

На правах рукописи



Константинова Анна Михайловна

**АЛГОРИТМЫ И ПРОГРАММНАЯ
ИНФРАСТРУКТУРА ДЛЯ РЕШЕНИЯ ЗАДАЧ
СПУТНИКОВОГО ОБЪЕКТНОГО МОНИТОРИНГА**

Специальность:

2.3.5. — Математическое и программное обеспечение
вычислительных систем, комплексов и компьютерных сетей

АВТОРЕФЕРАТ

диссертации на соискание учёной степени
кандидата технических наук

Москва — 2023

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте космических исследований Российской академии наук (ИКИ РАН)

Научный руководитель: **Лупян Евгений Аркадьевич**
доктор технических наук,
заведующий отделом «Технологии спутникового мониторинга» Федерального государственного бюджетного учреждения науки «Институт космических исследований Российской академии наук», Москва

Официальные оппоненты: **Александр Анатольевич Иванович**
доктор технических наук, заведующий лабораторией спутникового мониторинга Федерального государственного бюджетного учреждения науки «Институт автоматизации и процессов управления Дальневосточного отделения Российской академии наук», Владивосток

Макаренков Александр Алексеевич
кандидат технических наук, ведущий научный сотрудник, Федеральное государственное бюджетное образовательное учреждение высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», Рязань

Ведущая организация: Федеральное государственное автономное образовательное учреждение высшего образования «Сибирский федеральный университет», Красноярск

Защита диссертации состоится **5 апреля 2023 г. в 12 ч 00 мин** на заседании диссертационного совета 24.2.375.01 (Д 212.211.01) в Федеральном государственном бюджетном образовательном учреждении высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина» по адресу: **390005, Рязань, ул. Гагарина, д. 59/1**

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «РГРТУ» и на официальном сайте <http://www.rsreu.ru/>

Автореферат разослан «___» _____ 2023 г.

Учёный секретарь диссертационного совета 24.2.375.01 (Д 212.211.01)
доктор технических наук, доцент



А. В. Пруцков

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования и степень её разработанности. Современные спутниковые системы дистанционного зондирования Земли (ДЗЗ) представляют собой одно из самых востребованных средств наблюдения Земли. Они позволяют получать объективную информацию о состоянии земной поверхности, океана и атмосферы без доступа к территории исследования и на её основе принимать обоснованные решения в области устойчивого развития будущего человечества. В последние десятилетия наблюдается стремительный рост числа спутниковых систем ДЗЗ и повышение качества и детальности предоставляемых ими данных, что ведёт к резкому увеличению объёмов получаемой информации и требует разработки эффективных подходов для организации работы с ней. Вопросам создания систем для работы со спутниковой информацией посвящено большое число работ (Александрин А. И., Асмус В. В., Еремеев В. В., Кузнецов А. Е., Лупян Е. А., Милехин О. Е., Саворский В. П., Gorelick N. и др.). Несмотря на это, актуальной остаётся проблема быстрого извлечения полезной информации из большого объёма доступных спутниковых изображений, например, для построения временных рядов характеристик различных наблюдаемых объектов и явлений. Одним из способов решения данной проблемы представляется применение объектного подхода для организации мониторинга, отличающегося от традиционного попиксельного анализа данных тем, что спутниковая информация хранится, обрабатывается и анализируется не для каждого пикселя получаемых данных, а интегрируется на однородные области — объекты. Самым популярным примером применения такого подхода в области ДЗЗ выступает мониторинг сельскохозяйственных (с.-х.) полей, основанный на расчёте средних значений вегетационных индексов по спутниковым данным в границах полей, позволяющий оценить развитие с.-х. культуры на поле. Следует отметить, что сельское хозяйство не единственное перспективное направление, где может применяться объектный подход. Он может быть эффективно использован для исследования многих типов природных и антропогенных объектов, таких как водоёмы, участки леса, кратеры вулканов и другие. Поэтому актуальной задачей становится создание универсальных методов и инструментов, позволяющих единообразно организовывать объектный мониторинг для проведения новых тематических исследований на основе большой выборки данных, доступных в архивах ДЗЗ. Такие инструменты должны обеспечивать хранение и обработку спутниковой информации, управление наборами необходимых характеристик для расчёта в границах объектов, построение и анализ долговременных рядов характеристик и т. п.

Цель диссертации — разработка математических и программных средств для повышения доступности и автоматизации мониторинга разнообразных природных и антропогенных объектов с использованием пополняющихся многолетних архивов спутниковых данных различного пространственного разрешения. Для этого в работе решаются следующие задачи:

- анализируется текущий этап развития дистанционного спутникового мониторинга, приводится обзор возможностей, предоставляемых современными системами космических наблюдений Земли, формулируются основные проблемы, связанные с резким ростом количества получаемой спутниковой информации;
- проводится анализ способов работы со сверхбольшими объёмами спутниковых изображений для быстрого извлечения из них информации, приводятся области исследований, где такие подходы активно применяются, описывается необходимость масштабирования объектного подхода на различные предметные области;
- на основе анализа исследовательских задач, где применим объектный подход, разрабатывается концептуальная модель данных для организации дистанционного спутникового объектного мониторинга;
- разрабатывается алгоритм автоматической потоковой обработки спутниковых данных, реализующей расчёт характеристик объектов, разрабатывается алгоритм фильтрации облачности для задач объектного мониторинга и для формирования очищенных рядов значений характеристик объектов;
- с использованием разработанных модели, алгоритмов и подходов создаётся архитектура, на основе которой строится программная инфраструктура объектного мониторинга, а также описывается взаимодействие функциональных модулей в составе инфраструктуры, создаются программные инструменты, реализующие данные модули на основе программных решений, развиваемых в Институте космических исследований Российской академии наук (ИКИ РАН);
- на основе разработанной программной инфраструктуры создаются блоки различных специализированных информационных систем (ИС) дистанционного мониторинга, таких как: EcoSatMS (<http://suvo.geosmis.ru/>), VolSat-View (<http://kamchatka.volcanoes.smislab.ru/>), Vera-Science (<http://sci-vega.ru/>).

Научная новизна работы состоит в том, что в ней разработана архитектура, предложены новые алгоритмы обработки данных и на их основе создана программная инфраструктура организации дистанционного спутникового объектного мониторинга, позволяющая универсальным способом обеспечивать автоматизированный спутниковый мониторинг разнообразных природных и антропогенных объектов.

