

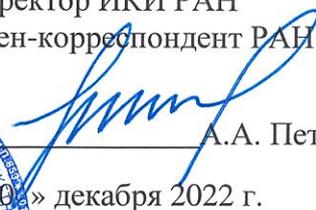
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ КОСМИЧЕСКИХ ИССЛЕДОВАНИЙ
РОССИЙСКОЙ АКАДЕМИИ НАУК
(ИКИ РАН)

УДК 528.85/87(15)

Номер государственной регистрации 122042500019-6

УТВЕРЖДАЮ
Директор ИКИ РАН
Член-корреспондент РАН




А.А. Петрукович
20 » декабря 2022 г.

О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

Большие данные в космических исследованиях: астрофизика, солнечная система,
геосфера

по теме:

КОСМОС-Д

FFWG-2022-0012

(промежуточный, этап 2022/1)

Научный руководитель

канд .техн. наук



А.А. Прошин

Москва

2022

СПИСОК ИСПОЛНИТЕЛЕЙ

СПИСОК ИСПОЛНИТЕЛЕЙ

Руководитель темы, старший научный сотрудник, зав. лаб., к.т.н.	 15.12.2022	А.А. Прошин (раздел 1-3, введение, заключение)
	Подпись, дата	
Ответственный исполнитель, ведущий научный сотрудник д.ф.-м.н.	 15.12.2022	Р.Ю. Лукьянова (раздел 1)
	Подпись, дата	
Ответственный исполнитель, старший научный сотрудник к.ф.-м.н.	 15.12.2022	А.В. Мешеряков (раздел 2)
	Подпись, дата	
Ответственный исполнитель, старший научный сотрудник, зав. Лаб., к.т.н.	 15.12.2022	А.А. Прошин (раздел 3)
	Подпись, дата	

Исполнители отдела 54

Исполнители отдела 54

Зав. лабораторией


15.12.22 г.

Подпись, дата

А.Т. Янаков
(раздел 1)

старший научный
сотрудник,


15.12.22 г.

Подпись, дата

А.М. Мерзлый
(раздел 1)

научный сотрудник, к.н.

15.12.22 г. 

Подпись, дата

В.А. Ожередов
(раздел 1)

ведущий инженер

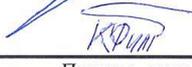
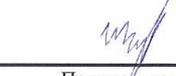
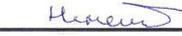

15.12.22 г.

Подпись, дата

О.В. Никифоров
(раздел 1)

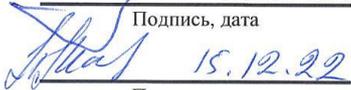
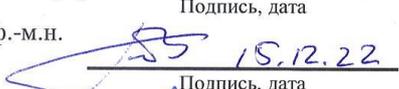
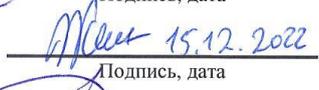
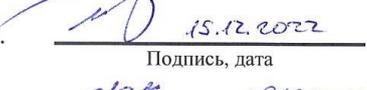
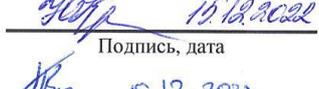
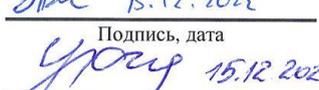
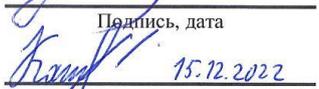
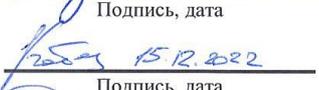
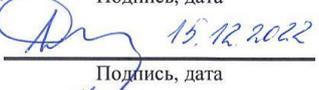
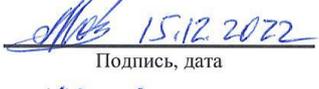
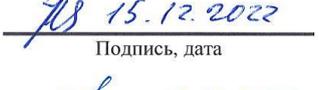
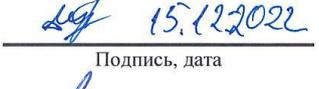
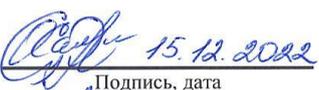
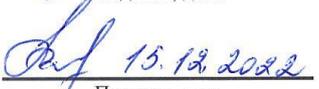
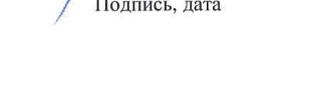
Исполнители отдела 52

