза о существовании городского «острова тепла» для г. Новосибирска и для г. Искитима. Однако для более мелких городских образований «острова тепла» не обнаружено.

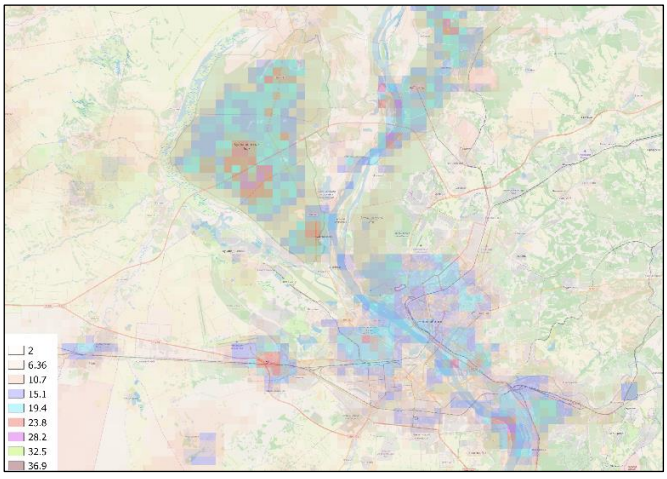



Рис. 5. Фрагмент карты г. Новосибирска с накопительной суммой аномальных значений индекса для периода с 1 апреля по 31 октября 2006 г.

**Заключение.** Проблема «тепловых островов» городов становится на современном этапе развития науки особенно актуальной в связи с вкладом городов в процессы, влияющие на экологию территории. Тепловые изображения позволяют выявить пространственную структуру «тепловых островов» городов. Такими данными могут быть снимки, полученные со спектрорадиометра MODIS, установленного на спутнике Terra. Несмотря на невысокое пространственное разрешение таких снимков, есть преимущество применимости этих материалов в исследовании температурного режима городской агломерации. Это высокая повторяемость и два съемочных канала тепловой инфракрасной съемки. А широкий охват и низкое пространственное разрешение дают возможность проводить исследования не только самого «теплового острова» города, но оценивать и его влияние на окрестности, общую мощность и протяжённость такого «теплового острова».

Впервые по данным космического мониторинга выполнен анализ температуры поверхности для Новосибирской городской агломерации. Установлены структурные особенности теплового поля в летнее и зимнее время, а также пределы территориальной изменчивости температуры городской поверхности. Тепловое поле в пределах города весьма неоднородно.

Определение структуры теплового поля городской агломерации позволит, по мнению авторов, обоснованному планированию развития территории с учетов требований создания комфортной среды проживания.

## ЛИТЕРАТУРА


[1] *Матузко А.К., Якубайлик О.Э.* Мониторинг температуры земной поверхности территории Красноярска и окрестностей на основе спутниковых данных Landsat-8 // Успехи современного естествознания. 2018. № 7 С. 177-182.

[2] *Кораблёва Е.Г., Ленская О.Ю.* Исследования острова тепла города Челябинска в зимний период // Вестник Челябинского государственного университета. Экология. Природопользование. 2010. № 8 (189). Вып. 4. С. 15-23.

[3] *Погорелов А.В., Липилин Д.А.* Тепловой «портрет» города Краснодара по данным спутниковых снимков // Вестник ПНИПУ. Прикладная экология. Урбанистика. 2016. № 4. С. 32-45.

[4] *Cheval S., Dumitrescu A.* The July urban heat island of Bucharest as derived from MODIS images // Theoretical and Applied Climatology. 2009. № 96. P. 145-153.

[5] *Pu R., Gong P., Michishita R., Sasagawa T.* Assessment of multi-resolution and multi-sensor data for urban surface temperature retrieval // Remote Sensing of Environment. 2006. № 104. P. 211-225.

[6] *Коновалова Т.И., Трофимова И.Е.* Картографирование экологического состояния урбанизированных территорий на основе материалов дистанционных исследований Земли // Исследования Земли из космоса. 2008. № 4. С. 36-44.

[7] *Voogt J.A., Oke T.R.* Thermal remote sensing of urban climates // Remote Sensing of Environment. 2003. N 86. P. 370-384.

[8] *Weng Q.* Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends // ISPRS Journal of Photogrammetry and Remote Sensing. 2009. N 64. P. 335-344.

[9] *Weng Q., Quattrochi D.A.* Thermal remote sensing of urban areas: An introduction to the special issue // Remote Sensing of Environment. 2006. N 104. P. 119-122.

[10] *Hung T., Uchihama D., Ochi S., Yasuoka Y.* Assessment with satellite data of the urban heat island effects in Asian mega cities // International Journal of Applied Earth Observation and Geoinformation. 2006. N 8. P. 34-48.

[11] *Stathopoulou M., Cartalis C.* Downscaling AVHRR land surface temperatures for improved surface urban heat island intensity estimation // Remote Sensing of Environment. 2009. N 113. P. 2592-2605.

[12] *Tramutoli V., Di Bello G., Pergola N. and Piscitelli S.* Robust satellite techniques for remote sensing of seismically active areas / // Annals of Geophysics. 2001. Vol. 44 (2). P. 295-312.

[13] *Corrado R., Caputo R., Filizzola C., Pergola N., Pietrapertosa C., Tramutoli V.* Seismically active area monitoring by robust TIR satellite techniques: a sensitivity analysis on low magnitude Earthquakes in Greece and Turkey // Natural Hazards and Earth System Sciences. 2005. N 5. P. 101-108.

[14] *Шокин Ю.И., Добрецов Н.Н., Мамаш Е.А., Кихтенко В.А., Воронина П.В., Смирнов В.В., Чубаров Д.Л.* Информационная система приема, обработки и доступа к спутниковым данным и ее

применение для решения задач мониторинга окружающей среды // Вычислительные технологии. 2015. Т. 20, № 5. С. 157-174.
[15] *Wan Z., Dozier J.* A generalized split-window algorithm for retrieving land-surface temperature from space // IEEE Trans. Geoscience and Remote Sensing. 1996. Vol. 34, N 4. P. 892-905.
[16] *Воронина П.В., Чубаров Д.Л., Добрецов Н.Н.* Развитие метода RST (Robust Satellite Technique) для выявления аномалий температуры поверхности на территории Горного Алтая // Интерэкспо Гео-Сибирь. 2018. Т. 4, № 1. С. 64-73.

# БИОГЕОХИМИЧЕСКИЕ ОСОБЕННОСТИ НАКОПЛЕНИЯ ZN, CU, PB И CD РАСТЕНИЯМИ ПОСТПИРОГЕННЫХ БОЛОТ

*Гашкова Л.П.*
Сибирский институт сельского хозяйства и торфа – филиал Сибирского федерального научного центра агробиотехнологий РАН, Томск;
Национальный исследовательский Томский государственный университет, Томск

В статье рассматривается изменение содержания в растениях Zn, Cd, Pb и Cu и их биогеохимической активности на постпирогенных участках болот. Отмечено, что на участках, пострадавших от пожара, происходит увеличение биогеохимической активности примерно в три раза. Также происходит увеличение концентрации Cu во всех видах растений и Pb в сфагновых мхах. Относительная биогеохимическая активность максимальна у багульника и осоки.
*Ключевые слова: цинк, кадмий, свинец, медь, относительная биогеохимическая активность видов*

Пожары на болотах приводят к выгоранию растительности и торфа, в результате чего в воздух высвобождается множество загрязняющих веществ, которые в течение многих лет связываются растениями и откладываются в торфяной залежи. Данные процессы приводят к изменению геохимической структуры болота. В частности, пожары ускоряют биогеохимический цикл микроэлементов на болотах, переводя их из органической в неорганическую, доступную для растений форму, благодаря чему улучшается обеспечение болотных растений микроэлементами [1]. Концентрация тяжёлых металлов в биодоступной форме в смоге от торфяных пожаров может превышать предельно допустимые нормы ВОЗ и представлять серьёзную угрозу для жизни человека [2].

Целью работы было оценить степень изменения состава растений постпирогенных болот и их биогеохимической активности.

Исследования включали в себя геоботаническое описание участков, измерение уровня болотных вод от средней поверхности болота [3], определение мощности торфяной залежи, отбор проб растений (Берёза - *Betula pubescens* Ehrh., Сосна – *Pinus sylvestris* L., Хамедафне – *Chamaedaphne calyculata* (L.) Moench, Багульник – *Ledum palustre* L., Пушица – *Eriophorum vaginatum* L., Осока – *Carex rostrata* Stokes, Морошка – *Rubus chamaemorus* L., Сфагнум – *Sphagnum fuscum* (Schimp.) Klinggr. и проб торфа для определения содержания тяжёлых металлов. Отбор проб торфа производился нами при помощи пробоотборочного бура ТБГ-1 с глубины 0-25 см.

Определение массовых концентраций Zn, Cd, Pb и Cu в растениях, торфе и воде проводилось методом инверсионной вольтамперометрии после предварительной подготовки проб. Пробоподготовка включает в себя: высушивание до абсолютно сухого веса; измельчение до порошкообразного состояния в лабораторной мельнице; озоление в муфельной печи с предварительным добавлением азотной кислоты и пероксида водорода; растворение полученной золы в концентрированной соляной кислоте. Анализ проводился по методике МУ 31-04/04 (ФР.1.31.2004.00986) в лабораторно-аналитическом центре СибНИИСХиТ (аналитик Кириллова М.Е.).

В качестве основной геохимической меры качества окружающей среды применялось определение массовой доли химического элемента [4].

Для установления связей между основными компонентами геосистемы, геохимических особенностей накопления тяжёлых металлов применялся сравнительный анализ на основе определения биогеохимических коэффициентов. В частности, определялся коэффициент биологического поглощения (КБП), отражающий отношение содержания элемента в золе растения к его содержанию в торфе [4]; биогеохимическая активность видов (БХА), представляю-

щая собой сумму коэффициентов биологического поглощения; а также относительная биогеохимическая активность видов, предложенная автором работы, которая рассчитывается как отношение БХА на нарушенных участках к БХА на естественных участках [5].

Исследования проводились на 5 участках осушенных болот, в различной степени пострадавших от пожара. В качестве фоновых (1) приняты средние значения для тех же видов растений болот, удалённых от всех видов загрязнения. Участок 2 представляет собой сосново-берёзово-вейниково-моховое переходное болото в междуречье рек Оби и Томи. После пожара стволы деревьев обгорели снизу до 1-1,5 м. Мощность торфа на данном участке составляет 4,5 м. Следующие три участка расположены на междуречье рек Икса и Шегарка, к северу от трассы Томск-Бакчар, пострадавшие от пожара в 1998 году. Участок 3 представляет собой сосново-кустарничково-сфагновое болото. После пожара остались погибшие стволы сосен высотой до 10 м. Мощность торфяной залежи составляет 1,4 м. Участок 4 представляет собой березово-сосново-кустарничково-сфагновое верховое болото, микрорельеф которого осложнён упавшими после пожара соснами. Оставшийся после выгорания слой торфяной залежи составляет 0,6 м. Участок 5 находится на окраине осушительной сети и наименее пострадал от пожара. Представляет собой сосново-кустарничково-сфагновое болото. Мощность торфяной залежи составляет 1,5 м. Участок 6 представляет собой сосново-берёзово-пушицево-моховое переходное болото, расположенное на террасе р. Обь. Мощность торфяной залежи составляет 2,1 м.

Сравнение биогеохимической активности видов на фоновых и постпирогенных болотах показало, что БХА увеличивается на постпирогенных участках почти в 3 раза (рис. 1).

Увеличение данного коэффициента происходит благодаря переходу элементов в биодоступную форму и накоплению их в торфе, благодаря чему растения лучше усваивают микроэлементы, чем на ненарушенных болотах, на которых содержание доступных форм гораздо ниже.

Сравнение рассчитанной нами биогеохимической активности видов на различных болотах показало, что данный показатель увеличивается на всех рассмотренных участках, но на участке 4 (березово-сосново-кустарничково-сфагновое верховое болото на междуречье рек Икса и Шегарка) увеличение статистически недостоверно, хотя относительная БХА на данном участке больше 1. Сравнение относительной БХА показало, что данный коэффициент наиболее высок на 3 и 6 участках (сосново-кустарничково-сфагновое верховое болото на междуречье рек Икса и Шегарка и сосново-берёзово-пушицево-моховое переходное болото террасе р. Обь) (рис. 2). Низкое значение относительной биогеохимической активности видов на участке 4 объясняется малой мощностью торфяной залежи на данном участке, оставшейся после пожара. Содержание элементов в оставшихся после пожара слоях низинного и переходного торфа выше, чем в верховом торфе с других участков, что является причиной низкой биогеохимической активности видов на участке 4.

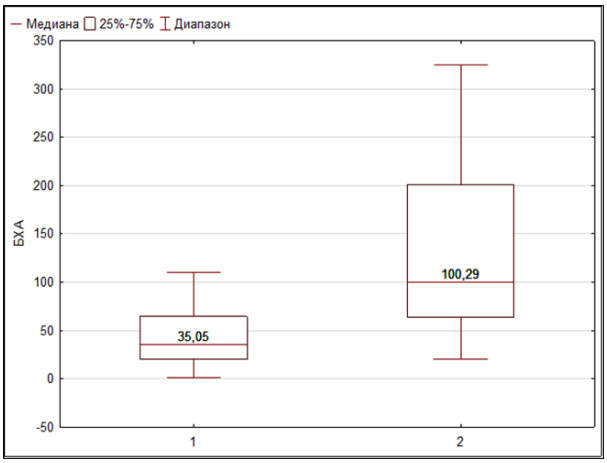



Рис. 1. Сравнение биогеохимической активности видов на фоновых (1) и постпирогенных (2) участках болот.

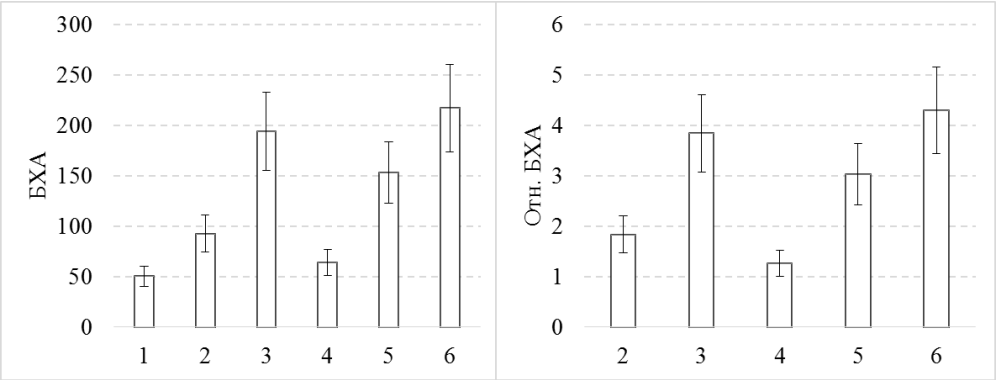



Рис. 2. Изменение биогеохимической активности видов (БХА) и относительной БХА на участках постпирогенных болот.

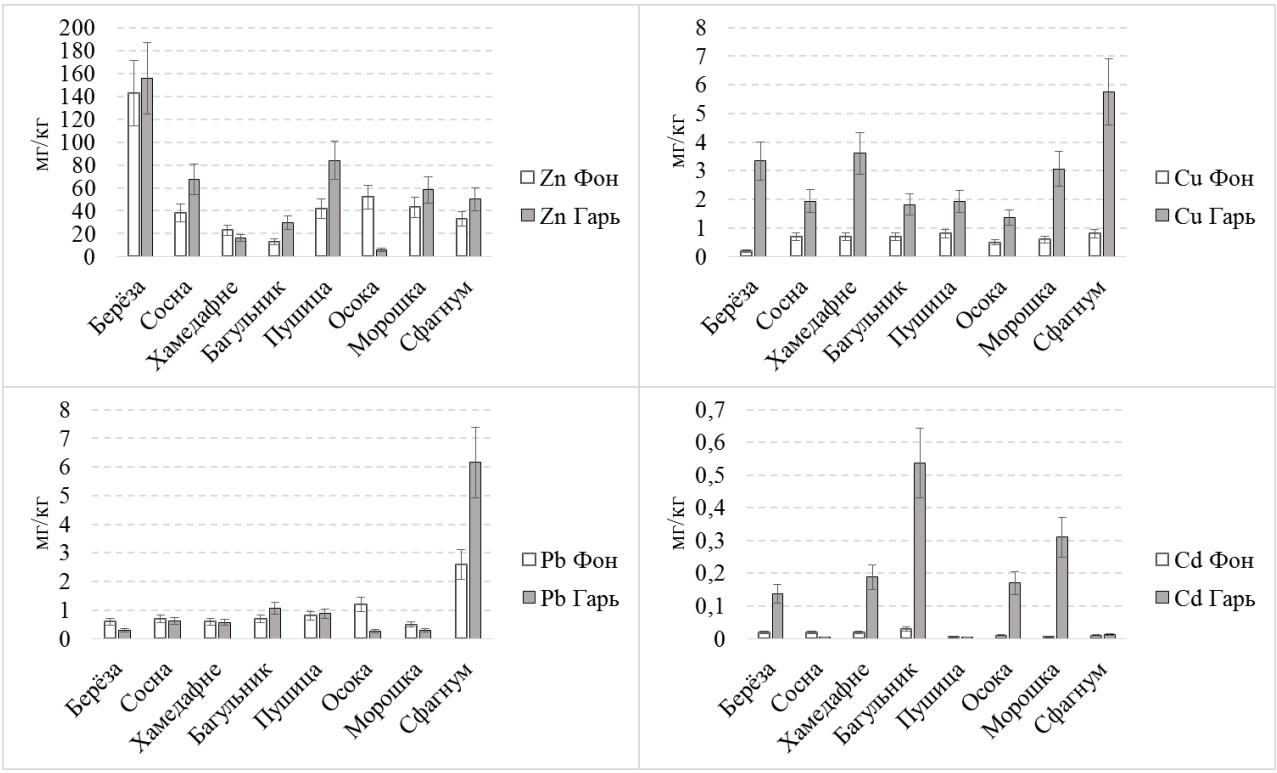



Рис. 3. Сравнение содержания элементов в растениях фоновых и постпирогенных участков.

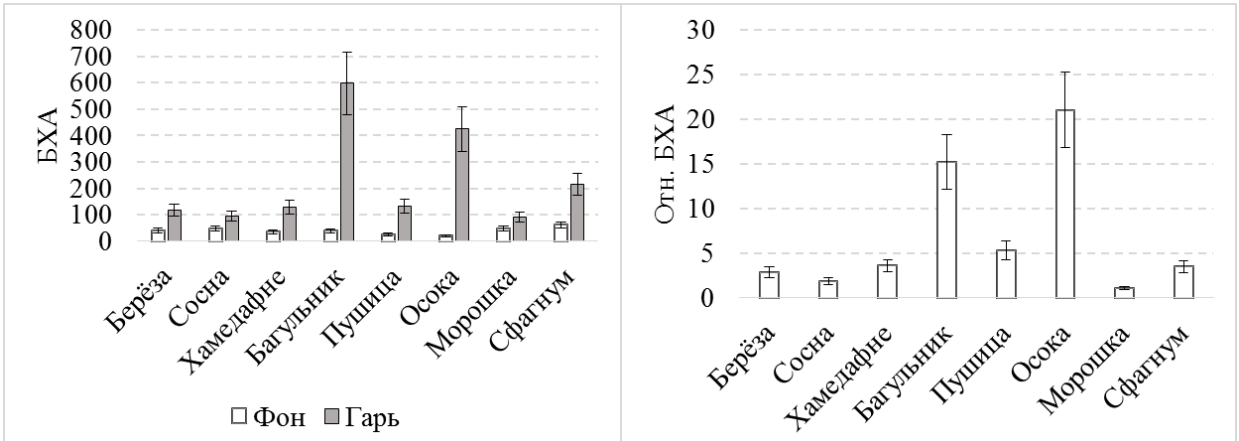



Рис. 4. Сравнение биогеохимической активности видов (БХА) и относительной БХА отдельных видов растений.

В результате сравнения полученных данных по содержанию Zn, Cu, Pb и Cd в растениях разных видов постпирогенных участков болот с растениями фоновых болот выяснилось, что содержание данных элементов не всегда увеличивается на постпирогенных участках, по сравнению с фоном. Содержание Zn увеличивается лишь у сосны, багульника, пушицы и сфагнума. Содержание Cu достоверно увеличивается у всех исследованных видов. Содержание Pb статистически значимо увеличивается только у сфагнума. Содержание Cd увеличивается у берёзы, хамедафне, осоки и морошки. Таким образом, у разных видов содержание элементов изменяется неодинаково, и колебания концентрации довольно значительны (рис. 3).

При рассмотрении биогеохимической активности разных видов выяснилось, что БХА увеличивается достоверно у всех рассмотренных видов. Наибольшая относительная БХА отмечена у багульника и осоки (рис. 4).

Таким образом, в результате проведённого исследования можно сделать вывод о том, что биогеохимическая активность видов увеличивается на постпирогенных участках болот почти в три раза. Наименьшая относительная биогеохимическая активность видов наблюдается на том участке, где после пожара осталась минимальная, по сравнению с другими участками, мощность торфяной залежи. Абсолютное содержание элементов в растениях на постпирогенных участках увеличивается неравномерно. Наиболее возрастает концентрация Cu во всех видах растений и Pb в сфагновых мхах. Относительная биогеохимическая активность на постпирогенных болотах максимальна у багульника и осоки.

*Работа выполнена в рамках проведения фундаментальных научных исследований по Государственному заданию № 007-00496-16 ПР.*

## ЛИТЕРАТУРА

[1] *Guoping W., Xiaofei Y., Kunshan B., Wei X., Chuanyu G., Qianxin L., Xianguo L.* Effect of fire on phosphorus forms in Sphagnum moss and peat soils of ombrotrophic bogs // Chemosphere Available. 2015. Vol. 119. P. 1329–1334. DOI: 10.1016/j.chemosphere.2014.01.084.

[2] *Bethaa R., Pradanib M., Lestarib P., Joshic U.M., Reidd J.S., Balasubramaniana R.* Chemical speciation of trace metals emitted from Indonesian peat fires for health risk assessment // Atmospheric Research 2013. Vol. 122, March. P. 571-578. DOI: 10.1016/j.atmosres.2012.05.024.

[3] Наставления гидрометеорологическим станциям и постам. Л.: Гидрометеоиздат, 1990. Вып. 8. Гидрометеорологические наблюдения на болотах. 360 с.

[4] *Перельман А.И. Касимов Н.С.* Геохимия ландшафта. М.: МГУ, 1999. 610 с.

[5] *Гашкова Л.П.* Биогеохимическая активность растений по отношению к тяжёлым металлам в условиях осушенных болот (на примере Бакчарского и Иксинского болотных массивов) // Западно-Сибирские торфяники и цикл углерода: прошлое и настоящее: материалы Пятого международного полевого симпозиума. Томск: Издательский Дом ТГУ. 2017. С. 137-140.

# ИНФОРМАТИВНОСТЬ NDVI В ИЗУЧЕНИИ ПРОСТРАНСТВЕННОЙ ИЗМЕНЧИВОСТИ СОДЕРЖАНИЯ МАГНИЯ В ПОЧВАХ И РАСТЕНИЯХ

*Гопп Н.В.[(1)], Нечаева Т.В.[(1)], Савенков О.А.[(1)], Смирнова Н.В.[(1)], Тархова В.П.[(2)], Смирнов А.В.[(3)]*

[(1)] Институт почвоведения и агрохимии СО РАН, Новосибирск
[(2)] Новосибирский государственный аграрный университет, Новосибирск
[(3)] Алтайский государственный университет, Барнаул

На основе регрессионных моделей и значений NDVI, рассчитанных по снимку Landsat 8 OLI (разрешение 30 м), построены прогнозные карты содержания магния в почвах и растениях. В регрессионных моделях показатель NDVI объяснил 51% вариации содержания обменного магния в почвах, 26% – общего магния в растениях. В среднегумусированных почвах (агрочерноземах, агротемно-серых) по сравнению с сильногумусированными агрочерноземами установлено более низкое содержание обменного магния (в 1.5 раза), при этом изученные почвы не отличались между собой по содержанию общего магния. Не обнаружено корреляции между запасами надземной фитомассы и содержанием магния в почвах и растениях.

*Ключевые слова: обменный магний, гумус, цифровое картографирование, запасы надземной фитомассы, овсяно-гороховая смесь, агрочернозем, агротемно-серая, Chernozems, Phaeozems.*

**Введение.** Магний – незаменимый макроэлемент в минеральном питании растений, так как является составной частью хлорофилла, в котором осуществляются фотохимические реакции. Недостаток магния тормозит синтез хлорофилла, поэтому главный внешний признак данного процесса – пятнистый межжилковый хлороз листьев [1]. Согласно зарубежным исследованиям [2], признаки дефицита магния в растениях обнаруживаются при его содержании в зрелых листьях менее 0.20% на сухую массу, достаточным считается содержание магния в диапазоне от 0.25 до 1.0%, чрезмерным или токсичным – более 1.5%.

Общее содержание магния в почвах колеблется в пределах от 0.1 до 1% [3]. Общий магний включает в себя четыре формы: водорастворимую, обменную, необменную и органическую. Доля этих форм от общего содержания магния в почве составляет 0.5-1; 5-10; 90-95 и 0.5-2% соответственно [4, 5]. Достаточная обеспеченность почв обменным магнием является залогом хорошего урожая, так как эта форма элемента доступна для поглощения растениями.

Значительное пространственное варьирование содержания магния в почвах и растениях обуславливает необходимость разработки методов картографирования и поиска информативных индикаторов, отображаемых на данных дистанционного зондирования Земли. Одним из самых известных индикаторов является вегетационный индекс (NDVI), использование которого позволяет не только картографировать, но и устанавливать закономерности пространственной изменчивости свойств почв и химического состава растений [6]. Известно, что более высокому значению NDVI посевов соответствует большее содержание хлорофилла в растениях [7, 8]. В исследованиях [8] показано, что NDVI чувствителен к изменению содержания магния в листьях различных растений. Таким образом, можно сказать, что NDVI, хлорофилл и магний являются взаимозависимыми характеристиками растительного покрова.

В литературе содержится довольно много данных о содержании магния в различных почвах, однако недостаточно информации о соответствии значений NDVI уровню содержания магния в почвах и растениях. Кроме этого недостаточно изучена пространственная неоднородность содержания магния в почвах различной гумусированности и растениях, произрастающих на них. Поэтому необходимо продолжить исследования в этом направлении.

**Задачи исследований:** 1) осуществить прогнозное картографирование содержания магния в почвах и растениях с использованием в качестве предиктора показателя NDVI, рассчитанного по снимку Landsat 8 OLI (разрешение 30 м); 2) провести сравнительную оценку

сильно- и среднегумусированных почв по содержанию магния в почвах и произрастающих на них растениях; 3) оценить корреляционную зависимость между содержанием магния в почвах, растениях и значениями NDVI.

**Объекты и методы исследований.** Исследование проводили на территории Предсалаирской дренированной равнины юго-востока Западной Сибири (Новосибирская обл., Тогучинский район, с. Усть-Каменка). Обследованный участок пахотного угодья площадью 100 га и протяженностью 2 км находится в пределах водосборных бассейнов рек Ирба и Хайрузовка. На первой высотной ступени (280-310 м) расположены преимущественно сильногумусированные почвы с содержанием гумуса 5-8%, на второй высотной ступени (190-280 м) – среднегумусированные с содержанием гумуса 3-5% (рис. 1).

Координаты точек опробования определены с помощью системы геопозиционирования (GPS, Garmin eTrex Vista). Отбор индивидуальных почвенных проб (n=35) произведен буром из слоя 0-30 см (пахотный горизонт) по нерегулярной сетке. Почвенные образцы проанализированы на содержание общего магния ($Mg_{общ}$) методом атомно-эмиссионной спектрометрии в сухих порошкообразных образцах, обменного магния ($Mg_{обм}$) – по ГОСТ 26487-85 (экстрагент 1н. KCl) с окончанием на атомно-абсорбционном спектрометре AAnalyst 400 (PerkinElmer Inc., США). Содержание гумуса в почвах определено мокрым озолением по Тюрину, физической глины – по Качинскому, pH водной суспензии ($pH_{вод}$) – потенциометрическим методом [9]. Содержание общего магния в растениях ($Mg_{раст}$) с расчетом на воздушно сухую массу выполнено методом сухого озоления с окончанием на атомно-абсорбционном спектрометре AAnalyst 400 [2]. Запасы надземной фитомассы овсяно-гороховой смеси в фазу кущения определены методом укосов с учетной площади 0.25 м$^2$. Прогнозные карты содержания магния в почвах и растениях составлены с использованием регрессионных моделей. В качестве предиктора использовали NDVI посевов овсяно-гороховой смеси, рассчитанного по многозональному снимку Landsat 8 OLI (разрешение 30 м, снимок от 14.07.2013 года). Оценка значимости различий средних значений параметров почв и растительности проведена с использованием t-критерия Стьюдента и U-критерия Манна-Уитни для нормально и ненормально распределенных данных соответственно.

**Результаты и обсуждение.** Установленные зависимости между NDVI и содержанием магния в почвах и растениях, а также рассчитанные уравнения регрессии (рис. 2) позволили построить карты исследуемых параметров на все обследуемое поле (рис. 3). NDVI объяснил 51% вариации содержания в почвах обменного магния и 26% вариации содержания общего магния в растениях (рис. 2).

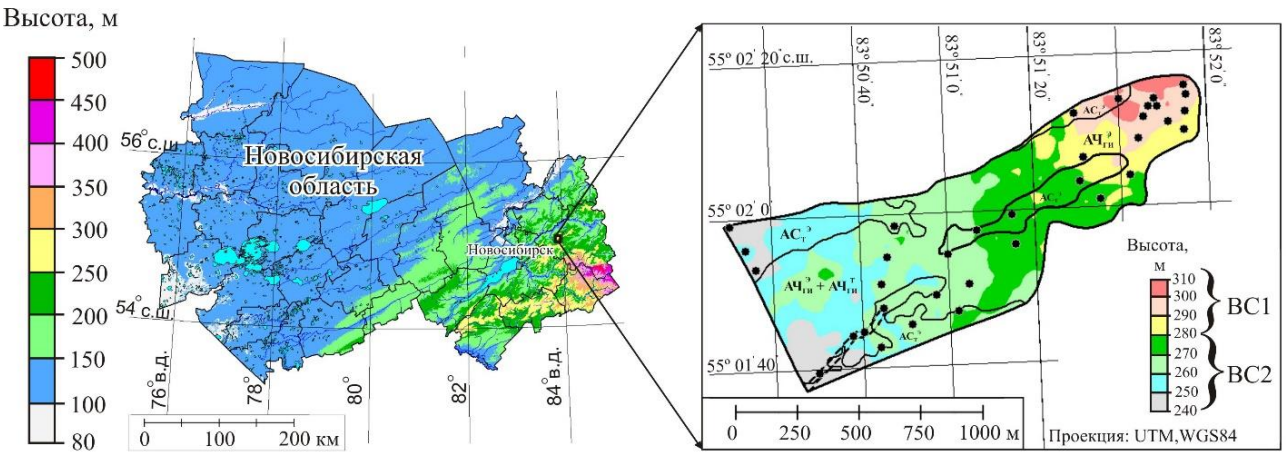



Рис. 1. Карта территории исследования и схема отбора почвенных проб (показана точками).
Условные обозначения: первая (ВС1) и вторая (ВС2) высотные ступени.
Почвы: $АЧ_{ГИ}^{Э}$ – агрочернозем глинисто-иллювиальный элювиированный;
$АЧ_{ГИ}^{Т}$ – агрочернозем глинисто-иллювиальный темноязыковатый; $АС_{Т}$ – агротемно-серая.

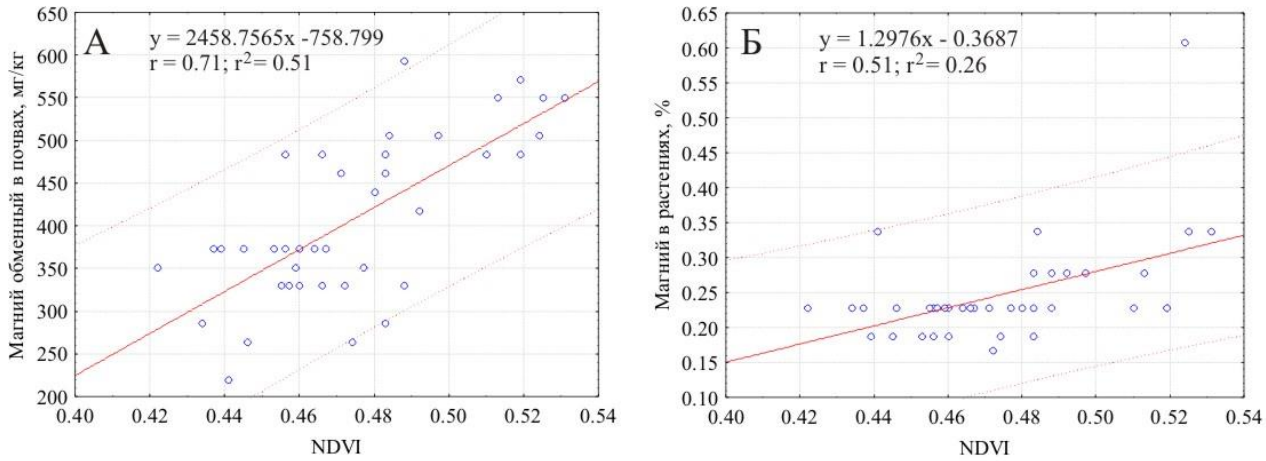



Рис. 2. Регрессионные модели зависимости между NDVI и содержанием магния в почвах (*а*) и растениях (*б*).

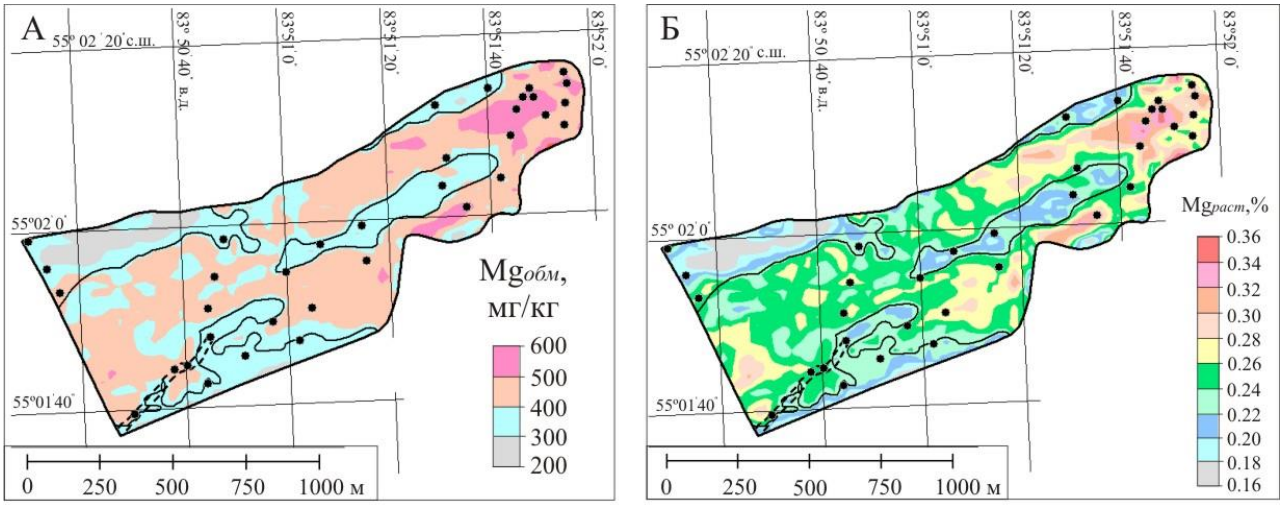



Рис. 3. Прогнозные карты содержания: *а* – обменного магния в почвах; *б* – общего магния в растениях.

Табл. 1. Параметры почв и растительности.

| **Параметры** | **Виды почв по содержанию гумуса** | | |
|---|---|---|---|
| | **сильногумусированные** | **среднегумусированные** | |
| | $АЧ_{ГИ}^{Э}$ (n = 13) | $АЧ_{ГИ}^{Э}+АЧ_{ГИ}^{Т}$ (n = 7) | $АС_{Т}$ (n = 15) |
| **Гумус, %** | 7.8 ± 1.2 | 5.0 ± 1.6* | 4.5 ± 1.2* |
| **$pH_{вод}$** | 5.85 ± 0.15 | 5.82 ± 0.11 | 5.68 ± 0.18 |
| **Ил, %** | 16.6; 15.4; 15.4 | 17.8; 13.5; 12.8 | 16.2 ± 3.3 |
| **Физ. глина, %** | 50.1 ± 3.3 | 45.2 ± 2.7* | 45.7 ± 3.4* |
| **$Mg_{общ}$, мг/кг** | 7935 ± 1555 | 8897 ± 1173 | 8514 ± 1750 |
| **$Mg_{обм.}$, мг/кг** | 504 ± 41 | 354 ± 45* | 326 ± 71* |
| **Доля $Mg_{обм}$ от $Mg_{общ}$, %** | 6.6 ± 1.5 | 4.1 ± 0.9* | 3.9 ± 1.1* |
| **$Mg_{раст}$, %** | 0.29 ± 0.1 | 0.24 ± 0.03 | 0.22 ± 0.04** |
| **ЗНФ, г/м$^2$** | 130.5 ± 27.6 | 155.5 ± 40.3 | 142.8 ± 30.4 |
| **NDVI** | 0.503 ± 0.02 | 0.476 ± 0.01** | 0.456 ± 0.01* |

Для нормально распределенных данных указаны среднее значение и стандартное отклонение (M ± s); для ненормально распределенных данных – среднее значение, медиана и мода (M; Me; Mo). * – показатели, статистически значимо ($p<0.01$) отличающиеся от соответствующих в сильногумусированных почвах, ** – отличия значимы при $p<0.05$. Обозначения почв: см. *рис. 1.*

Табл. 2. Показатели качества регрессионных моделей.

| r, n=35 | $r^2$ | F-критерий, p<0.05 | t-критерий для коэффициента регрессии (b), p<0.05 | t-критерий для коэффициента регрессии (a), p<0.05 | RS | MAPE, % |
|---|---|---|---|---|---|---|
| *Линейная модель «Обменный магний в почвах»*, y=2458.7565x - 758.799 | | | | | | |
| 0.71 | 0.51 | 37.76 | 6.15 | 3.99 | 4.4 | 15 |
| *Линейная модель «Общий магний в растениях»*, y=1.2976x - 0.3687 | | | | | | |
| 0.51 | 0.26 | 12.38 | 3.51 | 1.51 | 4.2 | 13 |

RS-критерий – тест на нормальность распределения остатков, MAPE – средняя абсолютная относительная ошибка, *y* – зависимая переменная (параметры почвы или растений), *x* – независимая переменная (значения NDVI), *n* – число наблюдений, *p* – вероятность ошибки.

Согласно коэффициенту корреляции (r=0.71), положительная связь между NDVI и содержанием обменного магния говорит о его большем содержании в сильногумусированных почвах, на которых произрастает овсяно-гороховая смесь с большими значениями NDVI (рис. 2,*а*, 3,*а*, табл. 1). Такая же зависимость отмечена для содержания общего магния в растениях, где коэффициент корреляции между изученными параметрами (r=0.51) положительный (рис. 2,*б*, 3,*б*, табл. 1).

Для оценки качества моделей рассчитаны различные показатели (табл. 2). Согласно *F*-критерию и его вероятности (*p*), модели и их коэффициенты детерминации статистически значимы ($\alpha$=0.05), так как $F_{факт} > F_{табл}$ (4.08).

Коэффициенты регрессии (*b*) статистически значимы (*t*-критерий больше критического значения 2.009). Свободный член уравнения в модели «Общий магний в растениях» статистически незначим, так как значение *t*-критерия Стьюдента меньше критического значения, поэтому при построении карты его не учитывали. Значение RS-критерия попадает в критический интервал (3.83-5.35), поэтому остатки имеют нормальное распределение. Значение MAPE лежит в интервале от 10 до 20%, что соответствует хорошей точности прогноза. Таким образом, на основании практически всех статистических показателей качества, рассматриваемые модели достаточно хорошо согласуются с фактически измеренными данными. Полученные в исследовании регрессионные модели некорректно использовать для соседних территорий, разных типов растительности и других сроков спутниковой съемки, так как эти модели актуальны только для данного обследованного участка.

Табл. 3. Коэффициенты корреляции Пирсона и Спирмена (выделены серым цветом) между параметрами почв и растительности (*n*=35).

| Параметры | $Mg_{общ}$ | $Mg_{обм}$ | $Mg_{раст}$ | Гу-мус | $pH_{вод}$ | Физ. глина | Ил | NDVI | ЗНФ |
|---|---|---|---|---|---|---|---|---|---|
| **$Mg_{общ}$** | – | | | | | | | | |
| **$Mg_{обм}$** | -0.38 | – | | | | | | | |
| **$Mg_{раст}$** | x | 0.32 | – | | | | | | |
| **Гумус** | x | **0.86** | 0.40 | – | | | | | |
| **$pH_{вод}$** | -0.35 | 0.45 | x | x | – | | | | |
| **Физ. глина** | x | 0.49 | x | x | **0.54** | – | | | |
| **Ил** | x | x | x | -0.38 | 0.43 | **0.78** | – | | |
| **NDVI** | x | **0.71** | **0.51** | **0.72** | x | x | x | – | |
| **ЗНФ** | x | x | x | x | x | x | x | x | – |

ЗНФ – запасы надземной фитомассы. Жирным шрифтом выделены коэффициенты корреляции заметной и высокой силы связи ($p$<0.01); подчеркиванием – умеренной силы связи ($p$<0.01); x – статистически незначимые ($p$>0.05).

На прогнозных картах (рис. 3) показаны закономерности изменения содержания магния в почвах и растениях вниз по склону. В среднегумусированных почвах (агрочерноземах, агротемно-серых) по сравнению с сильногумусированными агрочерноземами установлено более низкое содержание обменного магния (в 1.5 раза), при этом изученные почвы не отличались друг от друга по содержанию общего магния (табл. 1). Большее количество обменного магния в сильногумусированных почвах, по всей видимости, связано с более высоким содержанием физической глины (увеличивает сорбционную емкость) и гумусовых кислот, которые вытесняют катионы магния из почвенно-поглощающего комплекса и усиливают химическое выветривание магнийсодержащих первичных минералов. Этот вывод подтверждается положительными коэффициентами корреляции содержания обменного магния с гумусом и физической глиной высокой и умеренной силы связи соответственно (табл. 3).

Необходимо отметить, что согласно разработанным градациям [10], в сильногумусированных почвах содержание обменного магния в почвах было очень высоким (>485 мг/кг), а в среднегумусированных – повышенным (от 255 до 365 мг/кг). Содержание магния в растениях соответствует оптимальному уровню [2]. Однако содержание общего магния в растениях, произрастающих на сильногумусированных черноземах, в 1.2 раза выше по сравнению со среднегумусированными агротемно-серыми почвами (табл. 2).

Таким образом, использование NDVI в качестве косвенного индикатора содержания магния в почвах и растениях и рассчитанные на его основе цифровые карты позволили выявить пространственную изменчивость изученных параметров почв и растений, сэкономить время, снизить стоимость почвенной съемки и камеральных работ, увеличить информативность карт.

**Выводы.** 1. На основе регрессионных моделей и значений NDVI посевов овсяно-гороховой смеси, рассчитанных по снимку Landsat 8 OLI (разрешение 30 м), построены прогнозные карты содержания магния в почвах и растениях. В регрессионных моделях показатель NDVI объяснил 51% вариации содержания обменного магния в почвах, 26% – общего магния в растениях. Оценка точности прогноза содержания обменного магния в почвах хорошая (ошибка 15%), для общего магния в растениях ошибка прогноза составила 13%.

2. Содержание обменного магния в среднегумусированных почвах (агрочерноземах, агротемно-серых) составило в среднем 340 мг/кг и было в 1.5 раза ниже по сравнению с сильногумусированными агрочерноземами. Содержание обменного магния в сильно- и среднегумусированных почвах высокое и повышенное соответственно. Различия между почвами раной гумусированности по содержанию общего магния не выявлены.

3. Содержание общего магния в надземной массе овсяно-гороховой смеси соответствует оптимальному уровню (0.22-0.29%). Содержание общего магния в растениях, произрастающих на среднегумусированных агротемно-серых почвах, в 1.2 раза ниже, чем на сильногумусированных агрочерноземах.

4. Различия по запасам надземной фитомассы овсяно-гороховой смеси на сильно- и среднегумусированных почвах недостоверны. Корреляционные связи запасов надземной фитомассы с содержанием в почвах обменного магния и значениями NDVI не установлены.

## ЛИТЕРАТУРА


[1] *Минеев В.Г.* Агрохимия. М.: Издательство МГУ, Издательство «Колос», 2004. 720 с.

[2] Handbook of reference methods for plant analysis / Edited by Yash P. Kalra. Boca Raton, Boston, London, New York, Washington: CRC Press, 1998. 287 pp.

[3] *Тихонов В.Н.* Аналитическая химия магния. М.: Наука, 1973. 255 с.

[4] *Магницкий К.П.* Магниевые удобрения. М.: Колос, 1967. 200 с.

[5] *Шеуджен А.Х.* Агробиогеохимия. Краснодар: КубГАУ, 2010. 877 с.

[6] *Гопп Н.В., Нечаева Т.В., Савенков О.А., Смирнова Н.В., Смирнов В.В.* Индикационные возможности NDVI для прогнозного картографирования свойств пахотного горизонта почв склоновых позиций на юге Западной Сибири // Почвоведение, 2017. №11. С. 1377-1389.

[7] *Curran P.J., Windham W.R., Gholz H.L.* Exploring the relationship between reflectance red edge and chlorophyll content in slash pine leaves // Tree Physiology, 1995. V.15. 203-206 pp.

[8] *Im J., Jensen J.R., Coleman M., Nelson E.* Hyperspectral remote sensing analysis of short rotation woody crops grown with controlled nutrient and irrigation treatments // Geocarto International, 2009. Vol. 24. No. 4. 293-312 pp.
[9] Практикум по агрохимии / Под ред. академика РАСХН В.Г. Минеева. М.: Изд-во МГУ, 2001. 689 с.
[10] Методические указания по проведению комплексного мониторинга плодородия почв и земель сельскохозяйственного назначения. М.: Росинформагротех, 2003. 240 с.

# ОПИСАНИЕ УРБАНИЗИРОВАННОЙ ПОВЕРХНОСТИ С ИСПОЛЬЗОВАНИЕМ ГЕОПРОСТРАНСТВЕННЫХ ДАННЫХ ДЛЯ ЗАДАЧ МОДЕЛИРОВАНИЯ И ОЦЕНКИ КАЧЕСТВА ВОЗДУХА

*Гочаков А.В.[(1)], Пененко А.В.[(2)], Колкер А.Б.[(1)(3)], Антохин П.Н.[(4)]*
[(1)] Сибирский региональный научно-исследовательский гидрометеорологический институт, Новосибирск
[(2)] Институт вычислительной математики и математической геофизики СО РАН, Новосибирск
[(3)] Институт оптики атмосферы имени В.Е. Зуева СО РАН, Томск
[(4)] Новосибирский государственный технический университет, Новосибирск

Произведена оценка чувствительности прогноза температуры модели WRF к изменению внешних данных и горизонтального разрешения модельной ячейки. Исследовано влияние уточнения характеристики городской подстилающей поверхности и схемы городской параметризации на прогноз температуры. Показана возможность учета острова тепла при моделировании городского метеорологического режима.
*Ключевые слова: параметризация городской подстилающей поверхности WRF, остров тепла*

**Введение.** Успешность прогноза качества атмосферного воздуха в пространственном масштабе мегаполиса является комплексной задачей и зависит от сети наблюдений, сложности используемой модели численного прогноза погоды и её химико-транспортного блока. Возможности моделей, в том числе, зависят от качества описания задаваемых внешних данных (высота поверхности с учетом естественного рельефа, маска воды, учет категорий землепользования и городской подстилающей поверхности, источников эмиссий и пр.). Источники внешних данных должны иметь пространственное горизонтальное разрешение не ниже заданного для расчетного домена. Для описания подстилающей городской поверхности и источника эмиссий, оправдано разрешение от 1 километра и выше. Хотя современные мезомасштабные модели не способны полностью описать моделируемые процессы в таком масштабе, в настоящее время интенсивно развиваются параметризации, переходные между конвективным и LES масштабами.

Для учета и описания процессов влияния городской застройки и человеческой деятельности на климатический режим города используется класс моделей городской подстилающей поверхности [1]. Для модели WRF [2] это параметризация, требующая задания параметров, связанных с городской застройкой.

Целью работы является изучить чувствительность прогноза температуры при переходе к высокому разрешению горизонтальной сетки (460 метров) с использованием дополнительных внешних модельных данных (категории землепользования, городская застройка). Оценки температуры на уровне 2 метра выбрана как характеристики острова тепла, влияющей на качество прогноза загрязнения атмосферы.

Описание городской подстилающей поверхности в доступных наборах внешних данных для современных мезомасштабных моделей имеют грубое пространственное разрешение и только для общей категории городской застройки. Существует несколько дополнительных источников (NUDAPT, WUDAPT) подробного описания категорий застройки, городских каньонов, а также характеристик материалов, но, преимущественно, для территории Европы и Северной Америки. База данных WUDAPT позволяет получить классификацию локальных климатических зон (LCZ) для любого мегаполиса, основываясь на обработке спутниковых данных. Но первоначально, для создание классификаций, требуется экспертный анализ спутниковых снимков.

Другой источник информации с географической привязкой объектов и участков - проект OSM. В том числе данные проекта включают описание и городской, и естественной среды,

земель сельскохозяйственного назначения и пр. Так как данные общедостпны и классифицированы, они могут быть адаптированы на расчетную сетку модели (см. [3]). Внешние модельные данные, которые могут быть получены из OSM уступают по детализации NUDAPT и WUDAPT, но достаточны для использования в относительно простой и вычислительно эффективной модели городской подстилающей поверхности UCM. Для территории Российской Федерации только Москва имеет необходимые классифицированные данные для получения WUDAPT – категорий без экспертного анализа участков спутниковых снимков. А данные NUDAPT доступны только для территории Северной Америки. Поэтому оправдано использование интерпретации данных OSM в качестве внешних данных WRF для моделирования метеорологического режима города, качества атмосферного воздуха.

**Описание вычислительного эксперимента.** Произведена оценка чувствительности температуры на уровне 2 метра модели WRF к изменению внешних модельных данных - категорий землепользования, включая 3 категории городской застройки, а также использованию городской параметризации UCM для горизонтального разрешения модельной сетки 460 х 460 метров. Произведены сравнения с расчетами для разрешения модельной ячейки 1380 и 6900 метров. При моделировании был использован прямоугольный домен D1 в проекции Ламберта, ограниченный по горизонтали координатами 67,8-70,8° с.ш. и 84,1-92,1° в.д. и по высоте уровнем давления 50 гПа. Число узлов расчётной области составило 100х100х30 по долготе, широте и высоте, соответственно. Шаг сетки по горизонтали составил 6.9 км, по времени – 60 с, по высоте задавался ETA-координатой, учитывающей орографическую поверхность. Домен D1 включал два вложенных домена – D2 с шагом 1380 м и домен D3 совпадающий с территорией агломерации г. Новосибирска с шагом горизонтальной сетки 460 метров. Для D3 использована схема перехода к высокому горизонтальному разрешению, вертикальные уровни оставалось постоянным. Использовалась параметризация пограничного слоя YSU PBL [4]

Начальные и граничные условия для полей метеорологических величин задавались с использованием данных модели FNL (NCEP) с 6-часовым разрешением по времени. Поля метеорологических данных FNL получены на основе данных GFS. Вычислительный эксперимент был выполнен для зимнего периода с 11.12.2018 по 28.01.2019 г., для летнего периода с 10.06.2018 по 29.07.2018 г.

Параметры, чувствительность к которым изучалась:

- Разрешение модельной ячейки
- Категории землепользования, с использованием стандартного набора данных - по результатам снимков со спутника Modis, 21 категория.
- Категории землепользования, полученные в результате обработки данных проекта OSM, 33 категории, включая 3 категории городской застройки.
- Использование параметризации городской подстилающей поверхности UCM
- Процентная составляющая соответствующей категории городской подстилающей поверхности в модельной ячейке

Сравнительные оценки приведены для трех синоптических станций (рис. 1): Огурцово (29638), Учебная (29637) и Обская ГМО (29635). Все станции расположены на левом берегу Новосибирска. Метеоплощадка Учебная расположена в черте города, окружена плотной городской застройкой. Станция Огурцово размещена на фактической границе интенсивной городской застройки, участков малоэтажной застройки и земель сельскохозяйственного назначения. Метеостанция Обская находится вблизи обского водохранилища, на территории удаленного городского микрорайона.

Сценарные расчеты проводились для категорий MODIS Modified (21 категория) полученных категорий OSM (33 категории, включая 3 категории городской застройки). Для различных симуляций проводились расчеты для сценариев (в скобках приведено горизонтальное разрешение модельной ячейки), результаты приведены в таблице:

**Mod1:** D3, категории землепользования Modis, без UCM (460м)

**Mod2:** D3, UCM включена, категории землепользования Modis (460м)

**Mod3:** D3, UCM включена, категории землепользования OSM, информация о категориях городской застройки (460м)

**Mod4:** D2, категории землепользования Modis, без UCM (1380м)

**Mod5:** D1, категории землепользования Modis, без UCM (6900м)

Анализ полученных результатов позволяет сформулировать следующие выводы:

Ошибка модели в зимний период превышает летние ошибки для всех сценариев.

Для летних сценариев наименьшее отклонение для всех сценариев показала модельная ячейка, соответствующая станции Учебная, для зимних – станции Обская ГМО.

Летом лучший результат по среднеквадратическому отклонению для станций Обская ГМО и Учебная показал сценарий Mod4 (разрешение ячейки 1380 м, без городской параметризации). Для станции Огурцово – Mod1 (разрешение ячейки 460 м, без городской параметризации).

Для зимнего периода наименьшая СКО получена для сценария Mod3 (460 м, подготовленные данные застройки, городская параметризация)

Для разрешений модельной ячейки 1380 и 6900 метров тренд к занижению температуры для всех рассмотренных случаев. Использование городской параметризации в зимний период уменьшает отрицательный сдвиг модельных значений температур.

На рис. 2 и 3 приведен пример среднесуточных ошибок между различными сценариями WRF и данными наблюдений на станции Огурцово.

Станция Огурцово показывает наибольший положительный сдвиг для временного ряда температуры в сценарии Mod3 для зимнего периода времени, что согласуется с её расположением внутри плотной городской застройки. Для летнего периода, явно выраженного различия для одного разрешения модельной ячейки нет, ход температуры имеет зависимость от горизонтального масштаба расчетного домена.

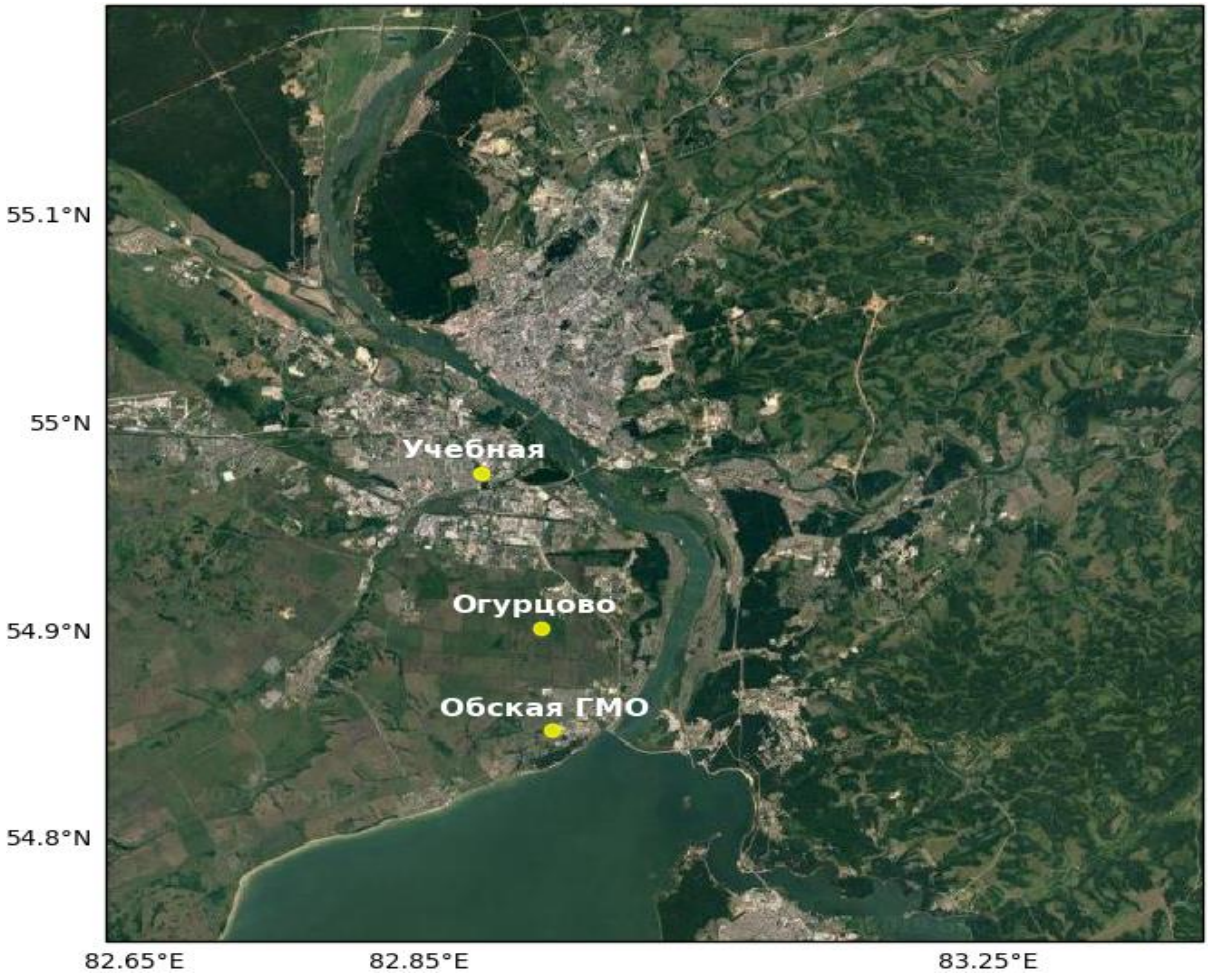



Рис. 1. Расположение метеорологических станций в пределах г. Новосибирск.

Результаты оценок вычислительного эксперимента для сценариев с двумя вариантами подстилающей поверхности и тремя разрешениями модельной ячейки для летнего и зимнего периодов для синоптических станций Огурцово (29638), Обская ГМО (29635), Учебная (29637). Оценки для среднесуточной ошибки, СКО, коэффициента корреляции.

| | сезон | *лето* | | | *зима* | | |
|---|---|---|---|---|---|---|---|
| | станция | **29638** | **29635** | **29637** | **29638** | **29635** | **29637** |
| **Mod1** | *смещение* | 0.01 | -0.32 | 0.28 | -1.14 | -0.45 | -2.14 |
| | *СКО* | 1.74 | 1.85 | 1.67 | 3.26 | 2.73 | 3.32 |
| | *корреляция* | 0.92 | 0.90 | 0.93 | 0.92 | 0.93 | 0.94 |
| **Mod2** | *смещение* | 0.04 | -0.38 | 0.33 | -0.99 | 0.08 | -0.95 |
| | *СКО* | 1.77 | 1.92 | 1.72 | 3.22 | 2.65 | 2.47 |
| | *корреляция* | 0.92 | 0.90 | 0.93 | 0.92 | 0.93 | 0.94 |
| **Mod3** | *смещение* | 0.36 | 0.08 | 0.26 | -0.30 | 0.50 | -0.11 |
| | *СКО* | 1.78 | 1.93 | 1.72 | 2.81 | 2.55 | 2.26 |
| | *корреляция* | 0.92 | 0.89 | 0.93 | 0.93 | 0.94 | 0.95 |
| **Mod4** | *смещение* | -0.37 | -0.43 | -0.21 | -1.93 | -0.24 | -2.97 |
| | *СКО* | 1.78 | 1.65 | 1.56 | 3.96 | 2.91 | 4.51 |
| | *корреляция* | 0.92 | 0.92 | 0.94 | 0.90 | 0.91 | 0.91 |
| **Mod5** | *смещение* | -0.50 | -0.93 | -0.76 | -1.95 | -1.21 | -2.91 |
| | *СКО* | 1.81 | 2.14 | 1.73 | 3.98 | 3.05 | 4.44 |
| | *корреляция* | 0.92 | 0.90 | 0.94 | 0.90 | 0.92 | 0.91 |

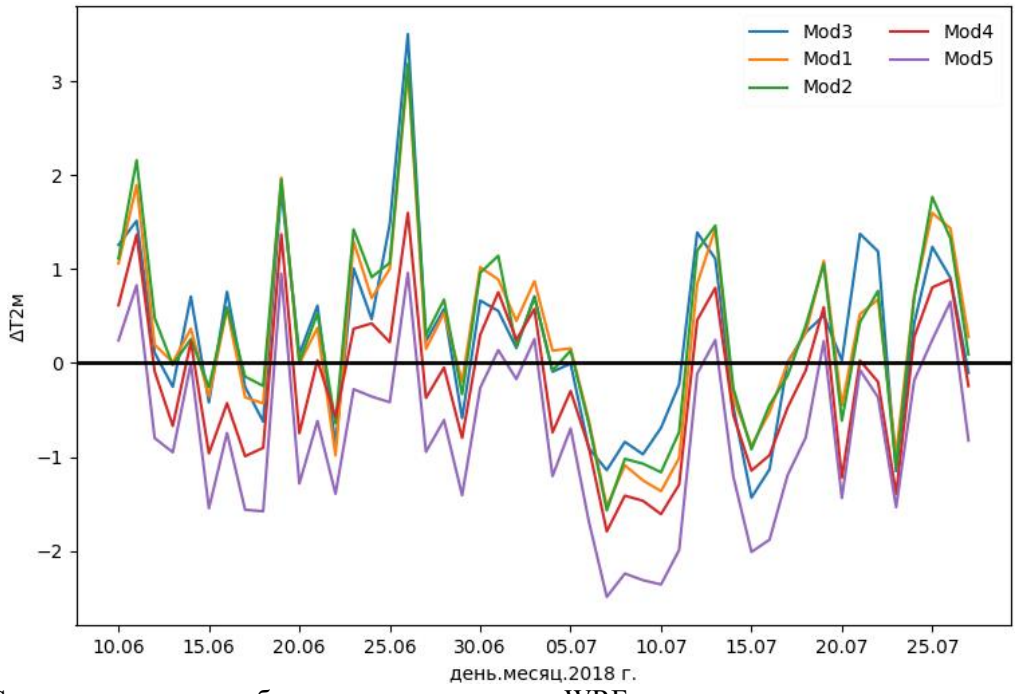



Рис. 2. Среднесуточные ошибки прогноза по модели WRF с двумя вариантами подстилающей поверхности и тремя разрешениями модельной ячейки для летнего периода.

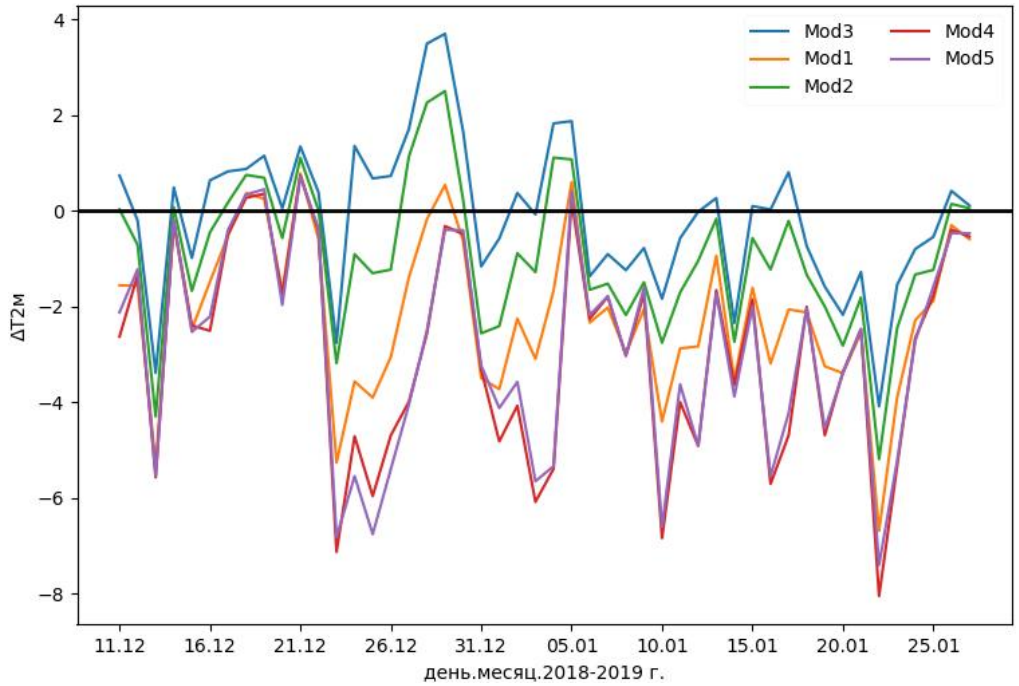



Рис. 2. Среднесуточные ошибки прогноза по модели WRF с двумя вариантами подстилающей поверхности и тремя разрешениями модельной ячейки для зимнего периода.

**Заключение.** По данным проекта OSM созданы категории землепользования, включая три категории городской застройки. Изучено влияние внешних модельных данных и горизонтального разрешения модельной сетки на прогноз температуры на уровне 2 метра. Результаты вычислительного эксперимента показали нецелесообразность использования вычислительно дорогостоящих расчетов высокого разрешения для летнего периода и физических параметризаций конвективного масштаба.

Для зимнего периода времени модельный прогноз оказывается чувствителен как к разрешению модельной ячейки, так и к качеству информации о городской подстилающей поверхности, показана чувствительность к городскому острову тепла в высоком разрешении, с использованием параметризации и полученных категорий городской подстилающей поверхности. Остров тепла оказывает значительный эффект на качество воздуха в условиях климатических особенностей Западной Сибири, развитием работы может стать изучение моделирование качества атмосферного воздуха в зимний период. Отдельно представляет интерес использование более сложных схем численного прогнозирования в пограничном слое, для которых качественное описания подстилающей поверхности является необходимым условием описания природных и антропогенных процессов в атмосфере.

*Работа выполнена при поддержке РФФИ (грант № 19-47-540011-р_а) и Правительства Новосибирской области.*

## ЛИТЕРАТУРА

[1] *Kusaka H., Kimura F.* Coupling a single-layer urban canopy model with a simple atmospheric model: Impact on urban heat island simulaton for an idealized case // J. Appl. Meteor. 2004. Vol. 43. P.1899-1910.

[2] *Grell G.A., Peckham S.E., Schmitz R. et al.* Fully coupled online chemistry within the WRF model // Atmospheric Environment. 2005. Vol. 39. P.6957-75.

[3] *Gochakov A.V., Penenko A.V., Antokhin P.N., Kolker A.B.* Air pollution modelling in urban environment based on a priori and reconstructed data // IOP Conf. Series: Earth and Env. Sci. 2018. Vol. 211.

[4] *Hong S.Y., Noh Y., Dudhia J.* A 2006 New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes // Monthly Weather Review. Vol. 134. P.2318-41.

# АНАЛИЗ КЛИМАТИЧЕСКИХ ХАРАКТЕРИСТИК ТЕРРИТОРИИ РАСПРОСТРАНЕНИЯ СИБИРСКОГО ШЕЛКОПРЯДА

*Дергунов А.В.[(1)(2)], Якубайлик О.Э.[(1)(2)(3)]*
[(1)] Федеральный Исследовательский Центр КНЦ СО РАН, Красноярск
[(2)] Институт Вычислительного Моделирования СО РАН, Красноярск
[(3)] Сибирский Федеральный Университет, Красноярск

Сибирский шелкопряд – один из главных вредителей хвойных лесов. В 2014 году произошла вспышка его численности в Енисейском районе Красноярского края. Лес левого берега реки Енисей, в отличие от правого берега, сильнее страдает от его воздействия. Цель работы состоит в анализе неоднородного поражения леса шелкопрядом на обоих берегах реки Енисей по метеоданным с 2009 по 2018 годы. Результаты показали, что левый берег сильнее прогрет по сравнению с правым берегом в среднем на 1-1,5 °С в течение рассматриваемого периода

*Ключевые слова: температура поверхности, сибирский шелкопряд, GFS, лесная нарушенность, река Енисей.*

**Введение.** Сибирский шелкопряд является одним из наиболее опасных вредителей хвойных лесов Сибири и Дальнего Востока, которые являются одним из важнейших природных ресурсов страны. Данный вид насекомых питается пихтой, сибирским кедром, елью и лиственницей. Низкий уровень осадков, засухи и относительно высокие температуры воздуха провоцируют массовое размножение этого насекомого в весенний и ранний летний периоды [1, 2]. На территории Красноярского края выделяются несколько районов вспышек численности шелкопряда: Чулымо-Кеский, Приангарский, Приенесейский, Канско-Бирюсинский, Кузнецко-Алатаусский, Западно-Саянский, Усинуский, Сибсим-Тубинский, Манско-Агульский. Наибольшей интенсивностью отличаются вспышки численности шелкопряда в темнохвойных лесах центральной части Красноярского края: Чулымо-Кетский, Приангарский и Приенесейские районы. Границей очагового распространения шелкопряда является температурный критерий – это средняя многолетняя температура августа, которая составляет +13.5°С [3].

Вспышки массового размножения этого насекомого приводят к серьезным экологическим и экономическим последствиям для лесного хозяйства регионов [3]. Так, в 1994-1996 гг. произошла одна из последних массовых вспышек численности сибирского шелкопряда на территории лесов Красноярского края, что привело к повреждению темнохвойных насаждений на площади около 700 тыс. га и гибели лесов на площади около 200 тыс. га [4, 5]. Экстремальным считается 2001 год, когда общая площадь очагов насекомых вредителей и болезней в России превысила 10 млн. га. На сибирского шелкопряда приходится почти 70% этой площади [6].

Выявлено, что после гибели темнохвойных лесов их восстановление затруднено. Сукцессии включают длительно-производную смену пород, а интенсивность процессов лесовосстановления зависит от размера очага повреждений и таких деструктивных факторов, как частые пожары. В ряде случаев темнохвойные леса не смогут восстановиться [7].

В 2014 году произошла очередная вспышка популяции сибирского шелкопряда в сибирских сосновых и пихтовых насаждениях в пределах Енисейской долины на юго-западе Красноярского края. Северная граница вспышки составила до 60°26’ с.ш., а общая площадь очага – 800 тыс. га. [8]. В данной работе рассматривается территория с координатами 89-92° в.д. и 58-60° с.ш., которая лежит между населенными пунктами: Подтесово и Ярцево. Выбранная территория разделена р. Енисей на два берега. На рис. 1 изображена рассматриваемая территория по данным портала Global Forest Change [9].

На рис. 1 красным цветом обозначены территории потери лесов за последние годы, а также четыре выбранные области и названия ближайших населенных пунктов. Здесь видно, что левый берег р. Енисей пострадал значительно больше от вредителя, чем правый.

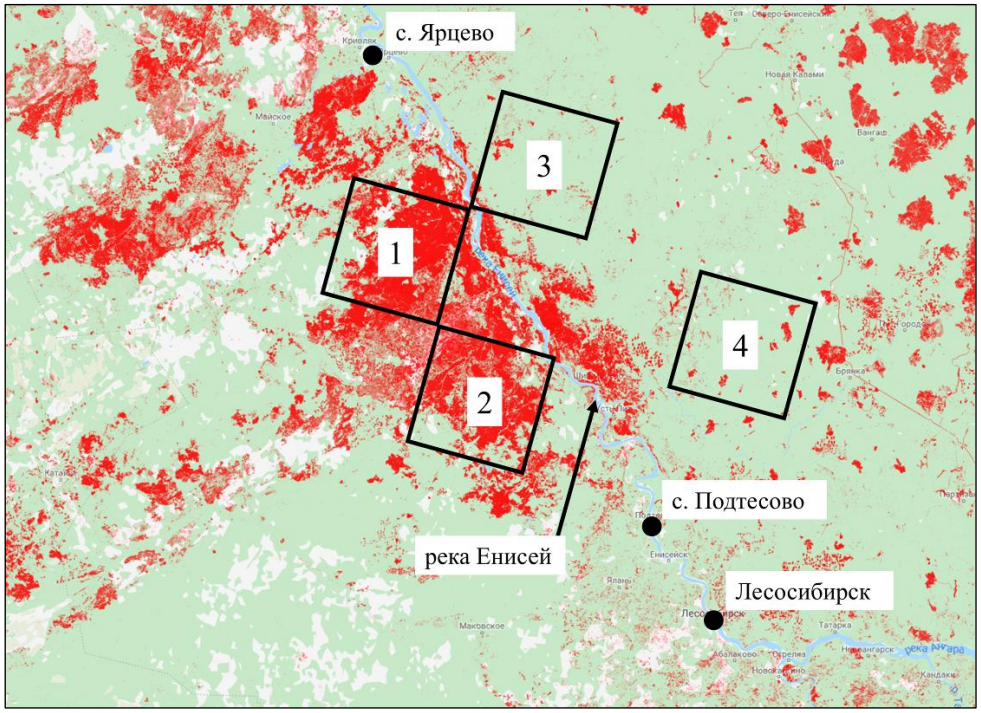


Рис. 1. Рассматриваемая территория.

Как известно, активация агрессивных видов насекомых-дендрофагов связана с изменением климата [10, 11]. Возникает вопрос: в чем отличительная особенность левого берега р. Енисей на рассматриваемой территории от правого с точки зрения климата? Представляет интерес пространственный анализ температуры поверхности для обоих берегов р. Енисей. В нашем исследовании были использованы метеорологические данные NOAA Global Forecast System [12] для решения этой проблемы.

Global Forecast System (GFS) – это глобальная система численного прогнозирования погоды, содержащая глобальную компьютерную модель и вариационный анализ, выполняемые Национальной метеорологической службой США (NWS). Данные GFS не защищены авторским правом и находятся в свободном доступе бесплатно. Это одна из самых известных мировых метеорологических моделей. Данные публикуются с интервалами 4 раза в день, а пространственное разрешение в настоящее время составляет 0.25 градуса (около 25 км для широты Красноярска). В 2019 году в результате недавнего десятикратного увеличения вычислительной мощности планируется обновить модель GFS, что увеличит ее горизонтальное разрешение в три раза до 9 км.

Для принятия эффективных управленческих мер в решении задач защиты леса и оптимизации расходов необходимо понимать причины вспышек численности насекомых-дендрофагов. Климатические данные могут служить ключевым инструментом в определении задач по выявлению этих причин. В связи с этим задачей данной работы является анализ ситуации с неоднородным поражением лесного массива сибирским шелкопрядом на обоих берегах р. Енисей в Красноярском крае по метеорологическим данным GFS.

Для выполнения поставленной задачи на рассматриваемой территории были выбраны 4 пробных области: две на левом берегу р. Енисей и две – на правом (см. рис. 1). Таким образом, были проанализированы две территории: левый берег р. Енисей, который характеризуется высокой степенью поражения шелкопрядом, и правый берег – с относительно невысокой лесной нарушенностью.

**Материалы и методы.** GFS (Global Forecast System) – это модель прогноза погоды, разработанная US National Centers for Environmental Prediction (NCEP). Десятки атмосферных и наземно-почвенных переменных доступны в этом наборе данных, от температуры, ветра и количества осадков, до показателей влажности почвы и концентрации атмосферного озона. GFS

покрывает весь земной шар с базовым горизонтальным разрешением 28 км между точками сетки, которое используется синоптиками для прогноза погоды до 16 дней вперед.

Точность модели постоянно улучшается. В частности, данные с пространственным разрешением 1° доступны с марта 2004 года, с разрешением 0.5° с января 2007 года и с разрешением 0.25° с января 2015 года [12, 13]. Данные GFS являются файлами специального формата *.grib2, который был разработан Всемирной Метеорологической Организацией (ВМО). Этот формат данных является стандартным в задачах хранения исторических и прогнозируемых данных о погоде. Например, каждый отдельный GFS файл типа «analysis», с пространственным разрешением 0.25° содержит 354 слоя различной метеорологической информации и обладает размером в среднем около 200 Mb.

Мы использовали фактические данные анализа метеорологических условий из архива NCEP, а не прогностическую информацию. Их пространственное разрешение составило 0.5° для 2009-2014 гг. и 0.25° для 2015-2018 гг.

Для реализации поставленной задачи все необходимые метеорологические данные GFS были загружены с официального сайта NOAA (National Oceanic and Atmospheric Administration) [12, 14].

После загрузки наборов данных GFS за 10 летний период они были предварительно преобразованы. А именно, все полученные данные были обрезаны по четырем выбранным областям рассматриваемой территории и слою «TMP:surface», то есть температуре поверхности. Для осуществления этой задачи была использована программа wgrib2. Данная программа специально разработана программистами NCEP для чтения и записи метеорологических данных формата *.grib2.

Для автоматизации обработки исходных данных был написан специальный скрипт, представляющий собой командный файл для ОС Windows. Данный скрипт инициирует циклическое чтение исходных файлов и запускает программу wgrib2 [15] с определенными параметрами для каждого из них. Эти параметры удовлетворяют задаче, а именно, они содержат координаты четырех выбранных областей и название нужного слоя данных. Выходные данные – это файлы, которые занимают размер более чем в сто раз меньше, чем исходные данные.

После предварительного преобразования результирующих данных GFS они должны быть приведены к табличному виду для последующего анализа. Для этого был также написан простой пакетный скрипт, который запускает программу wgrib2 с определенным параметром и преобразует каждый полученный файл в формат файла *.csv.

На языке программирования Си была написана специальная программа, которая считывает все отдельные файлы формата *.csv по четырем областям за 10 летний период и сводит их содержимое в общий файл того же формата.

Окончательная обработка и анализ преобразованных данных GFS проводились в Microsoft Excel.

**Результаты и обсуждение.** В результате обработки исходных данных был сформирован архив данных температуры поверхности за летние месяцы за 2009-2018 гг. для четырех выбранных областей рассматриваемой территории.

На рис. 2 представлены графики усредненных по летним месяцам значений температуры поверхности *T* по четырем выбранным областям за 10 лет. Здесь сплошными линиями обозначены области 1 и 2, которые соответствуют левому берегу р. Енисей, а прерывистыми линиями – области 3 и 4, которые расположены на правом берегу.

Следует отметить, что значения температуры поверхности *T* для областей 1 и 2 выше значений *T* для областей 3 и 4 в среднем на 1-1.5° в течение выбранного интервала времени с 2009 по 2018. Это означает, что поверхность зоны, на которую воздействует шелкопряд, прогревается сильнее, чем территория с относительно небольшим ущербом лесной растительности.

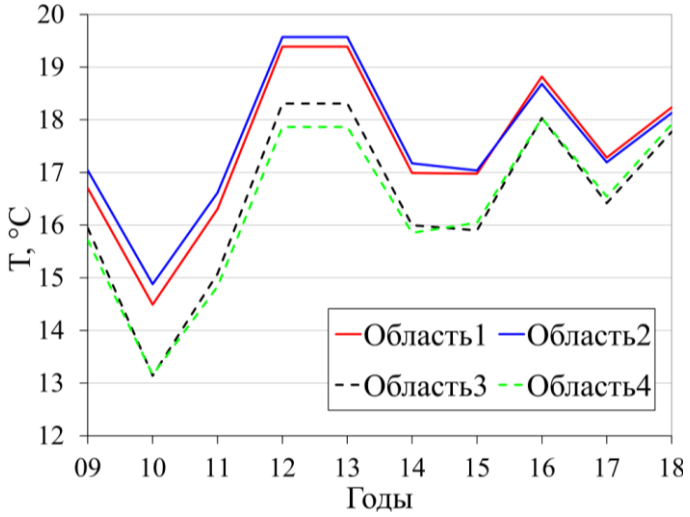



Рис. 2. Средне-летняя температура поверхности за 10 лет для четырех пробных областей.

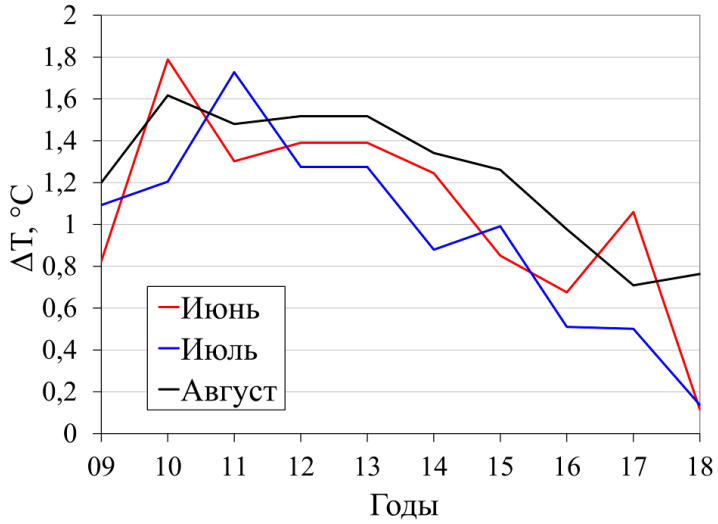



Рис. 3. Рассчитанная разица температур подслитающей поверхности $\Delta T$ левого и правого берегов р. Енисей за июнь, июль и август с 2009 по 2018 гг.

Для более детального анализа данных из сформированного архива было проведено усреднение значений температуры поверхности $T$ по летним месяцам двух выбранных областей на левом берегу р. Енисей и двух областей на правом берегу и подсчитана их разница $\Delta T$ (рис. 3).

На рис. 3 показано, что в год вспышки (2014 г.) популяции сибирского шелкопряда на рассматриваемой территории средняя температура поверхности значительно пострадавшего левого берега р. Енисей была выше в июне на 1.24°C, июле – 0.9°C, августе – 1.34°C, по сравнению с менее пораженным правым берегом. Графики также показывают, что разница средних температур поверхности $\Delta T$ между левым и правым берегами р. Енисей уменьшается с течением времени.

**Выводы.** В последнее время новые метеорологические данные с высоким пространственным разрешением открыли принципиально новые возможности анализа различных природных явлений и процессов. В частности, они помогают выявить возможные причины наблюдаемых аномалий размножения и распространения вредителей лесов. Анализ полученных данных показывает, что незначительное отклонение средней температуры поверхности может существенно повлиять на рост популяции и активность сибирского шелкопряда на левом берегу реки Енисей рассматриваемой территории, в отличие от правого берега. Вероятно, температура поверхности может быть одним из ключевых климатических факторов, влияющих на вспышки массового размножения сибирского шелкопряда.

ЛИТЕРАТУРА


[1] *Kolomiets N.G. and Maier E.I.* Principal Forest Pests of Tomsk Province. Tomsk: Knizhnoe Izdatelstvo, 1963

[2] *Kondakov Y.P.* Mass reproduction of Siberian silk moth in forests of Krasnoyarsk krai // Entomologicheskie issledovaniya v Sibiri. 2002. С. 25-74.

[3] *Стеганцов Р.И.* Насекомые и фитофаги // Молодежь и наука XXI века. 2018. С. 62-65

[4] *Барталев С.А., Ершов Д.В., Исаев А.С.* Оценка дефолиации лесов по многоспектральным спутниковым изображениям методом декомпозиции спектральных смесей // Исследование Земли из космоса. 1999. № 4. С. 78-86.