Основные положения, выносимые на защиту:

1. На основе системного анализа задач, возникающих при постоянном дистанционном мониторинге различных объектов (объектный мониторинг), **разработана архитектура** организации взаимодействия программных систем, обеспечивающих автоматизированный сбор, обработку, хранение, визуализацию и распределённый анализ данных. Разработанная архитектура позволила унифицировать процессы работы с данными для реализации дистанционного объектного мониторинга при решении широкого спектра научных и прикладных задач (*соответствует п. 3 паспорта научной специальности*).
2. На основе проведения вычислительных экспериментов по обработке больших архивов спутниковых наблюдений **предложен алгоритм** автоматической потоковой обработки изображений для оценки характеристик различных объектов, позволяющий минимизировать временные затраты с учётом особенностей структуры хранения спутниковых данных и пространственного распределения наблюдаемых объектов. В случаях обработки компактно

расположенных объектов предложенный алгоритм позволяет увеличить скорость обработки до 3,5 раз (*соответствует п. 7 паспорта научной специальности*).

3. **Предложен алгоритм** фильтрации зашумлённых спутниковых данных, основанный на анализе процента облачности на спутниковых изображениях в районах расположения конкретных объектов. Алгоритм обеспечил получение автоматически очищенных от шумов рядов характеристик различных объектов, вычисляемых на основе данных спутниковых наблюдений, даже в условиях низкого качества фильтрации облачности (*соответствует п. 7 паспорта научной специальности*).
4. На основе предложенных в работе архитектуры и алгоритмов **создана программная инфраструктура** для организации глобально распределённого сбора, обработки, хранения, визуализации и анализа данных в целях решения задач дистанционного спутникового объектного мониторинга. Созданная программная инфраструктура показала свою работоспособность и эффективность при реализации блоков и подсистем в различных научных и прикладных информационных системах дистанционного мониторинга (*соответствует п. 9 паспорта научной специальности*).

Теоретическая и практическая значимость диссертации состоит в том, что в ней выполнен системный анализ проблемы обработки и извлечения необходимой информации из сверхбольших непрерывно пополняющихся архивов спутниковых данных для решения задач мониторинга природных и антропогенных объектов. Полученные в работе результаты позволяют повысить эффективность работы со сверхбольшими объёмами спутниковой информации и достаточно легко организовывать мониторинг объектов для различных тематических направлений исследования окружающей среды.

Апробация работы. Основные положения и результаты диссертационной работы докладывались и обсуждались на 10 конференциях. Среди них всероссийские конференции: «Фундаментальные и прикладные космические исследования» (Москва, 2019–2022), «Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов» (Новосибирск, 2021) — и международные конференции: «Современные проблемы дистанционного зондирования Земли из космоса» (Москва, 2019–2022), «Региональные проблемы дистанционного зондирования Земли» (Красноярск, 2018). Также автор приняла участие во всероссийском молодёжном конкурсе исследовательских работ и инженерных проектов памяти лётчика-космонавта А. А. Сереброва (Королёв, 2019), где в соавторстве стала победителем в секции «Прикладная космонавтика и дистанционное зондирование Земли». В 2020 г. автор стала победителем конкурса молодых учёных Всероссийской конференции с международным участием «Современные проблемы дистанционного зондирования Земли из космоса», представив работу, которая легла в основу диссертации.

Публикации. Результаты диссертации опубликованы в 16 статьях (в том числе в 12 статьях в рецензируемых изданиях, включённых в перечень ВАК) и в 25 тезисах докладов в материалах международных и всероссийских конференций. Получено 1 свидетельство о регистрации программы.

Внедрение результатов работы. Диссертационная работа выполнена в ИКИ РАН. На проведение исследований, входящих в данную работу, автором был выигран грант РФФИ № 19-37-90114. Результаты работы использованы при выполнении проекта Минобрнауки в рамках темы «Мониторинг» № 122042500031-8, гранта РФФИ № 18-29-24121мк.

Достоверность результатов работы обеспечивается практическим использованием созданной программной инфраструктуры при решении задач мониторинга различных природных и антропогенных объектов, что подтверждается актами внедрения результатов работы.

Методология и методы исследования. Теоретические результаты получены с использованием методов обработки изображений, статистических методов обработки данных, методов классификации данных, методов оптимизации.

Соответствие паспорту специальности 2.3.5. Диссертация соответствует паспорту специальности 2.3.5. — «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» в части:

- пункта 3 «Модели, методы, архитектуры, алгоритмы, языки и программные инструменты организации взаимодействия программ и программных систем»;
- пункта 7 «Модели, методы, архитектуры, алгоритмы, форматы, протоколы и программные средства человеко-машинных интерфейсов, компьютерной графики, визуализации, обработки изображений и видеоданных, систем виртуальной реальности, многомодального взаимодействия в социкиберфизических системах»;
- пункта 9 «Модели, методы, алгоритмы, облачные технологии и программная инфраструктура организации глобально распределённой обработки данных».

Структура и объём работы. Диссертация состоит из введения, четырёх глав, заключения, списка сокращений, списка литературы и трёх приложений. Общий объём работы составляет 145 страниц, в том числе 61 рисунок и 6 таблиц. Список использованной литературы включает 81 источник.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе представлены результаты обзора текущего состояния и перспектив развития области ДЗЗ. Совместно с преимуществами, появляющимися с эволюцией спутниковых систем ДЗЗ, обозначены и сложности, которые возникают у исследователей при анализе возрастающих объёмов спутниковых данных. Отмечено, что всё чаще появляется необходимость мониторинга объектов различной природы. Сформулирована задача организации мониторинга различных объектов, которая бы решала подзадачи эффективной обработки, хранения и анализа сверхбольших объёмов спутниковых данных.

В первом параграфе представлены результаты анализа тенденций развития спутниковых систем ДЗЗ и существующих подходов к работе со спутниковой информацией. Проведён обзор мировых практик по работе со сверхбольшими объёмами растровых изображений, получаемых в результате обработки спутниковых данных. Также рассмотрены некоторые подходы для извлечения полезной информации из спутниковых данных, в частности путём расчёта спектральных индексов. На стыке приведённых подходов сформулировано понятие дистанционного объектного мониторинга (ДОМ).