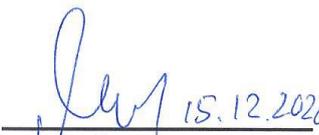
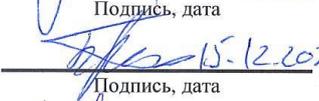
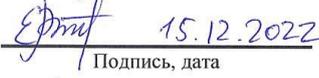
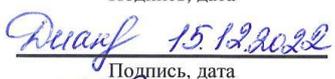
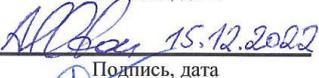
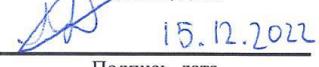
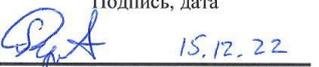
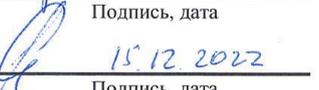
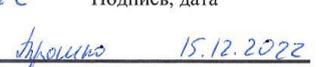
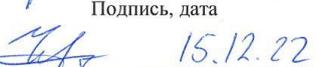
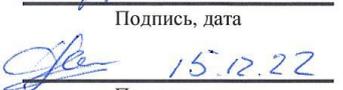
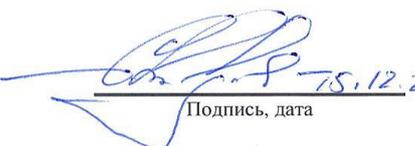
Исполнители отдела 52

старший научный сотрудник, к.ф.-м.н.	 _____ Подпись, дата	Р.А. Буренин (раздел 2)
научный сотрудник, к.ф.-м.н.	 _____ Подпись, дата	Е.В. Филиппова (раздел 2)
научный сотрудник, к.ф.-м.н.	 _____ Подпись, дата	Г.А. Хорунжев (раздел 2)
младший научный сотрудник	 _____ Подпись, дата	И.А. Зазнобин (раздел 2)
младший научный сотрудник	 _____ Подпись, дата	А.Р. Ляпин (раздел 2)
инженер	 _____ Подпись, дата	М.И. Бельведерский (раздел 2)
старший лаборант	 _____ Подпись, дата	Ф.Д. Губанов
старший лаборант	 _____ Подпись, дата	Н.В. Мальшева (раздел 2)
старший лаборант	 _____ Подпись, дата	А.А. Немешаева (раздел 2)

Исполнители отдела 56

Исполнители отдела 56

Зав. отделом, д.т.н.	 15.12.22 Подпись, дата	Е.А. Лупян (раздел 3)
ведущий научный сотрудник, к.ф.-м.н.	 15.12.22 Подпись, дата	А.А. Мазуров (раздел 3)
Зав. лабораторией, к.ф.-м.н.	 15.12.22 Подпись, дата	В.Н. Пырклов (раздел 3)
старший научный сотрудник, зав. лаб., к.т.н.	 15.12.2022 Подпись, дата	И.В. Балашов (раздел 3)
старший научный сотрудник, зав. лаб., к.т.н.	 15.12.2022 Подпись, дата	М.А. Бурцев (раздел 3)
старший научный сотрудник, к.ф.-м.н.	 15.12.2022 Подпись, дата	Ю.С. Крашенинникова (раздел 3)
старший научный сотрудник, к.т.н.	 15.12.2022 Подпись, дата	В.А. Толпин (раздел 3)
старший научный сотрудник, к.т.н.	 15.12.2022 Подпись, дата	И.А. Уваров (раздел 3)
старший научный сотрудник, к.т.н.	 15.12.2022 Подпись, дата	А.В. Кашницкий (раздел 3)
научный сотрудник, к.т.н.	 15.12.2022 Подпись, дата	Д.А. Кобец (раздел 3)
главный специалист	 15.12.2022 Подпись, дата	А.Ю. Дегай (раздел 3)
главный специалист, к.ф.-м.н.	 15.12.2022 Подпись, дата	А.А. Козочкина (раздел 3)
ведущий инженер	 15.12.2022 Подпись, дата	Ю.В. Артамонова (раздел 3)
ведущий инженер	 15.12.2022 Подпись, дата	А.М. Матвеев (раздел 3)
ведущий инженер	 15.12.2022 Подпись, дата	М.В. Радченко (раздел 3)
ведущий инженер	 15.12.2022 Подпись, дата	О.А. Суднева (раздел 3)
ведущий специалист	 15.12.2022 Подпись, дата	Г.С. Самиуллина (раздел 3)