[5] *Исаев А.С. и др.* Катастрофические воздействия антропогенных и природных факторов на лесные экосистемы // Изменение окружающей среды и климата. Природные и связанные с ними техногенные катастрофы. 8т. / Пред. ред. кол. НП Лаверов. РАН. М.: ИФЗ РАН, 2008. Т. 4. С. 66-79.

[6] *Брюханов А.* Экологическая оценка состояния лесов в Сибири: тревожные итоги // Устойчивое лесопользование. 2009. № 2. С. 21-31.

[7] *Гродницкий Д.Л.* Сибирский шелкопряд и судьба пихтовой тайги // Природа. 2004. №. 11. С. 49-56.

[8] *Kharuk V.I., Im S.T., Yagunov M.N.* Migration of the Northern Boundary of the Siberian Silk Moth // Contemporary problems of ecology. 2018. Vol. 11, N. 1. P. 26-34.

[9] Global Forest Change. **https://earthenginepartners.appspot.com/science-2013-global-forest** (дата обращения 30.05.2019).

[10] *Bentz B.J. et al.* Climate change and bark beetles of the western United States and Canada: direct and indirect effects // BioScience. 2010. Vol. 60, N 8. P. 602-613.

[11] *Weed A.S., Ayres M.P., Hicke J.A.* Consequences of climate change for biotic disturbances in North American forests // Ecological Monographs. 2013. Vol. 83, N 4. P. 441-470.

[12] Global Forecast System (GFS). **https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs** (дата обращения 30.05.2019).

[13] NCEP Global Forecast System (GFS). **https://www.emc.ncep.noaa.gov/GFS/doc.php** (дата обращения 30.05.2019).

[14] NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive. **https://rda.ucar.edu/datasets/ds084.1/index.html#!access** (дата обращения 30.05.2019).

[15] wgrib2: wgrib for GRIB-2. **https://www.cpc.ncep.noaa.gov/products/wesley/wgrib2/index.html** (дата обращения 30.05.2019).

# ДИФФЕРЕНЦИРОВАННЫЙ ПОДХОД К ОЦЕНКЕ ИНДИВИДУАЛЬНОГО РИСКА ЧС И ПРОИСШЕСТВИЙ

*Иванова У.С.[(1)(2)], Черных Д.А.[(1)(3)], Тасейко О.В.[(1)(3)], Никулина Т.Н.[(2)]*
[(1)] Красноярский филиал Института вычислительных технологий СО РАН, Красноярск
[(2)] Сибирский федеральный университет, Красноярск
[(3)] Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, Красноярск

Рассмотрена проблема устойчивого развития промышленных территорий на основе риск-ориентированного подхода. Цель работы – анализ методических рекомендаций по оценки индивидуального риска ЧС и происшествий. Актуальность определяется необходимостью реализации стратегии национальной безопасности на региональном уровне на основе разработки методов анализа техногенной безопасности. В условиях существующего уровня риска и негативных последствий необходима организация и модернизация системы управления территориальными рисками. Итогом работы является ранжирование территорий Красноярского края по уровню техногенной опасности (опасные, пограничные, безопасные) с использованием дифференцированного коэфициента.
*Ключевые слова: безопасность, устойчивое развитие, уровень риска, индивидуальные риски ЧС и проишествий*

**Введение.** При оценках комплексной безопасности территориальных образований большое внимание уделяется разработке и использованию критериев и методов анализа рисков. Особое внимание направлено на обеспечение безопасности жизнедеятельности человека и общества. В настоящее время ведется активная научная и организационная деятельность по вопросам снижения риска аварий и катастроф и сокращения потерь в результате чрезвычайных ситуаций (ЧС) [1-4]. Цель работы – анализ методических рекомендаций по оценки индивидуального риска ЧС и происшествий.

Актуальность определяется необходимостью реализации стратегии национальной безопасности на региональном уровне на базе разработки методов анализа техногенной безопасности.

В работе поставлены следующие задачи:

- расчет и анализ индивидуальных рисков ЧС и происшествий техногенного характера на примере Красноярского края по утвержденной методике;
- введение дифференцированного коэффициента в методику по оценки риска;
- пересчет индивидуальных рисков ЧС.

**Анализ метода оценки индивидуального риска ЧС и происшествий.** Официальной методикой по оценки индивидуального риска ЧС и происшествий является «Методические рекомендации по порядку разработки, проверки, оценки и корректировки электронных паспортов территорий (объектов)» (утв. МЧС России 15.07.2016 № 2-4-71-40).

При выявлении рисков возникновения ЧС и происшествий исходными данными являются: общая информация о субъекте Российской Федерации (муниципальном образовании), численность и плотность населения, характер местности, информация для оценки возможных последствий ЧС. Рассматриваются статистические показатели, отражающие определенный риск, оценивается ожидаемое влияние риска на определенную территорию или объект экономики, оцениваются ресурсы, необходимые для управления рисками.

Расчет индивидуального риска производится для установления реального уровня риска (приемлемый риск $R < 10^{-5}$), с целью определения необходимых мероприятий по управлению и снижению риска. Расчет индивидуального риска ЧС и происшествий производится по формуле:

$$R = \frac{N_{\text{п}}}{N_{\text{н}}} \qquad (1)$$

где $N_{\text{п}}$ – среднее количество погибших в год за последние 5 лет при определенном виде ЧС и происшествии на заданной территории; $N_{\text{н}}$ – количество населения, проживающего на данной территории.

Расчет индивидуального риска ЧС и происшествий выполнен на примере муниципальных образований Красноярского края на основании статистической информации базы данных АИУС РСЧС и методических рекомендаций [5]. Красноярский край – индустриально развитый субъект Российской Федерации в Сибирском федеральном округе, который подвержен рискам возникновения техногенных аварийных ситуаций, связанных с выбросом радиоактивных веществ, выбросом АХОВ, затоплением при разрушении гидротехнических сооружений, крупными производственными авариями и пожарами, лесными пожарами, наводнениями, крупными ДТП, авариями на авиа и железнодорожном транспорте и пр. В Красноярский край входят следующие административно-территориальные единицы: 13 краевых городов, 3 ЗАТО и 44 муниципальных районов и 1 поселок городского типа [6].

На рис. 1 представлено ранжирование территорий Красноярского края по уровню индивидуального риска (среднее значение за 20 лет). Для муниципальных образований получен диапазон значений индивидуального риска ЧС и происшествий техногенного характера от 0 до $2 \cdot 10^{-4}$. Наиболее опасными являются Боготольский, Большеулуйский, Козульский, Манский, Туруханский и Эвенкийский районы края (значение индивидуального риска превышает приемлемый уровень более чем в 10 раз).

По данным расчета видно, что только 14 муниципальных образований не превышают приемлемый уровень риска. В большинстве случаев высокое значение риска связано с низкой численностью населения на заданной территории, поэтому необходим дифференцированный подход к оценке индивидуального риска путем введения соответствующего коэффициента $k$. Приемлемый уровень риска равный $1 \cdot 10^{-5}$ соответствует гибели одного человека из ста тысяч населения, однако есть территории, численность населения которых ниже этого значения или превышает его во много раз. Введение коэффициента $k$ позволит более точно определять и анализировать территории по степени риска и безопасности.

Коэффициент определяется двумя способами: если $N \leq 100000$ и $N > 100000$. В табл. 1 представлены значения коэффициента для первого случая.

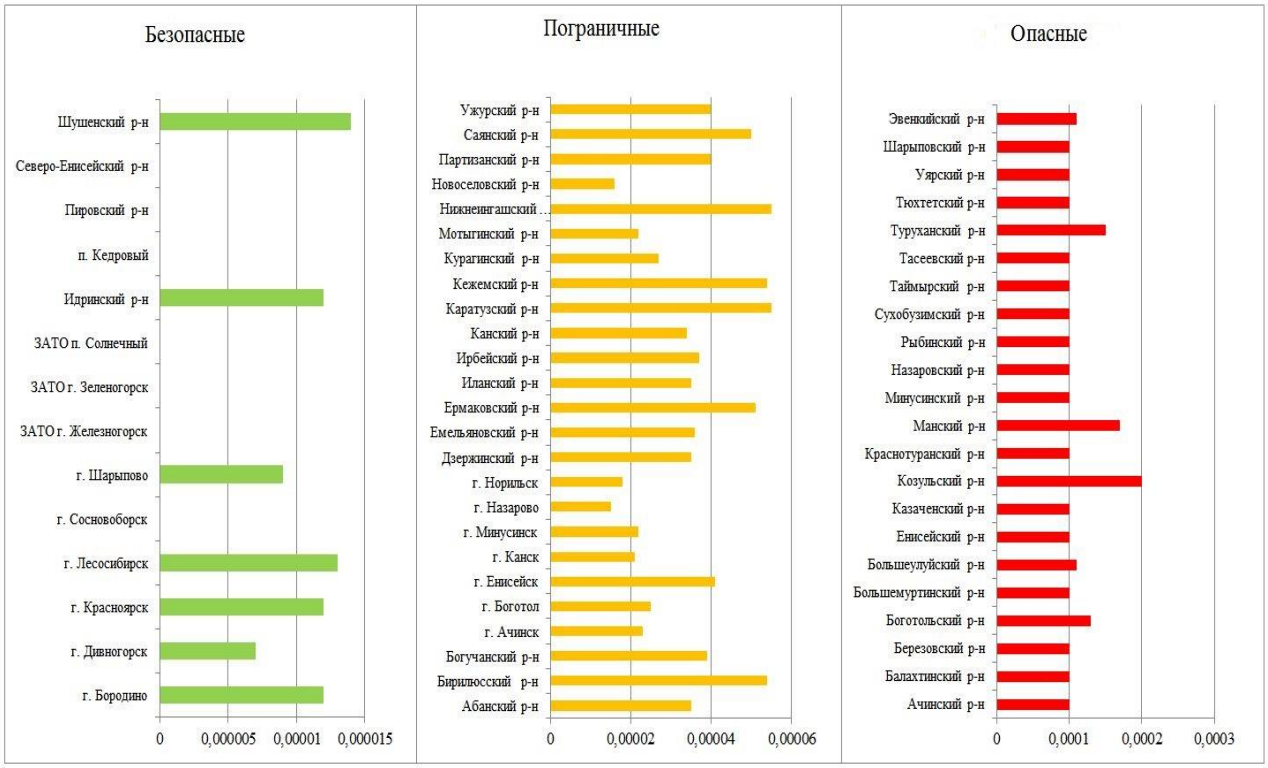



Рис. 1. Ранжирование муниципальных образований по уровню риска.

Для второго случая, коэффициент рассчитывается по формуле:

$$k = \frac{N}{100000}, \quad (2)$$

где $N$ – численность населения на заданной территории, превышающая 100000 человек.

В табл. 2 представлены исходные значения индивидуального риска ЧС и происшествий в пересчете с коэффициентом $k$ по районам Красноярского края, с последующим ранжированием территорий.

Табл. 1. Значения коэффициента при численности до 100000 человек.

| **Численность населения** | **Значения $k$** |
|---|---|
| 1-20000 | 0,1 |
| 20000-30000 | 0,2 |
| 30000-40000 | 0,3 |
| 40000-50000 | 0,4 |
| 50000-60000 | 0,5 |
| 60000-70000 | 0,6 |
| 70000-80000 | 0,7 |
| 80000-90000 | 0,8 |
| 90000-100000 | 0,9 |

Табл. 2. Дифференцированное ранжирование территорий Красноярского края по уровню индивидуального риска.

| **Район** | **Индивидуальный риск ЧС и происшествий ($R$)** | **Численность населения [7]** | **$k$** | **Пересчетное значение риска ($R'$)** | **Приемлемый риск** | **Ранжирование территорий** |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Абанский р-н | $3{,}5\cdot10^{-5}$ | 19951 | 0,1 | $3{,}5\cdot10^{-6}$ | $1\cdot10^{-5}$ | безопасная |
| Ачинский р-н | $1\cdot10^{-4}$ | 15213 | 0,1 | $1\cdot10^{-5}$ | | безопасная |
| Балахтинский р-н | $1\cdot10^{-4}$ | 18664 | 0,1 | $1\cdot10^{-5}$ | | безопасная |
| Березовский р-н | $1\cdot10^{-4}$ | 41727 | 0,4 | $4\cdot10^{-5}$ | | пограничная |
| Бирилюсский р-н | $5{,}4\cdot10^{-5}$ | 9649 | 0,1 | $5{,}4\cdot10^{-6}$ | | безопасная |
| Боготольский р-н | $1{,}3\cdot10^{-4}$ | 9751 | 0,1 | $1{,}3\cdot10^{-5}$ | | безопасная |
| Богучанский р-н | $3{,}9\cdot10^{-5}$ | 45525 | 0,4 | $1{,}6\cdot10^{-5}$ | | пограничная |
| Большемуртинский р-н | $1\cdot10^{-4}$ | 18243 | 0,1 | $1\cdot10^{-5}$ | | безопасная |
| Большеулуйский р-н | $1{,}1\cdot10^{-4}$ | 7525 | 0,1 | $1{,}1\cdot10^{-5}$ | | безопасная |
| г. Ачинск | $2{,}3\cdot10^{-5}$ | 106531 | 1,06 | $2{,}4\cdot10^{-5}$ | | пограничная |
| г. Боготол | $2{,}5\cdot10^{-5}$ | 20020 | 0,2 | $5\cdot10^{-6}$ | | безопасная |
| г. Бородино | $1{,}2\cdot10^{-5}$ | 16127 | 0,1 | $1{,}2\cdot10^{-6}$ | | безопасная |
| г. Дивногорск | $7\cdot10^{-6}$ | 33490 | 0,3 | $2{,}1\cdot10^{-6}$ | | безопасная |
| г. Енисейск | $4{,}1\cdot10^{-5}$ | 17826 | 0,1 | $4{,}1\cdot10^{-6}$ | | безопасная |
| г. Канск | $2{,}1\cdot10^{-5}$ | 89508 | 0,8 | $1{,}7\cdot10^{-5}$ | | пограничная |
| г. Красноярск | $1{,}2\cdot10^{-5}$ | 1091634 | 10,9 | $1{,}3\cdot10^{-4}$ | | опасная |
| г. Лесосибирск | $1{,}3\cdot10^{-5}$ | 64323 | 0,6 | $7{,}8\cdot10^{-6}$ | | безопасная |
| г. Минусинск | $2{,}2\cdot10^{-5}$ | 70910 | 0,7 | $1{,}5\cdot10^{-5}$ | | пограничная |
| г. Назарово | $1{,}5\cdot10^{-5}$ | 49991 | 0,4 | $6\cdot10^{-6}$ | | безопасная |
| г. Норильск | $1{,}8\cdot10^{-5}$ | 180239 | 1,8 | $3{,}2\cdot10^{-5}$ | | пограничная |
| г. Сосновоборск | 0 | 40128 | 0,4 | 0 | | безопасная |
| Енисейский р-н | $1\cdot10^{-4}$ | 22828 | 0,2 | $2\cdot10^{-5}$ | | пограничная |
| Ермаковский р-н | $5{,}1\cdot10^{-5}$ | 19334 | 0,1 | $5{,}1\cdot10^{-6}$ | | безопасная |

Табл. 2. Дифференцированное ранжирование территорий Красноярского края по уровню индивидуального риска (продолжение).

| **Район** | **Индивидуальный риск ЧС и происшествий (*R*)** | **Численность населения [7]** | ***k*** | **Пересчетное значение риска (*R'*)** | **Приемлемый риск** | **Ранжирование территорий** |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| ЗАТО г. Железногорск | 0 | 92851 | 0,9 | 0 | $1 \cdot 10^{-5}$ | безопасная |
| ЗАТО г. Зеленогорск | 0 | 62245 | 0,6 | 0 | | безопасная |
| ЗАТО п. Солнечный | 0 | 9950 | 0,1 | 0 | | безопасная |
| Идринский р-н | $1{,}2 \cdot 10^{-5}$ | 11183 | 0,1 | $1{,}2 \cdot 10^{-6}$ | | безопасная |
| Иланский р-н | $3{,}5 \cdot 10^{-5}$ | 23806 | 0,2 | $7 \cdot 10^{-6}$ | | безопасная |
| Ирбейский р-н | $3{,}7 \cdot 10^{-5}$ | 15 468 | 0,1 | $3{,}7 \cdot 10^{-6}$ | | безопасная |
| Казаченский р-н | $1 \cdot 10^{-4}$ | 9643 | 0,1 | $1 \cdot 10^{-5}$ | | безопасная |
| Канский р-н | $3{,}4 \cdot 10^{-5}$ | 25316 | 0,2 | $6{,}8 \cdot 10^{-6}$ | | безопасная |
| Каратузский р-н | $5{,}5 \cdot 10^{-5}$ | 14950 | 0,1 | $5{,}5 \cdot 10^{-6}$ | | безопасная |
| Кежемский р-н | $5{,}4 \cdot 10^{-5}$ | 20674 | 0,2 | $1{,}1 \cdot 10^{-5}$ | | безопасная |
| Козульский р-н | $2 \cdot 10^{-4}$ | 16246 | 0,1 | $2 \cdot 10^{-5}$ | | пограничная |
| Краснотуранский р-н | $1 \cdot 10^{-4}$ | 14067 | 0,1 | $1 \cdot 10^{-5}$ | | безопасная |
| Курагинский р-н | $2{,}7 \cdot 10^{-5}$ | 44977 | 0,4 | $1{,}1 \cdot 10^{-5}$ | | безопасная |
| Манский р-н | $1{,}7 \cdot 10^{-4}$ | 15668 | 0,1 | $1{,}7 \cdot 10^{-5}$ | | пограничная |
| Минусинский р-н | $1 \cdot 10^{-4}$ | 25944 | 0,2 | $2 \cdot 10^{-5}$ | | пограничная |
| Мотыгинский р-н | $2{,}2 \cdot 10^{-5}$ | 13891 | 0,1 | $2{,}2 \cdot 10^{-6}$ | | безопасная |
| Назаровский р-н | $1 \cdot 10^{-4}$ | 22063 | 0,2 | $2 \cdot 10^{-5}$ | | пограничная |
| Нижнеингашский р-н | $5{,}5 \cdot 10^{-5}$ | 29422 | 0,2 | $1{,}1 \cdot 10^{-5}$ | | безопасная |
| Новоселовский р-н | $1{,}6 \cdot 10^{-5}$ | 12969 | 0,1 | $1{,}6 \cdot 10^{-6}$ | | безопасная |
| п. Кедровый | 0 | 5450 | 0,1 | 0 | | безопасная |
| Партизанский р-н | $4 \cdot 10^{-5}$ | 9283 | 0,1 | $4 \cdot 10^{-6}$ | | безопасная |
| Пировский р-н | 0 | 6867 | 0,1 | 0 | | безопасная |
| Рыбинский р-н | $1 \cdot 10^{-4}$ | 30943 | 0,3 | 0 | | безопасная |
| Саянский р-н | $5 \cdot 10^{-5}$ | 10746 | 0,1 | $5 \cdot 10^{-6}$ | | безопасная |
| Северо-Енисейский р-н | 0 | 11090 | 0,1 | 0 | | безопасная |
| Сухобузимский р-н | $1 \cdot 10^{-4}$ | 20064 | 0,2 | $2 \cdot 10^{-5}$ | | пограничная |
| Таймырский Долгано-Ненецкий р-н | $1 \cdot 10^{-4}$ | 31762 | 0,3 | $3 \cdot 10^{-5}$ | | пограничная |
| Тасеевский р-н | $1 \cdot 10^{-4}$ | 11508 | 0,1 | $1 \cdot 10^{-5}$ | | безопасная |
| Туруханский р-н | $1{,}5 \cdot 10^{-4}$ | 15971 | 0,1 | $1{,}5 \cdot 10^{-5}$ | | пограничная |
| Тюхтетский р-н | $1 \cdot 10^{-4}$ | 8077 | 0,1 | $1 \cdot 10^{-5}$ | | безопасная |
| Ужурский р-н | $4 \cdot 10^{-5}$ | 31408 | 0,3 | $1{,}2 \cdot 10^{-5}$ | | безопасная |
| Уярский р-н | $1 \cdot 10^{-4}$ | 20715 | 0,2 | $2 \cdot 10^{-5}$ | | пограничная |
| Шарыповский р-н | $1 \cdot 10^{-4}$ | 14176 | 0,1 | $1 \cdot 10^{-5}$ | | безопасная |
| Шушенский р-н | $1{,}4 \cdot 10^{-5}$ | 32164 | 0,3 | $4{,}2 \cdot 10^{-6}$ | | безопасная |
| Эвенкийский р-н | $1{,}1 \cdot 10^{-4}$ | 15147 | 0,1 | $1{,}1 \cdot 10^{-5}$ | | безопасная |

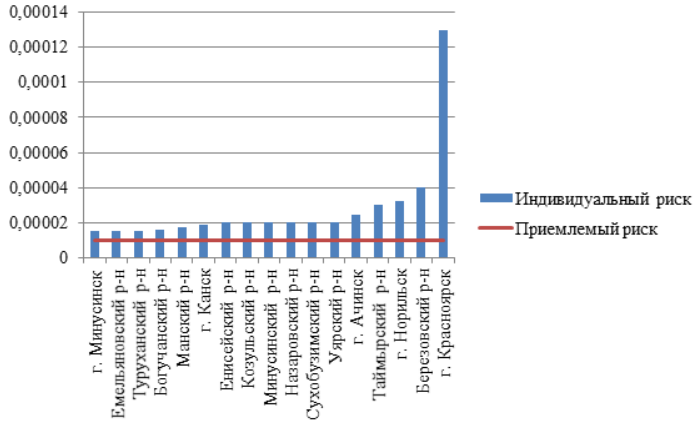



Рис. 2. Муниципальные образования Красноярского края с повышенным индивидуальным риском ЧС и происшествий.

При пересчете индивидуального риска при ЧС и происшествиях техногенного характера получен диапазон значений от 0 до $1{,}3\cdot10^{-4}$. На рис. 2 представлены муниципальные образования, в которых значения индивидуального риска превысили предельно-допустимый уровень. Наибольшее значение индивидуального риска при ЧС и происшествиях техногенного характера получено для г. Красноярска ($R'=1{,}3\cdot10^{-4}$). Данный показатель обусловлен наличием большого числа потенциально-опасных, критически и стратегически важных объектов.

**Заключение.** Управление техносферной безопасностью связано с прогнозированием возможных рисков, а также разработкой и совершенствованием моделей и методов анализа и оценки техногенных рисков. С этой целью в работе получены результаты:

- выполнен анализ методических рекомендаций по оценки индивидуального риска ЧС и происшествий;
- представлены расчет и анализ индивидуальных рисков ЧС и происшествий для территорий Красноярского края по методическим рекомендациям;
- проведен перерасчет индивидуального риска с дифференцированным коэффициентом.

Анализ техногенной опасности территорий показывает, что наибольшему риску подвержены крупные города с развитой инфраструктурой и наличием опасных объектов.

*Исследование выполнено при финансовой поддержке РФФИ, Правительства Красноярского края и Красноярского краевого фонда науки в рамках научного проекта: «Методы и информационные технологии оценки рисков развития социально-природно-техногенных систем промышленного региона».*

## ЛИТЕРАТУРА

[1] *Махутов Н.А., Гаденин М.М.* Экологическая безопасность и экологическое наследие в проблемах национальной безопасности // Экология и промышленность России. 2016. Т 20, № 3. С. 47-51.

[2] Безопасность России. Правовые, социально-экономические и научно-технические аспекты. Региональные проблемы безопасности. Красноярский край. М.: МГФ «Знание», 2001. 576 с.

[3] *Москвичев В.В., Бычков И.В., Потапов В.П., Тасейко О.В., Шокин Ю.И.* Информационная система территориального управления рисками развития и безопасностью // Вестник Российской академии наук. 2017. № 8. С. 696-705.

[4] Безопасность России. Правовые, социально-экономические и научно-технические аспекты. Системные исследования чрезвычайных ситуаций. М.: МГОФ «Знание», 2015. 864 с.

[5] Методические рекомендации по порядку разработки, проверки, оценки и корректировки электронных паспортов территорий (объектов) (утв. МЧС России 15.07.2016 № 2-4-71-40).

[6] О состоянии защиты населения и территорий Красноярского края от чрезвычайных ситуаций природного и техногенного характера: Государственный доклад. М.: МЧС России, 1996-2016 гг.

[7] **http://www.krskstate.ru/msu/terdel** [Электронный ресурс] (дата обращения 18.06.2019).

# АНАЛИЗ ДАННЫХ О ПРЕДВЕСТНИКАХ СИЛЬНЫХ ЗЕМЛЕТРЯСЕНИЙ

*Сибгатулин В.Г.[(1)], Перетокин С.А.[(2)], Кабанов А.А.[(1)(2)], Симонов К.В.[(3)]*
[(1)] НП «Экологический центр рационального освоения природных ресурсов», Красноярск
[(2)] Красноярский филиал Института вычислительных технологий СО РАН –
Специальное конструкторско-технологическое бюро «Наука», Красноярск
[(3)] Институт вычислительного моделирования СО РАН, Красноярск

Работа посвящена развитию методов геомониторинга для прогноза сильных землетрясений и оценки напряжённо-деформированное состояние геологической среды. Приводятся результаты наблюдений предвестников различной природы на геодинамических полигонах в Сибири и на Дальнем Востоке в период 2007-2018 гг.
*Ключевые слова: геомониторинг, землетрясение, предвестники, прогноз, напряжённо-деформированное состояние геосреды.*

**Введение.** Геодинамический мониторинг является обязательным элементом системы безопасности в сейсмически активных регионах. Научно-методические основы геодинамического мониторинга исторически развиваются в рамках РАН. Кроме мониторинговых сетей РАН, на территории РФ функционируют геодинамические полигоны различных ведомств (Минобороны, МЧС РФ, МПР, Минэнерго, Минатом и т.д.). Начиная с 2000 г., получили развитие и региональные наблюдательные геодинамические сети в различных субъектах федерации (Красноярский край, Кемеровская область, республика Тыва и др.). При этом, используются как сейсмологические сети, так и комплексные сети, регистрирующие различные геолого-геофизические параметры.

В МПР РФ мониторинг эндогенных процессов базируется на гидрогеодеформационном поле Земли (ГГД мониторинг). В различных сейсмоактивных регионах России (Кавказ, Алтае-Саянская область, Камчатка, Сахалин) получили развитие геофизические полигоны для контроля напряжённо-деформированного состояния (НДС) геологической среды в промышленных зонах, включая крупные города, для прогноза сейсмической опасности. При этом, идеология развития геодинамических полигонов МПР РФ базируется на комплексировании различных геолого-геофизических методов (сейсмология, электромагнитные измерения, ГГД мониторинг, мониторинг газов и т.д.). В связи с многообразием энергетических процессов в геологической среде, объективно необходима регистрация (мониторинг) различных по физической природе методов. Таким образом, комплексный подход к организации геодинамического мониторинга является своеобразным «стандартом» МПР РФ в отличие от «моносетей» геодинамического мониторинга различных ведомств.

Вместе с тем, несмотря на длительное использование комплекса геолого-геофизических методов, применяемых на геодинамическом мониторинге, нормативно-методическая основа упомянутого комплекса не разработана. В основном регламентированы такие виды геодинамического мониторинга как: ГГД мониторинг [1]; разработаны нормативные документы для организации и эксплуатации сейсмологических сетей в системе РАН; электромагнитный мониторинг, в частности на основе регистрации естественного импульсного электромагнитного поля Земли (ЕИЭМПЗ), осуществляемого на основе авторских разработок [2]; мониторинг радона и других газов, осуществляемый также на основе разработок различных исследователей [3].

За последние 10 лет накоплен значительный экспериментальный материал геолого-геофизических методов, используемых на геодинамических полигонах МПР РФ. В настоящей работе обобщены данные комплекса геолого-геофизических методов (сейсмология, ЕИЭМПЗ, эмиссия радона) на геодинамических полигонах в Сибири и на Дальнем востоке в 2007-2018 гг. для оценки изменения НДС геологической среды и прогноза сейсмических событий. По нашему мнению, несмотря на дискуссионность ряда выводов, настоящее обобщение может использоваться для разработки нормативно-методических рекомендаций комплексирования

геолого-геофизических методов на геодинамических полигонах с целью контроля НДС геологической среды. Разрядка полей напряжений способна спровоцировать техногенные катастрофы и, таким образом, влияет на безопасность людей и сооружений.

В течение 2007-2018 гг. на геодинамических полигонах в Алтае-Саянской сейсмоактивной области (АССО) Центральной Сибири, на п-ове Камчатка и о. Сахалине проводился мониторинг НДС геологической среды комплексом геолого-геофизических методов – сейсмический мониторинг, регистрация естественного импульсного электромагнитного поля Земли (ЕИЭМПЗ), газгидрогеохимический мониторинг на основе эмиссии радона. Научно-методические основы и технологические решения, используемые при мониторинге НДС геологической среды на геодинамических полигонах, прошли многолетнюю проверку. Ниже представлена оценка возможностей комплекса геолого-геофизических и газгидрогеохимических методов для прогноза напряжённо-деформированного состояния геологической среды по результатам геодинамического мониторинга на Камчатке, Сахалине и в Алтае-Саянской области в 2007-2018 гг.

**Сейсмический мониторинг.** Сейсмический мониторинг является основой для прогноза НДС геологической среды. Рассмотрим результаты использования сейсмологического мониторинга на геодинамических полигонах в Сибири и на Дальнем Востоке, приведем рекомендации по его совершенствованию. Напряжённо-деформированное состояние геологической среды определяет состояние сейсмической активности территории. Наибольшее влияние на сейсмический процесс оказывают эндогенные факторы – такие как блочные движения литосферных плит (от миллиметров до десятков сантиметров в год); термодинамические процессы в ядре и мантии Земли, включая развитие мантийных плюмов. Значительное влияние на сейсмические процессы Камчатки оказывает вулканизм. Геофизический мониторинг на Петропавловск-Шипунском, Южно-Сахалинском и Алтае-Саянском геодинамических полигонах МПР РФ в 2007-2018 гг. позволил выявить и оценить влияние гравитационных лунно-солнечных приливов и их резонансов в земной коре на циклические изменения НДС в пределах различных блоков земной коры.

По результатам многолетнего геодинамического мониторинга, экспериментально выявлены резонансы гравитационных приливов в земной коре [4, 5]. Ниже представлены условия и примеры проявлений резонансов гравитационных приливов в Земной коре. В 2009 г. была обоснована необходимость учета колебаний барицентра при оценке приливных воздействий на геологическую среду [6]. Для оценки координат барицентра использовался программный комплекс Калифорнийского технологического института (США) для расчета эфемерид. Для того, чтобы увязать внешние геодинамические факторы в единую систему отсчета, была использована нормировка значений этих параметров в интервале от нуля до единицы. При этом за единицу принято максимальное значение параметра в течение года, а за 0 – минимальное годовое значение. В результате получены графики изменения во времени нормированных к единице параметров изучаемых факторов (расстояние Земля-Луна, фазы Луны, расстояние барицентра от поверхности Земли и от наблюдателя). При этом максимальное влияние на НДС геологической среды Земного шара оказывают колебания барицентра в интервале ±28-30° относительно плоскости эклиптики. Приливные волны способны резонировать между собой при известных условиях близости частоты и фазы. Длиннопериодные (14-15 суток) гравитационные приливы удовлетворяют этим условиям, поэтому графики указанных факторов образуют точки пересечения, которые мы рассматриваем в качестве индикатора резонансов гравитационных приливов.

Обоснованность этого предположения подтверждена многолетними экспериментальными данными. В частности, авторами изучена корреляция между резонансами гравитационных приливов, сейсмической, электромагнитной эмиссией и эмиссией радона в подземных водах. Показано, что структура ЕИЭМПЗ, эмиссии радона, структура сейсмической эмиссии чётко коррелируются с пересечением графиков исследуемых факторов, т.е. с резонансами гравитационных приливов. Во всех регионах Сибири и Дальнего Востока – эта закономерность

проявления резонансов приливов в структуре геофизических данных при мониторинге устойчиво фиксируется. Сопоставление времени катастрофических (M≥6.0) землетрясений на земном шаре за XX и XXI век с расчетными резонансами гравитационных приливов для каждого сейсмического события, позволяет утверждать, что более 80% катастрофических землетрясений по времени с точностью ±12 часов совпадают с расчетными резонансами. Тесная корреляция (более 80%) между временем расчетных резонансов и срабатыванием сейсмических очагов зафиксирована в Алтае-Саянском сейсмоактивном регионе, на Камчатке и Сахалине для землетрясений с M≥5.0. Таким образом, резонансы гравитационных приливов являются триггерами подавляющего большинства сильных и катастрофических землетрясений на Земном шаре.

Не всякий резонанс (а их в течение лунно-солнечного месяца в любой точке земного шара происходит не менее 10) вызовет землетрясение в конкретном очаге. Для этого необходимо выполнение дополнительных условий. Собственные колебания конкретного сейсмического очага должны быть близки по частоте и фазе вынужденным колебанием геологической среды, возникающим в результате прохождения через среду приливных волн деформации. Накопленная (потенциальная) энергия в сейсмическом очаге достигла или превышает предел механической (энергетической) прочности горных пород, слагающих сейсмический очаг. При выполнении этих условий, резонанс гравитационных приливов способен вызвать землетрясение, т.е. разрядку накопленных напряжений.

Различными исследователями экспериментально зафиксированы медленные деформационные волны в геологической среде со скоростями порядка 0,1-10,0 м/сек, в том числе и при геодинамическом мониторинге в Сибири и на Дальнем Востоке. В случае если возникает одиночная деформационная волна типа солитона, возможны негативные последствия. Анализ возникновения деформационной волны при резонансе гравитационных приливов в августе 2009 г. в Алтае-Саянском регионе позволяет рассматривать в качестве причины известной аварии на Саяно-Шушенской ГЭС 17.08.2009 г. не только нарушение техники безопасности со стороны персонала станции, но и негативное влияние низкоростной (1,5 км/час) деформационной волны, образовавшейся за 5-6 суток до аварии на ГЭС. Таким образом, резонансы гравитационных приливов в Земной коре не только основной триггер для разрядки напряжений в сейсмических очагах, но и источник возникновения низкоскоростных (см/сек, м/сек) деформационных волн в геологической среде.

Существующие системы геодинамического мониторинга используются в настоящее время в основном для определения параметров землетрясений (координаты, время, магнитуда) на основе регистрации быстрых (5-7 км/сек) сейсмических волн. В тоже время, сети сейсмического мониторинга, при дополнительной спектральной обработке исходных данных, способны обеспечить также мониторинг медленных деформационных волн (1.0÷5.0 км/час). Это позволит заблаговременно предупреждать системы контроля техногенных объектов с близкими к волнам деформации резонансными свойствами о геодинамической опасности.

Отсюда следуют практически важные выводы.

1. Существующие системы геодинамического мониторинга (сеть сейсмических станций) помимо регистрации сейсмических событий могут быть использованы для оценки опасных резонансов и выявления генерируемых ими низкоскоростных (1,5-2 км/ч) деформационных волн. Деформационные волны постоянно регистрируются при проведении ГГД мониторинга. Для этого, помимо использования записей сейсмостанций для определения параметров землетрясений, необходимо проводить спектральный анализ шумов исходных записей. Это позволит выявить образование медленных деформационных волн и определить их основные моды. В связи с низкими скоростями распространения волн деформаций, всегда есть время (от часов до суток), чтобы принять меры по исключению резонансных явлений на контролируемых технических системах
2. Для обеспечения геодинамической безопасности особо ответственных объектов необходимо дополнить существующие сети мониторинга системой контроля резонансов грави-

тационных приливов. Должны быть созданы региональные системы прогноза геодинамической опасности, что позволит своевременно принимать меры по защите особо опасных объектов от влияния деформационных волн в геологической среде.

Сейсмический мониторинг обеспечивает среднесрочный (1-2 месяца) прогноз изменений НДС геологической среды. При прогнозе конкретных сейсмических событий авторами использована «энтропийная» модель развития энергетических процессов в сейсмическом очаге на стадии подготовки землетрясения [4]. Опыт применения «энтропийной модели» позволяет обеспечить прогноз сейсмических событий (M≥5,0 с ошибками по времени Δt ±3-15 суток, по магнитуде ±0,5. Что касается прогноза положения эпицентра, то выявлена следующая особенность – чем точнее прогнозируется магнитуда (ΔM±0,5) и время (Δt ±1-3 суток) сейсмического события, тем ненадёжнее прогнозируется положение эпицентра (ΔS±100-500 км). При использовании только сейсмических каталогов, в принципе невозможно достоверно определить (спрогнозировать) сразу три параметра – магнитуда, время и место (положение эпицентра) землетрясения.

Практика прогнозирования землетрясений на Камчатке, Сахалине и в Алтае-Саянском регионе в 2007-2018 гг. позволяет утверждать, что краткосрочный (с точностью ±1-3 суток, M±0,5) прогноз возможен, но при этом ошибка положения эпицентра возрастает до ±200-500 км. И, наоборот, при удачном прогнозе положения эпицентра (с точностью ±50-100 км) ошибка по времени может достигать ±10-15 суток, а по магнитуде ±1,0-1,5. Применительно к практике прогнозирования сейсмических событий, необходимо помимо сейсмических каталогов, использовать комплекс различных методов (имеющих отличную от сейсмического волнового процесса физическую природу). В частности, для прогноза положения эпицентра землетрясения, целесообразно использовать инфракрасное, инфразвуковое излучения, эмиссию газов в зонах сейсмических очагов на базе дистанционного зондирования Земли.

**Электромагнитный мониторинг.** Рассмотрим основные требования к электромагнитному мониторингу на основе регистрации естественного импульсного электромагнитного поля Земли (ЕИЭМПЗ). Природа ЕИЭМПЗ до сих пор вызывает дискуссии. Основатели метода (Воробьёв и другие – Томский политехнический институт, 70-80 гг. XX века), исследуя ЕИЭМПЗ в частотном диапазоне 10-20 кГц, пришли к выводу, что в спектре сигнала ЕИЭМПЗ присутствуют две основные составляющие: ионосферная (связанная с грозовой активностью и другими атмосферно-ионосферными процессами) и литосферная (имеющая механо-электромагнитную природу) [7, 8].

Результаты геодинамического мониторинга (сейсмического, электромагнитного, газгидрогеохимического) в 2007-2018 гг. на полигонах Петропавловск-Шипунского, Южно-Сахалинского, Алтае-Саянском (Красноярском и Тывинском) полигонах позволили выявить основной механизм формирования суточного хода ЕИЭМПЗ – волны сжатия-растяжения, связанные с гравитационными приливами в земной коре и литосфере. Форма и амплитуда суточных ходов ЕИЭМПЗ устойчиво (K>0,8) коррелируется с резонансами гравитационных приливов (суточных, месячных и годовых). Эффекты изменения структуры ЕИЭМПЗ на стадии подготовки землетрясений проявляются как в краткосрочном (1-10 суток), так и среднесрочном (до двух месяцев) плане нашли подтверждение в результатах геодинамического мониторинга в 2007-2018 гг. на Камчатке, Сахалине, АССО.

Следует учитывать, что в различных пунктах регистрации (в зависимости от особенностей геолого-тектонического строения и геодинамических режимов различных блоков земной коры) может наблюдаться различное по форме проявление краткосрочных признаков подготовки землетрясения в структуре суточных ходов ЕИЭМПЗ (вплоть до противофазных). Так при землетрясении 10 февраля 2011 года магнитудой 5,8 в Ермаковском районе на юге Красноярского края, ближайшая станция ЕИЭМПЗ «Ергаки» (100 км от эпицентра) за двое суток до землетрясения зафиксировала снижение уровня ЕИЭМПЗ, тогда как станция ЕИЭМПЗ «Нарва» (Заповедник «Столбы») на расстоянии около 400 км от эпицентра землетрясения зафиксировала чёткий максимум ЕИЭМПЗ в течении двух суток до события.

Летом признаки подготовки землетрясений проявляются на фоне повышенного в целом уровне интенсивности ЕИЭМПЗ и поэтому их сложнее идентифицировать в качестве признаки подготовки землетрясения. Аномальная структура ЕИЭМПЗ в АССО коррелируется с характером изменения концентрации радона в различных точках (Красноярск–Туманный–Аскиз–Кызыл). Регистрация радона в комплексе с геофизическими методами доказала свою эффективность в АССО для оценки НДС геологической среды. Подготовка слабых землетрясений (М около 3,0) также проявляется в структуре суточного хода ЕИЭМПЗ.

На структуру ЕИЭМПЗ оказывает существенное влияние подготовка сильных катастрофических землетрясений (М≥8) даже на значительном расстоянии от пункта регистрации (до 1000-3000 км). Выполнен анализ записей ЕИЭМПЗ на Сахалине при подготовке Японского катастрофического землетрясения (М=9,0, 11 марта 2011 года). Структура ЕИЭМПЗ на Сахалине (1000 км от эпицентра) изменилась за 6 месяцев до японской катастрофы. На Камчатке (3000 км от эпицентра Японского землетрясения) структура ЕИЭМПЗ за 3 месяца до катастрофы также явно изменилась. Аналогичная картина в структуре ЕИЭМПЗ на полигоне Камчатки зарегистрирована в мае 2013 года в ходе подготовки к известному Охотскому глубинному (глубина гипоцентра ≈ 600 км) землетрясению, которое проявилось макроэффектами на всей территории Азии (от Камчатки до Индии, Сибири, Урала).

Изменение структуры ЕИЭМПЗ при подготовке сильных землетрясений можно использовать для среднесрочного (несколько месяцев) прогноза катастрофических землетрясений в Курило-Камчатском регионе. Для этого необходимо установить аппаратуру МГР-01 не только на локальных геодинамических полигонах, но и вдоль Курило-Камчатского сейсмоактивного пояса от Алеутских островов до Японии, включая Курильские острова с интервалом между станциями около 500 км. Таким образом, результаты геодинамического мониторинга в 2007-2018 гг. на Камчатке, Сахалине, в Алтае-Саянском регионе подтвердили информативность ЕИЭМПЗ для прогноза изменений НДС геологической среды и в ряде случаев для прогноза землетрясений. Следует признать, что использование метода регистрации ЕИЭМПЗ без комплекса с другими (выше упомянутыми) методами, не позволяет однозначно давать оценку изменения НДС геологической среды и прогноз землетрясений.

Теория и аппаратурно-технологическая база метода ЕИЭМПЗ нуждается в дополнительном развитии. В частности, экспериментально установлен факт значительной глубинности метода ЕИЭМПЗ (десятки км), в то время как механо-электромагнитная модель (наиболее распространённое в настоящее время представление о природе ЕИЭМПЗ) не допускает излучение электромагнитных импульсов частотой 10-20 Гц с глубины более 1 км. Необходимо увеличить динамический диапазон регистрации ЕИЭМПЗ с 96 Дб минимум до 128 Дб. Разрядность счётчика импульсов необходимо увеличить с 16 разрядов до 24 разрядов минимум и расширить частотный диапазон регистрации в интервале 1-25 кГц.

При использовании имеющейся аппаратуры МГР-01 целесообразно на каждом пункте вести одновременную регистрацию ЕИЭМПЗ на одной «опорной» частоте (около 14 кГц) с различными коэффициентами усиления. Это позволит более надёжно идентифицировать сигналы литосферного происхождения на фоне влияния других видов источников (ионосферные, атмосферные и температурные шумы).

**Мониторинг радона.** Мониторинг радона в 2007-2018 гг. проводился в Алтае-Саянском регионе, где подтвердил свою эффективность в комплексе с другими геофизическими методами контроля НДС геологической среды. Характер изменения концентрации радона в пределах Алтае-Саянского геодинамического полигона связан также с длиннопериодными (3-6 месяцев) циклами гравитационных приливов. Радоновые предвестники подготовки сильного землетрясения проявляются в положительных или отрицательных аномалиях уровня радона в подземных водах [3]. Мониторинг содержания радона в подземных водах выполняется с целью выявления средне- краткосрочных (месяц, 2-10 дней) предвестников сильных землетрясений М>5.0-6.0 и оценки изменений напряженно-деформированного состояния геологической среды на основе анализа связи динамики колебаний концентрации радона в связи с сейсмической активностью региона.

Выполнен анализ и продемонстрированы радоновые предвестники сильных землетрясений на территории Республики Тыва в 2011 и 2012 гг. Получены оценки предвестника подготовки землетрясения 27 декабря 2011 года магнитудой 6,5. Эмиссия радона изменилась за 5-10 суток перед землетрясением во всех четырёх пунктах наблюдения Алтае-Саянского региона. Наиболее чётко предвестник проявился на п.н. Кызыл в 100 км от эпицентра землетрясения. За 4 суток до основного события в уровне концентрации радона наблюдалась аномалия, выраженная резким повышением уровня концентрации радона. На основе статьи [6] был выполнен численный анализ изменения концентрации радона в подземных водах. Оценки определяются тремя способами: по величине абсолютного максимального отклонения – положительного или отрицательного – от средне фоновой концентрации; по величине относительного максимального отклонения; по числу среднеквадратичных отклонений для фоновых концентраций.

На основе данного подхода выполнен анализ предвестников подготовки землетрясений в Каа-Хемском районе республики Тыва 27 декабря 2011 года магнитудой 6,5 и 26 февраля 2012 года магнитудой 6,7. Во время подготовки первого землетрясения средне фоновая концентрация радона была на уровне 69,6, тогда как среднее значение концентрации радона в аномальный период составила 88,2. Время предвестниковой радоновой аномалии оценивалось от начала устойчивого повышения концентрации радона величины $2*S_{фон}$ (среднеквадратичное отклонение фоновых концентраций) до момента землетрясения и составило 5 суток.

Во время подготовки второго землетрясения средне фоновая концентрация радона была на уровне 69,6, тогда как среднее значение концентрации радона в аномальный период составила 86,0. Время предвестниковой радоновой аномалии оценивалось от начала устойчивого повышения концентрации радона величины $2*S_{фон}$ (среднеквадратичное отклонение фоновых концентраций) до момента землетрясения и составило 3 суток. Данные расчёты проводились по данным пункта измерения «Кызыл», расстояние до очагов землетрясений составляет около 100 км.

Изменения концентрации радона в естественных водных источниках позволяют в комплексе с данными ЕИЭМПЗ, сейсмической эмиссии прогнозировать не только изменение НДС геологической среды, но также обеспечить среднесрочный (1-3 месяца) прогноз сильных сейсмических событий $M \geq 6,0$. Краткосрочный (1-3 суток) прогноз тоже возможен, но без комплексирования с другими методами (сейсмическим, ЕИЭМПЗ) достоверность интерпретации локальных аномалий радона при редкой сети наблюдений не позволяет уверенно прогнозировать даже сильные землетрясения ($M \geq 5,0$). Для решения задачи краткосрочного прогноза землетрясений с использованием данных мониторинга радона, необходимо увеличить сеть наблюдений, доведя число пунктов наблюдения уровня радона минимум до числа пунктов наблюдения ЕИЭМПЗ в пределах геодинамического полигона.

**Заключение.** Геодинамический мониторинг комплексом геолого-геофизических методов (сейсмология, ЕИЭМПЗ, радон) обеспечивает не только оценку изменения НДС геологической среды, но также среднесрочный (1-3 месяца) и краткосрочный (1-10 суток) прогноз сильных землетрясений с $M \geq 5,0$. Достигнутая точность прогнозов составляет по магнитуде $\Delta M \pm 0,5$; по времени $\Delta t \pm 1$-15 суток; по местоположению эпицентра $\Delta S$ от 100 до 500 км. Низкая точность определения положения эпицентров сейсмических событий связана с принципиальной невозможностью равноточного определения сопряжённых величин: магнитуды и положения эпицентра.

С другой стороны, уровень комплексирования (низкая плотность сетей регистрации ЕИЭМПЗ и измерения уровня концентрации радона в подземных водах) не обеспечивает надёжный прогноз положения эпицентра ($\Delta S \pm 50$-100 км). Для повышения надёжности определения положения эпицентров землетрясений, необходимо увеличить плотность сетей регистрации ЕИЭМПЗ и радона и дополнить применяемый геолого-геофизический комплекс данными спутниковых инфракрасной и геохимической ($CO_2$, метан) съёмок.

В связи с распространением в геологической среде наряду с высокоскоростными волнами сейсмического диапазона (скорость 5-7 км/сек), медленных волн деформации

(1-2 км/час и менее), необходимо проектировать геодинамические полигоны для мониторинга и медленных волн деформации. Целесообразно при организации геодинамических полигонов наряду с существующими сейсмическими сетями с большой апертурой (сотни км), использовать высокочувствительные малоапертурные (10-20 км) локальных группы.

## ЛИТЕРАТУРА


[1] *Вартанян Г.С., Гончаров В.С., Кривошеев В.П., Потемка Э.П., Стажило-Алексеев С.К.* Методические указания по ведению гидрогеодеформационного мониторинга для целей сейсмопрогноза (система RE-STEPS). М.: ВСЕГИНГЕО, 2000. 83 с.

[2] *Гордеев В.Ф., Малышков Ю.П., Шталин С.Г., Малышков С.Ю., Поливач В.И.* Оценка напряженно-деформированного состояния горного массива по параметрам ЕИЭМПЗ // Интерэкспо ГЕО-Сибирь. 2009. Т. 1, Ч. 2. С. 71-75.

[3] *Зубков С.И.* Радоновые предвестники землетрясений // Вулканология и сейсмология. 1981. № 6. С. 74-105.

[4] *Сибгатулин В.Г., Хлебопрос Р.Г., Перетокин С.А., Кабанов А.А.* Экологическая и инженерная геология: синергия процессов в сейсмических очагах и краткосрочный прогноз землетрясений // Инженерная экология. 2009. № 2. С. 32-42.

[5] *Sibgatulin V.G., Peretokin S.A., Kabanov A.A.* Resonances of gravitational tides and their effect on geological environment // Earth science frontiers. 2014. Vol. 21, N 4. P 303-310.

[6] *Сибгатулин В.Г., Перетокин С.А., Симонов К.В., Кабанов А.А.* Информационное и алгоритмическое обеспечение системы наблюдений флюидных геообъектов // Материалы XIII Российская конференция с участием иностранных ученых «Распределенные информационные и вычислительные ресурсы». Новосибирск: ИВТ СО РАН, 2010. С. 78-82.

[7] *Малышков Ю.П., Малышков С.Ю., Гордеев В.Ф.* Взаимосвязь импульсных электромагнитных полей Земли с движением тектонических блоков и землетрясениями. Томск, 2000. 38 с.

[8] *Малышков Ю.П., Малышков С.Ю.* Эксцентричное вращение ядра и оболочки Земли, возникновение деформационных волн, их практическое использование. Томск, 2010. 118 с.

# ПРОСТРАНСТВЕННО-ВРЕМЕННЫЕ ГРАДИЕНТЫ ПАРМЕТРОВ СОСТОЯНИЯ ПОВЕРХНОСТИ ЗЕМЛИ ДЛЯ СПУТНИКОВОГО МОНИТОРИНГА ПРИРОДНЫХ СИСТЕМ

*Картушинский А.В.*[(1)(2)], *Огородов Н.А.*[(3)], *Ларько А.А.*[(1)]
[(1)] Институт биофизики СО РАН, Красноярск
[(2)] Сибирский федеральный университет, Красноярск
[(3)] Федеральный исследовательский центр КНЦ СО РАН, Красноярск

Представлены результаты обработки спутниковых данных по расчету градиентных характеристик поверхности Земли. Рассматриваются структурные особенности динамики горизонтальных градиентных полей водных систем и поверхности суши по физическим и биологическим параметрам. Обсуждаются проблемы параметризации взаимодействия факторов окружающей среды для численного моделирования процессов физического и биологического уровня по данным дистанционного зондирования Земли. Обсуждаются пространственно-временное масштабирование и усреднение градиентных компонентов. Разработаны подходы для оценки поверхностных градиентов водной и земной поверхности в различных экосистемах.
*Ключевые слова: спутниковые данные, региональные особенности, градиентные поля, фронтальные зоны, экотоны, усреднение пространственное, среднесезонная изменчивость, синергетический эффект.*

**Введение**. Градиентные показатели измеряемых поверхностных параметров наземных и водных экосистем отражают некоторые пространственные границы динамических процессов различного свойства. Расчеты градиентных характеристик природных объектов необходимы для выявления зон с различной динамической активностью [1, 2]. Спутниковые измерения в видимом и инфракрасном диапазонах представляют собой изображения с пространственными полями различных природных объектов поверхности Земли. Основными параметрами в нашем случае являются отражательная способность поверхности и радиационная температура. Методологическая концепция, определена необходимостью расчета пространственных градиентов измеренных характеристик поверхности Земли. Рассчитанные градиенты с различным пространственным и временным усреднением необходимы для выявления зон с различной динамической активностью физических и биологических полей, растительности суши, а также геологических показателей, таких как рельеф или почвенный покров.

Градиентные характеристики параметров, которые используются для изучения поверхностной динамики пространственных неоднородностей экосистем представляют собой дифференцированные показатели скорости изменения величин. Величина градиента по пространственным координатам предоставляет возможность получить вектор развития или затухания процесса. Градиентные показатели пространственной изменчивости исследуемого объекта (поля), применяются зачастую не в полной мере, особенно для квазистатичных объектов, таких как экосистемы суши. При этом расчеты динамики исследуемых полей основываются на совмещении или композиции исходных изображений (снимков), сравниваемых в абсолютных количественных показателях, усредненных и интерполированных значениях [3, 4].

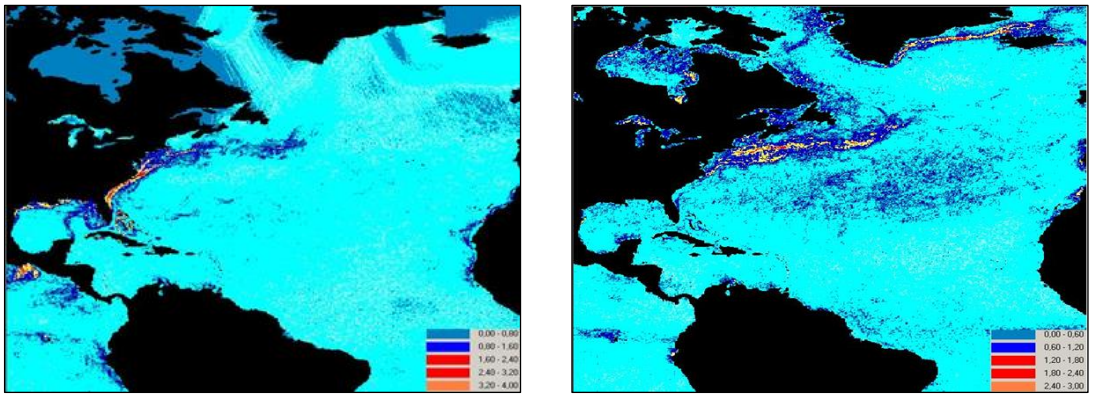
Определение условий развития и масштабов гидрологических процессов, необходимо для мониторинга экологического состояния водных систем включая океаны, моря и водные объекты суши. Расчет пространственных градиентов по спутниковым данным позволяет проводить оценку пространственных и временных масштабов формирования зон с высокими значениями градиентов. Возможные условия действующих факторов и физические механизмы взаимодействия гидробиологических и гидрологических компонент могут исследоваться с использованием методов численного моделирования на основе расчета частных производных уравнений переноса тепла, массы и примеси по спутниковым и контактным данным [5, 6, 7].

Анализ изменчивости пространственно-временных градиентов параметров состояния поверхности Земли необходим для оценки динамического проявления процессов как в водных

системах, так и на твердой земной поверхности. Таким образом, основной целью данной работы является реализация метода расчета градиентных характеристик земной поверхности по спутниковым данным с использованием разработанных информационно-программных средств и соответствующего программного обеспечения с применением методов математической статистики.

**Материалы и методы**. Спутниковые данные, полученные различными сканерами SeaWiFS, CZCS, AVHRR, MODIS, SPOT, LANDSAT, AQUARIUS, использовались для расчетов градиентных характеристик в различных районах Мирового океана и территории Евразии. Основное внимание уделяется выявлению закономерностей формирования горизонтальных высокоградиентных зон по физическим и биологическим компонентам поверхностного слоя Земли. Для этого используется градиентный метод обработки спутниковых данных при помощи соответствующего информационно-программного комплекса (ИПК) [1]. Применяется методология изучения изменчивости поверхностных градиентных полей водных и наземных экосистем, как динамических объектов на основе расчета пространственных градиентов (широтных, меридиональных и абсолютных) с различным периодом усреднения (неделя, месяц, сезон, год). В качестве исходных данных в работе используются спутниковые изображения и данные спетрорадиометров, полученные аппаратными средствами с высоким и низким разрешением AVHRR MCSST, CZCS, SeaWIFS, MIRAS AQUARIUS, MODIS, MSS, за различные периоды времени со спутников NOAA, TERRA, AQUA, SPOT, LANDSAT.

Для расчетов градиентных полей растительности суши использовался программный продукт ENVI, а для написания логики вычислений использовался язык программирования IDL.



*а* *б*

Рис. 1. Структура температурной фронтальной зоны системы взаимодействия теплого течения Гольфстрим и холодного Лабрадорского течения по расчету абсолютных (по модулю) градиентов поверхностной температуры океана для 2000 года; *а* – зима; *б* – лето.

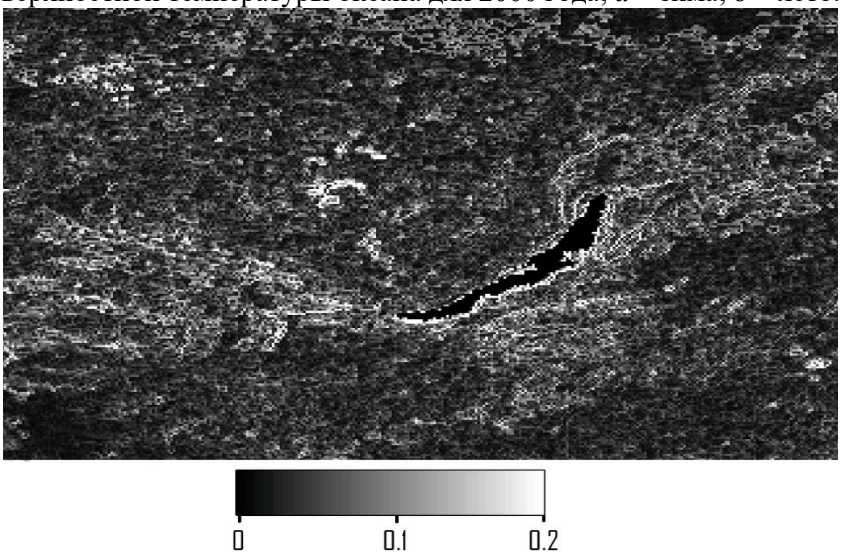



Рис. 2. Градиентное поле NDVI рассчитанное по модулю (абсолютные градиенты).

**Результаты и обсуждение.** Расчет пространственных градиентов характеристик земной поверхности в широтном и меридиональном направлении позволяет выявлять зоны со значительными различиями, которые могут трактоваться как граничные или аномальные. Градиентные поля гидрологических параметров в водных системах позволяют оценивать интенсивность изменчивости динамических процессов. Взаимодействие различных водных масс с отличающимися свойствами формирует гидрологические фронты, которые определяются именно по расчетам соответствующих градиентов (рис. 1). В данном случае Мировой океан представляет собой динамическую систему, в которой формируются гидрофизические и гидробиологические структуры, образованные процессами различных пространственных и временных масштабов.

В настоящее время значительное внимание уделяется обработке данных наблюдений и их реанализ с применением моделей общей циркуляции океана для совершенствования глобальных моделей взаимодействия в системе атмосфера-океан-суша [9-11]. Важно понять спусковой механизм для формирования таких явлений, как Североатлантические осцилляции (САО), Эль-Ниньо-Южные колебания (ЭНЮК). Анализ динамической активности высокоградиентных зон в океане и определение эффектов локального взаимодействия температурного поля океана и формирование соответствующего отклика атмосферных циркуляций (циклоны, антициклоны) на основе градиентных показателей, даст возможность понять пространственно-временные масштабы локальных процессов, определяющих синергизм глобальных (планетарных) явлений.

Применение в математических моделях, рассчитанных по спутниковым данным пространственных градиентных показателей, дает величину скорости изменения структуры поверхностного слоя. Анализ размерностей членов уравнений переноса тепла, массы и диффузии примеси, основанный на теории подобия, позволяет, при проведении численного моделирования, обоснованно вводить в математические модели начальные, граничные условия и сеточный масштаб расчета. Численное моделирование, основанное на 2D-модели переноса тепла, позволяет изучать изменчивость структуры температурных фронтальных зон в океане в зависимости от динамических условий переноса тепла, массы и концентрации примеси.

Что касается поверхности суши и наземных экосистем, важно определять пространственные масштабы экотонов (переходных границ между экосистемами) и временные масштабы изменчивости таких зон. Процессы с низкой скоростью динамической активности в пространстве, такие как эрозия почвы, растительный покров, морфология поверхности Земли, по градиентным полям практически не оцениваются [1]. Однако эти процессы оказываются значительно связанными на долговременных рядах наблюдений.

Так при оценке биомассы, валовой и чистой первичной продукции (ЧПП) экосистем суши при использовании спутниковых данных, иногда сложно определить площадь отдельных экосистем и их пространственные границы (экотоны). Кроме того, во многих работах не рассматривается пространственное распределение динамики ЧПП или другого показателя, а оцениваются тренды, усредненные по площади [8]. Полученные усредненные результаты, при дальнейшей верификации их наземными данными, зачастую очень неоднозначны. Таким образом, можно говорить о неприменимости линейных моделей для анализа временных рядов в сложных, пространственно-распределенных системах. Поэтому, возникает вопрос о проверке и анализе применимости различных методов получения временных трендов, подробное рассмотрение их особенностей, сильных и слабых сторон, применительно к данным дистанционного зондирования Земли, входящих в модели оценки ЧПП.

Расчет градиентных полей показателей растительности суши, в частности градиенты NDVI, могут помочь в определении пространственных границ экосистем, которые во многом соответствуют рельефу местности. Скорость и направление изменений почвенного и растительного покрова могут иметь разную когерентность, в том числе и как отклик на климатические изменения.

Пример градиентного поля вегетационного индекса, рассчитанного на основе использования космических данных, GIMMS (Global Inventory Modeling and Mapping Studies) NDVI с

1982 по 2006 гг., полумесячные композитные изображения, разрешение 8-км (на линии экватора), масштаб глобальный, спутник NOAA сканер AVHRR, представлен для Байкальского региона на рис. 2.

Таким образом результатом применения градиентного подхода возможно выделение крупных изменений граничного перехода между биомами и экосистемами, такими как переход тайга – тундра, или зон высокогорных регионов и низменностей с соответствующими типами растительности.

**Выводы.** Расчеты градиентных характеристик поверхности экосистем на основе спутниковых данных позволяют выявлять зоны с различной динамической активностью. Анализ распределения величины градиентных полей дает возможность выявлять гетерогенные и гомогенные зоны экологических систем, а при соответствующем периоде усреднения получать степень динамичности таких зон. Это в свою очередь позволяет изучать процессы структурной организации природных систем и пространственно-временные масштабы изменчивости, что может применяться при разработке математических прогностических моделей, как статистических, так и детерминированных.

Для водных систем, изучение структуры градиентных полей температуры воды, солености, концентрации хлорофилла, мутности, дает возможность выявлять масштабы зон неоднородного распределения характеристик, определять время существования граничных (переходных) зон по величине градиента в меридиональном, широтном направлении, а также по модулю. Такая методология в совокупности с применением соответствующих информационно-программных средств, позволяет оценивать действие факторов физической и биологической природы на формирование экосистемы в целом, либо оценивать цикличность действия этих факторов в отдельных частях системы. Для экосистем суши степень изменчивости поверхностных неоднородностей значительно ниже, чем в водных системах, однако градиентных подход может предоставлять дифференцированную картину границ экосистем на основе спутниковых данных. Таким образом можно оценивать синергетическую составляющую при самоорганизации системного уровня взаимодействия экосистем и элементов системы Биосфера на уровне атмосфера-океан-суша с соответствующими масштабами пространственно-временного усреднения.

## ЛИТЕРАТУРА


[1] *Картушинский А.В., Кукоба Н.А.* Изучение градиентных полей поверхности Земли по спутниковым данным // Вестник СибГАУ. 2015. Т.3 (16). С. 587-596.

[2] *Kartushinsky A.V., Sidorenko A.Y.* Analysis of the variability of temperature gradient in the ocean frontal zones based on satellite data // Advanced Space Research. 2013. Vol. 52, N 8. P. 1467-1475.

[3] *Коберниченко В. Г., Тренихин В. А.* Методы синтеза изображений на основе данных дистанционного зондирования Земли различного разрешения // Успехи современной радиоэлектроники. 2007. № 4. С. 22-31.

[4] *Асмус В.В., Бучнев А.А., Кровотынцев В.А. и др.* Planeta Monitoring Программный комплекс обработки спутниковых данных // Проблемы информатики. 2013. № 3. С. 85-99.

[5] *Kartushinsky A.V.* The investigation on the dynamics of frontal zones in the ocean based on the numerical modelling, using the AVHRR satellite data // Advances in Space Research. 2004. Vol. 33 (7). P. 1173-1178.

[6] *Kartushinsky, A.V.* Dynamic Features of Frontal Zones Structure in the Ocean for Using in the Numerical Models Based on Satellite Data // Journal of Siberian Federal University. Mathematics & Physics. 2011. Vol. 4 (2). P. 208-216.

[7] *Картушинский А.В.* Численное моделирование эффектов гидрофизического воздействия на распределение фитопланктона / Математическая биология и биоинформатика. 2012. Т. 7, № 1. С. 112-124.

[8] *Ларько А.А., Иванова Ю.Д., Шевырногов А.П.* Нелинейные тренды чистой первичной продукции растительности юга Красноярского края // Фундаментальные исследования. 2015. № 3. С. 106-110.

[9] *Казьмин А.С.* Изменчивость крупномасштабных океанических фронтальных зон: анализ глобальной спутниковой информации // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 9, № 1. С. 213-218.

[10] *Серых И.В.* О динамике и структуре глобальной атмосферной осцилляции в климатических моделях и реальности // Океанологические исследования. 2018. Т. 46, № 1. С. 14-28.

[11] *Rayner N.A., Parker D.E., Horton E.B.* Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century // JGR: Atmospheres. 2003. Vol. 108, N 14. P. 4407.

# ПРОГНОЗИРОВАНИЕ РАСПРОСТРАНЕНИЯ БОЛЕЗНЕЙ В ТРОПИЧЕСКИХ ЗОНАХ С ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

*Колесников А.А.[(1)], Кикин П.М.[(2)]*
[(1)] Сибирский государственный университет геосистем и технологий, Новосибирск
[(2)] Санкт-Петербургский Политехнический Университет Петра Великого, Санкт-Петербург

Заражение тропическими паразитарными болезнями, по данным ВОЗ, оказывают огромное влияние на здоровье более 40 миллионов человек во всем мире и являются второй по значимости причиной иммунодефицита. На количество заражений влияют многие факторы - климатические, демографические, растительный покров и ряд других. В статье представлено исследование и оценка степени влияния каждого из этих факторов, а также сравнение качества прогнозирования отдельными методами геоинформационного анализа и машинного обучения и возможности их ансамблирования.
*Ключевые слова: тропические болезни, лихорадка, машинное обучение, геостатистика, нейронные сети.*

**Введение.** Многие важные научные задачи связанны с обработкой данных, которые изменяются в пространстве и времени. В качестве примеров можно привести климатические, экологические, сейсмические исследования [1-3]. Корректно подобранные методы и алгоритмы моделирования и прогнозирования позволяют достоверно оценить тенденцию развития исследуемых показателей произвольных объектов и явлений. Автоматизировать процесс выбора алгоритмов, формул и их параметров способны технологии машинного обучения [4-6]. Формирование путей взаимодействия между технологиями геоинформационных систем и машинного обучения для пространственно-временного прогнозирования позволит разработать наиболее оптимальные решения для эффективного анализа и управления процессами и явлениями. Важно иметь информацию о методах и алгоритмах, которые хорошо работают на практике для обработки и прогнозирования данных максимально широкого диапазона пространственно-временных процессов [7, 8]. Значительный прирост в качестве построения математических моделей пространственно-временных процессов дало развитие концепции Deep Learning, в частности, различных вариантов рекуррентных нейронных сетей, например, LSTM. Также широко распространенные сверточные нейронные сети (CNN) используют слои со сверточными фильтрами для извлечения локальных объектов посредством скользящего окна и может моделировать близлежащие или долгосрочные пространственные зависимости (архитектура SRCN) [9]. Для оценки применимости, точности получаемых математических моделей, универсальности методов машинного обучения для решения задач геоинформатики была выбрана задача прогнозирования распространения тропической болезни денге по данным ДЗЗ, представленных в виде временных рядов [10]. Денге – это инфекционное заболевание, передающееся от человека к человеку посредством москитов Aedes aegypti и Aedes albopictus, которые является основным переносчиком вируса в различных частях земного шара. По оценкам Всемирной организации здравоохранения (ВОЗ), ежегодно во всем мире регистрируется около 50-100 миллионов случаев заболевания лихорадкой денге, а две пятых населения мира подвержены риску, а денге или DHF / DSS затронули более ста стран. С 1950 года было зарегистрировано более 500 000 случаев госпитализации и около 70 000 случаев смерти детей; уровень заболеваемости среди детей достигает 64 на 1000 человек населения. Согласно анализу глобального распространения вируса денге, число инфекций в год оценивается в 390 миллионов, из которых почти 96 миллионов являются симптоматическими. По оценкам, число инфекций денге резко возросло за последние 50 лет, что привело к огромному воздействию на здоровье человека во всем мире. Регионы распространения включают страны Юго-Восточной Азии, Латинской Америки, Африки, где лихорадка денге была гиперэндемической в течение десятилетий и представляет собой серьезную проблему [11-14].

Основная цель данного исследования – изучить возможность и оценить качество результатов применения методов машинного обучения для анализа и прогнозирования болезни денге

на основе измеренных параметров ДЗЗ, картографических данных и связанных статистических показателей.

**Методы и материалы.** Исходной информацией для обучения алгоритмов являлись данные об окружающей среде собранные Центром по контролю и профилактике заболеваний (https://www.cdc.gov/Dengue), Национальным управлением по вопросам океана и атмосферы в Министерстве торговли США (http://www.healthmap.org/dengue/en), Министерством здравоохранения Филиппин (https://www.doh.gov.ph).

В качестве метрики оценки точности была использована среднеквадратическая ошибка, для ее расчета и демонстрации качества моделей на основе общего рейтинга использовался сервис drivendata.org. Список параметров включал в себя следующие категории исходных данных: аббревиатуры городов, дату измерений, текущие климатические показатели и их прогноз (температуру, влажность, количество осадков) по данным NOAA's GHCN, PERSIANN, NOAA's NCEP, значения индекса NDVI рассчитанные для прилегающих к городу пикселей спутникового снимка [15-17].

Визуальное представление ряда климатических параметров и временного охвата имеющихся данных приведено на рисунке.

В работе ставились две основные цели: построить модель точного прогноза для отслеживания тенденции развития эпидемии денге путем сравнения различных современных алгоритмов и технологий прогнозирования и оценка степени влияния различных показателей исходного набора данных на итоговый результат прогноза. С точки зрения первой цели анализировался достаточно широкий набор инструментов прогнозирования - алгоритмы и технологии в геоинформационных системах, типовых библиотеках машинного обучения, средствах обучения и использования искусственных нейронных сетей, программном обеспечении, позволяющем использовать алгоритмы машинного обучения с помощью графического интерфейса, без необходимости программирования.

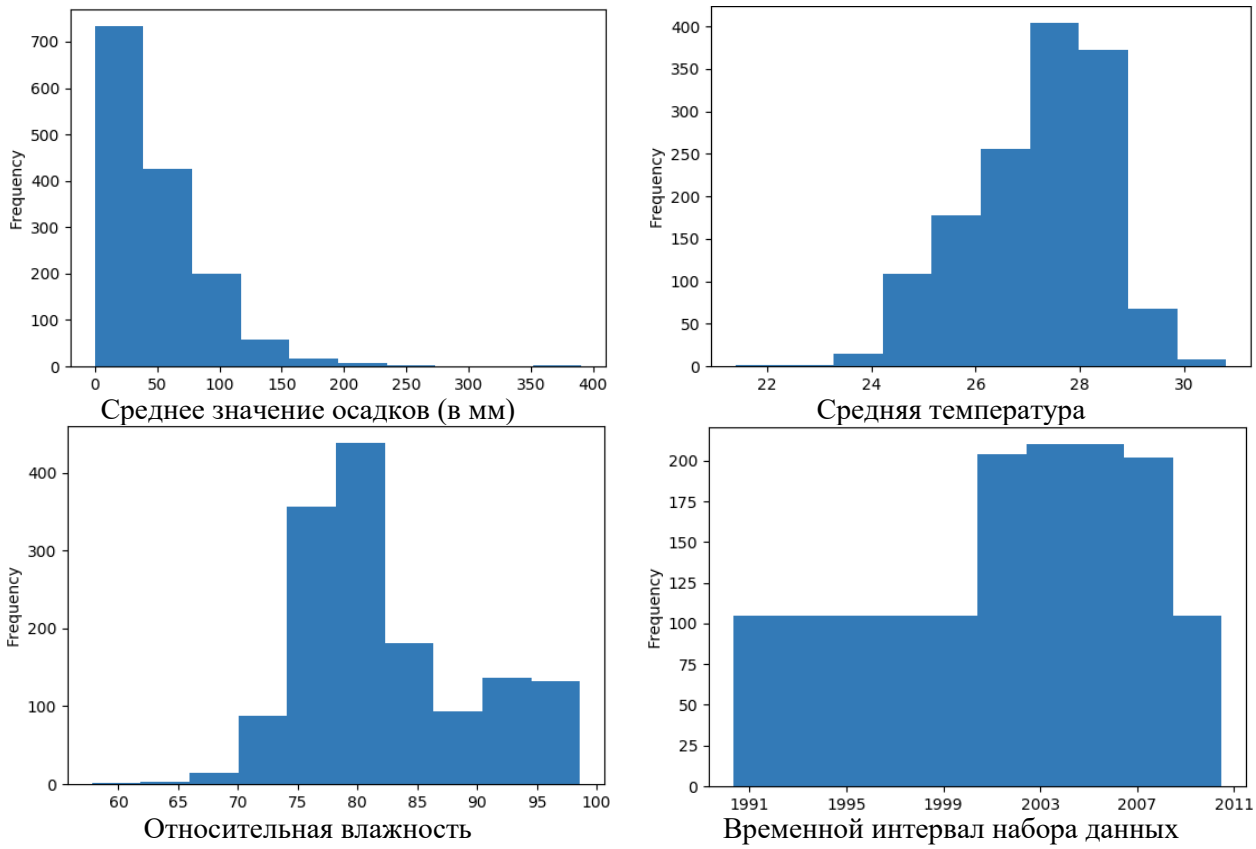



Рис.1. Графическое представление ряда параметров исходного набора данных

В современных геоинформационных системах (ArcGIS Pro, GRASS GIS, SAGA GIS) для моделирования пространственного размещения показателей объектов или явлений используются следующие методы пространственного прогнозирования: IDW, методы локальных и глобальных полиномов, кригинг. Проблема этих методов состоит в том, что они рассчитаны на пространственное прогнозирование, не учитывающее в явном виде временной составляющей. Для того чтобы учитывать данные о времени измерений нужно отдельно использовать либо специализированные алгоритмы (например, ARIMA/SARIMA), либо информацию о годе, месяце, квартале, дне и т.д. преобразовывать в отдельные параметры с числовыми значениями и добавлять их в атрибутивную таблицу внесения предыдущих измерений параметров в качестве отдельных дополнительных колонок анализируемых данных. Используя подобные модификации данных возможно учитывать временную составляющую для обычных алгоритмов регрессионного анализа, например, линейной регрессии, случайного леса. В качестве типовых методов машинного обучения были использованы наиболее популярные технологии – градиентный бустинг на базе деревьев решений (в реализации xgBoost, LightGBM, CatBoost), случайные лес, метод ближайших соседей, линейная регрессия. Кроме традиционной реализации данных технологий и алгоритмов в библиотеках языка python (scikit-learn) представляла интерес оценка возможности их использования в программном обеспечении с графическим интерфейсом. Этот вариант важен тем, что далеко не все специалисты по геоинформационным системам являются программистами в достаточной степени, чтобы использовать все последние достижения в области искусственного интеллекта [18]. Использование графического интерфейса позволяет значительно расширить набор способов анализа данных разных типов. Для оценки качества такого подхода использовался программный пакет (написать для чего этот пакет) Orange Lab (orange.biolab.si), позволяющий с помощью методов визуального программирования оперировать методами анализа данных и машинного обучения.

Указанные технологии хорошо изучены и являются негласными стандартами при построении математических моделей, если же говорить о перспективах развития, то самым новым направлением для задач прогнозирования пространственно-временных данных являются нейронные сети. Нейронные сети обладают способностью обучаться на имеющихся данных, что имеет большое теоретическое и практическое значение для создания моделей анализа и прогнозирования временных рядов. Дополнительными плюсами этой технологии являются способность эффективно работать в таких нестандартных условиях как недостаточность понимания структуры системы, ошибки и недостаточность в экспериментальных данных. Несмотря на то, что нейронные сети являются нелинейными структурами, они позволяют аппроксимировать произвольную непрерывную функцию. Модель на основе нейронной сети возможно обучить таким образом, чтобы она с высокой достоверностью определяла дальнейшее развитие изучаемого процесса или явления в указанный период. Поскольку временные ряды большинства показателей явлений и процессов представляют собой непрерывные функции, то применение нейронных сетей при их прогнозировании является вполне оправданным и корректным. Процесс использования нейронных сетей строится на использовании библиотек python – tensorflow, theano, pytorch и некоторых других. Кроме того, есть возможность использования созданных моделей и скриптов с интерфейсом ГИС, поскольку во многих распространенных настольных геоинформационных системах языком разработки модулей также является python [19-23].

Говоря о исходных данных для обучения вышеописанных алгоритмов, нужно отметить, что пространственно-временное моделирование описывает и моделирует процессы и явления в четырех измерениях – в трёх пространственных измерениях и во времени. В идеальном случае моделирования необходимо отслеживать все степени свободы (все доступные параметры), но в непрерывных системах каждая точка в пространстве прибавляет дополнительные степени свободы, что в итоге приводит к бесконечному числу измерений. В таких случаях необходимо провести дискретизацию, тем самым уменьшив число степеней свободы до допустимого в компьютерном моделировании. Таким образом, перед тем как перейти непосредственно к со-

зданию модели необходимо выполнить отбор параметров. Это может быть сделано как с помощью расчета числовых показателей корреляции и энтропии, так и визуально, с помощью диаграмм, либо на карте. Сравнение параметров между собой может быть выполнено как с помощью традиционного расчета значения корреляции (например, коэффициенты Пирсона, Спирмена), так и с помощью специализированных расчетных показателей, ориентированных на анализ пространственных данных и временных рядов. Пространственную корреляцию обычно измеряют с помощью индекса Морана, показывающего присутствует ли кластеризация объектов, либо они расположены хаотично. Расчет этого показателя реализован, например, в ArcGIS Pro, GRASS GIS, PySAL. Для анализа энтропии временного ряда наиболее универсальным является показатель Ляпунова. Также для это цели могут быть использованы коэффициент Хёрста, detrended fluctuation analysis. Наибольшее количество инструментов расчета показателей хаотичности временного ряда реализовано в библиотеке nolds (https://pypi.org/project/nolds) для языка python. Также для оценки значимости конкретных параметров на результат предсказания были использован параметр feature importances (присутствует в большинстве реализаций алгоритмов scikit-learn) и специализированный алгоритм для отбора параметров – Boruta. При анализе имеющихся данных выявилась высокая корреляция между значениями температуры, влажности и количества осадков, полученных с метеостанций и по результатам анализа спутниковых измерений. По результатам выполнения оценки важности с помощью Boruta и feature importances наиболее приоритетными являются температура, значения индекса NDVI и сезон.

Сводная таблица результатов оценки прогностических моделей.

| **Место** | **Модель** | **MAE** |
|---|---|---|
| 1 | Ансамблирование SARIMA и xgBoost | 25.8 |
| 2 | Случайный лес, подбор гиперпараметров, отсечение параметров с корреляцией более 0.9 | 26.4 |
| 3 | Случайный лес, подбор гиперпараметров | 26.6 |
| 4 | Случайный лес в Orange | 26.6130 |
| 5 | Случайный лес, подбор гиперпараметров, отсечение атрибутов с рангом Boruta менее 20 | 26.9 |
| 6 | Случайный лес, подбор гиперпараметров, отсечение атрибутов с Feature importances <0.3 | 27.1 |
| 7 | нейронная сеть с двумя скрытыми слоями | 27.4 |
| 8 | xgBoost, параметры по умолчанию | 27.9 |
| 9 | LightGBM, параметры по умолчанию | 28.7 |
| 10 | CatBoost, подбор гиперпараметров, 5 категориальных переменных | 29.561 |
| 11 | Линейная регрессия в Orange | 29.8173 |
| 12 | SARIMA | 30.3 |
| 13 | Keras, без скрытых слоев | 32.5 |
| 14 | KNN в Orange | 33.8774 |
| 15 | CatBoost, параметры по умолчанию, 4 категориальных переменных | 37.1 |
| 16 | CatBoost, параметры по умолчанию, категориальная переменная - город | 37.2 |

Наилучшим образом себя обычно проявляют следующие алгоритмы машинного обучения: случайный лес, градиентный бустинг на основе деревьев решений (в реализации xgBoost, LightGBM, CatBoost), нейронная сеть (в реализации Tensorflow и Keras) с двумя архитектурами, различающимися наличием скрытых слоев, SARIMA и ансамблирование результатов работы алгоритмов SARIMA и xgBoost. Эти алгоритмы и были выбраны для анализа данных. Для всех этих алгоритмов выполнялся подбор гиперпараметров и анализировались варианты с отсечением маловажных атрибутов (по результатам анализа корреляционных матриц, параметра feature importances, сводной таблицы Boruta).

**Результаты.** При проведении экспериментов также учитывались следующие особенности:

- для использования нейронных сетей выполнялась предварительная нормализация данных;
- в процессе обучения для алгоритмов в Orange Lab автоматически выполняется подбор гиперпараметров, а для реализации на python это настраивалось вручную;
- для обоих вариантов нейронных сетей использовался метод активации «relu» и оптимизация – «adam»;
- скрытые слои для нейронной сети состояли из 5 и 13 элементов;
- для CatBoost использовались варианты с указанием только города в качестве категориального параметра и дополнительно добавлением к нему года, сезона и недели

Результаты качества построения математических моделей описанными алгоритмами приведены в таблице.

**Заключение.** По результатам проведенных исследований были сформулированы следующие выводы:

- для пространственно-временного прогнозирования с большим количеством параметров необходимо сочетание различных алгоритмов обработки данных методами бустинга или бэггинга;
- предобработка данных как правило является не менее сложным и продуктивным процессом, чем построение моделей;
- графические инструменты для обработки данных и построения моделей, такие как Orange уже практически не уступают по точности скриптам на языке python при этом превосходя их в скорости создания моделей, но необходимо отметить, что пока им не достает гибкости в обработке данных и представлении результатов;
- алгоритм случайного леса наряду с градиентным бустингом являются наиболее универсальными для задач пространственно-временного прогнозирования;
- в случае большого числа входов нейронной сети практически обязательны скрытые слои.

Также для увеличения точности и большей универсальности в дальнейших исследованиях планируется расширить исходные параметры открытыми данными о мониторинге москитов Ae. aegypti и Ae. Albopictus, дополнительными индексами, рассчитываемыми на основе спутникового мониторинга и для построения моделей временного ряда использовать нейронные сети архитектуры LSTM и ее вариаций.