Во втором параграфе более подробно раскрыто, что собой представляет объектный подход к анализу спутниковых данных и в чём заключаются его преимущества. Приведён пример использования объектного подхода для мониторинга с.-х. полей, а также проанализировано, в каких направлениях тематических исследований воз-

никает задача организации ДОМ, и обозначена проблема необходимости создания унифицированной программной инфраструктуры, реализующей объектный подход.

В *третьем параграфе* сформулированы требования, предъявляемые к создаваемой программной инфраструктуре объектного мониторинга.

В *четвёртом параграфе* проведён обзор программных решений, разработанных в отделе «Технологии спутникового мониторинга» ИКИ РАН, которые могли бы быть задействованы в создании описываемой программной инфраструктуры.

В *пятом параграфе* приведено краткое описание центра коллективного пользования системами архивации, обработки и анализа данных спутниковых наблюдений ИКИ РАН (ЦКП «ИКИ-Мониторинг»), который предложен в качестве технической основы для реализации программной инфраструктуры ДОМ.

Во второй главе предложена архитектура построения программной инфраструктуры (положение 1, выносимое на защиту), реализующей объектный подход на основе программных решений, развиваемых в ИКИ РАН, а также разработана концептуальная модель данных для организации дистанционного спутникового объектного мониторинга, с использованием которой созданы базы данных (БД) объектов и их наблюдений.

В *первом параграфе* на основании сформулированных в первой главе требований и проведённого анализа задач, для решения которых создаётся программная инфраструктура, были описаны основные инструменты, необходимые для её реализации. Среди них:

- инструменты создания объектов совместно с атрибутами: геометрическим описанием и правилами мониторинга (наблюдений);
- инструменты автоматического расчёта показателей по полигонам объектов на основе оперативной и исторической спутниковой информации;
- инструменты ведения БД с информацией об объектах, содержащей как статическое описание объектов, так и информацию о показателях, рассчитанных для объектов на основе спутниковых данных;
- различные инструменты представления и визуализации данных, предназначенные для анализа информации, получаемой для объектов на основе спутниковых данных, в том числе для представления её в картографическом интерфейсе и в виде графиков, а также экспорта данных для их последующего анализа во внешних системах обработки данных;
- инструменты администрирования, управления и контроля работоспособности всех элементов создаваемой программной инфраструктуры.

Для реализации необходимых инструментов была разработана архитектура программной инфраструктуры, структурные элементы которой представлены на рис. 1. При этом структурные элементы архитектуры можно разделить на две основные группы: элементы, отвечающие за интерфейсную часть, с которой будут работать пользователи ИС, и элементы средств ведения БД и обработки.

Во *втором параграфе* на основе анализа тематических задач, в которых необходимо обеспечить мониторинг объектов, составлен ряд требований, предъявляемых к БД ДОМ. Для организации хранения информации ДОМ разработана концептуальная модель данных. Базовым элементом модели выступает исследовательский проект, в котором решается конкретная научная или прикладная задача. В рамках проекта выделяются различные районы исследований, внутри которых выбираются объекты, исследуемые в данном проекте. В зависимости от типа объекта (промышленный объект, участок, занятый естественной растительностью, водоём и т. п.) для проведения его мониторинга необходимо осуществлять получение тех или иных показателей, рассчитываемых по спутниковым данным. В рамках одного проекта могут

использоваться различные типы объектов. Создаваемые объекты определяются контуром (границей), набором показателей и периодом наблюдения (либо требованием к постоянному оперативному получению данных).

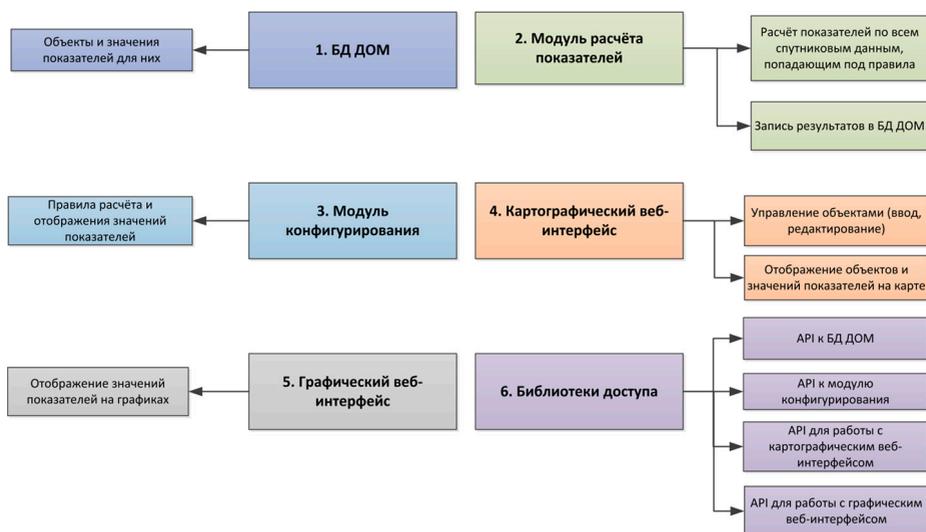


Рис. 1. Структурные элементы архитектуры программной инфраструктуры ДОМ

В *третьем параграфе* подробно описан программный комплекс (ПК) Objects-SurveysSMIS_les, ставший технической основой для построения и ведения БД ДОМ. Данный ПК реализовывался для хранения геометрической информации об объекте, полученной по конкретной сцене спутниковых данных (наблюдению) с привязкой к метainформации об этой сцене. Рассматриваемый ПК нацелен на хранение информации о разнообразных объектах, поэтому БД, входящая в состав ПК, имеет гибкую структуру и позволяет хранить произвольное количество атрибутов как для объекта, так и для его наблюдения, а также обеспечивать связи с внешними объектами, хранящимися в других БД. Была проведена имплементация концептуальной модели данных ДОМ в структуру хранения рассматриваемого ПК. При этом функциональность ПК была значительно расширена для того, чтобы полностью покрыть задачи организации хранения ДОМ. В частности, были реализованы следующие возможности:

- добавлен тип наблюдений, представляющий собой не геометрическое описание, а набор числовых значений показателей, рассчитанных по конкретной сцене спутниковых данных в границах объекта;
- добавлен специальный тип наблюдения, отвечающий за базовый контур объекта, где хранится геометрическое описание объекта (его границы), причём таких наблюдений у объекта может быть несколько, за счёт чего можно организовать мониторинг объекта с меняющимися во времени границами;
- поддержано хранение различных описаний объектов (например, периода наблюдения) в атрибутах объекта;
- реализован модуль визуализации значений рассчитанных показателей в границах объектов в картографических веб-интерфейсах согласно легенде показателя;

- поддержано заведение исследовательских проектов с возможностью разграничения доступа к проектам в зависимости от авторизации пользователя.