ведущий конструктор	 15.12.2022	В.В. Марченков (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.2022	А.А. Бриль (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.2022	Е.Е. Волкова (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.2022	П.В. Денисов (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.2022	Д.С. Дианова (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.2022	А.М. Константинова (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.2022	Д.В. Лозин (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.22	Б.П. Руткевич (раздел 3)
инженер	Подпись, дата	
инженер, к.г.н.	 15.12.2022	К.С. Сенько (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.2022	К.А. Трошко (раздел 3)
инженер	Подпись, дата	
инженер	 15.12.22	В.Н. Черных (раздел 3)
инженер	Подпись, дата	
старший лаборант	 15.12.22	Д.А. Юдин (раздел 3)
старший лаборант	Подпись, дата	
нормоконтролер	 15.12.22	Е.А. Лупян
нормоконтролер	Подпись, дата	

РЕФЕРАТ

Отчет 75 стр., 31 рис., 38 источников.

НАСТОЯЩИЙ ОТЧЕТ ПОСВЯЩЕН ОПИСАНИЮ ОСНОВНЫХ РЕЗУЛЬТАТОВ РАБОТ, ПОЛУЧЕННЫХ В 2022 ГОДУ В РАМКАХ ТЕМЫ "КОСМОС-Д" (FFWG-2022-0012), ВЫПОЛНЯЮЩЕЙСЯ В РАМКАХ ГОСУДАРСТВЕННОГО ЗАДАНИЯ ИНСТИТУТОМ КОСМИЧЕСКИХ ИССЛЕДОВАНИЙ РОССИЙСКОЙ АКАДЕМИИ НАУК.

Тема «Космос-Д» «Большие данные в космических исследованиях: астрофизика, солнечная система, геосфера» сформирована для развития новой парадигмы «больших данных» в науках о космосе и получения качественных новых результатов на этой основе. Тема впервые сформирована в 2021 г в рамках программы Министерства высшего образования и науки по созданию новых лабораторий, на основе задела, созданного в других «профильных» темах государственного задания ИКИ РАН.

В ОТЧЕТЕ ПРИВОДЯТСЯ ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТ, ЗАПЛАНИРОВАННЫХ НА 2022 ГОД ПО СЛЕДУЮЩИМ НАПРАВЛЕНИЯМ:

Развитие методов моделирования высокоширотной ионосферы; построение численных и гибридных (включающих элементы ассимиляции данных измерений ионосферно-магнитосферных параметров) моделей распределения электронной концентрации в полярной ионосфере и зависимости от изменяющихся гелио-геофизических условий

Исследование и разработка алгоритмов машинного обучения для измерения расстояний (красных смещений) астрономических объектов по данным многоволновых фотометрических обзоров неба

Разработка новых подходов и методов организации сверхбольших архивов данных спутниковых наблюдений Земли и доступа к ним для обеспечения эффективной обработки и анализа данных ДЗЗ для решения научных и прикладных задач