## ЛИТЕРАТУРА


[1] *Москвичев В.В.* Риски развития и мониторинг социально-природно-техногенных систем – основа безопасности, стратегического планирования и управления промышленными регионами страны // Сборник трудов всероссийской конференции (29-31 августа 2017 г., г. Бердск). Новосибирск: ИВТ СО РАН, 2017, С.22-26.

[2] *Колесников А.А., Кикин П.М., Комиссарова Е.В., Грищенко Д.В.* Использование машинного обучения для построения картографических изображений // Международная научно-практическая конференция «От карты прошлого – к карте будущего», 28-30 ноября 2017, г. Пермь – г. Кудымкар. С. 110-120.

[3] *Brown, F.J., Reed C.B., Hayes J.M., Wilhite A.D., Hubbard K.* A prototype drought monitoring system integrating climate and satellite data // Proceedings of the Pecora L5/land satellite information 1V/IS-PRS commission I/FIEOS, 2002, Colarado, USA.
[4] *Breiman L.* Random forests // Machine Learning. 2001. Vol. 45, N. 1. P. 5-32.
[5] *Колесников А.А., Кикин П.М., Комиссарова Е.В., Грищенко Д.В.* Анализ и обработка данных ДЗЗ методами машинного обучения // Сборник материалов V Международной научной конференции «Региональные проблемы дистанционного зондирования Земли» (РПДЗЗ-2018). Красноярск, 2018. С. 130-134.
[6] *Goodfellow I., Bengio Y., Courville A.* Deep Learning. MIT Press, 2016. 800 p. ISBN: 9780262035613
[7] *Ничепорчук В.В., Чернякова Н.А.* Использование инфраструктур данных для оценивания рисков чрезвычайных ситуаций // Сборник трудов всероссийской конференции (29-31 августа 2017 г., г. Бердск). Новосибирск: ИВТ СО РАН, 2017. С. 280-285.
[8] The Influence of Global Environmental Change on Infectious Disease Dynamics: Workshop Summary. Washington (DC): National Academies Press (US), 2014. N 3. **https://www.ncbi.nlm.nih.gov/books/NBK241611** (дата обращения 11.06.2019).
[9] *Chen L.-C., Papandreou G., Kokkinos I., Murphy K., Yuille A.L.* Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. arXiv preprint arXiv:1606.00915, 2016.
[10] *Kuno G.* Research on dengue and dengue-like illness in East Asia and the Western Pacific during the First Half of the 20th century // Reviews in medical virology. 2007. Vol. 17, N 5. P. 327-341. DOI: 10.1002/rmv.545.
[11] *Xiao J.P., He J.F., Deng A.P., Lin H.L., Song T., Peng Z.Q.* Characterizing a large outbreak of dengue fever in Guangdong Province // China. Infectious diseases of poverty. 2016. PubMed Central PMCID: PMCPmc4853873. DOI: 10.1186/s40249-016-0131-z.
[12] *Shepard D.S., Undurraga E.A., Halasa Y.A.* Economic and disease burden of dengue in Southeast Asia // PLoS neglected tropical diseases. 2013. Vol. 7, N 2. P. e2055. PubMed Central PMCID: PMCPmc3578748. DOI: 10.1371/journal.pntd.0002055.
[13] *Ooi E., Gubler D.* Dengue in Southeast Asia: epidemiological characteristics and strategic challenges in disease prevention // Cadernos de saude publica. 2009. Vol. 25, Suppl 1. P. S115-24.
[14] *Halstead S.* Dengue in the Americas and Southeast Asia: do they differ? // Revista panamericana de salud publica. 2006. Vol. 20, N 6. P. 407-415.
[15] *Haug S., Ostermann J.* A Crop Weed Field Image Dataset for the Evaluation of Computer Vision Based Precision Agriculture Tasks // Computer Vision – ECCV 2014 Workshops. Zurich: Springer, 2014. P. 105-116.
[16] *Peters, J.A., Walter-Shea A.E., Ji L., Vina A., Hayes M., Svoboda D.M.* Drought monitoring with NDVI-based standardized vegetation index // Photogrammetric Engineering and Remote Sensing. 2002. Vol. 68. P. 7175.
[17] *Бериков В.Б., Пестунов И.А., Караев Н.М., Тевари А.* Распознавание гиперспектральных изображений с использованием кластерного ансамбля и частично контролируемого обучения // Сборник трудов всероссийской конференции (29-31 августа 2017 г., г. Бердск). Новосибирск: ИВТ СО РАН, 2017. С. 60-65.
[18] *Bottou L.* Large-scale machine learning with stochastic gradient descent // Proceedings of COMPSTAT' 2010. Springer, 2010. P. 177-186.
[19] *Badrinarayanan V., Kendall A., Cipolla R.* Convolutional encoder-decoder architecture for image segmentation // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. Vol. 39, N 12. P. 2481-2495.
[20] *Dai J., He K., Sun J.* Boxsup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation // IEEE International Conference on Computer Vision. 2015. P. 1635-1643.
[21] *Eigen D., Fergus R.* Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture // IEEE International Conference on Computer Vision. 2015. P. 2650-2658.
[22] *Hariharan B., Arbelaez P., Girshick R., Malik J.* Simultaneous detection and segmentation // European Conference on Computer Vision. Springer, 2014. P. 297-312.
[23] *Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A. C., Fei-Fei L.* ImageNet Large Scale Visual Recognition Challenge // International Journal of Computer Vision (IJCV). 2015. Vol. 115, N 3. P. 211-252.

# ПОСТРОЕНИЕ АНАЛИТИЧЕСКОЙ МОДЕЛИ ДЛЯ СПЕЦИАЛИЗИРОВАННОЙ МОДЕЛЬНО-ОРИЕНТИРОВАННОЙ СИСТЕМЫ СБОРА ДАННЫХ НАУЧНЫХ ИССЛЕДОВАНИЙ

*Коробко А.В., Коробко А.А.*
Институт вычислительного моделирования СО РАН, Красноярск

В основе эффективной обработки экспериментальных данных и данных натурных исследований лежит качество представления первичных результатов экспериментов и наблюдений и эффективная аналитическая обработка результатов исследования. В работе предложен алгоритм автоматического формирования аналитической модели для специализированной модельно-ориентированной системы сбора и анализа данных научных исследований, построенной на базе оригинальной программной платформы.
*Ключевые слова: аналитическая модель, сбор и обработка данных, данные научных исследований, модельно-ориентированная разработка.*

**Введение.** Безусловно, основная ценность проводимых экспериментальных и натурных исследований заключается в полученных результатах, выявленных закономерностях и обнаруженных связях. Другими словами, ученые нацелены на получение новых знаний об исследуемом объекте. В основе эффективной обработки экспериментальных данных лежит качество представления первичных результатов экспериментов, заключающееся в полноте и детальности описания, выборе единых шкал и единиц измерения, использовании одних нормативно-справочных данных и систематизированном хранении. Именно продуманная систематизация и консолидация экспериментальных данных обеспечивает получение достоверных знаний об объекте исследования.

Наиболее распространённые в научной среде инструменты, используемые для хранения и анализа данных – это табличные редакторы [1]. Основные их преимущества – это доступность, популярность в научной среде, наличие стандартных статистических инструментов и возможность визуализации данных. Наряду с достоинствами, табличные редакторы имеют ряд существенных недостатков. Как правило, инструменты имеют ограничения на размеры строк и столбцов, отсутствует контроль ввода, проверка непротиворечивости и целостности данных. Хранение информации в виде файлов исключает возможность одновременной работы с данными нескольких пользователей с различных устройств и платформ. Анализ данных часто ведёт к их изменению или дублированию, что затрудняет сравнительный и ретроспективный анализ данных при сложных экспериментах [2]. Эти недостатки делают невозможным его систематическое использование в серьёзной продолжительной исследовательской работе.

В ответ на запрос ученых, в ИВМ СО РАН разработана программная платформа построения систем сбора и анализа данных исследований [3]. Возможности платформы позволяют создавать специализированные модельно-ориентированные системы учёта (ввода) и анализа первичных данных научных исследований. Создание и наполнение систем выполняется с помощью интерфейса платформы и не требует от пользователя глубоких знаний в области информационных технологий. Однако, построение аналитических отчетов в настоящее время сопряжено с привлечением IT-специалиста. Актуальна задача по повышению доступности средств аналитической обработки данных в рамках программной платформы. Современные тенденции развития технологии анализа данных заключаются в снижении требований к квалификации конечного пользователя [4]. Одним из перспективных направлений «демократизации» аналитики стала разработка средств нативного формирования аналитических запросов к данным, заключающаяся в построении прозрачной аналитической модели [5].

Уникальность платформы построения систем сбора и анализа данных исследований заключается в том, что в ходе создания специализированной системы формируется управляющая модель, которая описывает состав консолидируемых данных и определяет интерфейс системы. Управляющая модель, являясь тематическим ядром системы, хранится в виде метаданных и может стать основой для построения аналитической модели.

Настоящая статья призвана представить оригинальный подход к формированию аналитической модели для специализированной модельно-ориентированной системы на основе управляющей модели. Первый раздел посвящен концепции MDD (модельно-ориентированной разработки) и формальному описанию управляющей модели, формируемой в рамках рассматриваемой программной платформы, на примере Системы исследования состояния почвенного покрова. Во втором разделе представлена теория построения аналитической модели в соответствии со спецификацией CWM. Третий раздел содержит алгоритм формирования аналитической модели специализированной системы сбора данных научных исследований на основе её управляющей модели. В заключении подведены итоги статьи, сформулированы задачи для продолжения настоящего исследования.

**Управляющая модель специализированной модельно-ориентированной системы сбора данных.** Классический модельно-ориентированный подход (MDD) к разработке программного обеспечения (ПО) подразумевает построение набора моделей разного уровня абстракции в ходе проектирования и реализации ПО [6]. Разработка моделей высокого уровня абстракции включает процесс построения мета-метамодели (M3), модели языка моделирования, и процесс построения метамодели (M2), логической модели предметной области приложения в нотации мета-метамодели. Проектирование моделей низкого уровня абстракции состоит из процесса построения моделей прикладного уровня (M1) и этапа формирования экземпляров концептов (M0) определённых на уровне M1 [7].

Авторами настоящей работы ранее предложена оригинальная реализация подхода к модельно-ориентированной разработке. Основные преимущества предложенной реализации по сравнению с классическим подходом заключается в формировании управляющей модели (модели уровня M1) в ходе построения системы средствами платформы и автоматическом формировании прикладных моделей на основе управляющей. Такая реализация помогает существенно снизить требования к квалификации пользователя в области информационных технологий, позволяя ему сосредоточиться на исследованиях, и сохраняет при этом гибкость и универсальность модельно-ориентированной разработки. Включение управляющей модели в состав системы в виде метаданных позволяет оперативно реагировать на изменяющиеся требования к тематическому наполнению и перманентно развивать систему в соответствии с растущим охватом проводимых научных исследований. Преимущества предложенной реализации модельно-ориентированного подхода достигнуты благодаря умышленному сужению функционала платформы до построения систем сбора и консолидации исследовательских данных и за счет разработки собственной метамодели (уровень M2), соответствующей назначению систем. В качестве мета-матамодели (уровень M3) используется нотация Unified Modeling Language (UML).

Метамодель содержит три класса объектов: класс «Объект» – N, класс «Атрибут» – F и класс «Группа» – G. Объекты класса «Атрибут» описываются тройкой F=(A, T, D), где A – наименование атрибута, T – обозначает имя специализированного типа атрибута, D – флаг темпоральности атрибута. Пример управляющей модели, построенной в соответствии с предложенной метамоделью, приведен на рис. 1.

В рамках метамодели задан ряд отношений, определяющий связи между экземплярами (объектами) классов модели. Между объектами определены два вида отношений, «Вложенность» и «Зависимость». Одно-многозначное отношение «Вложенность» – $\varphi$, задано на множестве N, $\varphi \subseteq N \times N$ и предназначено для задания организационной иерархии объектов. Одно-многозначное отношение «Зависимость», обозначим его, как $\chi$, задано на множестве N, $\chi \subseteq N \times N$. Отношение позволяет связывать объекты друг с другом, реализуя различные функциональные взаимодействия. Много-многозначное соответствие между объектами и атрибутами – «Обладание», обозначим как $\theta$, где $\theta \subseteq N \times F$. Одно-многозначное соответствие между объектами и группами – «Группировка», обозначим как $\psi$, где $\psi \subseteq N \times G$. Более подробно метамодель описана в работе [8].

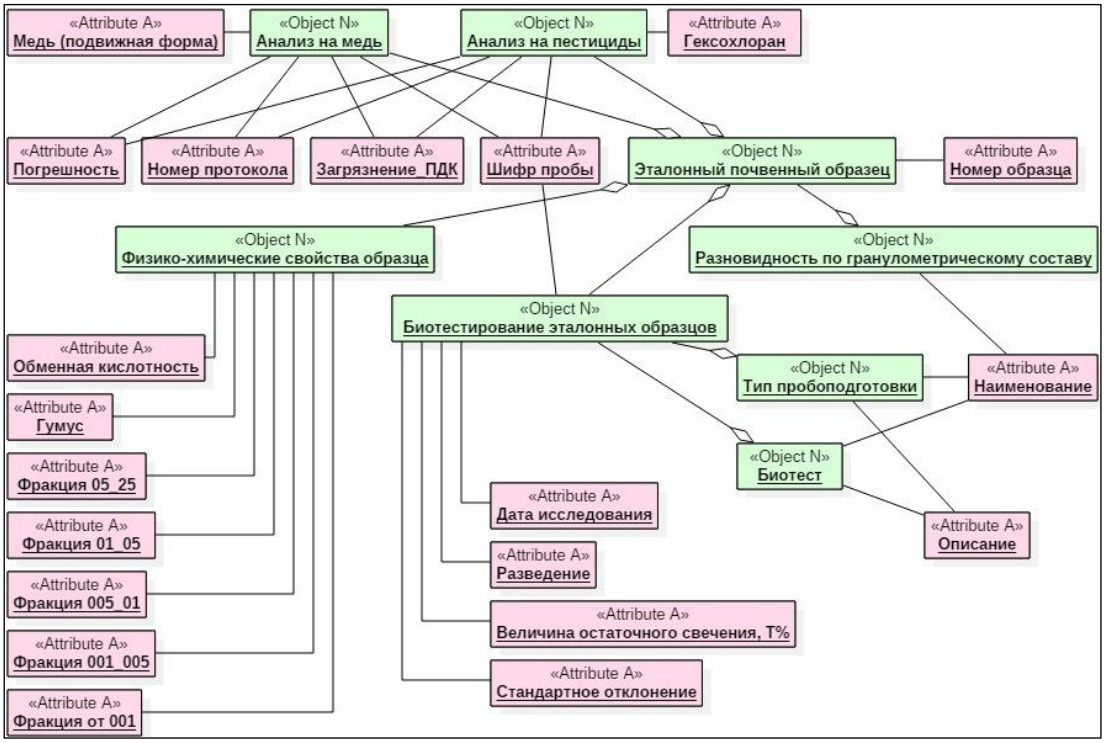


Рис. 1. Фрагмент управляющей модели Системы исследования состояния почвенного покрова.

Управляющая модель, формируемая в соответствии с предложенной метамоделью, формально описывает как предмет изучения, так и результаты, полученные в ходе научного исследования. В работе [3] предложен порядок формирования управляющей модели и требования к ее элементам, обеспечивающие консистентное хранение данных и достоверность результатов их анализа.

**Аналитическая модель.** Концепция модельно-ориентированной разработки, описанная выше, нашла широкое применение во многих областях информационных технологий. В частности, консорциум Object Management Group (OMG) в 2003 году предложил собственную метамодель для обеспечения информационного обмена (взаимодействия) разнородных систем с целью аналитической обработки данных – спецификацию CWM (Common Warehouse Metamodel – общая метамодель хранилищ данных) [9].

Спецификация CWM представляет собой набор моделей уровня M2 (в соответствии с подходом MDD), обеспечивающих описание реляционных и XML источников данных, аналитической модели (рис. 2) в терминах технологии On-Line Analytical Processing (OLAP) и структуры Хранилища данных, предназначенного для консолидации «операционных» данных из разнородных источников и подготовки их к оперативной обработке.

Ключевым требованием OLAP является представление данных в многомерном виде, интуитивно понятном пользователю. Многомерность заключается в разделении данных на аспекты анализа – *измерения* – и агрегируемые числовые характеристики исследуемого процесса – *показатели*. В основе аналитической модели спецификации CWM лежит предложенная Матео Голфарелли (и др.) в 1998 году модель измерение-факт (dimensional fact model) [10] (Рис. 2). Важно, что показатели, описывающие один анализируемый факт, группируются в куб («Cube»). На пример, показатели «Обменная кислотность», «Гумус» и «Фракция от 0,01» характеризуют факт «Физико-химические свойства почвенного образца». Измерения («Dimension») имеют иерархическую структуру и объединяют несколько уровней анализа («Hierarchy»). Так, измерение «Время» включает такие уровни иерархии как «Дата», «День недели», «Месяц», «Квартал», «Год», «Сезон» и т.д. В рамках одной схемы кубы и измерения связаны отношением ассоциации («CubeDimensionAssociation»), с учетом разных уровней иерархии. Принимаемые значения измерения обозначаются как отдельный класс «MemberSelection».

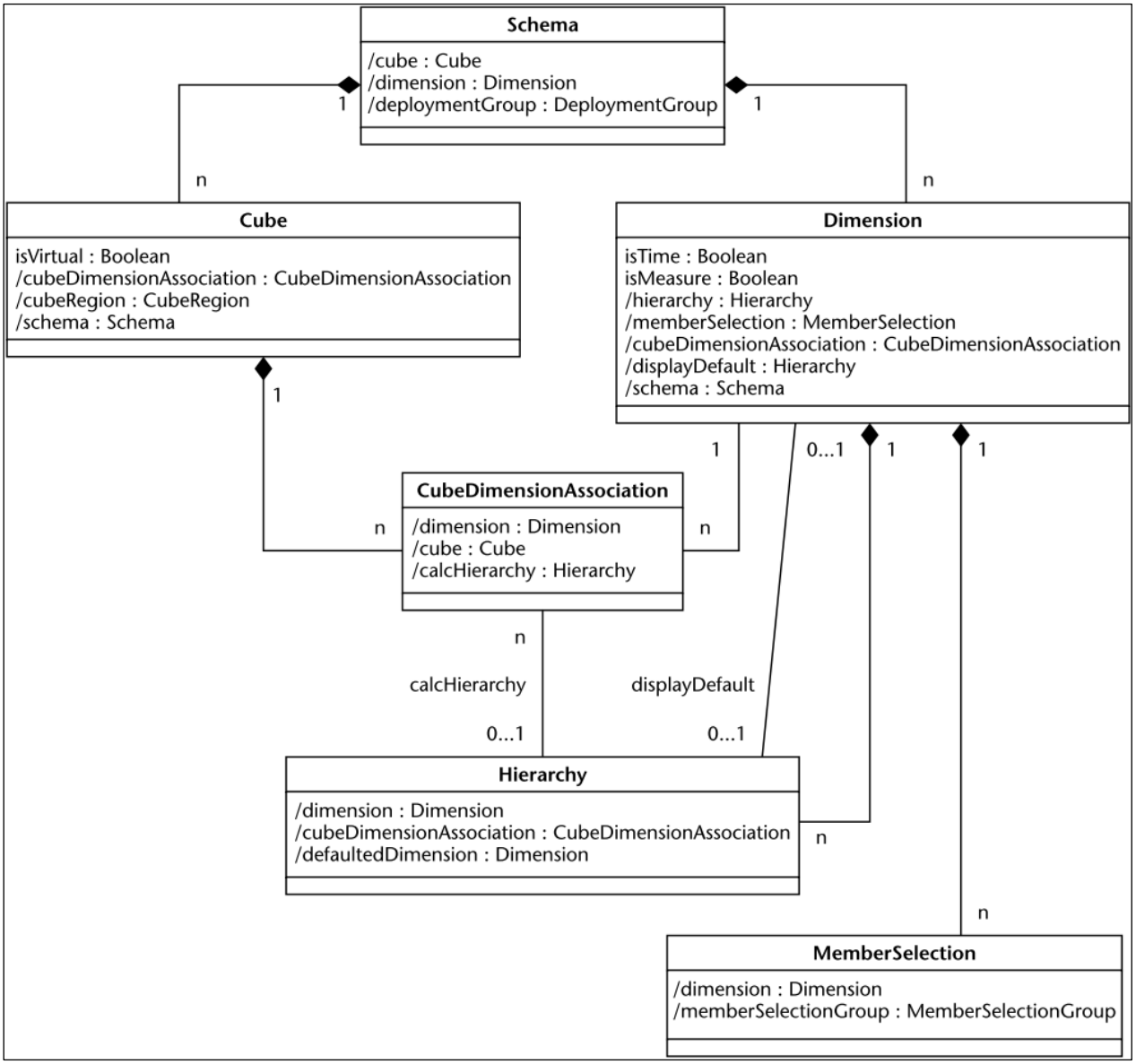



Рис. 2. Метамодель аналитической модели спецификации CWM.

Построение многомерной аналитической модели для модельно-ориентированной системы обеспечивает возможность обработки результатов аналитических запросов в распространенных BI-системах (Business Intelligence – бизнес-аналитика) поддерживающих информационный обмен в соответствии со спецификацией CWM. Актуальной является задача разработки алгоритма формирования многомерной модели на основе свойств оригинальной метамодели платформы.

**Алгоритм формирования аналитической модели.** Для описания алгоритма воспользуемся приведенным выше формальным описанием метамодели управляющей модели специализированной модельно-ориентированной системы в терминах теории множеств.

Алгоритм (рис. 3) заключается в последовательном рассмотрении экземпляров класса «Объект» управляющей модели ($\forall n \in N$) и их атрибутов ($\forall f \in F \mid f\theta n$). В теории многомерного моделирования атрибуты делятся на «не агрегируемые» (dimensional) и «агрегируемые» (non dimensional). «Не агрегируемые» атрибуты участвуют в формировании «измерения», а «агрегируемые» атрибуты одного объекта (или поля одной таблицы, в классическом случае) составляют «куб». Dimensional атрибуты являются не агрегируемыми условно, так как к ним может быть применима агрегатная функция «count». Исключение составляют измерения, которые функционально не зависят от других измерений. В таком случае свойство «isMeasure» измерения аналитической модели устанавливается как «false». В алгоритме стандартная проверка

типа атрибута дополнена процедурой синтаксического анализа наименования атрибута. Несмотря на числовой тип, атрибут не является агрегируемым, если в его названии присутствуют слова-маркеры, свидетельствующие о его «справочном» характере. Множество слов-маркеров может расширяться путем накопления прецедентов. «Куб» каждого рассмотренного объекта добавляется в множество кубов аналитической модели C. Между объектом n и порожденным им кубом устанавливается ассоциативная связь, записанная в виде отношения σ. Аналогично формируется множество измерений D аналитической модели и устанавливается ассоциативная связь между объектом и измерением – δ. Эти связи нужна для работы алгоритма (рис. 4) формирования ассоциаций между кубами и измерениями.

Отношение χ описывает связь между объектами управляющей модели. Запись nχk, где n,k∈N, означает, что у объекта n есть атрибут, связанный с объектом k. На пример объект «Биотестирование эталонных образцов» связан с объектом «Биотест». Это значит, что при вводе данных о результатах биотестирования почвенных образцов необходимо выбрать конкретный вид используемого биотеста. В свою очередь «Биотест» является самостоятельным объектом – пользователь может расширять набор его атрибутов и добавлять при необходимости новые значения.

При построении аналитической модели отношение χ трактуется как внешний ключ или отношение функциональной зависимости. Ассоциативные связи между кубами и измерениями создаются в двух случаях. Во-первых, куб σ(n), порожденный некоторым объектом n, должен быть связан с измерением δ(n), порожденным этим же объектом. Во-вторых, σ(n) должен быть связан со всеми измерениями, порожденными объектами, напрямую и транзитивно связанные отношением χ с рассматриваемым объектом n – элементами множества χ*(n). Множество χ*(n) представляет собой транзитивное замыкание отношения χ для объекта n.

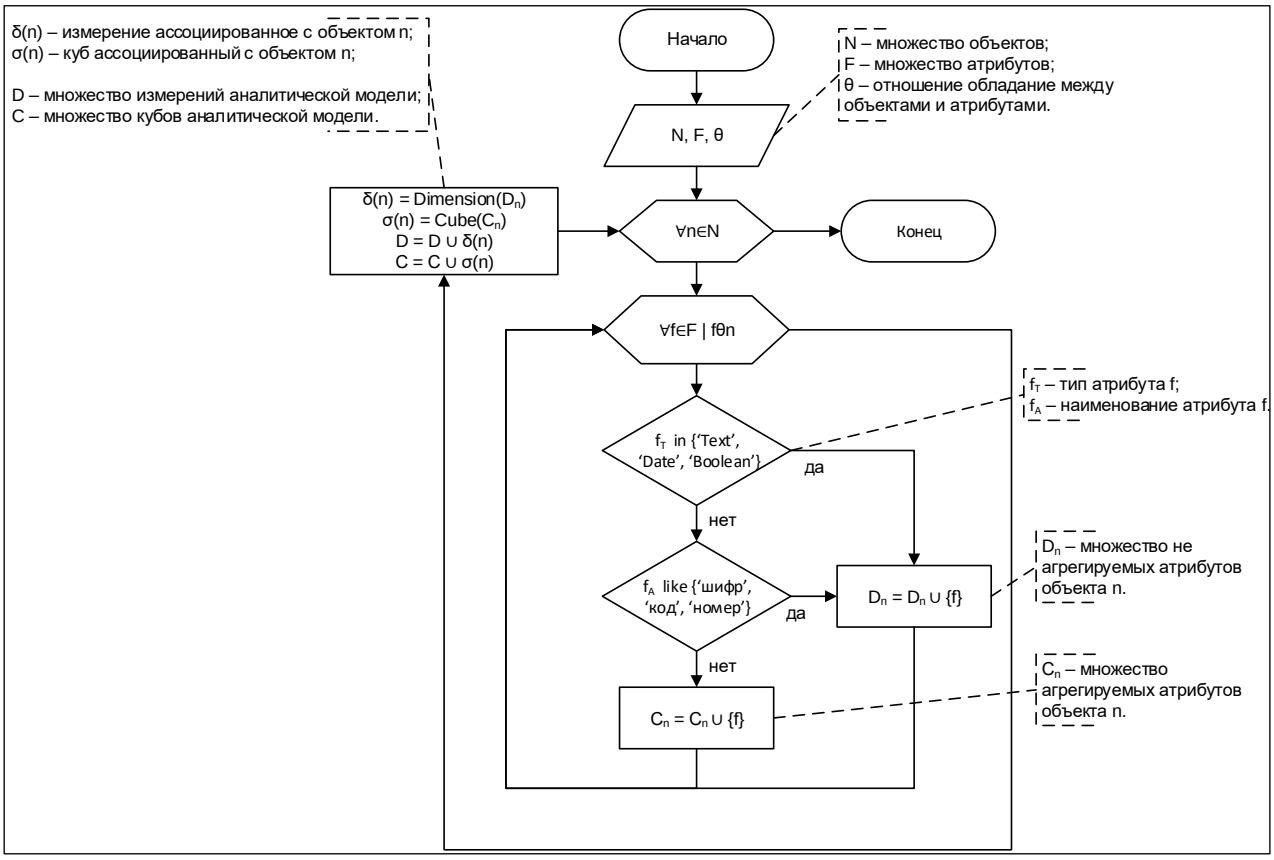



Рис. 3. Алгоритм формирования множеств кубов и измерений аналитической модели.

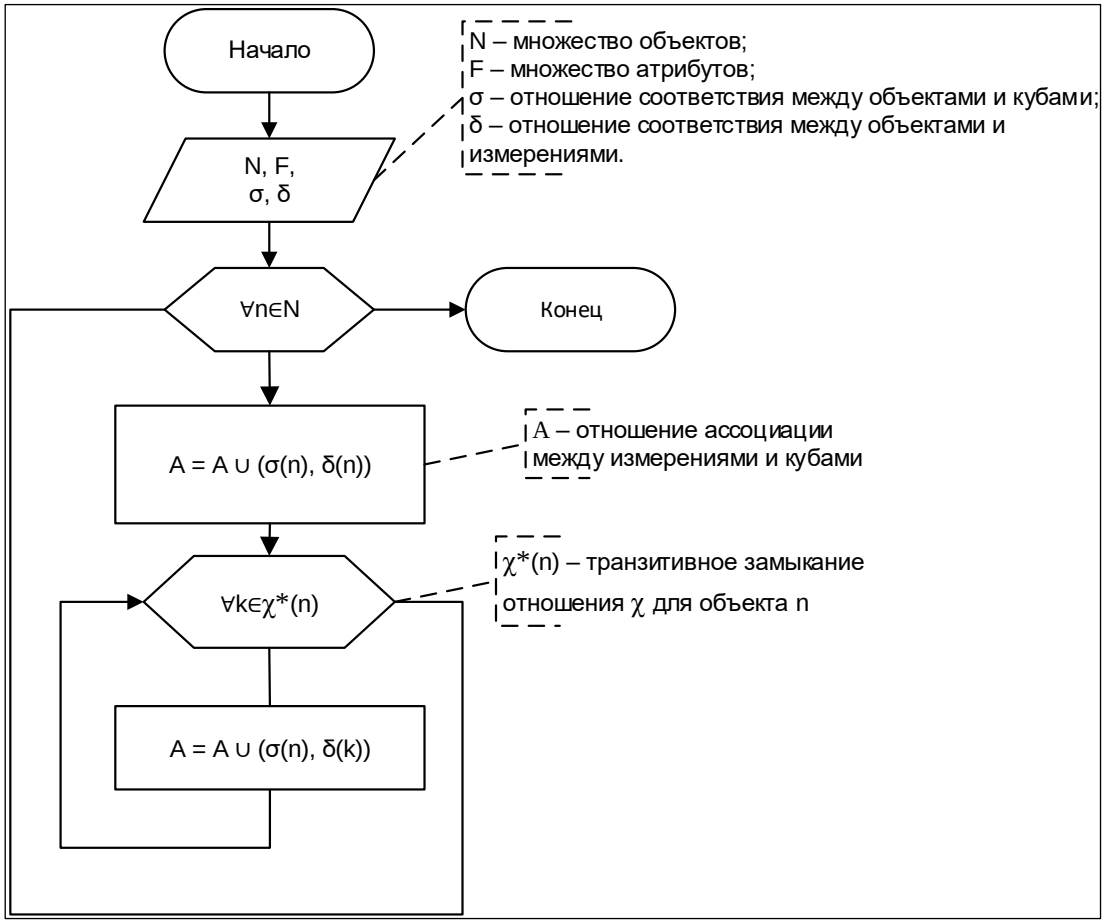



Рис. 4. Алгоритм формирования отношения ассоциации между измерениями и кубами аналитической модели.

Последовательное применение двух алгоритмов позволяет сформировать аналитическую модель в составе: измерения (Dimensions), кубы (Cubes) и связь между ними (CubeDimensionAssociation). В рамках настоящей статьи мы не рассматриваем задачу построения иерархий измерения, оставляя ее для будущих исследований.

**Заключение.** Платформа построения модельно-ориентированных систем разработана сотрудниками ИВМ СО РАН для поддержки консолидации, хранения и аналитической обработки данных научных исследований. Технологическая и методическая основа платформы позволили создать Систему исследования состояния почвенного покрова силами самих исследователей-биофизиков без привлечения IT-специалистов. Платформа обеспечивает возможность консистентного непротиворечивого хранения исследовательских данных, выполнение анализа в рамках долгосрочных многоэтапных научных проектов и сопоставление результатов исследований, выполняемых различными группами ученых. С целью повышения доступности аналитической обработки, в частности инструментов формирования аналитических запросов, в работе была поставлена и решена задача разработки алгоритма для формирования аналитической модели специализированной системы на основе управляющей модели. Представленный алгоритм является теоретической основой для разработки мастера построения аналитических отчетов для специализированных модельно-ориентированных систем. Следующим этапом создания доступного инструмента анализа данных научных исследований должен стать алгоритм генерации SQL-запросов к базе данных системы в соответствии с пользовательским запросом в терминах аналитической модели.

*Исследование выполнено при финансовой поддержке РФФИ и Правительства Красноярского края в рамках научного проекта №18-47-240005.*

## ЛИТЕРАТУРА


[1] *Broman K.W., Woo K.H.* Data Organization in Spreadsheets // Am. Stat. 2018.

[2] *Panko R.R.* What We Know About Spreadsheet Errors // J. Organ. End User Comput. 2014.

[3] *Korobko A., Korobko A., Kolosova E.* Constructing the model-driven system for scientific researches support on the original software platform for primary data consolidation // Surveying Geology & Mining Ecology Management (SGEM). 2018. Vol. 18, N 2.1. P. 255-262.

[4] *Alpar P., Schulz M.* Self-Service Business Intelligence // Bus. Inf. Syst. Eng. 2016. Vol. 58, N 2. P. 151-155.

[5] *AnHai D., Alon H., Zachary I.* Principles of Data Integration. Elsevier, 2012.

[6] *Seidewitz E.* What models mean // IEEE Softw. 2003. Vol. 20, N 5. P. 26-32.

[7] *Atkinson C., Kühne T.* Model-driven development: A metamodeling foundation // IEEE Softw. 2003. Vol. 20, N 5. P. 36-41.

[8] *Korobko A.A.* Algorithm of Interface Generation for Model-Driven Data Consolidation System // 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC). 2018. P. 1-4.

[9] Peyton L. Common Warehouse Metamodel. 2016.

[10] *Golfarelli M., Maio D., Rizzi S.* The Dimensional Fact Model: a Conceptual Model for Data Warehouses // Int. J. Coop. Inf. Syst. 1998. Vol. 07, N 02n03. P. 215-247.

# МЕСТО ГЕОИНФОРМАТИКИ В ЦИФРОВОЙ ЭКОНОМИКЕ РОССИИ

*Кошкарев А.В.*
Институт географии РАН, Москва

Дан детальный анализ Плана мероприятий по направлению «Информационная инфраструктура» национального проекта «Цифровая экономика Российской Федерации», который предусматривает создание отечественных цифровых платформы сбора, обработки и распространения пространственных данных и данных дистанционного зондирования Земли из космоса. План нацелен на реализацию законодательных инициатив в сфере обращения пространственных данных и построения российской инфраструктуры пространственных данных.
*Ключевые слова: цифровая экономика, пространственные данные, инфраструктура пространственных данных, геоинформационные системы, геопорталы.*

**Введение.** Несостоявшаяся модернизация экономики России перешла в новую фазу – фазу ее цифровизации. Охватывая все сферы деятельности – науку, культуру, образование, производство, – она действительно способна на многое. В отличие от большинства других наук и технологий геоинформатика изначально опиралась на цифровую среду обращения пространственных данных, и в этом следует видеть перспективы естественного и эффективного ее развития в новых цифровых реалиях. Процесс цифровизации российской экономики инициирован распоряжением Правительства РФ от 28 июля 2017 г. № 1632-р, утвердившем программу (ныне национальный проект) «Цифровая экономика РФ» [1], вслед за которой были разработаны планы ее реализации, включая направление (федеральный проект) «Информационная инфраструктура», часть разделов которого имеет прямое отношение к обсуждаемой теме [2]. Она представлена многочисленными публикациями в СМИ и в научной литературе, но работ о состоянии и перспективах развития геоинформатики, в том числе с учетом содержания федерального проекта, не так много [3-6].

Состоялся, однако, ряд научных мероприятий, либо целиком посвященных этой проблеме, либо включавших доклады на эту тему. Можно назвать VI Научно-практическую конференцию «Актуальные вопросы геодезии и геоинформационных систем» (Казань, 6-7 сентября 2017 г.), Национальную картографическую конференцию (Москва, 16–19 октября 2018 г.), Международную научно-техническую конференцию «Пространственные данные как основа развития цифровой экономики России» (Москва, 28 мая 2018 г.), форсайт-сессию, проведенную 29 ноября 2018 г. Росреестром совместно с НИУ «Высшая школа экономики» [7]. В 2019 г. прошли два мероприятия: Международная научно-практическая конференция «ГИС для цифрового развития. Применение ГИС и ДЗЗ в науке и управлении» (Якутск, 9-10 апреля 2019 г.) и круглый стол «Будущее «бумажных» карт в цифровую эпоху» Комиссии картографии и аэрокосмических методов МГО РГО по инициативе «Федерального центра геодезии, картографии и ИПД». Из их резолюций, в частности, следует, что внедрение технологий ГИС и дистанционного зондирования в экономику регионов является одним из фундаментальных механизмов цифровизации страны.

Осмысление новых реалий и среднесрочное планирование действий, нацеленных на геоинфраструктурное обеспечение цифровой экономики, представляет собой одну из актуальных проблем. Цель работы – анализ текущего состояния и путей развития отечественной геоинформатики с учетом новых задач, поставленных перед нею документами стратегического планирования и новыми нормативными правовыми актами.

**Геоинформатика в эпоху инфраструктур пространственных данных.** Понимание сегодняшней роли геоинформатики и ее окружения, включая картографию и технологии дистанционного зондирования Земли (ДЗЗ), невозможно без учета современного этапа перехода от эпохи географических информационных систем (ГИС) к инфраструктурам пространственных

данных (ИПД). Имеются в виду цифровые пространственные данные и системы, оперирующие ими. Под этим углом зрения следует рассматривать все недавние инициативы, обсуждаемые ниже.

Нужно сослаться, прежде всего, на указ Президента РФ «О стратегии развития информационного общества в Российской Федерации на 2017–2030 годы» [8] и уже упомянутую программу «Цифровая экономика РФ» [1], с февраля 2019 г. утратившую силу и получившую статус национального проекта. Направление «Информационная инфраструктура» приобрело статус федерального проекта в ее составе [9]. Паспорт проекта наследует основные положения более раннего документа, Плана мероприятий [2], который предусматривает создание отечественных цифровых платформ сбора, обработки и распространения пространственных данных и данных ДЗЗ из космоса, обеспечивающих потребности граждан, бизнеса и власти, содержит мероприятие 04.03.001.012, в результате чего будут «обеспечены способы предоставления в электронном виде пространственных данных и материалов, содержащихся в федеральном фонде пространственных данных, создана единая электронная картографическая основа (ЕЭКО) и государственная информационная система ведения ЕЭКО; создана государственная информационная система Федеральный портал пространственных данных». Кроме того, будут «определены условия и порядок создания инфраструктуры и цифровой платформы приема, сбора, хранения и обработки данных ДЗЗ из космоса на базе межведомственной единой территориально-распределенной информационной системы ДЗЗ», здесь не рассматриваемому, и еще несколько мероприятий, связанных с получением и обработкой ДЗЗ. Эти инициативы могут рассматриваться в качестве мер по реализации Федерального закона «О геодезии, картографии и пространственных данных...» (431-ФЗ), который вступил в силу с 2017 г. [10].

Проанализируем, в какой мере упомянутые документы соответствуют задаче создания ИПД РФ, одного из немногих развитых государств, до сих пор не приступивших к реализации федеральной ИПД и ее региональных узлов, хотя концепция ее создания и развития была подготовлена около пятнадцати лет назад и утверждена в 2006 г. [11].

Прежде всего обратим внимание на исчезновение из лексикона этих документов самого понятия «инфраструктура пространственных данных». Оно заменено на достаточно неопределенный термин «цифровая платформа» или на словосочетание «инфраструктура сбора, хранения и обработки пространственных данных». Рассмотрим компоненты ИПД, именуя их так, как это принято в международном и российском профессиональном геоинформационном сообществе. В их архитектуре принято выделять, четыре компонента: нормативную правовую базу и стандарты; базовые пространственные данные; метаданные, описывающие пространственные данные и сервисы; геопорталы для их поиска.

Назначение *стандартов* в области геоинформатики – обеспечить интероперабельность (взаимосовместимость) пространственных данных и связанных с ними сетевых сервисов (геосервисов). О стандартизации как непременном условии успешности реализации информационных систем в выше упомянутых документах ни слова, однако, к счастью, в 2016 г. после многолетнего перерыва продолжил свою деятельность российский технический комитет по стандартизации ТК 394 «Географическая информация/геоматика», подготовивший несколько национальных стандартов, имеющих отношение к тематике ИПД (http://tc394.ru/DEVELOPMENT_STANDARDS/standarts-tk-394). Среди них ГОСТ Р 57668-2017 (ИСО 19115-1:2014) «Пространственные данные. Метаданные. Часть 1. Основные положения», ГОСТ Р 57656-2017 (ИСО 19115-2:2009) «Пространственные данные. Метаданные. Часть 2. Расширения для изображений и матричных данных», ГОСТ Р 57657-2017 (ИСО 19131:2007) «Пространственные данные. Спецификация информационного продукта», ГОСТ Р 57773-2017 (ИСО 19157:2013) «Пространственные данные. Качество данных». Кроме того, в 2018 г. комитетом подготовлены проекты (первые редакции) двух стандартов новой серии «Инфраструктура пространственных данных», а также, в 2019 г., их окончательные версии; в их числе: ГОСТ Р XXXX «Инфраструктура пространственных данных. Общие требования» и ГОСТ Р XXXX «Инфраструктура пространственных данных. Требования к информацион-

ному обеспечению». Заметим, что в двух последних стандартах речь идет об устройстве российской ИПД, создание которой не предусмотрено федеральным законодательством. В целом стандарты комитета не покрывают полностью предметные области ГИС и ИПД, и разработчики технологических решений в этих областях вынуждены использовать оригинальные стандарты серии ИСО 19100.

Среди пространственных данных разного происхождения, типа и назначения выделяют *базовые пространственные данные* (БПД). Это один из компонентов ИПД, представляющем собой общедоступный набор наиболее востребованных данных многократного использования. Вместо них предлагается создание Единой электронной картографической основы (ЕЭКО). Согласно 431-ФЗ под ЕЭКО понимается пакет открытых цифровых карт и планов разных масштабов. Учитывая, что модели данных ГИС, оперирующих цифровыми пространственными данными, и цифровых карт принципиально различны, о чем говорилось и писалось не раз, речь идет об эрзаце БПД в общепринятом смысле.

Для достижения одной из целей ИПД – инвентаризации национальных ресурсов пространственных данных и геосервисов для обеспечения свободного сетевого доступа к ним на поисковых геопорталах – служат *пространственные метаданные*, данные о свойствах наборов данных, важных для их поиска и использования, упорядоченные некоторым стандартизованным образом. Достаточно давно был разработан международный стандарт на содержание пространственных метаданных ISO 19115:2003 «Geographic information – Metadata», которому соответствовал российский профиль ГОСТ Р 52573-2006. На основе двух новых стандартов ИСО разработаны уже упоминавшиеся стандарты ГОСТ Р 57668-2017 и ГОСТ Р 57656-2017. Подготовка метаданных предполагает наличие системы управления ими, включающей редактор метаданных и валидатор, проверяющий их на соответствие стандартам [12]. В законе 431-ФЗ пространственным метаданным посвящена ст. 14 «Сведения о пространственных данных (пространственные метаданные)». Обязанность их ведения лежит на держателях фондов пространственных данных. С недавних пор метаданные о цифровой и аналоговой картографической продукции, хранимой в Федеральном фонде пространственных данных, доступны на сайте Центра геодезии, картографии и ИПД (http://cgkipd.ru/opendata/spatial_metadata), увы, только в формате электронных таблиц MS Excel, а не в соответствии со стандартами. Задача инвентаризации пространственных данных выходит далеко за пределы фондов того или иного типа, уровня или ведомственной принадлежности; международный опыт свидетельствует о необходимости тотальной инвентаризации распределенных (а не централизованных в фондах, архивах или иных их коллекциях) национальных информационных ресурсов, связывая напрямую производителей (правообладателей) с их пользователями (потребителями) на геопорталах.

Геопортал – еще один ключевой компонент ИПД, представляющий собой сайт для поиска пространственных данных, их наборов, серий и геосервисов по метаданным, или, в терминологии Плана мероприятий и 431-ФЗ (ст. 19), – «федеральный портал пространственных данных», или, пользуясь общепринятой терминологией, национальный геопортал и его региональные узлы (региональные порталы). В 431-ФЗ под ними понимаются «информационные системы» для доступа к картам, планам, снимкам и метаданным. Напомним, что в свое время в сети существовал Федеральный геопортал ИПД, многие годы пребывавший в статусе тестируемого, а сейчас, в мае 2019 г., как и два года назад, недоступный. Российские геопорталы – одно из немногих реализаций компонентов ИПД. В русскоязычной терминологии под ними понимаются как собственно геопорталы ИПД, так и средства и результат веб-картографирования. Поэтому их принято делить на поисковые, визуализационно-поисковые и визуализационные. Ранее опубликованные обзоры российских региональных научно-образовательных и ведомственных (отраслевых) геопорталов позволили выявить ряд достоинств и недостатков в их устройстве и сформулировать некоторые рекомендации относительно возможностей их совершенствования [13-15]. Некоторые примеры российских геопорталов, вошедшие в оконча-

тельную редакцию ГОСТ Р XXXX «Инфраструктура пространственных данных. Общие требования», в их сравнении с международными, зарубежными и региональными геопорталами могут быть рекомендованы читателю для анализа (см. таблицу).

Из российских примеров в полной мере отвечает всем требованиям к функциональности геопортала ИПД лишь геопортал ИВиС ДВО РАН.

**Заключение.** В очередной раз [4, 5] обратим внимание на список организаций-исполнителей Плана мероприятий, который включает Росреестр, Роскосмос, ПАО «Ростелеком», Росстандарт, МГУ им. М. В. Ломоносова, РГ НТИ «Аэронет». По меньшей мере странно, что в нем нет Российской академии наук. Налицо невостребованность отечественной академической и вузовской науки в реализации национального проекта. Как показано нами на примере Института географии РАН [16], он являлся и продолжает оставаться лидером в ряде исследований, непосредственно связанных с геоинформационным картографированием в цифровой среде, созданием новой цифровой и электронной картографической продукции, с использованием инструментальных средств ГИС, проводя исследования по теме 138 «Научные основы разработки методов, технологий и средств исследования поверхности и недр Земли, атмосферы, включая ионосферу и магнитосферу Земли, гидросферы и криосферы; численное моделирование и геоинформатика (инфраструктура пространственных данных и ГИС-технологии)».

Примеры геопорталов.

| **Наименование ИПД/геопортала** | **Тип/уровень** | **Унифицированный указатель ресурсов, URL (дата обращения 29.05.2019 г.)** |
|---|---|---|
| Инфраструктура пространственной информации Европейского союза INSPIRE (Infrastructure for Spatial Information in the European Community) | Межнациональный | http://inspire-geoportal.ec.europa.eu |
| ИПД Арктического региона (Arctic SDI) | Межнациональный | https://geoportal.arctic-sdi.org/# |
| Национальная ИПД США NSDI (National Spatial Data Infrastructure) | Национальный | https://www.geoplatform.gov |
| Геопортал национальной ИПД ФРГ (GDI-DE) | Национальный | http://www.geoportal.de |
| Геопортал ИПД земли Бавария (GDI-BY) | Региональный | https://www.gdi.bayern.de |
| Геопортал национальной ИПД Испании IDEE (Infraestructura de Datos Espaciales de España) | Национальный | http://www.idee.es |
| Геопортал ИПД Каталонии IDEC (Infraestructura de Dades Espacials de Catalunya) | Региональный | https://ide.cat |
| Геоинформационный портал Республики Коми | Региональный | http://gis.rkomi.ru |
| Геопортал электронного правительства Самарской области | Региональный | http://geoportal.samregion.ru |
| Государственная геоинформационная система «Енисей-ГИС» | Региональный | http://www.24bpd.ru |
| Геопортал Института вулканологии и сейсмологии ДВО РАН | Научный | http://geoportal.kscnet.ru |

*Работа выполнена в рамках темы Госзадания № АААА-А19-119022190168-8.*

## ЛИТЕРАТУРА


[1] Программа «Цифровая экономика Российской Федерации». **http://www.consultant.ru/document/cons_doc_LAW_221756** (дата обращения: 30.05.2019).

[2] План мероприятий по направлению «Информационная инфраструктура» программы «Цифровая экономика Российской Федерации». **http://www.consultant.ru/document/cons_doc_LAW_287865** (дата обращения: 30.05.2019).

[3] *Зимин М.В., Кошкарев А.В.* Рынок ДЗЗ в эпоху цифровой экономики // Земля из космоса – наиболее эффективные решения. 2019. № 26. С. 5-8.

[4] *Кошкарев А В.* Геоинформатика в инфраструктурном обеспечении цифровой экономики // Геодезия и картография. 2019. Т. 80, № 1. С. 119-126. DOI: 10.22389/0016-7126-2019-943-1-119-126.

[5] *Кошкарев А.В.* Геоинформатика и цифровая экономика // VI Междунар. науч.-практ. конф. «Актуальные вопросы геодезии и геоинформационных систем» (6-7 сентября 2017 г., Казань). Программа. Тезисы и доклады. Казань, 2017. С. 14-23.

[6] *Карпик А.П., Лисицкий Д.В.* Перспективные направления развития геодезической отрасли в условиях постиндустриальной эпохи и цифровой экономики // Геодезия и картография. 2019. Т. 80, № 4. С. 55-64. DOI: 10.22389/0016-7126-2019-946-4-55-64.

[7] Аналитический доклад по результатам форсайт-сессии «Исследование и прогнозирование потребностей экономики в пространственных данных, данных дистанционного зондирования Земли и геоинформационных технологиях, а также услугах, сервисах и продуктах, созданных на их основе». **https://pd.hse.ru/data/2019/01/15/1146735060/doklad.pdf** (дата обращения: 30.05.2019).

[8] Указ Президента РФ «О стратегии развития информационного общества в Российской Федерации на 2017-2030 годы» от 09.05.2017 № 203. **http://www.kremlin.ru/acts/bank/41919** (дата обращения: 26.05.2019).

[9] Паспорт федерального проекта «Информационная инфраструктура» национальной программы «Цифровая экономика Российской Федерации». **http://майскийуказ.рф/upload/iblock/5dd/Utv_FP_2_II.pdf** (дата обращения: 23.05.2019).

[10] Федеральный закон «О геодезии, картографии и пространственных данных и о внесении изменений в отдельные законодательные акты Российской Федерации» от 30.12.2015 № 431-ФЗ. **http://www.consultant.ru/document/cons_doc_LAW_191496** (дата обращения: 26.05.2019).

[11] Концепция создания и развития инфраструктуры пространственных данных Российской Федерации. **http://www.zakonprost.ru/content/base/part/488442** (дата обращения: 30.05.2019).

[12] *Кошкарев А.В.* Пространственные метаданные и геопорталы как средства интеграции геоинформационных ресурсов и сервисов // Изв. РАН. Серия геогр. 2009. № 1. С. 121-123.

[13] *Кошкарев А.В., Ротанова И.Н.* Российские научно-образовательные и отраслевые геопорталы как элементы инфраструктуры пространственных данных // Вестник НГУ. Сер.: Информационные технологии. 2014. Т.12, вып. 4. С. 38-52.

[14] *Алексеенко Н.А., Кошкарев А.В., Курамагомедов Б.М., Медведев А.А*. Геопорталы российских особо охраняемых природных территорий // Геодезия и картография. 2019. Т. 80, № 5. С. 34-46. DOI: 10.22389/0016-7126-2019-947-5-34-46.

[15] *Кошкарев А.В.* Образовательные геопорталы: данные и сервисы // Геодезия и картография. 2017. Спецвыпуск. С. 33-40. DOI: 10.22389/0016-7126-2017-33-40.

[16] *Кошкарев А.В.* Картография в эпоху инфраструктур пространственных данных // Вопросы географии / Московский филиал ГО СССР / Русское геогр. об-во / Сб. 144: Картография в цифровую эпоху. М.: Издательский дом Кодекс, 2017. С. 29-41.

# ИСПОЛЬЗОВАНИЕ СПУТНИКОВОЙ ИНФОРМАЦИИ В ИССЛЕДОВАНИИ ПРОЦЕССОВ РАСПРОСТРАНЕНИЯ ПРИМЕСЕЙ В ПОГРАНИЧНОМ СЛОЕ АТМОСФЕРЫ

*Леженин А.А.[(1)], Рапута В.Ф.[(1)], Ярославцева Т.В.[(2)]*
[(1)] Институт вычислительной математики и математической геофизики СО РАН, Новосибирск
[(2)] Новосибирский институт гигиены Роспотребнадзора, Новосибирск

Обсуждаются вопросы использования спутниковых данных в моделях оценивания параметров источников примесей. Предложены методы определения высоты дымовых шлейфов. По зимним спутниковым снимкам г. Барнаула выполнены оценки эффективной высоты подъема факелов для крупной ТЭЦ города. Точность оценивания дополнительной высоты подъема дымового факела проведена на основе общепринятых соотношений, включающих в себя тепловой поток плавучести и динамический импульс выброса.
*Ключевые слова: спутниковые снимки, нижняя атмосфера, дымовой шлейф, эффективная высота источника, поток плавучести, динамический импульс, оценка, параметр.*

**Введение.** Для решения проблем окружающей среды важную роль играют как экспериментальные, так и теоретические исследования загрязнения атмосферы. Они позволяют определять уровни воздействия на биосферу и климатическую систему, также проводить оценки риска здоровья населения [1-3].

Дымовые шлейфы выбросов от труб промышленных предприятий и крупных ТЭЦ хорошо прослеживаются на спутниковых снимках [4, 5]. Направления и скорости ветра на высотах выбросов визуализируются траекториями дымовых факелов. Снимки из космоса дают возможность отслеживать шлейфы дыма на большие расстояния. Такая картина наиболее характерна для зимних условий [2].

Модели оценивания высот подъёма дымовых шлейфов базируются на уравнениях гидротермодинамики атмосферы [6-8]. Для получения оценок высоты необходимо привлечение информации о метеорологических условиях и параметрах источников выбросов примеси [9-11]. Отметим, что интенсивность выброса и эффективная высота подъема дымового шлейфа требуют большей точности определения. Важное значение имеет информация о вертикальном профиле ветра. Она может быть получена с использованием как расчетных методов, так и данных спутниковых, аэрологических и метеорологических наблюдений. Для решения поставленных задач необходима разработка математических моделей оценивания параметров с привлечением информации с различных систем наблюдений состояния атмосферы и модельных описаний процессов переноса и диффузии примеси [12]. В конечном итоге это приведет к повышению точности определения полей атмосферного загрязнения.

Целью работы является разработка методов оценки параметров источников дымовых выбросов с использованием спутниковой информации.

**Материалы и методы исследований.** Зимние спутниковые снимки территории г. Барнаула послужили базой для проведения исследований. На них хорошо фиксируются выбросы от основных источников атмосферных примесей, включая ТЭЦ-2, ТЭЦ-3.

На рис. 1 представлен спутниковый снимок г. Барнаула, полученный в Сибирском центре ФГБУ «Научно-исследовательский центр космической гидрометеорологии «ПЛАНЕТА» (http://www.rcpod.ru).

Предварительный анализ показывает, что направления выносов примеси от труб ТЭЦ-2, ТЭЦ-3 совпадают, несмотря на существенное различие в высоте источников. Перенос дымовых шлейфов происходит в северо-восточном направлении. Тени шлейфов на земную поверхность значительно смещены от их траекторий. Это указывает на наличие слабого ветра в атмосферном пограничном слое, что приводит к значительной высоте подъема газо-воздушной смеси, достигающей нескольких сотен метров. Таким образом, дымовые струи находятся на высотах зимнего пограничного слоя атмосферы.

На рис. 2 приведен спутниковый снимок г. Барнаула за 2 декабря 2018 года (ФГБУ «Научно-исследовательский центр космической гидрометеорологии «ПЛАНЕТА», http://www.rcpod.ru).

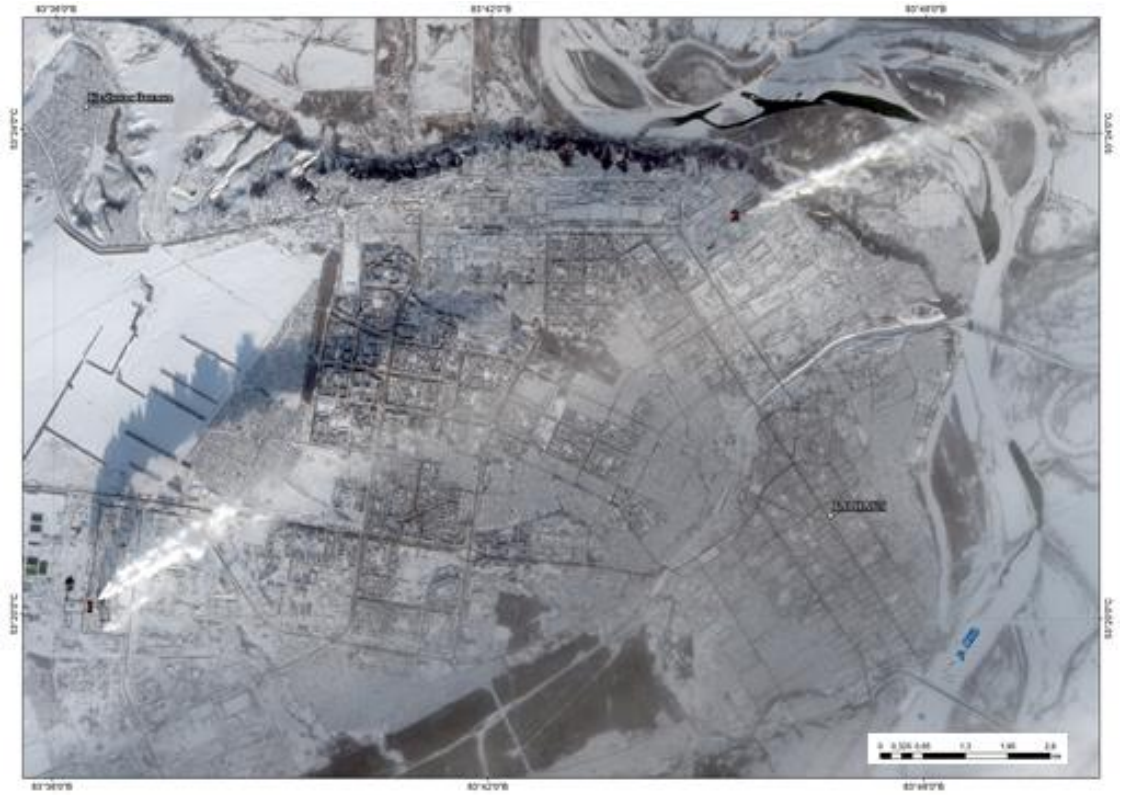


Рис. 1. Спутниковый снимок г. Барнаула за 19 ноября 2018 года на 12 часов 41 мин. местного времени с ИСЗ «Sentinel-2B» № 1 (1-ТЭЦ-3, 2-ТЭЦ-2).




Рис. 2. Спутниковый снимок г. Барнаула за 2 декабря 2018 года на 11 часов местного времени с ИСЗ «Ресурс-П» № 1.

Анализ рис. 2 показывает схожий характер переноса примеси, изображенном на рис. 1. В то же время, наблюдается заметное различие в направлениях выноса выбросов от разновысоких труб ТЭЦ-2, ТЭЦ-3. Это объясняется правым поворотом ветра с высотой в пограничном слое атмосферы [13]. Угол между траекториями шлейфов составляет 30-35 градусов. Тени шлейфов на земную поверхность также значительно удалены от их траекторий.

Для анализа метеорологических условий в течение 19 ноября и 2 декабря 2018 г. использованы данные аэрологической станции г. Барнаула. Станция расположена на высоте 159 м над уровнем моря. В табл. 1 и 2 представлены распределения по высоте основных метеорологических величин: давления, температуры, направления и скорости ветра в 07 и 19 часов местного времени (00 и 12 UTC).

Из табл.1, 2 следует, что в течение дня в пограничном слое атмосферы температурная стратификация была близка к нейтральной и наблюдались слабые ветра юго-западного направления, в то же время, в приземном слое прослеживалась температурная инверсия. В табл. 3 приведены основные данные метеостанции г. Барнаула за 19 ноября 2018 г.

Табл. 1. Параметры нижней атмосферы по данным аэрологической станции г. Барнаула на 19 ноября 2018 г.

| **Время, UTC** | **Высота над уровнем моря, м** | **Давление, hPa** | **Температура, °C** | **Направление ветра, град** | **Скорость ветра, м/с** |
|---|---|---|---|---|---|
| **00** | 159 | 1012 | - 20,1 | 220 | 1 |
| | 232 | 1000 | - 16,1 | 225 | 1 |
| | 355 | 984 | - 10,1 | 232 | 2 |
| | 497 | 966 | - 10,1 | 240 | 3 |
| | 658 | 946 | - 10,1 | 235 | 4 |
| | 691 | 942 | - 10,7 | 237 | 4 |
| | 831 | 925 | - 11,1 | 245 | 4 |
| **12** | 159 | 1017 | - 11,7 | 230 | 1 |
| | 275 | 1000 | - 10,1 | 225 | 3 |
| | 376 | 987 | - 9,3 | 220 | 4 |
| | 702 | 946 | - 10,6 | 250 | 7 |
| | 784 | 936 | - 10,9 | 250 | 8 |
| | 834 | 930 | - 11,1 | 250 | 8 |
| | 975 | 925 | - 10,9 | 250 | 8 |

Табл. 2. Параметры нижней атмосферы по данным аэрологической станции г. Барнаула на 2 декабря 2018 г.

| **Время, UTC** | **Высота над уровнем моря, м** | **Давление, hPa** | **Температура, °C** | **Направление ветра, град** | **Скорость ветра, м/с** |
|---|---|---|---|---|---|
| **00** | 159 | 1019 | - 19,9 | 190 | 1 |
| | 283 | 1000 | - 17,7 | 195 | 3 |
| | 434 | 980 | - 16,9 | 195 | 4 |
| | 511 | 970 | - 16,7 | 211 | 5 |
| | 620 | 956 | - 16,7 | 235 | 7 |
| | 779 | 936 | - 17,6 | 260 | 10 |
| | 867 | 925 | - 18,1 | 255 | 8 |
| **12** | 159 | 1023 | - 20,7 | 220 | 2 |
| | 303 | 1002 | - 16,9 | 235 | 4 |
| | 317 | 1000 | - 16,5 | 235 | 5 |
| | 332 | 998 | - 16,1 | 236 | 5 |
| | 561 | 968 | - 16,7 | 255 | 5 |
| | 821 | 935 | - 17,3 | 240 | 6 |
| | 902 | 925 | - 17,5 | 250 | 6 |

Табл. 3. Данные метеорологической станции г. Барнаула (индекс ВМО 29838) за 19 ноября 2018 г.

| **Время UTC** | **Давление, hPa** | **Температура, °C** | **Направление ветра, град** | **Скорость ветра, м/с** |
|---|---|---|---|---|
| **00** | 1032,0 | -14,7 | штиль | 0 |
| **03** | 1033,5 | -14,4 | штиль | 0 |
| **06** | 1035,0 | -12,3 | штиль | 0 |
| **09** | 1035,5 | -11,0 | штиль | 0 |
| **12** | 1036,8 | -12,9 | 180 | 1 |
| **15** | 1037,0 | -12,1 | 180 | 1 |
| **18** | 1036,3 | -10,2 | штиль | 0 |
| **21** | 1034,3 | -8,4 | 180 | 1 |

Из табл. 3 следует, что в течение суток наблюдались слабые ветра и штилевые условия. Суточный ход атмосферного давления и температуры был слабо выражен. Аналогичная ситуация наблюдалась 2-го декабря 2018 г.

Для моделирования полей атмосферного загрязнения требуется задание характеристик источников примесей и информация о метеорологических условиях [9, 13]. Основными параметрами источника являются геометрическая высота, диаметр трубы, мощность эмиссии, температура и скорость выбрасываемой газо- воздушной смеси. Большое влияние на процессы переноса и диффузии примеси оказывают метеорологические условия: скорость ветра, температурная и влажностная стратификация атмосферы [3, 9, 10].

Существует множество подходов для определения эффективной высоты источника [9, 14, 15]. Фактической высотой источника выброса является его эффективная высота

$$h_{эф} = h_{тр} + \Delta h, \tag{1}$$

где $h_{тр}$ – высота трубы; $\Delta h$ – добавка, обусловленная скоростью выброса примеси и тепловым потоком плавучести.

При нейтральной стратификации вычисление $\Delta h$ производится по формуле [9]

$$\Delta h = a\frac{w_0}{U}D_0 + b\frac{\Phi}{U^3}, \qquad \Phi = g w_0 \frac{D_0^2 \Delta T}{4T_a}, \tag{2}$$

где $g$ – ускорение силы тяжести; $w_0$ – скорость выброса примеси в атмосферу; $U$ – скорость ветра; $D_0$ – диаметр трубы; $\Phi$ – поток плавучести; $\Delta T$ – разность между температурой исходящих газов и температурой атмосферы $T_a$.

При измерении скорости ветра на уровне флюгера: $a = 1{,}9$, $b = 4{,}95$. При измерении скорости ветра на высоте источника: $a = 1{,}5$, $b = 400$ [9].

**Результаты и обсуждение.** В табл. 4 приведены результаты оценивания высот подъёма дымового шлейфа, выполненные на основе спутниковой информации, представленной на рис. 1 и 2. Эффективная высота источника рассчитывается по высоте солнца над горизонтом и положением тени дымовой струи на спутниковом снимке.

Табл. 4. Оценки по спутниковым снимкам высот подъёма дымовой струи от высотной трубы ТЭЦ – 3 г. Барнаула.

| **Время снимка** | **Угловая высота Солнца, град** | **Расстояние до верхней кромки тени струи, км** | **Оценка высоты подъёма струи, м** |
|---|---|---|---|
| **19.11.2018 г. 5:41 UTC** | 17.2 | 2.44 | 745 |
| **02.12.2018 г. 4:07 UTC** | 9.88 | 2.65 | 467 |

Табл. 5. Оценки высот подъема дымовых шлейфов ТЭЦ-3, выполненные по формулам (1)-(2).

| **Скорость ветра, м/с** | **Дополнительная высота, м** | **Эффективная высота, м** |
|---|---|---|
| **1** | 3276 | 3506 |
| **2** | 459 | 685 |
| **3** | 160 | 390 |
| **4** | 82 | 312 |

Анализ табл. 4 показывает, что угловая высота Солнца в рассматриваемые моменты времени была относительно небольшой. В следствие этого измеренные проекции дымовых шлейфов на земную поверхность превышали 2 км. По ним были вычислены высоты подъёма шлейфов.

Расчёт дополнительных высот подъема выбросов ТЭЦ проводился также по формулам (1)-(2). Параметры выбросов в атмосферу от высотной трубы ТЭЦ-3 г. Барнаула взяты из тома ПДВ. Высота источника составляет 230 м., диаметр трубы равен 7,2 м, объем выходящей газо-воздушной смеси – 392,3 м$^3$/с, температура смеси – 120 °С. Рассчитанная по этим данным скорость исходящих газов из трубы $w_0$ =9,6 м/с. Результаты расчетов высоты подъема дымового шлейфа представлены в табл. 5.

Исходя из данных о метеорологических условиях, приведенных в табл. 1 и 2, величина $\Delta T$ была задана 135 °С. Расчеты выполнены для значений скоростей ветра 1, 2, 3, 4 м/с, которые типичны для зимних метеоусловий г. Барнаула.

Совместный анализ табл. 4 и 5 показывает, что приемлемые значения дополнительных высот подъема шлейфа получаются лишь при скоростях ветра больше 2 м/с.

**Заключение.** Проведённое исследование показало возможность использования спутниковой информации о дымовых шлейфах для анализа процессов распространения примесей в атмосфере. Предложен косвенный метод определения эффективной высоты подъёма примесей от дымовых труб. Использование этого метода не предполагает наличия информации о текущих внутренних характеристиках источника: динамического импульса и температуры исходящей из трубы газо-воздушной смеси. На ряде примеров показано, что в условиях слабых ветров предпочтительно применение предложенного метода оценивания, поскольку расчёты по общепринятым методикам дают явно нереальные оценки высот подъёма и не могут быть полезны при проведении моделировании процессов распространения примесей в атмосфере. Следует также отметить наибольшую эффективность метода применительно к зимним условиям. В этот период времени предприятия теплоэнергетики работают с максимальной нагрузкой, а наличие снежного покрова повышает контрастность проекции шлейфа на поверхность земли.

Для более полного использования спутниковой информации о дымовых шлейфах необходимо проведение дальнейших исследований по численному анализу процессов на участках активного подъёма дымовых смесей. В итоге это позволит уточнить внутренние параметры источников выбросов и повысить эффективность моделирования процессов антропогенного загрязнения атмосферы.

*Работа выполнена в рамках Госзадания № 0315-2019-0004, программы президиума РАН № 51 «Изменение климата: причины, риски, последствия, проблемы адаптации и регулирования» (№ 0315-2018-0016), финансовой поддержке РФФИ и Правительства Новосибирской области в рамках научного проекта № 19-47-540008.*

ЛИТЕРАТУРА


[1] *Obolkin V.A., Potemkin V.L., Makukhin V.L., Chipanina Y.V., Marinayte I.I.* Low-level atmospheric jets as main mechanism of long-range transport of power plant plumes in the Lake Baikal Region // Int. J. Environ. Studies. 2014. Vol. 71 (3). P. 391–397.

[2] *Балтер Б.М., Балтер Д.Б., Егоров В.В., Стальная М.В.* Использование данных ИСЗ Landsat для определения концентрации загрязнителей в шлейфах от продувки газовых скважин на основании модели источника // Исследование Земли из космоса. 2014. № 2. С. 55–66.

[3] *Lezhenin A.A., Raputa V.F., Yaroslavtseva T.V.* Numerical Analysis of Atmospheric Circulation and Pollution Transfer in the Environs of Norilsk Industrial Region // Atmospheric and Oceanic Optics. 2016. Vol. 29 (6). P. 565–569.

[4] *Solomos S., Amiridis V., Zanis P., Gerasopoulos E., Sofiou F.I., Herekakis T., Brioude J., Stohl A, Kahn R.A., Kontoes C.* Smoke dispersion modeling over complex terrain using high resolution meteorological data and satellite observations – The FireHub platform // Atmos. Environ. 2015. Vol. 119. P. 348–361.

[5] *Кондратьев К.Я., Григорьев Ал.А., Покровский О.М., Шалина Е.В.* Космическое дистанционное зондирование атмосферного аэрозоля. Л.: Гидрометеоиздат, 1983. 216 с.

[6] *Gao Z., Bresson R., Qu Y., Milliez M., Munck C., Carissimo B.* High resolution unsteady RANS simulation of wind, thermal effects and pollution dispersion for studying urban renewal scenarios in a neighborhood of Toulouse // Urban Climate. 2018. Vol. 23. P. 114–130.

[7] *Li H., Cui G., Zhang Z.* A New Scheme for the Simulation of Microscale Flow and Dispersion in Urban Areas by Coupling Large-Eddy Simulation with Mesoscale Models // Boundary-Layer Meteorol. 2018. Vol. 167. P. 145–170.

[8] *Fallah-Shorshani M., Shekarrizfard M., Hatzopoulou M.* Evaluation of regional and local atmospheric dispersion models for the analysis of traffic-related air pollution in urban areas // Atmos. Environ. 2017. Vol. 167. P. 270–282.

[9] *Бызова Н.Л., Гаргер Е.К., Иванов В.Н.* Экспериментальные исследования атмосферной диффузии и расчеты рассеяния примеси. Л.: Гидрометеоиздат, 1991. 278 с.

[10] *Mikhailuta S.V., Lezhenin A.A., Pitt A., Taseiko O.V.* Urban wind fields: Phenomena in transformation // Urban Climate. 2017. Vol. 19. P. 122–140.

[11] *Islam M.A., Roy G.D.* A mathematical model in locating an unknown emission source // Water, Air, and Solid Pollut. 2002. Vol. 136. N 1/4. P. 331–345.

[12] *Рапута В.Ф., Шлычков В.А., Леженин А.А., Романов А.Н., Ярославцева Т.В.* Численный анализ данных аэрозольных выпадений примесей от высотного источника // Оптика атмосферы и океана. 2014. Т. 27. № 8. С. 713–718.

[13] *Зилитинкевич С.С.* Атмосферная турбулентность и планетарные пограничные слои. М.: ФИЗМАТЛИТ, 2013. 252 с.

[14] Методы расчетов рассеивания выбросов вредных (загрязняющих) веществ в атмосферном воздухе // Утверждены приказом Минприроды России от 06.06.2017 № 273.

[15] *Рапута В.Ф., Леженин А.А, Ярославцева Т.В.* Оценка параметров выбросов Новосибирских ТЭЦ с использованием спутниковой информации // Интерэкспо Гео-Сибирь. 2018. Т. 1. № 4. С. 137-146.

# МЕТОДЫ АВТОМАТИЗИРОВАННОЙ ОБРАБОТКИ МНОГО- И ГИПЕРСПЕКТРАЛЬНЫХ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ В ИНТЕРЕСАХ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ ОЦЕНКИ СОСТОЯНИЯ ПРИРОДНЫХ И АНТРОПОГЕННЫХ ОБЪЕКТОВ

*Марков А.В., Григорьева О.В., Жуков Д.В., Иванец М.О.*
Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург

В работе рассматриваются разработанные в Военно-космической академии имени А.Ф. Можайского подходы к разработке методик и программ автоматизированной тематической обработки много- и гиперспектральных данных дистанционного зондирования Земли. Приводятся примеры их практической реализации в задачах классификации древесной растительности по породному составу, обнаружения гарей и загрязнения почв нефтепродуктами.
*Ключевые слова: дистанционное зондирование Земли, многоспектральные и гиперспектральные данные, тематическая обработка, спектральные признаки, оценка состояния природных и антропогенных объектов.*

**Введение.** Анализ отечественного и иностранного опыта в области дистанционного зондирования Земли (ДЗЗ) показывает, что на современном этапе развития проявляется устойчивый интерес к разработке и внедрению новых методов и средств регистрации изображений, в том числе с помощью много- и гиперспектральной (МГС) авиационной и космической аппаратуры наблюдения. Применение этих инновационных технологий позволяет при тематической обработке материалов аэрокосмической съемки использовать новую более информативную систему идентификационных признаков объектов, дающую возможность в автоматизированном режиме решать целый ряд задач социально-экономического, научного и оборонного назначения.

Вместе с тем практическому внедрению МГС технологий препятствуют сложности, связанные с необходимостью анализа больших потоков информации, и имеющиеся ограничения стандартных программ обработки данных ДЗЗ (ENVI, Erdas Imagine, ScanEx Image Processor и др.). При всем многообразии предоставляемых функций эти программы не содержат четких рекомендаций по действиям оператора при решении конкретных тематических задач. Многочисленность предлагаемых алгоритмов обработки в ряде случаев не облегчает, а затрудняет процесс дешифрирования, так как оператор зачастую не знает, какой именно алгоритм наиболее эффективен для распознавания искомого объекта или идентификации его состояния.

Отмеченные проблемы обусловливают актуальность создания специализированных, адаптированных к прикладным задачам, методик и программ обработки МГС данных.

**Общий подход к созданию методик тематической обработки МГС данных ДЗЗ.** Технологический цикл разработки методик и программ тематической обработки МГС данных ДЗЗ в общем случае предусматривает выполнение следующих основных этапов:

- анализ оптических свойств объектов и сред распространения излучения, рассматриваемых в исследуемой задаче, с целью выявления закономерностей формирования сигналов, регистрируемых от объектов интереса и сопутствующих фонов в различных участках электромагнитного спектра;
- проведение экспериментальных измерений для изучения или уточнения (в том числе проверки теоретических гипотез) спектральных характеристик (СХ) объектов и фонов;
- определение на основе анализа СХ информативных спектральных признаков, являющихся характерными для рассматриваемых объектов и фонов;
- выбор оптимальных (с точки зрения достижения наилучших показателей достоверности получаемых результатов и оперативности вычислительного процесса) методов классификации МГС данных;
- формирование общей методики тематической обработки, объединяющей все обоснованные частные методы, и создание компьютерной программы, реализующей эту методику;

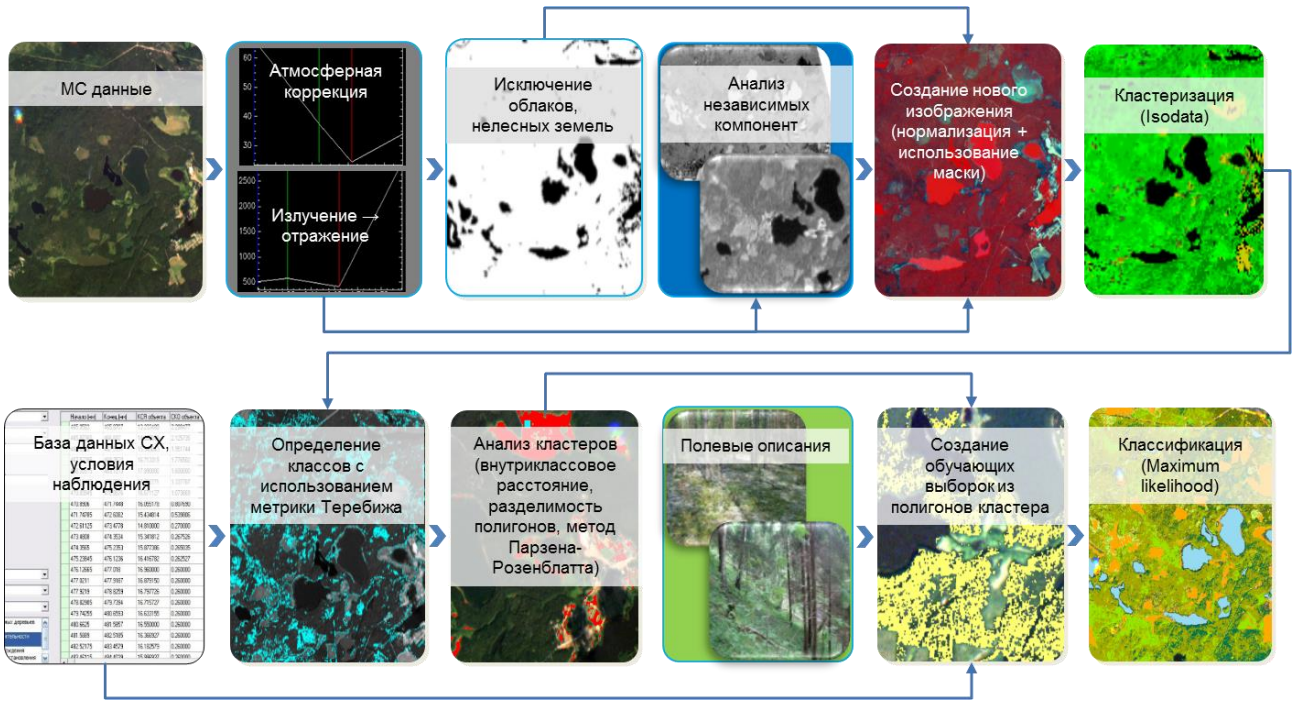



Рис. 1. Ансамблевый алгоритм обработки многоспектральных (МС) данных для решения задачи классификации породного состава лесных массивов.

– верификация разработанной методики по независимым (не использовавшимся при определении информативных спектральных признаков) данным авиационной или космической МГС съемки.

Спектральные признаки могут отражать как отличия в значениях яркости объекта интереса и фона (или отличия в значениях яркости разных классов объектов) в определенных участках спектра, так и характеризовать форму спектральной сигнатуры объекта интереса. В первом случае они традиционно выражаются либо в виде яркостных контрастов объекта и фона в фиксированной полосе спектра, либо в виде различных индексов (например, вегетационных), задаваемых математическими соотношениями между значениями яркости в нескольких информативных спектральных каналах. Во втором случае признаки несут информацию о локальных участках спектра объекта, на которых наблюдаются изменения СХ (экстремумы, перегибы и др.), являющиеся для объекта уникальными. Для расчета признаков формы используются градиентный метод, пространственно-масштабируемая фильтрация, секвентный анализ и др. [1].

В качестве методов автоматического анализа МГС данных могут применяться как параметрические (максимального правдоподобия, минимального расстояния, параллелепипеда, спектрального угла и др.), так и непараметрические (искусственные нейронные сети, SVM, Fuzzy и др.) методы классификации изображений, алгоритмы кластеризации и морфологической сегментации (методы высокочастотной фильтрации, пространственное дифференцирование, метод водораздела). В ряде случаев для достижения требуемого результата разрабатываются ансамблевые алгоритмы, включающие в себя комбинации различных методов. Например, для классификации лесных массивов по породному составу (рис. 1) рекомендуется алгоритм, который заключается в использовании методов параметрической классификации с построением обучающей выборки по результатам сегментации или неуправляемой кластеризации [2].

Для идентификации объектов по гиперспектральным (ГС) данным среднего и низкого пространственного разрешения применяются алгоритмы спектрального разделения смеси. В этом случае в зависимости от объекта исследования используются линейные или нелинейные модели формирования смешанного спектра пикселя. Для определения доли (относительной распространенности) объектов, составляющих пиксель, используются алгоритмы анализа

спектральных компонент (LSMA, линейного спектрального разложения и др.), учитывающие ограничения положительности и аддитивности относительных распространенностей. При этом набор эталонных спектров объектов может быть получен из специализированных баз данных (БД) СХ или инициализирован с помощью методов определения спектральных сигнатур, составляющих изображения: методов симплекса минимального объема, анализа зависимых спектральных компонент и т.п. [3].

В соответствии с описанным подходом в Военно-космической академии имени А.Ф. Можайского создана серия методик и компьютерных программ автоматизированной тематической обработки МГС данных в интересах решения прикладных задач оценки состояния природных и антропогенных объектов. Далее приводятся примеры выполненных разработок и кратко рассматриваются методы, положенные в их основу.

**Классификация древесной растительности по породному составу.** Разработанная программа определения кода древесной породы в процессе обработки МГС данных на первоначальном этапе производит разделение лесных и нелесных земель с созданием маски леса с помощью алгоритма релаксационной разметки, где в качестве признака используется вегетационный индекс NDVI. Это позволяет исключить влияние достаточно высоких значений метрических показателей в пространстве признаков при дальнейшей классификации оставшихся фрагментов.

Далее для классификации древесных пород используются два варианта обработки.

Первый алгоритм применяется при наличии обучающих выборок и основывается на методе параметрической классификации изображений, в котором при распознавании применяется байесовский подход на основе дискриминантных функций. При этом для ГС данных проводится предварительная обработка, направленная на оптимизацию спектральных каналов для минимизации ошибок идентификации объектов при использовании метода максимального правдоподобия. Оптимизация заключается в устранении корреляции между каналами путем преобразования фрагментов анализируемой сцены методом главных компонент, который позволяет определить весовой коэффициент каждого спектрального канала (элементы собственного вектора ковариационной матрицы), свидетельствующий о его информативности.

Второй вариант классификации породного состава предусматривает использование встроенной в программу БД СХ основных лесообразующих пород и их статистических оценок (коэффициентов вариации или среднеквадратических отклонений). Каждая спектральная сигнатура в БД сопровождается информацией о проективном покрытии (сомкнутости); фенологическом периоде наблюдения (предусмотрено 8 фенофаз); преобладающей породе или составе древостоя для смешанных лесов. Дополнительно учитывается тип леса, характеризующийся общностью лесорастительных условий (комплексом климатических, гидрологических и почвенных факторов), и бонитет (показатель роста и потенциальной производительности насаждений для данных условий местопроизрастания).