В третьей главе описаны функциональные блоки для построения программной инфраструктуры ДОМ. В частности, предложены алгоритмы, реализованные в блоке автоматического расчёта показателей объектов: алгоритм, позволяющий минимизировать временные затраты на обработку спутниковой информации с учётом особенностей структуры хранения спутниковых данных и пространственного распределения наблюдаемых объектов (положение 2, выносимое на защиту), и алгоритм фильтрации зашумлённых данных в рядах рассчитанных показателей (положение 3, выносимое на защиту). В результате реализации функциональных блоков была создана программная инфраструктура ДОМ (положение 4, выносимое на защиту).

В силу того, что программная инфраструктура ДОМ должна обеспечивать мониторинг разных объектов в рамках исследовательских проектов (причём наблюдения объектов должны быть организованы исключительно однотипным образом) необходимо создать модуль управления правилами наблюдения в зависимости от конкретного объекта. Для организации работы с такой информацией была создана унифицированная схема описания показателей, реализованная на основе специализированного каталога показателей. Описание структуры каталога, список требований к его инструментам и правила работы с показателями приведены в *первом параграфе* главы.

Второй параграф посвящён описанию автоматического сервиса аннотирования (расчёта значений показателей объектов), предназначенного для организации непосредственно мониторинга объектов по спутниковым данным. При этом особое внимание уделено модулю, запрашивающему исходные данные для расчёта показателей из распределённых архивов ЦКП «ИКИ-Мониторинг», так как эти запросы можно сформировать двумя способами: для каждого объекта по отдельности или для группы объектов, попадающих на одну единицу хранения данных (фрагмент). В случае группировки объектов запрашиваются элементы спутниковых сцен для области интереса, накрывающей все сгруппированные объекты, и если один из объектов находится в отдалении от других, большая часть запрошенных данных оказывается фактически лишней. Поэтому в общем случае при произвольной конфигурации объектов необходимо найти их разбиение по группам для общих запросов, которое давало бы минимальное суммарное время выполнения. В настоящей работе был предложен алгоритм, осуществляющий быстрое разбиение объектов по группам для достижения минимального времени их аннотирования по данным конкретной спутниковой системы. Данный алгоритм включает в себя два основных шага:

1. В каждой группе объектов, попадающих на одну единицу хранения данных, выделить объекты, обработку которых целесообразнее проводить отдельно, если такие имеются.
2. Оставшуюся группу объектов разделить на кластеры для минимизации времени выполнения запросов.

Для реализации шага 1 алгоритма в данной работе предлагается субоптимальный итерационный жадный алгоритм оптимизации поиска наибольшей области интереса, который позволяет найти локально оптимальное решение задачи на каждом шаге. В общем виде алгоритм формулируется следующим образом:

- пусть $A_j = \{a_1, a_2, \dots, a_n\} \in F_j$ — рассматриваемое подмножество объектов, а F_j — множество всех объектов, лежащих на j -м фрагменте, тогда необходимо сравнить время выполнения общего запроса по всем объектам из A — T_i

и суммы времени выполнения запросов для каждого объекта из A по отдельности — $Tsum$;

- в случае если $Tu > Tsum$, т. е. выполнение общего запроса для A невыгодно, то формируется множество B такое, что

$$B \subset A, \quad |B| = |A| - 1 \quad \text{и} \quad B = \arg \min_i t(A \setminus a_i), \quad (1)$$

где $t(Q)$ — оценка времени выполнения общего запроса для множества объектов Q . То есть B является таким подмножеством A , что в нём ровно на один элемент меньше, и из всех таких подмножеств время выполнения общего запроса для него минимально. Далее алгоритм повторяется снова для полученного B , пока его мощность больше 1;

- в случае если $Tu \leq Tsum$, выполняется общий запрос для A .

Для шага 2 предложенного алгоритма применяется алгоритм кластеризации `AgglomerativeClustering` библиотеки Python `sklearn`, рекурсивно объединяющий пару кластеров выборочных данных, используя расстояние связи (<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AgglomerativeClustering.html>).

Чтобы проводить оценку времени, необходимо найти экспериментальную зависимость времени выполнения запроса от площади интереса (в пикселях) для схемы хранения данных конкретной спутниковой системы в архивах ЦКП «ИКИ-Мониторинг».

С целью оценки эффективности предложенного алгоритма было проведено исследование на данных прибора MSI (*англ.* Multispectral Instrument) (Sentinel-2), имеющих фиксированное разбиение на фрагменты. Была получена экспериментальная зависимость нормированного времени обработки запросов для областей интересов разного размера, попадающих на один фрагмент данных, которая подтвердила гипотезу о линейности зависимости времени выполнения запросов от площади областей интереса:

$$t(S) = t_1 S + t_0 = 10^{-8} S + 0,1, \quad (2)$$

где S — площадь области интереса (количество пикселей); t — нормированное время выполнения запроса; t_0 и t_1 — коэффициенты линейной регрессии. Нормировка проводилась делением на время выполнения запроса для области максимального размера.

Результаты экспериментов на десятках фрагментов данных показали, что применение алгоритма позволяет существенно сократить время обработки. В случае если объекты локализуются по областям, при этом отдельные объекты расположены на больших расстояниях от этих локализаций, применение алгоритма позволяет уменьшить время обработки до 3,5 раз.

Третий параграф посвящён задаче построения временных рядов показателей объектов, очищенных от шумов. Под шумами понимаются значения показателей, рассчитанные по спутниковым изображениям, на которых над объектами явно наблюдается облачность. При этом использование стандартных средств фильтрации облачности, а именно попиксельного анализа маски облачности не всегда эффективно из-за недостаточной точности масок.

В работе предлагается алгоритм фильтрации спутниковых данных для задачи объектного мониторинга, основанный на анализе доли площади, занятой облачностью (ПЗО) в окрестности объекта на конкретном спутниковом изображении. Данный алгоритм заключается в следующей последовательности шагов:

- для исследуемого объекта и рассматриваемой сцены спутниковых данных формируется маска облачности (это может быть стандартная маска облач-

ности, предоставляемая поставщиками спутниковых данных, или результат детектирования облачности в пикселе любым другим способом);

- по сформированной маске рассчитывается ПЗО в окрестности анализируемого объекта;
- если ПЗО превышает значение порога, то считается, что выбранная сцена непригодна для дальнейшего расчёта показателей объекта, и она не используется для формирования ряда характеристик наблюдаемого объекта.