КЛЮЧЕВЫЕ СЛОВА

СИСТЕМЫ ПАРАЛЛЕЛЬНОЙ ОБРАБОТКИ СПУТНИКОВЫХ ДАННЫХ, СВЕРХБОЛЬШИЕ ДОЛГОВРЕМЕННЫЕ АРХИВЫ СПУТНИКОВЫХ ДАННЫХ, БЛОЧНЫЙ ДОСТУП К ДАННЫМ В АРХИВАХ, АНАЛИЗ ДОЛГОВРЕМЕННЫХ СЕРИЙ НАБЛЮДЕНИЙ, ТЕХНОЛОГИЯ UNISAT, ОБРАБОТКА СПУТНИКОВЫХ ДАННЫХ, УПРАВЛЕНИЕ ОБРАБОТКОЙ СПУТНИКОВЫХ ДАННЫХ, ПОСТРОЕНИЕ БЕЗОБЛАЧНЫХ КОМПОЗИТНЫХ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ, ДОСТУП К АРХИВАМ СПУТНИКОВЫХ ДАННЫХ, ИНФОРМАЦИОННАЯ СИСТЕМА "ИС VSV", МОНИТОРИНГ ВУЛКАНИЧЕСКОЙ АКТИВНОСТИ

ИОНОСФЕРНО-МАГНИТОСФЕРНОЕ ВЗАИМОДЕЙСТВИЕ И ЭЛЕКТРОДИНАМИКА, ВЫСОКОШИРОТНАЯ ИОНОСФЕРА, АВРОРАЛЬНЫЕ ДУГИ, ОБРАБОТКА МАССИВОВ СПУТНИКОВЫХ ДАННЫХ, ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ

МАШИННОЕ ОБУЧЕНИЕ, ГЛУБОКОЕ ОБУЧЕНИЕ, ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ, ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ, КОСМИЧЕСКАЯ ОБСЕРВАТОРИЯ "СПЕКТР-РГ", РЕНТГЕНОВСКАЯ АСТРОФИЗИКА, ОБЗОРЫ НЕБА

СОДЕРЖАНИЕ

СПИСОК ИСПОЛНИТЕЛЕЙ.....	2
РЕФЕРАТ.....	7
СОДЕРЖАНИЕ.....	8
ВВЕДЕНИЕ.....	10
ГЛАВА 1 Развитие методов моделирования высокоширотной ионосферы; построение численных и гибридных (включающих элементы ассимиляции данных измерений ионосферно-магнитосферных параметров) моделей распределения электронной концентрации в полярной ионосфере и зависимости от изменяющихся гелио-геофизических условий.....	12
Введение.....	12
1.1 Разработка подходов к построению ассимиляционных блоков для ионосферного моделирования.....	12
1.2 Параметризация и валидация моделей с использованием массивов данных спутниковых измерений и наземных измерительных комплексов.....	14
1.3 Развитие программных средств, основанных на эффективных численных схемах и пригодных для оперативного прогноза основных характеристик полярной ионосферы.....	16
1.4 Развитие методик глобальной и локальной оптической диагностики распределений интенсивности авроральных эмиссий в видимой и УФ-областях спектра для оценки динамики аврорального овала.....	17
1.5 Исследование статистических связей глобальных и локальных геомагнитных вариаций с процессами в ионосфере, магнитосфере и солнечном ветре с целью выявления закономерностей появления экстремальных событий и их прогноза.....	22
1.6 Создание деревьев принятия решений для классификации изменений состояния организма человека в связи с воздействием на него факторов космической и земной погоды на основе алгоритмов дифференциальной эволюции.....	23
1.7 Разработка методов диагностики источников и эффектов локальной ионизации в верхних слоях атмосферы Земли на основе анализа больших массивов данных радио наблюдений.....	27
1.8 Развитие методов диагностики метеорных потоков на основе анализа их ионосферных эффектов.....	31
Заключение.....	34
ГЛАВА 2 Исследование и разработка алгоритмов машинного обучения для измерения расстояний (красных смещений) астрономических объектов по данным многоволновых фотометрических обзоров неба.....	36
Введение.....	36
2.1 Исследование и разработка алгоритмов глубокого обучения для сегментации и детектирования объектов, а также для измерения их физических характеристик в данных космических и наземных обсерваторий в радио, рентгеновском и оптическом диапазонах спектра.....	37
2.2 Исследование и развитие алгоритмов машинного обучения для классификации астрономических объектов по комбинированным данным астрономических каталогов и фотометрическим изображениям неба.....	38
2.3 Исследование и разработка вероятностных моделей кросс-отождествления астрономических объектов в многоволновых обзорах неба на основе пространственной и фотометрической информации.....	38
2.4 Разработка технологий и систем для обработки и интеллектуального анализа больших астрономических данных в различных спектральных диапазонах.....	39