Алгоритм обработки в этом случае заключается в последовательном использовании метода классификации с построением обучающей выборки по результатам кластеризации. В качестве методов кластеризации в зависимости от типа данных используются методы на подобии Isodata или нечеткой кластеризации, а в качестве методов классификации – методы максимального правдоподобия или нейронных сетей. Обучающие выборки назначаются по кластерам, для этого оценивается компактность кластера путем анализа индексов внутриклассового расстояния и количества локальных максимумов функции плотности распределения, описываемой с помощью многомерной оценки Парзена-Розенблатта. При этом объем выборки определяется пространственным разрешением данных и количеством спектральных каналов, участвующих в классификации (например, для ГС данных КА «Ресурс-П» с разрешением 30 м объем выборки должен быть не менее 4×4 пикселя при числе спектральных каналов не более 16).

Установление соответствия между пикселями класса и породным составом осуществляется с использованием специальных критериев принадлежности, выраженных в виде вероят-

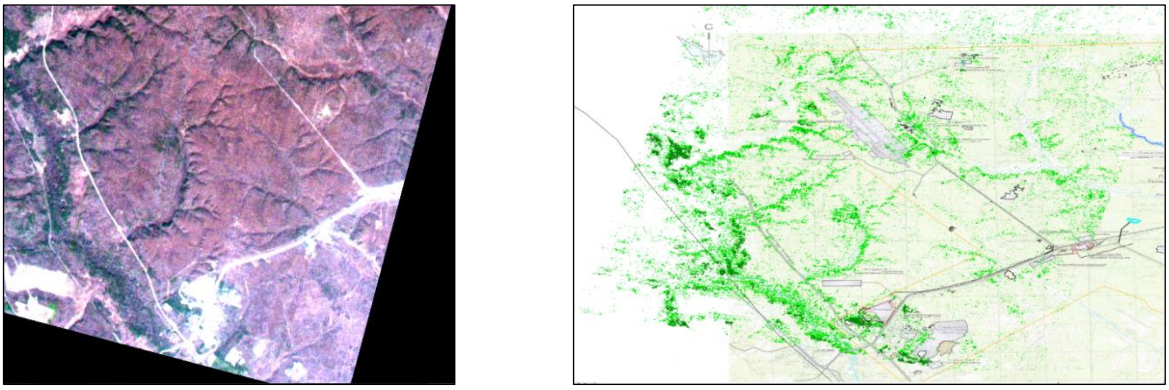
ности правильного отождествления спектров, специальных метрик (Махаланобиса и Теребижа) и коэффициента корреляции. Для ГС данных дополнительно применяется субпиксельный метод разделения смеси или проводится моделирование эталонной сигнатуры при наличии данных о сомкнутости крон и коэффициенте спектральной яркости (КСЯ) подстилающей поверхности, так как в данном случае в пиксель попадает несколько крон деревьев, а также лесная подстилка или верхний горизонт почв [4].

Для определения сомкнутости крон используются частотно-контрастные характеристики пространственного профиля яркости участка леса на изображении, которые определяются на основе использования прямого преобразования Фурье и метода построения автоковариационной функции (АКФ). Фазовая составляющая преобразования Фурье содержит важную информацию о взаимном расположении элементов на изображении и пространственной структуре изображения, которая значительно отличается для выделов с разной сомкнутостью и диаметром крон за счет наличия теней в межкроновом пространстве и наблюдения освещенных участков травянисто-кустарниковой растительности в редколесье. Показатели АКФ (максимум функции и диапазон частот, при которых АКФ становится меньше нуля) эффективно отражают расположение и размеры структурных элементов на изображении.

Каждой установленной в результате проведенной классификации древесной породе присваивается код (условное буквенное обозначение, например, Б – береза). В атрибутивные данные электронного слоя выдела записывается формула породного состава. Например, для лесного насаждения, в котором присутствует 40 % сосны и 60 % березы, береза признается преобладающей породой и формула состава имеет вид 4С6Б.

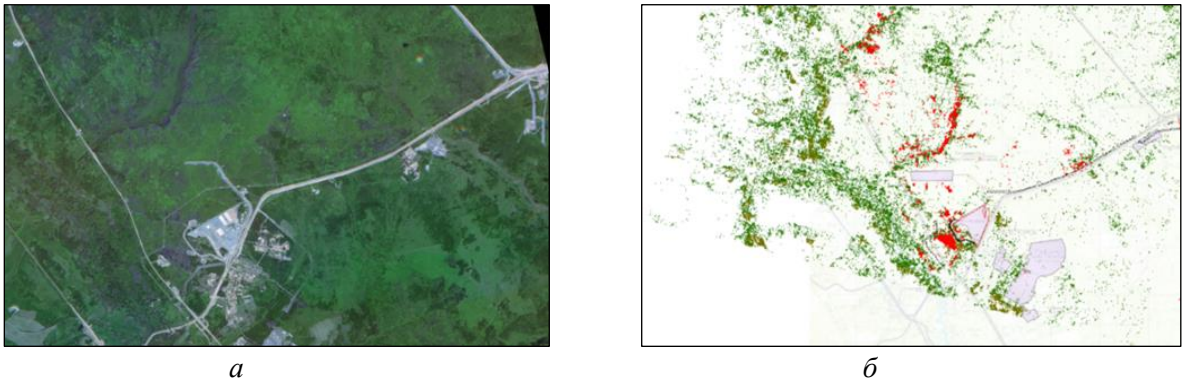
В соответствии с установленным породным составом выделов производится раскраска полигонов электронного слоя карты с использованием цветовой шкалы лесных страт, приведенной в «Методических рекомендациях по проведению государственной инвентаризации лесов», утвержденных приказом Рослесхоза от 10 ноября 2011 г. № 472.

Пример работы программы, реализующей описанный алгоритм, приведен на рис. 2.



*а* *б*

Рис. 2. Определение породного состава леса (хвойных и смешанных участков) по ГС данным:
*а* – исходное изображение; *б* – результат обработки.



*а* *б*

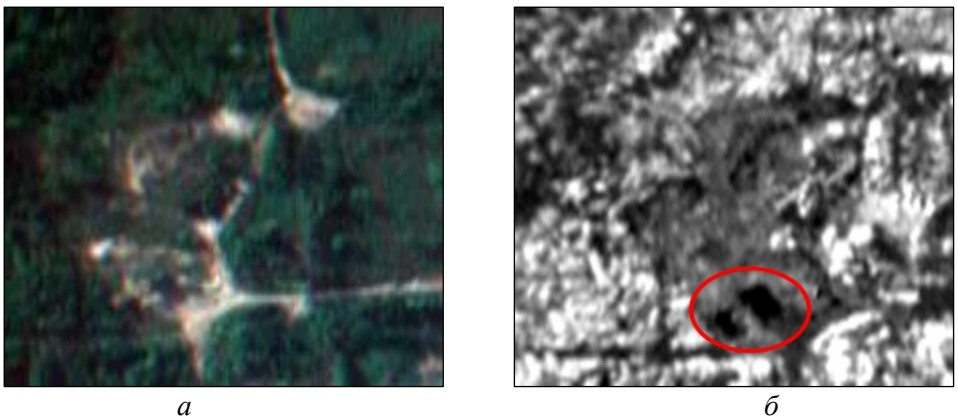
Рис. 3. Обнаружение гарей и горельников по МС данным:
*а* – исходное изображение; *б* – результат обработки.

**Обнаружение гарей.** Методика обнаружения гарей при обработке МГС данных основана на использовании вегетационных индексов. При этом состав индексов зависит от спектрального диапазона обрабатываемых МС или ГС изображений. При наличии данных в полосах спектра 1,55-1,75 или 2,10-2,35 мкм достаточно расчета только двух индексов NDVI и SWVI [5]. В случае, если правая граница спектрального диапазона не превышает 1,0 мкм, требуется использование трех вегетационных индексов NDVI, PSRI, TCHVI, отражающих изменения физиологических характеристик растительности в виде снижения концентрации хлорофилла [5]. Для ГС данных дополнительно рассчитываются индексы, характеризующие содержание влаги в растительном покрове (например, WBI). Кроме того, усложняется сам алгоритм обнаружения, который построен в виде дерева решающих правил и включает в себя пороговую обработку по указанным вегетационным индексам и значениям КСЯ в спектральных диапазонах 540-600; 620-680; 720-800 и 800-900 нм. Использование этих признаков позволяет отбраковать фрагменты (вырубки, болота, торфяники, редколесья), имеющие схожие с гарями спектральные характеристики, и повысить достоверность классификации. Пример применения предложенной методики приведен на рис. 3.

**Обнаружение загрязнений почв нефтепродуктами.** Методика обнаружения загрязнений почв нефтепродуктами (НП) предусматривает выполнение следующих действий [6].

1) Проводится неконтролируемая кластеризация трех индексных изображений NDVI методом нечетких множеств, где в качестве нормы разделения кластеров используется расстояние Евклида, разбивающее анализируемую область на три кластера. Индексные изображения NDVI рассчитываются в трех спектральных интервалах: 700-740, 720-800 и 800-900 нм.
2) Из зоны дальнейшего поиска по значениям вегетационного индекса исключаются кластеры, соответствующие растительному покрову.
3) Для оставшихся кластеров проводится повторная кластеризация методом Isodata по всем исходным МГС данным и выделение кластеров, соответствующих по своим спектральным признакам нефтезагрязненным почвам. В качестве критериев отнесения кластера к участку нефтезагрязнения используются следующие меры сходимости спектральных сигнатур: расстояние Евклида, учитывающее разницу средних значений КСЯ кластеров в спектральном канале, и коэффициент корреляции эталонных спектров НП и анализируемых спектров кластеров.
4) Осуществляется идентификация выделенных участков нефтезагрязнения по типу НП (битумизированный кир, жидкая фаза НП, нефтяная пленка и водонефтяная эмульсия) и степени загрязнения почв (полное разрушение, высокий, средний, низкий уровень) с использованием эталонных спектральных признаков почв и НП. При этом уровень загрязнения почв определяется по индексу нефтезагрязненности IS, рассчитываемому в ближнем инфракрасном и синем спектральных каналах.

На рис. 4 приведен пример обнаружения загрязнения почв нефтепродуктами с помощью разработанной методики.



*а* *б*

Рис. 4. Обнаружение нефтезагрязненных почв по ГС данным: *а* – исходное ГС изображение, синтезированное в естественных цветах 517, 569 и 677 нм; *б* – результат обработки.

**Заключение.** Представленные в статье примеры практической реализации методов автоматической обработки МГС данных ДЗЗ демонстрируют высокий потенциал использования в деятельности, связанной с природопользованием, информации об особенностях СХ природных и антропогенных объектов. В настоящее время круг тематических задач, решаемых с помощью МГС данных, постоянно расширяется. В частности, результаты проведенных в ВКА имени А.Ф. Можайского исследований показывают возможность их эффективного использования при идентификации загрязнений водных объектов [7], определении глубин прибрежных акваторий [8], оценивании степени проходимости местности [9], анализа состояния растительности [10] и др.

## ЛИТЕРАТУРА


[1] *Григорьева О.В., Марков А.В., Иванец М.О., Жуков Д.В.* Методы подготовки формализованных эталонных признаков для идентификации объектов по данным гиперспектральной космической съемки // Материалы V Всероссийской науч.-техн. конф. «Актуальные проблемы ракетно-космической техники» (V Козловские чтения). Самара: АО «РКЦ «Прогресс», изд. СамНЦ РАН, 2017. Т. 1. С. 281-286.

[2] *Григорьева О.В., Мочалов В.Ф., Рогачев С.А.* Экологический мониторинг особо охраняемых природных территорий на основе материалов съемки отечественных космических аппаратов // Сборник трудов II Всероссийской научной конференции «Экология и космос» имени академика К.Я. Кондратьева. Санкт-Петербург: ВКА имени А.Ф. Можайского, 2015. С. 104-109.

[3] *Григорьева О.В.* Субпиксельная идентификация объектов местности по гиперспектральным данным на основе последовательного квадратичного программирования и метода анализа формы спектральных компонент // Цифровая обработка сигналов. 2018. № 3. С. 26-32.

[4] *Марков А.В., Григорьева О.В., Чапурский Л.И., Мочалов В.Ф.* Методы подготовки эталонной информации для обработки материалов много- и гиперспектральной аэрокосмической съемки в задаче мониторинга природной среды // Сборник трудов III Всероссийской научной конференции «Проблемы военно-прикладной геофизики и контроля состояния природной среды». Санкт-Петербург, 2014. Т. 2. С. 525-534.

[5] http://www.indexdatabase.de

[6] *Григорьева О.В., Саидов А.Г., Панин А.В.* Индикатор состояния почво-грунтов в районах обращения нефти и нефтепродуктов, полученный неконтактными средствами наблюдений // Экология и промышленность России, октябрь 2010 г. С. 2-5

[7] *Григорьева О.В., Жуков Д.В., Марков А.В., Мочалов В.Ф.* Методика оценивания эколого-экономического ущерба водным объектам по данным дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса. 2018. Т. 15, № 1. С. 233-244.

[8] *Григорьева О.В., Жуков Д.В., Марков А.В., Мочалов В.Ф.* Восстановление глубин прибрежных акваторий по данным много- и гиперспектральной съемки // Оптика атмосферы и океана. 2016. Т. 29, № 7. С.553-559.

[9] *Саидов А.Г., Григорьева О.В., Марков А.В.* Использование геопространственных данных для оценки проходимости приграничных территорий // Материалы I межведомственной НПК «Специальные геоинформационные технологии и аэрокосмическое зондирование Земли». Калининград: Калининградский пограничный институт ФСБ России, 2016. С. 33-36.

[10] *Григорьева О.В.* Наблюдение деградации лесов по данным гиперспектрального аэро- и космического зондирования // Исследование Земли из космоса. 2014. № 1. С. 43-48.

# ПОИСК ИЗМЕНЕНИЙ ТЕМПЕРАТУРЫ ГОРОДСКОЙ СРЕДЫ ПО СПУТНИКОВЫМ ДАННЫМ НА ПРИМЕРЕ ГОРОДА КРАСНОЯРСКА

*Гостева А.А.[(1)], Матузко А.К.[(1)(2)], Якубайлик О.Э.[(1)(2)]*
[(1)] Сибирский федеральный университет, Институт космических и информационных технологий, Красноярск
[(2)] Институт вычислительного моделирования СО РАН, Красноярск

Изменение городской среды можно выявлять, используя космические снимки различного пространственного разрешения в видимом и дальнем инфракрасном диапазоне. Для изучения этих изменений наиболее доступными, полными и открытыми в настоящее время являются данные Landsat. Тепловая съёмка широко применяется для исследования и мониторинга антропогенных объектов, таких как трубопроводы, объекты городского хозяйства, промышленные объекты и загрязнения. В работе представлены результаты оценки температуры поверхности земли в городе Красноярске за двухлетний период с сентября 2016 года по сентябрь 2018 года на основе анализа спутниковых снимков Landsat-8 и Planet Scope.
*Ключевые слова: тепловые космические снимки, Landsat, Planet Scope, температура поверхности земли, климат городской среды.*

**Введение.** Микроклимат территории, в условиях естественного ландшафта, формируется под влиянием местных природных факторов: рельефа, близости к водоемам, характера подстилающей поверхности. В условиях искусственного ландшафта большое влияние оказывают антропогенные факторы, такие как: городская застройка, покрытие асфальтом растительного покрова и естественных почв, искусственные потоки тепла и загрязнение воздуха.

Влияние на экологическую обстановку в городе является важной особенностью изменения городского микроклимата, и как следствие изменения температуры поверхности. При устойчивой стратификации атмосферы, в особенности при инверсиях температуры, дым может накапливаться в приземном слое атмосферы в таком количестве, что оказывает вредное физиологическое воздействие. Автомобильный транспорт также вносит вклад в загрязнение воздуха городов являясь частой причиной смога [1].

Вследствие изменений городской среды происходит изменение микроклимата города, что влечет за собой непосредственное изменение температуры поверхности земли, оценить эти изменения можно как используя натурные измерения, так и дистанционные методы. В качестве дистанционные методов исследования температуры применяются данные космической съемки Земли в дальнем инфракрасном диапазоне. Такие снимки содержат информацию о температуре, которую практически невозможно получить каким-либо иным способом, например, с помощью снимков в видимом и ближнем инфракрасном диапазоне [2]. Тепловая съёмка широко применяется для исследования и мониторинга антропогенных объектов, таких как трубопроводы, объекты городского хозяйства, промышленные объекты и загрязнения [3].

В работе рассмотрены примеры изменений температуры поверхности земли в городе Красноярске за двухлетний период с сентября 2016 года по сентябрь 2018 года на основе анализа спутниковых снимков Landsat-8 и Planet Scope. Выбор данного временного периода связан с активной застройкой города, а также с началом территории города Красноярска группировкой Planet Scope. Применение спутниковых снимков высокого разрешения Planet Scope позволяет повысить разрешение тепловых карт, получаемых на основе тепловых снимков Landsat-8, что важно, так как тепловые снимки содержат информацию, которую практически невозможно получить каким-либо другим способом, например, с помощью снимков в видимом и ближнем инфракрасном диапазоне [4]. Представлена методика для выявления изменений температуры в рамках одной территории за определенный временной интервал и показана ее применимость для определения антропогенных изменений ландшафта.

**Увеличение точности пространственного разрешения температурных каналов.** В феврале 2013 года состоялся запуск современного спутника Landsat-8, данные представлены на онлайн-сервисе (https://www.earthexplorer.usgs.gov), который получает данные, используя два различных сенсора – Operational Land Imager (OLI) и Thermal Infrared Sensor (TIRS). Сканер теплового диапазона TIRS ведет съемку в двух каналах 10,6-11,2 мкм (10 канал) и 11,5-12,5 мкм (11 канала) с пространственным разрешением 100 метров [5].

Пространственное разрешение температурной карты, полученной по данным мультиспектральных каналов Landsat 8 составляет 30 метров. Для изучения городской территории существует необходимость повышения пространственного разрешения. Поэтому было принято решение использовать для создания классификации изображения снимки высокого разрешения PlanetScope (www.planet.com). Снимки высокого разрешения Planet Scope представлены в онлайн-сервисе (https://www.planet.com/products/explorer). Группировка Planet Scope насчитывает 120 спутников, равномерно распределённых по орбите, и предназначена для регулярного космического мониторинга любой территории в высоком пространственном разрешении. Четыре спектральных канала, с разрешением 3,7 метра – синий (0,450-0,515 мкм), зеленый (0,515-0,595 мкм), красный (0,605-0,695 мкм), ближний инфракрасный (0,740-0,900 мкм).

На следующем этапе происходит синтез космических изображений Landsat-8 и PlanetScope. Традиционно, используется подход, основанный на улучшении пространственного разрешения одного изображения за счет другого. Так, Landsat-8 позволяет получать цифровые изображения земной поверхности с пространственным разрешением 30 м в видимом диапазоне и 100 м в дальнем инфракрасном диапазоне, а Planet Scope получает изображения с пространственным разрешением 3,7 м в видимом диапазоне. Синтез этих изображений позволяет увеличить пространственное разрешение Landsat-8 до 3 метров и вычислить температуру поверхности земли по данным 10-го теплового канала Landsat-8 и классификации поверхности земли Planet Scope (рис. 1). Для каждой даты исследования необходимо составление своей карты температуры поверхности земли. Методика вычисления температуры поверхности земли подробно описана авторами в предыдущей работе [6].

**Методика выявления изменений городской среды по температуре поверхности земли.** Влияние большого современного города на формирование его микроклимата значительно, особенно это видно на отдельных кварталах, где изменяется городская застройка, вносятся существенные изменения в природный ландшафт.

Именно в результате изменений городской застройки, искусственного изменения рельефа, вырубки лесных насаждений часто возникают изменения микроклимата городской среды. Соответственно, происходит непосредственное изменение температуры поверхности города.

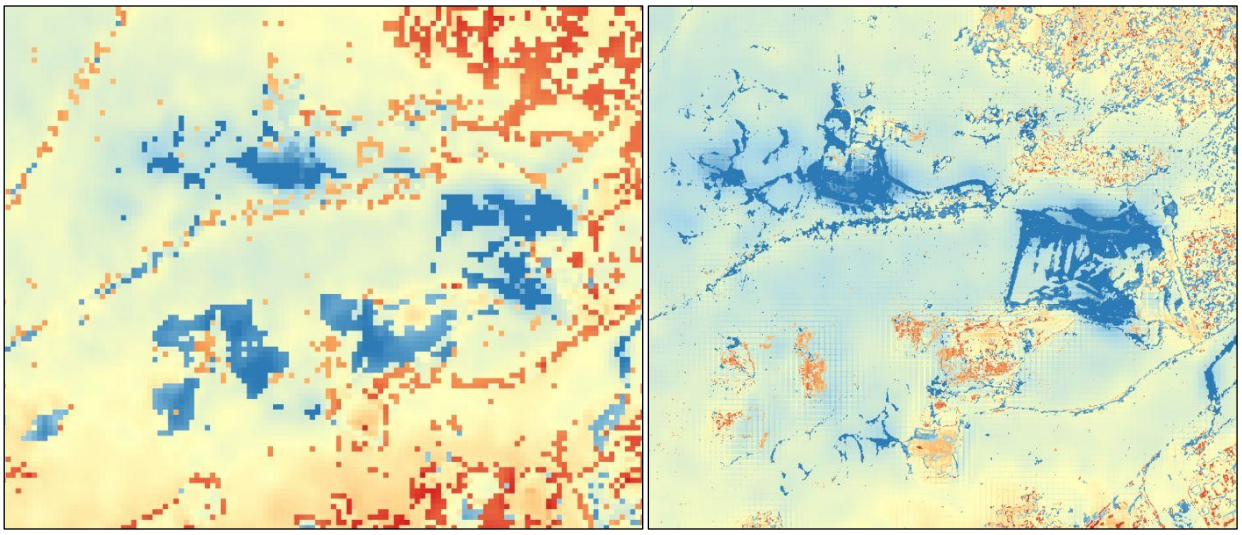


Рис. 1. Пример повышения детальности изображения за счет использования снимка высокого разрешения PlanetScope.

Таким образом, изменение температуры поверхности города является признаком изменений городской среды.

Методика выявления изменений городской среды основана на использовании температурных карт, полученных по спутниковым данным 10-го теплового канала Landsat-8 [7] и спутниковым данным Planet Scope видимого диапазона. При реализации выделяют следующие этапы:

Обязательным этапом является предварительная обработка спутниковых данных, которая включает атмосферную коррекцию исходных слоев.

Следующий шаг классификация изображения Planet Scope по типу подстилающей поверхности: вода, растительность, строения, открытая почва. В работе применяется классификация изображения с обучением.

Территория исследования ограничивается границами города Красноярска, поэтому далее необходимо проведение операции оверлея по границе города для каждого слоя.

Важным этапом является нормирование данных для возможности сравнения между собой полученных температурных карт. Используя статистические характеристики каждого изображения, а именно по среднестатистическому значению температуры карты, вычисляется добавочное значение для операции нормирования. После этого проводится математическая операция алгебры карт, позволяющая проводить вычитание одной карты из другой. В результате получается новое изображение содержащее разницу между температурными значениями в каждом пикселе.

Для корректного выявления изменений по двум спутниковым снимкам необходимо использовать безоблачные односезонные космические снимки с максимально равнозначными метеорологическими условиями. Для апробации методики выявления изменений городской среды авторами были подобраны спутниковые снимки на территорию города Красноярска с идентичными метеоусловиями на сентябрь 2016 года и сентябрь 2018 года. Набор спутниковых изображений представлен на даты 20.09.2016 и 05.09.2018 со спутника Landsat-8 и снимками высокого разрешения Planet Scope на 20.09.2016 и 20.09.2018. В таблице представлена температура воздуха в даты исследования и температура поверхности земли, полученная по космическим снимкам. Температура на снимке соответствует времени съемки Landsat-8 в 12 часов дня.

Рассмотрим подробнее проведение нормирования для выявления изменений температуры за два года. Все вычисления проводятся в границах города Красноярска. Сравнивая средние значения температуры поверхности земли 20.09.2016 (ТПЗ2016) и 05.09.2018 (ТПЗ2018), указанные в таблице, получаем, что значение в 2016 году больше, чем в 2018. Поэтому вычитаем среднее значение температуры 2018 из среднего в 2016 и получаем разницу, которую добавляем к карте ТПЗ2018, в результате получаем ТПЗ2018плюс. После нормирования, средние значения ТПЗ2016 и ТПЗ2018плюс стали равными. Далее проводится математическая операция алгебры карт из ТПЗ2018плюс вычитаем ТПЗ2016.

В результате вычитания получается новое изображение, содержащее разницу между температурными значениями в каждом пикселе для исследуемых дат, из значений 2018 года вычитаются 2016 года. Полученные значения для наглядности разделяются на 3 группы: увеличение температуры, незначительные изменения/без изменений и уменьшение температуры.

**Результаты.** Температурные изменения наблюдаются как в сторону увеличения температуры, так и в сторону её уменьшения. Разница температуры больше нуля говорит об увеличении температуры в 2018 году, относительно температуры в 2016, и наоборот, если разница температур меньше нуля, то произошло понижение температуры в 2018 году относительно 2016.

Сравнение значений температуры воздуха и полученной температуры.

| **Дата** | **Температура воздуха, °С** | | **Средняя температура на снимке, °С** |
|---|---|---|---|
| **05.09.2016** | Днем +19 | ночью +14 | 19.6880 |
| **20.09.2018** | Днем +15 | ночью +4 | 15.8686 |



Рис. 2. Изменение температуры поверхности земли возле новой «Платинум Арены».



Рис. 3. Изменение температуры поверхности земли на территории «Комбайнового завода» в г. Красноярске с 2016 по 2018 год.



Рис. 4. Изменение температуры поверхности земли в лесном массиве города Красноярска с 2016 года до 2018 года

На рис. 2, показано изменение температуры поверхности в районе строительства многофункциональный спортивно-зрелищный комплекс «Платинум Арена» с ледовой ареной на 7000 зрителей. Общая площадь объекта – 27 040 кв. м. В 2016 году строительство объекта только начиналось, а в 2018 уже было завершено. Красным цветом показаны участки с увеличением температуры поверхности земли, которые совпадают с контуром построенного СЗК «Платинум Арены».

На рис. 3 показана территория бывшего Комбайнового завода, где в 2017 году начался снос промышленных корпусов под строительство жилых домов. На изображении 2016 года до начала подготовки строительной площадки. На рисунке 2018 уже завершен снос зданий и идет подготовка строительной площадки. На третьем фрагменте карты изменений температуры поверхности земли за два года, красным цветом показано увеличение температуры поверхности земли, возникшее в результате сноса промышленных зданий.

В связи с проведением XXIX Всемирной зимней универсиады 2019 года в Красноярске, в лесных массивах в черте города проводилась активная стройка горнолыжных трасс, были вырублены большие площади древесных насаждений. На рис. 4 показано изменение температуры поверхности земли в зоне строительства объектов Универсиады на Сопке и в Академгородке города Красноярска, на рисунке 2016 году лес до начала строительства, а на следующем

представлен завершающий этап строительства в 2018 году. На рисунке синим цветом обозначено изменение температуры в сторону понижения, что соответствует контурам горнолыжных трасс, построенных для проведения соревнований в лесных массивах города во время проведения Универсиады.

**Выводы.** Метод визуального анализа по видимым спектральным каналам затрудняет оценку изменений, особенно на больших территориях. Выявление изменений по видимым каналам ограничивает качество анализа человеческим фактором. Так же визуальные изменения не дают понимания об их скрытых характеристиках.

В работе проведена апробация методики выявления изменений городской среды по температуре поверхности земли на основе спутниковых данных инфракрасного диапазона. В результате создана карта разницы температур за два года по двум датам исследования (05.09.2016 и 20.09.2018). Для анализа участков с изменением температуры поверхности города Красноярска все значения разделяются на 3 класса: увеличение температуры, незначительные изменения/без изменений и уменьшение температуры.

При анализе участков с максимальными изменениями температуры, отмечено, что это связано с изменением городской застройки, а именно со строительством новых микрорайонов, торговых центров. Так же изменение температуры поверхности связано с изменением лесных ландшафтов, например, в местах вырубки древесных насаждений. Изменения температуры поверхности земли за временной период соответствуют изменению городской застройки, лесных массивов, ландшафтов, которые влияют на формирование климата. Данная методика подходит для решения задач дистанционного мониторинга изменения городской территории, легка в организации и хорошо автоматизируется.

## ЛИТЕРАТУРА


[1] *Хромов С.П., Петросянц М.А.* Метеорология и климатология. М.: Изд-во Московск. ун-та., 2006. 528 с.

[2] *Merlin O., Duchemin B., Hagolle O., Jacob F., Coudert B., Chehbouni G., Dedieu G., Garatuza J., Kerr Y.* Disaggregation of MODIS surface temperature over an agricultural area using a time series of Formosat-2 images // Rem. Sens. of Env. 2010. Vol. 114. P 2500-12.

[3] *Windahl E., Beurs K.* An intercomparison of Landsat land surface temperature retrieval methods under variable atmospheric conditions using in situ skin temperature // Int. J Appl. Earth Obs. Geoinf. 2016. Vol. 51. P. 11-27.

[4] *Frey C.M. Parlow E.* Flux measurements in Cairo. Part 2: On the determination of the spatial radiation and energy balance using aster satellite data // Remote Sens. 2012. Vol. 4. P. 2635-2660.

[5] *Niclos R., Valiente J., Barbera M., Caselles V.* Land surface air temperature retrieval from EOS-MODIS images // IEEE Geosci. Rem. Sens. Let. 2014. Vol. 11. P. 1380-84.

[6] *Matuzko A.K., Yakubailik O.E.* Urban heat island effects over Krasnoyarsk obtained on the basis of Landsat 8 remote sensing data // IOP Conference Series: Earth and Environmental Science. 2018. Vol. 211, N 1. P. 012010.

[7] *Effat H., Hassan O.* Change detection of urban heat islands and some related parameters using multi-temporal Landsat images; a case study for Cairo city // Egypt Urban Climate. 2014. Vol. 10. P. 171-88.

[8] PlanetSope satellite imagery. **https://www.planet.com/products/explorer** (дата обращения 23.05.2019).

# ИНТЕГРАЛЬНОЕ АНАЛИТИЧЕСКОЕ ОЦЕНИВАНИЕ СОСТОЯНИЯ ПРИРОДНО-ТЕХНОГЕННОЙ БЕЗОПАСНОСТИ ТЕРРИТОРИЙ КРАСНОЯРСКОГО КРАЯ

*Метус А.М., Пенькова Т.Г.*
Институт вычислительного моделирования СО РАН, Красноярск

Природно-техногенная безопасность территории определяется совокупностью факторов окружающей среды и техносферы. В работе выполнено оценивание состояния природно-техногенной безопасности территорий Красноярского края на основе метода интегрального аналитического оценивания, который позволяет сформировать комплексный показатель безопасности территории и рассчитать интегральную оценку комплексного показателя, являющуюся количественным выражением уровня безопасности.
*Ключевые слова: природно-техногенная безопасность, интегральное аналитическое оценивание, комплексный показатель, управление территориальной безопасностью.*

**Введение.** Одним из важнейших факторов защиты населения и территорий от чрезвычайных ситуаций является стратегическое управление территориальной безопасностью [1], которое направлено на совершенствование деятельности подразделений, отвечающих за ликвидацию последствий ЧС, планирование мероприятий и разработку управляющих рекомендаций по уменьшению риска. Безусловной необходимостью в процессе стратегического управления является всесторонний анализ показателей безопасности территории, их временной и пространственной динамики, обеспечивающий информационно-аналитическую поддержку лиц, принимающих решения.

В условиях множественности сфер мониторинга и роста объема накапливаемых данных требуется применение методов комплексного оценивания безопасности территории. В настоящее время для оценивания состояния безопасности используется три основных подхода [2]. Вероятностный подход подразумевает построение математической модели риска. Методы данного типа используются для расчета индивидуальных, коллективных и социальных рисков и ориентированы, как правило, на конкретный производственный объект. Статистический подход позволяет формировать количественную оценку на основе анализа данных за определенный период наблюдения. Методы данного типа дают возможность объективно исследовать динамику изменений наблюдаемых параметров и формировать сводные показатели. Однако такие методы не могут быть применены для редко наблюдаемых событий и не позволяют получать интерпретации количественных оценок. Эвристический подход позволяет формировать качественные оценки, когда формальные методы слишком сложны, а исходная база данных недостаточна для получения однозначного аналитического решения. Однако применение методов данного типа без аналитической поддержки ведет к ошибкам субъективного характера. Таким образом, для оценивания природно-техногенной безопасности территории необходим комбинированный подход, позволяющий опираться на статистические наблюдения и учитывать экспертные знания о характере протекающих процессов.

В работе выполнен анализ состояния природно-техногенной безопасности территорий Красноярского края с использованием метода интегрального аналитического оценивания [3] на основе мониторинговых данных Главного управления МЧС России по Красноярскому краю. Метод обеспечивает формирование комплексного показателя на основе многомерного аналитического моделирования состояния окружающей среды и объектов техносферы и позволяет получить интегральную оценку природно-техногенной безопасности территорий.

**Метод интегрального аналитического оценивания.** Общий подход к анализу состояния природно-техногенной безопасности территорий, согласно используемому методу, предусматривает формирование стандарта безопасности территории, а затем – интегральное оценивание состояния объектов окружающей среды и техносферы на основе разработанного стандарта.

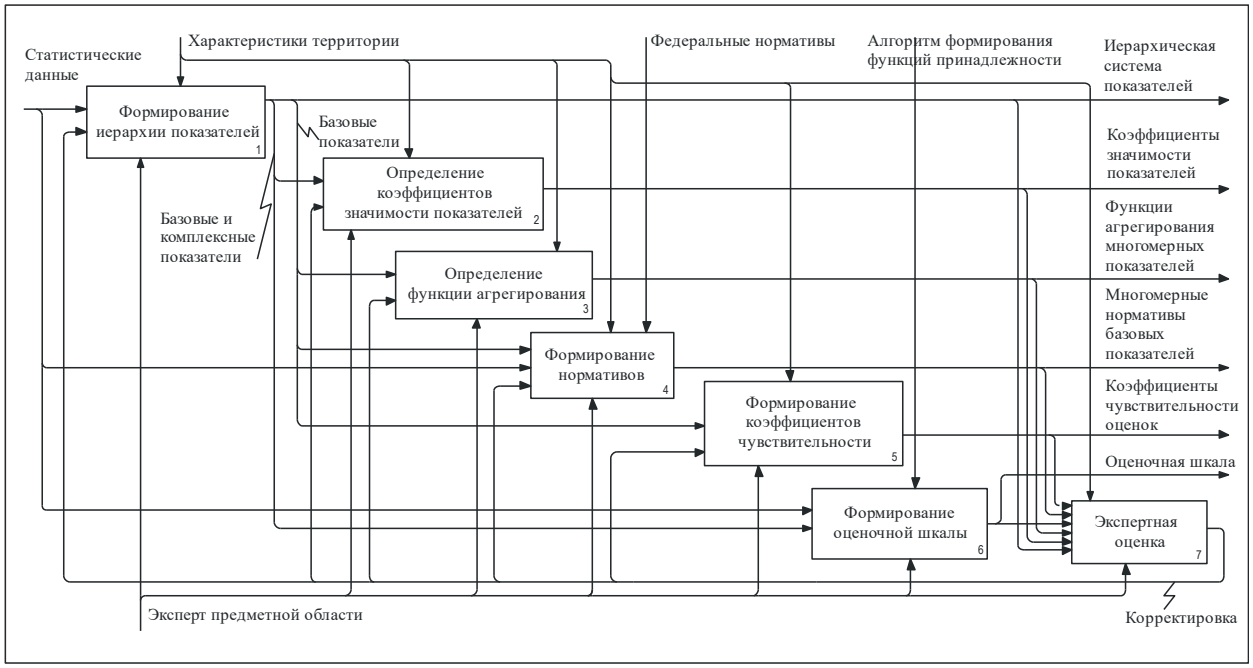



Рис. 1. Формирование стандарта безопасности территории.

Стандарт природно-техногенной безопасности территорий представляет собой территориально-ориентированную нормативную модель, множество переменных которой описывают «желаемый» уровень безопасности с учетом индивидуальных особенностей территории и реальных возможностей его достижения. На рис. 1 представлена функциональная диаграмма формирования стандарта природно-техногенной безопасности в нотации IDEF0.

На первом шаге формируется иерархическая система показателей – набор показателей, характеризующих природные и техногенные факторы риска возникновения чрезвычайных ситуаций [4]. Иерархия содержит два типа показателей: на нижнем уровне находятся базовые показатели, имеющие распределение по территории; на уровнях выше – комплексные, обобщенные по территории, показатели. На втором шаге определяются коэффициенты значимости показателей – весовые коэффициенты, характеризующие относительный вклад показателей нижнего уровня иерархии в показатели верхнего уровня. Коэффициенты значимости показателей определяются для каждой территории с учетом её физико-географических и социально-экономических особенностей. На третьем шаге определяются функции агрегирования оценок – функции, обеспечивающие переход от многомерных оценок показателей, рассчитанных по нескольким пунктам наблюдения, к одномерным значениям, рассчитанным по территории в целом. Функции агрегирования определяют, как отражается разброс оценок в пунктах наблюдений на оценке показателя территории. На четвертом шаге определяются нормативные значения показателей – интервалы значений, определяющих статистически нормальную величину показателей с учетом их распределения по территории. На пятом шаге определяются коэффициенты чувствительности оценок – коэффициенты, регулирующие скорость изменения оценки при отклонении фактического значения показателя от установленного норматива. На шестом шаге формируется оценочная шкала – отображение количественного значения оценки в ее качественное выражение с использованием методов нечеткой логики. Для этого определяется лингвистическая переменная «Уровень безопасности», принимающая множество значений: «Улучшенный», «Хороший», «Приемлемый», «Удовлетворительный», «Пониженный», «Низкий», «Критический». Каждому значению лингвистической переменной ставится в соответствие значение оценки показателя, которое со 100% вероятностью соответствует данному значению лингвистической переменной и которое служит для определения вероятности принадлежности любого другого значения оценки этому значению переменной. В завершение выполняется экспертиза, в ходе которой проводится оценка адекватности сформированных параметров стандарта и необходимая корректировка.

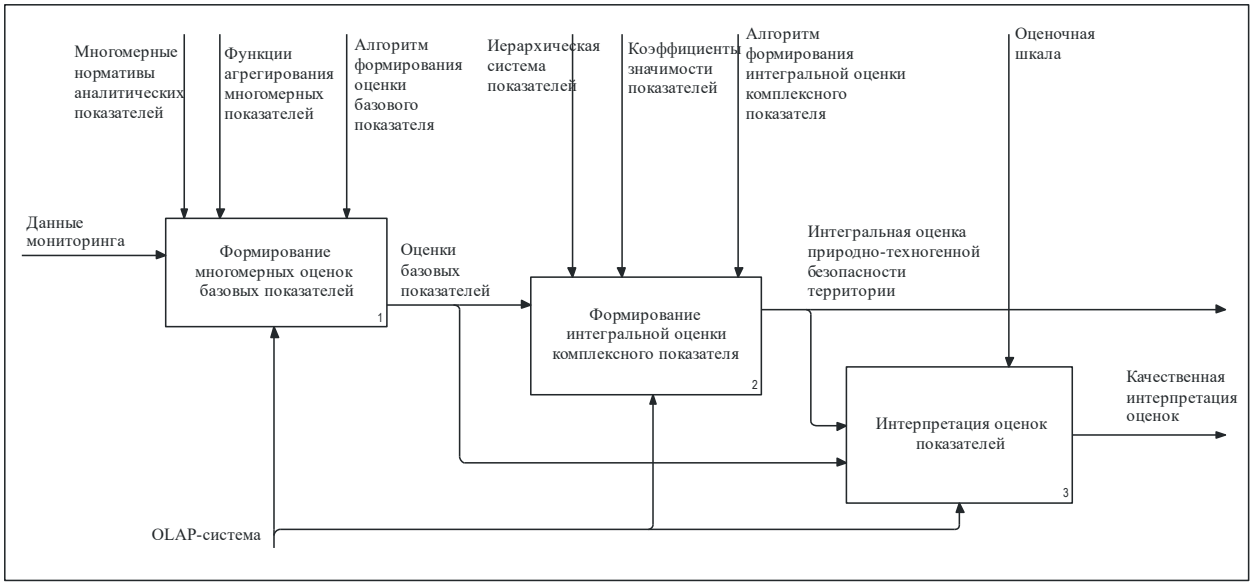



Рис. 2. Оценивание состояния природно-техногенной безопасности территории.

На втором этапе метода проводится интегральное оценивание состояния природно-техногенной безопасности территории путем последовательного расчета оценок уровней иерархии показателей с применением выработанного стандарта. На рис. 2 представлена функциональная диаграмма процессов оценивания состояния природно-техногенной безопасности территории.

На первом шаге рассчитываются оценки показателей нижнего уровня иерархии – базовых показателей. Вначале рассчитываются многомерные оценки характеризующие соответствие показателя нормативу в разрезе отдельных пунктов наблюдения. Многомерные оценки рассчитываются следующим образом:

$$i_j^k = 1 + \Delta P_j^k S_j^k$$

где $S_j^k = \pm 1$ – коэффициент, отражающий тенденцию $k$-го показателя в $j$-м пункте наблюдения, и принимает значение 1, если состояние безопасности улучшается при увеличении значения показателя, -1 в ином случае; $\Delta P_j^k$– коэффициент соответствия фактического значения $k$-го показателя нормативу в $j$-м пункте наблюдения. Коэффициент соответствия рассчитывается следующим образом:

$$\Delta P_j^k = \begin{cases} 0, \text{если } P_j^k \in \left[N_j^k, Z_j^k\right] \\ \left(\dfrac{P_j^k - Z_j^k}{Z_j^k - N_j^k}\right)^{q_k}, \text{если } P_j^k > Z_j^k \\ -\left(\dfrac{N_j^k - P_j^k}{Z_j^k - N_j^k}\right)^{q_k}, \text{если } P_j^k < N_j^k \end{cases}$$

где $q_k$ – коэффициент чувствительности оценки к отклонению $k$-го показателя от норматива, заданный в стандарте безопасности; $\left[N_j^k, Z_j^k\right]$– диапазон нормативных значений $k$-го показателя в $j$-м пункте наблюдения; $P_j^k$ – фактическое значение $k$-го показателя в $j$-м пункте наблюдения. Значение коэффициента $\Delta P_j^k$ в совокупности со значением коэффициента $S_j^k$ позволяют получить количественную оценку $i_j^k$ показателя. Значения оценки $i_j^k$, превышающие единицу, демонстрируют улучшение показателя.

Затем, проводится агрегация многомерных оценок базовых показателей по территории:

$$I_k = f_{agr}^k(i_1^k, \dots, i_m^k),$$

где $i_j^k, j \in \{1 \dots m\}$ – многомерная оценка $k$-го базового показателя в $j$-м пункте наблюдения; $f_{agr}^k$ – функция агрегирования для $k$-го базового показателя, заданная в стандарте безопасности.

На втором шаге рассчитывается интегральная оценка комплексного показателя на основе оценок базовых показателей с учетом их значимости в показателе верхнего уровня иерархии:

$$I_\Sigma = \sum_{k=1}^{n} u_k I_k$$

где $I_\Sigma$ – интегральная оценка комплексного показателя; $I_k$– оценка $k$-го базового, или промежуточного комплексного, показателя по территории; $u_k$– коэффициент значимости $k$-го базового показателя, заданный в стандарте безопасности.

На третьем шаге проводится интерпретация количественного значения оценок показателей с помощью шкал, заданных в стандарте безопасности, следующим образом:

$$\mu_i = \frac{1}{\sum_{k=1}^{C} \left( \frac{\|x - c_j\|}{\|x - c_k\|} \right)^2}$$

где $\mu_i$ – функция принадлежности $i$-го значения лингвистической переменной; $c_i$ – значение оценки показателя, которое со 100% вероятностью соответствует $i$-му значению лингвистической переменной; $c_k$ – значение оценки показателя, которое со 100% вероятностью соответствует $k$-му значению лингвистической переменной; $x$ – текущее значение оценки показателя; $C = 7$ – число значений лингвистической переменной. Значением функции принадлежности будет вероятность, с которой данное значение оценки показателя соответствует $i$-му значению лингвистической переменной.

Таким образом, метод интегрального аналитического оценивания позволяет сформировать комплексный показатель природно-техногенной безопасности, имеющий иерархичную структуру, рассчитать интегральную оценку комплексного показателя, являющуюся количественным выражением уровня безопасности и провести ее качественную интерпретацию.

**Интегральное аналитическое оценивание состояния безопасности территорий Красноярского края.** Территория Красноярского края характеризуется повышенным уровнем чрезвычайных ситуаций, во многом определяющимся социально-экономическими причинами, большим ресурсным потенциалом, географическим положением и климатическими условиями. По данным Государственного доклада Главного управления МЧС РФ по Красноярскому краю за 2017 год [5], на территории региона расположено большое количество потенциально опасных объектов техносферы как регионального, так и федерального значения, большое количество объектов жизнеобеспечения, включая котельные, водозаборы, очистные сооружения. Территория края занимает более 1 млн. кв. км. и расположена в семи климатических зонах. Крупные ЧС природного характера фиксируются практические ежегодно.

Для всех территорий Красноярского края сформирован стандарт природно-техногенной безопасности [6]. Часть стандарта безопасности приведена в табл. 1. Оценивание природно-техногенной безопасности территорий Красноярского края осуществляется в разрезе территорий по различным сферам мониторинга. В качестве примера рассмотрим формирование интегральной оценки комплексного показателя «Пожарная обстановка» городского округа г. Красноярска, в состав которого входят населенные пункты г. Красноярск и д. Песчанка.

На первом этапе рассчитываются оценки базовых показателей пожарной обстановки – определяются многомерные оценки (в разрезе двух пунктов наблюдения – г. Красноярск и д. Песчанка) и выполняется их агрегирование по всей территории – городскому округу г. Красноярска. Фактические значения показателя «Количество бытовых и производственных пожаров на 10000 населения» для г. Красноярска составляет $P_{1.4.1} = 3{,}721$, для д. Песчанка – $P_{1.4.1} = 1{,}346$. Поэтому, коэффициент соответствия для г. Красноярск рассчитывается по условию выхода за верхнюю границу норматива: $\Delta P_{1.4.1} = \left((3{,}721 - 2{,}52)/2{,}52\right)^1 = 0{,}48$; для д.

Песчанка коэффициент рассчитывается по условию выхода за нижнюю границу норматива: $\Delta P_{1.4.1} = -\big((1{,}882 - 1{,}346)/(7{,}54 - 1{,}882)\big)^1 = -0{,}09$. Далее, с учетом полярности показателя, оценка для г. Красноярска составляет $i_{1.4.1} = 1 + 0{,}48(-1) = 0{,}52$, для д. Песчанка – $i_{1.4.1} = 1 + (-0{,}09)(-1) = 1{,}09$. Аналогично, для показателя «Количество пожаров с погибшими на 10000 населения» оценки составляют: $i_{1.4.2} = 0{,}43$ для г. Красноярска и $i_{1.4.2} = 1{,}84$ для д. Песчанка, для показателя «Количество пожаров с пострадавшими на 10000 населения» – для $i_{1.4.3} = 0{,}44$ и $i_{1.4.3} = 2{,}38$ соответственно.

Агрегированные оценки базовых показателей рассчитываются исходя из наихудших значений, зарегистрированных в отдельных пунктах наблюдения. Так, оценка базового показателя «Количество бытовых и производственных пожаров на 10000 населения» составляет $I_{1.4.1} = \min(0{,}52; 1{,}09) = 0{,}52$. Оценки базовых показателей «Количество пожаров с погибшими на 10000 населения» и «Количество пожаров с пострадавшими на 10000 населения» составляют – $I_{1.4.2} = 0{,}43$ и $I_{1.4.3} = 0{,}44$ соответственно.

Табл. 1. Стандарта природно-техногенной безопасности территорий Красноярского края.

| Иерархия показателей | Коэффициент чувствительности, $q_k$ | Нижняя граница норматива, $N_k$ | Верхняя граница норматива, $Z_k$ | Коэффициент значимости, $u_k$ |
|---|---|---|---|---|
| | **г. Красноярск** | | | |
| 1. Техногенная безопасность | | | | 0,8 |
| 1.1 Обстановка на объектах техносферы | | | | 0,35 |
| 1.2 Транспортная обстановка | | | | 0,3 |
| 1.3 Обстановка на объектах ЖКХ | | | | 0,1 |
| 1.4 Пожарная обстановка | | | | 0,1 |
| Количество бытовых и производственных пожаров на 10000 населения,$\times 10^{-8}$ | 1,0 | 0 | 2,520 | 0,4 |
| Количество пожаров с погибшими на 10000 населения,$\times 10^{-8}$ | 0,3 | 0,012 | 0,036 | 0,3 |
| Количество пожаров с пострадавшими на 10000 населения,$\times 10^{-8}$ | 0,8 | 0,005 | 0,041 | 0,3 |
| 1.5 Радиационная обстановка | | | | 0,15 |
| 2.Природная безопасность | | | | 0,2 |
| 2.1 Метеорологическая обстановка | | | | 0,3 |
| 2.2 Гидрологическая обстановка | | | | 0,2 |
| 2.3 Геофизическая обстановка | | | | 0,3 |
| 2.4 Лесопожарная обстановка | | | | 0,2 |

Табл. 2. Оценочная шкала комплексного показателя природно-техногенной безопасности.

| | $l_i$ | $c_i$ | $\mu_i$ |
|---|---|---|---|
| 1 | Улучшенный | 0.9093267666288 | 0,986412206 |
| 2 | Хороший | 0.8136647929115 | 0,008926737 |
| 3 | Приемлемый | 0.7275432132811 | 0,002325299 |
| 4 | Удовлетворительный | 0.6561627709198 | 0,00114636 |

| | | | |
|---|---|---|---|
| 5 | Пониженный | 0.6003893396118 | 0,0 |
| 6 | Низкий | 0.4804927665189 | 0,0 |
| 7 | Критический | 0.1016071897181 | 0,0 |

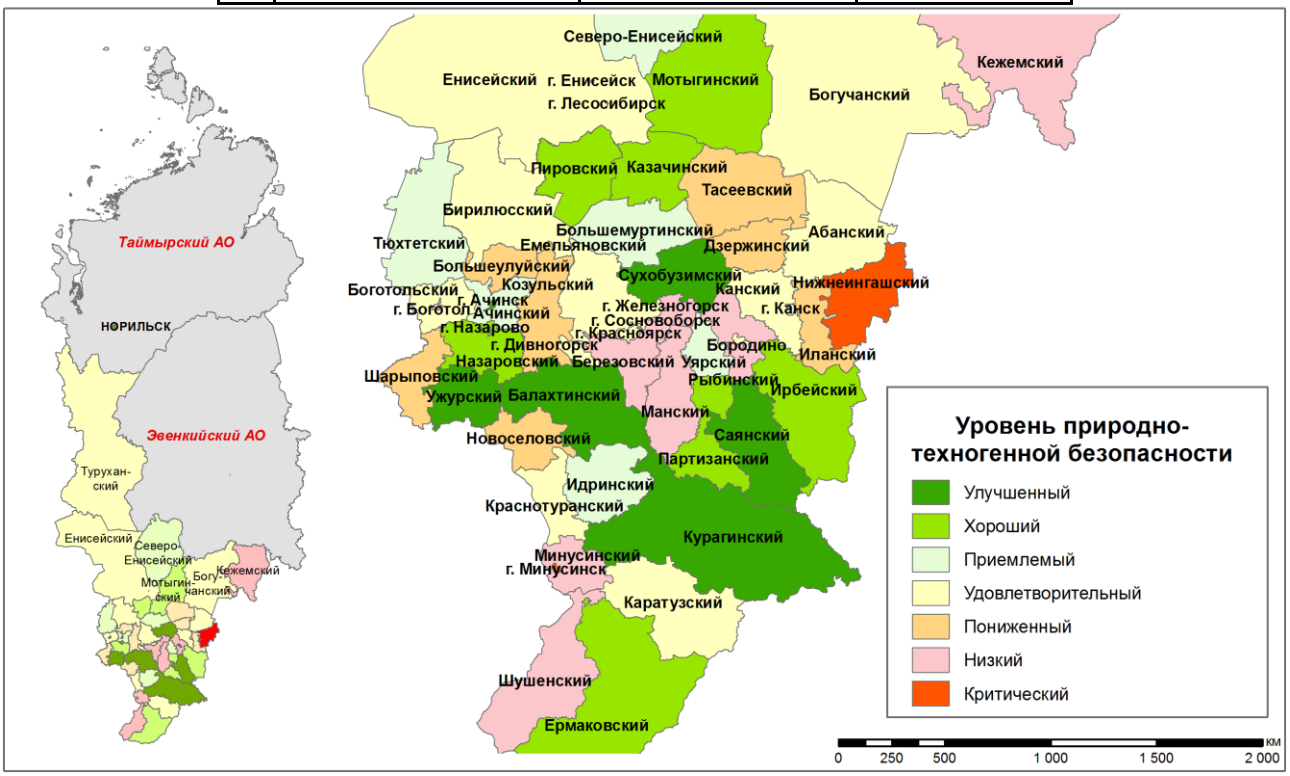



Рис. 3. Уровни природно-техногенной безопасности территорий Красноярского края.

На втором этапе рассчитывается интегральная оценка комплексного показателя «Пожарная обстановка» по городскому округу г. Красноярска на основе коэффициентов значимости входящих в него показателей. Оценка комплексного показателя «Пожарная обстановка» составляет $I_{\text{Пожар.}} = 0{,}4 * 0{,}52 + 0{,}3 * 0{,}43 + 0{,}3 * 0{,}44 = 0{,}47$. Аналогичным образом рассчитываются остальные комплексные показатели: «Обстановка на объектах ЖКХ» – $I_{\text{ЖКХ}} = 0{,}43$; «Транспортная обстановка» – $I_{\text{Трансп.}} = 1{,}0$; «Обстановка на объектах техносферы» – $I_{\text{Техносф.}} = 1{,}23$; «Радиационная обстановка» –$I_{\text{Радиац.}} = 1{,}0$; «Лесопожарная обстановка» – $I_{\text{Лесопожар.}} = 1{,}05$; «Геофизическая обстановка» –$I_{\text{Геофизич.}} = 1{,}0$; «Гидрологическая обстановка» – $I_{\text{Гидролог.}} = 1{,}0$; «Метеорологическая обстановка» –$I_{\text{Метеор.}} = 1{,}17$.

Комплексные показатели более высокого уровня рассчитываются на основе коэффициентов значимости комплексных показателей нижнего уровня: «Техногенная безопасность» – $I_{\text{Техног.}} = 0{,}47 * 0{,}1 + 0{,}43 * 0{,}1 + 1 * 0{,}3 + 1{,}23 * 0{,}35 + 1 * 0{,}15 = 0{,}97$; «Природная безопасность» – $I_{\text{Природ.}} = 1{,}05 * 0{,}2 + 1 * 0{,}3 + 1 * 0{,}2 + 1{,}17 * 0{,}3 = 1{,}06$; «Природно-техногенная безопасность» – $I_{\text{Природ.техног.}} = 0{,}97 * 0{,}8 + 1{,}06 * 0{,}2 = 0{,}99$.

Рассмотрим процесс интерпретации оценки на примере показателя «Природно-техногенная безопасность». Оценочная шкала для данного показателя представлена в табл. 2, где $l_i$ представляет значения лингвистической переменной «Уровень природно-техногенной безопасности», $c_i$ – значения показателя «Природно-техногенная безопасность», которые со 100% вероятностью соответствуют значению переменной, $\mu_i$ – значения функции принадлежности оценочной шкалы. Значение оценки г. Красноярска $I_{\text{Природ.техног.}} = 0{,}99$ характеризуется как улучшенный уровень природно-техногенной безопасности территории.

Аналогично выполнен расчет и интерпретация оценок иерархии показателей для остальных территорий Красноярского края. Пространственная динамика уровня природно-техногенной безопасности по территориям края показана на рис. 3.

Наиболее высокий уровень природно-техногенной безопасности демонстрируют такие районы как Сухобузимский, Ужурский, Саянский, а так же города Енисейск, Лесосибирск, Красноярск. К территориям с наиболее низким уровнем природно-техногенной безопасности относятся Нижнеингашский район и город Минусинск, что обусловлено неблагоприятными значениями показателя «Количество аварий на системах водоснабжения на 10000 чел. населения» в г. Минусинске и показателями «Количество аварий на системах электроснабжения на 10000 чел. населения», «Количество бытовых и производственных пожаров на 10000 чел. населения» в населенных пунктах Нижнеингашского района.

**Заключение.** Выполнен анализ состояния природно-техногенной безопасности территорий Красноярского края с использованием метода интегрального аналитического оценивания. Представлены этапы формирования стандарта безопасности, учитывающего индивидуальные особенности территории, приведен алгоритм расчета интегральной оценки состояния безопасности на основе сформированного стандарта, являющийся количественным выражением уровня безопасности, приведен процесс интерпретации интегральной оценки, позволяющий получить качественную интерпретацию интегральной оценки. Полученные оценки состояния территорий служат в качестве информационно-аналитической поддержки в процессе стратегического управления территорией.

## ЛИТЕРАТУРА


[1] *Ямалов И.У.* Моделирование процессов управления и принятия решений в условиях чрезвычайных ситуаций. М.: Лаборатория базовых знаний, 2013. 288 с.

[2] *Метус А.М.* Актуальные задачи комплексного оценивания природно-техногенной безопасности территории // Молодой ученый. 2015. № 11. С. 89-92.

[3] *Пенькова Т.Г., Метус А.М., Ничепорчук В.В.* Метод интегрального аналитического оценивания природно-техногенной безопасности территорий (на примере Красноярского края) // Проблемы анализа риска. 2018. Т.15, № 5. С. 16-25.

[4] *Ничепорчук В.В., Пенькова Т.Г.* Система аналитических показателей для стратегического контроля природно-техногенной безопасности территорий // Проблемы анализа риска. 2018. Т. 15, № 1. С. 70-77.

[5] Государственный доклад о состоянии защиты населения и территорий Красноярского края от чрезвычайных ситуаций природного и техногенного характера: Главное управление МЧС России по Красноярскому краю в 2017 году. Красноярск, 2018. 248 с.

[6] *Ничепорчук В.В., Пенькова Т.Г., Метус А.М.* Формирование стандарта природно-техногенной безопасности территорий Красноярского края // Проблемы безопасности и чрезвычайных ситуаций. 2018. № 2. С.41-52.

# ВЕРОЯТНОСТНЫЙ АНАЛИЗ СЕЙСМИЧЕСКОЙ ОПАСНОСТИ ПЛОЩАДОК ОСОБО ОТВЕТСТВЕННЫХ ОБЪЕКТОВ

*Миронов В.А.[(1)(2)], Перетокин С.А.[(3)], Симонов К.В.[(2)]*
[(1)] Институт вычислительного моделирования СО РАН, Красноярск
[(2)] Красноярский филиал Института вычислительных технологий СО РАН – Специальное конструкторско-технологическое бюро «Наука», Красноярск
[(3)] НП «Экологический центр рационального освоения природных ресурсов», Красноярск

Работа посвящена развитию методов вероятностного анализа сейсмической опасности, как одного из основных этапов в инженерно-сейсмологических изысканиях для площадок особо ответственных объектов, приведены обзор программных средств и примеры расчетов.
*Ключевые слова: ВАСО, анализ сейсмической опасности, землетрясение, сейсмические параметры опасности, программное обеспечения для ВАСО.*

**Введение.** Оценка сейсмической опасности площадки строительства – это неотъемлемая часть комплекса инженерно-геологических изысканий при проектировании ответственных объектов. В Российской федерации, как и в большинстве стран мира, нормативные оценки сейсмической опасности носят вероятностный характер. На картах общего сейсмического районирования (ОСР) РФ приведены расчетные воздействия с вероятностью их непревышения 90, 95 и 99% за 50 лет.

С 26.11.2018 г. в действие вступил свод правил СП 14.13330.2018 «Строительство в сейсмических районах». В соответствии с пунктом 4.3, для уточнения сейсмичности района строительства объектов повышенного уровня ответственности, необходимо проводить специализированные сейсмологические и сейсмотектонические исследования – детальное сейсмическое районирование (ДСР). ДСР, как и ОСР, включает в себя комплекс исследований, которые условно можно объединить в две группы: составление и параметризация моделей возникновения очагов землетрясений (ВОЗ) и подготовка уравнений прогнозирования движений грунта (УПДГ). Непосредственно расчет вероятных сейсмических воздействий на основе этих групп данных выполняется в процедурах вероятностного анализа сейсмической опасности (ВАСО). При этом корректность получаемых оценок существенно зависит от программных средств, используемых для расчета.

В данной работе проведен расчёт сейсмической опасности для Северомуйского тоннеля – железнодорожного тоннеля в Республике Бурятия на Байкало-Амурской магистрали протяженностью более 15 км. Расчет выполнен в специализированных программных комплексах, а именно, в российском программном комплексе EAST-2016 [1], где ВАСО реализуется на основе метода Монте-Карло, и в европейском программном комплексе OpenQuake Engine [2], где ВАСО реализуется на основе *классического* похода Корнелла.

**Основные подходы для вероятностного анализа сейсмической опасности.** Оценка опасности методом ВАСО состоит в определении уровня сейсмических воздействий на площадке, который не будет превышен с фиксированной вероятностью в течение заданного промежутка времени [3]. Сейсмическая опасность по ВАСО тесно связана с понятием сейсмической сотрясаемости, введенной Ю. В. Ризниченко в 1965 году. Как вычислительная процедура, подход ВАСО впервые был представлен Корнеллом (1968) для пикового ускорения грунта [3].

Вероятностный подход, который предложил Корнелл для оценки сейсмической опасности, в современной литературе трактуется как *классический* ВАСО. Согласно данному подходу, чтобы вычислить вероятность превышения заданной амплитуды движения грунта на исследуемой площадке, вклады опасности интегрируются по всем магнитудам и расстояниям, для всех зон ВОЗ, согласно теореме полной вероятности [4]. Классический подход для ВАСО включает четыре основных шага.

1. Идентификация и параметризация сейсмических источников. Источники могут быть представлены как площади, разломы, точки и др.
2. Характеристика временного и магнитудного распределения сейсмичности для источника. Повторяемость землетрясений во времени статистически независима, землетрясения происходят с постоянной частотой, а повторяемость будущих событий не зависит от последнего произошедшего землетрясения.
3. Подготовка УПДГ. Выбор УПДГ является важным моментом, поскольку часто УПДГ – это основной фактор вклада неопределенностей при выполнении ВАСО.
4. Реализация и учет различных видов неопределенностей. Например, в положении землетрясения в зоне ВОЗ, в магнитудах, в графиках повторяемости.

Конечным результатом ВАСО являются так называемые кривые опасности, определяемые в координатах параметра движения грунта и вероятности его превышения в заданный интервал времени.

В последнее время наблюдается тенденция к решению проблемы ВАСО с помощью моделирования методом Монте – Карло. Метод Монте – Карло также известен как *стохастический* метод, ранее использовался для численной оценки сейсмической опасности в восьмидесятых годах. Однако широкое распространение получил после работ Роджера Муссона. Процедуру ВАСО на основе метода Монте Карло можно условно разделить на два этапа.

1. На основе модели зон ВОЗ генерируется каталог землетрясений для заданного периода времени *T* (лет). Каждое землетрясение характеризуется набором параметров: магнитудой; длиной и шириной площадки очага; азимутом ее верхней кромки, задающим направление простирания очага; углом ее падения (углом между горизонтальной плоскостью и площадкой очага); географическими координатами и глубиной гипоцентра.
2. Вычисляется сейсмический эффект от каждого землетрясения в расчетной точке на основе принятой УПДГ и набирается статистика по количеству сотрясений разной интенсивности. Полученная статистика переводится в кумулятивный вид, из которого уже рассчитываются кривые опасности.

Использование весьма длительного синтетического каталога землетрясений позволяет учесть неопределенности в случайном характере параметров возможных землетрясений, что позволяет существенно упрощать логическое дерево в части учета неопределенностей параметров зон ВОЗ.

**Обзор современных программных средств ВАСО.** Работа по оценке сейсмической опасности и современному сейсмическому районированию не может быть выполнена без адекватного программно-математического обеспечения (ПМО), совершенствуемого по мере возникновения новых задач. В настоящее время в мире активно разрабатываются и широко используются компьютерные программы для проведения ВАСО на основе классического подхода, такие как, комплекс программных модулей OpenSHA, комплекс программных модулей R-CRISIS, расчетный модуль OpenQuake Engine.

Расчетный модуль OpenQuake Engine, используемый в рамках данной работы, разрабатывался на основе программных модулей OpenSHA (Университет Южной Калифорнии (USC) [5]) и реализован на языке программирования Python [2]. Поводом к разработке послужила реализация европейской научной программы Seismic Hazard Harmonization in Europe (SHARE). Основная цель SHARE заключалась в предоставлении модели сейсмической опасности для европейско-средиземноморского региона. Проект был направлен на установление новых стандартов в практике вероятностной оценки сейсмической опасности путем тесного сотрудничества ведущих европейских геологов, сейсмологов и инженеров [6].

Российский пакет программ PRB-60 разработан в 1994-1995 гг. в рамках исследований по созданию комплекта карт ОСР-97. Реализованная в нем процедура ВАСО базируется на методе Монте Карло. Методические основы пакета и базовые алгоритмы были разработаны А.А. Гусевым с участием Л.С. Шумилиной и В.М. Павлова, программная реализация пакета осуществлена В.М. Павловым. Актуализация PRB-60 выполнялась в 2003, 2010, 2016 гг. Последняя модификация этого ПО выполнена в программной среде Microsoft Visual Studio 2008

и получила наименование EAST-2016 [1]. Основные характеристики OpenQuake Engine и EAST-2016 представлены в таблице [1].

**Тестовый расчет сейсмической опасности для района Северомуйского тоннеля.** Линеаментно-доменная модель зон возникновения очагов землетрясений (зоны ВОЗ) содержит два основных структурных элемента. Линеаменты (линейные источники или плоскости, падающие под заданным углом) служат основным каркасом ЛДФ-модели. По сути, линеаменты – это представленные в генерализованном виде сейсмически активные разломы, которые несут основную часть сейсмического потенциала. Домены (площадные источники) охватывают объемы геологической среды, внутри которых с равной плотностью расположены (рассеяны) очаги землетрясений умеренных и невысоких магнитуд.

Для демонстрации возможностей программных комплексов ВАСО была подготовлена тестовая модель зон ВОЗ Байкальской рифтовой зоны. В качестве основы для построения линеаментных структур была принята база данных плиоцен-четвертичных разломов байкальской рифовой зоны [7], разработанная ИЗК СО РАН г. Иркутска. Визуализация элементов модели зон ВОЗ в районе Северомуйского тоннеля приведена на рис. 1.

Основные характеристики используемых программных комплексов.

| **Характеристики, методы** | | **Программное обеспечение** | |
|---|---|---|---|
| | | **OQ Engine** | **EAST-2016** |
| Подход к вероятностной оценке сейсмической опасности | | Классический | Монте-Карло |
| **Тип сейсмического источника** | | | |
| Площадной | | да | да |
| Разлом | Линейный (2D) | да | да |
| | Объемный (3D) | да | да |
| Точечный | | да | да |
| Возможность задавать распределение глубины по каждому источнику | | да | да |
| Возможность задавать механизм очага по каждому источнику | | да | да |
| Моделирование длины и ширины разрыва | | да | да |
| **Распределение повторяемости магнитуд (сейсмический режим)** | | | |
| Распределение Гутенберга-Рихтера | | да | да |
| Распределение Гаусса | | да | да |
| Настраиваемое интервальное распределение | | да | да |
| **Уравнения прогнозирования движения грунта (УПДГ)** | | | |
| Встроенные | | да | да |
| Задаваемые пользователем | | да | да |
| Возможность задавать различные уравнения для отдельного (одного) типа сейсмического источника | | да | да |
| Возможность задавать параметры уравнения для различных грунтов | | да | да |
| **Логическое дерево** | | | |
| Возможность установки логического дерева | | да | нет |
| **Выходные данные** | | | |
| Кривые опасности (график вероятности превышения интенсивности) | | да | да |
| Расчет длительности сотрясении | | нет | да |
| Карты опасности (карты значения интенсивности (какой-либо величины) с заданной вероятностью превышения) | | да | да |
| Однородные спектры опасности | | да | да |

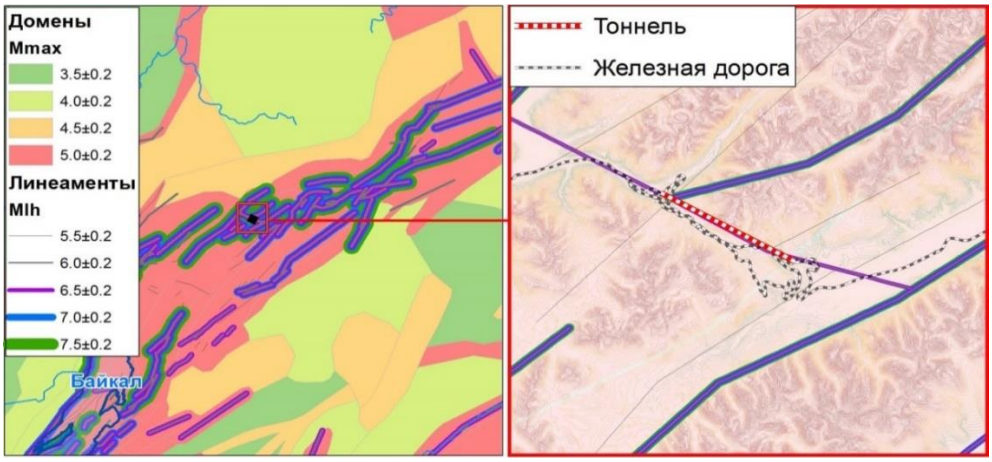


Рис. 1. Элементы модели зон ВОЗ в районе Северомуйского тоннеля.

Для каждой зоны ВОЗ, на основании региональных каталогов землетрясений, задано магнитудно-частотное распределение землетрясений (т.е. число ожидаемых событий в год для магнитудных интервалов).

Для выполнения ВАСО в отечественном ПО использовалось УПДГ Ф.Ф. Актикаева [8]. Для расчета ВАСО в OpenQuake Engine использовались с равными весами УПДГ: Campbell and Bozorgnia 2014, Chiou and Youngs 2014, Abrahamson et al. 2014. Расчет ВАСО в используемых ПО выполнен в привязке к грунтам I категории по сейсмическим свойствам (СП 14.13330.2018). В результате расчета ВАСО для Северомуйского тоннеля были получены спектры реакции колебаний грунта в численных характеристиках – спектральных ускорениях грунта. На рис. 2 представлены расчетные спектры реакции в привязке к скальным грунтам для периода повторяемости воздействий 1000 лет (вероятность не превышения 95% в течение 50 лет), рассчитанные с помощью OpenQuake Engine и EAST-2016.

Пиковое ускорение грунта ($PGA_{1000}$) составила 0,545 и 0,408 единиц g, максимальные значения спектра реакции на периодах колебаний 0,125 сек. и 0,4 сек. по OpenQuake Engine и EAST-2016 соответственно. Дополнительно в EAST-2016 рассчитана длительность сотрясений по Ф.Ф. Аптикаеву. Для пересчета PGA в макросейсмическую интенсивность использованы соотношения ГОСТ Р 57546-2017. На рисунке 3 показаны кривые опасности в долях балла, полученные в результате ВАСО по EAST-2016 и OpenQuake Engine.

Из рис. 3 видно, что оценки, полученные по обеим программам, в целом согласуются друг с другом. Максимальное отклонение на заданных периодах соответствует 0,2 доли балла. Следует отметить, что проведенные расчеты ВАСО для площадки Северомуйского тоннеля носят тестовый характер и предназначены для демонстрации возможностей программных средств, поэтому полученные оценки не следует рассматривать как абсолютно достоверные.

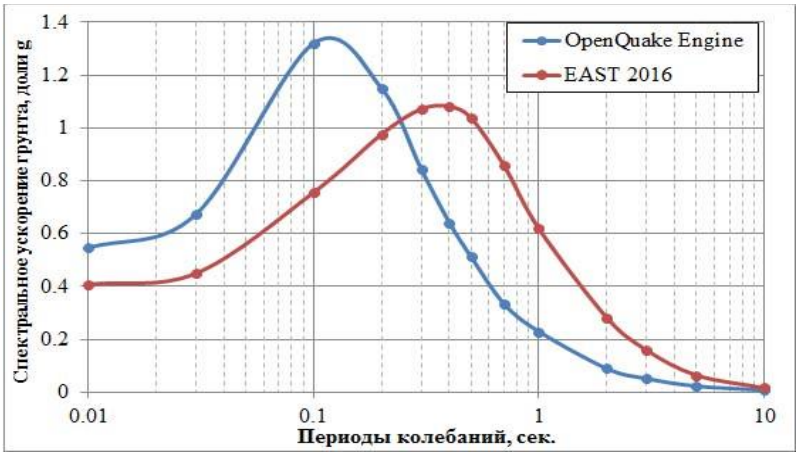


Рис. 2. Расчетные спектры реакции в привязке к скальным грунтам для периода повторяемости воздействий 1000 лет.

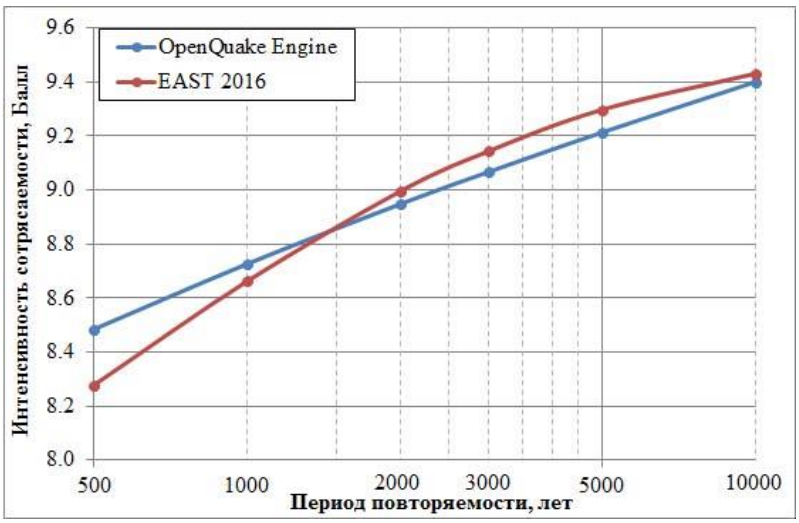


Рис. 3. Кривые опасности в долях балла.

**Заключение.** Поскольку получение корректных оценок сейсмической опасности территорий под строительство и эксплуатацию объектов разной степени ответственности является актуальной задачей, выбор программы и алгоритмов, на основе которых выполняется ВАСО, является очень важным моментом в исследованиях. Можно утверждать, что выбор инструмента ВАСО сказывается на корректности и надежности полученных результатов.

Сравнивая характеристики и полученные результаты двух программ, можно сделать вывод, что EAST-2016 обеспечивает надежную оценку сейсмической опасности и не уступает по функциональным возможностям современному зарубежному ПО OpenQuake Engine. При этом программный комплекс EAST-2016 ориентирован на отечественные нормативные документы (ГОСТ Р 57546-2017, СП 14.13330.2018), что делает его использование предпочтительнее при выполнении.

## ЛИТЕРАТУРА

[1] *Завьялов А.Д. Перетокин С.А. Данилова Т.И. Медведева Н.С. Акатова К.Н.* Общее сейсмическое районирование – от карт ОСР-97 к картам ОСР-2016 и картам нового поколения в параметрах физических характеристик // Вопросы инженерной сейсмологии. 2018. Т. 45, № 4. С. 47-68.

[2] OpenQuake Engine – Program for computing seismic hazard. **https://www.globalquakemodel.org/openquake** (дата обращения 01.06.2019).

[3] *Gupta I.D.* Probabilistic seismic hazard analyses method for mapping of spectral amplitudes and other design-specific quantities to estimate the earthquake effects on manmade structures // ISET Journal of Earthquake Technology. 2007. Vol. 44, N 1. Paper No. 480. P. 127-167.

[4] *Atkinson G.M.* The Integration of Emerging Trends in Engineering Seismology // 13th World Conference on Earthquake Engineering. Lisbon, Portugal, 2012.

[5] OpenSHA – Program for computing seismic hazard. **http://www.opensha.org** (дата обращения 01.06.2019).

[6] SHARE – European project for probabilistic seismic hazard. **http://www.share-eu.org** (дата обращения 01.06.2019).

[7] *Лунина О.В.* Цифровая карта разломов для плиоцен-четвертичного этапа развития земной коры Юга Восточной Сибири и сопредельной территории Северной Монголии // Геодинамика и тектонофизика. 2016. Т. 7, № 3. С. 407-434.

[8] *Аптикаев Ф.Ф.* Инструментальная шкала сейсмической интенсивности. М.: Наука и образование, 2012. 176 с.

# МЕТАН В АТМОСФЕРЕ ЗАПАДНОЙ СИБИРИ: РЕЗУЛЬТАТЫ СПУТНИКОВЫХ НАБЛЮДЕНИЙ И МОДЕЛИРОВАНИЯ

*Мордвин Е.Ю.[(1)], Лагутин А.А.[(1)(2)], Волков Н.В.[(1)(2)], Макушев К.М.[(1)], Прилипкова А.С.[(1)]*
[(1)] Алтайский государственный университет, Барнаул
[(2)] Институт вычислительных технологий СО РАН, Новосибирск

В работе исследуется поведение полного содержания метана в атмосфере Западной Сибири по данным спутниковых наблюдений и математического моделирования. По данным AIRS/AQUA установлен рост полного содержания метана в 2003-2018 гг. при тренде ~3.3 ±0.2 млрд$^{-1}$/год. С использованием глобальной химической транспортной модели MOZART4 получены оценки влияния удаленных источников на полное содержание метана в регионе. По данным климатических моделей установлено среднее значение эмиссии метана болотными комплексами Западной Сибири в период 2000-2013 гг, получены прогностические оценки эмиссии до 2050 г.
*Ключевые слова: атмосферный метан, Западная Сибирь, AIRS/AQUA, удаленные источники, болотные комплексы.*

**Введение.** Атмосферный метан является одним из наиболее важных парниковых газов, вносящий существенный вклад в изменение климата Земли. По данным Межправительственной группы экспертов по изменению климата 2013 года его вклад в радиационный форсинг составляет ~17% [1]. За последние 270 лет отношение смеси метана в приземном слое фоновой атмосферы возросло примерно на 257% и составляет на начало 2018 г. порядка 1860 млрд$^{-1}$ [2]. Особенно быстро рост метана происходил в 2014-2017 гг. [3].

В связи с этим актуальным является задача мониторинга содержания метана в атмосфере, а также выполнение прогностических оценок эмиссии метана.

Целью работы является исследование поведения полного содержания метана в атмосфере Западной Сибири по данным спутниковых наблюдений, а также установление трендов эмиссии метана до 2050 г. с использованием результатов климатических моделей.

**Содержание метана в атмосфере Западной Сибири.** Оценка полного содержания метана [$CH_4$] в атмосфере Западной Сибири проведена с использованием исследовательского продукта уровня 2 гиперспектрометра AIRS/AQUA, а также регрессионного метода, разработанного авторами ранее [4]. Для оптимизации статистической обработки использовалась база данных «Содержание метана в свободной тропосфере Западной Сибири», созданная в АлтГУ [5]. Анализ результатов проводился в двух зонах исследуемого региона: 1 – (65-55º с.ш., 60-90º в.д.); 2 – (55-45º с.ш., 60-90º в.д.). Северная зона содержит основной естественный источник метана региона – Большое Васюганское болото.

На рис. 1 показано полное содержание метана в двух зонах Западной Сибири, полученное с использованием регрессионной модели для всего рассматриваемого периода 2003-2018 гг. Точки на рисунке показывают усредненные по каждой зоне значения содержания метана, а линия - скользящее среднее за 30 дней. Видно, что в годовом цикле [$CH_4$] в атмосфере Западной Сибири присутствуют два максимума: зимний (январь-февраль) и летний (июль-сентябрь).

Анализ аномалий полного содержания метана, полученных в рамках подхода [6] и показанных на рис. 2, позволяет установить, что рост [$CH_4$] в 2003-2018 гг. проходил при тренде ~3.3 ±0.2млрд$^{-1}$/год.

**Влияние удаленных источников.** Оценка чувствительности полного содержания метана в атмосфере Западной Сибири к удаленным источникам выполнялась с использованием химической транспортной модели MOZART-4 [7]. В зоне расположения предполагаемого источника искусственно в 10 раз увеличивалась приземная концентрация метана. Запуск модели MOZART-4 проводился для модельного периода в 3 месяца, продолжительность эмиссии равнялась одному месяцу. Термин «чувствительность» в проведенном исследовании обозначает

отношение результатов моделирования с включенным удаленным источником и без него. В качестве зон расположения удаленных источников были выбраны регионы в западной Европе, юго-восточной Азии, на восточном побережье Северной Америки, на Аравийском полуострове и на севере Южной Америки. Анализ содержания метана при распространении от удаленных источников выполнялся на высотах атмосферы 300, 500 и 700 гПа.

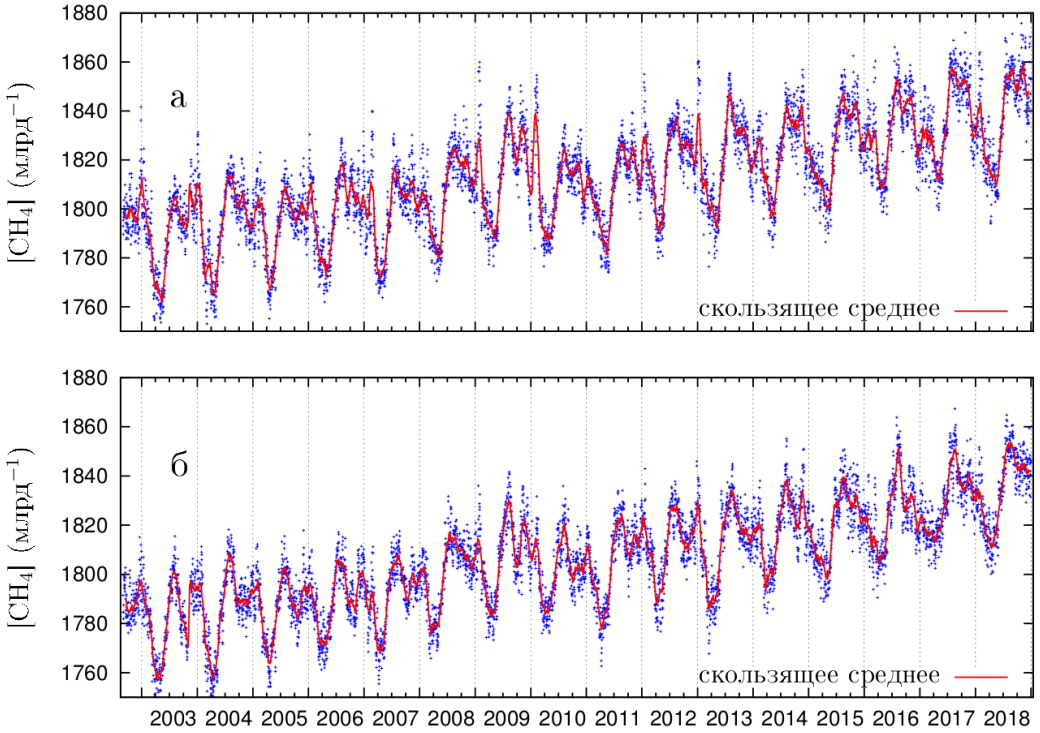



Рис. 1. Годовой ход и межгодовая изменчивость полного содержания метана в атмосфере Западной Сибири для зон 1 (*а*) и 2 (*б*).