При этом оптимальным значением порога ПЗО становится значение, при котором минимизирована интегральная ошибка (ИО), включающая в себя ошибки как первого, так и второго рода. В рассматриваемой задаче к ошибкам первого рода относятся облачные сцены, оставшиеся после фильтрации в ряду наблюдений, а к ошибкам второго рода — пропуски (невключения в ряд) сцен, на которых непосредственно над объектом нет облачности. Оптимальное значение порога рассчитывается по формуле (3):

$$\varepsilon^* = \arg \min_{\varepsilon} \left[\sum_{i=1}^n \Phi_+ (\varepsilon_{a_i} - \varepsilon) + \sum_{j=1}^m \Phi_- (\varepsilon_{b_j} - \varepsilon) \right] : a_i \in A, b_j \in B, \quad (3)$$

где A — множество облачных изображений; B — множество безоблачных изображений; $n = |A|$ — количество облачных изображений; $m = |B|$ — количество безоблачных изображений; Φ_+ и Φ_- — функции потерь (ФП) при классификации; ε_{a_i} — ПЗО на i -м облачном изображении; ε_{b_j} — ПЗО на j -м безоблачном изображении; ε^* — оптимальное значение порога ПЗО.

Для пороговой ФП формулу (3) можно свести к виду:

$$\varepsilon^* = \arg \min_{\varepsilon} \left[\frac{N_e(\varepsilon) + N_m(\varepsilon)}{N} \right], \quad (4)$$

где N_e — количество лишних (extra) некачественных изображений, попавших в выборку после фильтрации; N_m — количество пропущенных (missed) в результате фильтрации качественных изображений; N — размер выборки.

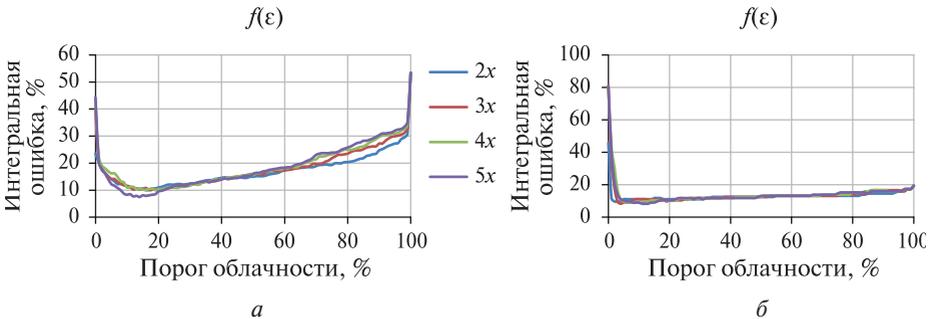


Рис. 2. Пример функции ИО, сформированной для объектов, расположенных в зонах, для которых число облачных дней в году: a — 50%; b — менее 20%. 2–5х — размеры окрестностей объекта. Ряды сформированы на основе анализа данных прибора MSI

При этом проведённый анализ различных наборов данных для ряда объектов, расположенных на географически разных территориях, показал, что для всех объектов независимо от размера окрестности наблюдается диапазон ПЗО (10–20 %), при котором обеспечиваются низкие значения ИО (рис. 2).

В случаях, когда стандартные маски имеют значительные ошибки и не могут использоваться для расчёта ПЗО, применяется специально реализованный алгоритм детектирования облаков, основанный на значениях коэффициента спектральной яркости (КСЯ) в голубом канале и нормализованного разностного индекса снега NDSI (*англ.* Normalized Difference Snow Index). Если значение КСЯ в голубом канале ниже выбранного порога, то пиксель относится к безоблачному, в противном случае он считается «потенциально облачным». При этом если в «потенциально облачном» пикселе значение индекса снега NDSI не превышает заданный порог, то пиксель считается облачным, иначе он относится к снежному покрову.

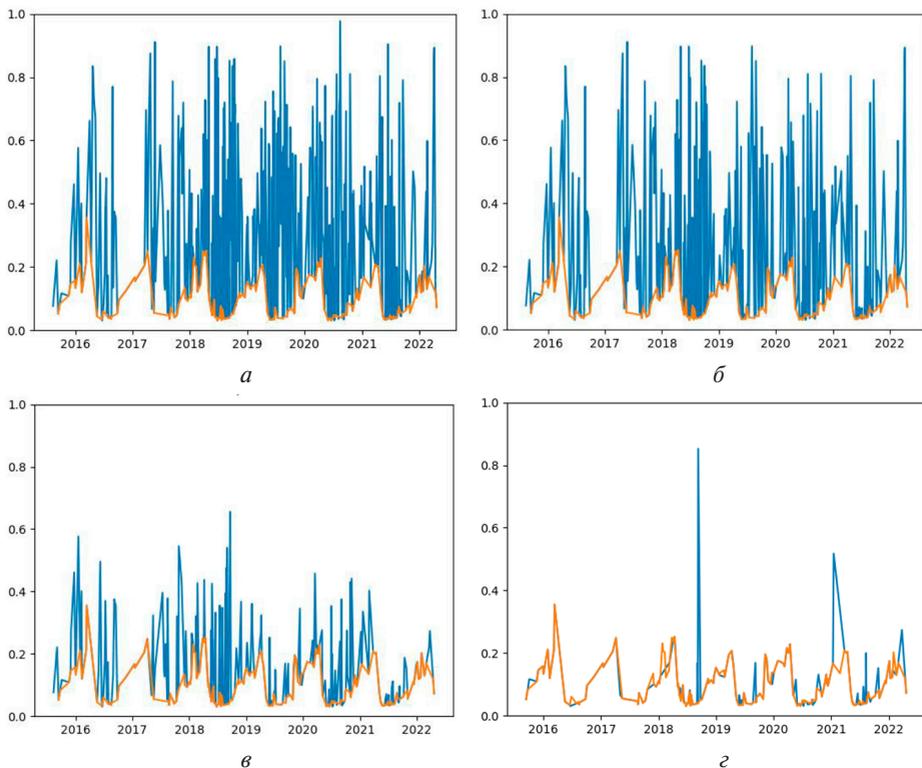


Рис. 3. Ряды со значениями конечного показателя для объекта (оранжевый цвет — очищенные вручную ряды, синий цвет — ряды на различных этапах фильтрации): *а* — до фильтрации; *б* — после «маскирования» по маске облачности; *в* — после фильтрации выбросов для «маскированных» данных; *г* — после фильтрации «маскированных» данных по порогу облачности с помощью алгоритма по NDSI

На рис. 3 приведён пример применения предложенного алгоритма фильтрации для одного из исследуемых объектов. В качестве размера окрестности было приня-

то четыре величины объекта по каждой оси, а значение порога облачности выбрано равным 15 %. На рисунках представлены очищенные вручную (тестовые) ряды характеристик объектов (оранжевый цвет), сформированные на основе данных Sentinel-2 без атмосферной коррекции, совместно с рядами характеристик, полученных на различных этапах фильтрации (синий цвет).