2.5 Исследование и разработка алгоритмов инкрементального моделирования с применением в области наблюдательной астрофизики. Исследование и разработка моделей обучения с подкреплением для решения задач комбинаторной оптимизации ..40	40
Заключение.....	40
ГЛАВА 3 Разработка новых подходов и методов организации сверхбольших архивов данных спутниковых наблюдений Земли и доступа к ним для обеспечения эффективной обработки и анализа данных ДЗЗ для решения научных и прикладных задач ..42	42
Введение	42
3.1 Разработка новых методов и алгоритмов организации динамического блочного доступа к сверхбольшим разнородным архивам данных спутниковых наблюдений Земли для повышения эффективности распределенной многопоточковой обработки информации, поступающей от различных систем ДЗЗ.....	44
3.1.1 Модернизация программной реализации технологии динамического блочного доступа.....	44
3.1.2 Разработка новых подходов к организации обработки для получения ежедневных безоблачных композитов по данным прибора КМСС	45
3.1.3 Подсистема подготовки блоков данных для проведения дальнейшей их обработки	49
3.2 Разработка подходов и методов организации динамического блочного доступа к сверхбольшим разнородным архивам данных спутниковых наблюдений Земли для повышения эффективности инструментов онлайн доступа и анализа спутниковых данных и результатов их обработки.....	51
3.2.1 Модернизация картографического веб-интерфейса, использующего динамическое формирование тайлов изображений	51
3.2.2 Модернизация модуля растеризации изображений	53
3.2.3 Профилирование времени выполнения запросов на получение тайлов изображений.....	53
3.3 Развитие с использованием созданных методов и подходов различных специализированных научных информационных систем, в том числе. ИС ВЕГА-Лес (комплексный научный мониторинг лесных ресурсов России), ИС STS (комплексный научный мониторинг различных явлений в пограничных морях России), ИС VSV (комплексный научный мониторинг вулканической активности Камчатки и Курил).....	54
3.3.1 Разработка новой системы управления комплексными задачами по обработке данных	54
3.3.2 Разработка сервисов для эффективного предоставления доступа к временным сериям данных сверхбольших распределенных архивов.....	57
3.3.3 Технология создания безоблачных композитных изображений по данным спутников серии Sentinel-2	62
Заключение.....	67
ЗАКЛЮЧЕНИЕ.....	68
ПУБЛИКАЦИИ.....	69

ВВЕДЕНИЕ

Институт космических исследований Российской академии наук проводит научно-исследовательские и опытно-конструкторские работы в соответствии с государственным заданием ЧАСТЬ 2: ГОСУДАРСТВЕННЫЕ РАБОТЫ по следующим научным направлениям:

- Фундаментальные и прикладные научные исследования в области астрофизики и радиоастрономии;
- Фундаментальные и прикладные научные исследования в области Физики космической плазмы, энергичных частиц, Солнца и солнечно-земных связей;
- Фундаментальные и прикладные научные исследования планет и малых тел Солнечной системы;
- Фундаментальные и прикладные научные исследования планеты Земля;
- Фундаментальные и прикладные научные исследования в области механики, систем управления и информатики;
- Развитие исследовательской, конструкторской, опытно-экспериментальной базы научного космического приборостроения и методов экспериментальной физики.