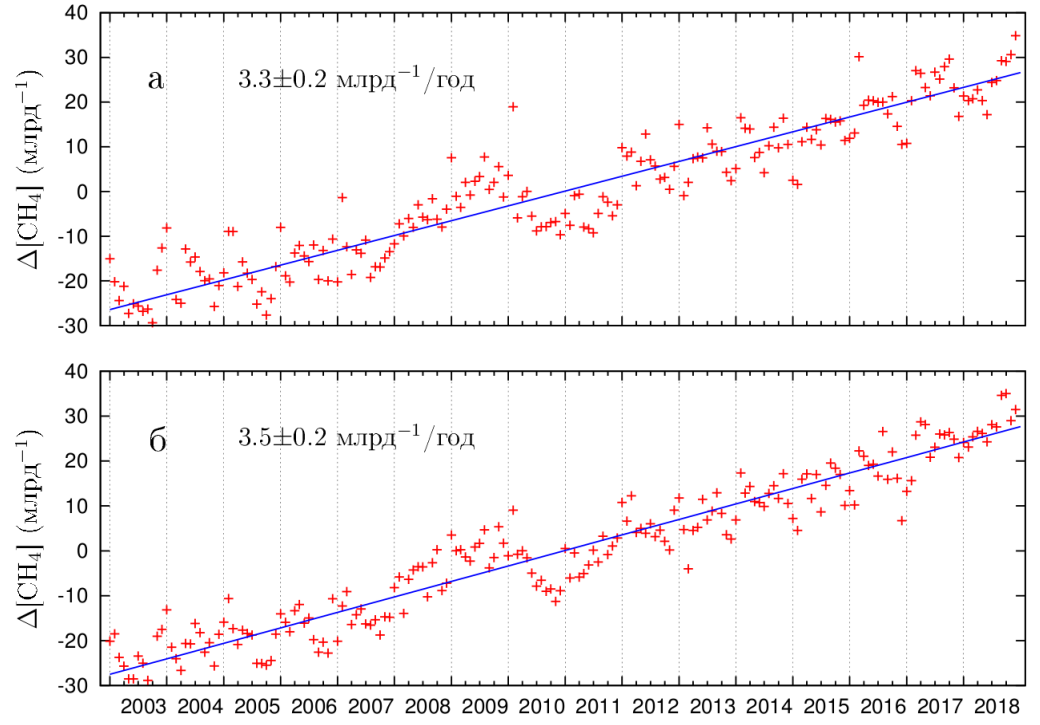



Рис. 2. Аномалии полного содержания метана в атмосфере Западной Сибири для зон 1 (*а*) и 2 (*б*).

На рис. 3 показана чувствительность полного содержания метана [$CH_4$] в атмосфере Западной Сибири в летние месяцы, полученная по данным MOZART-4 для источников, расположенных на территории Европы, Северной Америки и Азии. Видно, что в исследуемый период вклад в [$CH_4$] а атмосфере Западной Сибири могут давать источники расположенные, прежде всего, в Европе и Северной Америке. Максимум чувствительности для этих источников наблюдается на 15 и 25 день после начала эмиссии. На рис. 4 показана область распространения метана от источника в Западной Европе на территорию Западной Сибири на высоте 300 гПа.

Анализ аналогичных результатов для других источников показал, что эмиссия метана из Азии и с территории Аравийского полуострова может оказать влияние на содержание метана в верхней тропосфере Западной Сибири. Такие события происходят в случае обмена субтропических и среднеширотных воздушных масс. Источники, расположенные на севере южной Америки практически не оказывают влияния на [$CH_4$].

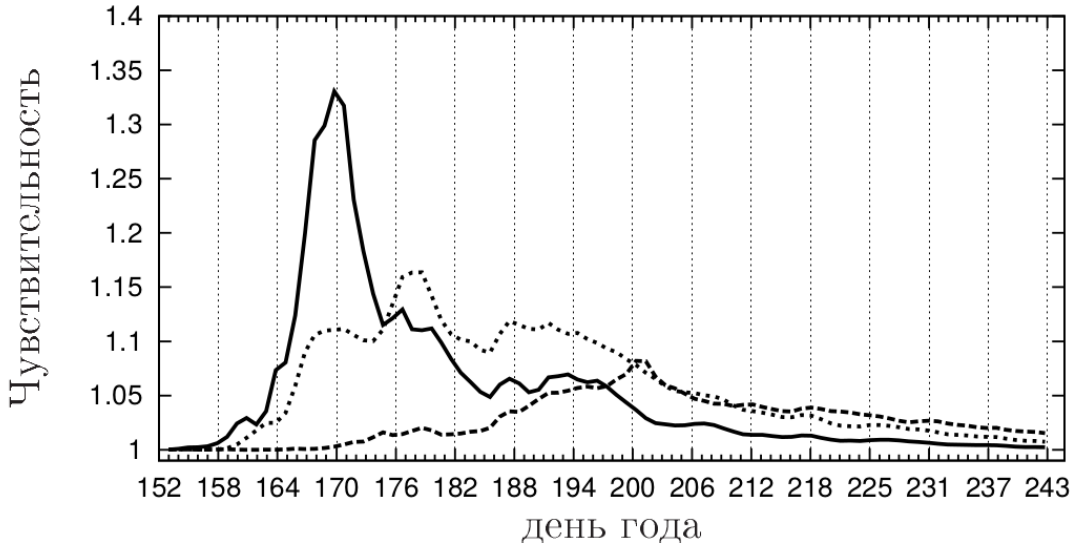



Рис. 3. Чувствительность полного содержания метана [$CH_4$] в атмосфере Западной Сибири в летние месяцы, полученная по данным MOZART-4 для источников, расположенных на территории Европы (сплошная линия), Северной Америки (точечная линия) и Азии (пунктирная линия).

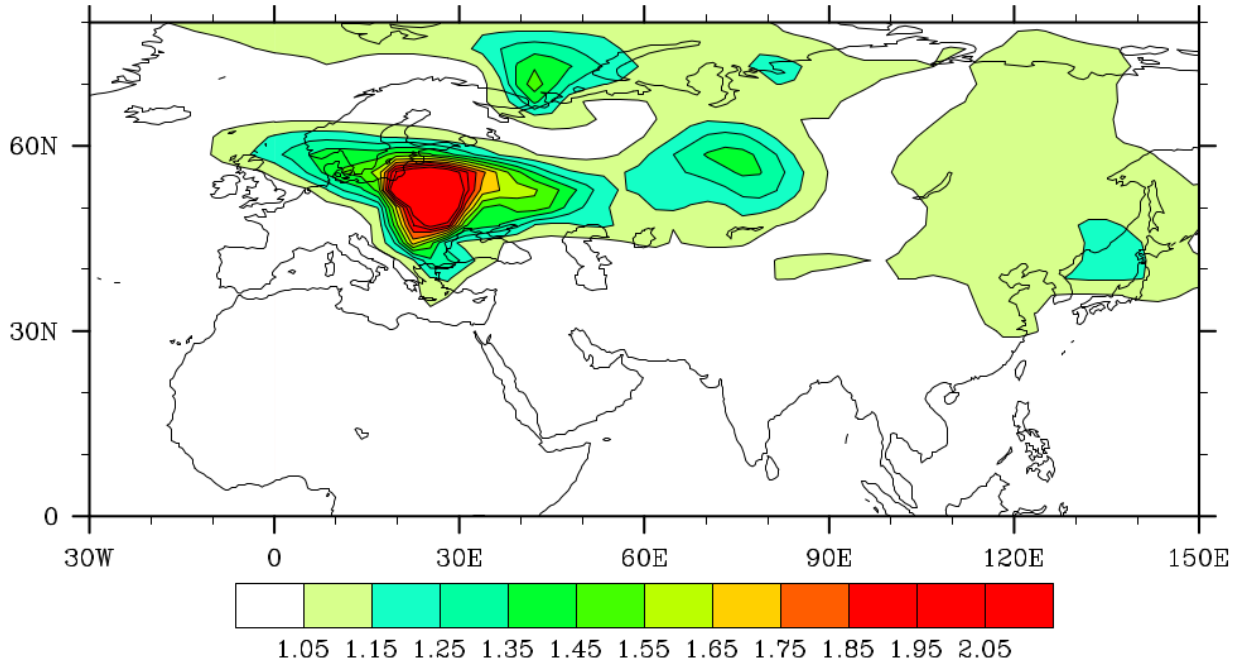



Рис. 4. Чувствительность содержания метана на уровне 300 гПа для источника на территории Европы в июне 2007 г.

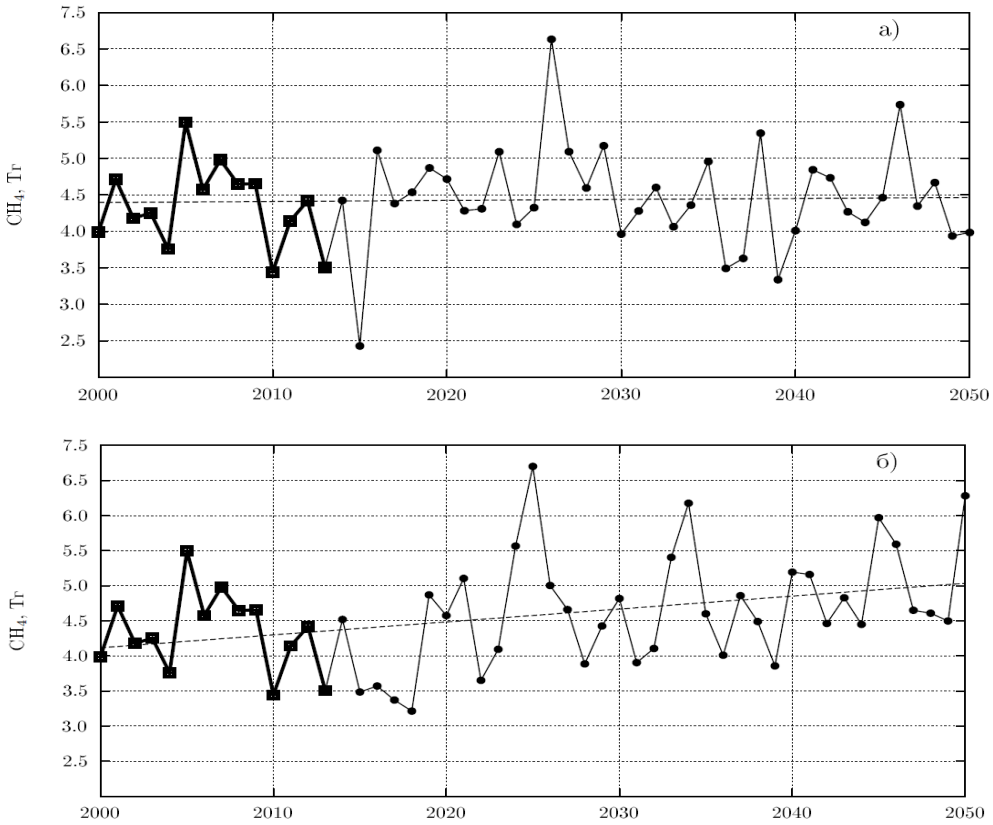



Рис. 5. Эмиссия метана болотными комплексами Западной Сибири в 2000-2050 гг. Жирные линии – данные для современного периода; граничные условия задавались с использованием реанализа NCEP-DOE AMIP-II (R-2). Тонкие линии – прогностические оценки, нормированные на данные современного периода; граничные условия по данным глобальной модели HadGEM2-ES в рамках двух сценариев: *а* – RCP 4.5; *б* – RCP 8.5.

**Эмиссия метана болотными комплексами.** Моделирование эмиссии метана болотными комплексами Западной Сибири проведено в рамках подхода, описанного в [8]. В качестве входных данных использовались результаты региональной климатической модели RegCM4[9], модели переноса тепла и влаги в почве CLM4.5 [10], а также данные базы [11], содержащей информацию о болотных экосистемах Западной Сибири. При моделировании характеристик климатической системы для задания начальных и граничных условий использовались данные реанализа NCEP-DOE AMIP-II (R-2) и глобальной модели HadGEM2-ES для сценариев радиационного форсинга RCP 4.5 и RCP 8.5.

На рис. 5 показаны результаты моделирования эмиссии метана для периода 2000-2050 гг. Прогностические значения для периода 2013-2050 гг. были нормированы с учетом данных за современный период 2010-2013 гг.

В результате анализа полученных данных установлено, что среднее значение эмиссии метана болотными комплексами Западной Сибири в период 2000-2013 гг. составляет ~4.34 Тг/год. Скорость изменения эмиссии метана в этот период практически равна нулю. Рост эмиссии $CH_4$ установлен лишь в зонах тундры и лесотундры. Прогнозируемое увеличение эмиссии метана в 2041-2050 гг. по отношению к периоду 2001-2010 гг. составит 0.58 Тг/год при тренде 0.18 ± 0.06 Тг/10 лет для сценария RCP 8.5. Для сценария RCP 4.5 установлен нейтральный тренд эмиссии метана. Это связано с тем, что рост температуры компенсируется уменьшением влагосодержания почвы.

**Результаты и выводы.** Проведены исследования поведения полного содержания метана в атмосфере Западной Сибири. Основные результаты работы заключаются в следующем.

1. В результате анализа данных гиперспектрометра AIRS/AQUA установлен тренд полного содержания метана в атмосфере Западной Сибири для 2003-2018 гг. ~3.3 ±0.2млрд$^{-1}$/год.

2. Показано, что на полное содержанием метана могут оказывать влияние удаленные источники расположенные, прежде всего, в Европе и Северной Америке.
3. С использованием данных климатических моделей установлено, что среднее значение эмиссии метана болотными комплексами Западной Сибири в период 2000-2013 гг. составляет ~ 4.34 Тг/год. Прогнозируемое увеличение эмиссии метана в 2041-2050 гг. по отношению к периоду 2001-2010 гг. составит 0.58 Тг/год при тренде 0.18 ± 0.06 Тг/10 лет для сценария RCP 8.5. Для сценария RCP 4.5 увеличение эмиссии незначительно.

## ЛИТЕРАТУРА


[1] IPCC 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change / T.F. Stoker, D. Qin, G.-K. Plattner et al. (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. 2013. 1535 pp.

[2] WMO. The state of greenhouse gases in the atmosphere based on global observations through 2017 // WMO greenhouse gas bulletin. 2018. No 14.

[3] *Nisbet E.G., Manning M. R., Dlugokencky E.J. et al.* Very strong atmospheric methane growth in the 4 years 2014–2017: Implications for the Paris Agreement // Global Biogeochemical Cycles, 2019. V. 33. Pp. 318–342.

[4] *Мордвин Е.Ю., Лагутин А.А.* Метан в атмосфере Западной Сибири (монография) / Барнаул: Азбука, 2016. 148 с. ISBN 978-5-93957-893-6 .

[5] *Мордвин Е.Ю., Лагутин А.А.* Содержание метана в свободной тропосфере Западной Сибири / Федеральная служба по интеллектуальной собственности (Роспатент). Свидетельство о государственной регистрации базы данных №2014621683 от 5 декабря 2014 г.

[6] *Щелканов Н.Н.* Обобщенный метод построения линейной регрессии и его применение для построения однопараметрических моделей аэрозольного ослабления // Оптика атмосферы и океана. 2005. Т. 18(1-2). С. 86-90.

[7] *Emmons L.K., Walters S., Hess P.G. et al.* Description and evaluation of the Model for Ozone and Related Chemical Tracers, version 4 (MOZART-4) // Geosci. Model Dev. 2010. Vol. 3. P. 43-67.

[8] *Christensen T. R., Cox P.* Response of methane emission from arctic tundra to climatic change: Results from a model simulation // Tellus. 1995. V. 47B. Pp. 301-310.

[9] *Giorgi F., Coppola E., Solmon F. et al.* RegCM4: Model description and preliminary tests over multiple CORDEX domains // Clim. Res. 2012. V. 52. Pp. 7-29.

[10] *Lawrence D.M., Oleson K.W., Flanner M.G. et al.* Parameterization improvements and functional and structural advances in version 4 of the Community Land Model // J. Adv. Model. Earth Sys. 2011. V. 3. M03001. Doi: 10.1029/2011MS000045.

[11] *Sheng Y., Smith L.C., MacDonald G.M. et al.* A high-resolution GIS-based inventory of the west Siberian peat carbon pool // Global Biogeochem. Cycles. 2004. V. 18. Pp. GB3004. DOI: 10.1029/2003GB002190.

# ЧУВСТВИТЕЛЬНОСТЬ УХОДЯЩЕГО ДЛИННОВОЛНОВОГО ПОТОКА К ВАРИАЦИЯМ КОЭФФИЦИЕНТА ИЗЛУЧЕНИЯ ПОДСТИЛАЮЩЕЙ ПОВЕРХНОСТИ

*Мордвин Е.Ю., Лагутин А.А., Кравченко Р.В.*
Алтайский государственный университет, Барнаул

Исследуется чувствительность потока уходящего из атмосферы длинноволнового излучения к вариациям коэффициента излучения подстилающей поверхности $\varepsilon_c(\nu)$. Реализованный в работе подход к решению рассматриваемой проблемы основан на функциональной теории чувствительности. С использованием полученного в работе аналитического результата по коэффициенту дифференциальной чувствительности уходящего излучения и созданного на базе модели LBLRTM вычислительного комплекса показано, что максимальная чувствительность уходящего излучения к вариациям $\varepsilon_c(\nu)$ наблюдается в диапазонах 780-1000 см$^{-1}$ и 1015-1200 см$^{-1}$.
*Ключевые слова: уходящее длинноволновое излучение, чувствительность, коэффициент излучения, подстилающая поверхность, LBLRTM.*

**Введение.** Поток уходящего длинноволнового излучения является одной из ключевых составляющих радиационного баланса Земли [1, 2]. Он характеризует количество энергии, сформированное излучением подстилающей поверхности Земли и восходящим излучением атмосферы, которое уходит в космос из системы «подстилающая поверхность – атмосфера».

Если для температуры, давления и коэффициента излучения подстилающей поверхности (ПП) использовать обозначения $T_s$, $P_s$ и $\varepsilon_s(\nu)$, для функции Планка – $B(\nu, T_s)$, а под $\tau(\nu, P \to 0; \theta)$ понимать функцию пропускания атмосферой излучения с частотой $\nu$ на трассе «уровень атмосферы с давлением $P$ – спутник», то спектральная интенсивность уходящего из безоблачной нерассеивающей атмосферы под зенитным углом $\theta$ излучения $L_{CLR}(\nu, \theta)$ может быть представлена в виде [3]

$$L_{CLR}(\nu, \theta) = \varepsilon_s(\nu) B(\nu, T_s) \tau(\nu, P_s \to 0; \theta) + \int_{P_s}^{0} B\big(\nu, T(P)\big) \frac{d\tau(\nu, P \to 0; \theta)}{dln(P)} dln(P). \quad (1)$$

Отметим, что вкладами солнечного излучения и процессов перерассеяния ПП нисходящего излучения в спектральную интенсивность (1) уходящего длинноволнового излучения (УДИ) в данной работе пренебрегаем.

Можно также показать, что если доля $\alpha$ пикселя закрыта облаками одного яруса, спектральный коэффициент излучения которых есть $\varepsilon_c(\nu)$, а температура и давление атмосферы на уровне верхней кромки равны, соответственно, $T_c$ и $P_c$, то спектральная интенсивность уходящего из облачной атмосферы излучения описывается выражением

$$L(\nu, \theta) = \big(1 - \alpha\varepsilon_c(\nu)\big) L_{CLR} + \alpha\varepsilon_c(\nu) L_{CLD}(P_c), \quad (2)$$

где

$$L_{CLD}(P_c) = B(\nu, T_c) \tau(\nu, P_c \to 0; \theta) + \int_{P_c}^{0} B\big(\nu, T(P)\big) \frac{d\tau(\nu, P \to 0; \theta)}{dln(P)} dln(P)$$

есть интенсивность излучения в случае непрозрачного облака.

Поток уходящего длинноволнового излучения $F$ находится интегрированием интенсивности (2) по углам и частоте. В случае азимутальной симметрии уходящего излучения поток УДИ определяется равенством

$$F = 2\pi \int_0^{\pi/2} d\theta \int_0^{\infty} L(\nu, \theta) sin\theta cos\theta d\nu.$$

В данной работе интеграл от спектрального УДИ по углу $\theta$ вычисляется в рамках «приближения эффективного угла», который определяется выражением [4, 5]

$$F_\nu = 2\pi L\left(\nu, \theta_{ef}(\nu)\right) \int_0^{\pi/2} sin\theta cos\theta d\theta = \pi L\left(\nu, \theta_{ef}(\nu)\right).$$

Зависимость эффективного угла $\theta_{ef}$ от частоты [4].

| $\Delta\nu$, см$^{-1}$ | $\theta_{ef}$, град | $\Delta\nu$, см$^{-1}$ | $\theta_{ef}$, град | $\Delta\nu$, см$^{-1}$ | $\theta_{ef}$, град |
|---|---|---|---|---|---|
| 2–150 | 50.36 | 650–800 | 52.94 | 1500–1800 | 50.36 |
| 150–250 | 50.64 | 800–950 | 55.86 | 1800–2100 | 52.51 |
| 250–350 | 51.62 | 950–1100 | 53.08 | 2100–2400 | 52.74 |
| 350–500 | 52.57 | 1100–1250 | 55.21 | 2400–2750 | 55.08 |
| 500–650 | 52.53 | 1250–1500 | 51.64 | | |

С использованием этого подхода для потока $F$ получаем:

$$F = \pi \int_0^\infty L\left(\nu, \theta_{ef}(\nu)\right) d\nu. \tag{3}$$

Значения эффективных углов для диапазона 2 - 2750 см$^{-1}$, полученных в [4], приведены в таблице.

Равенства (2)-(3) показывают, что поток УДИ определяется коэффициентом излучения и температурой подстилающей поверхности, профилями температуры и влажности, свойствами облаков, концентрациями парниковых газов и аэрозолей в атмосфере. В силу наличия нелинейных связей между этими характеристиками системы и УДИ для интерпретации экспериментальных результатов, получаемых спутниковыми приборами, необходимы данные по влиянию вариаций характеристик системы на уходящее из атмосферы излучение.

В нашей работе [6] для решения таких задач предложен подход, основанный на функциональной теории чувствительности [7]. Получены выражения для коэффициентов дифференциальной чувствительности показаний каналов спутникового гиперспектрометра ИК-диапазона к вариациям газового состава атмосферы. Показано, что этот коэффициент выражается через массовый коэффициент поглощения исследуемого газа и универсальную функцию, определяемую интенсивностями уходящего излучения для невозмущенной атмосфере.

В данной работе этот подход применяется для анализа влияния вариаций коэффициента излучения ПП на поток уходящего из атмосферы длинноволнового излучения. Актуальность этого исследования обусловлена необходимостью оценки влияния на УДИ изменений структуры ПП, вызванных как изменениями глобального и регионального климата, так и режимом использования земель, учета погрешностей в зависимости коэффициента излучения ПП от частоты при интерпретации спутниковых данных, верификации корректности описания $\varepsilon_s(\nu)$ в климатических моделях.

**Чувствительность уходящего потока к вариациям коэффициента излучения подстилающей поверхности.** Следуя [6], вариацию потока $\Delta F$,

$$\Delta F\big(\nu(\cdot) \to \nu'(\cdot)\big) = F\big(\nu'(\cdot)\big) - F\big(\nu(\cdot)\big),$$

обусловленную изменением коэффициента излучения $\varepsilon_s(\nu) \to \varepsilon_s'(\nu) = \varepsilon_s(\nu) + \Delta\varepsilon_s(\nu)$, представим в виде

$$\Delta F = \int_0^\infty \frac{\delta F\big(\varepsilon_s(\cdot)\big)}{\delta\varepsilon_s(\nu_0)d\nu_0}\Delta\varepsilon_s(\nu_0)d\nu_0. \tag{4}$$

Входящую в (4) вариационную производную принято называть коэффициентом дифференциальной чувствительности (см., [6, 7]). Функция

$$S = \frac{\varepsilon_s(\nu_0)}{F}\frac{\delta F\big(\varepsilon_s(\cdot)\big)}{\delta\varepsilon_s(\nu_0)d\nu_0}$$

описывает изменение $F$ в процентах, обусловленное изменением $\varepsilon_s$ в единичном интервале около $\nu_0$ на 1%.

Вычисляя вариационную производную, находим коэффициент дифференциальной чувствительности и вариацию потока УДИ:

$$\frac{\delta F\big(\varepsilon_\nu(\cdot)\big)}{\delta\varepsilon_s(\nu_0)d\nu_0} = \pi\big(1 - \alpha\varepsilon_c(\nu_0)\big)B(\nu_0, T_s)\tau\left(\nu_0, P_s \to 0; \theta_{ef}(\nu_0)\right), \tag{5}$$

$$\Delta F = \pi \int_0^{\infty} \left(1 - \alpha\varepsilon_c(\nu_0)\right) B(\nu_0, T_s)\tau\left(\nu_0, P_s \to 0; \theta_{ef}(\nu_0)\right) \Delta\varepsilon_\nu(\nu_0) d\nu_0 . \qquad (6)$$

**Результаты.** Расчеты функции пропускания атмосферы $\tau(\nu, P_s \to 0; \theta_{ef})$, потока УДИ и коэффициента дифференциальной чувствительности выполнялись с использованием модели LBLRTM (Line-By-Line Radiative Transfer Model) [8]. Для подготовки необходимых для модели характеристик атмосферы и подстилающей поверхности, а также формирования конфигурационного файла для запуска LBLRTM использовалась авторская программа. Эти данные извлекались из исследовательского продукта AIRS версии 6, который содержит измерения на 100 уровнях атмосферы в диапазоне высот от поверхности до ~ 60 км. Для задания коэффициента излучения подстилающей поверхности использовалась данные MODIS UCSB Emissivity Library of the MODIS LST group at University of California, Santa Barbara (UCSB) (http://www.icess.ucsb.edu/modis/EMIS/html/em.html).

Рис. 1 и 2 демонстрируют зависимость чувствительности потока $S\Delta\nu$ от частоты в условиях безоблачной атмосферы для песчаной поверхности и для поверхности, покрытой растительностью. Нетрудно видеть, что максимумы чувствительности находятся в диапазонах 780-1000 см$^{-1}$ и 1015-1200 см$^{-1}$. Существенно меньшая чувствительность наблюдается в диапазонах 2100-2200 см$^{-1}$ и 2480-2550 см$^{-1}$.

В докладе будут представлены дополнительные материалы, характеризующие чувствительность УДИ к вариациям коэффициента излучения ПП для условий безоблачной и облачной атмосфер.

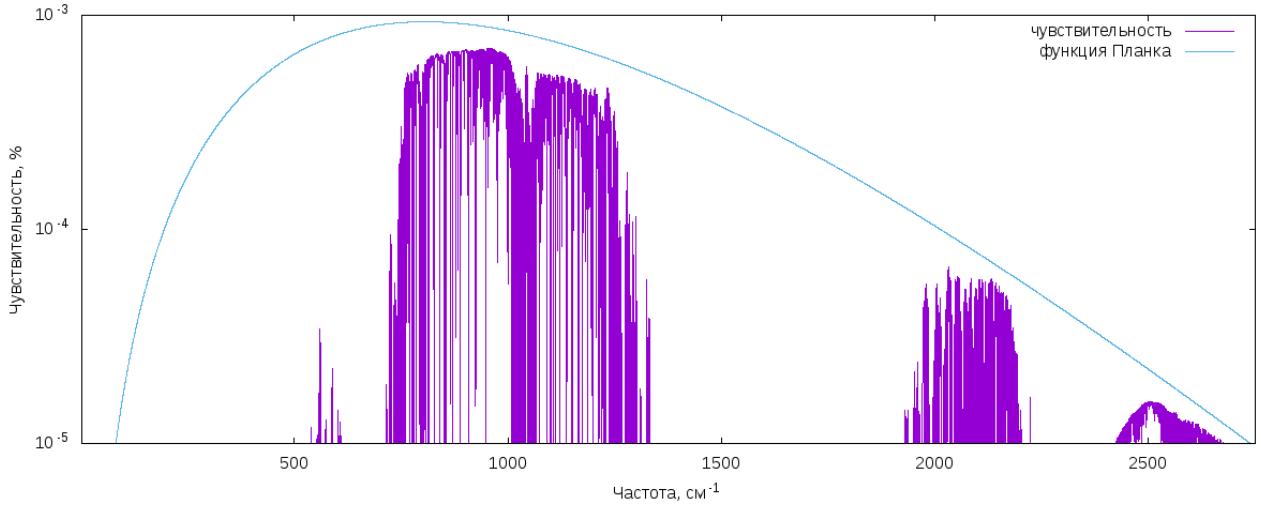



Рис. 1. Чувствительность потока $S\Delta\nu$ к изменению коэффициента излучения $\varepsilon_\nu$ на 1% в интервале $\Delta\nu = \nu/1200$ для песчаной поверхности.

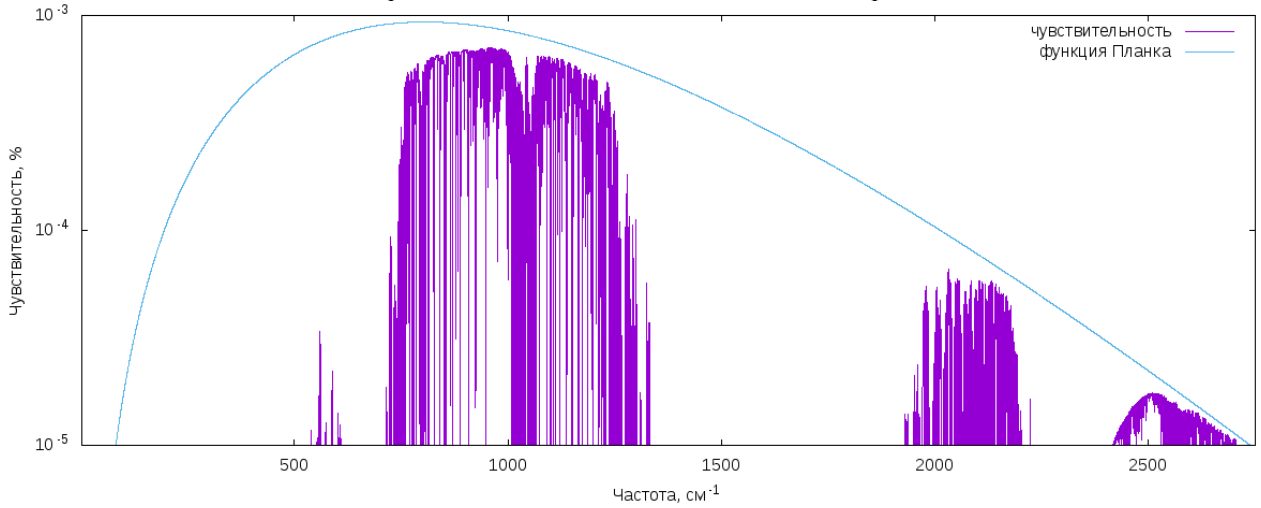



Рис. 2. Чувствительность потока $S\Delta\nu$ к изменению коэффициента излучения $\varepsilon_\nu$ на 1% в интервале $\Delta\nu = \nu/1200$ для поверхности, покрытой растительностью.

**Заключение.** В работе проведено исследование влияния вариаций коэффициента излучения ПП на поток уходящего из атмосферы длинноволнового излучения. Реализованный в работе подход базируется на функциональной теории чувствительности. С использованием полученного в работе аналитического результата по коэффициенту дифференциальной чувствительности УДИ и созданного на базе модели LBLRTM вычислительного комплекса показано, что максимальная чувствительность УДИ к вариациям коэффициента излучения ПП наблюдается в диапазонах 780-1000 $см^{-1}$ и 1015-1200 $см^{-1}$.

ЛИТЕРАТУРА


[1] *Trenberth K.E., Fasullo J.T., Kiehl J.* Earth's global energy budget // Bull. Am. Meteorol. Soc. 2009. V. 90. Pp. 311-323.

[2] *Stephens G.L., Li J., Wild M. et al.* An update on Earth's energy balance in light of the latest global observations // Nat. Geosci. 2012. V. 5. Pp. 691-696.

[3] *Тимофеев Ю.М., Васильев А.В.* Теоретические основы атмосферной оптики. СПб: Наука, 2003.

[4] *Mehta A., Susskind J.* Outgoing longwave radiation from the TOVS Pathfinder Path A data set // J. Geophys. Res. 1999. V.104. N.D1, 2.193-12.212.

[5] *Susskind J., Blaisdell J.M., Iredell L.* Improved methodology for surface and atmospheric soundings, error estimates, and quality control procedures: the atmospheric infrared sounder science team version-6 retrieval algorithm // J. Applied Remote Sensing. 2014. V.8. Pp. 1-33.

[6] *Sarmisokov Z.T., Lagutin A.A., Mordvin E.Yu.* Sensitivity of the satellite thermal infrared hyperspetrometer to variations of atmospheric characteristic // Proc. SPIE. 2017. 10466, 104661Z.

[7] *Лагутин А.А., Учайкин В.В.* Метод сопряженных уравнений в теории переноса космических лучей высоких энергии. Барнаул: Изд-во Алт. ун-та, 2013.

[8] *Clough S.A., Shephard M.W., Mlawer E.J. et al.* Atmospheric radiative transfer modeling: a summary of the AER codes // JQSRT. 2005. V. 91. P. 233-244.

# МЕТОД ИСПОЛЬЗОВАНИЯ КРИТЕРИЕВ ОПАСНОСТЕЙ ДЛЯ ИДЕНТИФИКАЦИИ РИСКООПАСНЫХ СИТУАЦИЙ

*Ничепорчук В.В.*
Институт вычислительного моделирования СО РАН, Красноярск

Разработаны критерии идентификации опасностей и угроз, основанные на систематизации параметров комплексного мониторинга безопасности территорий. Числовые значения интервалов безопасного функционирования территориальных социально-природно-техногенных систем разработаны с учётом требований нормативных и методических документов, экспертного оценивания. Предложен метод использования критериев для идентификации рискоопасных ситуаций, заключающийся в комплексной аналитической обработке мониторинговых данных.
*Ключевые слова: комплексный мониторинг, оперативная обстановка, критерии опасностей*

**Введение.** Решение задач обеспечения безопасного функционирования территориальных С-П-Т систем требует установления допустимых интервалов параметров, контролируемых разными системами мониторинга природных и антропогенных процессов [1]. Большой перечень возможных рисков, характерных для территории Сибири, их взаимовлияние, усиливающее негативные эффекты проявления, обосновывает необходимость сбора, консолидации, комплексной обработки данных всего спектра оперативных наблюдений [2, 3]. Разработка унифицированного подхода к идентификации опасностей и угроз позволяет создавать интегрированные информационно-аналитические системы, позволяющие решать оперативные и стратегические задачи управления безопасностью территорий [4].

В отличие от интегральных оценок рисков, оценка которых проводится ежегодно, органы управления территориальной безопасностью решают задачи идентификации опасностей и угроз в непрерывном режиме [5]. При этом используются практически те же информационные ресурсы комплексного мониторинга, а методы аналитической обработки данных аналогичны применяемым при решении задач информационной поддержки превентивных мероприятий по снижению рисков.

Функционал большинства систем мониторинга сводится к on-line визуализации контролируемых параметров и аналитической обработки архивов наблюдений [6]. Идентификация опасностей и угроз, как правило, проводится в ручном режиме путём выборки максимальных (или минимальных) значений за исследуемый период. Дальнейшее принятие решений реализуется с привлечением дополнительных информационных ресурсов, систем ситуационного и аналитического моделирования. Использование большого количества систем при формировании решений существенно снижает эффективность управления. Фрагментарная программная реализация подзадач безопасного развития территорий приводит к дублированию информации, а различие в регламентах актуализации – к противоречиям, разрешение которых снижает оперативность и достоверность управленческих решений.

С целью построения универсального подходы идентификации опасностей и угроз разработаны критерии предельных значений параметров для доступных видов мониторинга С-П-Т систем. Предложен метод использования критериев для раннего обнаружения предвестников опасных ситуаций и своевременного оповещения аварийно-спасательных формирований, населения, руководящего состава.

**Систематизация параметров мониторинга.** С целью систематизации параметров мониторинга *PR* введём понятия «комплексный мониторинг» и «оперативная обстановка» и обозначения составляющих их элементов [7]. *Комплексный мониторинг* – система наблюдений и контроля, проводимых регулярно по определенной программе для оценки состояния окружающей среды и объектов техносферы, являющихся источниками опасностей $O_1$, анализа происходящих в ней процессов и своевременного выявления тенденций ее изменения, а также сбор данных о характеристиках и текущем состоянии защищаемых объектов $O_2$ и объектов

управления $O_3$ для последующей комплексной обработки [8]. *Оперативная обстановка ST* – совокупность факторов, условий и обстоятельств, в которых осуществляется подготовка и выполнение определённых действий в сфере обеспечения безопасности объектов и территорий, описываемых значениями элементов множества видов информационных ресурсов *S* для определённого вида ситуаций *H* [9].

Основными видами *ST* являются метеорологическая, гидрологическая, сейсмическая, радиационная, лесопожарная, санитарно-эпидемическая и ряд других. Консолидация параметров разных видов мониторинга позволяет учёсть влияние *ST* друг на друга. Например, неблагоприятные метеорологические факторы усиливают масштаб опасностей практически для всех видов опасных ситуаций, а факторы фитопатологической обстановки определяют опасность природных пожаров.

Первичным источником параметрических данных контроля $O_1$ является инструментальный мониторинг. Приборы позволяют измерять значения физических параметров, таких как температура, давление, скорость, концентрация вещества, мощность излучения, изменение размеров и др., и передавать данные для обработки в сколь угодно малые промежутки времени. К приборному мониторингу относятся и средства аварийной сигнализации, передающие сигнал об опасном событии на основе внутренней обработки результатов измерений. Информационные ресурсы мониторинга содержат результаты инструментальных измерений различной степени обработки (исправление ошибок, агрегация и др.). Однако, для непосредственного контроля функционирования оборудования на объектовом уровне приборы являются единственным методом оперативной идентификации опасностей и угроз.

Важнейшим источником информации о состоянии больших территорий являются данные дистанционного зондирования Земли [10]. Помимо перечисленных выше параметров космические аппараты контролируют широкий спектр электромагнитного излучения, альбедо, пространственные характеристики поверхности [11]. В зависимости от пространственного разрешения, принимаемых каналов, периодичности съёмки системы ДЗЗ используются для информационной поддержки управления практически во всех видах опасных ситуаций [12]. В отличии от наземного мониторинга, для использования в процессе идентификации опасностей и угроз данные космической съёмки требуют сложной обработки с использованием специализированных программных комплексов [13].

**Критерии обнаружения опасностей и угроз.** Решение задачи идентификации опасностей основывается на оперативной аналитической обработке потоков данных комплексного оперативного мониторинга состояния окружающей среды и объектов техносферы. Идентификация опасностей и угроз заключается в сравнении наблюдаемых или вычисляемых параметров критериям, уникальным для вида ситуаций *H*.

Основным документом, на основе которого разработаны критерии опасностей и угроз является Приказ МЧС России № 329 от 8.07.2004 г. «О введении в действие перечня критериев информации о чрезвычайных ситуациях». Использованы также ведомственные методики и нормативно-правовые акты Росгидромета, Министерства природных ресурсов РФ, Министерства транспорта РФ, Министерства энергетики РФ и другие.

Общие критерии используются для всех пунктов наблюдений за сейсмической, радиационной обстановками, уникальные – для гидрологической обстановки (данные критических уровней воды), функционирования объектов техноосферы и т.п. Групповые критерии метеорологической обстановки применяются для различных климатических зон – территорий, расположенных в умеренных широтах, Арктической зоны, горной местности.

Критерии опасностей пересматриваются с различной периодичностью: лимиты расходов сброса ГЭС устанавливаются ежемесячно в зависимости от гидро- и метеорологической обстановки; критические уровни воды – раз в пять лет. Более длинные периоды действия у критерием радиационной опасности, зависящих от длительных медицинский исследований, а также критериев идентификации техногенных происшествий. Поскольку некоторые виды ситуаций *H* имеют сезонный характер, их контроль происходит в определенные месяцы годового

цикла. Для природных ЧС, зависящих от температуры воздуха существуют суточные интервалы контроля опасностей. Например, подъём уровня воды и возгорания в лесу наиболее вероятны в период с 13 до 17 часов. Обоснование сезонных и суточных интервалов для других видов ситуаций целесообразно реализовывать на основе аналитической обработки каталогов чрезвычайных ситуаций и происшествий.

Значения критериев угроз при отсутствии конкретных данных в литературе или результатов натурных наблюдений определялись экспертным путём как доля от опасного значения. Большинство критериев опасности представляют собой максимально или минимально возможное значение контролируемого параметра, а критериев угроз – процент от критического значения. Кроме этого, критерии могут представлять собой сумму значений за определенный период, разность соседних измерений, а также количество объектов имеющих определённые свойства, либо изменивших свойства за период. Для техногенных объектов критерии опасностей и угроз могут быть в середине рабочий интервалов (например, скорости вращения гидроагрегатов ГЭС).

Принятие решений по экстренному реагированию на опасности и угрозы имеет многоуровневый характер. Оперативный мониторинг состояния окружающей среды проводится на региональном уровне. Мониторинг объектов техносферы проводится дежурно-диспетчерскими службами объектового уровня, а в вышестоящие органы управления данные поступают в агрегированном виде. Это связано не только с большим объёмом информации контроля узлов и участков промышленных объектов и инфраструктуры территорий, но и с необходимостью специальных знаний для интерпретации сигналов в качестве признаков нештатной или аварийной ситуации.

Таким образом, критерии опасности в большинстве в случаев являются необходимыми, но не достаточными условиями реализации опасного события. После перепроверки сигнала по другим источникам информации и в зависимости от вида ситуации *H* принимается решение по реагированию оперативных служб или переводе их в режим повышенной готовности. Полный цикл информационной поддержки управления с использованием критериев идентификации заключается в последовательном решении следующих задач:

- идентификации опасностей или угроз в параметрах текущего или прогнозируемого состояния какой-либо обстановки *ST* на пункт наблюдений *PO,* с оповещением (информированием) оперативных служб и населения для конкретной территории;
- идентификации превышения показателей риска сверх допустимых значений с выполнением задач планирования превентивных мероприятий для территории с повышенным уровнем риска ЧС [14].

В таблице приведены значения критериев опасностей для параметрического мониторинга разных видов опасностей.

Таблица дополняется по мере появления новых данных мониторинга опасностей $O_1$. Например, для обнаружения термических точек на лесных территориях по данным ДЗЗ с инфракрасном диапазоне используется критерий предельной температуры +60°…+80° в зависимости от класса пожарной опасности, типа приёмника, характеристик территории [13].

Помимо приведенных предельных значений параметров PR разработаны критерии опасностей и угроз в системах мониторинга событий. Для опасностей техногенного характера это факт срабатывания систем сигнализации, например ЭРА-ГЛОНАСС, Гонец, КОСПАС-САРСАТ, пожарная и т.п. Критерием угрозы является время, прошедшее с момента остановки функционирования системы. Например, решения об экстренной эвакуации населения принимаются на основании длительности остановки системы теплоснабжения в сочетании с масштабом аварии и температурой наружного воздуха. Проявление опасностей природного характера (морось, гололёд, затор льда, сухие грозы и т.п.) затруднительно измерить в числовых шкалах. Однако факт возникновения угроз повышает вероятность возникновения опасных ситуаций других видов – аварий автомобильного транспорта, затоплений территорий, природных пожаров.

Критерии опасностей и угроз.

| Параметр, ***PR*** | Критерии угроз, $KR_2$ | Критерии опасностей, $KR_1$ | Примечание |
|---|---|---|---|
| **ОПАСНОСТИ ТЕХНОГЕННОГО ХАРАКТЕРА** | | | |
| **Радиационная обстановка** | | | |
| МЭД (*Md*, мЗВ) | *Md > 0,6* | *Md > 1,2* | Разовое значение |
| **Обрушение зданий и сооружений** | | | |
| Уровень снега на крупнопролётных конструкциях *(LS,* см) | *LS > 20* | *LS > 30* | Может изменяться в зависимости от конкретного объекта |
| **Аварии на ЖКХ** | | | |
| Давление горячей воды теплокамере снабжения домов (*PW*, атм) | *PW < 6* | *PW < 4* | На прямой подаче |
| Температура горячей воды в теплокамере снабжения жилых домов (*TW*, °С) | *TW < 90* | *TW < 70* | На прямой подаче |
| **ОПАСНОСТИ ПРИРОДНОГО ХАРАКТЕРА** | | | |
| **Опасные метеорологические явления** | | | |
| Температура (*t*, °С) | *30 ≤ t < 35 или -40 < t ≤ -35* | *переход через 0* | Может изменяться в зависимости от климатической зоны |
| | *t ≥ 35* | *t ≤ -40* | |
| Скорость ветра (*SW*, м/с) | *15 ≤ SW < 25* | *SW ≥ 25* | |
| Осадки (*Pre*, мм) | *Pre ≥ 30 за 1 час* / *Pre ≥ 50 за 12 час.* / *Pre ≥ 120 за 42 час.* | | Для твёрдых и жидких осадков |
| Диаметр отложения мокрого снега (*DS*, мм) | *DS > 20* | *DS > 30* | |
| Класс пожарной опасности, *CFD* | *CFD = 3* | *CFD ≥ 4* | Вычисляется на основе температуры и влажности воздуха, продолжительности периода без осадков |
| Уровень снега на лавиноопасном участке (*LS,* см) | *LS > 30* | *LS > 50* | Может изменяться в зависимости от уклона и экспозиции склона |
| Продолжительная засуха (*DD,* суток) | *DD > 7* | *DD > 14* | Может изменяться в зависимости от климатической зоны |
| **Опасные гидрологические явления** | | | |
| Уровень воды в реках (*RL*, см) | *RL* ≥ Уровень начала подтопления $Kh_1 \times 0{,}8$ | | Уникальные критерии $Kh_i$ для каждого пункта наблюдений *PO* |
| | *RL* ≥ Уровень начала подтопления $Kh_1$ | | |
| Суточное изменение уровня воды (*dRL*, м) | *dRL > 100* | | |
| **ОПАСНОСТИ БИОЛОГО-СОЦИАЛЬНОГО ХАРАКТЕРА** | | | |
| **Массовые заболевания людей** | | | |
| Концентрация загрязняющего вещества, *Cp* | $Cp \geq maxD$ | $Cp \geq maxD_{24}$ | *maxD* – максимальная разовая концентрация; $maxD_{24}$ – макс. суточная экспозиционная концентрация |
| | $C \geq maxD \times Nc$ *или* $C \geq maxD_{24} \times Nc$ | | $Nc \in [3; 10]$ в зависимости от контр-го параметра |

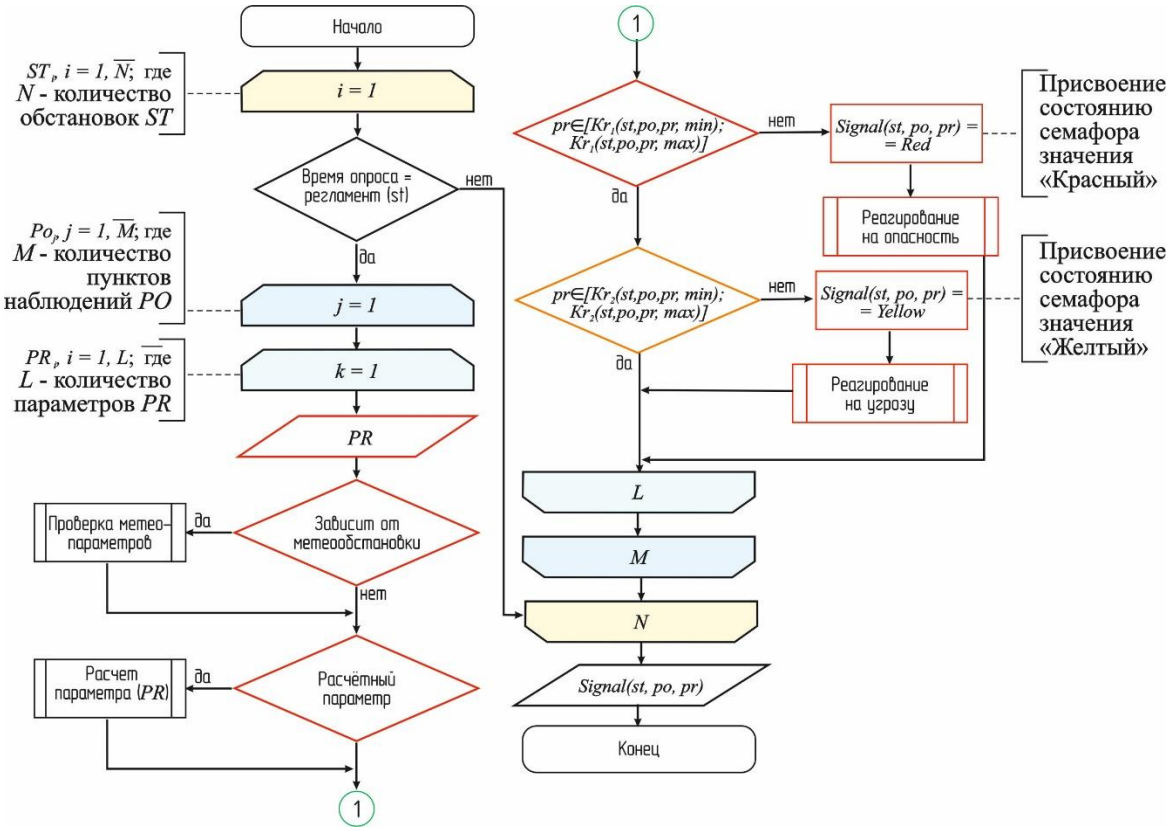
**Использование критерии опасностей и угроз.** Автоматическая индикация угроз ЧС и неблагоприятных параметров обстановки основана на «принципе семафора». Обнаружение угроз реализовано путём проверки данных мониторинга на соответствие или выход за пределы установленных пороговых значений. Оценивание осуществляется путём сравнения текущих значений контролируемых параметров с критериями, приведёнными в таблице. Фильтрация осуществляется для определения уровней состояния объектов контроля. В зависимости от её результатов контролируемые параметры раскрашиваются в следующие цвета:

- «зелёный» – обстановка в норме, значения контролируемых параметров находятся в пределах допустимых значений;
- «угроза» – идентификация угрозы, когда значения параметров приближаются к критическим, либо наблюдается резкая динамика изменений, либо фиксируется длительное отклонение от среднемноголетних нормативных значений;
- «красный» – идентификация опасности, повышенный риск возникновения ситуации, связанной с ущербом либо нарушением жизнедеятельности территории. При этом значения параметров равны или превышают критические.

Серый цвет индикатора свидетельствует об отсутствии оперативных данных для оценки и является сигналом к необходимости проверки источника информации и канала передачи данных.

Практика показала оптимальное применение трёхцветной градации для оперативного принятия решений с использованием заранее разработанных сценариев реагирования. Сценарии описывают последствия и действия в опасной ситуации в зависимости от вида, масштаба и корректируются с учётом конкретной территории и складывающихся условий.

Алгоритм использования критериев для идентификации опасностей и угроз представлен на рисунке. Процесс включает три цикла: по видам обстановок *ST*, по пунктам наблюдений *PO* и по измеряемым параметрам *PR*. Состав мониторинговой информации зависит от вида источника информации. Пакет данных приборного мониторинга можно представить вектором параметров *PR*. Данные систем наблюдений или веб-порталов интегрируют сведения со всех пунктов наблюдений *PO* (матрица *PO*× *PR)*.




Алгоритм идентификации опасностей и угроз.

Большинство видов обстановок *ST* зависит от состояния и прогноза погоды, поэтому при анализе обстановок необходим учёт метеорологических параметров. Диапазон допустимых значений обстановок при неблагоприятных значениях метеорологических параметров уменьшается. Некоторые параметры *PR* требуют предварительного расчёта, например, сумма осадков *Pre* за период или класс пожарной опасности в лесах *CFD*. В процессе сравнения текущих значений *PR* с критериями опасностей и угроз *Kr* происходит заполнение массива *Signal*. В случаях изменения значений элементов массива *Signal[st, po, pr]* решаются задачи оперативного реагирования $t_{21}$ и $t_{22}$ согласно заранее разработанных сценариев действий.

С помощью операции агрегации разработана автоматическая индикация опасностей и угроз для обстановки в целом. Обстановке присваивается уровень опасности, соответствующий наихудшему из уровней множества аналитических индикаторов. Сигнализация опасностей и угроз происходит на уровне территории в целом, а функция детализации аналитической модели OLAP позволяет просмотреть пункт наблюдения и параметр, значения которого вышли за пределы допустимых. Данная функция реализована в системе ЭСПЛА-М в составе АРМ диспетчера Территориального центра мониторинга Красноярского края [15].

**Заключение.** Разработаны критерии идентификации опасностей и угроз, основанные на систематизации параметров комплексного мониторинга безопасности территорий. Числовые значения интервалов безопасного функционирования территориальных социально-природно-техногенных систем разработаны с учётом требований нормативных и методических документов, экспертного оценивания. Предложен метод использования критериев для идентификации рискоопасных ситуаций, заключающийся в комплексной аналитической обработке мониторинговых данных.

## ЛИТЕРАТУРА


[1] *Москвичёв В.В., Бычков И.В., Потапов В.П., Тасейко О.В., Шокин Ю.И.* Информационная система территориального управления рисками развития и безопасностью // Вестник РАН. 2017. Т. 87, № 8. С. 696-705.

[2] *Москвичев В.В., Шокин Ю.И.* Антропогенные и природные риски на территории Сибири // Вестник РАН. 2012. № 2. С. 131-140.

[3] *Москвичев В.В., Тасейко О.В., Иванова У.С., Черных Д.А*. Базовые региональные риски развития территорий Сибирского федерального округа // Вычислительные технологии. 2018. Т. 23. № 4. С.95-109.

[4] *Москвичёв В.В., Ничепорчук В.В., Потапов В.П., Тасейко О.В., Фалеев М.И.* Информационное обеспечение мониторинга и рисков развития социально-природно-техногенных систем // Проблемы анализа риска. 2018. Т. 15, № 2. С. 56-77.

[5] *Пенькова Т.Г., Метус А.М., Ничепорчук В.В*. Метод интегрального аналитического оценивания природно-техногенной безопасности территорий (на примере Красноярского края) // Проблемы анализа риска. 2018. Т. 15, № 5. С. 16-25. DOI: 10.32686/1812-5220-2018-15-5-16-25.

[6] *Фалеев М.И., Соколов Ю.И. и др.* Раннее предупреждение о чрезвычайных ситуациях. М.: МЧС России, 2015. 232 с.

[7] *Ничепорчук В.В., Ноженков А.И.* Архитектура территориальной системы мониторинга чрезвычайных ситуаций // Информатизация и связь. 2018. № 2. С. 35-41.

[8] ГОСТ Р 22.1.02-95. Безопасность в чрезвычайных ситуациях. Мониторинг и прогнозирование. Термины и определения. М.: СтандартИнформ, 2017. 5 с.

[9] Гражданская защита: Энциклопедический словарь; под общей ред. В.А. Пучкова / МЧС России. М.: ФГБУ ВНИИ ГОЧС (ФЦ), 2015. 664 с.

[10] *Резников В.М.* Аэрокосмическая система мониторинга: состояние, проблемы, перспективы. М.: ФГУ ВНИИ ГОЧС (ФЦ), 2009. 200 с.

[11] *Якубайлик О.Э., Павличенко Е.А., Ромасько В.Ю.* Организация оперативной обработки данных дистанционного зондирования Земли //Международный научно-исследовательский журнал. 2018. № 11 (77), ч. 1. С. 79-81. DOI: 10.23670/IRJ.2018.77.11.014.

[13] *Михайлов С.И.* Классификатор тематических задач МЧС России, решаемых с использованием данных ДЗЗ из космоса // Земля из космоса. Наиболее эффективные решения. 2010. Вып. 4. С. 64-71.
[14] *Бондур В. Г., Крапивин В. Ф., Савиных В. П.* Мониторинг и прогнозирование природных катастроф. М.: Научный мир, 2009. 692 с.
[15] Атлас природных и техногенных опасностей и рисков чрезвычайных ситуаций: научное издание / МЧС России. М.: Феория: Объединённая редакция МЧС, 2011. 652 с.
[16] *Ноженков А.И., Ноженкова Л.Ф., Ничепорчук В.В., Пенькова Т.Г., Коробко А.В., Марков А.А., Метус А.М*. Информационно-аналитическая система мониторинга и прогнозирования чрезвычайных ситуаций ЭСПЛА-М». Свидетельство об официальной регистрации программы для ЭВМ № 2018661879 от 20.09.2018 г.

# ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ОБРАБОТКИ СПУТНИКОВЫХ ДАННЫХ ДЛЯ ВЫЧИСЛЕНИЯ ГРАДИЕНТНЫХ ХАРАКТЕРИСТИК ПОВЕРХНОСТИ МИРОВОГО ОКЕАНА

*Огородов Н.А.*[(1)(2)], *Картушинский А.В.*[(3)(2)]
[(1)] Федеральный исследовательский центр КНЦ СО РАН, Красноярск
[(2)] Сибирский федеральный университет, Красноярск
[(3)] Институт биофизики СО РАН, Красноярск

Рассматриваются структурные компоненты информационно-программных средств для расчета градиентных характеристик поверхности Мирового океана на основе спутниковых данных. Представлены примеры расчета градиентных полей поверхности Мирового океана с использованием информационно-программных средств на основе спутниковых данных. Обсуждается пространственно-временное усреднение градиентов в соответствии с масштабами гидрофизических процессов. Рассматриваются проблемы пространственно-временного масштабирования и усреднения расчетных областей Мирового океана.
*Ключевые слова: глобальные данные, преобразование формата, региональные особенности, градиентныеполя, фронтальные зоны, усреднение пространственное, среднесезонная изменчивость.*

**Введение**. Проблема глобального изменения климата привела к необходимости поиска причинно-следственных связей между опасными погодными явлениями и глобальными процессами в Мировом океане. В последнее время для решения задачианализа проявления тенденций изменения климата используются глобальные данные, представляющие собой результаты усреднения данных метеостанций, судовых, буйковых и спутниковых измерений и других источников, которые применяются в моделях ре-анализов [1, 2]. На основе этих данных проводятся исследования глобальных климатических изменений с использованием данных температуры поверхности океана (ТПО) на крупномасштабных расчетных сетках, и сравнения глобальных различий давления на уровне моря и температуры воздуха у поверхности (ТВП), возникающих при Эль-Ниньо и Ла-Нинья [3]. Особое внимание уделяется расчетам временных энергетических спектров индексов Эль-Ниньо–Южное колебание (ЭНЮК) и Глобальной атмосферной осцилляции, рассчитанных по данным наблюдений и их ре-анализам [3], с соответствующими характеристиками, полученными при интегрировании моделей CMIP5 [4], с целью выявления недостатков современных климатических моделей.

Постоянное усовершенствование глобальных наборов данных позволяет работать с доступными данными измерений, которые включают в себя интерполяционные корректировки и усреднение данных судовых и буйковых измерений для восстановления последующего периода и оценки неопределенности, связанных с внутренними параметрами ре-анализа. ТПО демонстрирует, например, существенно более реалистичное поведение Эль-Ниньо/Ла-Нинья в ранний период фиксации, когда данные являются малочисленными, нопозволяют получить улучшенную оценку долгосрочной изменчивости в зоне ключевого режиме внутренней изменчивости климата [5].

В глобальных наборах данных улучшается представление локальной детализации ТПО, используя наложение исходных данных ТПО с привязкой к градусной сетке, в том числе и с подавлением случайных ошибок выборки и контактных измерений. Здесь обеспечивается уникальный глобальный анализ исторических данных ТПО с середины 19-го века 1871 года [6]. Однако данный набор данных не обладает достаточно высоким пространственным разрешением, чтобы исследовать региональные особенности динамики ТПО, которые можно фиксировать по спутниковым данным и выявлять локальные особенности структурной организации температурных фронтальных зон (ТФЗ) при условии меандрирования и вихреобразования в системе взаимодействия холодного Лабрадорского течения и теплого струйного течения Гольфстрим [6, 11].

**Материалы и методы**. В работе применяется информационная технология и информационно-программные средства (ИПС) для изучения изменчивости ТФЗ поверхностного слоя океана на основе спутниковых данных. Данная информационная технология состоит из двух частей: это конвертор для преобразования исходных данных формата hdf и информационной системы. Основой ИПС для расчета и построения градиентных полей поверхности Мирового океана является база спутниковых данных, включающая в себя информацию из банков данных PODAAC и Ocean Color [7, 8]:

- спутниковые данные температуры поверхностного слоя мирового океана (NOAA) за период с 1981 по 2002 в формате HDF. Пространственный масштаб 18км/пикс -среднене-дельный композит;
- спутниковые данные температуры поверхностного слоя мирового океана (MODISA) в период с 2002 по 2014 в формате HDF. Масштаб 9км/пикс - среднемесячный композит.

HDF- многообъектный формат файла для формирования данных в распределенной среде, разработанный в 1992 году в Национальном центре по применению суперкомпьютеров для того, чтобы использовать данные учеными, работающими над одним проектом, но в разных научных направлениях [9]. Основным преимуществом формата является самоописываемость формата. Термин «самоописывающийся» подразумевает, что для каждой структуры HDF данных в файле содержится необходимая информация о местоположении точек измерения и их распределении в файле. Эта информация указывается как метаданные (metadata). Доступ к различным частям информации осуществляется посредством метаданных, которые задает пользователь при создании файла [9]. Наличие метаданных в формате несколько осложняет обработку информации, в виду того, что структура метаданных в различных банках данных может существенно отличаться. Поэтому данные перед обработкой подготавливаются и приводятся к определенному формату.

Для преобразования в необходимый формат данных разработан конвертор PARSER, который представляет собой консольное приложение, работающее в тандеме с утилитой HDFTOBIN. Основная задача приложения вызывать конвертор HDFTOBIN на каждый файл из выборки, передавая ей в качестве атрибута имя файла с дальнейшей постобработкой *.data файла и формировании выборки в формате SSV. Финальным этапом формирования выборки является формализация выборки, в процессе которой файлы-заготовки переименовываются в соответствии с требованиями формата SSV. В зависимости от выбранного пункта меню запускается та или иная функция обработки данных. На рис. 1 представлена структурная схема работы конвертора PARSER.

В качестве инструмента для исследования изменчивости структурных компонент поверхности океана используется информационная система (ИС) SSVOSatD (Sea Surface Variability of Oceanbased on Satellite data), которая позволяет работать с данными на любом носителе, так как жесткой привязки к директории с входными данными нет.

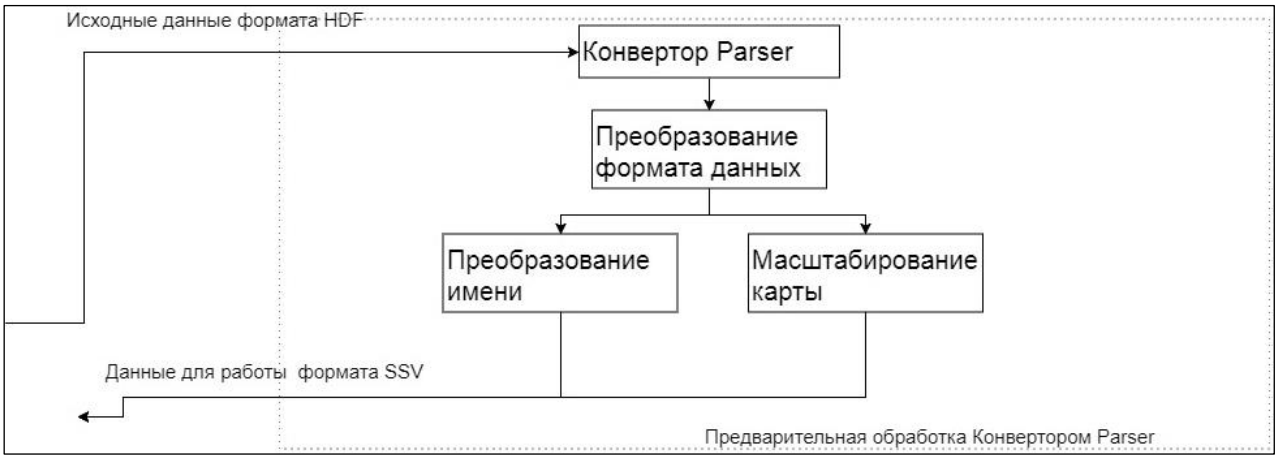



Рис. 1. Схема элементных связей конвертора PARSER

При разработке программного обеспечения использована интегрированная среда разработки (IDE) приложений BorlandBuilder C++ 6.0 Enterprise. IDE Builder 6.0 относится к RAD системе (система быстрого проектирования). Проблема быстроты обработки входной информации решена за счет предварительного конвертирования входных данных в специальный формат SSV, основанный на представлении информации в бинарном виде. Такой подход существенно сокращает время обработки информации [10].

**Результаты и обсуждение.** Основа технологии обработки спутниковых данных для расчета градиентных характеристик, проводится при помощи разработанной программно-информационной системы реализующей программные процедуры и имеющей ряд возможностей. Функциональные возможности ИС заключаются в следующем:

- обеспечение усреднения входных данных по времени (месячное, сезонное, годовое) и масштабу расчетной сетки (8, 16, 32, 64, 128), которое позволяет выделять крупномасштабные и долговременные гидрофизические процессы;
- расчет градиентов и визуализация кластеров по значениям градиентов, что подразумевает процесс разбиения выборки на однородные группы с указанием количества выделяемых кластеров и максимального числа итераций расчета, (позволяет визуально выделить зоны в океане со схожими изменениями значений градиентов), что дает наглядное отображение структуры температурных фронтальных зон. Сектора, в которых изменения значений градиентов существенно отличаются, автоматически разбиваются на интервалы значений градиентов, что приводит к формированию карты кластеров. Пример кластеризации представлен на рис. 2.

Одной из важных особенностей ИС является кросс-корреляция по выделяемым пространственным областям. Расчет коэффициентов корреляции позволяет выявлять наличие или отсутствие связи между соответствующими заданными переменными и необходимы для анализа динамики процессов в поверхностном слое Мирового океана. Анализируемые секторы выбираются с учетом нескольких постановочных задач: кластеры с близким временным ходом изменчивости градиентов в широтном (зональном), меридиональном и градиенты по модулю, что позволяет выявлять энергоактивные зоны, пространственно-временные масштабы динамических процессов в Мировом океане, в том числе определять механизмы глобальной циркуляции (конвейерная лента).

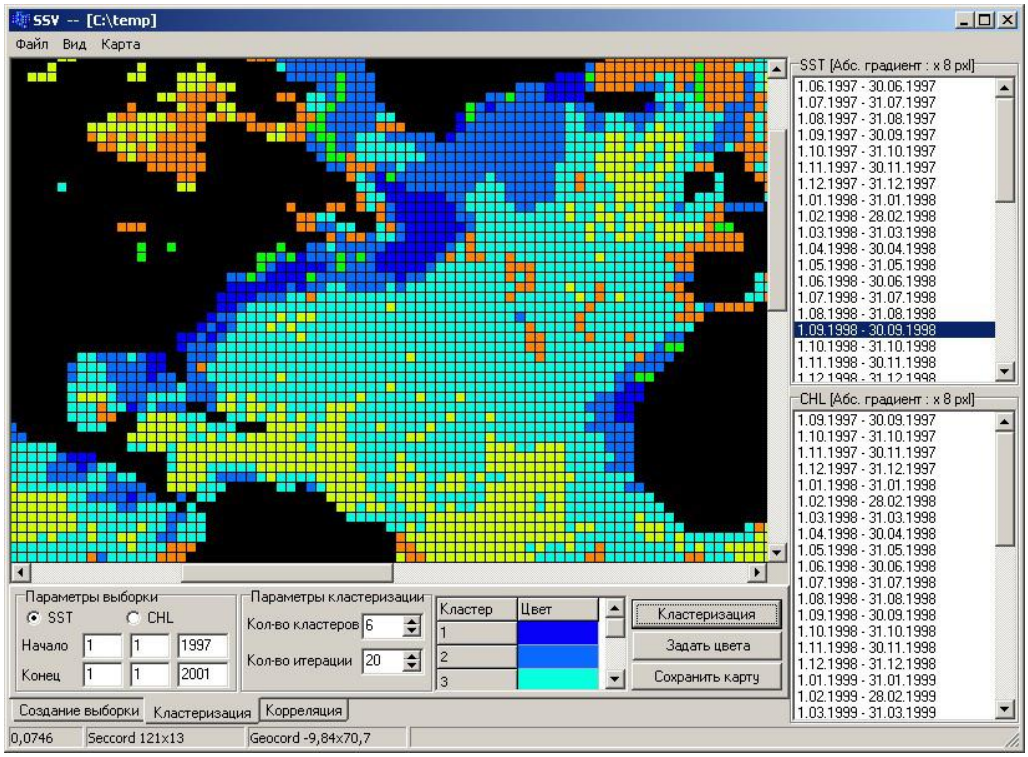



Рис. 2. Пример вывода информации – «Панель Кластеризация».

Дополнительными функциональными возможностями ИС является отображение графиков временного хода градиентных показателей для нескольких выбранных секторов, отражающего динамику значений градиентов. Так же реализована возможность пересчета коэффициента корреляции после временного «смещения» (лага) значений градиентов второго сектора относительно первого, что позволяет выявлять связь между процессами с учетом временной задержки. Пример представлен на рис. 3.

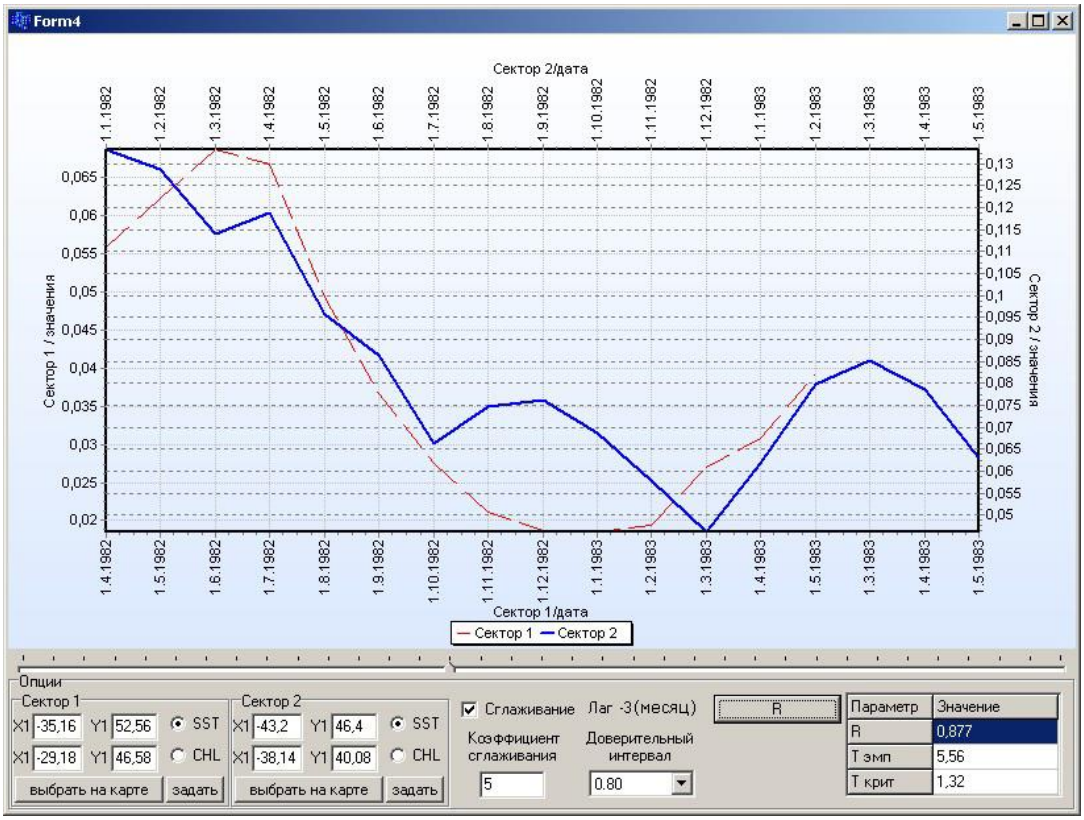



Рис. 3. Окно «Корреляция секторов».
Данные по температуре поверхностного слоя океана (с лагом 3 месяца).

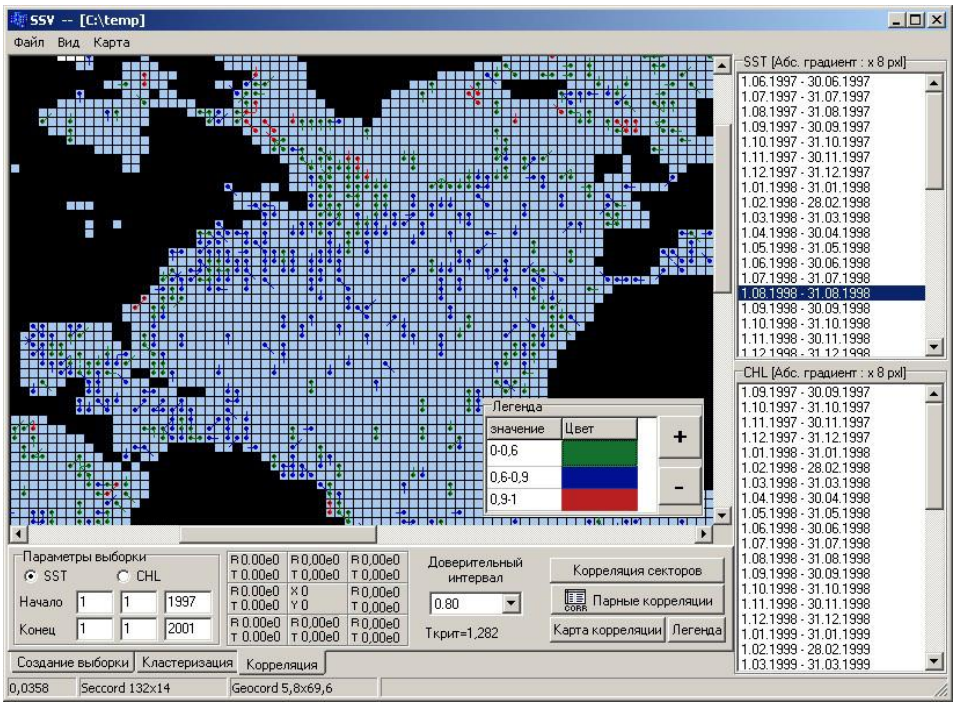



Рис. 4. Пример вывода информации - Окно «Карта корреляций».

Также при использовании ИС предусмотрена возможность вывода изображений в виде таблицы корреляции и корреляционной карты, полученной на основе вычислений парных корреляций, отображающих связи между соседними секторами (рис. 4).

В качестве примера можно рассмотреть последовательность процедур, которые необходимо реализовать, чтобы представить расчетные характеристики в удобном для анализа виде. На рис. 5 представлен алгоритм работы информационной системы.

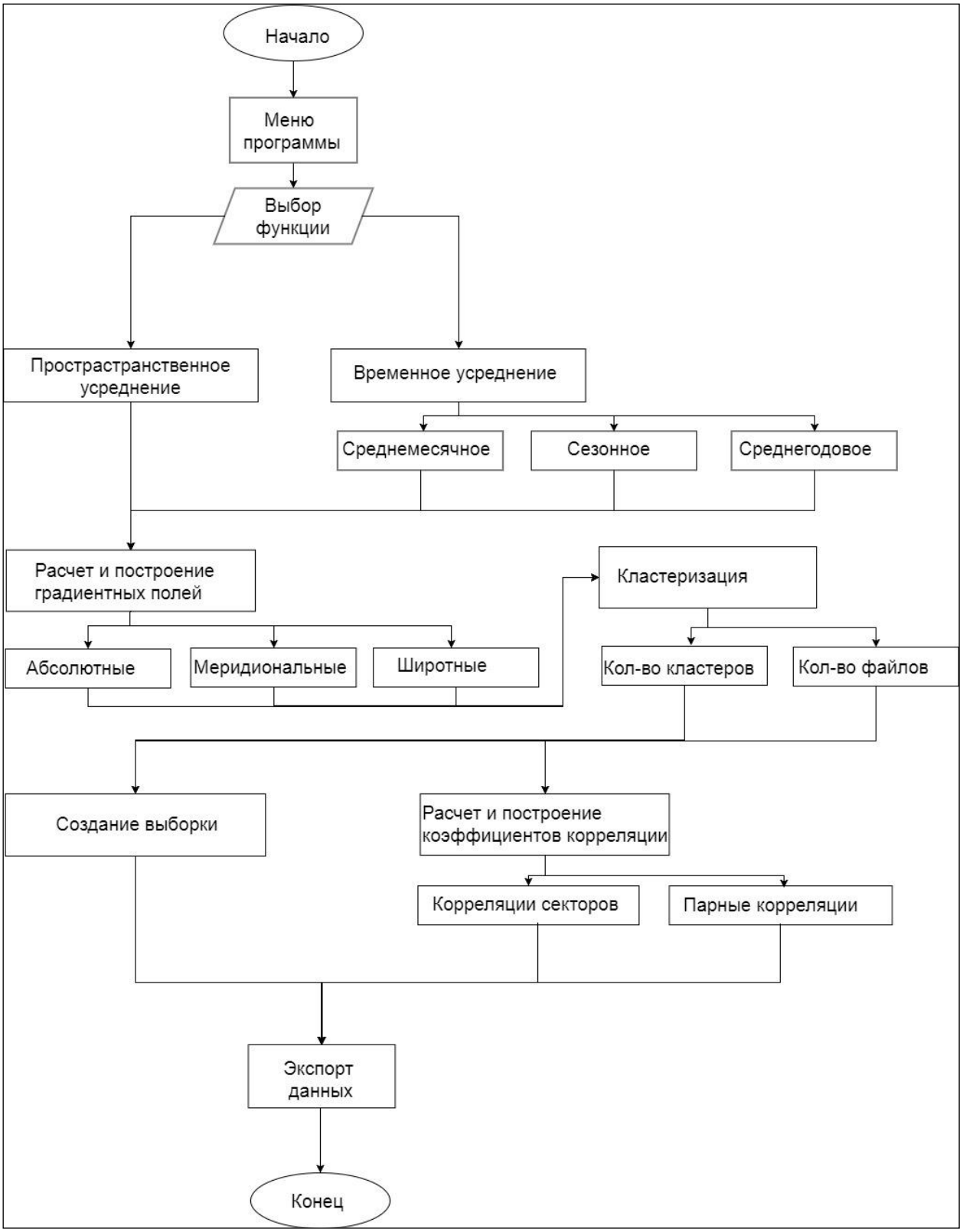



Рис. 5. Алгоритм реализации информационной технологи использования спутниковых данных.

**Заключение.** Одним из основных выводов по результатам работы можно сформулировать следующее: в последнее время наблюдается чрезмерное увлечение объединением разнородных измеряемых данных о состоянии климатической системы Земли, включая данные контактные и данные дистанционного зондирования. Глобальное потепление, Глобальные атмосферные осцилляции, Глобальные климатические модели используются для изучения фундаментальных закономерностей сложной нелинейной комбинации взаимодействия в системе атмосфера-океан-суша и дают возможность совершенствовать прогностические оценки состояния системы. Однако, нельзя оставлять без внимания региональные особенности проявления динамики энергоактивных зон океана и структурных особенностей локального взаимодействия процессов в отдельных районах океана и потоков тепла и влаги в тропосфере.

Расчеты градиентных характеристикповерхности Мирового океана на основе спутниковых данных позволяют выявлять зоны с различной динамической активностью, которые играют ключевую роль в определении местоположения зон с проявлением синергизма. Анализ распределениявеличины градиентных полей позволяет выявлять гетерогенные и гомогенные зоны в океане, а при соответствующем периоде усреднения получать степень динамичности таких зон. Это, в свою очередь, позволяет изучать процессы формирования естественной и аномальной структурной организации градиентных зон, масштабы ее изменчивости, и вектор развития динамических процессов, что может применяться приразработке математических прогностических моделей.

Статистический анализ градиентных показателейпозволяет получить степень вероятности развития событий на основе технологии обработки спутниковой информации и использовать данные для изучения глобальных и региональных особенностей поведения природных систем.

## ЛИТЕРАТУРА


[1] *Huang B.* Набор данных ERSSTv4. **https://www.ncdc.noaa.gov/data-access/marineocean-data/extended-reconstructed-sea-surface-temperature-ersst-v4** (дата обращения 10.06.2019).

[2] *Rayner N.A.* Набор данных HadISST. **https://www.metoffice.gov.uk/hadobs/hadisst** (дата обращения 10.06.2019).

[3] *Серых И.В.* О динамике и структуре глобальной атмосферной осцилляции в климатических моделях и реальности // Океанологические исследования. 2018. Т. 46, № 1. С. 14-28.

[4] *Taylor K.E., Stouffer R.J., Meehl G.A.* Overview of CMIP5 and the experiment design // Bull. Am. Meteor. Soc. 2012. Vol. 93. P. 485-498.

[5] *Huang B., Banzon V.F., Freeman E.* Extended reconstructed sea surface temperature version 4 (ERSST.v4). Part I: Upgrades and intercomparisons // J. Clim. 2015. Vol. 28, No. 3. P. 911-930.

[6] *Rayner N.A., Parker D.E., Horton E.B.* Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century // J. geophys. Res. 2003. Vol. 108. Article D144407.

[7] Физический океанографический распределенный активный архивный центр. **https://podaac.jpl.nasa.gov** (дата обращения 10.06.2019).

[8] Банк данных Oceancolor. **https://oceancolor.gsfc.nasa.gov** (дата обращения 10.06.2019).

[9] HDF спецификация и руководство разработчика. **http://gis-lab.info/docs/hdf/SpecDevG.pdf** (дата обращения 10.06.2019).

[10] *Адаменко О.С., Анкудинов А.А., Картушинский А.В.* Программные средства для исследования структурной организации полей температуры и хлорофилла поверхностного слоя океана на основе спутниковых данных // Сборник материалов VII Всероссийской научно-технической конференции «Молодёжь и наука» **http://conf.sfu-kras.ru/sites/mn2011/thesis/s3/s3_001.pdf** (дата обращения 10.06.2019).

[11] *Kartushinsky A. V.,Sidorenko A. Y.* Analysis of the variability of temperature gradient in the ocean frontal zones based on satellite data // Advanced Space Research. 2013. Vol. 52, N 8. P. 1467-1475.

[12] *Картушинский А.В., Кукоба Н.А.* Изучение градиентных полей поверхности земли по спутниковым данным // Вестник СибГАУ. 2015. Т. 16, № 3. С. 587-596.