Фильтрация выбросов проводилась путём расчёта верхней границы доверительного интервала значений показателя внутри скользящего окна (см. формулу (5)) и затем отбрасыванием всех значений показателя, превосходящих эту границу.

$$l_{top} = m + t_b \sigma_m, \quad (5)$$

где l_{top} — верхняя граница доверительного интервала; m — оценка математического ожидания; $t_b = f(\beta/2)$ — значение функции, обратной функции Лапласа, где β — доверительная вероятность ($t_b = 1,28$ для $\beta = 0,8$; $t_b = 1,96$ для $\beta = 0,95$); $\sigma_m = \sqrt{D/N}$ — среднеквадратичное отклонение оценки математического ожидания, где D — оценка дисперсии; N — размер выборки (ширина скользящего окна).

В таблице приведена ИО в процентах, полученная в результате применения двух способов фильтрации для трёх объектов, расположенных в зонах с разным количеством облачных дней в году.

Интегральная ошибка, рассчитанная по результатам использования различных видов фильтрации

Объект	Кол-во изображений	ИО фильтрация выбросов, %	ИО фильтрация по порогу ПЗО, %
1	344	44	17
2	394	29	11
3	129	28	12

В *четвёртом параграфе* рассматривается создание инструментов, позволяющих работать с данными ДОМ в интерфейсах анализа: картографическом веб-интерфейсе и интерфейсе графиков, функционирующих в рамках различных тематических ИС, развиваемых в ИКИ РАН (рис. 4). Данные интерфейсы нацелены решить две задачи: управление объектами (ввод, редактирование), а также непосредственную оценку их состояния. Инструменты работы с описываемыми интерфейсами разрабатывались с использованием программных модулей, разработанных в ИКИ РАН, и принципов унификации подходов и простоты масштабируемости.

В *пятом параграфе* описан сценарий взаимодействия с программной инфраструктурой ДОМ. Администратор заводит новый тип объектов в служебном интерфейсе каталога показателей и определяет, какие показатели и по данным каких спутниковых систем необходимо рассчитывать для мониторинга такого типа объектов. В картографическом веб-интерфейсе создаётся новая тематическая вкладка, либо поддерживается работа с данным типом объектов на основе уже имеющейся вкладки, если в рамках исследовательского проекта существует необходимость осуществлять мониторинг нескольких типов объектов. Для создания нового объекта пользователю, работающему в картографическом веб-интерфейсе, достаточно оконтурить объект на карте, а при занесении объекта в БД ДОМ — указать его тип и временной период наблюдения за ним. Далее новый объект подхватывается сервисом аннотирования и спустя некоторое время расчёта пользователь имеет возможность визуализировать рассчитанные значения на карте, выбрав объект из списка и соот-

ветствующую сцену спутниковых данных, анализировать динамику рассчитанных характеристик объекта в модуле анализа временных рядов и выгружать значения показателей для объекта в табличные интерфейсы для дальнейшего анализа.

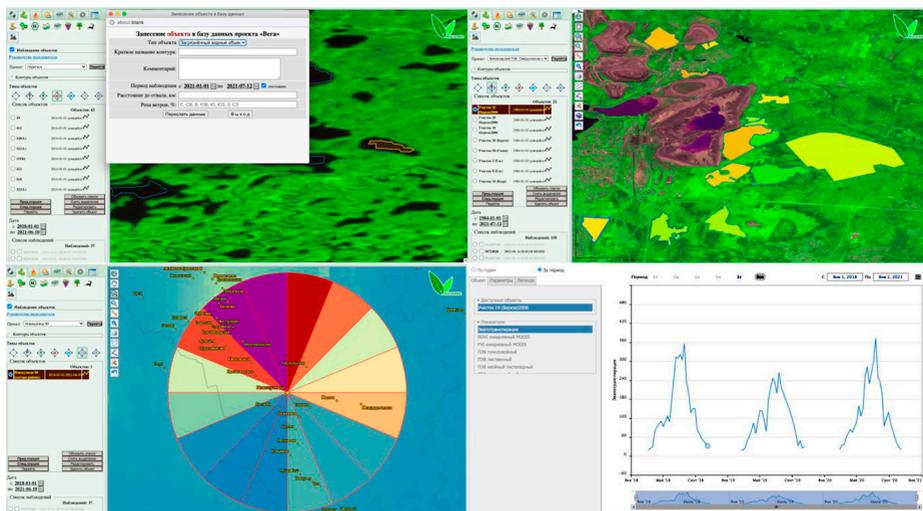


Рис. 4. Специализированные интерфейсы для работы с данными ДОМ

В четвёртой главе представлены примеры применения разработанной программной инфраструктуры для решения реальных задач.