# ВЕБ-СЕРВИС КЛАССИФИКАЦИИ СЕЙСМИЧЕСКИХ СОБЫТИЙ НА БАЗЕ СИСТЕМЫ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ APACHE SPARK

*Попов С.Е., Замараев Р.Ю.*
Институт вычислительных технологий СО РАН, Новосибирск

В статье описаны ключевые моменты процесса разработки сервиса для быстрой автоматической классификации сейсмических сигналов на основе диагностических шаблонов. Представлены программные решения для предварительной обработки сигнала и алгоритмизации параллельных вычислений на математической модели выработки конечных заключений с использованием базы рейтингового голосования. Показаны возможности интеграции таких решений с системой распределенных вычислений Apache Spark. Проведены тесты производительности алгоритма классификации для набора суточных сигналов в различных программных средах.
*Ключевые слова: веб-сервис, распределенные вычисления, Apache Spark, классификация сейсмических событий.*

**Введение.** Региональный мониторинг и анализ региональной геодинамической ситуации характеризуется сложностью решаемых на этом направлении задач. Причина в том, что наряду с мощными возмущениями из известных очаговых зон приходится анализировать и классифицировать разнородный поток событий, среди которых промышленные взрывы различной мощности и глубины заложения, региональные и местные сейсмические события. В горнопромышленных регионах функционирует большое количество предприятий, регулярно проводящих массивные взрывные работы. В общей сложности за год может регистрироваться более 2,5 тысяч сейсмических событий. Накапливаются огромные массивы информации, где только для одной станции недельный пул суточных записей составляет около 150Мб (≈174 млн. отсчетов), с нескольких – более 1 Гб (1 млрд отсчетов). Учитывая данный фактор, процесс детектирования и классификации различных возмущений в наборе даже суточных записей с нескольких станций требует значительных вычислительных ресурсов и затрат времени. Что подтверждается анализом большинства исследований в области обработки сейсмических сигналов [1-17]. Однако в большинстве исследований все апробации алгоритмов на реальных сигналах осуществляются на малых таймфреймах размером не более 60-80 секунд, рассматриваются отрезки сигнала с априори достоверной информацией о присутствии на них существенных (детектируемых) возмущений. Время работы представленных алгоритмов на полном сигнале (суточной записи) не приводится. В некоторых работах [16, 17] представлены так называемые быстрые алгоритм (Fingerprint And Similarity Thresholding (FAST)) для детектирования природных землетрясений. Однако тестирование на реальных сигналах недельного пула записей демонстрирует неутешительное время исполнения (около 1 часа 36 минут, для автокорреляционного метода – 9 дней (данные взяты с одной станции)). Таким образом при анализе даже 2-3 недельных таймфреймов с нескольких сейсмопостов время работы может увеличиваться до 1 дня, а с использованием другим методов и до месяца.

Рассматривая программные средства информационно-аналитического обеспечения процессов регионального сейсмического мониторинга можно обозначить основные требования: стабильное поступление массивов актуальных сейсмических данных; их оперативная обработка; анализ и выработка экспертных заключений с использованием быстрых алгоритмов. В мировой практике насчитывается большое количество программных решений, реализующих различные функции обработки и анализа сейсмической информации [18-25].

Однако функциональные возможности существующих программных решений не поддерживают классификацию сейсмических событий, основанную на обработке множества реальных суточных записей, полученных с разных станций наблюдений. Обработка данных ведется в ручном режиме, с последовательной загрузкой файлов, и выделением мелких таймфреймов интересующего события. Интегрированные в такие программные средства алгоритмы

классификации не поддерживают запуск в параллельном режиме обработки полного временного отрезка сигнала. В большинстве случаев программные решения представляют собой статические приложения, ориентированные на работу со специализированной аппаратной частью.

Учитывая выше изложенное, возникает актуальная задача разработки программного обеспечения для классификации сейсмических событий, поддерживающего высокопроизводительную обработку больших массивов данных и открытый доступом к ним на базе веб-технологий.

**Исходные данные и шаблоны.** Источником сейсмических данных служит региональная сеть из 8 станций с международными кодами: (ASR1, ELT, BRCR, KEM, LUZB, NVS, SALR, TASR). Сигналы поступают в формате miniSEED, данные предоставляются по трем каналам (например, EHE, EHN, EHZ).

Суточная запись с каждого канала содержит около 8,5 млн отсчетов (замеров с датчиков) и время каждого отсчета, т.е. (24 часа) х (3600 сек в часе) х (100 количество отсчетов в секунде (sample rate)). Начальные записи в каналах могут быть сдвинуты относительно начала суток (00:00:00). Для их синхронизации выбирается самый поздний по времени начальный отсчет (по максимальному времени от начала), и от этого времени извлекаются все значения каждого канала. Далее, выбирается минимальная длина из получившихся массивов (по минимальному времени от конца сигнала), оставшиеся массивы «обрезаются справа» до этой длины. Таким образом получаем матрицу со значениями типа Double

$$\mathrm{CH} = \{ch_{ij}, i \in [0, L_{ch}] \in \mathrm{Z}, j \in [0, 2] \in \mathrm{Z}\}, \quad (1)$$

где $L_{ch}$ – длина массива данных для каждого синхронизированного канала (в среднем 8,3-8,5 элементов); $j$ – номер канала.

Типовые шаблоны построены на примере сигналов с одной станции (BRCR). Использовались сейсмограммы (часть суточного сигнала) 35 достоверных промышленных взрывов и 19 региональных землетрясений с магнитудами от 1,8 до 2,4 балла. Длина расчетного окна в среднем берется $m = 6145$ отсчетов или 61,45 секунды при частоте дискретизации сигнала (sample rate) 100 Гц.

Для каждой сейсмограммы вычисляются значения характеристической функции (см. раздел алгоритм классификации). Из них для совокупностей взрывов и землетрясений были выделяются по три шаблона: средний и две его границы: для взрывов Blast (Blast±S), для землетрясений EarthQuake (EarthQuake±S). Используя выражение $f(t) = A(t/Tn^2)^n \exp(n - t/Tn) + \varepsilon(t)$, где $\mathrm{A}, \mathrm{T}, \mathrm{n}$ – числовые параметры (A=1 T=3800 n=44, подбираются индивидуально, для станций в зависимости от расстояния до очаговых зон, $\varepsilon(t)$-вектор случайных чисел, соответствует статистическому распределению модели сигнала «Белый шум» (WhiteNoise)) добавляются абстрактные шаблоны. Они показывают последовательное прохождение сейсмического возмущения со сдвигом 100 отсчетов (1 сек.) через расчетное окно. Таким образом получаем шаблоны со следующими названиями: на входе «WaveFront-I», «WaveFront-II» и «WaveFront-III»; на выходе «WaveRear-I», «WaveRear-II» и «WaveRear-III»; в середине окна со смещениями влево и вправо «WaveMiddle», «WaveLeft» и «WaveRight» соответственно. Всего выделяется 16 шаблонов – матрица со значениями типа Double ($C_{ij}^t, i \in [0, 6144] \in \mathrm{Z}, j \in [0, 15] \in \mathrm{Z}$).

**Алгоритм классификации.** Алгоритм классификации является оригинальной разработкой авторов [27]. Алгоритм позволяет анализировать полные трехкомонентные суточные сигналы.

На вход алгоритму подается сформированная матрица (1) ($\mathrm{CH} = \{ch_{ij}, i \in [0, L_{ch}] \in \mathrm{Z}, j \in [0, 2] \in \mathrm{Z}\}$). Задается скользящее окно размером в $m = 6145$ отсчетов, с шагом сдвига $step = 100$ отсчетов. На каждом шаге формируется сейсмограмма в виде матрицы $\mathrm{X} = \{x_{ij}, i \in [0, m] \in \mathrm{Z}, j \in [0, 2] \in \mathrm{Z}\}$, содержащая часть сигнала CH (1). Алгоритм определяет (классифицирует) тип сейсмограммы согласно шаблонам, следующим образом:

Шаг 1. Компоненты матрицы $X = \{x_{ij}\}$ заменяем квадратами размахов $sw_{i,j}$, что обеспечивает неотрицательность значений для дальнейших вычислений

$$sw_{i,j} = (x_{i,j} - x_{i+1,j})^2, i \in [0, m-1], j \in [0, 2]. \quad (2)$$

Шаг 2. Вычисляем матрицу весов (3) и, затем, матрицу энтропий (4)

$$q_{i,j} = \frac{sw_{i,j}}{\sum_{i=0}^{m-1} sw_{i,j}}, i \in [0, m-1], j \in [0, 2]. \quad (3)$$

$$E_{i,j} = -q_{i,j} l\, n\left(q_{i,j}\right), i \in [0, m-1], j \in [0, 2]. \quad (4)$$

Шаг 3. Вычисляем вектор обобщенной информации по трем каналам измерений

$$H_i = E_{i,0} + E_{i,1} + E_{i,2}, i \in [0, m-1]. \quad (5)$$

Шаг 4. Строим характеристическую функцию в расчетном окне (6). Данный процесс называется аккумулированием сигнала

$$\mathrm{C}_i^s = \sum_{l=0}^{i} H_l, i \in [0, m-1]. \quad (6)$$

За счет аккумулирования три компоненты сигналов приводятся к одномерной стационарной форме [29]. Стационарность модели (6) обеспечивает хорошую аппроксимацию по осредненным (сглаженным) данным (шаблонам с известными характеристиками).

Шаг 5. Добавляем справа к матрице $C_{ij}^t, i \in [0, 6144] \in \mathrm{Z}, j \in [0, 15] \in \mathrm{Z}$ вектор-столбец $\mathrm{C}^s$, получаем матрицу $C_{ij}, i \in [0, m-1], j \in [0,\ n], \mathrm{n} = 16$.

Согласно (6), все шаблоны находятся в одном метрическом пространстве. Полагая шаблоны признаками, а отсчеты объектами (независимыми наблюдениями), можем дополнить их набор выборочной характеристической функцией и вычислить аналог диагностической матрицы по Байесу путем стандартизации в объектах

$$S_{i,j} = \left(\mathrm{C}_{i,j} - \mu_i\right)/\sigma_i, \text{ где } \mu_i = \frac{1}{n}\sum_{j=1}^{n} C_{i,j} \text{ и } \sigma_i = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(C_{i,j} - \mu_i)^2}. \quad (7)$$

Теперь в матрице (7) можно оценивать подобие характеристической функции (6) каждому шаблону из набора, как расстояние между двумя признаками (одномерными векторами). Для этого фиксируем $j = 16$, и для каждого $S_{i,j}, j \in [0, 15]$ формируем пару с $S_{i,16}$, получаем 16 пар.

Шаг 6. Для каждой пары рассчитываем значения следующих статистических расстояний расстояний [26]: Bray-Curtis, Canberra, CityBlock, «Корреляция (нормированная)», «Евклидово», «Евклидово второй степени», «Минковского третьей степени», «Косинусное», а также их варианты с весовыми коэфициентами.

Получаем матрицу расстояний

$$D = \left\{D_{k,j}, k \in [0{,}11], j \in [0, 15]\right\}, \quad (8)$$

где весовой коэффициент $w_i = \begin{cases} 1, i \in [0, 2(m-1)/3] \\ 0, \text{ в остальных случаях} \end{cases}$.

Априорных суждений о преимуществах тех или иных дистанций не существует, поэтому в текущей версии алгоритма используются все, пригодные для номинальных признаков с различными вариациями [30].

Теперь можно реализовать простейшую систему голосования, в которой каждая дистанция имеет один голос. Каждый голос отдается шаблону с минимальной дистанцией (8) до выборочной характеристической функции. Простое суммирование голосов у каждого шаблона определяет его простой рейтинг.

Шаг 7. Преобразуем матрицу $D$ следующим образом: для каждого $k \in [0{,}11]$

$$D_{k,j} = \begin{cases} 1, D_{k,j} = \min(D_k), \\ 0, \text{ в остальных случаях.} \end{cases} \quad (9)$$

Шаг 8. Рассчитываем рейтинги $R_j = \sum_{k=0}^{11} D_{k,j}$ (10) и находим максимум $\{R_j\}$

Шаг 9. Формируем заключение классификации по следующей схеме, назовем ее рейтинговым голосованием:

а) если не существует единственного максимума, то заключение «**не определено (undefined)**»;

б) если единственный максимум больше 10, то заключение «**строгое (strictly)**» соответствие шаблону;
в) если единственный максимум больше 8, но меньше 11, то заключение «**нестрогое (not strictly)**» соответствие шаблону;
г) если единственный максимум меньше 9, то заключение «**возможное (perhaps)**» соответствие шаблону.

Таким образом, сдвигая окно от начала сигнала в конец, алгоритмом детектируются и идентифицируются согласно шаблону любые значимые возмущения суточного таймфрейма.

**Оптимизация алгоритма.** Для уменьшения времени работы программной реализации алгоритма и его адаптации к запуску в среде массово-параллельного исполнения заданий были проведены перечисленные далее действия (итерации).

1. Предварительно рассчитываем элементы $sw_{i,j}$ (2) для трех каналов суточной записи. Получаем матрицу $swO_{ij}, i \in [0, L_{ch} - 1]$. Заранее рассчитываем $\sum_{i=1}^{m-1} sw_{i,j}$ (2) для каждого шага сдвига. Получаем массив

$$sw_sums_{j,s} = sw_sum_{j,s-1} \pm \sum_{l=0}^{100} swO_{100s \pm l,j}, s \in \left[1, \frac{L_{ch}}{100}\right],$$

где $sw_sum_{j,0} = \sum_{i=0}^{m-1} sw_{i,j}$.

Массивы $swO$ и $sw_sums$ являются общедоступными константами для всех заданий в среде Apache Spark, и передаются при помощи специального объекта **Broadcast**. Данная оптимизации позволит существенно сократить время расчета формул (2) и (3), т.к. на каждом сдвиге окна вычисляются суммы только предыдущих и последующих 100 элементов матрицы $sw$.

2. Для матрицы $C^t$ рассчитываем сумму $CSum_i = \sum_{j=1}^{n} C_{i,j}^t$. $CSum_i$ объявляем общедоступной константой (**Broadcast**). Тогда на каждом шаге выражение $\mu_i = \frac{1}{n}\sum_{j=1}^{n} C_{i,j}$ в формуле (7) можно заменить на следующее $\mu_i = \frac{1}{n}(C_{i,16} + CSum_i)$, что также позволяет сократить количество операций с суммой элементов.

3. Используемые алгоритмы нахождения расстояний согласно [14] в своих расчетах содержат повторяющиеся выражения. Например, расстояние Bray-Curtis будет содержать выражение $S_{i,16} - S_{i,j}$, которое также есть в Canberra, или $\sum \left| S_{i,16} - S_{i,j} \right|$, $j \in [0, 15]$ в City Block. Расчет корреляции через ковариацию, позволит получить выражения: $\sum {S_{i,16}}^2, \sum {S_{i,j}}^2, \sum S_{i,16} S_{i,j}$, которые присутствуют в расчете евклидова и косинусного расстояний. Учитывая, что весовой коэффициент принимает значение 1 при двух третьих от общего количества итераций, а остальные значения равны нулю, то при расчете сумм в соответствующих расстояниях без весового коэффициента, необходимо зафиксировать значении суммы на номере итерации равном $2(m-1)/3$, и использовать ее при получении значения расстояния с весовым коэффициентом. Т.е. если евклидово расстояние есть $\sum {S_{i,16}}^2 + \sum {S_{i,j}}^2 - 2\sum S_{i,16} S_{i,j}, i \in [0, m-1]$, то евклидово с весовым коэффициентом будет вычисляться также, только для $i \in [0, 2(m-1)/3]$.

4. Анализ алгоритма классификации (шаги 1-9) показал независимость расчетов на каждой итерации сдвига расчетного окна. Этот факт означает, что заключение классификации получается, как единичное абстрактное значение одного из признаков (см. шаг 9, раздел **Алгоритм классификации**). Таким образом, возможно разделить всю суточную запись длинной $L_{ch}$ на части (**partitions**), и вычислять модель (2)-(10) параллельно. Назовем этот этап **Map**. Затем после завершения всех заданий Map, запускать процесс объединения полученных заключений, сортируя их по времени в начальном сигнале (получение карты классификаций). Данный этап назовем **Reduce**. Таким образом возможно организовать вычисления по классической схеме MapReduce.

Такого рода оптимизации позволяют значительно сократить время работы алгоритма,

т.к. при каждом сдвиге приходится выполнять одни и те же вычислительные операции на одних и тех же наборах данных по нескольку раз. Учитывая количество шагов (в среднем $\frac{L_{ch}}{step}$) и размер окна ($m - 1 = 6144$ отсчета) потери производительности могут быть существенными.

**Технологический стек.** Вычислительное ядро сервиса (**BACKEND**), который предлагается в настоящей статье, реализовано на базе Apache Spark API (Java), менеджера ресурсов Apache YARN и сервиса удаленного запуска заданий Apache Livy. Система обработки HTTP-запросов пользователя (**MIDDLEWARE**) функционирует под управлением NodeJS, Chart.js (ECMAScript 6 (ES6)). Графический интерфейс сервиса построен с применением библиотек React+Redux, Plot.ly, Semantic UI (ES6) под управлением сервера Nginx (**FRONTEND**). Операции ввода/вывода **BACKEND**- и **MIDDLEWARE**-компонентов реализуются на базе распределенной файловой системы HDFS (ОС Ubuntu 16.04).

**Описание веб-сервиса.** Разработанный веб-сервис логически разделен на три составляющие: **BACKEND**, **MIDDLEWARE**, **FRONTEND**.

**BACKEND**-компонент представлен java-классами, реализующими непосредственно расчетную часть алгоритма классификации и дополнительными классами, имплементирующими методы предварительной обработки сейсмического сигнала и работы с объектами Apache Spark API (рис .1).

**ClassificationProcessor** основной класс, который отвечает за запуск процесса классификации (рис. 2). Содержит подкласс **classify** наследующий объект **Function** Spark API для передачи его в функцию **map**. Подкласс **classify** реализует программный алгоритм классификации (классы **SignalProcessor** и **DistanceClassifier**), адаптированный для работы в распределенном режиме на узлах кластера. Класс **ClassificationProcessor** обеспечивает настройку среды исполнения (**Executor**) заданий посредством объекта **SparkContext**. В **ClassificationProcessor** реализована процедура размещения неизменяемых объектов-констант (**Broadcast**), содержащих предварительно рассчитанные данные (см. **Оптимизация алгоритма**) в классах **TemplateProcessor** и **MiniSEEDProcessor**. Данные константы доступны со всех узлов кластера в общей памяти текущего контекстного объекта.

Данные шаблонов хранятся в HDFS в CSV-файле. Класс **TemplateProcessor** поддерживает методы чтения, обработки данных CSV и построение массива $C_{ij}^{t}$ и $\sum_{j=1}^{n} C_{i,j}$ для каждого $i$-го отсчета, для добавления в пул общедоступных констант (**Broadcast**, см. **Оптимизация алгоритма**).

Класс **MiniSEEDProcessor** содержит методы работы с файлами miniSEED-формата при помощи библиотеки iris-WS.jar. Она позволяет открывать, декодировать и считывать данные из файлов каналов сейсмических записей. **MiniSEEDProcessor** реализует процедуру синхронизации каналов по времени и формирования константы **Broadcast** для матрицы CH (1), Так же в данном классе рассчитываются вспомогательные данные: длинна сигнала и метаинформация по каждому каналу (название, частота дискретизации, начальное/конечное время записи).

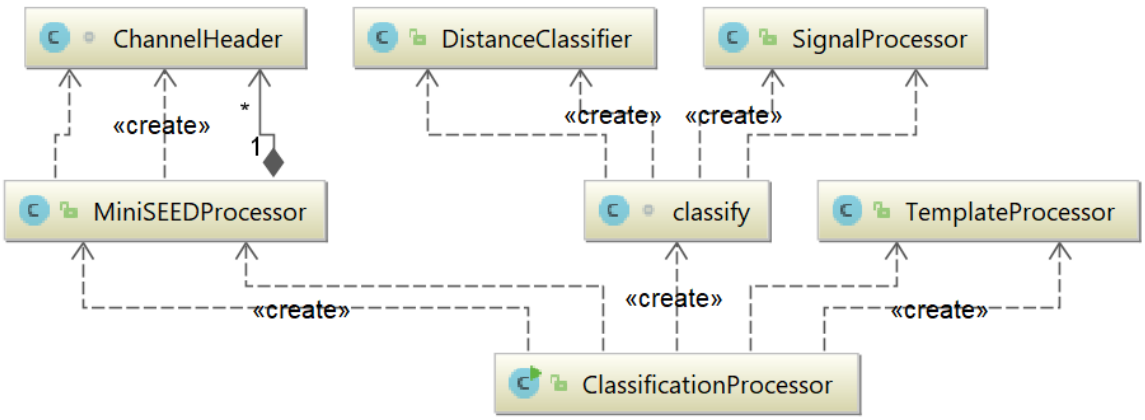



Рис. 1. Диаграмма объектов **BACKEND**-компонента.

```
// Обработчик шаблона
TemplateProcessor templateProcessor = new TemplateProcessor();
templateProcessor.process(templatesFile);
// Обработчик каналов
MiniSEEDProcessor miniSEEDProcessor = new MiniSEEDProcessor(...);
miniSEEDProcessor.process(chFiles[0], chFiles[1], chFiles[2], true);
// Настройка контекста запуска
SparkConf sparkConf = new SparkConf();
JavaSparkContext sc = new JavaSparkContext(sparkConf);
// Размещение общедоступных данных
Broadcast<double[][]> CBroadcast = sc.broadcast(templateProcessor.C);
...
// Запуск задания
List<String> classificationMapParts = sc.parallelize(IntStream.range(0,partitionCount)
         .boxed().collect(Collectors.toList()), partitionCount)
         .map(new classify(...,..., CBroadcast,...
        )).collect();
...
// Имплементация объекта Function для метода map (параллельно для каждого задания)
class classify implements Function<Integer, String> {
...
public classify(...,...,Broadcast<double[][]> CBroadcast,...) {
        this.CBroadcast = CBroadcast;}

    @Override
    public String call(Integer core) {

        SignalProcessor signalProcessor = new SignalProcessor(...);
        DistanceClassifier distanceClassifier = new DistanceClassifier(...);
        ...
        // Доступ к общим данным внутри задания SparkContext
        double[][] C = CBroadcast.value();
        ...
        ...
        // Расчет алгоритма классификации
        double[][] D = signalProcessor.process(ch1RawData, ch2RawData, ch3RawData, C, CSums);
        conclusionResult = distanceClassifier.process(D);
        ...
        return conclusionResult // Результат в виде JSON-строки
...
```

Рис. 2. Фрагмент кода настройки контекста Spark и запуска расчетного задания класса **ClassificationProcessor.**

```
{
  ...
  // Номер части (соответствует номеру задания Spark)
  "partition004": {...},
  "partition005": {
    "undefined": {...},
    // Тип заключения классификации
    "strictly": {
      // Координаты для отображения на карте классификаций
      "x": [15952,16030,16031],
      "y": [1, 1, 1],
      // Временная метка, соответствует окну сдвига в 1 сек.
      "times": [
        "2013-10-08 04:26:01",
        "2013-10-08 04:27:19",
        "2013-10-08 04:27:20"
      ]
    },
    "notstrictly": {...},
    "perhaps": {...}
  },
  "partition006": {...},
  "partition007": {...},
  // Названия каналов классифицируемого сигнала
  "channel1": "AN.BRCR.81.EHE.D.2013.281",
  "channel2": "AN.BRCR.81.EHN.D.2013.281",
  "channel3": "AN.BRCR.81.EHZ.D.2013.281",
  // Параметры фильтрации, используются на стороне FRONTEND
  "filter": {
    "startTime": "", "endTime": "", "isApply": false,
    "onlyBlastStrictly": false
  },
  // Начальное и конечное время сигнала, используется для фильтрации
  // на стороне FRONTEND
  "signalStartTime": "2013-10-08 00:00:09",
  "signalEndTime": "2013-10-09 00:00:01"
}
```

Рис. 3. Результат классификации суточной записи сейсмического сигнала в виде карты классификаций в формате JSON

Результатом работы метода **classify** является JSON-файл (рис. 3), хранящийся в HDFS, где каждая часть (выполненное задание) идентифицируется ключом **partition**, содержащим свойства того или иного заключения карты классификаций.

**MIDDLEWARE-**компонент реализован на базе объектов языка ES6. Имплементирует методы программного каркаса библиотеки NodeJS API. Функционирует как прокси-уровень между **FRONTEND** и **BACKEND**, выступает в качестве обработчика пользовательских HTTP-запросов для вызовов их методов.

```
{
  // Основной java-класс
  "className": "org.myapp.seismatica.classifiers.ClassificationProcessor",
  // Путь к файлу программы
  "file": "hdfs://10.101.81.203/user/jars/seismatica/seismatica-classifier-1.0.jar",
  "name": "Seismatica - Classifier",
  "args": [
    username, // Имя пользователя
    "DistanceClassifier", // Название классификатора
    chFilesArg, // Массив путей к файлам каналов
    templateFile, // Путь к файлу шаблонов
    "100", // Шаг сдвига окна
    "8" // Количество задач (исполняются параллельно)
  ],
  "conf": {
    "spark.executor.instances": "4", // Количество Executor-объектов (работают параллельно)
    "spark.task.cpus": "1", // Количество CPU на одно задание
    "spark.executor.cores": "2", // Количество задач на один Executor
    "spark.executor.memory": "2g", // Выделяемая память для одного Executor
    "spark.driver.memory": "3g", // Выделяемая память для управляющей программы
    "spark.driver.extraClassPath": "/mnt/hdfs/user/jars/seismatica/*", // Путь к java-классам
    "spark.executor.extraClassPath": "/mnt/hdfs/user/jars/seismatica/*" // Путь к java-классам
  }
}
```

Рис. 4. JSON-объект (параметр **body**) в POST-запросе к сервису Apache Livy на удаленный запуск Spark-задания.

```
function getResultAsJSON(resultFile) {

    // Чтение файла результата из HDFS, преобразование в JSON-объект
    let result = JSON.parse(fs.readFileSync(resultFile, 'utf8'));
    // Получение частей
    let partitions = Object.keys(result).filter(key => key.includes('partition'))
        .filter(key => Object.keys(result[key]).length > 0).sort();

    // Объединение частей (для каждого из заключений классификации)
    let reducedResult = {
        undefined: {
            x: reduceResultPartition(result, partitions, 'undefined', 'x'),
            y: reduceResultPartition(result, partitions, 'undefined', 'y'),
            times: reduceResultPartition(result, partitions, 'undefined', 'times')
        },strictly: {...}, notstrictly: {...}, perhaps: {...}
    };

    // Объединение временных меток (для каждого из заключений классификации)
    reducedResult.times = [...reducedResult.undefined.times,
    ...reducedResult.strictly.times,...reducedResult.notstrictly.times,
    ...reducedResult.perhaps.times].sort();

    // Добавление остальных параметров в JSON-объект
    reducedResult.channel1 = result.channel1;
    reducedResult.channel2 = result.channel2;
    reducedResult.channel3 = result.channel3;
    ...
    return reducedResult;
}

// Функция редукции одного из элементов из каждой части JSON-объекта
function reduceResultPartition(result, resultParts, conclusion, conclusionKey) {
    return resultParts.reduce((accumulator, currentValue, currentIndex, array) => {
            return [...accumulator, ...result[currentValue][conclusion][conclusionKey]];
        }, []);
}
```

Рис. 5. Фрагмент кода метода **getResultAsJSON**.

Для запуска задания на стороне **BACKEND**-компонента **MIDDLEWARE**-компонент использует объекты **router** и **request,** предоставляемые стандартной библиотекой NodeJS Express API. Данные объекты перенаправляют POST-запросы со стороны **FRONTEND**-компонента сервису Apache Livy (http://livy-server:8998/batches). POST-запрос содержит настройки параметров среды Apache Spark, менеджера ресурсов Apache YARN, параметров расчетного модуля и управляющей программы-драйвера (рис. 4).

Ответом на POST-запрос является JSON-объект, состоящий из частей (partition) по количеству параллельно выполненных задач. Для их объединения используется метод **getResultAsJSON** (рис. 5).

Для простоты взаимодействия объектов **MIDDLEWARE**-компонента с распределенной файловой системой HDFS, она монтируется, как обычная директория в операционной системе с использованием программного интерфейса FUSE (Filesystem in Userspace). Это позволиляет осуществлять операции ввода/вывода на базе методов объекта **NodeJS FileSystem** и обрабатывать соответствующие GET/POST-запросы со стороны **FRONTEND** непосредственно в **MIDDLEWARE**-компоненте**.**

**FRONTEND**-компонент представлен программными объектами-классами отвечающими за графический интерфейс и взаимодействие с пользователями веб-сервиса.

Основной функцией **FRONTEND** является удаленный запуск расчета алгоритма классификации в системе Apache Spark и визуализация результатов в виде карты классификаций (рис. 7). Программные объекты **FRONTEND** взаимодействуют с компонентами **MIDDLEWARE** через протокол HTTP, с помощью библиотеки axios-js (рис. 6).

Веб-сервис (рис. 7) поддерживает следующие функциональные возможности: добавление/удаление файлов каналов сейсмостанции в формате miniSEED (сохраняются в HDFS, доступны со всех узлов кластера) и формирование сейсмического сигнала; запуск процесса классификации для выбранного сигнала, получение результатов с сервера, удаление результата из HDFS; фильтрация отображения результата по времени; построение карты классификации выбранного результата; аутентификации на базе системы Apache Hue; просмотр данных каналов в соответствии с выбранным уровнем масштабирования карты классификаций; поддержка русского/английского языка интерфейса.

```
class SubmitButton extends React.PureComponent {
  ...
  onSubmitHandler() {
     ...
     // Передача параметров HTTP-запросу
     let templateFile = '/mnt/hdfs/user/seismo_usr/seismatica' +
                        ' /templates/SeisPatterns.csv';
     axios.get('http://' + settings.MIDDLEWARE_IP + ':'
         + settings.MIDDLEWARE_PORT + '/' + settings.SPARK_PROCESSOR
         + '/sparkSubmit?' + 'username=' + this.props.login.userName + '&'
         + 'chFiles='+ signal.channels[0] + ',' + signal.channels[1] + ','
         + signal.channels[2] + '&' + 'templateFile=' + templateFile
     // Выполнение GET-запроса, получение ответа и его обработка
     ).then((response) => {
        this.props.onSubmitSignal(response.data.id);})}

     // Размещение элемента на странице
     render() {
        return (
         <Button floated='right' icon labelPosition='left' size='small'
                 onClick={this.onSubmitHandler.bind(this)}
                 ...
         </Button>);}
```

Рис. 6. Фрагмент кода отправки задания в систему Apache Spark в компоненте **FRONEND** (SubmitButton.js).

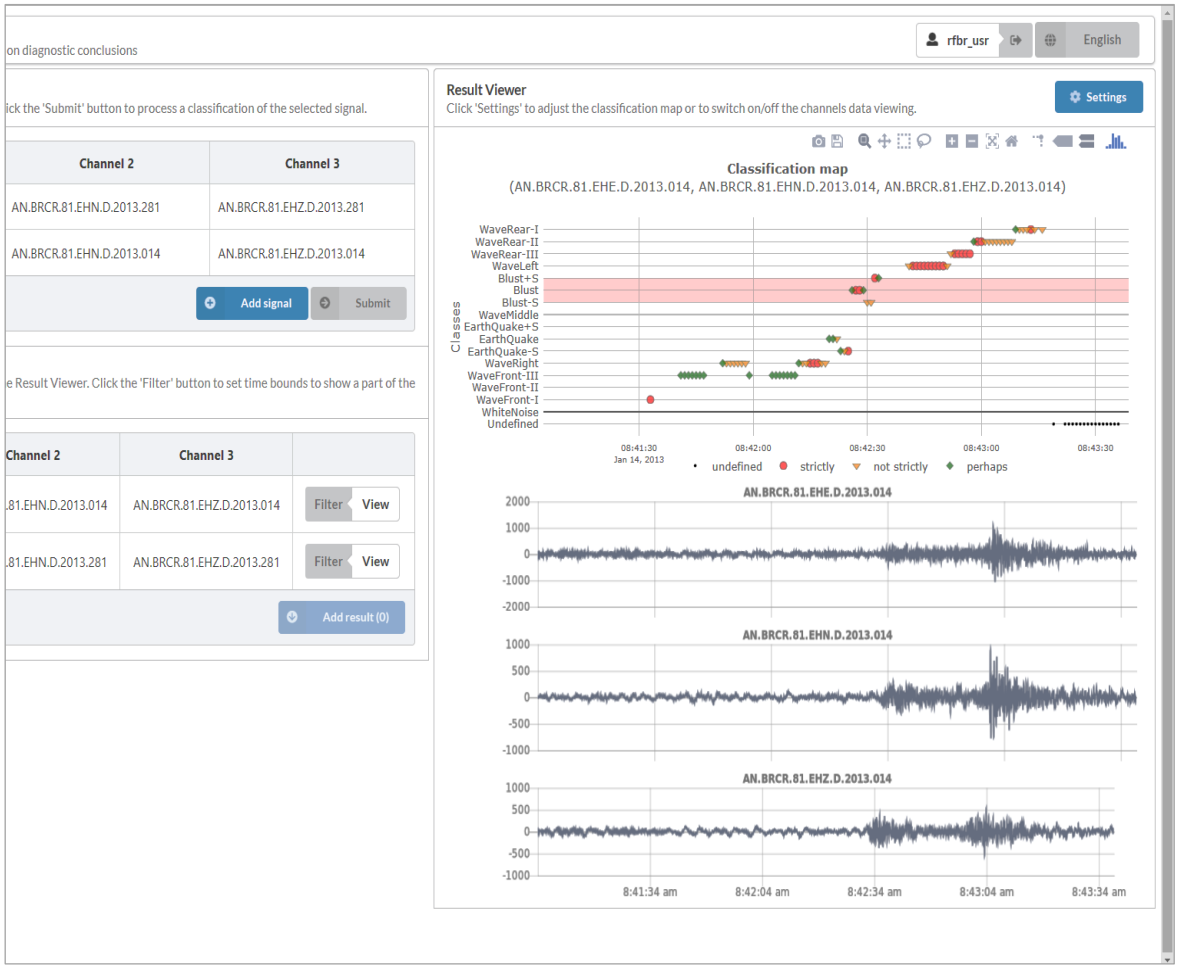



Рис. 7. Общий вид графического интерфейса веб-сервиса.

Тест производительности программной реализации представленного алгоритма классификации (время работы).

| | **Java Spark API** | **Python Spark API** | **Matlab** | **Java** |
|---|---|---|---|---|
| | **Программный код запускается параллельно на 30 вычислительных ядрах, с разбивкой на partitions** | | **Программный код запускается последовательно, без применения Apache Spark API** | |
| | **Суточная запись** | | | |
| **Время работы (сек.)** | 27 | 34 | 3574 | 801 |
| **Время работы (сек.)** | **Недельный таймфрейм** | | | |
| | 224 | 283 | 25145 | 5599 |

Суточная запись, синхронизированная по каналам в среднем, составляла 8355839 отчетов с интервалом в 10 мс или 83558 сдвигов. Аппаратное обеспечение: 2 сервера (AMD Ryzen 1700 (8+8 cores (Simultaneous Multi-Threading)) 3.2 GHz, 16Gb RAM, 1Gb/s скорость передачи д анных между серверами). Локальный тест проводился на одном сервере.

**Тест по оценке производительности.** Тест, направленный на оценку производительности системы, производился на примере запуска процесса классификации набора суточных записей со станции BRCR. Проведено 150 запусков. Файлы сигналов (3 канала) не повторялись. В таблице указано среднее время работы алгоритма. Представлены 4 программные реализации

алгоритма в средах Matlab, Java (консольное приложение) – локальный тест, Java, Python (Spark API приложение) – распределенный тест. Фиксировалось только время расчета от подачи входных параметров (файлы каналов, файл шаблонов, параметры настройки Spark и др.), до получения JSON-файлы карты классификаций.

Авторами проведены сравнения с протоколами наблюдений службы геофизического мониторинга Кемеровской области. Полученные результаты в 95% случаев (выборка 2013 года, всего около 500 событий (промышленный взрыв), с двух станций) полностью совпадали по типам заключений.

**Заключение.** Разработан веб-сервис для детерминирования и идентификации сейсмических событий с возможностью построения визуальных карт классификаций. Графическое представление элементов (заключений), распределенных во времени прохождения волны в заданном интервале, описывается классами возмущений сейсмического сигнала на базе характеристических функций (шаблонов). Алгоритм классификации, применяемый в вычислительном ядре сервиса успешно адаптирован для его запуска в распределенном режиме реализации на основе массово-параллельного исполнения заданий в среде Apache Spark. Проведенные тесты производительности показали, что предложенный подход к оптимизации математической модели и программной реализации алгоритма, позволяет применять его для потоковой обработки сигнала в виду очень малого времени выполнения программного кода.

В работе продемонстрированы механизмы интеграции современных веб-технологий построения интернет-приложений с элементами кластерной инфраструктуры. По мнению авторов, такой подход позволит разрабатывать подобные решения и в других областях научно-технической деятельности.

*Исследование выполнено при финансовой поддержке РФФИ (грант № 18-07-00013-а).*

## ЛИТЕРАТУРА

[1] *Scarpetta S., Giudicepietro F., Ezin E. C., Petrosino S., Del Pezzo E., Martini M., Marinaro M.* Automatic Classification of Seismic Signals at Mt. Vesuvius Volcano, Italy, Using Neural Networks // Bulletin of the Seismological Society of America, 2005. Vol. 95, No. 1. P. 185-196.

[2] *Benbrahim M., Daoudi A., Benjelloun K., Ibenbrahim A.* Discrimination of Seismic Signals Using Artificial Neural Networks // Proceedings of world academy of science, engineering and technology. 2005. Vol. 4. P. 4-7.

[3] *Diersena S., Leeb E-J., Spearsc D., Chenb P., Wanga L.* Classification of Seismic Windows Using Artificial Neural Networks // Procedia Computer Science. 2011. Vol. 4. P. 1572-1581.

[4] *Hamer R.M., Cunningham J.W.* Cluster analyzing profile data confounded with interrater differences: A comparison of profile association measures. Applied Psychological Measurement. 1981. Vol. 5. P. 63-72.

[5] *Kedrov E.O., Kedrov O.K.* Spectral time method of identification of seismic events at distances of 15°-40° // Izvestiya, Physics of the Solid Earth. 2006. Vol. 42. No. 5. P. 398-415.

[6] *Langer H., Falsaperla S., Powell T., Thompson G.* Automatic classification and a-posteriori analysis of seismic event identification at Soufrière Hills volcano, Montserrat // Journal of Volcanology and Geothermal Research. Elsevier 2006. Vol. 153 (1). P. 1-10.

[7] *Lyubushin Jr. A.A., Kaláb Z., Častová N.* Application of Wavelet Analysis to the Automatic Classification of Three-Component Seismic Records // Izvestiya, Physics of the Solid Earth, 2004. Vol. 40. No.7. P. 587-593.

[8] *Musil M., Pleginger A.* Discrimination between Local Microearthquakes and Quarry Blasts by Multi-Layer Perceptrons and Kohonen Maps // Bulletin of the Seismological Society of America, 1996. Vol. 86. No. 4. P. 1077-1090.

[9] *Ryzhikov G.A., Biryulina M.S., Husebye E.S.* A novel approach to automatic monitoring of regional seismic events // IRIS Newsletter. 1996. Vol. XV, No. 1. P. 12-14.

[10] *Shimshoni Y., Intrator N.* Classification of Seismic Signals by Integrating Ensembles of Neural Networks // IEEE transactions on signal processing. 1998. Vol. 46. No. 5. P. 1194-1201

[11] *Ryan T.M., Borisov D., Lefebvre M., Tromp J.* SeisFlows – Flexible waveform inversion software // Computers & Geosciences. 2018. Vol. 115. P. 88-95.
[12] *Lesage P.* Interactive Matlab software for the analysis of seismic volcanic signals // Computers & Geosciences. 2009. Vol. 35. Iss. 10. P. 2137-2144.
[13] *Jiang W., Yu H., Li L., Huang L.* A Robust Algorithm for Earthquake Detector // Proceedings of the 15 World Conference on Earthquake Engineering. Lisbon. Portugal. 2012.
[14] *Álvarez I., García L., Mota S. et al.* An Automatic P-Phase Picking Algorithm Basedon Adaptive Multiband Processing // IEEE Geosci. and Remote Sens. Lett. 2013. Vol. 10, No. 6, P. 1488-1492.
[15] *Guilherme M., António R.* A neural network seismic detector // IFAC Proceedings Volumes. 2009. Vol. 42, Iss. 19. P. 304-309
[16] *Clara E.Y., Ossian O.R., Karianne J.B., Beroza G.C.* Earthquake detection through computationally efficient similarity search // Science Advances. 2015. Vol. 1. P. e1501057(1-13)
[17] *Paul B.Q., Pierre G., Yoann C., Munkhuu U.* Detection and classification of seismic events with progressive multichannel correlation and hidden Markov models // Computers & Geosciences. 2015. Vol 83. P. 110-119
[18] IRIS. Incorporated Research Institutions for Seismology. **https://www.iris.edu/hq** (дата обращения: 04.05.2018).
[19] *Romero J.E., Titos M., Bueno Á. et al.* APASVO: A free software tool for automatic P-phase picking and event detection in seismic traces // Computers & Geosciences. 2016. Vol. 90. Part A. P. 213-220.
[20] GeoSeisQC. **http://www.geoleader.ru/index.php/ru/produkty-ru/geoseicqc** (дата обращения: 07.05.2018).
[21] ZETLAB Детектор STA/LTA. **https://zetlab.com/shop/programmnoe-obespechenie/funktsii-zetlab/analiz-signalov/detektor-sta-lta** (дата обращения: 07.05.2018).
[22] Stratimagic. **http://www.pdgm.com/products/stratimagic** (дата обращения: 07.05.2018).
[23] Разработка и создание Грид-приложений для решения прикладных задач геофизики. (грант РФФИ № 10-07-00491-а). **http://www.rfbr.ru/rffi/ru/project_search/o_49145** (дата обращения: 07.05.2018).
[24] Использование слабо связанных вычислительных систем для решения обратных задач геофизики (грант РФФИ № 11-05-00988-а). **http://www.rfbr.ru/rffi/ru/project_search/o_43212** (дата обращения: 07.05.2018).
[25] Разработка GRID-системы и вычислительных сервисов для исследования геодинамических пространственно-временных процессов по данным ДЗЗ» (грант РФФИ № 11-07-12045-офи). **http://www.rfbr.ru/rffi/ru/project_search/o_46676** (дата обращения: 07.05.2018).
[26] Distance computations. **https://docs.scipy.org/doc/scipy/reference/spatial.distance.html** (дата обращения: 11.05.2018).
[27] *Замараев Р.Ю., Попов С.Е., Логов А.Б.* Алгоритм классификации сейсмических событий на основе энтропийного отображения сигналов // Физика Земли. 2016. № 3. С. 31-37.
[28] *McKay D.* Information Theory, Inference, and Learning Algorithms // Cambridge: Cambridge University Press, 2003. 631 P.
[29] *Kortström J., Uski M., Tiira T.* Automatic classification of seismic events within a regional seismograph network // Computers & Geosciences. 2016. No. 87. P. 22-30.
[30] *Gan G., Ma C., Wu J.* Data clustering: theory, algorithms, and applications (ASA-SIAM series on statistics and applied probability) // Society for Industrial and Applied Mathematics Philadelphia. PA. USA. 2007. 451 P.

# ОЦЕНКА СНЕГОЗАПАСОВ ПО ДАННЫМ МИКРОВОЛНОВОГО РАДИОМЕТРА AMSR-2

*Ромасько В.Ю.*[(1)(2)], *Гордеев И.Н.*[(2)(3)], *Бураков Д.А.*[(2)(4)]
[(1)] Сибирский центр ФГБУ «НИЦ «Планета», Новосибирск
[(2)] ФГБУ «Сибирский региональный научно-исследовательский гидрометеорологический институт», Новосибирск
[(3)] ФГБУ Среднесибирское управление по гидрометеорологии и мониторингу окружающей среды, Красноярск
[(4)] ФГБОУ ВО Красноярский государственный аграрный университет, Красноярск

Рассмотрена стандартная методика оценки снегозапасов по данным микроволнового радиометра AMSR-2, установленного на спутнике GCOM-W1. Проведён сравнительный анализ результатов оценки снегозапасов с наземными данными маршрутной снегосъёмки, показавший наличие существенных ограничений для применения спутниковых данных в гидрологической практике.
*Ключевые слова: запасы воды в снеге, дистанционное зондирование, микроволновая съёмка, AMSR.*

**Введение.** Снежный покров со спутников наблюдается практически с первых попыток взглянуть на Землю с орбиты. И картографирование снежного покрова на регулярной основе ведётся уже с 1966 г. [1]. Однако, задача определения высоты снежного покрова или запасов воды в снеге оказалась существенно сложнее. Первая предложенная методика была опубликована только в 1982 г. [2] и рассчитана на использование данных микроволнового радиометра SMMR со спутника Nimbus 7. После ряда последующих работ [3-5], основанных на данных того же и последующих микроволновых радиометров, таких как SSM/I на КА DMSP и AMSR-E на КА Aqua, сформировался современный подход к оценке высоты снежном покрова и запасов воды в снеге по измерениям радиационной температуры в различных диапазонах сантиметровых волн.

Запасы воды в снеге являются ключевым параметром во всех прогностических гидрологических моделях стока, включающих моделирование снеготаяния. Прямая дистанционная оценка пространственного распределения запасов воды в снеге в пределах речного бассейна могла бы заменить ныне применяющиеся точечные (в масштабах бассейна) данные наземных маршрутных снегосъёмок и исключить процедуру их экстраполяции на весь бассейн.

**Описание метода.** В настоящее время на орбите действует КА GCOM-W1, имеющий на борту микроволновой радиометр AMSR-2, данные которого централизованно принимаются и обрабатываются в Японском космическом агентстве, и доступны для свободного использования. Этот радиометр имеет коническую развёртку с углом к вертикали 55°, гарантирующую постоянство размеров мгновенного поля зрения на местности. Главное параболическое зеркало диаметром 2 м обеспечивает хорошую фокусировку и малое мгновенное поле зрения. Радиометр имеет 8 каналов в диапазоне от 6 до 90 ГГц. Разрешение на местности пропорционально длине волны и составляет от 3x5 до 35x62 км. Угол сканирования в 122° с солнечно-синхронной орбиты высотой 700 км обеспечивает полосу захвата шириной 1600 км, чего, однако, недостаточно для полного покрытия поверхности Земли каждые сутки. Полное покрытие за сутки обеспечивается только для широт выше 51°.

Данные AMSR-2 собираются и обрабатываются централизованно и доступны в трёх вариантах обработки. Формат обмена данными – HDF5. В настоящем исследовании использовались данные 3-го уровня обработки на сетке 0,1° в географической системе координат на весь земной шар в виде наложения витков друг на друга за 12 часов.

Применяемый алгоритм оценки высоты снега [6] несложен и содержит две ветви. Сначала оценивается наличие глубокого снега по условиям:

$$T_{36H} < 245 \text{ K} \text{ и } T_{36V} < 255 \text{ K}, \text{ или } T_{10V} > T_{36V}, \text{ или } T_{10H} > T_{36H},$$

где $T_{NH/V}$ – это радиационная температура на частоте N горизонтальной (H) или вертикальной (V) поляризации. Если условия не выполняются, то оценивается наличие тонкого снега по выполнению условий:

$T_{89V}$ <= 255 K, и $T_{89H}$ <= 265 K, и $T_{23V}$ > $T_{89V}$, и $T_{23H}$ > $T_{89H}$, и T < 267 K,

где $T = 58.08 - 0.39T_{18V} + 1.21T_{22V} - 0.37T_{36H} + 0.36T_{89H}$. Если условия наличия тонкого снега выполняются, то высота снежного покрова принимается равной 5 см. Иначе оценивается высота глубокого снежного покрова по формуле:

$$SD = f_f SD_f + (1 - f_f) SD_o$$

где $f_f$ – доля площади, покрытая лесом; $SD_f$ и $SD_o$ – оценки высоты снега для лесистой и безлесной части пикселя:

$$SD_f = \frac{T_{18V} - T_{36V}}{log_{10}(pol_{36})(1 - 0.6 f_d)}$$

$$SD_o = \frac{T_{10V} - T_{36V}}{log_{10}(pol_{36})} + \frac{T_{10V} - T_{18V}}{log_{10}(pol_{18})}$$

Здесь $f_d$ – плотность древостоя по карте покрытия растительности (Vegetation Continuous Fraction) Университета штата Мериленд; а $pol_{36}$ и $pol_{18}$ – поляризационные множители (разница между радиационной температурой вертикальной и горизонтальной поляризаций на 36 и 18 ГГц), помогающие в некоторой степени учитывать средний размер зёрен снега.

Запасы воды в снеге затем вычисляются с использованием карты плотности снега, построенной вычислением среднемноголетней плотности снега из [7, 8] для каждого класса снега карты сезонной классификации снега [9]. Причина первичного вычисления высоты снежного покрова, а не запасов воды в снеге, заключается в том, что количество наземных данных о высоте снежного покрова значительно больше, чем о запасах воды в снеге, что по мнению авторов алгоритма потенциально позволяет добиться большей точности.

Авторы алгоритма отмечают [6] потенциальные неучтённые или недоучтённые факторы, искажающие оценки по этому алгоритму, такие как: ослабление излучения в атмосфере, ослабление излучения высокой вертикальной растительностью, изменение плотности и грануляции снега со временем, влияние наличия в пределах пикселей мелких водных объектов, а также невозможности оценки высоты тающего снега. Так, ослабление излучения неучтёнными вариациями в оптической плотности атмосферы, приведённые к ошибке запасов воды в снеге, могут достигать 25-50% [10]. Применяемая методика учёта влияние древостоя снижает, но исключает полностью его влияние, то же относится и к учёту размера гранул с помощью поляризационных множителей. Влияние мелких озёр в пределах пикселей в принципе не учитывается, тогда как, фактически, их эффект для тундры и бореальных лесов имеет даже разный знак. Наличие тающего снега только приблизительно обнаруживается, в случае чего алгоритм просто не применяется.

**Описание методики сравнения.** Для оценки точности измерений запасов воды в снеге по данным прибора AMSR-2 использовались данные наземной съёмки с 96 снегомерных маршрутов и 11 снегомерных пунктов Красноярского края и бассейнов Енисея и Верхней Оби, всего с 94 метеостанций. Маршрутная снегомерная съёмка проводится каждую декаду в зимний период и каждые 5 суток в весенний. Вне графика обычно также отмечается дата схода снега. Для сравнения использовались только данные по запасам воды в снеге. Всего для сравнения использовано более 6398 наземных измерений за период с 2012-2017 гг.

Для сбора и обработки данных AMSR-2 была разработана автоматическая система с расчётом на её дальнейшее использование в оперативной практике. Собираемые данные накапливаются в базе данных под управлением СУБД PostgreSQL. Собираемая информация AMSR-2 включает за каждый момент наблюдения (дважды в сутки) матрицу 4x4 из 16 ближайших к каждой метеостанции узлов сетки, а также результаты пространственной билинейной и бикубической интерполяции в точке метеостанции. По ходу исследования система усложнялась добавлением новых видов обработки, например, таких как вычисление максимума, среднего и медианы в скользящем окне за 16 суток.

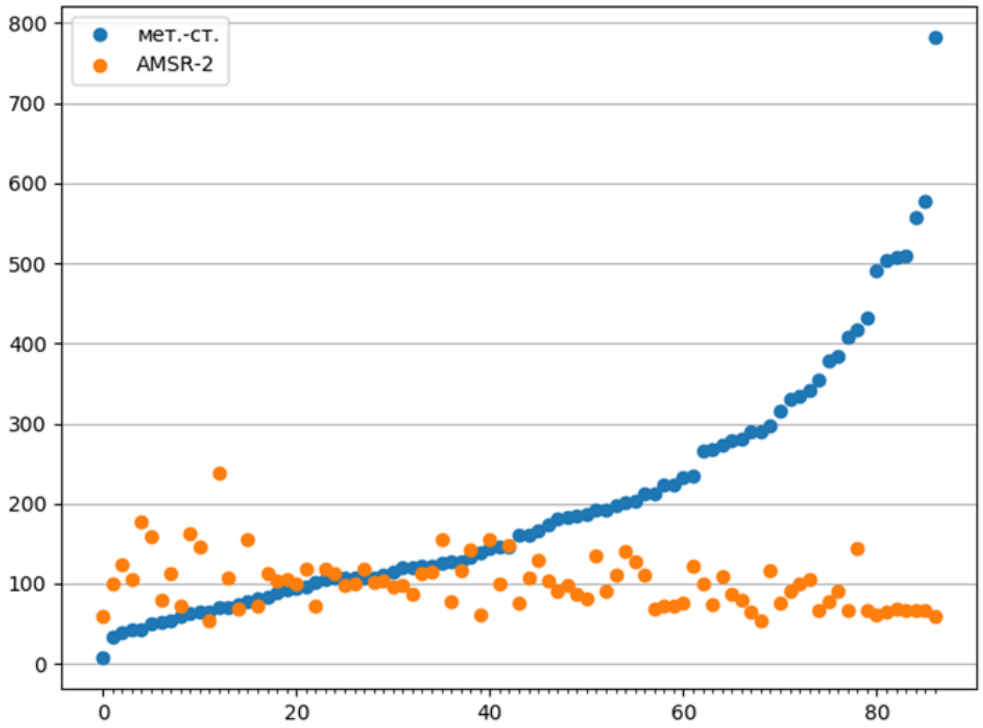
**Результаты сравнения.** В результате анализа полученных наборов данных было выявлено несколько основных особенностей данных AMSR-2, а именно:

– наличие провалов до нуля в данных среди зимнего сезона;
– неполное покрытие за одни сутки;
– ограничение сверху на максимальный обнаруживаемой снегозапас;
– бесполезность пространственной интерполяции.

Наличие нулей в данных обусловлено вышеупомянутой чувствительностью к наличию воды на поверхности снега, в случае наличия которой метод полностью теряет чувствительность до нуля. Несмотря на наличие оценки применимости метода через эмпирическую оценку температуры, сам факт неприменимости метода в данных никак не отмечается. Соответственно, нули в данных могут появиться в любой момент, например, в период оттепели зимой. При этом они, естественно, не имеют смысла, как все другие данные AMSR-2, фиксирующие снижение снегозапасов. Таким образом, требуется обязательная внешняя проверка применимости данных AMSR-2, показывающих снижение запасов воды в снеге.

Ширина полосы захвата прибора AMSR-2 составляет около 1600 км, что не обеспечивает ежедневного покрытия всей территории земного шара. Для получения непрерывного временно̀го ряда наблюдений на широтах меньше 51° необходимо использовать данные, как минимум, за двое суток. Учитывая склонность к занижению оценок при наличии снеготаяния, а также для целей определения максимального зимнего снегозапаса целесообразно применять функцию максимума из данных за двое суток.

Для метеостанций с большим зимними накапливаемыми снегозапасами проявляется ограничение на максимальный снегозапас, который может быть оценён по данным AMSR-2. В описании методики приводятся верхние оценки высоты снежного покрова, обнаруживаемые методом, в 1 м, что соответствует снегозапасу в 150-400 кг/м$^2$. И, действительно, максимальное зарегистрированное значение в окрестностях метеостанций по данным AMSR-2 составило 565 кг/м$^2$. Однако, фактически по результатам сравнения максимально обнаруживаемый снегозапас по данным AMSR-2 для многоснежных метеостанций значительно ниже – в районе 75-200 кг/м$^2$ при фактических максимумах по наземным данным 200-950 кг/м$^2$. Общий вид соотношения максимумов спутниковых и наземных данных представлен на рисунке.




Максимальные наблюдавшиеся снегозапасы (кг/м$^2$) по спутниковым и наземным данным (по горизонтальной оси — ранг метеостанции по наземным данным).

Квадраты коэффициента корреляции в узлах интерполяции м.-ст. Курагино.

| | | | |
|---|---|---|---|
| 0,757 | 0,688 | 0,577 | 0,466 |
| 0,814 | 0,679 | 0,629 | 0,528 |
| 0,797 | 0,807 | 0,792 | 0,710 |
| **0,825** | 0,818 | 0,790 | 0,657 |

Неприятным является также тот факт, что из самих данных AMSR-2 признаки его ограничения никак не вытекают, и чтобы это ограничение обнаружить требуются независимые наземные данные. Поэтому из дальнейшего исследования пришлось исключить метеостанции, для которых наблюдавшийся максимум оценки снегозапаса по AMSR-2 не достигал наблюдаемых максимумов по наземным данным. После исключения осталось 34 пункта из 106. Карта распределения оставленных и исключённых метеостанций не показывает какого-либо значимого географического фактора, по которому можно было бы без использования наземных данных судить о применимости спутниковых оценок снегозапасов.

Линейные регрессионные состоятельные модели оценки наземных данных по AMSR-2, отдельные для каждой метеостанции, построены для 24 станций из 34 по условию наличия 20 и более пар точек. Независимой переменной считались данные AMSR-2, прошедшие фильтрацию медианой в окне 16 суток (привязанную к 8 суткам окна) и интерполированные билинейной интерполяцией в точку метеостанции по ближайшим 4 узлам. Квадрат коэффициента корреляции этих моделей ($R^2$) составил от 0,29 до 0,76, для 50% станций – более 0,46. Линейный коэффициент пропорциональности в моделях от 0,156 до 0,86, константа от –7,29 до 41,86 кг/м$^2$. Попытка использовать в качестве данных не интерполированные данные узлов, показала, что всегда можно найти узел, дающий более точную модель, чем получается построить по интерполированным данным. В случае поиска наилучшего узла, обеспечивающего лучшую корреляцию, квадрат коэффициента корреляции этих моделей составил от 0,33 до 0,83, для 50% станций – более 0,55. В таблице для примера приведены квадраты коэффициентов корреляции в узлах интерполяции для метеостанции Курагино, для которой он равен 0,756 для линейной интерполяции по 4 узлам в точку метеостанции и 0,747 – для бикубической интерполяции (север вверху, м.-ст. находится между 4 внутренних узлов). Видно, что наибольший квадрат коэффициента корреляции 0,825 (выделен в табл.) наблюдается даже не в ближайшем к метеостанции узле.

Таким образом, процедура пространственной интерполяции представляется бесполезной из-за сильной неоднородности поверхности внутри площадки пикселя в несколько километров в диаметре. Получается, что окрестные пиксели могут иметь случайно более близкие интегральные оценки снегозапаса к точечным наземным наблюдениям, чем результаты формальной интерполяции по ближайшим узлам.

Для двух станций не удалось построить состоятельную модель даже при наличии более 20 пар данных. Для метеостанции Кресты Таймырские коэффициент пропорциональности получился отрицательным, а квадрат коэффициента корреляции 0,139, для Тембенчи – $R^2 = 0{,}046$. Облако точек на скаттерограмме у них, фактически, круглое и не имеет выраженной ориентации. Это можно объяснить случайным фактическим распределением типа поверхности и рельефа местности в пределах пикселей AMSR-2. Так, например, Кресты Таймырские расположены на берегу р. Хатанга в окружении тундровых озёр, негативное влияние которых отмечалось в литературе выше.

**Заключение.** Таким образом использование данных микроволнового радиометра AMSR-2, установленного на КА GCOM-W1, для оценки запасов воды в снеге возможно только в период снегонакопления и невозможно в период снеготаяния. Для обеспечения полного покрытия территории, в том числе для широты меньше 51°, необходимо использовать, как минимум, данные за двое суток. При необходимости связать данные о запасах воды в снеге с AMSR-2 с данными наземных измерений требуется построение отдельной регрессионной зависимости для каждого пункта наземных измерений с подбором узла сетки данных AMSR-2 с

наибольшим коэффициентом корреляции в качестве независимой переменной. Построение регрессионной зависимости не всегда возможно по условиям местности вокруг метеостанции. Также максимальные возможные снегозапасы в точке метеостанции не должны превышать величины около 200 кг/м$^2$, выше которой снегозапасы по данным AMSR-2 оценить невозможно по условиям применимости методики. Фактические максимальные регистрируемые снегозапасы в данном месте необходимо определять по наземным данным.

## ЛИТЕРАТУРА


[1] *Matson M.* NOAA satellite snow cover data // Paleogeography and Paleoecology. 1995. N 90. P. 213-218.

[2] *Kunzi K.F., Patil S., Rott H.* Snow-cover parameters retrieved from Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR) data // IEEE Transactions on Geoscience and Remote Sensing GE20. 1982. P. 452-467.

[3] *Chang A.T.C., Foster J.L., Hall D.K.* Nimbus-7 SMMR derived global snow cover parameters // Ann. Glaciol. 1987. Vol. 9. P. 39-44.

[4] *Kelly R.E.J., Chang A.T.C.* Development of a passive microwave global snow depth retrieval algorithm for SSM/I and AMSR-E data // Radio Science. 2003. Vol. 38.

[5] *Kelly R.* The AMSR-E snow depth algorithm: description and initial results // Journal of the Remote Sensing Society of Japan. 2009. Vol. 29(1). P. 307-317.

[6] Descriptions of GCOM-W1 AMSR2 Level 1R and Level 2 Algorithms. Japan Aerospace Exploration Agency Earth Observation Research Center. 2013. **http://suzaku.eorc.jaxa.jp/GCOM_W/data/doc/NDX-120015A.pdf** (Accessed at 13.07.2018).

[7] *Brown R.D., Braaten R.O.* Spatial and temporal variability of Canadian monthly snow depths, 1946-1995 // Atmosphere-Ocean. 1998. Vol. 36. P. 37-45.

[8] *Krenke A.* Former Soviet Union hydrological snow surveys, 1966-1996 / Edited by NSIDC. Boulder, CO: National Snow and Ice Data Center/World Data Center for Glaciology. Digital media.

[9] *Sturm M., Taras B., Liston G. et al.* Estimating Snow Water Equivalent Using Snow Depth Data and Climate Classes // Journal of Hydrometeorology. 2010. Vol. 11. P. 1380-1394. **http://journals.ametsoc.org/doi/abs/10.1175/2010JHM1202.1** (Accessed at 30.05.2019).

[10] *Wang J., Tedesco M.* Identification of atmospheric influences on the estimation of snow water equivalent from AMSR-E measurements // Remote Sensing of Environment. 2007. Vol. 111 (2-3). P. 398-408.

# СЕГМЕНТАЦИЯ ОЛИВКОВЫХ ДЕРЕВЬЕВ С ИСПОЛЬЗОВАНИЕМ АРХИТЕКТУРЫ MASK R-CNN

*Сафонова А.Н.[(1)(2)], Табик С.[(2)], Маглинец Ю.А.[(1)]*
[(1)] Институт космических и информационных технологий СФУ, Красноярск
[(2)] Центр исследований в области информационных технологий и коммуникаций Университета Гранады, Гранада, Испания

В этой работе мы представляем результаты сегментации экземпляров оливковых деревьев с использованием изображений высокого разрешения, индексов растительности NDVI и GNDVI и современной архитектуры Mask R-CNN.
*Ключевые слова: сегментация экземпляра, машинное обучение, глубокие нейронные сети, оливковые деревья, снимки высокого разрешения.*

**Введение.** Методы локализации крон деревьев и определения параметров древостоев являются востребованным направлением ДЗЗ и решаются многими научными коллективами. В работе рассматриваются задачи локализации на изображениях земной поверхности деревьев оливы, а также измерения их параметров, с точностью до отдельного дерева. Практическая значимость работы определяется следующими соображениями. Испания занимает лидирующую позицию в мире по производству оливкового масла (44% от общемирового объема производства) [1], площади посадок таковы, что наземные измерения требуют привлечения значительных ресурсов в то время, как своевременное измерение характеристик деревьев позволяет осуществлять прогнозирование урожайности и расчет других экономических параметров.

Методы машинного обучения развиваются бурными темпами. В частности, они демонстрируют высокие результаты в задачах распознавания, классификации и сегментации отдельных объектов на данных дистанционного зондирования Земли [2-4]. В последнее время указанные методы широко используются для решения задач агромониторинга. В настоящей работе представлены результаты применения сверточных нейронных сетей для определения попиксельной маски объектов оливковых деревьев по данным сверхвысокого разрешения.

**Метод исследования.** Одной из ключевых задач мониторинга оливковых полей является задача сегментации изображений, направленной на выявление крон деревьев и других элементов, важных при их распознавании. В работе выполняется сегментация экземпляров (instance segmentation) – задача определения контуров объектов на уровне пикселей анализируемого изображения. Для ее решения мы использовали сверточную нейронную сеть Mask R-CNN (regions with convolutional neural networks) [5]. В данной СНС реализован двухэтапный подход. На первом этапе сканируется входное изображение и генерируются предложения (области, которые могут содержать объект). На втором этапе выполняется классифицирование предложения и формирование ограничивающих рамок и масок.

Mask R-CNN состоит из описанных следующих структурных компонентов.

1. Основа – стандартная СНС ResNet101, на ранних слоях которой обнаруживаются объекты низкого уровня (края и углы), а на более поздних слоях сети обнаруживаются элементы боле высокого уровня (дерево, человек, здание). Проходя через нейронную сеть, изображение преобразуется из матрицы вида 1024×1024px×3 (RGB) в карту объектов формы 32×32×2048. Эта карта характеристик становится входом для следующих уровней сети. Для улучшения качества извлечения объектов сетью мы использовали метод Feature Pyramid Network (далее FPN), который берет обнаруженные СНС элементы высокого уровня и передает их вниз в нижние слои. Это позволяет функциям на каждом уровне иметь доступ к функциям как нижнего, так и более высокого уровня.

2. Region Proposal Network – СНС, которая сканирует изображение в режиме скользящего окна и находит области, содержащие объекты. В результате сеть генерирует: класс привязки

(передний план и фон, где в первом подразумевается присутствие классифицируемого объекта) и ограничительный фрейм с оценкой дельты (процент относительно точности измерения координат, ширины, высоты) для уточнения поля привязки для лучшего соответствия объекту.

3. Классификатор и ограничивающее окно-регрессор. Данный компонент генерирует два выхода для каждого объекта: класс объекта и координаты ограничивающего фрейма, необходимые для дальнейшего уточнения расположения и размера объекта.

Так как для подачи информации на вход классификатора нейронной сети требуются изображения фиксированного размера, мы использовали дополнительный слой подвыборки (субдискретизации), который позволяет обрезать часть карты объекта до фиксированного размера.

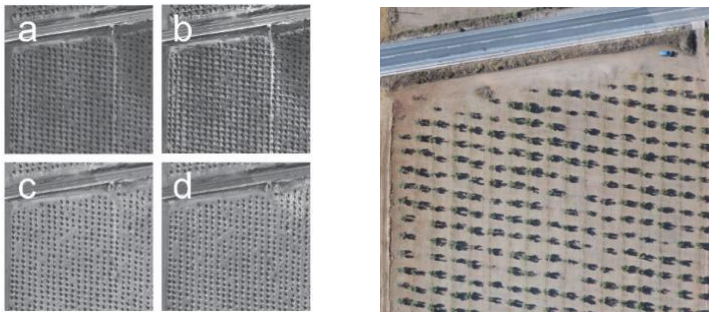
4. Сегментные маски – сверточная сеть, которая берет выделенные классификатором области объектов, и генерирует для них маски. Сгенерированные маски имеют низкое разрешение: 28×28 пикселей. Во время обучения мы уменьшаем размер маски до 28×28 пикселей, чтобы вычислить потери, а во время вывода мы увеличиваем предсказанные маски до размеров ограничивающего фрейма объекта, и они являются окончательными масками, по одной на объект.

**Исходные данные для эксперимента.** Тестовый участок расположен на севере города Гранады, Андалусия, Испания. Объектом исследования выступает дерево *Picular*. Это культивированный сорт оливок, используемый для изготовления оливкового масла (*Marteño, Nevadillo, Lopereño*). Деревья были высажены в 2006 году. Количество деревьев на испытательном участке составляло 2700 штук, которыми занималась площадь 12 Га, но в общей сложности на плантации произрастает 11000 деревьев (общая площадь 50 Га). Для проведения эксперимента использовались материалы съемки с беспилотного летательного аппарата (далее БПЛА). Мы использовали два БПЛА: PARROT DISCO-PRO AG с многоспектральным датчиком Parrot Sequoia в четырех диапазонах (зеленый (Green), красный (Red), крайний красный (Red_ Edge), ближний инфракрасный каналы (NIR)) и DL-Phantom 4 Pro с RGB камерой (рис. 1).

Четырехканальные изображения сверхвысокого разрешения были получены в феврале 2019 года с разрешением 13 см/пиксель, а цветное изображение было получено в июле 2019 года с разрешением 3 см/пиксель. Оба полета дронов производились на высоте 120 метров. Также в качестве эксперимента мы провели расчеты вегетационных индексов (далее ВИ): нормализованный вегетационный индекс растительности NDVI (1) и зеленый нормализованный вегетационный индекс растительности GNDVI (2) по данным четырехканального изображения. GNDVI является показателем фотосинтетической активности растительного покрова, наиболее часто он используется при оценке влагосодержания и концентраций азота в листьях растений. GNDVI более чувствителен к концентрациям хлорофилла по сравнению с индексом NDVI. (рис. 2) [6].

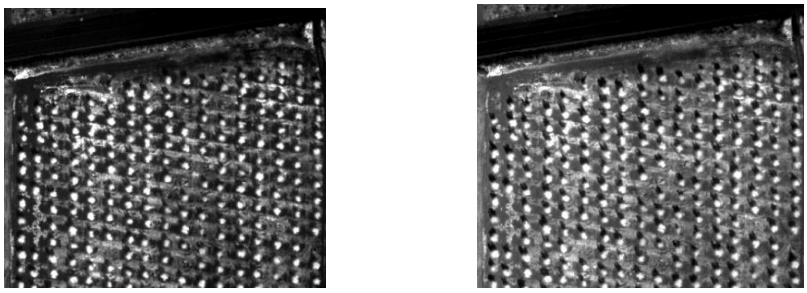
$$NDVI = \frac{NIR-Red}{NIR+Red} \quad (1)$$

$$GNDVI = \frac{NIR-Green}{NIR+Green}. \quad (2)$$




*а* *б*

Рис. 1. Фрагменты ортоизображений тестового участка:
*а* – четырехканальное изображение (a – Green, b – Red, c – Red_ Edge, d – NIR),
*б* – RGB изображение.



*а* *б*

Рис. 2. Фрагмент ортоизображений рассчитанных ВИ NDVI (*а*) и GNDVI (*б*).

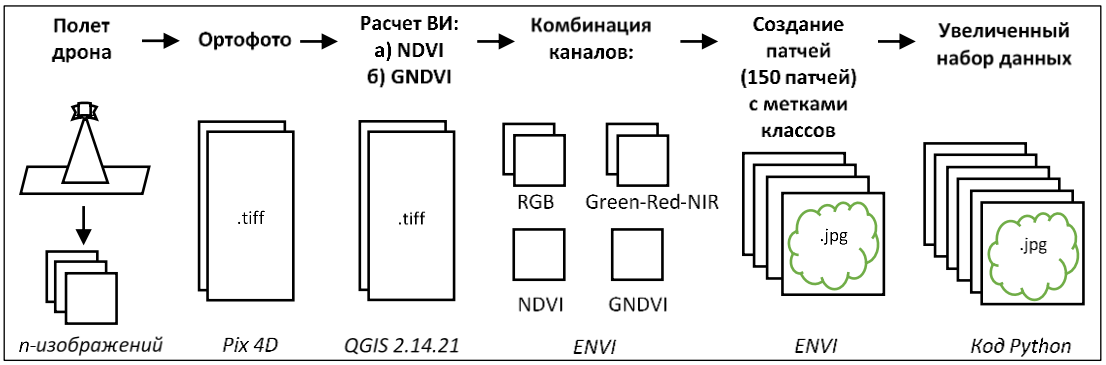



Рис. 3. Процесс получения и создания набора данных для обучения и тестирования модели.

Таким образом, для проведения экспериментов мы использовали следующие комбинации каналов: а) RGB изображение, б) Green-Red-NIR, в) NDVI, и г) GNDVI.

На выбранных комбинациях далее велась подготовка набора данных для обучения и тестирования архитектуры Mask R-CNN. Было подготовлено 150 изображений (патчей) из которых 80% было предназначено для обучения модели и 20% для валидации и тестирования обученной модели. Патч представляет собой изображение, на котором представлено от одного до восьми деревьев. Общая схема создания набора данных представлена на рис. 3.

Подготовленный набор патчей был конвертирован в одноканальные изображения в формате .jpg для дальнейшего искусственного увеличения данных с помощью применения таких функций как: поворот изображения по горизонтали на 99%, поворот изображения по вертикали на 99%, улучшение или ухудшение контрастности изображения, увеличение резкости, изменение яркости изображения на 10-150% от исходного изображения. Тестовый набор данных был искусственно увеличен до 1500 изображений для каждой комбинации набора данных.

**Результаты экспериментов.** В данном разделе представлены результаты экспериментальных исследований применения Mask R-CNN в задаче сегментации оливковых деревьев на снимках сверхвысокого разрешения, полученных с помощью БПЛА.

Основная работа проводилась с помощью языка программирования Python и модуля машинного обучения TensorFlow Object Detection API. TensorFlow – программная библиотека с открытым исходным кодом для высокопроизводительных численных вычислений, которая работает в различных гетерогенных системах, включая кластеры распределенных графических процессоров (GPU). Расчеты выполнялись на ПК с процессором Intel Xeon E5-2630v4, ускоренным с помощью графического процессора NVIDIA Titan Xp в качестве платформы для обучения и тестирования предложенной методики. Для подготовки набора изображений были использованы следующие программы: Pix 4D, QGIS 2.14.21, ENVI, VGG Image Annotator. VGG Image Annotator является автономным программным обеспечением для ручного аннотирования изображений, аудио и видео. Создание меток классов осуществлялось в программе VGG Image Annotator с сохранением их в формате JSON (пример ручной отрисовки меток на изображении RGB представлен на рис. 4).

Как видно из рис. 5, результаты сегментации и построения пиксельных масок для каждого объекта на изображениях дают среднюю точность 99.66%.




Рис. 4. Создание меток классов «Оливковые деревья» и «Тень» на изображении RGB.

Далее архитектура Mask R-CNN обучалась в течении 100 эпох для каждой комбинации каналов в течении 60 часов. Результаты сегментации экземпляров оливковых деревьев для каждой комбинации изображений представлены на рис. 5.

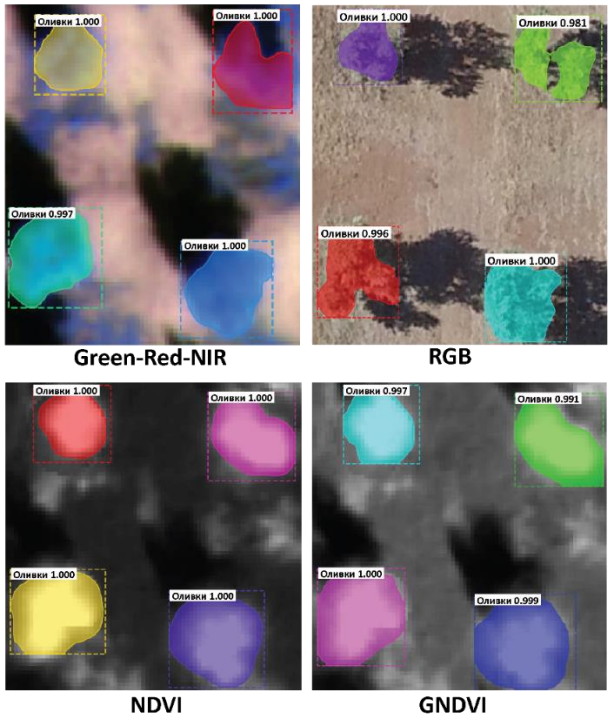



Рис. 5. Результат сегментации класса «Оливковые деревья» с использованием Mask R-CNN на увеличенных наборах данных.

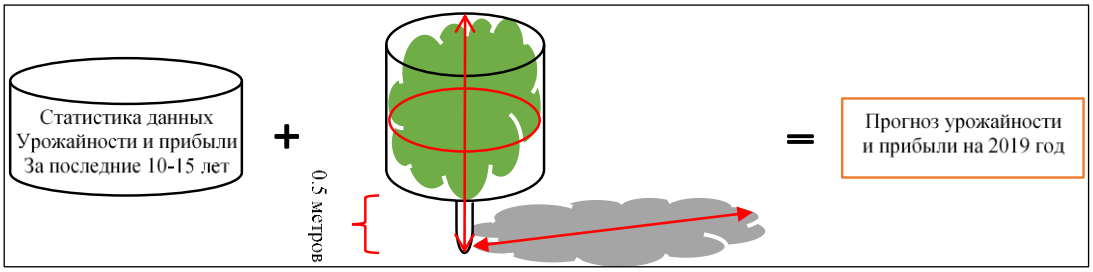



Рис. 6. Графическое представление расчета параметров оливковых деревьев.

В качестве эксперимента нами было проведено дополнительное обучение СНС для сегментации класса «тени оливковых деревьев» и расчета размеров каждого дерева. Результаты выполненных расчетов совпадают с наземными измерениями 10 тестовых деревьев в среднем на 95%. Таким образом, нам удалось рассчитать:

- попиксельно площадь объекта, учитывая, что размер одного пикселя в маске соответствует 3 см для RGB изображений и 13 см для изображения Green-Red-NIR, NDVI, GNDV;

- высоту дерева ($h$) по размеру тени дерева в метрах ($l_T$) и с учетом высоты солнца над горизонтом в зависимости от географического положения (широты и долготы), даты и времени съемки ($tg\gamma$) (3) [7].

$$h = l_T * tg\gamma \quad (3)$$

- объем кроны дерева, без учета нижней части ствола дерева, на котором нет веток (в среднем до 0.5 метров) (рис. 6).

Предложенные расчеты возможно использовать для проведения дополнительных полевых измерений и экспериментов по прогнозированию урожайности и прибыли, получаемой с каждого дерева, на основе статистических данных, существующих за последние 10-15 лет.

**Заключение.** Осуществлено исследование применимости одного класса свёрточных нейронных сетей к задачам обнаружения на изображениях земной поверхности деревьев оливы, а также расчету ряда их параметров, с точностью до отдельно стоящего дерева. Полученные результаты показывают, что использование машинного обучения, и, в частности, глубоких нейронных сетей, дает возможность решить поставленные задачи с высокой точностью. Предложенная методика сегментации оливковых деревьев с помощью современной архитектуры Mask R-CNN и модуля TensorFlow продемонстрировала среднюю точность распознавания в 99.66%. Таким образом, данную методику пиксельной сегментации с использованием машинного обучения можно с высокой эффективностью использовать в задачах сельского и лесного хозяйства. В дальнейшем планируется провести тестирование обученных СНС на спутниковых данных среднего разрешения, что представляет наибольший интерес для использования возможных результатов на больших территориях, а также прогнозирования урожайности и прибыли, полученной с оливковых деревьев.

*Работа выполнена при финансовой поддержке стипендии Президента Российской Федерации для проведения научной стажировки за рубежом на 2018/19 учебный год от 07.06.2018 г. № 05-1997 и Российского фонда фундаментальных исследований (номер проекта: 18-47-242002 р_мк), Правительства Красноярского края, Красноярского краевого фонда науки в рамках научного проекта: «Разработка технологии создания интеллектуальных информационных систем объектно-ориентированного мониторинга территорий по данным дистанционного зондирования».*

## ЛИТЕРАТУРА

[1] EU olive oil farms report // European commision. Greece, Italy, Spain. 2012 **http://www.fao.org/family-farming/detail/en/c/326556** (дата обращения 25.06.2019).

[2] *Зарубин О.А.* Применение нейронных сетей для целей анализа данных дистанционного зондирования Земли // Современные научные исследования и инновации. 2016. № 8. **http://web.snauka.ru/issues/2016/08/70887** (дата обращения 25.06.2019).

[3] *Gonzalez-Fernandez I., Iglesias-Otero M.A, Esteki M. et al.* A critical review on the use of artificial neural networks in olive oil production, characterization and authentication // Critical Reviews in Food Science and Nutrition. 2018. Vol. 59. **https://doi.org/10.1080/10408398.2018.1433628** (дата обращения 25.06.2019).

[4] *Safonova A., Tabik S., Alcaraz-Segura D. et al.* Detection of Fir Trees (Abies sibirica) Damaged by the Bark Beetle in Unmanned Aerial Vehicle Images with Deep Learning // Remote Sensing. 2019. Vol. 11, N. 6. P. 643

[5] *Ren S., He K., Girshick R., Sun J.* Faster R-CNN: Towards real-time object detection with region proposal networks // Neural Information Processing Systems (NIPS). 2015.

[6] Растительность // Геоаналитика.Агро. **http://agro.geoanalitika.com/ru/products/rastitelnost/#** (дата обращения 25.06.2019).

[7] *Selly H.E.* Tree Heights from shadows. **https://www.asprs.org/wp-content/uploads/pers/1942journal/jun/1942_jun_100-109.pdf** (дата обращения 25.06.2019).

# КЛАССИФИКАЦИЯ ПРОСТРАНСТВЕННЫХ СТРУКТУР ОТРОГОВ БОЛЬШОГО ВАСЮГАНСКОГО БОЛОТА С ИСПОЛЬЗОВАНИЕМ СПУТНИКОВЫХ ДАННЫХ

*Синюткина А.А.*
Сибирский институт сельского хозяйства и торфа – филиал Сибирского федерального научного центра агробиотехнологий РАН, Томск

В статье представлена классификация пространственных структур болотных массивов на примере северо-восточных отрогов Большого Васюганского болота. Выделены 4 класса пространственных структур, определяющие стадию развития и форму поверхности болотных массивов и, как следствие, интенсивность поверхностного стока с болота на прилегающие территории. Выделенные классы являются индикаторами протяженности зоны влияния болот при низких уклонах поверхности. Средняя протяженность зоны влияния в пределах рассматриваемой территории составила 3,9 км с колебаниями в пределах 0,6-8,4 км.
*Ключевые слова: заболачивание, верховой болотный массив, Landsat, болотные микроландшафты, уклон поверхности.*

В условиях современного изменения климата является актуальной проблема определения тренда и оценка зоны потенциального заболачивания прилегающих к болотам лесов [1]. Известно, что протяженность зоны влияния болота главным образом определяется геолого-геоморфологическими факторами, среди которых выделяются угол наклона, горизонтальное и вертикальное расчленение поверхности. На интенсивность развития процесса заболачивания территории может оказать влияние смена уклона поверхности даже на десятые доли градуса. На территориях с низкими значениями уклонов поверхности, к каким относится окраинные части Большого Васюганского болота, зависимости протяженности зоны влияния болота от уклона поверхности не всегда определяются при среднемасштабном картографировании в связи с недостаточной детальностью и точностью данных топографических карт и SRTM, а оценка влияние морфометрических показателей рельефа на процесс заболачивания становится возможной только при проведении детальных полевых исследований на ключевых участках [2]. Поэтому для моделирования зон потенциального заболачивания необходим поиск доступных индикаторов протяженности зоны влияния болотных массивов, являющимися удобными в использовании и позволяющими проводить оценку гидроморфной трансформации на обширных площадях. Одним из таких индикаторов является характеристика пространственных структур верховых болотных массивов, отражающая стадию развития болота, форму его поверхности и, как следствие, интенсивность поверхностного стока на прилегающие территории. Целью исследования является разработка классификации пространственных структур болотных массивов и прилегающих территорий по набору и закономерностям смен болотных фаций от центральной части болотного массива к окраинам, протяженности зоны влияния болота на прилегающие территории.

Объектами исследования являются северо-восточные отроги Большого Васюганского болота (Иксинский, Бакчарский болотные массивы) (рис. 1).

Оценка пространственных структур болотных массивов и определение протяженности влияния болота проведена с использованием методов геоинформационного моделирования. Всего заложено 40 профилей от центра к периферии болотных массивов параллельно линиям стока. На каждом профиле поведено измерение протяженности болотного массива, расстояния от водораздела до окраины болота, протяженности зоны влияния болота, угла наклона поверхности в пределах зоны влияния, определен набор и протяженность основных микроландшафтов (грядово-озерковый и грядово-мочажинный комплексы, сосново-кустарничково-сфагновые болота с низкой и высокой сосной, травяно-моховые топи). Карта углов наклона поверхности построена методом растровой интерполяции *Topo to Raster* на основе высотных отметок топографических карт с коррекцией по линям гидрографической сети. Определение границ зоны потенциального заболачивания проведено на основе классификации данных Landsat 8 (дата съемки 17.07.2018). Обучающая выборка составлена с использованием данных полевых

ландшафтных исследований и включала в себя участки, индицирующие развитие процессов гидроморфной трансформации геосистем: 1 – сильно переувлажненные участки лесов, характеризующиеся массовым повалом деревьев; 2 – сосново-кустарничковое травяно-сфагновое верховое болото за пределами основного контура болотных массивов; 3, 4 – участки заболоченных сосново-березовых и кедрово-березовых лесов. Согласно данным проведенных исследований, участки характеризуются накоплением торфа мощностью до 1 м и более, увлажнение изменяется в пределах 14,2-15,7 баллов и характеризуется как сыровато-лесолуговое, сыролесолуговое и мокро-лесолуговое согласно оценке условий местообитаний методом стандартных экологических шкал [3, 4]. Для классификации пространственных структур болотных массивов использован метод кластерного анализа, сравнение отличий значений протяженности зоны влияния болота между классами проведено с использованием тестов Mann-Whitney и Kruskal-Wallis. Выявление зависимостей протяженности зоны влияния болотных массивов от характеристик пространственной структуры и углов наклона поверхности проведено с использованием метода главных компонент.

Для большей части крупных болотных массивов изучаемой территории характерна смена растительных сообществ от кедрово-соснового травяно-сфагнового верхового болота, протягивающегося узкой полосой вдоль границы болотного массива на сосново-кустарничково-сфагновое болото, занимающее значительные площади на склонах болотных массивов. В центральной части болотных массивов широкое распространение получили грядово-мочажинные и грядово-озерковые комплексные сосново-кустарничковые травяно-сфагновые болота [5]. Северные части крупных болотных массивов, а также более мелкие болота характеризуются отсутствием комплексных микроландшафтов и более простой пространственной структурой. Ширина болотных массивов параллельно направлению стока в среднем составляет около 11 км, на отдельных участках достигает 24 км. Углы наклона поверхности болот изменяются в пределах 0-0,2°, но на большей части поверхности не превышают 0,05°. В пределах водораздельных равнин, занятых лесами, углы наклона поверхности обычно не превышают 0,3° и в среднем составляют около 0,1° (рис. 2).

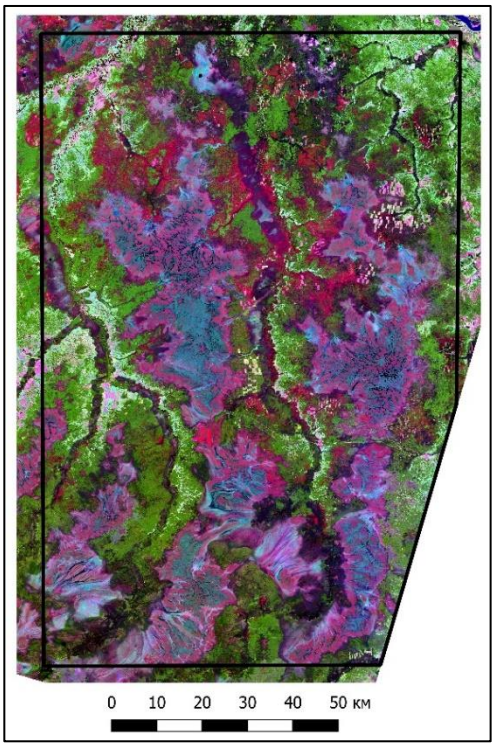



Рис. 1. Схема расположения объектов исследования (КС Landsat 8, 17.07.2018).

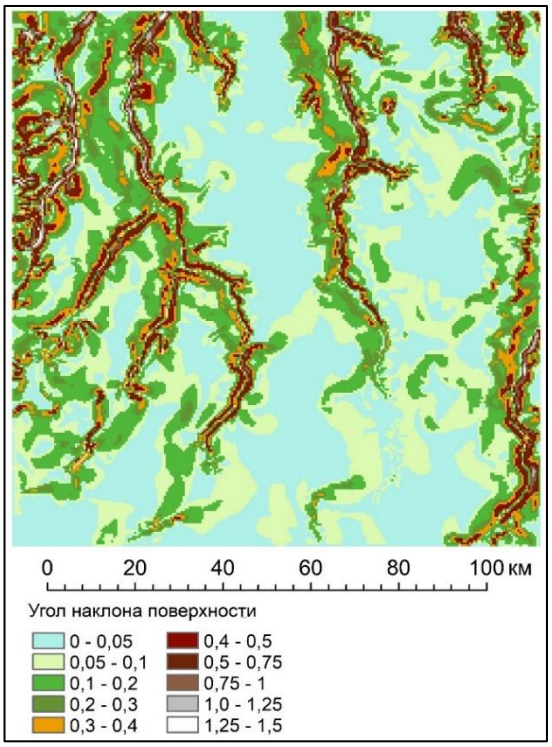



Рис. 2. Карта углов наклона ключевого участка (градусы).

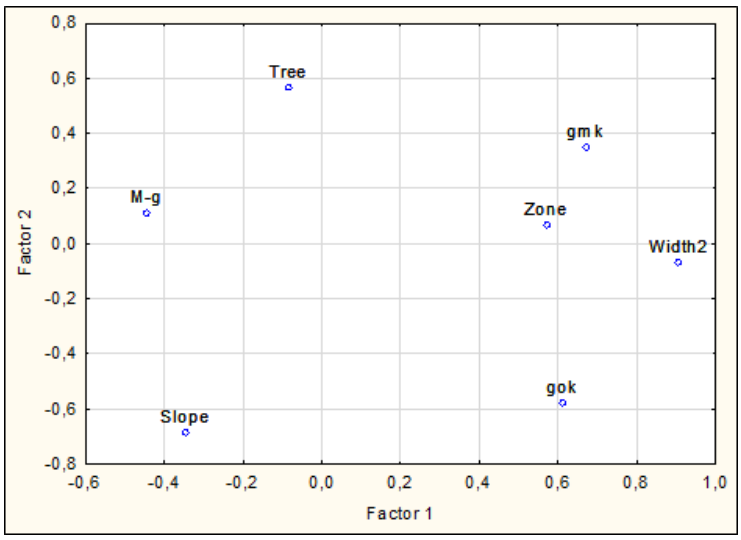



Рис. 3. Результаты факторного анализа показателей пространственной структуры, уклона и протяженности зоны влияния болот: Zone – протяженность зоны влияния болота; Width2 – расстояния от водораздела до окраины болота; Slope – угол наклона поверхности; протяженность микроландшафтов: M-g – травяно-моховые топи; Tree – древесные и древесно-моховые; gmk – грядово-мочажинные комплексы; gok – грядово-озерковые комплексы.

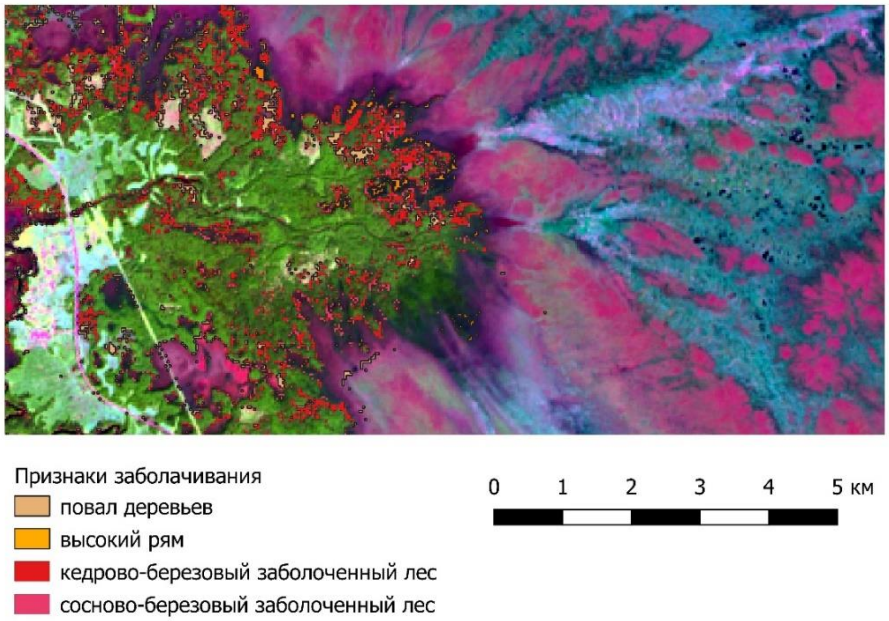



Рис. 4. Зона влияния болотного массива с пространственной структурой 1 класса (КС Landsat 8, 17.07.2018).

Проведенный факторный анализ показал, что протяженность зоны влияния болота в большей степени определяется шириной болотного массива (0,92), протяженностью грядово-мочажинных (0,64) и грядово-озерковых (0,63) комплексных микроландшафтов. Меньшее влияние оказывают углы наклона поверхности прилегающих к болоту территорий (0,30), что связано с низкой вариацией значений при практически плоском рельефе изучаемого участка (рис. 3).

В результате кластерного анализа массива данных о протяженности болотных микроландшафтов рассмотренные профили объединены в 4 кластера, соответствующие классам пространственных структур.

Первый класс пространственных структур отличается широким распространением комплексных болотных, в том числе грядово-озерковых микроландшафтов, протяженность которых достигает своего максимального значения в пределах рассматриваемого участка в 15 км,

со средним значением 8,7 км (рис. 4). В отличие от первого, второй класс характеризуется большим распространением грядово-мочажинных микроландшафтов, при практически полном отсутствии грядово-озерковых. В большинстве случаев окраины болот заняты древесными и древесно-моховыми микроландшафтами. Для первого и второго класса пространственных структур характерны максимальные значения протяженности зоны влияния болота в пределах ключевого участка, достигающей 8,4 км от границы болотного массива со средним значением 4,6 км (рис. 6).

Третий класс характеризуется широким развитием топяных микроландшафтов. Средняя протяженность грядово-мочажинных болот не превышает 1 км, грядово-озерковые комплексы в большинстве случаев отсутствуют. Средняя протяженность зоны влияния болота составляет 3,5 км. Четвертый класс пространственных структур отличается полным отсутствием комплексных и широким развитием древесных и древесно-моховых микроландшафтов. Класс характеризуется минимальной протяженностью зоны влияния болот в пределах рассматриваемого участка, среднее значение которой составило 1,7 км.

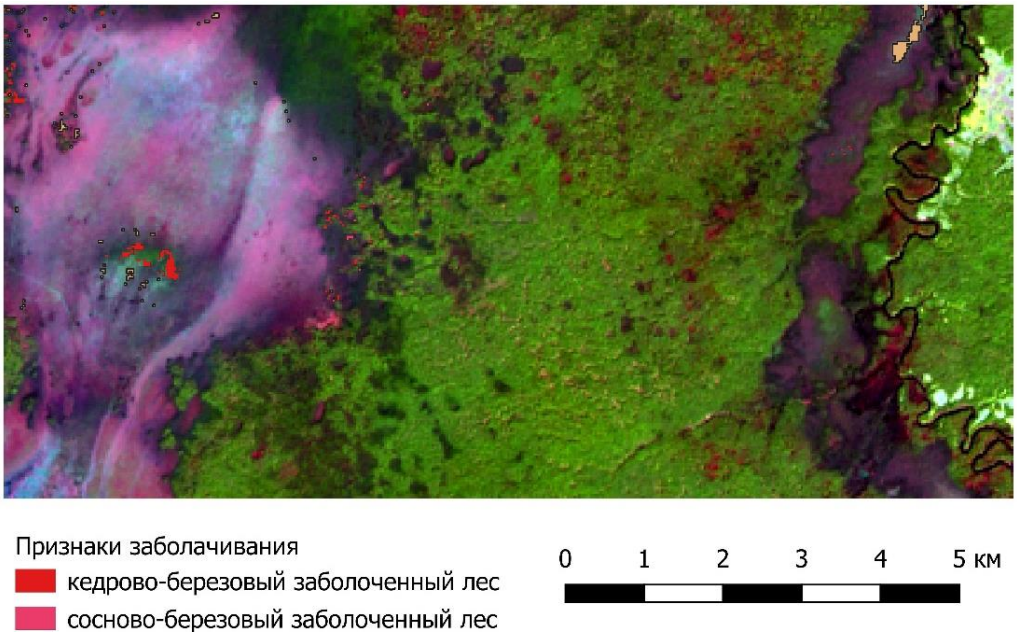



Рис. 5. Зона влияния болотного массива с пространственной структурой 4 класса (КС Landsat 8, 17.07.2018)

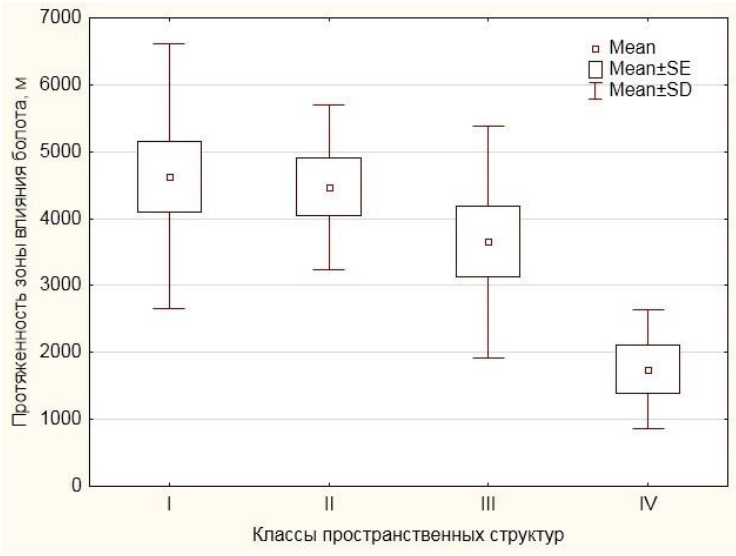



Рис. 6. Сравнение протяженности зоны влияния болотных массивов с разными пространственными структурами (I–IV – номера классов)

Таким образом, в результате исследования проведена оценка зоны влияния северо-восточных отрогов Большого Васюганского болота на прилегающие территории. Средняя протяженность зоны составила 3,9 км с колебаниями в пределах 0,6–8,4 км. Статистически подтверждена зависимость протяженности зоны от характеристик пространственных структур болотного массива. Выделенные классы пространственных структур отражают стадию развития болотного массива, как следствие, интенсивность поверхностного стока с болота на прилегающие территории и являются индикаторами протяженности зоны влияния болот при низких уклонах поверхности. Протяженность зоны влияния главным образом определяется распространением комплексных болотных микроландшафтов, наличие которых отражает заключительную стадию развития болотного массива.

*Работа выполнена при финансовой поддержке РФФИ (грант № 18-35-00387 –мол_а).*

## ЛИТЕРАТУРА

[1] *Бабиков Б.В., Кобак К.И.* Поглощение атмосферного углекислого газа болотными экосистемами территории России в голоцене. Проблемы заболачивания // Известия высших учебных заведений. Лесной журнал. 2016. № 1(349). С. 6-36.

[2] *Синюткина А.А., Гашкова Л.П.* Оценка состояния окраинных частей Большого Васюганского болота (на примере Бакчарского болотного массива) // Почвы и окружающая среда. 2018. Том 1. № 4. С. 243-255.

[3] *Прокопьев Е.П.* Экология растений (особи, виды, экогруппы, жизненные формы). Томск: Томский государственный университет, 2001. 340 с.

[4] *Цыганов Д.Н.* Фитоиндикация экологических режимов в подзоне хвойно-широколиственных лесов. М: Наука, 1983. 197 с.

[5] *Чудиновская Л.А., Синюткина А.А.* Картографирование растительных сообществ окраинных частей Большого Васюганского болота // Перспективы развития и проблемы современной ботаники: Материалы IV(VI) Всероссийской молодежной конференции с участием иностранных ученых (8-12 октября 2018 г., г. Новосибирск) / Отв. ред. А.П. Беланова. Новосибирск: Изд-во «Академиздат», 2018. С. 244-247.

# СИГНАЛЫ GPS, ГЛОНАСС И GALILEO В МОНИТОРИНГЕ СОСТОЯНИЯ ДРЕВОСТОЯ

*Сорокин А.В.[(1)], Иванов В.А.[(2)], Пономарев Е.И.[(1)], Макаров Д.С.[(1)], Харламов Д.В.[(1)], Голубев Д.В.[(3)]*

[(1)] ФИЦ КНЦ СО РАН, Красноярск

[(2)] СибГУ им. М.Ф. Решетнева, Красноярск

[(3)] Центр защиты леса Красноярского края, Красноярск

Рассмотрен вариант использования сигналов навигационных спутников в мониторинге состояния древостоя, пораженных полиграфом уссурийским. Апробирован метод радиоскопии для определения состояния древостоя. Проведены тестовые измерения погонных коэффициентов ослабления сигналов навигационных спутников диапазона L1 на участках смешанного леса. Выявлены возможности локализации участков с пожароопасным состоянием древостоя.

*Ключевые слова: радиоскопия, сигналы навигационных спутников, мониторинг, лес, полиграф уссурийский, влажность, пожароопасность.*

**Введение.** Изучение, оценка текущего и прогнозного состояния обширных лесных массивов требует использования объективных средств контроля и долговременного мониторинга, что в современных условиях перспективно проводить, используя спутниковые мониторинговые комплексы, в том числе глобальные навигационные спутниковые системы (GPS, ГЛОНАСС, GALILEO).

К факторам, определяющим состояние лесной территории, можно отнести наличие ветровалов, вырубок, пожаров, поражений энтомовредителями (сибирский шелкопряд, полиграф уссурийский и т.д.).

Другой аспект проблемы заключается в том, что климатические изменения, увеличивающаяся антропогенная нагрузка, нарушение лесных территорий (рубки, воздействие энтомовредителей) [1, 2], приводят к повышению природной пожарной опасности и, следовательно, к увеличению горимости лесов [3-5]. В мониторинге пожароопасности важную роль играет возможность контроля влагосодержания лесных горючих материалов, что определяет риски возникновения и распространения горения. Оценки и мониторинг влагосодержания лесных массивов также перспективно проводить, используя радиометрические данные различных спутниковых систем [6, 7].

В лесных районах Сибири в качестве альтернативы можно использовать сигналы навигационных спутников для непрерывного мониторинга состояния древостоев и лесных горючих материалов. Например, методы радиоскопии древостоя и рефлектометрии лесного опада с использованием автономных специализированных приемников сигналов навигационных спутников позволяют эффективно оценивать влагосодержание лесных горючих материалов [8]. В качестве критерия динамики пожаропасного состояния можно использовать данные о последовательных съемках лесных участков. Периодичность повторения съемки может определяться скоростью достижения равновесного влагосодержания в зависимости от типа лесной подстилки. В среднем величина периода высыхания материалов лесного опада до равновесного значения влагосодержания составляет 100-150 часов [7].

В данной работе анализируется метод использования радиоскопии в диапазоне L 1 для определения коэффициентов погонного ослабления сигналов GPS, ГЛОНАСС, Galileo и восстановления состояние древостоя, подвергнувшегося воздействию энтомовредителей.

**Источники зондирующего излучения.** Сигналы глобальных навигационных спутниковых систем (ГНСС) высокотехнологичны и позволяют в непрерывном режиме, наряду с координатными измерениями, восстанавливать в реальном времени электрофизические параметры сред, с которыми взаимодействуют в процессе распространения. Системы ГНСС имеют значимый для мониторинга земных покров ресурс – когерентные и поляризованные импульсные радиосигналы широкого диапазона частот, проходящие околоземное пространство с ионосфе-

рой и атмосферой, отраженное и рассеянное поверхностью покровов Земли. В каждой группировке по 24 рабочих спутников (27 для Galileo), которые работают на круговых орбитах с высотой 19100, 20183, 23224 метра для ГЛОНАСС, GPS, Galileo соответственно. Несущие частоты L1=1575.42МГц для GPS и Galileo, и L1 = 1602 МГц ± K*562,5 кГц, где K = –7, –6, … 6 для ГЛОНАСС. Радиометрия изменений параметров сигнала, трансформируемого средами, позволяет восстанавливать значимые для практики электрофизические характеристики этих сред. Это коэффициенты отражения от границ раздела земных поверхностей, фазовые задержки и коэффициенты пропускания объемных сред – ионосферы, атмосферы, растительных, снежных и ледовых покровов.

Фактически навигационные спутниковые системы позволяют реализовать глобальный радиометр на рабочих частотах действующих группировок GPS, ГЛОНАСС и GALILEO. В течение одного часа приемная аппаратура регистрирует около 10-12 сканирующих серий сигналов, трансформируемых исследуемым объектом. Частота регистрируемых изменений параметров составляет 1 Гц [9]. Восстановление временных изменений характеристик сред земных покровов теоретически возможно во временных интервалах с дискретностью в 1 секунду. Пространственная локализация этих изменений ограничена точностью координатных измерений действующих навигационных систем.

Специфика взаимодействия сигналов ГНСС с объектами зондирования определяется параметрами зондирующего излучения и электрофизическими характеристиками земных покровов.

**Метод измерений коэффициентов погонного ослабления.** Радиоскопия лесных массивов осуществляется регистрацией сигналов, изменяющих свои характеристики в процессе прохождения древостоя. На рис. 1 представлена схема измерений в варианте расположения антенны вблизи границы лесного массива.

Пространственно-временные координаты положения навигационных спутников в сочетании с координатами расположения антенны и границы лесного массива позволяют однозначно определить траекторию трассы прохождения сигнала сквозь лесной массив и её перемещения в объеме древостоя. Это позволяет получать коэффициент погонного ослабления с последующим восстановлением фитомассы и влагосодержания древостоя с координатной привязкой [10]. Длина трассы (L), которую прошел сигнал в лесу, рассчитывается по формуле:

$$L = \frac{(H - h)\cos(\alpha) - dtg(\varphi)}{\cos(\alpha)\sin(\varphi)}, \quad (1)$$

где H – средняя высота деревьев; h – высота расположения антенны; d – расстояние от антенны до кромки леса; α – азимут спутника; φ – угол возвышения спутника.

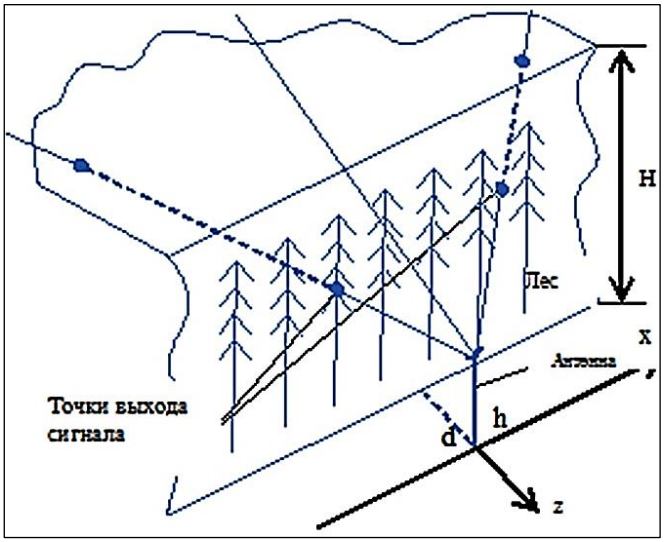



Рис. 1. Схема эксперимента по радиопросвечиванию древостоя.

Сигнал, прошедший через свободное пространство, аппроксимируем с учетом формы полной кривой, полученной ранее для каждого спутника в базе данных характеристик прямых сигналов группировок ГНСС. Далее, с учетом положения максимума амплитуды при максимальном значении угла возвышения, экстраполируем функцию зависимости амплитуды прямого сигнала от времени на время прохождения сигнала через древостой. Для выбранных точек (кадров) на рис. 3,*б* определяем амплитуды сигнала и соответствующие им длины трасс в лесу и рассчитываем значения коэффициента погонного ослабления γ:

$$\gamma = -\frac{10ln\frac{P}{P_0}}{L}, \quad (2)$$

где P – амплитуда сигнала, прошедшего через древостой; $P_0$ – амплитуда сигнала свободного пространства.

Изменения амплитуды сигналов ГНСС, возникающих при взаимодействии с объемом древостоя и учет вариаций характеристик прямых сигналов, позволяют восстанавливать его фитомассу и влагосодержание.

**Натурные измерения.** Измерения коэффициентов ослабления сигналов GPS и ГЛОНАСС проведены на участке Мининского лесничества 29.05.2019 г. Участок смешанного леса, с преобладанием в древостое пихты и кедра. Категория защитности: леса зеленой зоны поселения, хозяйственный объект. Средний возраст пихты и кедра 120 и 140 лет, высоты в интервале 18-24 м, объем древесины 210 м$^3$/га. Место расположения антенны выбрано на просеке шириной 8 м в массиве леса, пораженного полиграфом уссурийским, с координатами 56°02'25.7"N 92°25'17.9"E. Спутниковый снимок соответствует здоровому состоянию лесного массива до 2012 г. (рис. 2,*а*). Современное состояние фрагмента древостоя представлено на рис. 2,*б*.

В измерениях использована антенна АМ 475 ГНСС L1/L2/L3, и специализированный автономный приемник-регистратор НСРП-04 Регистрировался сигнал диапазона L1 систем GPS, ГЛОНАСС, GALILEO. Высота расположения антенны над уровнем лесной подстилки 0,25 м. Ось диаграммы направленности антенны ориентирована вертикально. Антенна находилась на середине просеки, разделяющий лесной массив.

Продолжительность непрерывной регистрации составляла 2 часа 55 минут. Частота отсчетов составляла 1 Гц, число спутников от каждой группировки GPS, ГЛОНАСС, GALILEO за сеанс регистрации составило 18, 13 и 11 соответственно. Расположение антенны обеспечивало регистрацию сигналов ГЛОНАСС, GPS и Galileo, прошедших через массив леса и частично свободное пространство, в интервале углов азимута 0-360$^o$ и углов возвышения 10-90$^o$. Максимальная длина трассы прохождения сигнала в лесном массиве составляет около 100 м. Полярная диаграмма траекторий трех спутников и графики изменения амплитуды сигналов от времени (шкала кадров, шаг 1 с.) представлены на рис. 3.




*а* *б*

Рис. 2. Тестовая площадка «Мининское лесничество»: *а* – спутниковый снимок неповрежденного древостоя, июнь 2012 года; *б* – фотография текущего состояния 29.05.2019 г.

**Результаты и обсуждение.** Анализ записей сигналов радиопросвечивания позволил выявить динамику амплитуды сигналов спутников в соответствии со степенью нарушенности древостоя восточного (ГЛОНАСС 10) и западного (GPS 18, GALILEO 11) участков. Отличие древостоя участков состояло в различной плотности деревьев пораженной пихты и здоровых кедров. Изменения амплитуды зондирующего сигнала определяются плотностью фитомассы стволовой древесины и её влагосодержания. Среднее значение влагосодержания образцов древесины пораженных деревьев по зондируемым участка составило 27%.

Угловые координаты источника сигналов азимут α, угол возвышения φ, средняя высота древостоя *H*; высота расположения антенны *h* и расстояние от антенны до кромки леса *d* позволяют определить погонный коэффициент ослабления γ с координатной привязкой траектории прохождения сигнала к точке размещения антенны.

Результаты обработки данных трех спутников для траекторий сигнала, дающих значения амплитуд в точках 1-8, указанных на рис. 3,*б*, представлены в таблице. Относительные изменения этих коэффициентов позволяют фиксировать границу здоровых и пораженных участков древостоя, их временную динамику. Имеется возможность оперативного восстановления влажности древостоя с использованием диэлектрической модели, адаптированной к видовой и возрастной особенностям древостоя, его состояния и рельефу почвы.

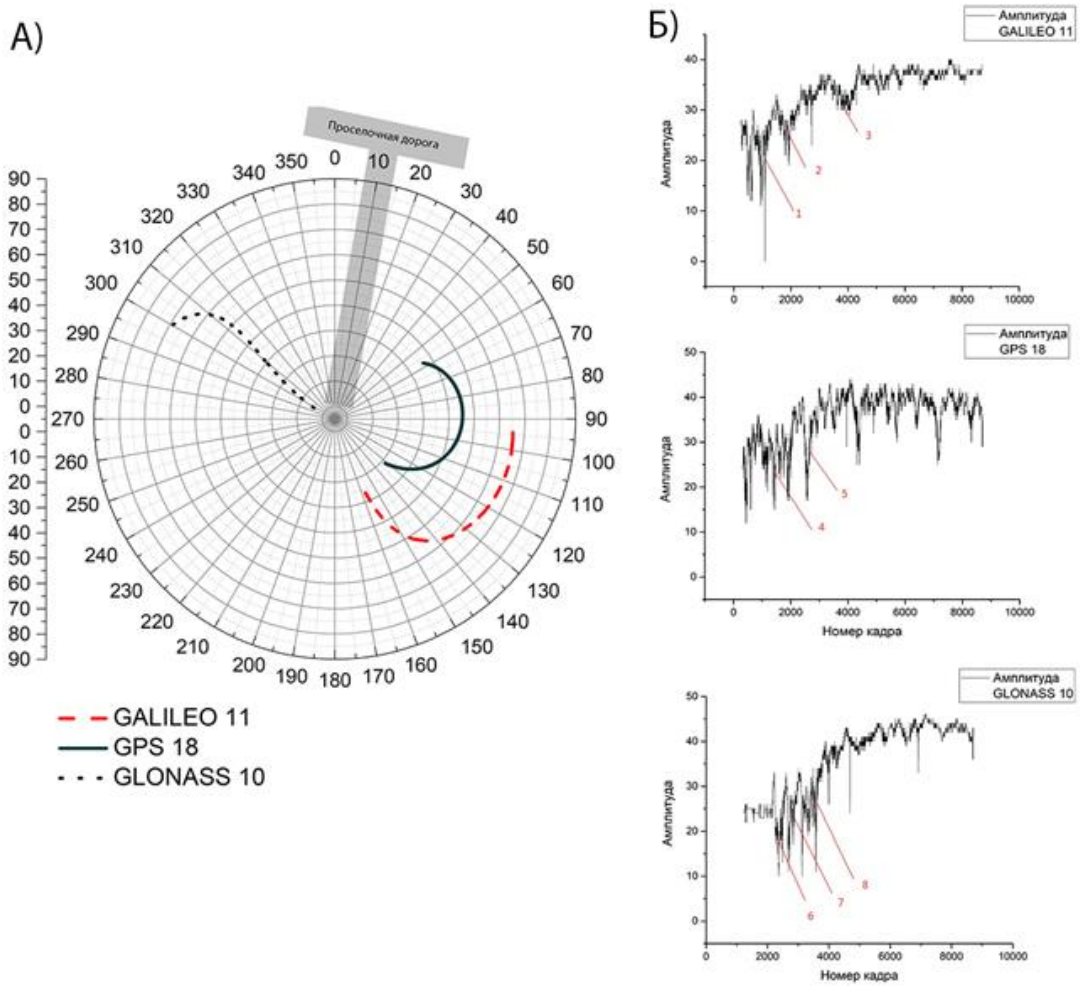



Рис. 3. Полярная диаграмма траектории спутников GALILEO 11, GPS18, GLONASS 10 (*а*) и амплитуда сигналов спутников при прохождении через древостой (*б*).

Значения коэффициентов погонного ослабления γ.

| **Спутники** | **GALILEO 11** | | | **GPS 18** | | **ГЛОНАСС 10** | | |
|---|---|---|---|---|---|---|---|---|
| Траектория | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| γ | 0,11 | 0,14 | 0,08 | 0,19 | 0.10 | 1,01 | 0,63 | 0,37 |

Вариации коэффициента погонного ослабления, представленные в таблице, связаны с наличием неоднородностей эффективной диэлектрической проницаемости древостоя, формируемых живыми кедрами и разряженными участками древостоя пораженных деревьев пихты.

**Заключение.** Рассмотрен вариант использования сигналов навигационных спутников в мониторинге состояние лесных массивов, пораженных полиграфом уссурийским. Проведены тестовые измерения погонных коэффициентов ослабления сигналов навигационных спутников диапазона L1 участков смешанного леса. Сделан вывод о возможности использования относительных изменений коэффициентов погонного ослабления с временной динамикой для определения границ здоровых и пораженных участков древостоя. Перспективна разработка диэлектрических моделей древостоя, адаптированных к его состоянию, видовым и возрастным особенностям и рельефу почвы.

## ЛИТЕРАТУРА


[1] *Иванов В.А., Москальченко С.А., Пономарёв Е.И.* Влияние нарушенных лесных территорий на частоту пожаров в Нижнем Приангарье // Хвойные бореальной зоны. СибГТУ, 2009. Т. 26, № 2. С. 249-254.

[2] *Харук В.И., Антамошкина О.А.* Воздействие сибирского шелкопряда на горимость лесных территорий // Сибирский экологический журнал. 2017. Т. 24, № 5. С. 647-654.

[3] *Conard S.G., Ivanova G.A.* Wildfire in Russian boreal forests-potential impacts of fire regime characteristics on emissions and global carbon balance estimates // Environmental Pollution. 1997. Vol. 98. P. 305-313.

[4] *Loupian E.A., Mazurov A.A., Flitman E.V. et al.* Satellite Monitoring Of Forest Fires in Russia at Federal And Regional Levels // Mitigation and Adaptation Strategies for Global Change. 2006. Vol. 11, N 1. P. 113-145.

[5] *Пономарев Е.И., Харук В.И.* Горимость лесов Алтае-Саянского региона Сибири в условиях наблюдаемых изменений климата // Сибирский экологический журнал. 2016. № 1. С. 38-46. DOI: 10.15372/SEJ20160104.

[6] *Сухинин А.И., Пономарев Е.И.* Картирование и краткосрочное прогнозирование пожарной опасности в лесах Восточной Сибири по спутниковым данным. // Сибирский экологический журнал. 2003. Т. 10, № 6. С. 669-675.

[7] *Сухинин А., Пономарев Е.* Оценка влагосодержания лесных горючих материалов по радиационной температуре // Деп. в ВИНИТИ. 1998. № 1144-В98. 26 с.

[8] *Сорокин А.В., Подопригора В.Г., Фомин С.В., Савин И.В.* Эволюция характеристик сигналов навигационных спутников в лесу // Известия вузов. Физика. 2017. Т. 60, № 12/2. С. 19-22.

[9] *Jin S., Cardellach E., Xie F.* GNSS Remote Sensing. New York, London: Springer Dordrecht Heidelberg, 2014. 286 p.

[10] Пат. 2 682 718 Российской Федерации. МПК G01S 19/03 (2010.01) Способ определения коэффициентов погонного ослабления сигналов навигационных космических аппаратов в лесном массиве с координатной привязкой // *Макаров Д.С., Савин И.В., Сорокин А.В., Фомин С.В. Харламов Д.В.* № 2018119448; заявл.25.05.2018; опубл. 21.03.2019, Бюл. № 9. 8 с.

# ВЛИЯНИЕ АНТРОПОГЕННЫХ ФАКТОРОВ НА КАЧЕСТВО ВОДЫ В РЕКЕ В РАЙОНАХ С ИНТЕНСИВНОЙ УГЛЕДОБЫЧЕЙ

*Счастливцев Е.Л., Быков А.А., Юкина Н.И.*
Институт вычислительных технологий СО РАН, Россия

В работе рассмотрено влияние различных антропогенных факторов на качество воды в реке. В качестве примера, оценки загрязнений, рассмотрен бассейн реки Мрас-Су (Южный Кузбасс). Вычислены вклады антропогенных факторов: сбросов карьерных вод, выпадение загрязняющих веществ из атмосферы и вклад других антропогенных источников загрязнения.
*Ключевые слова: бассейн реки, загрязнение, ингредиенты, атмосферное выпадение, качество вод.*

Основными источниками загрязнения воды в реках являются: прямые сбросы угольно-добывающих предприятий, промпредприятий, смыв с территорий, загрязненных продуктами хозяйственной деятельности людей, а также выпадение и оседание атмосферных примесей, выбрасываемых источниками загрязнения воздуха на поверхность бассейна реки. Для примера, рассмотрим бассейн реки Мрас-Су. Река Мрас-Су является левым притоком реки Томь, берет начало с Абаканского хребта, течёт в глубокой долине по Горной Шории по территории Шорского национального парка. В 635 км от устья по левому берегу расположен г. Мыски. На рис. 1 представлен бассейн реки Мрас-Су, точками отмечены места сброса сточных вод угольно-добывающими предприятиями.

Одним из источников загрязнения реки, являются смыв с осадками с территории бассейна продуктов жизнедеятельности людей. На территории бассейна р.Мрас-Су находится 37 сельских и 2 городских поселения (Мыски и Шерегеш), всего 55388 жителей [1]. Сельские поселения в основном размещены по берегам реки. Ориентировочные поступления твердых бытовых отходов от сельских поселений можно количественно оценить на основе [2], согласно которому отходы из жилищ несортированные (исключая крупногабаритные) составляют 210-225 кг на 1 чел/год, а мусор от бытовых помещений несортированный – 40-70 кг на 1 чел/год. В табл. 1 приведена ориентировочная оценка поступления на территории бассейнов твердых отходов от сельских поселений.

Естественно, эти отходы не попадают напрямую в водоемы, но загрязняющие вещества частично смываются в реки осадками. Не исключена и фильтрация через почву, поскольку сельские поселения находятся в непосредственной близости к водоемам.

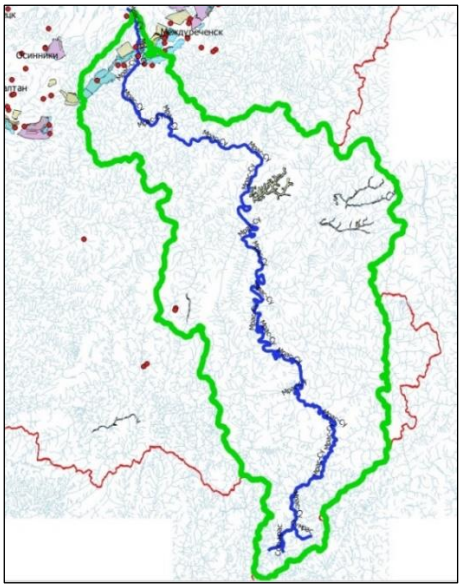


Рис. 1. Бассейн реки Мрас-Су.

Табл. 1. Оценка поступления на территории бассейнов твердых отходов от сельских поселений.

| **Бассейн реки** | **Число поселений и городов** | **Суммарное население, чел** | **Отходы, т в год** | | |
|---|---|---|---|---|---|
| | | | из жилищ | Из бытовых помещений | В сумме |
| Мрас-Су | 39 | 55 388 | 12185 | 3046 | 15231 |

Весьма значимым источником загрязнения бассейнов рек, является выпадение промышленных выбросов из атмосферы. Вклад выпадения аэрозолей из атмосферы в суммарное загрязнение бассейна реки Мрас-Су оценен с применением долгосрочной модели расчета выпадения на подстилающую поверхность [3]. Модель встроена в состав программного комплекса «ЭРА-ВОЗДУХ» (www.lpp.ru) широко используемого для проектных работ в Сибирском регионе. Это позволяет использовать накопленные в форматах комплекса нормативные исходные данные по источникам выбросов для научных исследований.

В представленных ниже расчетах использованы данные инвентаризаций источников загрязнения (4504 источника), проведенные для разработки сводных томов ПДВ городов Киселевск, Прокопьевск, Новокузнецк, Мыски и Междуреченск. Детальные данные по источникам достаточно крупных населенных пунктов от Осинников до Таштагола авторам недоступны, поэтому суммарные выбросы по этим населенным пунктам приближенно оценено на основе данных Государственного доклада [4].

Изолинии выпадения суммы всех пылевых частиц показаны на рис. 2. Видно, что области со значительным выпадением примесей находятся в северной части бассейна, вблизи устья реки Мрас-Су, где сосредоточены промышленные города и индустриальные районы. Центральная и южная части бассейна реки Мрас-Су подвержена незначительному загрязнению атмосферы и выпадению атмосферных примесей. На рис. 2 северный участок отделен от всего бассейна жирной линией, и на территорию северного участка (6% от общей площади) выпадает почти 20% всех пылевых частиц промышленного происхождения.

Аналогичные расчеты проделаны для нитратов (азотосодержащие примеси $NO_2$ + NO + $NH_3$) и сульфатов (серосодержащие примесей $SO_2$+$H_2S$). Данные примеси имеют малые скорости выпадения (точнее – имеющие размерность скорости коэффициенты улавливания поверхностью), поэтому их суммарные и удельные показатели выпадения на порядок меньше, чем для пылевых частиц. Однако интенсивность затухания загрязнения в зависимости от удаления от источников для таких примесей значительно меньше, что приводит к более равномерному загрязнению бассейна реки Мрас-Су сульфатами и нитратами.

Программа «ЭРА-ВОЗДУХ» позволяет рассчитывать распределение загрязнения атмосферы и выпадения на почву не только по регулярной сетке, но и по любому полигону. Кроме того, имеется функция по расчету интеграла выпадения частиц на полигон, что использовано для определения суммарного поступления промышленных примесей из атмосферы в бассейн реки Мрас-Су и выделенный на рис. 2 северный участок. Результаты для промышленной пыли, сульфатов и нитратов представлены в табл. 2. Для северного участка указан процентный вклад для каждого показателя по всему бассейну.

Авторами обобщены все данные о загрязнении воды в реке Мрас-Су с 2000 по 2017 года, полученные по литературным данным и в результате собственных экспедиционных исследований. По результатам анализов 12-ти ингредиентов, рассчитан ассоциативный показатель (АП) [5, 6], который равен сумме концентраций всех веществ, деленных на соответствующие предельно допустимые концентрации. Для реки Мрас-Су он равен 1.7, что соответствует III классу качества воды «умеренно загрязненная».

Ниже для сравнения приведена табл. 4 с результатами ранее проведенных авторами [7] аналогичных расчетов антропогенной нагрузки на бассейны рек Кемеровской области, расположенных в идентичных с Мрас-Су природно-климатических условиях и подверженных влиянию одних и тех же промышленных источников загрязнения атмосферы.

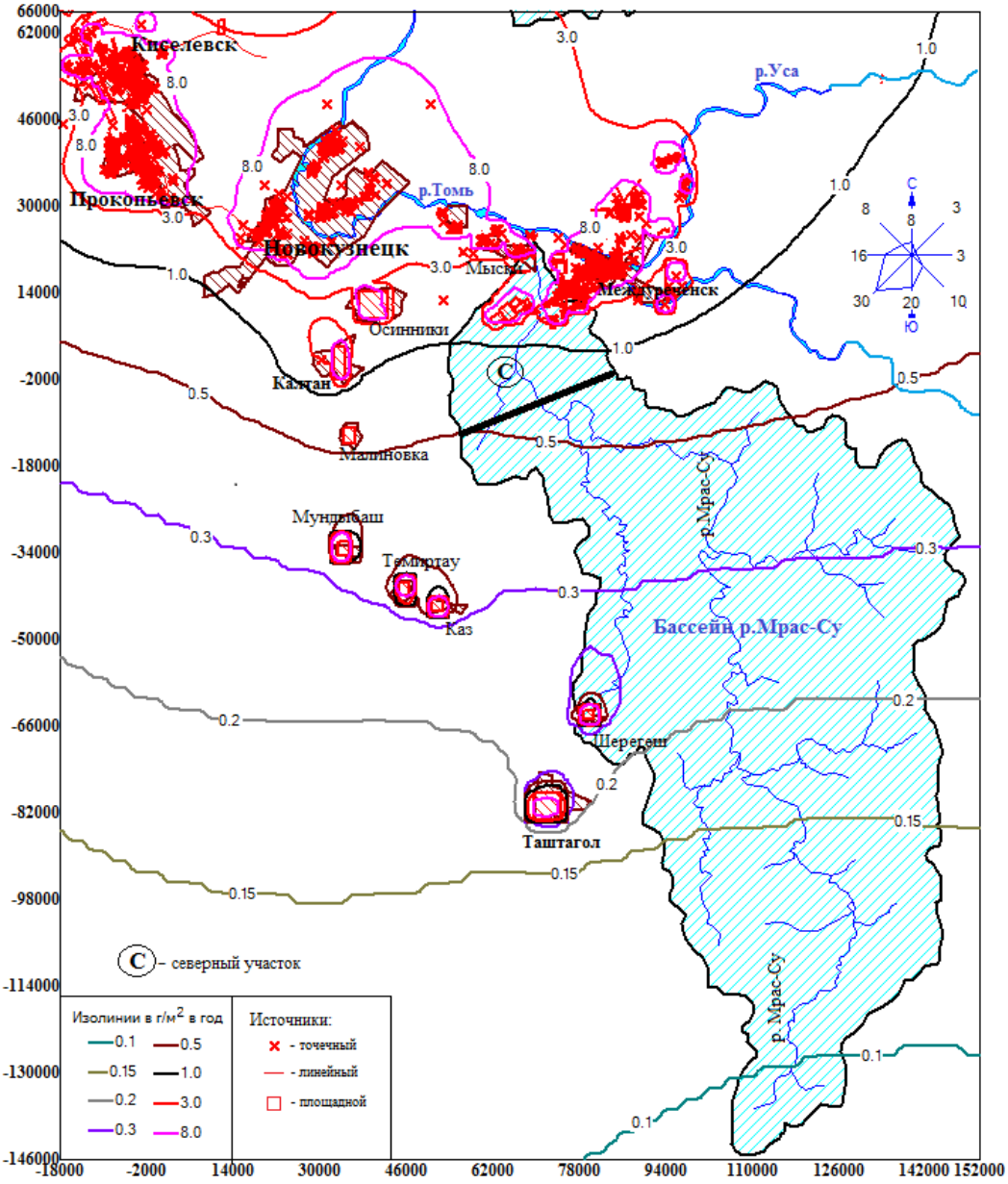



Рис. 2. Изолинии осаждения на почву промышленной пыли всех фракций и составов, выбрасываемой в атмосферу в районе бассейна реки Мрас-Су.

Табл. 2. Интегральное выпадение атмосферных промышленных выбросов на поверхность всего бассейна реки Мрас-Су и северного (индустриально нагруженного участка).

| **Участок Бассейна Мрас-Су** | **Площадь участка, км$^2$** | **Выпадение, т в год (г/м$^2$)** | | | **Суммарное выпадение на участок бассейна, т в год (г/м$^2$)** |
|---|---|---|---|---|---|
| | | Пыль | Нитраты $NO_2$+NO+$NH_3$ | Сульфаты $SO_2$+$H_2S$ | |
| Весь бассейн | 8840 | 5002 (0.56) | 185 (0.021) | 242 (0.027) | 5429 (0.61) |
| Северный участок | 530 (6%) | 932 (1.759) (19%) | 17 (0.028) 9% | 21 (0.033) 9% | 970 (1.743) 18% |

Табл. 3. Суммарная антропогенная нагрузка на бассейн р. Мрас-Су и ассоциативный показатель (АП) качества воды.

| **Река** | **Антропогенная нагрузка, тонн в год (%)** | | | **АП** |
|---|---|---|---|---|
| | Промышленность, население | Выпадение из атмосферы | Сумма | |
| Мрас-Су | 15231 (73) | 5429 (27) | 20660 (100) | 1,7 |

Табл. 4. Сравнение антропогенной нагрузки на бассейны некоторых рек Кемеровской области.

| Река | Площадь бассейна, км$^2$ | Население, чел | Атмосферное выпадение, т в год (г/м$^2$) | Суммарная антропоген. нагрузка, т в год | АП |
|---|---|---|---|---|---|
| Мрас-Су | 8840 | 55388 | 5429 (0,61) | 20626 | 1,7 |
| Верхняя Терсь | 1068 | 998 | 1497 (1,40) | 1744 | 4.8 |
| Средняя Терсь | 2022 | 124 | 2184 (1.08) | 2218 | 2.4 |
| Тайдон | 2258 | 0 | 1363 (0,60) | 1363 | 2.4 |

**Заключение.** В работе рассмотрено влияние различных антропогенных факторов на качество воды в реках. Предложен подход к расчету показателей интегральной антропогенной нагрузки на водные бассейны рек, расположенных вблизи индустриально развитых территорий. Приведен пример практических расчетов суммарной и удельной нагрузки для бассейна реки Мрас-Су, расположенной в южной части Кемеровской области. Дано сравнение показателей нагрузки на бассейн с обобщенной оценкой загрязнения воды в реке.

Установлено, что наличие угледобывающих предприятий и городской промышленной агломерации (Мыски, Междуреченск) вблизи устья реки, вызывает достаточно высокую локальную антропогенную нагрузку на северный участок бассейна, качество воды в р. Мрас-Су остается достаточно чистым и оценивается, как «умеренно загрязненное». Этому способствует большая площадь бассейна и незначительная нагрузка в его центральной и южной частях.

Сравнение расчетов для Мрас-Су с другими реками Кемеровской области позволяет заметить, что наилучшим образом ассоциативный показатель загрязнения воды коррелирует с показателем суммарного удельного выпадения (г/м$^2$) промышленных атмосферных выбросов на площади бассейнов рек.

## СПИСОК ЛИТЕРАТУРЫ


[1] Всероссийская перепись населения 2010 года. Кемеровская область. 1.6. Численность населения городских округов, муниципальных районов, городских и сельских поселений, городских населённых пунктов, сельских населённых пунктов.

[2] СанПиН 2.1.7.1322-03. Гигиенические требования к размещению и обезвреживанию отходов производства и потребления от 15.06.2003г. **http://docs.cntd.ru/document/901862232** (Дата обращения 15.11.2018).

[3] *Быков А.А., Счастливцев Е.Л., Пушкин С.Г.* Особенности построения и практического применения локальной модели загрязнений почвы техногенными выбросами пылевых частиц // Проблемы безопасности и чрезвычайных ситуаций. 2007. № 4. С. 74-82.

[4] Доклад о состоянии и охране окружающей среды Кемеровской области в 2017 году. Кемерово, 2018г. 483 с.

[5] *Потапов В.П., Мазикин В.П., Счастливцев Е.Л., Вашлаева Н.Ю.* Геоэкология угледобывающих районов Кузбасса. Новосибирск: Наука, 2005. 660 с.

[6] *Счастливцев Е.Л., Юкина Н.И., Харлампенков И.Е.* Информационно-аналитическая система геоэкологического мониторинга водных ресурсов угледобывающего региона // Вестник КузГТУ. Кемерово. 2016. № 2 (114). С. 157-164.

[7] *Счастливцев Е.Л., Быков А.А., Юкина Н.И.* Расчетная оценка антропогенной нагрузки на бассейны рек с учетом осаждения из атмосферы промышленных выбросов // VI Всероссийская конференция «Безопасность и мониторинг техногенных и природных систем»: Материалы и доклады. Красноярск. 2018. С. 446-451.

# ОТСЛЕЖИВАНИЕ ПЕРЕМЕЩЕНИЙ ДИКИХ ЖИВОТНЫХ ПО ВИЗУАЛЬНЫМ ДАННЫМ ФОТОЛОВУШЕК НА ОСНОВЕ СЕТЕЙ ГЛУБОКОГО ОБУЧЕНИЯ

*Фаворская М.Н., Пахирка А.И., Зотин А.Г.*
ФГБОУ ВО Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева, Красноярск

Отслеживание перемещений диких животных в природных парках предполагает анализ большого объема визуальных данных, полученных от нескольких фотоловушек. Ставится задача сопоставления изображений конкретной особи с видеоматериалами данного вида животных. Накопленный объем видеоматериалов позволяет решить задачу с использованием технологии Deep Learning, в частности, на основе Siamese Neural Network. Метод протестирован с использованием набора изображений, полученных на территории природного парка «Ергаки», Красноярский край, Россия.
*Ключевые слова: мониторинг диких животных, фотоловушки, обработка изображений, глубокое обучение.*

**Введение.** Использование фотоловушек для мониторинга диких животных становится все более популярным способом неинвазивного наблюдения за дикими животными в их среде обитания. Проблема заключается в обработке больших объемов информации, полученных, например, в течение полугода от нескольких десятков фотоловушек, установленных в национальных парках и на особо охраняемых природных территориях. Выбор мест установки определяется работниками парков в местах троп, водопоев, солонцов и т.д. Ставится задача сопоставления визуальных данных конкретной особи с видеоматериалами, полученными от распределенной сети фотоловушек, с целью нахождения локальных миграций данной особи. Такая задача является одной из завершающих задач мониторинга диких животных при условии, что выполнен предварительный отсев малоинформативных снимков и распознан вид животного. Большой накопленный объем видеоматериалов позволяет решить эту проблему с использованием технологии Deep Learning, в частности с использованием сети глубокого обучения Siamese Neural Network. Метод протестирован с использованием набора данных изображений, полученных на территории природного парка «Ергаки», Красноярский край, Россия.

**Постановка задачи.** Известно, что каждое животное имеет свой ареал обитания, и поэтому его появление может быть зафиксировано ограниченным количеством близлежащих фотоловушек $N_F$. В силу удаленности территорий национальных парков фотоловушки не связаны в единую информационную сеть и не могут передавать текущую информацию на сервер. Они способны сохранять серию снимков в имеющемся накопителе после срабатывания датчика движения. Обычно длительность между отснятыми снимками составляет 3-5 с, что позволяет получить несколько изображений особи в разных положениях. Фотоловушки срабатывают в любое время суток и при любых метеорологических условиях, что приводит к появлению плохих по качеству снимков или снимков с отсутствием животного. Следует отметить, что появление человека фиксируется аналогичным образом. Серия снимков $S_{im} = \{Im_1, Im_2, Im_3, \ldots, Im_{Ns}\}$, где $IM_i$ – $i$-й снимок, $Ns$ – количество снимков в серии, имеет атрибуты дата $D_s$, время $T_s$, температура $K_s$. Примем, что срабатывание фотоловушки инициируется одним визуальным объектом $VO_i \in \{VO_1, VO_2, \ldots, VO_{Na}\}$, где $VO_{Na}$ – количество классов, включая класс неопознанного визуального объекта. Таким образом, мощность множества полученных снимков $IM_T$ за установленный временной интервал определяется как сумма:

$$\{IM_T\} = \sum_{i=1}^{N_F} count\left(S_{im_i}\right),$$

где $count\,(\cdot)$ – функция подсчета количества снимков в $i$-й серии.

Примем, что из полученного множества снимков удалены неинформативные снимки [7], содержащие отсутствие визуального объекта или значительные артефакты, вызванные условиями съемки. При этом неиформативные снимки могут составлять от 60% до 75% от общего

объема видеоматериалов [1]. К оставшемуся множеству снимков $IM_S$ применяется процедура составления модели фона (для конкретной фотоловушки), что позволяет достаточно быстро и точно находить визуальный объект [2]. Отобранное множество снимков содержит подмножества изображений видов $IM_A$ и подмножество изображений человека $IM_H$:

$$IM_S = \left( IM_{A_1}, IM_{A_2}, \ldots, IM_{A_M} \right) \bigcup IM_H ,$$

где $M$ – количество видов (классов) животных.

Авторами разработан метод распознавания видов животных с использованием объединенной сверточной нейронной сети, включающей две структуры VGG16 для распознавания изображений головы и части тела животного и одну структуру VGG19 для распознавания изображения животного в целом [3]. Выходы трех структур объединяются, что позволило получить точность распознавания 80,6% Top-1 и 94,1% Top-5 на сбалансированном наборе данных. Для несбалансированного набора данных (что является характерной ситуацией для данной задачи в силу количества животных на конкретной территории и особенностей их поведения) были получены худшие результаты, а именно 38,7% Top-1 и 54,8% Top-5. Тем не менее, данный подход позволяет разделить подмножества изображений по видам животных. Таким образом, мы получаем подмножество заданного вида $IM_{Ai}$, которое содержит изображения нескольких особей, относящихся к этому виду. Данная информация является исходной для идентификации конкретной особи и отслеживания ее перемещения с использованием видеоматериалов, полученных от $N_F$ фотоловушек.

**Технологии Deep Learning.** Технологии глубокого обучения (Deep Learning) позволяют автоматически извлекать абстрактные признаки из исходных («сырых») данных за счет того, что каждый слой нейронов использует операции свертки на перекрывающихся небольших по размеру регионах, поступающих от предыдущих слоев. Причем, на первых сверточных слоях размер таких регионов больше, чем на последующих сверточных слоях сети, так как пространственная область сокращается от слоя к слою и применение большой по размеру маски приведет к пропуску структурных особенностей. Последним слоем глубокой нейронной сети, как правило, является слой softmax, выполняющий роль классификатора. Однако для обучения таких сетей требуется большой объем исходных данных, промаркированный вручную. Отметим, что задача обнаружения и распознавания диких животных удовлетворяет таким условиям и поэтому может быть решена с помощью сверточных нейронных сетей.

Применение технологий глубокого обучения для классификации изображений, полученных от фотоловушек, началось несколько лет назад, и с появлением новых архитектур сверточных нейронных сетей точность распознавания повышалась. Одной из первых была работа [8], в которой ставилась задача автоматической идентификации 20 классов животных при наличии 20 000 изображений. Была достигнута точность распознавания около 38%, что, тем не менее, превышало точность классификации с применением традиционной технологии на основе «портфеля слов». В [9] было показано, что технологию Deep Learning можно успешно применять на небольших наборах изображений, если требуется отделить птиц от млекопитающих (набор данных из 1572 изображений) и распознать два вида млекопитающих (набор данных из 2597 изображений). Сеть предварительно была обучена на наборе данных ImageNet [10]. Как для обнаружения животных на снимках, так и для распознавания их видов применяются различные сети глубокого обучения, например, AlexNet (8 слоев) [11], NiN (Network in Network) (16 слоев) [12], VGG (Visual Geometry Group, Department of Engineering Science, University of Oxford) (22 слоя) [4], GoogLeNet (32 слоя) [13], а также ResNet (18, 34, 50, 101 и 152 слоя) [14]. Точность детектирования при использовании современных сверточных нейронных сетей высока и может достигать 93,8% Top-1 и 98,8% Top-5 [5], однако следует понимать, что такие высокие значения получаются для набора изображений животных, обитающих на конкретной территории в конкретных климатических условиях и при наличии хорошо обученной сети.

**Метод идентификации особи по изображениям.** Помимо сверточных нейронных сетей с одним входом в последние годы появились такие архитектуры, которые предполагают подачу двух или даже трех изображений одновременно на две или три ветви сети. Такие сети называются двойными (сиамскими) или тройными. При идентификации конкретной особи из ограниченного набора изображений животных данного вида сиамская структура сети будет полезной.

Архитектура сиамской сети впервые была предложена в 1994 г. для верификации подписи, представленной временным рядом [15]. Позже эта архитектура усилена сверточными нейронными сетями для верификации лиц с использованием техники уменьшения размерности [16]. В последние годы сиамские сверточные сети стали использоваться для повторной идентификации людей, лиц, а также как системы сопровождения объектов.

Архитектура сиамской сети такова, что обучение осуществляется путем подачи на входы двух изображений одновременно. Причем, используются две идентичные параллельные нейронные сети, содержащие одинаковые наборы весовых коэффициентов $W$. Общая структура сиамской нейронной сети представлена на рис. 1. Однако следует отметить, что имеются другие решения, когда традиционная структура данной сети изменяется путем применения неодинаковых наборов весовых коэффициентов [6].

На входы обоих сетей поступают входные изображения $X_1$ и $X_2$, их выходами являются векторы признаков $O_X1$ и $O_X2$. Далее вычисляется расстояние между этими векторами и формируется функция $E_W(X_1, X_2)$, которая и характеризует сходство или различие входных изображений. В качестве сетей выбрана архитектура VGG19 (Visual Geometry Group), которая хорошо зарекомендовала себя в задачах распознавания благодаря однотипным конволюционным слоям 3×3 элементов (Conv), слоям подвыборки (Maxpooling) и полносвязным слоям (Fullyconnected) из 4096 нейронов, как показано на рис. 2.

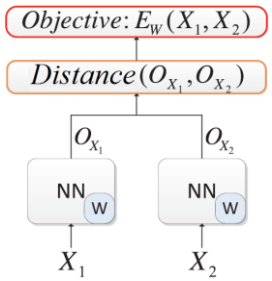



Рис. 1. Общая структура сиамской нейронной сети.

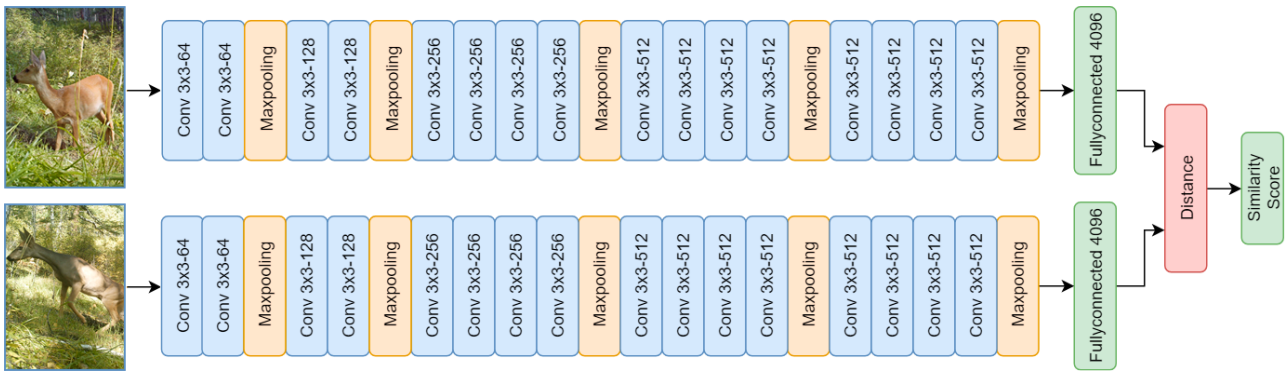



Рис. 2. Структура сиамской нейронной сети на основе VGG19.

Результаты точности сопоставления особей косули.

| | **По одному снимку, %** | **По серии из 5 снимков, %** |
|---|---|---|
| **Фотоловушка 1** | 41,2 | 56,2 |
| **Фотоловушка 2** | 42,4 | 58,4 |
| **Фотоловушка 3** | 43,1 | 57,5 |
| **Фотоловушка 4** | 44,6 | 59,0 |

**Экспериментальные результаты.** Для обучения сиамской сверточной сети использовался набор из 842 изображений косуль, полученных на территории природного парка «Ергаки», Красноярский край, Россия. Данный набор содержит изображения 8 особей косули, полученных от 4 фотоловушек, расположенных в разных местах природного парка. На основе исходного набора была сгенерирована обучающая выборка размером порядка 12000 изображений, содержащая случайным образом выбранные пары изображений одной и той же особи, а также разных особей. Далее обучающая выборка была искусственно расширена до 40000 изображений с помощью аугментации (augmentation). Для аугментации применялись следующие действия:

- Поворот изображения на [–45, +45] градусов случайным образом.
- Горизонтальное отражение изображения со случайным сдвигом.
- Масштабирование изображения с коэффициентом [0,75; 1,25].
- Кадрирование изображения.
- Сдвиг каждого цветового канала на [–25,5; +25,5] единиц случайным образом.

Далее полученный набор изображений был разделен на обучающую и тестовую выборки в соотношении 80% на 20%. Сиамская сверточная нейронная сеть была реализована на языке программирования Python 3.5 с использованием фреймворка TensorFlow 1.5. Эксперименты проводились на ПК с процессором Intel i7 3.2GHz, 16GB оперативной памяти, видеокартой 8GB Nvidia GeForce GTX 1070 под операционной системой Windows 7.

Для отслеживания перемещения животного по природному парку требовалось сопоставить изображение конкретной особи с изображениями всех особей данного вида, имеющихся в базе данных. Таким образом, результатом сиамской сверточной нейронной сети являлась оценка подобия двух входных изображений на основе эвклидовой метрики. Результаты сопоставления конкретных особей косули с привязкой к четырем фотоловушкам представлены в таблице.

Среднее значение точности сопоставления особей по одному снимку составляет 42,8%, а по серии из пяти снимков – 57,8%. Более точного сопоставления можно добиться увеличением серии тестовых снимков для конкретной особи.

**Заключение.** Предложенный метод сопоставления визуальных данных конкретной особи с видеоматериалами, полученными от распределенной сети фотоловушек, основан на использовании сиамской сверточной нейронной сети. Предполагается, что изображений особей конкретного вида отобраны, и требуется провести сопоставление снимков для определения миграции особи по природному парку с точностью до мест, в которых установлены фотоловушки. Задача является сложной в силу различных условий съемки и малого количества изображений, пригодных для распознавания конкретной особи. Относительно низкая точность распознавания конкретной особи обусловлена небольшим набором исходных изображений и может быть повышена за счет большей исходной выборки.

*Исследование выполнено при финансовой поддержке РФФИ, Правительства Красноярского края, Красноярского краевого фонда науки в рамках научного проекта (грант № 18-47-240001-а).*

## ЛИТЕРАТУРА

[1] *Newey S., Davidson P., Nazir S., Fairhurst G., Verdicchio F., Irvine R.J., van der Wal R.* Limitations of recreational camera traps for wildlife management and conservation research: A practitioner's perspective // Ambio. 2015. Vol. 44, pp. 624-635.

[2] *Favorskaya M.N., Buryachenko V.V.* Background extraction method for analysis of natural images captured by camera traps // Informatsionno-upravliaiushchie sistemy [Information and Control Systems]. 2018. Vol. 6, pp. 35-45.

[3] *Favorskaya M., Pakhirka A.* Animal species recognition in the wildlife based on muzzle and shape features using joint CNN. Procedia Computer Science. 2019. (in print).

[4] *Simonyan K., Zisserman A.* Very deep convolutional networks for large-scale image recognition. 2014. arXiv preprint arXiv:1409.1556
[5] *Norouzzadeh M.S., Nguyen A., Kosmala M., Swanson A., Palmer M., Packer C., Clune J.* Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning // PNAS. 2018. Vol. 115, no. 25. pp. E5716-E5725.
[6] *Shaham U., Lederman R.R.* Learning by coincidence: Siamese networks and common variable learning // Pattern Recognition. 2018. Vol. 74. pp. 52-63.
[7] *Favorskaya M.N., Buryachenko V.V.* Selecting informative samples for animal recognition in the wildlife // In: Czarnowski I., Howlett R., Jain L. (eds) Intelligent Decision Technologies 2019. SIST, vol 143, Malta: Springer, 2019. pp. 65-75.
[8] *Chen G., Han T.X., He Z., Kays R., Forrester T.* Deep convolutional neural network based species recognition for wild animal monitoring // Proceedings of the 2014 IEEE International Conference on Image Processing. Paris, France: IEEE, 2014. pp. 858-862.
[9] *Gomez A., Diez G., Salazar A., Diaz A.* Animal identification in low quality camera-trap images using very deep convolutional neural networks and confidence thresholds // Proceedings of the International Symposium on Visual Computing. Las Vegas, Nevada, USA: Springer, 2016. pp. 747-756.
[10] *Deng J., Dong W., Socher R., Li L.-J., Li K., Fei-Fei L.* Imagenet: A large-scale hierarchical image database // Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA: IEEE, 2009. pp. 248-255.
[11] *Krizhevsky A., Sutskever I., Hinton G.E.* Imagenet classification with deep convolutional neural networks // Proceedings of the 25th International Conference on Neural Information Processing Systems, Vol. 1, Nevads, USA: IEEE, 2012. pp. 1097-1105.
[12] *Lin M., Chen Q., Yan S.* Network in network // Proceedings of the International Conference on Learning Representations, Banff, Canada: IEEE, 2014. pp. 1-10.
[13] *Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A.* Going deeper with convolutions // Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA: IEEE, 2015. pp. 1-9.
[14] *He K. Zhang X., Ren S., Sun J.* Deep residual learning for image recognition // Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition Las Vegas, NV, USA: IEEE, 2016. pp. 770-778.
[15] *Bromley J., Guyon I., Lecun Y., Säckinger E., Shah R.* Signature verification using a "Siamese" time delay neural network // Proceedings of the Neural Information Processing Systems, Denver, Colorado, USA: IEEE, 1994. pp. 737–744.
[16] *Chopra S., Hadsell R., LeCun Y.* Learning a similarity metric discriminatively, with application to face verification // Proceedings of the 2005 IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, San Diego, CA, USA: IEEE, 2005. pp. 539–546.

# АНАЛИЗ ВЛИЯНИЯ ФАКТОРОВ ОКРУЖАЮЩЕЙ СРЕДЫ НА СМЕРТНОСТЬ НАСЕЛЕНИЯ Г. КРАСНОЯРСКА

*Черных Д.А.[(1)(2)], Тасейко О.В.[(1)(2)], Иванова У.С.[(1)(3)]*
[(1)] Красноярский филиал Института вычислительных технологий СО РАН, Красноярск
[(2)] Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, Красноярск
[(3)] Сибирский федеральный университет, Красноярск

В работе выполнен анализ влияния факторов окружающей среды на смертность населения г. Красноярска методами многомерного статистического анализа. Объектом исследования являлась смертность жителей г. Красноярска в возрастной группе от 60 до 74 лет от наиболее распространенных заболеваний, характеризующихся чувствительностью к климатическим факторам за период с 2010 по 2014 гг. Цель исследования – оценить степень совместного влияния климатических параметров и факторов окружающей среды на смертность населения г. Красноярска.
*Ключевые слова: качество атмосферного воздуха, климатические параметры, многомерная регрессионная модель, смертность населения.*

Рост городских территорий позволил расширить доступ ко многим объектам, таким как транспорт, энергетика, образование, водоснабжение и т.д. Как следствие, автомобильный и промышленный рост в сочетании с неблагоприятными метеорологическими условиями вызвал несколько всемирных эпизодов чрезмерного загрязнения воздуха с потерями жизней и ущербом здоровью. В качестве примера можно привести хорошо известную катастрофу в американском городе Донора, штат Пенсильвания в 1948 г., в результате которого погибли 20 человек, и заболело более 5910 человек [1]. С тех пор все большее внимание стало уделяться изучению воздействия загрязнения воздуха на здоровье населения. Было проведено множество эпидемиологических исследований, которые показали, что загрязнение воздуха влияет на здоровье человека, особенно при респираторных и сердечно-сосудистых заболеваниях, даже если уровни концентрации загрязнителей ниже стандартных уровней качества воздуха [2-7].

При взаимодействии антропогенных и природных факторов нарастают риски устойчивого развития территориальных образований, при этом активно проявляются тенденции, урбанизирующие (преобразующие) природную среду, создавая новые виды природных и техногенных угроз. Оценки темпов и уровня устойчивого развития требуют использования методологии учета различных факторов индивидуального и коллективного рисков [8-11].

Климат Красноярского края характеризуется наличием высоких температур летом и экстремально низких зимой и относится преимущественно к резко континентальному типу. В связи с такими климатическими особенностями, а также наличием промышленных агломераций для Красноярского края актуален анализ синергетического влияния климатических параметров и факторов окружающей среды на уровни смертности населения. Для решения данной проблемы возникает необходимость в разработке методологии оценки рисков от многофакторного воздействия на смертность населения, с использованием многомерных статистических моделей, учитывающие особенности региона и структуру населения, а также в разработке мер защиты населения от влияния факторов окружающей среды.

Оценка воздействия загрязнения воздуха на здоровье человека является сложной в связи с тем, что некоторые характеристики населения (возраст, генетика, социальные условия и т.д.) влияют на реакцию организма под воздействием загрязненного воздуха. Например, было показано, что более высокая концентрация загрязнения воздуха увеличивает количество респираторных заболеваний у пожилых и детей, при этом дети более восприимчивы, так как им нужно вдвое больше воздуха, вдыхаемого взрослыми [2, 3]. В то время как для пожилых людей более значимое влияние оказывают слабые иммунная и дыхательная системы, а также тот факт, что они подвергались значительному загрязнению воздуха на протяжении всей жизни.

Другим важным фактором является генетика. Исследования показали, что люди с хроническими заболеваниями или аллергиями, такими как бронхит и астма, более чувствительны к загрязнению воздуха.

Цель исследования – оценить степень совместного влияния климатических параметров и факторов окружающей среды на смертность населения г. Красноярска.

**Материалы и методы.** Объектом исследования являлась смертность жителей г. Красноярска в возрастной группе от 60 до 74 лет от наиболее распространенных заболеваний, характеризующихся чувствительностью к климатическим факторам за период с 2010 по 2014 гг.:

- инфаркты (I20 – I25);
- инсульты (I60 – I69);
- органы дыхания (J00 – J22, J30, J40 – J45);
- внешние причины (V01 – Y98).

Для решения поставленной задачи использовалась база данных показателей смертности, предоставленная Территориальным органом Федеральной службы государственной статистики по Красноярскому краю. Ежедневная смертность в г. Красноярске изучалась за 5 лет (с 1 января 2010 по 31 декабря 2014 года).

Идентификация климатических характеристик выполнялась с применением методов дисперсионного анализа независимых выборок для оценки влияния коротких (дискретных) погодных эпизодов – волн жары и холода и резких перепадов температуры в течение суток на смертность.

Зависимость смертности от загрязнения атмосферного воздуха и климатических факторов рекомендуется изучать с помощью Пуассоновской регрессионной модели (1-2) [12]

$$ln(\mu_t) = \beta_0 + \beta_1 \cdot X_{1,t-L} + \cdots + \beta_k \cdot X_{k,t-L} + \beta_{k+1} \quad (1)$$

$$\mu_t = e^{\beta_0} \cdot e^{\beta_1 \cdot X_{1,t-L}} \cdot \ldots \cdot e^{\beta_k \cdot X_{k,t-L}} \cdot e^{\beta_k} \quad (2)$$

где $\mu_t$ – прогнозируемый параметр (пуассоновская переменная); $x_{1,t} \ldots x_{k,t}$ – влияющие переменные; $L$ – отставание негативного эффекта во времени (лаг); $\beta_1 \ldots \beta_{k+1}$ – параметры модели; $\beta_0$ – свободный член. Свободный член не связан с влияющими переменными и характеризует фоновые колебания пуассоновской переменной.

Модель предусматривает расчет коэффициента при различных лагах (временной задержки реакции организма на негативное воздействие) от нуля до 14 дней.

Для оценки коэффициента предполагаемой пуассоновской модели также использовались среднесуточные концентрации загрязняющих веществ и метеорологических параметров за период с 2010 по 2014 гг.:

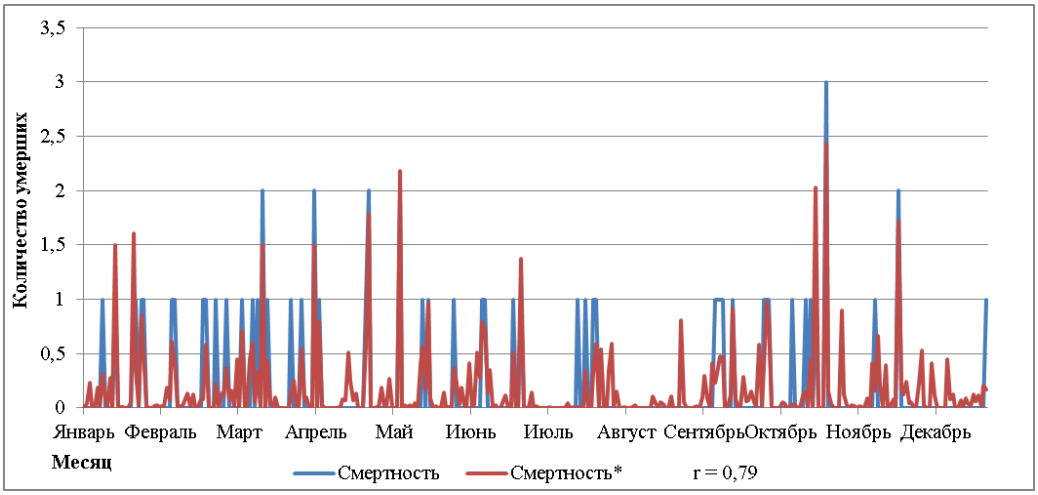
1. Концентрации загрязняющих веществ:
   - диоксид азота;
   - формальдегид;
   - пыль. Данные о ежедневной концентрации загрязняющих веществ города Красноярска были предоставлены Федеральным государственным бюджетным учреждением «Среднесибирское управление по Гидрометеорологии и Мониторингу Окружающей Среды».
2. Метеорологические показатели:
   - температура атмосферного воздуха;
   - относительная влажность;
   - экстремальные перепады температур в течение суток;
   - температурные волны. Для оценки метеорологических параметров использовались данные государственной наблюдательной сети метеостанции «Опытное поле», которая является фоновой для территории г. Красноярска [13].

**Результаты исследования.** В табл. 1 приведены параметры пуассоновской регрессионной модели для оценки влияния факторов окружающей среды на смертность от болезней органов дыхания у женщин.

Табл. 1. Расчетные коэффициенты модели смертности от болезней органов дыхания у женщин.

| **Влияющие переменные** | **Оценка параметра, $\beta_i$** | **Лаг, дни** | **Стандартная ошибка** | **p-value** |
|---|---|---|---|---|
| Свободный член | -7,6 | - | 5,5 | 0,2* |
| Формальдегид (*F*) | 77,2 | 12 | 33,7 | 0,022 |
| Температура (*Temp*) | -0,4 | 2 | 0,1 | 0,012 |
| Температура (*Temp*) | -0,3 | 12 | 0,1 | 0,023 |
| Перепад (*Diff*) | 0,1 | 1 | 0,06 | 0,025 |
| Перепад (*Diff*) | 0,1 | 12 | 0,05 | 0,017 |
| Влажность (*Hum*) | 0,1 | 11 | 0,04 | 0,024 |
| Волны (*W*) | 0,3 | 9 | 0,1 | 0,035 |

* – Зависимость статистически не значима (p > 0,05).




Фактическая и расчетная смертность от болезней органов дыхания у женщин.

Табл. 2. Расчетные коэффициенты модели прогноза смертности населения.

| **Влияющие факторы** | **Пол** | **Инфаркт** | **Инсульт** | **Органы дыхания** | **Внешние причины** |
|---|---|---|---|---|---|
| Корреляция | Ж | 0,69 | 0,65 | 0,79 | 0,72 |
| | М | 0,6 | 0,62 | 0,6 | 0,68 |
| Коэффициент модели, $\beta_i$ / Лаг, L | | | | | |
| Диоксид азота | Ж | 16,5 L=5 | 22,9 L=15 | | 63,2 L=14 |
| | М | 10,9 L=5 | 16,7 L=13 | 26,3 | 20,3 L=5 |
| Формальдегид | Ж | | | 77,2 L=12 | 72 L=9 |
| | М | | 26,3 L=4 | | |
| Взвешенные вещества | Ж | 1,2 L=5 | 1,9 L=8 | | 5 L=8 |
| | М | 1,1 L=15 | 1,7 L=5 | | |
| Температура | Ж | | | | |
| | М | | | | |
| Влажность | Ж | 0,02 L=6 | | 0,1 L=11 | 0,07 L=5 |
| | М | | | | 0,03 L=7 |
| Резкие перепады температуры | Ж | | | 0,1 L=1 | 0,1 L=1 |
| | М | | | 0,04 L=5 | 0,04 L=14 |
| Температурные волны | Ж | 0,06 L=2 | | 0,3 L=9 | 0,3 L=7 |
| | М | | | | |

Для полученных параметров Пуассоновская регрессионная модель примет вид

$$\ln(\mu_t) = -7{,}6 + 77{,}2 \cdot F_{t-12} - 0{,}4 \cdot Temp_{t-2} - 0{,}3 \cdot Temp_{t-12} + 0{,}1 \cdot Diff_{t-1} + 0{,}1 \cdot Diff_{t-12} - 0{,}1 \cdot Hum_{t-11} + 0{,}3 \cdot W_{t-9} \quad (3)$$

где $F$ – формальдегид; $Temp$ – температура; $Diff$ – перепад температуры в течение суток; $Hum$ – влажность; $W$ – температурные волны.

Рассчитанный показатель хорошо описывает фактическую смертность от болезней органов дыхания у женщин с коэффициентом корреляции 0,79.

Сводный анализ коэффициентов всех реализаций многомерных моделей, полученных согласно формулам (1)-(2), представлен в табл. 2. Пустые ячейки в таблице означают, что факторы воздействия, указанные в строках, не оказывают значимого влияния на смертность от соответствующих причин, указанных в столбцах.

При анализе полученных моделей, были выявлены следующие результаты (см.табл. 2).

- Наибольший вклад в смертность населения вносит загрязнение атмосферного воздуха, что нашло отражение во всех полученных моделях независимо от причин смертности;
- При анализе моделей с разделением по полу, выявлено, что воздействию факторов окружающей среды больше подвержены женщины.
- Наибольшему влиянию факторам окружающей среды подвержено население, умирающее от внешних причин. В наименьшей степени от инсультов.
- Что касается степени тесноты связи между фактической и рассчитанной смертностью, то наибольшие коэффициенты корреляции наблюдаются от внешних причин и от болезней органов дыхания.
- Анализ распределения лагов показывает, что большая часть негативных эффектов у мужчин проявляется до шести дней после воздействия неблагоприятных факторов, в то время как у женщин негативные эффекты для здоровья проявляются до десяти дней.

Дальнейшее развитие этого подхода будет связано с внесением в модель корректировок на сезонные, недельные или суточные тренды и уточнение способов введения в модель таких климатических факторов как волны тепла и температурные контрасты.

*Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований, Правительства Красноярского края, Красноярского краевого фонда науки в рамках научного проекта «Методы и информационные технологии оценки рисков развития социально-природно-техногенных систем промышленного региона».*

ЛИТЕРАТУРА

[1] *Lipfert F.W.* A critical review of studies of the association between demands for hospital services and air pollution // Environmental Health Perspectives Supplements. 1993. Vol. 101, Suppl. 2. PP. 229-268.

[2] *Braga A.L.F., Conceição G.M.S., Pereira L.A.A et al*. Air pollution and pediatric respiratory hospital admissions in Sao Paulo, Brazil // Journal of Occupational and Environmental Medicine. 1999. Vol. 1. PP. 95-102.

[3] *Braga A.L.F., Saldiva P.H.N., Pereira L.A.A. et al*. Health effects of air pollution exposure on children and adolescents in Sao Paulo, Brazil // Pediatric Pulmonology. 2001. Vol. 31. PP. 106-113.

[4] *Burnett R.T., Cakmak S., Brook J.R.* The effect of the urban ambient air pollution mix on daily mortality rates in 11 Canadian cities // Canadian Journal of Public Health. 1998. Vol. 89. PP. 152-156.

[5] *Ibald-Mulli A., Timonen K.L., Peters A. et al*. Effects of particulate air pollution on blood pressure and heart rate in subjects with cardiovascular disease: A multicenter approach // Environmental Health Perspectives. 2004. Vol. 112, No. 3. PP. 369-377.

[6] *Peng R.D., Dominici F., Louis T.A.* Model choice in time series studies of air pollution and mortality // Journal of the Royal Statistical Society Series A. 2006. Vol. 169. PP. 179-203.

[7] *Pope III C.A., Burnett R.T., Thun M.J. et al*. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. Journal of the American Medical Association. 2002. Vol. 287, No. i9. PP. 1132-1142.

[8] Безопасность России. Региональные проблемы безопасности. Красноярский край. М.: МГФ «Знание», 2001. 576 с.
[9] *Винокуров Ю.И., Лепихин А.М., Москвичев В.В., Чернякова Н.А.* Геоинформационные технологии и математические модели для мониторинга и управления экологическими и социально-экономическими системами. Барнаул: ИВЭП СО РАН. Пять плюс, 2011. 250 с.
[10] *Левкевич В.Е., Лепихин А.М., Москвичев В.В., Никитенко П.Г. и др.* Безопасность и риски устойчивого развития территорий. Красноярск: Сибирский федеральный ун-т, 2014. 224 с.
[11] *Акимов В.А., Быков А.А., Шевченко А.В., Стрелко С.В.* Развитие научно-методических основ государственного управления с использованием показателей стратегических рисков, в том числе и рисков чрезвычайных ситуаций (на примере Ульяновской области) // Стратегия гражданской защиты: проблемы и исследования. 2013. Т. 3, № 2. С. 736-780.
[12] МР 2.1.10.0057-12.2.1.10. Состояние здоровья населения в связи с состоянием окружающей среды и условиями проживания населения. Оценка риска и ущерба от климатических изменений, влияющих на повышение уровня заболеваемости и смертности в группах населения повышенного риска (утв. 17.01.2012). 36 с.
[13] Всероссийский научно-исследовательский институт гидрометеорологической информации – мировой центр данных. **http://meteo.ru** (Дата обращения 11.11.2016).