В первом параграфе кратко изложено применение программной инфраструктуры ДОМ для комплексного спутникового мониторинга водных объектов на примере водных ресурсов Узбекистана и Центральной Азии в рамках специально созданной для этой цели ИС EcoSatMS (<http://suvo.geosmis.ru/>). Контроль водных ресурсов на основе использования только данных наземных гидропостов (ГП), которые управляются службами различных государств, сопряжён с трудностями в получении одноуровневой информации на протяжении всего течения трансграничных рек. Поэтому получение объективной информации о состоянии рек региона на основе данных спутниковых наблюдений представляется актуальной задачей. В данном проекте предлагается измерять показатель, который бы отражал водность реки на исследуемом участке и расчёт которого по спутниковым данным возможно автоматизировать. Таким показателем стала площадь, занятая водой в границах некоторого полигона, расположенного на участке реки, — «космического» гидропоста (КГП). Расчёт площади осуществляется на основе модифицированного нормализованного разностного водного индекса MNDWI по данным спутниковых приборов MSI (Sentinel-2), OLI (англ. Operational Land Imager) (Landsat-8), ETM+ (англ. Enhanced Thematic Mapper) (Landsat-7), TM (англ. Thematic Mapper) (Landsat-4, -5), при этом пиксели с положительным значением индекса принимаются за водную поверхность. Применение инструментов разработанной программной инфраструктуры позволило пользователям ИС добавлять новые КГП, полностью автоматизировать процесс расчёта площади водного зеркала в их границах и, как следствие, получать порядка 100 наблюдений в год для каждого КГП без существенных трудозатрат. Ряд экспериментов позволил установить линейную зависимость рассчитываемой площади

водного зеркала от суточных стоков на наземных ГП. В настоящее время с использованием созданной программной инфраструктуры ведутся исследования по построению сети КГП путём последовательной связи КГП друг с другом и с наземным ГП выше или ниже по течению реки, чтобы иметь возможность восстанавливать суточные стоки на некотором отдалении от наземных ГП.

Во *втором параграфе* кратко изложены подходы к использованию программной инфраструктуры ДОМ для мониторинга вулканической активности Камчатки и Курил в ИС VolSatView (<http://kamchatka.volcanoes.smislab.ru/>). Инструменты программной инфраструктуры ДОМ используются для наблюдений за кратерами действующих вулканов с целью мониторинга и изучения температурных аномалий и потенциальных извержений, возникающих на этих участках. Такой подход позволил обеспечить быстрый сбор и извлечение необходимой информации, получаемой от различных спутниковых систем, как низкоорбитальных, так и геостационарных. Поскольку спутниковые системы имеют разные углы визирования, один и тот же объект на поверхности Земли имеет разные координаты на снимках геостационарного и низкоорбитального спутников. В особенности это актуально для объектов, которые находятся на большой высоте над уровнем моря, в частности, для кратеров вулканов. Поэтому инструменты программной инфраструктуры, используемые в картографическом веб-интерфейсе, позволяют добавлять несколько контуров для одного исследуемого участка. Расчёт показателей на основе спутниковых данных осуществляется в границах контура, введённого с использованием снимка соответствующей спутниковой системы.

В *третьем параграфе* описывается использование программной инфраструктуры ДОМ для организации долговременного постоянного мониторинга окружающей среды в районах расположения крупных источников антропогенных загрязнений (КИАЗ) на примере промышленных предприятий, добывающих руду. Это исследование ведётся на базе уникальной научной установки (УНУ) «Vega-Science» (<http://sci-vega.ru/>). Следует отметить, что влияние КИАЗ распространяется на значительные территории (сотни квадратных километров), а также имеет долгосрочный характер, его последствия проявляются на больших временных промежутках. Поэтому оценить такое влияние без использования дистанционных методов практически невозможно. Архивы спутниковых наблюдений, накопленные за десятилетия, могут быть эффективно задействованы для мониторинга таких районов и анализа возникающих в них долговременных изменений состояния окружающей среды. На данный момент с помощью программной инфраструктуры ДОМ в УНУ Vega-Science ведётся постоянный дистанционный мониторинг 60 районов, в которых расположены КИАЗ. В основу мониторинга таких районов лёг подход сравнительного анализа потенциально загрязнённых участков, расположенных в зоне влияния КИАЗ, совместно с потенциально чистыми (контрольными) участками, расположенными вне зоны влияния КИАЗ. В силу комплексности воздействия КИАЗ на окружающую среду в данной задаче изучается их влияние как на растительность, так и на водные объекты. Для каждого района исследования (промышленного предприятия) выбираются потенциально загрязнённые и контрольные участки однородного леса и воды. Совместный анализ полученных с помощью инструментов созданной программной инфраструктуры долговременных рядов значений вегетационных индексов для объектов с однородным лесом на потенциально загрязнённых и контрольных территориях позволил выявить деградацию растительности в зонах влияния КИАЗ. Оценка значений водных индексов позволила выявить аномальные состояния некоторых водоёмов вблизи хвостохранилищ и других источников загрязнений.

Четвёртый параграф посвящён применению программной инфраструктуры ДОМ для изучения концентрации газовых составляющих в атмосфере над террито-

риями крупных городов на примере исследования уровня загрязнённости городов диоксидом азота. Данная работа ведётся на основе УНУ «Вега-Science». В крупных городах основным источником диоксида азота становится автотранспорт, и по концентрации данного газа можно в целом судить о качестве городского воздуха, которое напрямую влияет на здоровье и продолжительность жизни населения городов. С использованием инструментов программной инфраструктуры ДОМ были описаны крупные города России, затем осуществлён расчёт среднего значения концентрации диоксида азота внутри оконтуренных территорий городов по всем еженедельным, ежемесячным, ежегодным и многолетним композитным изображениям, построенным на основе данных прибора TROPOMI (*англ.* Tropospheric Monitoring Instrument) (Sentinel-5P). В результате был составлен рейтинг загрязнённости городов России диоксидом азота, выявлены сезонные и межгодовые тренды значений концентрации, в том числе тренды, связанные с введением ограничительных мер на передвижение и антропогенную деятельность во время распространения COVID-19. Следует отметить, что программная инфраструктура ДОМ позволяет организовывать постоянный дистанционный мониторинг качества воздуха любого города или района. Ограничением представляется лишь пространственное разрешение используемых для расчёта спутниковых данных относительно размеров исследуемой территории.

В заключении представлены основные результаты работы.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. На основе системного анализа исследовательских задач, возникающих при изучении природных и антропогенных объектов, а также тенденций развития современных систем ДЗЗ, разработана концептуальная модель данных для организации дистанционного спутникового объектного мониторинга.
2. Разработана архитектура взаимодействия программ и систем сбора, обработки, архивации и представления спутниковых данных для организации дистанционного мониторинга различных объектов, в том числе созданы программные инструменты, реализующие данную архитектуру.
3. Создана программная инфраструктура распределённой потоковой обработки спутниковых данных для автоматического формирования статистики по объектам и ведения (пополнения) БД объектного мониторинга. Для созданной программной инфраструктуры разработан алгоритм, позволяющий минимизировать временные затраты на расчёты характеристик объектов с учётом особенностей структуры хранения спутниковых данных и пространственного распределения наблюдаемых объектов.
4. Разработан алгоритм фильтрации зашумлённых спутниковых данных, основанный на количественном анализе процента облачности на спутниковых изображениях в районах расположения конкретных объектов, который позволяет получать очищенные ряды данных в условиях низкого качества фильтрации облачности стандартными алгоритмами.
5. Разработаны специализированные интерфейсы, позволяющие управлять БД объектного мониторинга, визуализировать и анализировать результаты обработки спутниковой информации.
6. На основе разработанных в работе архитектуры, инструментов и алгоритмов создана программная инфраструктура для организации дистанционного спутникового объектного мониторинга.