# ИСПОЛЬЗОВАНИЕ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ ДЛЯ СЪЕМКИ МИКРОРЕЛЬЕФА ПРИ МОНИТОРИНГЕ АГРОЦЕНОЗОВ

*Шевырногов А.П., Мальчиков Н.О., Ботвич И.Ю., Емельянов Д.В., Ларько А.А.*
Институт биофизики СО РАН, Красноярск

Показано применение БПЛА для получения данных о микрорельефе поля, относящегося к сельскохозяйственным угодьям. В работе используются данные о значениях интеграла NDVI в течение вегетационного периода, полученные с помощью спутниковой и наземной спектрометрии. Установлено, что изменение микрорельефа оказывает существенное воздействие на величину интеграла NDVI и урожая. Предложенный подход можно использовать в точном земледелии для планирования будущих сельскохозяйственных работ.
*Ключевые слова: БПЛА, сельское хозяйство, вегетационный индекс, ячмень, точное земледелие, спектрометрия.*

**Введение.** Картирование урожайности – это элемент системы точного земледелия, который позволяет определить неоднородность одного из важнейших показателей – урожайности сельскохозяйственных культур [1]. Карта урожайности сельскохозяйственных культур является основой дифференцированного внесения удобрений при точном земледелии [2]. Она позволяет идентифицировать зоны с невысокой урожайностью для целенаправленного изучения причин её снижения на данном участке поля и принятия соответствующих мер для решения данной проблемы [3].

Построение цифровых карт сельскохозяйственных полей возможно с использованием спутниковой информации высокого пространственного и временного разрешения. Сегодня непрерывный мониторинг посевов сельскохозяйственных культур с высоким пространственным разрешением (3 метра) и ежедневными измерениями обеспечивает самая большая в мире спутниковая группировка PlanetScope компании Planet [4]. Система предоставляет данные, позволяющие оценить состояние сельскохозяйственных культур в течение всего периода вегетации.

В настоящее время в мониторинге территорий, занятых посевами сельскохозяйственных культур широкое применение находят беспилотные летательные аппараты (БПЛА) [5]. Использование беспилотных летательных аппаратов в земледелии и, в целом, в сельском хозяйстве – одно из наиболее перспективных направлений применения этой технологии. БПЛА могут эффективно использоваться для планирования и контроля этапов сельскохозяйственного производства, а также для химической обработки посевов и других растений [6]. Помимо этого, сверхвысокое разрешение изображений, полученных с помощью бесплотных летательных аппаратов, позволяет исследовать микрорельеф и почвенную структуру полей [7-8]. Это, в свою очередь, позволяет учитывать большое количество новых, ранее недоступных, параметров, в том числе, определение участков с повышенной и пониженной влажностью, которая напрямую влияет на состояние посевов [9].

**Цель исследования:** Изучение возможности применения БПЛА для оценки микрорельефа территорий, относящихся к сельскохозяйственным угодьям, и сопоставление результатов с данными пространственного распределения урожайности на исследуемой территории. Полученные результаты должны иметь значение для точного земледелия.

**Объект и методы исследования.** Исследования проводились на землях ООО учхоз «Миндерлинское» Сухобузимского района Красноярского края до начала и в течение вегетационного периода. Первые съемки с помощью БПЛА проведены в период схода снежного покрова. Опытное поле состоит из 5 полос, средняя протяженность которых составляет около 650 метров. Ширина полос равна 20 метрам, расположение с запада на восток. В северной части каждой полосы вносились удобрения, в южных частях удобрения не вносились. Каждая, из 5 полос, засеивалась сельскохозяйственными культурами. В работе использовался анализ и результаты обработки данных по полосе № 5 – засеянной ячменем *«Ача»*.

Посев ячменя произведен 19 мая 2018 года. Все полосы разделены на 4 тестовых участка в соответствии с четырьмя видами обработки почв: *a* – вспашка (ПН-5-35 на 20-22 см); *b* –

плоскорезная обработка (КПШК-3,8 на 20-22 см); *c* – поверхностная обработка (дискатор БДШ-5,6 на 8-10 см); *d* – прямой посев при нулевой обработке почвы (Агратор 4,8). Для чистоты проведения эксперимента между тестовыми участками проложены защитные полосы шириной 5 метров.

Для выполнения высотной съемки использовался БПЛА серии DJI Phantom4 PRO+ v2.0. Получаемые данные – фотоизображения с сантиметровым разрешением и планово-высотной привязкой координат.

Также, в работе использовался набор следующих данных:

- спутниковые данные PlanetScope, с предварительной атмосферной коррекцией. Корректирующие коэффициенты для каждого канала присутствуют в метаданных каждой сцены (данные снимки обладают пространственным разрешение в 3 метра);
- данные полевого спектрометрирования, которое выполнялось 5 раз в течение вегетационного периода 2018 г. Расстояние между измерениями составляло около 50 метров.

При проведении полевой спектрометрии применялся спектрорадиометр Spectral Evolution PSR-1100F. Получаемые данные – коэффициенты спектральной яркости (КСЯ) объекта в диапазоне от 320 до 1100 нм.

По спутниковым и наземным спектрофотометрическим данным проводился расчет спектральных индексов NDVI (Normalized Difference Vegetation Index) [10], VARI (Visible Atmospherically Resistant Index) [11], MSAVI2 (Modified Soil Adjusted Vegetation Index) [12], ClGreen (Green chlorophyll index) [13].

По данным, полученным с помощью БПЛА, выполнялось построение ортофотоплана местности, 3D-модели местности и построение профиля высот, для исследуемой полосы № 5.

**Результаты и обсуждение.** При исследовании опытного поля с помощью БПЛА перед началом полетов составлялось полетное задание. Для получения адекватного ортофотоплана соблюдалось перекрытия снимков: 60% в продольном и 30% в поперечном направлениях [14]. Съемка выполнялась в безветренную погоду, с высоты 50 метров, объектив камеры направлен в надир. В программном обеспечении, поставляемом в комплекте с БПЛА (DJI GO) установлены параметры съемки с интервальная в 5 с, и скоростью полета 15 м/с. Затем, уже без участия оператора, выполнялись полет и съемка, занимавшие, в среднем, около 30 минут.

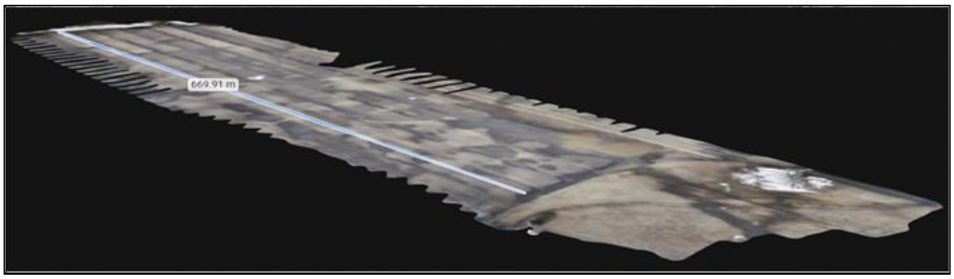
В результате съемки получен набор перекрывающихся фотоизображений, в количестве 294 шт. Полученный набор данных обрабатывался в ПО Pix4D mapper. Такая обработка позволяет получать из набора отдельных изображений полноценную ортофотомозаику и 3D-модель местности, которая несет в себе информацию о планово-высотном положении каждого пикселя.

На рис. 1 представлены результаты обработки полученных данных. На изображении, исследуемой полосы № 5, выделяется зона с остатками снежного покрова (на рис. 1,*а*). На других частях полосы остатки снежного покрова отсутствуют, заметны только зоны повышенной влажности, выделяющиеся более темными тонами. На построенном профиле высот по полосе № 5, видно, что в зоне с остатками снежного покрова наблюдается понижение рельефа на 1 метр (рис. 1,*в*). В дальнейшем, сопоставление данных спектрометрии и результатов съемки с помощью БПЛА показало, что зона понижения рельефа выделяется на картограмме повышенными значениями NDVI.

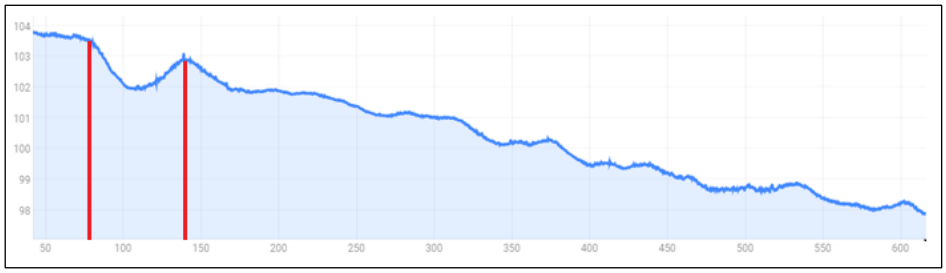
Фактическая урожайность ячменя варьируется в зависимости от наличия удобрительного фона и вида обработки. Наибольшая урожайность ячменя зафиксирована на варианте с отвальной вспашкой (25,9 ц/га (удобренный фон), и 23,7 ц/га (не удобренный фон), наименьшая – на варианте с поверхностной обработкой (19,7 ц/га (удобренный фон) и 15,8 ц/га (не удобренный фон). По всем видам обработок внесение минеральных удобрений привело к повышению урожайности зерна ячменя.



*а*




*б*




*в*

Рис. 1. Обработанные результаты съемки: *а* – ортофотомозаика; *б* – 3D-модель; *в* – профиль высот по полосе № 5. Красным кругом и линиями выделен участок с понижением высот на 1 м.

Оценка фактической урожайности на поле является важным показателем эффективности сельскохозяйственного производства. Определение пространственного распределения (картограмм) урожайности является более информативным признаком, позволяющим выявлять неоднородность уровня урожайности в пределах одного поля. Количественной характеристикой состояния посевов служат вегетационные индексы, в том числе NDVI, MSAVI2, ClGreen. В ряде работ подтверждена устойчивая корреляционная связь между NDVI и урожайностью [1, 2, 15-16].

Использование спутниковой информации PlanetScope обеспечивает определение участков с устойчиво низкими или высокими величинами фитомассы в отдельные фазы роста и развития растений. Построение карт пространственного распределения индексных величин высокого пространственного разрешения в течение периода вегетации является неотъемлемой частью точного земледелия.

В качестве базовой характеристики, которая используется в работе, является динамика нарастания растительного агробиоценоза в течение вегетационного периода. Основным параметром, который используется в дальнейшей работе является NDVI. В течение вегетационного периода величина NDVI претерпевает значительные изменения от 0,1 до 0,8. Типичная кривая NDVI в течение вегетационного периода – это рост в начале периода, максимальные значения – в середине, и снижение – в конце периода вегетации. Такая форма представления этой величины позволяет рассчитать интеграл под кривой вегетационного индекса.

В практических целях возможно использование не полного интеграла от начала до конца вегетационного периода, а его начальной части. Чем раньше рассчитывается прогноз урожайности, тем он более значим для практических целей. Отсюда следует, что необходимо найти оптимальную дату конечного времени для расчета интеграла вегетационного периода, которая к тому же обеспечивает достаточную точность прогноза.

| дата | | d | c | b | a | | |
|---|---|---|---|---|---|---|---|
| начало | конец | | | | | мин | макс |
| 04.06.2018 | 13.06.2018 | | | | | 1,55 | 2,15 |
| 04.06.2018 | 18.06.2018 | | | | | 3,47 | 4,85 |
| 04.06.2018 | 23.06.2018 | | | | | 4,9 | 7 |
| 04.06.2018 | 26.06.2018 | | | | | 6,95 | 9,55 |
| 04.06.2018 | 02.07.2018 | | | | | 8,35 | 11,1 |
| 04.06.2018 | 16.07.2018 | | | | | 10,9 | 14,4 |
| 04.06.2018 | 31.07.2018 | | | | | 16,85 | 21,7 |
| 04.06.2018 | 08.08.2018 | | | | | 22,55 | 28,45 |
| 04.06.2018 | 17.08.2018 | | | | | 25,1 | 31,45 |
| 04.06.2018 | 21.08.2018 | | | | | 27,4 | 33,7 |
| 04.06.2018 | 26.08.2018 | | | | | 27,9 | 34,5 |
| 04.06.2018 | 05.09.2018 | | | | | 28,4 | 35,8 |

Рис. 2. Картограммы значений интеграла NDVI посевов ячменя в течение периода вегетации (с 4 июня по 5 сентября 2018 г.), полученные по данным PlanetScope. Белыми линиями обозначено разделение на виды обработки (a, b, c, d). Раскраска картограмм выполнена с использованием одной цветовой шкалы для всех дат исследований. Минимальные и максимальные значения цветовых шкал представлены в таблице – напротив соответствующей картограммы.

Анализ связи между урожайностью и величиной интеграла в разные периоды вегетации показал, что с середины июля величина r достигает достаточно высоких величин (более 0,7) для индексов NDVI, MSAVI2, ClGreen (рис. или ссылка). Максимальные величины достигнуты к началу августа. Коэффициенты корреляции, рассчитанные с 2 июля по 5 сентября, являются достоверными при уровне значимости 0,05. Таким образом, показано, что использование интегральных показателей позволяет выполнять прогноз урожайности ячменя в фазе цветения – молочной спелости.

На рис. 2 представлены картограммы значений интеграла NDVI посевов ячменя в течение периода вегетации с 4 июня по 5 сентября 2018 г., полученные по данным PlanetScope.

Представление информации в форме накопленных значений интегралов NDVI показывает, что по сравнению с пространственным изменением прямых значений NDVI, появляется устойчивая во времени структура при определенных видах обработки (*d* и *a*). В то же время, проявляется четкая зависимость изменение пространственной структуры интегралов NDVI при видах обработки *c* и *b*. При обработке вида **c** ярко выраженная величина интеграла с 4 по 13 июня уменьшается к 8 августа, исчезает, и весь участок становится почти равномерным. То же самое происходит при обработке типа *b*, но разница между началом и концом меньше, чем при обработке вида *c*.

Отмечается, что на картограмме, в участке с типом обработки *d*, выделяется красная зона (с повышенным значением интеграла NDVI). При экспертном сравнении построенной картограммы и результатов съемки с БПЛА установлено, что зона с повышенным значением NDVI соответствует понижению рельефа на 1 метр, которое было детектировано на ортофотомозаике.

Можно предположить, что понижение высотного профиля приводит к повышению влагозапаса в почве. Следствием этого являются повышенные значения интеграла NDVI и соответственно повышение урожайности на этом участке.

Также следует отметить подобную связь на участке А. На рис. 1,*a* видны зоны повышенной влажности – темные пятна, которые при визуальной оценке наблюдались как насыщенные талой водой участки почвы. На картограмме интегралов NDVI этот участок также имеет повышенные значения NDVI и, следовательно, урожайности. В этом случае относительное понижение микрорельефа незначительно, но, тем не менее, приводит к накоплению влаги.

Из приведенной информации можно сделать вывод, что съемка с помощью БПЛА в поздний период снеготаяния может дать весьма существенные данные для оценки последующей динамики вегетационного процесса на посевных площадях.

Таким образом, сопоставление картограммы значений интеграла NDVI и результатов съемки с БПЛА показывает, что использование подобного типа данных позволяет учитывать все особенности рельефа поля и детектировать выделяющиеся участки, а следовательно – требующие особого подхода в проведении соответствующих мероприятий.

В дальнейшем, полученная картограмма, и данные о микрорельефе, полученные с помощью БПЛА, планируется использовать при дифференцированном внесении удобрений и мониторинге роста культур. Дифференцированное внесение удобрений и посадка сельскохозяйственных культур с учетом микрорельефа поля позволят повысить эффективность применяемых удобрений и выровнять урожайность культуры в пределах поля. Такой подход позволит улучшить экономические показатели сельскохозяйственного производства.

**Заключение.** Использование БПЛА при съемке микрорельефа агроценозов позволяет детектировать выделяющиеся участки, связанные с различным уровнем влагозапаса в почве.

Установлено, что интеграл NDVI, вычисляемый в течение вегетационного периода можно рассматривать как параметр, связанный с урожайностью. Анализ связи между урожайностью и величиной интеграла в разные периоды вегетации показал, что с середины июля величина коэффициента корреляции достигает достаточно высоких величин (более 0,7, при уровне значимости 0,05) для индексов NDVI, MSAVI2, ClGreen. Максимальные величины достигнуты к первым числам августа.

В результате проведенных исследований выполнено построение ортофотомозаики и 3D-модели местности по данным полученным с БПЛА. Выполнено построение пространственного изменения значений NDVI поля ячменя с 2 июня по 20 сентября 2018 г. по спутниковым и наземным данным. Получено пространственное изменение интеграла NDVI поля ячменя в течение периода вегетации (с 4 июня по 5 сентября 2018 г.) по спутниковым данным.

Разработан метод оценки пространственного распределения урожайности ярового ячменя, реализованный на основе использования оптических наземных и спутниковых спектральных данных (PlanetScope с пространственным разрешением 3 метра).

Установлена связь изменения значений интеграла NDVI от изменения микрорельефа поля и пространственного распределения зон с повышенной влажностью в начале вегетационного периода.

Полученные карты урожайности и данные о микрорельефе поля будут использоваться при планировании будущих сельскохозяйственных работ.

*Исследование выполнено при поддержке Красноярского краевого фонда науки (проект «Разработка и апробация методов контроля земель сельскохозяйственного назначения для создания системы точного земледелия»), и в рамках Комплексной программы фундаментальных наук СО РАН «Междисциплинарные интеграционные исследования» на 2018-2020 гг. (проект № 74.2).*

## ЛИТЕРАТУРА

[1] *Акинчин А.В., Левшаков Л.В., Линков С.А., Ким В.В., Горбунов В.В.* Информационные технологии в системе точного земледелия // Вестник Курской государственной сельскохозяйственной академии. 2017. № 9. С. 16–21.

[2] *Антонов С.А.* Роль геоинформационных технологий и дынных дистанционного зондирования Земли для оценки состояния и продуктивности агроландшафтов // Известия Оренбургского государственного аграрного университета. 2018. Т. 73. № 5. С. 10–14.

[3] *Бобкова Ю.А., Лобков В.Т.* Использование технологий точного земледелия при создании агрохимических картограмм // Вестник ОрелГАУ. 2017. Т. 67. № 4. С. 25–31.

[4] *Planet Team.* Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. 2018. **https://api.planet.com**.

[5] *Wang X., Sun H., Long Y., Zheng L., Liu H., Li M.* Development of Visualization System for Agricultural UAV Crop Growth Information Collection // IFAC-PapersOnLine. 2018. Vol. 51. P. 631-636.
[6] *Марченко Л.А., Личман Г.И., Смирнов И.Г., Мочкова Т.В., Колесникова В.А.* Дифференцированное внесение удобрений и пестицидов с использованием беспилотных летательных аппаратов // Сельскохозяйственные машины и технологии. 2017. № 3. С. 17-23.
[7] *Korzeniowska K., Pfeifer N., Landtwing S.* Mapping gullies, dunes, lava fields, and landslides via surface roughness // Geomorphology. 2018. Vol. 301. P. 53-67.
[8] *Guimarães R.M.L., Lamandé M., Munkholm L.J., Ball B.C., Keller T.* Opportunities and future directions for visual soil evaluation methods in soil structure research // Soil and Tillage Research. 2017. Vol. 173. P. 104-113.
[9] *Pineux N., Lisein J., Swerts G., Bielders C. L., Lejeune P., Colinet G., Degré A.* Can DEM time series produced by UAV be used to quantify diffuse erosion in an agricultural watershed? // Geomorphology. 2017. Vol. 280. P. 122-136.
[10] *Jordan C.F.* Derivation of leaf-area index from quality of light on the forest floor // Ecology. 1969. Vol. 50. P. 663–666.
[11] *Gitelson A.A., Stark R., Grits U., Rundquist D., Kaufman Y., Derry D.* Vegetation and soil lines in visible spectral space: a concept and technique for remote estimation of vegetation fraction // International Journal of Remote Sensing. 2002. Vol. 23. P. 2537–2562.
[12] *Qi J., Chehbouni A., Huete A.R., Y.H.Kerr Y.H., Sorooshian S.* A modified soil adjusted vegetation index // Remote Sensing of Environment. 1994. Vol. 48. P. 119–126.
[13] *Dall'Olmo G., Gitelson A.A., Rundquista D.C., Leavitt B., Barrow T., Holz J.C.* Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands // Remote Sensing of Environment. Vol. 96. P. 176–187.
[14] *Tu Y., Bian M., Wan Y., Fei T.* Tea cultivar classification and biochemical parameter estimation from hyperspectral imagery obtained by UAV // PeerJ. 2018. P. 1-15.
[15] *Бондур В.Г., Гороховский К.Ю., Игнатьев В.Ю., Мурынин А.Б., Гапонова Е.В.* Метод прогнозирования урожайности по космическим наблюдениям за динамикой развития вегетации // Известия высших учебных заведений. Геодезия и аэрофотосъемка. 2013. № 6. С. 61–68.
[16] *Толпин В.А., Барталев С.А., Бурцев М.А., Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Матвеев А.М., Прошин А.А., Флитман Е.В.* Оценка состояния сельскохозяйственных культур на основе межгодовой динамики с использованием данных MODIS // Современные проблемы дистанционного зондирования Земли из космоса. 2007. Т. 4. № 2. С. 380–389.

# ПРОСТРАНСТВЕННЫЙ АНАЛИЗ ЗАГРЯЗНЕНИЯ АТМОСФЕРЫ КРАСНОЯРСКА

*Якубайлик О.Э.[(1)(2)(3)], Заворуев В.В.[(2)(3)], Малимонов М.И.[(1)], Пушкарев А.А.[(1)]*
[(1)] Федеральный исследовательский центр Красноярский научный центр СО РАН, Красноярск
[(2)] Институт вычислительного моделирования СО РАН, Красноярск
[(3)] Сибирский федеральный университет, Красноярск

Рассматривается загрязнение Красноярска взвешенными частицами PM2.5 по данным сети научно-исследовательского экологического мониторинга ФИЦ КНЦ СО РАН. Анализируется качество данных, измеряемых станциями мониторинга CityAir, проводится их сравнение с данными сертифицированных приборов. Особое внимание уделяется периоду неблагоприятных метеоусловий в первой половине февраля 2019 г. Делается вывод о значительном влиянии печного отопления на загрязнение атмосферы Красноярска.
*Ключевые слова: экологический мониторинг, взвешенные частицы, PM2.5, загрязнение воздуха, аэрозоль, геопортал ИВМ СО РАН, экологические веб-сервисы, печное отопление.*

**Введение.** Задачи оперативного мониторинга и оценки загрязнения атмосферного воздуха на территориях с высокой плотностью населения, прежде всего – в крупных городских агломерациях, в настоящее время являются предметом многочисленных научных и прикладных исследований во многих странах мира. Рассматривается широкий спектр задач – построение математических и информационных моделей различного типа [1-3], оценка загрязнения [4, 5], создание и эксплуатация сетей мониторинга [6-8], вычислительное моделирование протекающих процессов [9]. Значительное внимание данным проблемам уделяется государственными структурами и профильными ведомственными организациями, которые занимаются организацией и проведением соответствующих мероприятий по данной проблематике. Традиционно активно действуют независимые общественные экологические организации. Благодаря современным информационным технологиям, социальным сетям в Интернете все более значимую роль в обсуждении проблем экологического мониторинга играют инициативные граждане – блоггеры, ведущие популярные интернет-дневники в ЖЖ, Инстаграме, ВКонтакте, Facebook, и проч. [10].

Нормативно-правовое обеспечение решаемых задач основано на экологическом законодательстве РФ, прежде всего – Федеральном законе от 10.01.2002 № 7-ФЗ (ред. от 27.12.2018) «Об охране окружающей среды», Федеральном законе от 04.05.1999 № 96-ФЗ (ред. от 29.07.2018) «Об охране атмосферного воздуха». Нормативно-правовое обеспечение также формируется на региональном и местном уровне.

Мониторинг состояния загрязнения атмосферного воздуха в г. Красноярске ведется рядом организаций федерального и региональных уровней – ФГБУ «Среднесибирское управление по гидрометеорологии и мониторингу окружающей среды» (УГМС), КГБУ «Центр реализации мероприятий по природопользованию и охране окружающей среды Красноярского края» (КГБУ «ЦРМПиООС»), Главное управление МЧС по Красноярскому краю, Сибирский федеральный университет, Федеральный исследовательский центр «Красноярский научный центр СО РАН» (ФИЦ КНЦ СО РАН), и другие. Каждая из этих организаций обладает собственными методиками, технологиями и системами сбора, хранения и обработки данных. Разнообразие используемых решений, межведомственная и организационная разобщенность приводят к тому, что комплексный анализ и оперативная оценка всего массива регистрируемой информации представляются в настоящее время технически затруднительными и практически не проводятся. С учетом запланированного рядом уполномоченных организаций расширения числа стационарных постов наблюдений за состоянием природной среды в ближайшие годы, ситуация будет только усугубляться [11].

Исследования и разработки для мониторинга загрязнения атмосферы, выполняемые в ФИЦ КНЦ СО РАН, направлены на решение указанной проблемы. Создаваемое программно-технологическое обеспечение обеспечивает прием и интеграцию информации, поступающей

из разных источников данных. Формируемая база данных обеспечена средствами оперативной агрегации и визуализации данных, их аналитической обработки [10, 11].

Представленные ниже результаты посвящены оценке динамики загрязнения г. Красноярска взвешенными частицами по данным сети научно-исследовательского экологического мониторинга ФИЦ КНЦ СО РАН, построенной с использованием станций мониторинга CityAir. Качество получаемых данных верифицировалось по данным КГБУ «ЦРМПиООС».

**Объект исследования.** Предметом исследования в настоящей работе стала динамика загрязнения г. Красноярска взвешенными частицами.

Взвешенные частицы, также известные как PM2.5 и PM10 (PM – particulate matter), представляют собой сложную смесь чрезвычайно мелких твердых частиц. Состав их разнообразен: кислоты (например, нитраты и сульфаты), выбросы производств, органические химические вещества, металлы и частицы почвы, пыли, сажа выхлопных газов, частицы асфальта и стертых шин и т.д. Источниками взвешенных частиц являются автомобили и дороги, электростанции, заводы и производства, дым от пожаров и костров. Взвешенные частицы так же появляются в результате испарения воды и сжигания любых видов топлива, например – вследствие печного отопления в частном секторе г. Красноярска в зимний период. Природными источниками взвешенных частиц являются растения. Именно благодаря своему размеру вызывают проблемы со здоровьем, потому что частицы, попадают через горло и нос попадают в легкие. После вдыхания, эти частицы влияют на сердечно сосудистую и легочную системы и могут вызвать серьезные последствия для здоровья. Некоторые частицы, такие как пыль, грязь, копоть, дым достаточно большие и если таких частиц достаточно много, то их увидеть невооруженным глазом. Другие частицы настолько малы, что они могут быть обнаружены только с помощью электронного микроскопа. Взвешенные частицы принято разделять на несколько категорий: с диаметром меньше, чем 10 мкм (PM10) и частицы с диаметром в 2,5 микрометра и меньше (PM2.5).

В настоящей работе мы рассматривали период времени с декабря 2018 г. по май 2019 г., для которого были проанализированы имеющиеся данные по уровню концентраций взвешенных частиц PM2.5 в атмосфере Красноярска, полученные сетью научно-исследовательского экологического мониторинга ФИЦ КНЦ СО РАН. Особое внимание было уделено первой половине февраля, т.к. в этот период Среднесибирским УГМС был объявлен режим неблагоприятных метеоусловий (НМУ) – единственный раз за 2019 г.

Режим НМУ – краткосрочное особое сочетание метеорологических факторов, способствующих накоплению вредных (загрязняющих) веществ в приземном слое атмосферного воздуха. В соответствии со статьей 19 Федерального закона «Об охране атмосферного воздуха» № 96-ФЗ юридические лица, имеющие источники выбросов вредных (загрязняющих) веществ в атмосферный воздух, при объявлении режима НМУ обязаны проводить мероприятия по уменьшению выбросов вредных (загрязняющих) веществ в атмосферный воздух. В зависимости от ожидаемого уровня загрязнения атмосферы составляются предупреждения трех степеней, которым соответствуют три режима работы предприятий в периоды НМУ.

Сеть экологического мониторинга ФИЦ КНЦ СО РАН работает на основе станций мониторинга воздуха «CityAir», создаваемых группой компаний из новосибирского технопарка (CityAir, Унискан, Тион) [12]; в настоящее время в Красноярске установлено 20 указанных станций, преимущественно на левобережье Енисея. Эти станции мониторинга воздуха предназначены для сбора данных о состоянии окружающего воздуха (массовая концентрация аэрозольных частиц, температура, влажность и атмосферное давление) и передачи их на сервер посредством беспроводного канала связи. Практика эксплуатации станций мониторинга подтвердила их возможность использования, работоспособность при низких температурах воздуха вплоть до -40°C [13].

Вся информация со станций сначала передается в облачное хранилище данных компании-разработчика и затем, через веб-сервисы, поступает в базу данных оперативного мониторинга геопортала Института вычислительного моделирования СО РАН, ФИЦ КНЦ СО РАН.

Специализированное программное обеспечение на основе веб-сервисов геопортала используется для визуализации и анализа данных.

**Анализ качества измеряемых данных.** Одной из первых задач стало сравнение данных по взвешенным частицам PM2.5, получаемых станциями мониторинга CityAir, с результатами измерений тех же самых концентраций на постах КГБУ «ЦРМПиООС», оснащенных сертифицированными приборами-пылемерами. В данной работе мы использовали информацию с автоматизированного поста наблюдений (АПН) в мкр. Покровка (АПН «Покровка-Красноярск»), где, по согласованию с КГБУ «ЦРМПиООС», была также установлена станция CityAir. В течение полугода станции, расположенные в одном месте, работали параллельно, одновременно измеряя концентрации частиц. Пример результатов этих измерений в течение одной недели (во время проведения Универсиады-2019 в Красноярске) показан на рис. 1.

Следует отметить, что прямое сравнение получаемых данных затруднительно, т.к. в рассматриваемых устройствах используются разные физические принципы измерений. В приборах, установленных на АПН КГБУ «ЦРМПиООС» измеряются концентрации твердых частиц (принцип действия – радиоизотопный, основан на поглощении β-излучения частицами пыли, осажденными на фильтрующую ленту), а станция CityAir измеряет концентрацию аэрозоли (в приборах CityAir – оптический сенсор. Воздух, проходящий через этот оптический сенсор непрерывно просвечивается лазером, что позволяет высокочувствительному фотоэлементу распознавать количество взвешенных частиц, содержащихся в воздухе).

Сравнительный анализ графиков концентраций, измеренных разными приборами (Рис. 1), показывает, что оба они в целом похожим образом отражают качественную динамику изменения концентраций, но имеют при этом существенные количественные различия. В качестве первого приближения в решении этой проблемы можно рассматривать введение понижающего «калибровочного коэффициента», на который нужно умножить значения концентраций CityAir, чтобы получить величины, сопоставимые с данными сертифицированных устройств КГБУ «ЦРМПиООС». Расчеты по имеющемуся массиву данных на АПН «Покровка-Красноярск» за полгода показали, что среднем значением такого коэффициента является величина

$$k = 0{,}46$$

Попытки найти какую-либо зависимость этого коэффициента от метеоусловий (температура, влажность, точка росы) или каких-либо других параметров на данном этапе к успеху не привели (рис. 2).

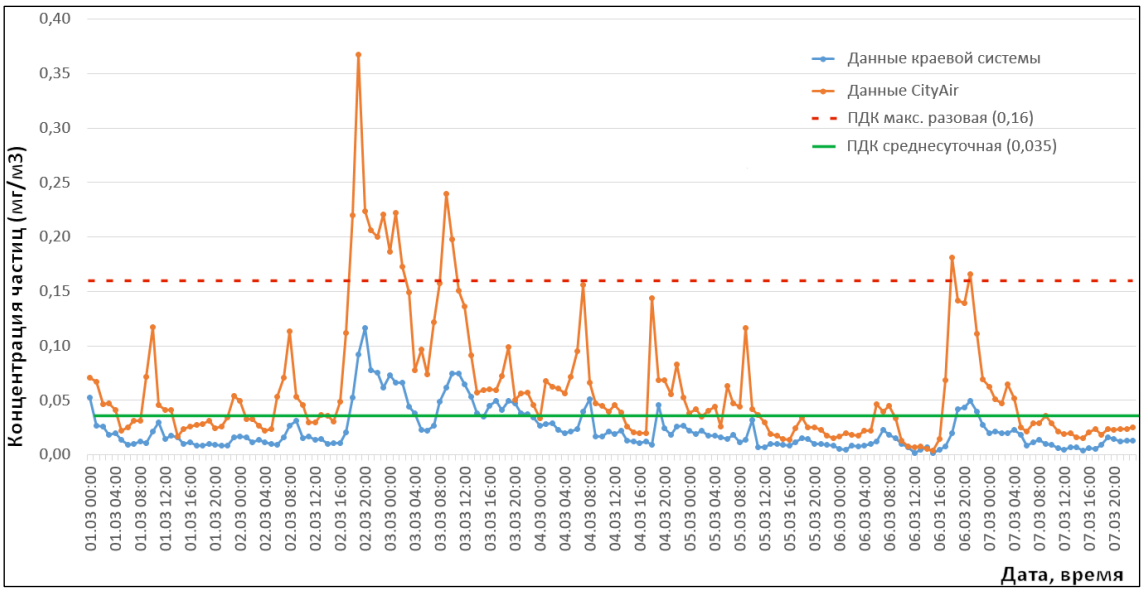



Рис. 1. Сравнение данных по концентрации PM2.5 за первую неделю марта 2019 г. на посту мониторинга «Покровка-Красноярск».

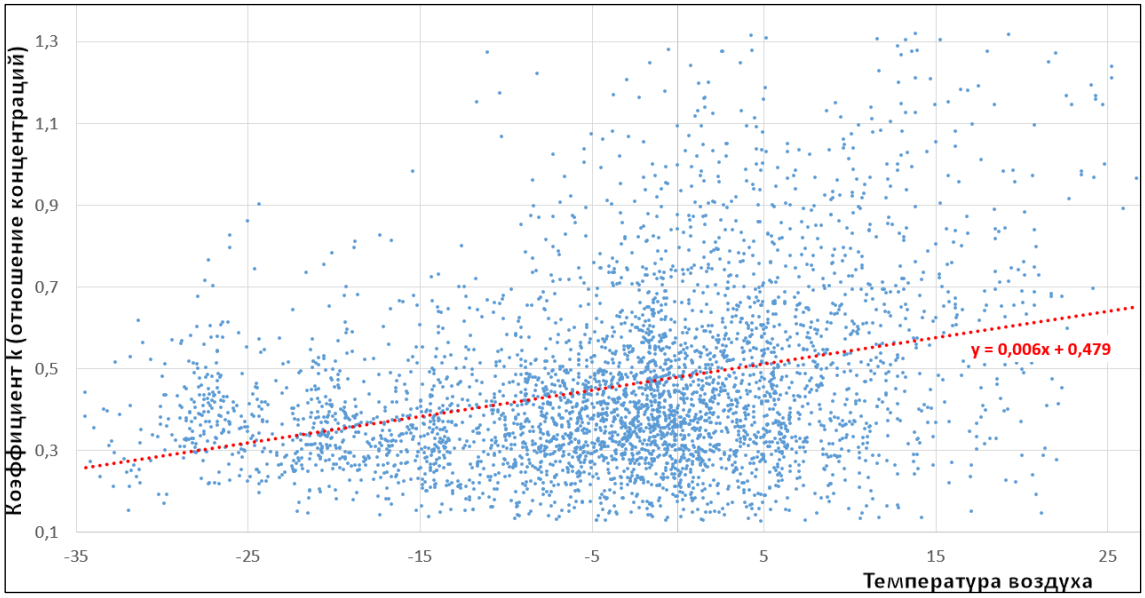



Рис. 2. Зависимость понижающего коэффициента k от температуры воздуха.

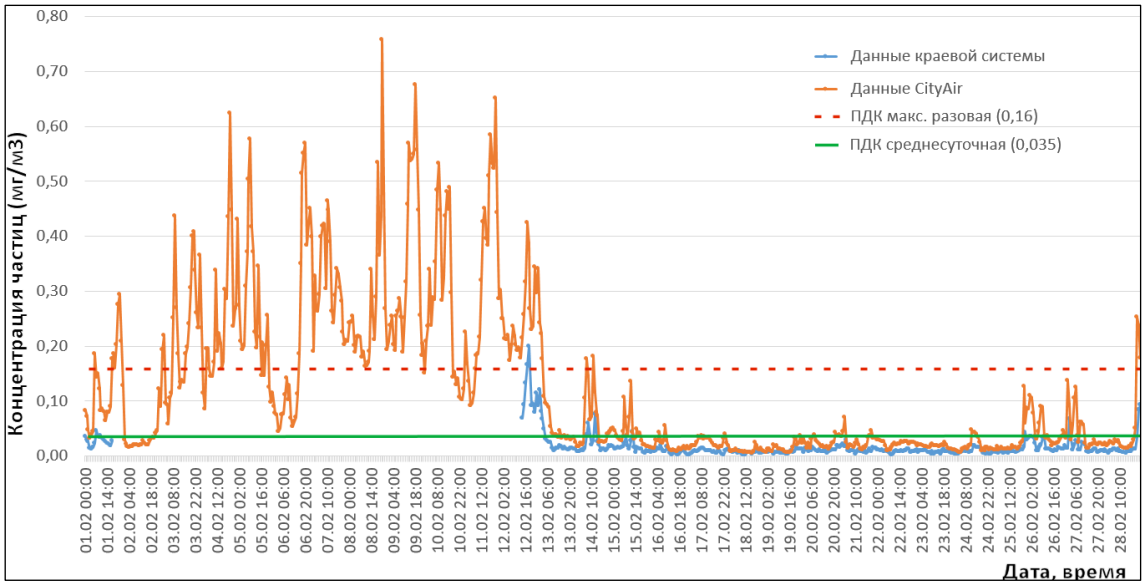



Рис. 3. Результаты измерений концентрации PM2.5 в феврале 2019 г.
на посту мониторинга «Покровка-Красноярск».

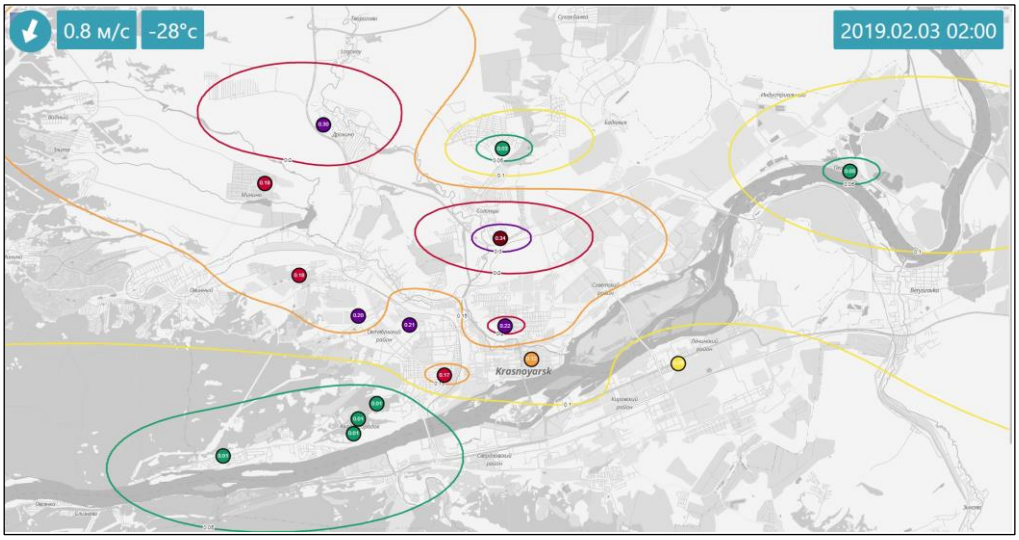



Рис. 4. Распределение концентраций взвешенных частиц при загрязнении Красноярска.

Режим НМУ объявлялся в г. Красноярске в период с 19:00 08 февраля 2019 г. по 19:00 13 февраля 2019 г. Как оказалось, в этот период измерение концентраций взвешенных частиц PM2.5 на АПН «Покровка-Красноярск» приборами КГБУ «ЦРМПиООС» не проводилось – они были отправлены на поверку (рис. 3). Тем не менее, станции мониторинга CityAir в это время работали, что позволило проанализировать динамику распространения загрязнения над городом и его последующей очистки.

**Динамика загрязнения Красноярска взвешенными частицами**. Анализ данных показал, что загрязнение города началось в ночь со 2 на 3 февраля. Еще вечером 2 февраля преобладающее направление ветра было юго-западным, со скоростью 2-3 м/с, и в этот момент город был абсолютно чистым. Ориентировочно в полночь его направление поменялось на юго-восточное, со скоростью порядка 1 м/с. Уже в 2 часа ночи (рис. 4) половина города была относительно грязной. Пространственное распределение загрязнения в этот момент времени показывает максимальные концентрации PM2.5 в частном секторе города (Николаевка, Покровка) и в расположенных в непосредственной близости от города поселках (Солонцы, Дрокино, Минино). Все перечисленные территории можно охарактеризовать как малоэтажную застройку с большим количеством печного отопления, которое, скорее всего, и является основной причиной высоких концентраций PM2.5, учитывая, что температура воздуха в этот момент была около -30°C и все жители активно отапливали (преимущественно используя бурый уголь) свои дома в это время.

По динамике распространения загрязнения по времени видно, что смог из этих районов частного сектора плавно «накрывает» значительную часть города, в соответствии с ветром, который балансирует между юго-восточным и южным направлениями, частично наблюдаются штилевые условия. К позднему вечеру 4 февраля город уже становится практически полностью грязным, с высокими значениями концентраций PM2.5. На протяжении ряда последующих дней существенных изменений в направлении/скорости ветра не наблюдается.

Обратный процесс («очистка города») начинается в полночь с 12 на 13 февраля. Ветер меняет свое направление на юго-западное, со скоростью 1-4 м/с. Уже к полудню 13 февраля город становится абсолютно чистым. Обращает на себя внимание неравномерность очистки города, видимо связанная с рельефом – в низинах, в том числе в центральной части города, загрязнение задерживается дольше.

Подробности пространственного распределения загрязнения и динамика его изменения, исследованная в данной работе, доступны в видеороликах [14, 15].

**Заключение.** Сеть научно-исследовательского мониторинга загрязнения атмосферы, основанная на большом количестве распределенных по пространству станций, предоставляет возможность проведения детального пространственно-временного анализа данных. Предварительные результаты обработки полученной информации показывают существенное влияние печного отопления на загрязнение Красноярска. Необходимы дальнейшие исследования в данном направлении.

## ЛИТЕРАТУРА


[1] *Sunil Gulia, S.M. Shiva Nagendra, Mukesh Khare, Isha Khanna* Urban air quality management – a review // Atmospheric Pollution Research. 2015. V. 6, Is. 3. pp. 286-304.

[2] *Xuezhen Qiu et al* Development of an integrated policy making tool for assessing air quality and human health benefits of air pollution control // Frontiers of Environmental Science & Engineering. 2015. V.9, Is. 6. pp. 1056-1065.

[3] *Stefan Wiemann, Johannes Brauner, Pierre Karrasch, Daniel Henzen, Lars Bernard* Design and prototype of an interoperable online air quality information system // Environmental Modelling & Software. 2016. V. 79. pp. 354-66.

[4] *Mehran Ghodousi, Farideh Atabi, Jafar Nouri, Alireza Gharagozlou* Air quality management in Tehran using a multi-dimensional decision support system // Polish Journal of Environmental Studies. 2017. V. 26, Is. 2. pp. 593-603.

[5] *Hua Wang et al* Design and demonstration of a next-generation air quality attainment assessment system for $PM_{2.5}$ and $O_3$ // Journal of Environmental Sciences. 2015. V. 29. pp. 178-188.
[6] *Ашитко А.Г., Маньшина И.В.* Система мониторинга состояния качества атмосферного воздуха в г. Калуге // Вестник Калужского университета. 2014. Т. 22, № 1. С. 5-9.
[7] *Tom Savu, Bogdan Alexandru Jugravu, Daniel Dunea* On the development of a PM2.5 monitoring network for real-time measurements in urban environments // REVISTA DE CHIMIE. 2017. V. 68, № 4. pp. 796-801.
[8] *Lei Hu, Peng Yue, Mingda Zhang, Jianya Gong, Liangcun Jiang, Xining Zhang* Task-oriented sensor web data processing for environmental monitoring // Earth Science Informatics. 2015. V. 8, Is. 3. pp. 511-525.
[9] *Hrebtov M, Hanjalic K.* Numerical Study of Winter Diurnal Convection Over the City of Krasnoyarsk: Effects of Non-freezing River, Undulating Fog and Steam Devils // Boundary Layer Meteorology. 2017. V. 163, Is. 3. pp. 469-495.
[10] *Kadochnikov A.A., Tokarev A.V., Zavoruev V.V., Yakubailik O.E.* Prototype of city environmental monitoring system based on geoportal technologies // IOP Conference Series: Materials Science and Engineering, 2019, V. 537, 062052.
[11] *Yakubailik O.E., Kadochnikov A.A., Tokarev A.V.* Development of computational technologies and software for assessment of air pollution in Krasnoyarsk // IOP Conference Series: Earth and Environmental Science, 2018, V. 211, 012080.
[12] City quality monitoring system CityAir. **http://cityair.ru**.
[13] *Zavoruev V.V., Zavorueva E.N., Kadochnikov A.A., Tokarev A.V., Yakubailik O.E.* Assessment of the possibility of using CityAir air monitoring station in environmental engineering // IOP Conference Series: Materials Science and Engineering, 2019, V. 537, 062053.
[14] Dynamics of pollution of Krasnoyarsk by particulate matter in early February 2019. DOI: 10.6084/m9.figshare.8343944.v1.
[15] Dynamics of cleaning of Krasnoyarsk after particulate matter pollution in mid-February 2019. DOI: 10.6084/m9.figshare.8343962.v1.

# ОБНАРУЖЕНИЯ БОЛЕЗНЕЙ НА ПОСЕВАХ ПШЕНИЦЫ ПО ДАННЫМ ГИПЕРСПЕКТРАЛЬНОЙ СЪЕМКИ

*Дубровская О.А.[(1)], Гурова Т.А.[(2)], Пестунов И.А.[(1)]*
[(1)] Институт вычислительных технологий СО РАН, Новосибирск
[(2)] Сибирский федеральный научный центр агробиотехнологий РАН, Краснообск

Представлены спектральные кривые для различных сортов пшеницы, полученных с помощью гиперспектральной камеры Spesim IQ. В результате анализа гиперспектральных данных определены наиболее информативные вегетационные индексы при обнаружении патогена.
*Ключевые слова: заболевания пшеницы, корневая гниль, гиперспектральные данные, вегетационные индексы, спектральные характеристики.*

**Введение.** По данным ФГБУ «Россельхозцентр» по Новосибирской области в 2018 году в результате фитосанитарной оценки на посевах зерновых колосовых культур отмечено наличие следующих листо-стеблевых инфекций: септориоза, мучнистой росы, бурой листовой ржавчины, а также тёмно-бурой пятнистостью листьев пшеницы (одна из форм гельминтоспориозной корневой гнили). Вредоносность заболевания выражается в снижении урожая, уменьшении числа и массы зерен в колосе [1].

Наиболее вредоносными заболеваниями на пшенице в Новосибирской области считаются корневые гнили. Из них существенное значение имеют гельминтоспориоз (возбудитель гриб *Bipolaris sorokiniana* Shoem. = *Drechslera sorokiniana Subram. et Jain, Helminthosporium sativum Pam.*), поражающий практически все органы растения (первичные, вторичные корни, колеоптиле, стебель, листья, зерно), и фузариозы (возбудители грибы рода *Fusarium*), поражающие корневую систему и зерно. Болезнь приводит к гибели всходов, отставанию в росте, отмиранию продуктивных стеблей, пустоколосице, щуплости зерна. Потери урожая в среднем составляют 15% в результате снижения продуктивной кустистости, озернѐнности колоса и массы зерна [2, 3].

В отдельные годы недобор урожая от септориоза может составлять 10% и более, от мучнистой росы – 15-20% и более. В годы массового развития бурой ржавчины недобор зерна яровой пшеницы достигает 25-40% и более, стеблевой ржавчины – 60-70% и более [2].

Обследование посевов на заражённость корневыми гнилями приурочивают к периодам всходы – формирование зерновки – созревание зерна. Второй учёт корневых гнилей, а также пустоколосости и белостебельности проводят в фазу молочной спелости зерна. Степень развития заболевания устанавливают, как средний показатель поражённости растений, в котором учтено число (%) больных растений и степень (балл) поражения.

Цель данной работы – проведение экспериментальных исследований по выявлению информативности оптических характеристик различных сортов пшеницы при действии возбудителя обыкновенной корневой гнили злаков.

**Лабораторные эксперименты и данные.** В настоящее время для фитосанитарного мониторинга посевов различных сельскохозяйственных культур применяются RGB, мульти- и гипеспектральные камеры. Наиболее точная диагностика заболеваний получена при использовании гиперспектрометров.

Гиперспектральные данные были получены с помощью спектрометра Specim IQ по контрольным и зараженным проросткам пшеницы (экспериментальных образцов). Спектрометр позволяет получать спектры отражения в диапазоне 400-1000 нм (204 канала).

Исследования проводились в лабораторных условиях (вегетационный опыт – водные культуры) на проростках районированных сортов яровой пшеницы с разной восприимчивостью к обыкновенной корневой гнили злаков (возбудитель *Bipolaris sorokiniana* Shoem., сокр. *B. Sorokiniana*).

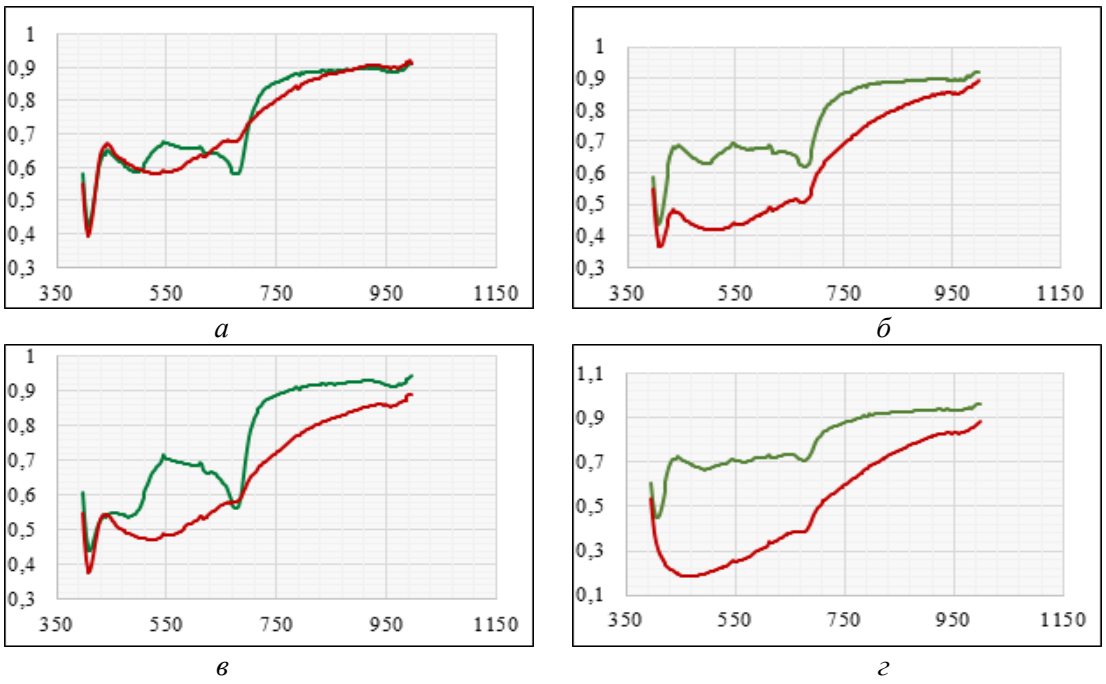



Рис. 1. Спектральные кривые здоровых проростков (зеленая кривая) и зараженных корневой гнилью *B. sorokiniana* (красная кривая) для различных сортов пшеницы: *а* - Новосибирская 18; *б* – Омская 18; *в* - Новосибирская 44; *г* - Сибирская 21.

Варианты опытов: 1) контроль, 2) инфекционный фон (возбудитель обыкновенной гнили *B. Sorokiniana*) 5000 конидий на зерно.

Проростки культивировали в рулонной культуре в универсальной камере для моделирования климатических условий среды «Биотрон 7» при температуре 18 и 22 ºС, влажности 60%, фотопериоде «день-ночь» (16 и 8 часов соответственно), освещенности (день) – 20000 лк до фазы 1-2 листа. Инфицированные и контрольные растения выращивали на водопроводной воде.

**Результаты обработки гиперспектральных данных.** Анализ спектральных кривых, полученных для проростков сортов Новосибирская 18, Омская 18, Новосибирская 44 и Сибирская 21 показал, что отражательная характеристика проростков пшеницы при действии возбудителя обыкновенной корневой гнили злаков выражена в двух частях спектра: видимой (400-700 нм) и ближней инфракрасной области (700-900 нм) (рис. 1). У проростков контрольных вариантов в видимой части спектра наблюдается возрастание отражательной способности с небольшим пиком в зеленой области (около 550 нм), затем идёт понижение из-за сильного поглощения света пигментами растений с экстремумом при длине волны около 680 нм. В ближней инфракрасной области, отражательная способность проростков контрольных и опытных вариантов повышается, что связано с внутренним рассеянием света мезофиллом [4]. Снижение отражательной способности у зараженных проростков, наиболее выраженное у сорта Сибирская 21, возможно связано с формированием адаптационных реакций при патогенезе, приводящих к увеличению синтеза пигментов.

Различия отражательных характеристик в определенных зонах спектра послужили основой для использования вегетационных индексов для обнаружения и диагностики корневой гнили на посевах и распознавания особенностей здоровых и пораженных заболеванием всходов пшеницы. В результате анализа различных вегетационных индексов, ранее используемых при диагностике и мониторинге развития других заболеваний пшеницы [5], а также на основе анализа спектральных характеристик, полученных при лабораторном эксперименте, было выбрано 13 вегетационных индексов для идентификации корневой гнили (*B. sorokiniana*), которые представлены в таблице.

Вегетационные индексы, используемые для идентификации заболеваний пшеницы.

| Индекс | Название индекса | Формула* |
|---|---|---|
| ***m*SR$_{705}$ (**Modified Red Edge Simple Ratio Index) | Модифицированный относительный индекс в крайнем красном спектре | $(R_{750} - R_{445})/(R_{750} + R_{445})$ |
| **NDVI$_{705}$ (**Red Edge Normalized Difference Vegetation Index) | Нормализованный разностный вегетационный индекс в крайнем красном спектре | $(R_{750} - R_{705})/(R_{750} + R_{705})$ |
| ***m*NDVI$_{705}$ (**Modified Red Edge Normalized Difference Vegetation Index) | Модифицированный нормализованный разностный вегетационный индекс | $(R_{750} - R_{705})/(R_{750} + R_{705} - 2R_{445})$ |
| **NBNDVI (**Narrow-Band Normalized Difference Vegetation Index) | Узкополосный нормализованный разностный вегетационный индекс | $(R_{850} - R_{680})/(R_{850} + R_{680})$ |
| **RVSI (**Red-Edge Vegetation Stress Index) | Вегетационный индекс состояния в крайнем красном спектре | $[(R_{712} + R_{752})/\ 2] - R_{732}$ |
| **PSRI (**Plant Senescence Reflectance Index) | Индекс отражения огрубевшего углерода в растительных тканях | $(R_{Red} - R_G)/R_{Nir}$<br>$(R_{678} - R_{500})/R_{750}$ |
| **ARI (**Anthocyanin Reflectance Index) | Антоциановый отражательный индекс | $(R_{550})^{-1} - (R_{700})^{-1}$ |
| **PRI (**Photochemical/Physiological Reflectance Index) | Фотохимический индекс отражения | $(R_{531} - R_{570})/(R_{531} + R_{570})$ |
| **SIPI (**Structural Independent Pigment Index) | Структурный индекс интенсивности пигментов | $(R_{800} - R_{445})/(R_{800} + R_{680})$ |
| **PhRI (**Physiological Reflectance Index) | Индекс физиологического отражения | $(R_{550} - R_{531})/(R_{550} + R_{531})$ |
| **NPCI (**Normalized Pigment Chlorophyll Index) | Нормализованный индекс поглощения в хлорофилле | $(R_{680} - R_{430})/(R_{680} + R_{430})$ |
| **MCARI (**Modified Chlorophyll Absorption Ratio Index) | Модифицированный относительный индекс поглощения в хлорофилле | $3 \cdot [(R_{701} - R_{671}) - 0.2 \cdot (R_{701} - R_{549}) \cdot \left(\frac{R_{701}}{R_{671}}\right)]$ |
| **TVI (**Triangular Vegetation Index) | Треугольный вегетационный индекс | $0.5 \cdot [120 \cdot (R_{Nir} - R_G) - 200 \cdot (R_{Red} - R_G)]$<br>$0.5 \cdot [120 \cdot (R_{750} - R_{550}) - 200 \cdot (R_{670} - R_{550})]$ |

∗ $R_X$ – отражательная способность на соответствующей длине волны или в соответствующем диапазоне спектра: B (Blue – 450-520 нм), G (Green – 520-600 нм), Red (630-690 нм), RE (700-730 нм), Nir (760-900 нм).

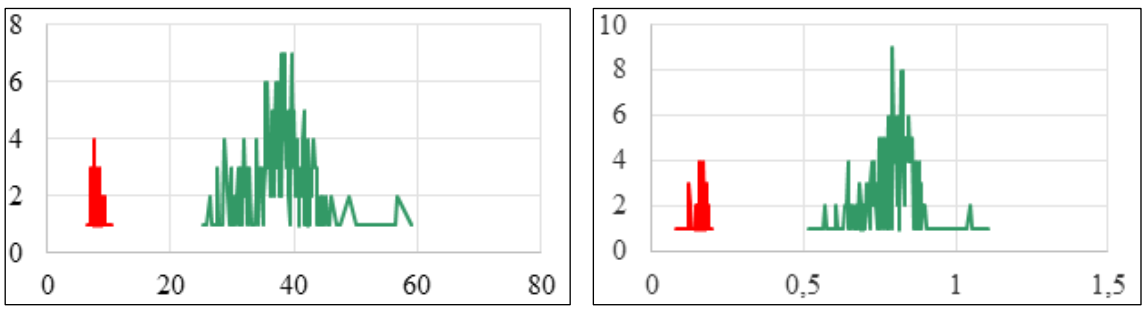



Рис. 2. Гистограммы значений вегетационных индексов TVI (слева) и MCARI (справа) для здоровых проростков (зеленая кривая) и зараженных корневой гнилью (*B. sorokiniana*) (красная кривая).

Анализ гистограмм значений вегетационных индексов показал (рис. 2), что индексы TVI и MCARI являются наиболее информативными для обнаружения патогена на проростках пшеницы по данным гиперспектральной съемки.

**Заключение.** В ходе проведенных лабораторных экспериментов были получены спектральные характеристики для здоровых проростков и зараженных корневой гнилью (*B. sorokiniana*) различных сортов пшеницы: Новосибирская 18, Омская 18, Новосибирская 44, Сибирская 21. Статистический анализ гистограмм 13-ти вегетационных индексов, которые широко используются при обнаружении различных болезней посевов пшеницы, показал, что индексы TVI и MCARI являются наиболее информативными признаками при обнаружении корневой гнили (*B. sorokiniana*) на посевах пшеницы по гиперспектральным данным, полученных с помощью спектрометра Specim IQ.

*Работа выполнена в рамках проекта комплексной программы фундаментальных исследований СО РАН «Междисциплинарные интеграционные исследования» на 2018-2020 гг. № 49 «Разработка цифровых технологий раннего обнаружения и локализации поражений посевов сельскохозяйственных культур».*

## ЛИТЕРАТУРА

[1] *Койшыбаев М., Муминджанов Х.* Методические указания по мониторингу болезней, вредителей и сорных растений на посевах зерновых культур. Анкара: Продовольственная и сельскохозяйственная организация Объединенных Наций, 2016. 28 с.

[2] *Долженко В.И., Власенко Н.Г., Власенко А.Н. и др*. Зональные системы защиты яровой пшеницы от сорняков, болезней и вредителей в Западной Сибири. Новосибирск: ГНУ СибНИИЗиХ, 2014. 124 с.

[3] *Гурова Т.А., Денисюк С.Г., Луговская О.С., Свежинцева Е.А., Минеев В.В.* Методические положения ранней диагностики устойчивости сортов яровой пшеницы и ячменя к совокупному действию стрессоров. Новосибирск: СФНЦА РАН, 2017. 62 с.

[4] *Jan B., Jörg S., Lutz P.* Detection of early plant stress responses in hyperspectral images // ISPRS Journal of Photogrammetry and Remote Sensing. 2014, Vol. 93. P. 98-111. [Электронный ресурс]. 2014. DOI: 10.1016/j.isprsjprs.2014.03.016.

[5] *Дубровская О.А., Гурова Т.А., Пестунов И.А., Котов К.Ю.* Обзор современных методов обнаружения болезней на посевах яровой пшеницы по данным дистанционного зондирования // Сибирский вестник сельскохозяйственных науки. 2018. Т. 48, № 6. С.76-89.

# ВЫЧИСЛЕНИЕ СОЧЕТАНИЯ ГЕНЕТИЧЕСКИХ МАРКЕРОВ В ОПРЕДЕЛЕНИИ ИДЕНТИЧНОСТИ ОБРАЗЦОВ ДРЕВЕСИНЫ

*Пятаев А.С.[(1)(2)], Ибе А.А.[(2)], Шилкина Е.А.[(2)]*
[(1)] Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева, Красноярск
[(2)] Филиал ФБУ «Рослесозащита» – «ЦЗЛ Красноярского края», Красноярск

В настоящей работе предложен метод определения сочетания минимального оптимального набора микросателлитных маркеров, применяемых для изучения генетической структуры популяции древесных растений на примере сосны обыкновенной. Разработанный метод позволяет оптимизировать проведение генетических анализов в задачах мониторинга состояния лесных генетических ресурсов.
*Ключевые слова: генетическая структура, микросателлитные маркеры, сосна обыкновенная.*

**Введение.** В течение последнего десятилетия в мире все большее внимание уделяется происхождению лесоматериалов из-за негативных экономических и экологических последствий незаконных рубок [1, 2]. Статистика импорта и экспорта, проведенная экспертами WWF в 2008 г., показала, что значительное количество незаконно заготовленной древесины поступает на европейский и китайский рынки из России и стран Восточной Европы. В Российской Федерации только за 2008-2016 гг. зафиксировано 197 228 случаев незаконных рубок, общий ущерб составил 104,5 млрд. руб., возмещенный – 2,83 млрд. руб. (2,7% суммы начисленного ущерба). Нелегальное использование лесов выявлено практически во всех регионах Российской Федерации [3]. Чтобы пресекать эти правонарушения, сотрудники органов внутренних дел должны иметь возможность доказательно выявлять истинное происхождение перевозимой древесины [4].

Одним из перспективных направлений доказательства законности или незаконности нахождения древесины в обороте является применение молекулярно-генетических методов анализа. Данные методы основаны на использовании генетических маркеров – микросателлитов – варьирующих участков (локусов) в ядерной ДНК и ДНК органелл (митохондрий и пластид), состоящих из тандемно повторяющихся нуклеотидных последовательностей. Эти маркеры характеризуются высоким уровнем полиморфизма и часто встречаются в геноме [5]. Тем не менее, в настоящее время не существуют методов, позволяющих проводить подбор оптимальных минимальных сочетаний молекулярных праймеров, дающих наименьшую вероятность случайного совпадения родственных многолокусных генотипов. Таким образом, задача разработки метода определения оптимальной последовательности микросателлитных маркеров, применяемых для изучения генетической структуры популяции древесных растений на примере сосны обыкновенной, является актуальной.

**Метод определения сочетания генетических маркеров**. В качестве эталонной выборки послужили образцы древесины сосны обыкновенной (Pinus sylvestris L.), отобранные в насаждении, произрастающем в Балахтинском районе Красноярского края. Объем выборки составил 29 образцов, что является репрезентативным и принятым в анализе ядерных кодоминатных ядерных маркеров. При молекулярно-генетических исследованиях количество образцов в выборке берется в среднем не более 30 особей на ценопопуляцию [6-7]. Древесину тщательно измельчали, гомогенизировали и выделяли ДНК CTAB-методом [8]. В основе метода лежит распад клеток под влиянием цетилтриметиламмонийбромида (CTAB), удаление белков с помощью хлороформа и осаждение ДНК изопропанолом. В результате предварительной работы по выявлению наиболее полиморфных и стабильно амплифицирующихся локусов было отобрано 10 микросателлитных маркеров. В табл. 1 представлены характеристики рекомендованных для работы ядерных микросателлитных локусов для сосны обыкновенной [9-12].

Табл. 1. Характеристика отобранных для работы ядерных микросателлитных локусов для сосны обыкновенной.

| п/п | Локус | Мотив | Размеры ампликона, п.н. | Температура отжига, °С | Число аллелей | Источник |
|---|---|---|---|---|---|---|
| **1** | psy119 | $(GCT)_7$ | 315-324 | 55 | 3 | [9] |
| **2** | psy157 | $(ACC)_7$ | 187-202 | 55 | 6 | [9] |
| **3** | PtTx3116 | $(TTG)_7(TTG)_5$ | 122-226 | 55 | 8-10 | [10] |
| **4** | PtTx3107 | $(CAT)_{14}$ | 150-182 | 55-45↓1 | 5-6 | [10] |
| **5** | PtTx4001 | $(CA)_{15}$ | 201-224 | 60-50↓1 | 4-7 | [10] |
| **6** | lw_isotig21953 | $(ATGGG)_7$ | 208 | 60 | 7 | [11] |
| **7** | PtTx4011 | $(CA)_{20}$ | 230-284 | 60-50↓1 | 21 | [10] |
| **8** | SPAC11.4 | $(AT)_5(GT)_{19}$ | 130-170 | 65-55↓1 | 38 | [12] |
| **9** | lw_isotig04306 | $(TCC)_7$ | 196 | 55 | 3 | [11] |
| **10** | lw_isotig27940 | $(TGGA)_5$ | 231 | 55 | 3 | [11] |

Табл. 2. Вероятность идентичности особей с одинаковым многолокусным генотипом при увеличении комбинации генетических маркеров.

| | **Комбинации локусов** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **1+2** | **1+2+3** | **1+2+3+4** | **1+2+3+4+5** | **1+2+3+4+5+6** | **1+2+3+4+5+6+7** | **1+2+3+4+5+6+7+8** | **1+2+3+4+5+6+7+8+9** | **1+2+3+4+5+6+7+8+9+10** |
| **Вероятность идентичности** | 1 | 0,76 | 0,059 | $6{,}3 \cdot 10^{-3}$ | $1{,}3 \cdot 10^{-3}$ | $6{,}6 \cdot 10^{-5}$ | $2{,}3 \cdot 10^{-5}$ | $1{,}5 \cdot 10^{-6}$ | $2{,}3 \cdot 10^{-7}$ | $2{,}2 \cdot 10^{-8}$ |

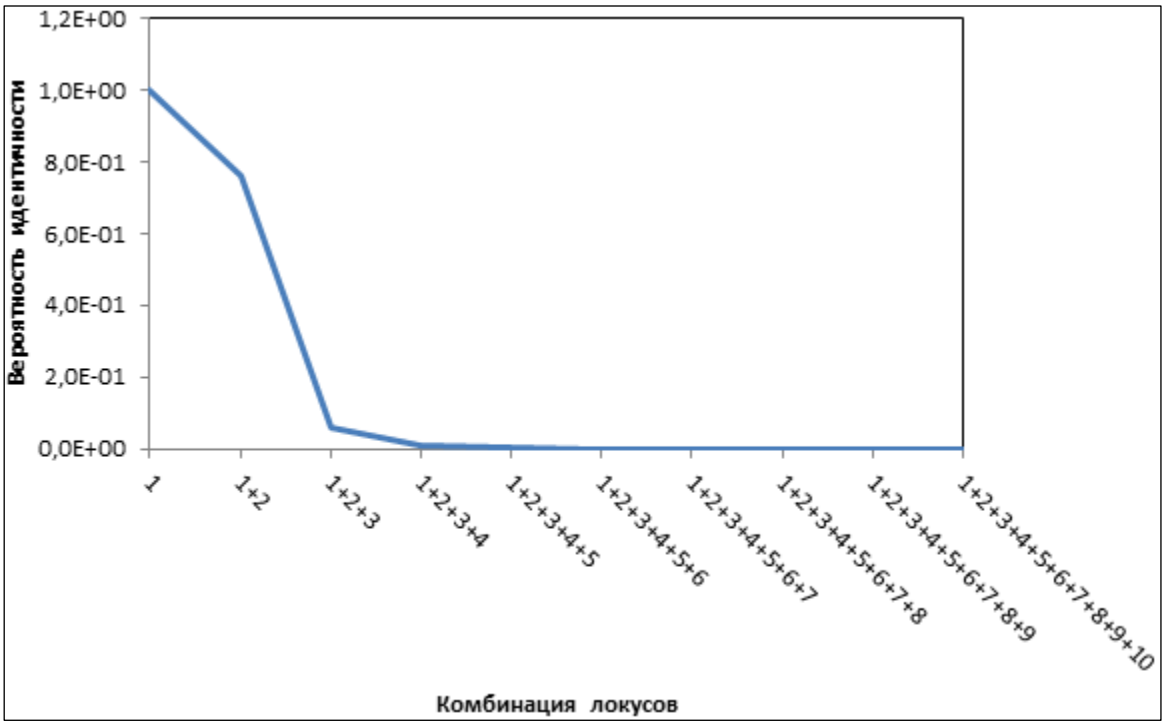



Рис. 1. Вероятность идентичности особей с одинаковым многолокусным генотипом при увеличении комбинации генетических маркеров.

В настоящее время обработка результатов генетических анализов осуществляется с помощью проприетарного программного обеспечения GenAlEx [13], свободно распространяемой надстройки для MS Excel. Полученные с помощью программы GenAlEx результаты приведены в табл. 2. Наименьшая вероятность случайного совпадения многолокусных генотипов (2,2E-08) достигается только при комбинации 10 генетических маркеров: psy119, psy157, PtTx3116, PtTx3107, PtTx4001, lw_isotig21953, PtTx4011, SPAC11.4, lw_isotig04306, lw_isotig27940. Данное значение говорит о низкой вероятности случайного совпадения генотипов [14]. Порядок генетических маркеров представлен в табл. 1 и на рис. 1.

Проведение генетического анализа для определения идентичности образцов древесины, используя все 10 маркеров, весьма трудозатратно и дорого. Для снижения этих затрат предложен метод определения оптимальной минимальной последовательности маркеров для исключения ложноположительной идентификации образцов. Цель метода – подобрать такое количество и последовательность маркеров, которые дадут минимальную вероятность ложноположительной идентификации образцов.

Метод, предложенный в данной работе, основан на анализе встречаемости пар аллелей каждого локуса среди образцов заведомо уникальной выборки. В качестве эталонной выборки использованы образцы сосны обыкновенной, отобранные в естественном насаждении, произрастающем в Балахтинском районе Красноярского края. Эталонная выборка состоит из 29 пронумерованных уникальных образцов и на начальном этапе проверяется 10 маркерами.

Обозначим

$$M = \{M_i, i = 1..10\},$$

как множество результатов действия маркеров. Каждый маркер по каждому образцу дает пару значений:

$$M_i = \{m_j = (a_j, b_j) : j = 1..29, a_j \in \mathbb{N}, b_j \in \mathbb{N}\}.$$

В табл. 3 продемонстрированы значения молекулярного веса пар аллелей выборки, измеряемые в «bp» (от англ. base pair), либо в п.н. – парах нуклеотидов в русскоязычной литературе.

Идентификаторы образцов группируются по приведенным в таблице 3 парам значений по каждому маркеру:

$$G_i = \{g_k(i) = \{m_n\} : k \in \mathbb{N}, k < 30, n \in \mathbb{N}, n < 30\}.$$

Результат группировки образцов по значениям аллелей, на примере маркера lw_isotig04306 приведен в табл. 4.

Табл. 3. Пары аллелей внутри локусов по образцам.

| № обр. | lw_isotig 21953 | lw_isotig0 4306 | lw_isotig 27940 | PtTx3107 | … | SPAC 11.4 |
|---|---|---|---|---|---|---|
| 1 | 258/258 | 187/187 | 235/255 | 159/165 | … | 146/156 |
| 2 | 223/263 | 187/193 | 247/247 | 153/159 | … | 138/142 |
| 3 | 203/223 | 187/187 | 247/247 | 159/165 | … | 152/152 |
| 4 | 248/258 | 184/187 | 247/247 | 180/180 | … | 142/152 |
| 5 | 258/258 | 187/193 | 247/247 | 159/165 | … | 146/146 |
| 6 | 203/203 | 175/187 | 235/235 | 162/165 | … | 142/146 |
| … | … | … | … | … | … | … |
| 24 | 248/263 | 184/187 | 235/235 | 153/153 | … | 138/146 |
| 25 | 248/248 | 178/187 | 255/255 | 159/168 | … | 146/160 |
| 26 | 223/243 | 193/193 | 239/247 | 159/159 | … | 138/160 |
| 27 | 248/253 | 187/193 | 239/255 | 165/165 | … | 138/142 |
| 28 | 248/253 | 187/187 | 239/255 | 165/165 | … | 138/146 |
| 29 | 203/203 | 193/193 | 255/263 | 162/165 | … | 138/138 |

Табл. 4. Группировка образцов по значениям аллелей маркера lw_isotig04306.

| **lw__isotig04306** | **ids** |
|---|---|
| 175/187 | {6,19} |
| 178/187 | {18,25} |
| 181/187 | {16,17,22} |
| 184/184 | {7,10,13} |
| 184/187 | {4,11,12,20,24} |
| 184/193 | {15} |
| 187/187 | {1,3,8,21,23,28} |
| 187/193 | {2,5,9,14,27} |
| 193/193 | {26,29} |

Следующим этапом работы метода является ранжирование множества сгруппированных по парам из множеств $M_i$ идентификаторов $G_i$ по количеству уникальных пар, т.е. в приоритете $G_i$, у которых максимальное количество $|g_k(i)| = 1$.

Далее идет итеративный процесс пересечения множеств $G_i$. Каждое множество сгруппированных генотипов взаимно пересекается с другими множествами сгруппированных генотипов. На первом шаге итерации происходит попарное пересечение множеств $G_i$. Из полученных результатов пересечений остаются только те множества, у которых существуют подмножества мощностью не меньше двух:

$$R = \{R_l = g_m(i) \cap g_n(i) : |R_l| > 1;\ l, m, n, i, j \in \mathbb{N};\ m, n < 30;\ i, j \leq 10\}.$$

Затем $R_l$ ранжируются по возрастанию мощности. На следующих итерациях происходит пересечение $R$ с оставшимися $G_i$:

$$R = R \cap G_i.$$

Процесс идет до тех пор, пока пересечение не станет пустым множеством. Таким образом, анализируя тестовую выборку подобным способом, можно получить оптимальную последовательность маркеров, которая однозначно идентифицирует образцы проверяемой породы.

Для образцов сосны обыкновенной, отобранных в естественном насаждении, произрастающем в Балахтинском районе Красноярского края, оптимальным минимальным сочетанием генетических маркеров оказалась последовательность lw_isotig21953, SPAC11.4, PtTx3107.

Для проверки метода была отобрана и проанализирована контрольная группа образцов древесины сосны обыкновенной из этого же насаждения. В табл. 5 представлены результаты группировки образцов контрольной группы по значениям аллелей маркера lw_isotig04306.

Анализ образцов контрольной группы показал эффективность подобранной последовательности маркеров. В случае ложноположительной идентификации образцов контрольная группа дополнительно анализируется маркером lw_isotig27940.

Табл. 5. Группировка образцов контрольной группы по значениям аллелей маркера lw_isotig04306.

| **lw_isotig04306** | **ids** |
|---|---|
| 175/187 | {15,16} |
| 178/187 | {24,23} |
| 181/187 | {14,13,22} |
| 184/184 | {9,19,21} |
| 184/187 | {29,10,11,12,20} |
| 184/193 | {28} |
| 187/187 | {26,27,5,6,8,18} |
| 187/193 | {25,3,4,7,17} |
| 193/193 | {1,2} |

Табл. 6. Вероятность идентичности особей с одинаковым многолокусным генотипом при увеличении комбинации генетических маркеров.

| | **Комбинации локусов** | | | |
|---|---|---|---|---|
| | **6** | **6+8** | **6+8+4** | **6+8+4+10** |
| **Вероятность идентичности** | $5{,}2 \cdot 10^{-2}$ | $3{,}4 \cdot 10^{-3}$ | $3{,}6 \cdot 10^{-4}$ | $3{,}5 \cdot 10^{-5}$ |

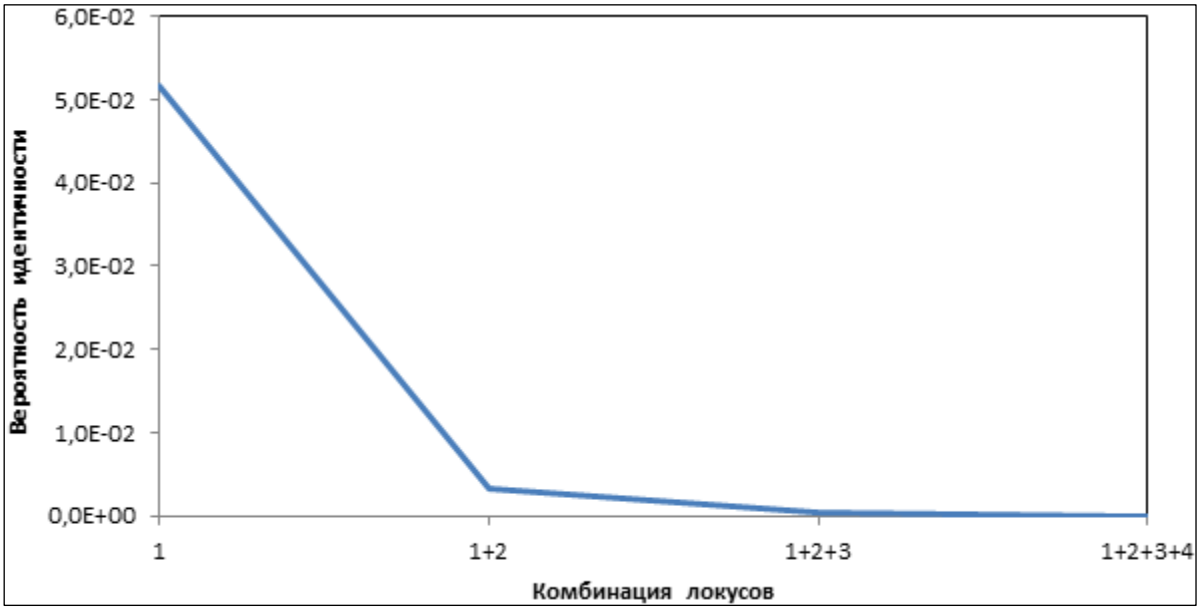



Рис. 2. Вероятность идентичности особей с одинаковым многолокусным генотипом при увеличении комбинации генетических маркеров.

Результаты анализа контрольной группы образцов были проверены программой GenAlEx с указанной последовательностью маркеров. Вероятность идентичности особей с одинаковым многолокусным генотипом при увеличении комбинации генетических маркеров приведены в табл. 6 и на рис. 2.

Порядок генетических маркеров представлен в табл. 1.

Сравнивая результаты табл. 2 и 6, видно, что при использовании отсортированной последовательности маркеров вероятность случайного совпадения многолокусных генотипов оказалась ниже и, для достижения приемлемого результата, оказалось достаточным использование всего трех маркеров.

**Заключение.** Для определения идентичности образцов сосны обыкновенной, отобранных в естественном насаждении, произрастающем в Балахтинском районе Красноярского края, минимальным оптимальным сочетанием генетических маркеров оказалась последовательность lw_isotig21953, SPAC11.4, PtTx3107. Эффективность подобранной последовательности маркеров проверена на контрольной группе. Использование предложенного метода подбора оптимальной последовательности микросателлитных маркеров позволяет существенно сократить трудовые, временные и материальные затраты при определении идентичности образцов древесины. В дальнейших исследованиях планируется использовать данный алгоритм подбора маркеров применительно к другим выборкам, подлежащим аналогичному генетическому анализу.

## ЛИТЕРАТУРА


[1] *Céline Blanc-Jolivet, Yulai Yanbaev, Birgit Kersten, Bernd Degen*. A set of SNP markers for timber tracking of *Larix* spp. in Europe and Russia // Forestry. 2018. С. 1–15.

[2] WWF World Wide Fund For Nature 2008 Illegal wood for the European market. Frankfurt a M: WWF Germany, 43p.

[3] *Кузьмичев, Е. П., Трушина И. Г., Лопатин Е. В*. Объемы незаконных рубок лесных насаждений в Российской Федерации // Лесохозяйственная информация. 2018. № 1. С. 63–77.
[4] *Латов Ю.В., Жаворонков Ю.М.* Достижения и перспективы применения дендрохронологической экспертизы для борьбы с незаконными рубками лесных насаждений // Теория и практика правоохранительной деятельности: Труды Академии управления МВД России. 2013. № 4 (28). С. 44-48.
[5] *Ильинов А.А., Раевский Б.В.* Состояние генофонда сосны обыкновенной *Pinus sylvestris* L. в Карелии // Сибирский лесной журнал. 2016. № 5. С. 45-54.
[6] *Nei M.* Estimation of average heterozygosity and genetic distance from a small number of individuals // Genetics. 1978. V. 89. P. 583–590.
[7] *Shurkhal A. V., Podogas A. V., Zhivotovsky L. A.* Allozyme differentiation in the genus Pinus // Silvae Genetica. 1992. V. 41. P. 105–109.
[8] *Devey M. E., Bell J. C., Smith D. N., Neal D. B., Moran G. F.* A genetic linkage map for Pinus radiata based on RFLP, RAPD, and microsatellite markers // Theor. Appl. Genet. 1996. V. 92. Iss. 6. P. 673–679.
[9] *Sebastiani F., Pinzauti F., Kujala S. T., Gonzalez-Martinez S. C., Vendramin G. G.* Novel polymorphic nuclear microsatellite markers for Pinus sylvestris L. // Conservation Genet. Resour. 2012. V.4. Iss. 2. P. 231–234.
[10] *Belletti P., Ferrazzini D., Piotti A., Monteleone I., Ducci F.* Genetic variation and divergence in Scots pine (Pinus sylvestris L.) within its natural range in Italy // European Journal of Forest Research. 2012. V. 131. Iss. 4. P. 1127–1138.
[11] *Fang P., Niu Sh., Yuan H., Li Zh., Zhang Yu., Yuan L., Li W.* Development and characterization of 25 EST-SSR markers in Pinus sylvestris var. mongolica (Pinaceae) //Applications in Plant Sciences. 2014. V. 2. Iss. 1. P. 1–4.
[12] *Soranzo N., Provan J., Powell W.* Characterization of microsatellite loci in Pinus sylvestris L. // Molecular Ecol. 1998. V. 7. P. 1247–1263.
[13] *Peakall R., Smouse P. E.* GenAlEx v. 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – an update // Bioinformatics. 2012. V. 28. Iss. 19. P. 2537–2539.
[14] *Brown S. M.* Methods of genome analysis in plants // Ed. P.P. Jauhar. N.-Y., London. Tokyo. 1996. P. 147–159.