7. Созданная программная инфраструктура показала свою работоспособность и эффективность при реализации блоков и подсистем в различных научных и прикладных информационных системах дистанционного мониторинга. Таким образом, цель работы достигнута и поставленные в ней задачи полностью решены.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в рецензируемых изданиях, включённых в перечень ВАК:

1. **Константинова А. М.** Алгоритм автоматической фильтрации облачных данных для решения задач объектного дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 4. С. 88–99. DOI: 10.21046/2070-7401-2022-19-4-88-99. (Scopus)
2. **Константинова А. М.**, Балашов И. В., Кашницкий А. В., Лупян Е. А. Унифицированная технология дистанционного мониторинга природных и антропогенных объектов // Современные проблемы дистанционного зондирования Земли из космоса. 2021. Т. 18. № 4. С. 41–52. DOI: 10.21046/2070-7401-2021-18-4-41-52. (Scopus)
3. **Константинова А. М.**, Лупян Е. А. Анализ последствий прорыва дамбы Сардобинского водохранилища 1 мая 2020 г. // Современные проблемы дистанционного зондирования Земли из космоса. 2020. Т. 17. № 3. С. 261–266. DOI: 10.21046/2070-7401-2020-17-3-261-266. (Scopus)
4. Loupian E. A., Bourtsev M. A., Proshin A. A., Kashnitskii A. V., Balashov I. V., Bartalev S. A., **Konstantinova A. M.**, Kobets D. A., Radchenko M. V., Tolpin V. A., Uvarov I. A. Usage Experience and Capabilities of the VEGA-Science System // Remote Sensing. 2022. V. 14. No. 1. Art. No. 77. DOI: doi.org/10.3390/rs14010077. (WoS, Scopus)
5. Мухамеджанов И. Д., **Константинова А. М.**, Лупян Е. А., Умирзаков Г. У. Оценка возможностей спутникового мониторинга динамики речного стока на примере анализа состояния реки Амударья // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 1. С. 87–103. DOI: 10.21046/2070-7401-2022-19-1-87-103. (Scopus)
6. Лупян Е. А., **Константинова А. М.**, Кашницкий А. В., Ермаков Д. М., Саворский В. П., Панова О. Ю., Бриль А. А. Возможности организации долгосрочного дистанционного мониторинга крупных источников антропогенных загрязнений для оценки их влияния на окружающую среду // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 1. С. 193–213. DOI: 10.21046/2070-7401-2022-19-1-193-213. (Scopus)
7. Гирина О. А., **Константинова А. М.**, Крамарева Л. С., Сорокин А. А., Маневич А. Г., Мельников Д. В., Романова И. М., Уваров И. А., Мальковский С. И., Королев С. П. Эксплозивное событие 19 апреля 2022 г. вулкана Карымский (Камчатка) по спутниковым данным // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 2. С. 255–260. DOI: 10.21046/2070-7401-2022-19-2-255-260. (Scopus)
8. Лупян Е. А., Прошин А. А., Бурцев М. А., Кашницкий А. В., Балашов И. В., Барталев С. А., Бриль А. А., Егоров В. А., Жарко В. О., **Константинова А. М.**, Кобец Д. А., Мазуров А. А., Марченков В. В., Матвеев А. М., Миклашевич Т. С., Плотников Д. Е., Радченко М. В., Стыценко Ф. В., Сычугов И. Г., Толпин В. А., Уваров И. А., Хвостиков С. А., Ховратович Т. С. Система «Vega-Science»: особенности построения, основные возможности и опыт использования // Со-

- временные проблемы дистанционного зондирования Земли из космоса. 2021. Т. 18. № 6. С. 9–31. DOI: 10.21046/2070-7401-2021-18-6-9-31. (Scopus)
9. Лупян Е.А., **Константинова А.М.**, Балашов И.В., Кашницкий А.В., Саворский В.П., Панова О.Ю. Разработка системы анализа состояния окружающей среды в зонах расположения крупных промышленных объектов, хвостохранилищ и отвалов // Современные проблемы дистанционного зондирования Земли из космоса. 2020. Т. 17. № 7. С. 243–261. DOI: 10.21046/2070-7401-2020-17-7-243-261. (Scopus)
 10. Гирина О.А., Лупян Е.А., Мельников Д.В., Кашницкий А.В., Уваров И.А., Бриль А.А., **Константинова А.М.**, Бурцев М.А., Маневич А.Г., Гордеев Е.И., Крамарева Л.С., Сорокин А.А., Мальковский С.И., Королев С.П. Создание и развитие информационной системы «Дистанционный мониторинг активности вулканов Камчатки и Курил» // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 3. С. 249–265. DOI: 10.21046/2070-7401-2019-16-3-249-265. (Scopus)
 11. Саворский В.П., Лупян Е.А., Горный В.И., Ермаков Д.М., Панова О.Ю., **Константинова А.М.** Методы и инструменты анализа данных ДЗЗ для выявления изменений растительного покрова, вызванных техногенными отходами и отвалами // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 6. С. 31247. DOI: 10.21046/2070-7401-2019-16-6-31-47. (Scopus)
 12. Kashnitskii A. V., Loupian 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. Iss. 1. P. 84–88. DOI: 10.1134/S1024856017010080. (WoS, Scopus)

Константинова Анна Михайловна

**АЛГОРИТМЫ И ПРОГРАММНАЯ ИНФРАСТРУКТУРА
ДЛЯ РЕШЕНИЯ ЗАДАЧ СПУТНИКОВОГО
ОБЪЕКТНОГО МОНИТОРИНГА**

Автореферат

диссертации на соискание учёной степени
кандидата технических наук

055(02)2

ИКИ РАН
Москва, 117997, Профсоюзная ул., 84/32

Подписано к печати 24.01.2023

Заказ 4202

Формат 60×84/16

Тираж 100

0,93 усл.-печ. л.