Эти направления НИР и ОКР соответствуют следующим направлениям фундаментальных исследований, указанными в Программе фундаментальных исследований в Российской Федерации на долгосрочный период 2021-2030 гг., утвержденной распоряжением Правительства РФ от 31 декабря 2020 г. № 3684-р

п/п	Направление фундаментальных исследований	Номер направления в «Программе»
	Пламенные процессы в геофизике и астрофизике	1.3.4.3.
	Фундаментальные проблемы распространения радиоволн, в том числе в ионосфере	1.3.6.4.
	Происхождение, строение и эволюция Вселенной	1.3.7.1.
	Физика галактик и межгалактической среды	1.3.7.2.
	Развитие методов наземной и внеатмосферной астрономии	1.3.7.6.
	Информационно-вычислительные системы и среды в науке и образовании	1.1.8
	Алгоритмы и программные системы в космическом мониторинге Земли и экологии	1.1.7.6
	Метрология и цифровизация в науках о Земле	1.5.12

Работы по теме сфокусированы на новой парадигме «больших данных». В космических исследованиях переход к новой парадигме обусловлен беспрецедентным ростом количества и разнообразия данных наблюдений. Существующие и будущие дальние космические миссии и системы наблюдения Земли из космоса становятся все более сложными, представляя высокое пространственное/временное/спектральное разрешение. Такое развитие экспериментальной базы требует адекватного усовершенствования подходов к анализу и моделированию соответствующих физических

процессов. Перспективным представляется развитие интеллектуальной автоматизированной аналитики, усовершенствование методов моделирования с применением ассимиляции и глубокого машинного обучения, использование новых подходов организации хранения и управления данными (в том числе с использованием облачных технологий и высокопроизводительных вычислительных систем для масштабируемого анализа). Также, результативными могут быть и новые применения традиционных методов анализа с учетом специфики больших данных.

В работах по теме «Космос-Д» предусматривается создание новых подходов к анализу данных космических экспериментов и мониторинга в астрофизике, солнечно-земной физике, физике солнечной системы, дистанционном зондировании Земли. Работы по этим направлениям объединяются разработкой уникальной единой системы хранения и анализа информации.

Работы по теме ведутся по следующим основным направлениям:

- Развитие методов моделирования высокоширотной ионосферы; построение численных и гибридных (включающих элементы ассимиляции данных измерений ионосферно-магнитосферных параметров) моделей распределения электронной концентрации в полярной ионосфере и зависимости от изменяющихся гелиогеофизических условий.
- Исследование и разработка алгоритмов машинного обучения для измерения расстояний (красных смещений) астрономических объектов по данным многоволновых фотометрических обзоров неба.
- Разработка новых подходов и методов организации сверхбольших архивов данных спутниковых наблюдений Земли и доступа к ним для обеспечения эффективной обработки и анализа данных ДЗЗ для решения научных и прикладных задач.

Представлению результатов, полученных в данных направлениях, посвящены соответствующие главы настоящего отчета.

ГЛАВА 1 Развитие методов моделирования высокоширотной ионосферы; построение численных и гибридных (включающих элементы ассимиляции данных измерений ионосферно-магнитосферных параметров) моделей распределения электронной концентрации в полярной ионосфере и зависимости от изменяющихся гелио-геофизических условий

Введение

На 2022 г. было запланировано выполнение комплекса работ по следующим тематическим пунктам: - Разработка подходов к построению ассимиляционных блоков для ионосферного моделирования; - Параметризация и валидация моделей с использованием массивов данных спутниковых измерений и наземных измерительных комплексов; - Развитие программных средств, основанных на эффективных численных схемах и пригодных для оперативного прогноза основных характеристик полярной ионосферы; - Развитие методик глобальной и локальной оптической диагностики распределений интенсивности авроральных эмиссий в видимой и УФ-областях спектра для оценки динамики аврорального овала; - Исследование статистических связей глобальных и локальных геомагнитных вариаций с процессами в ионосфере, магнитосфере и солнечном ветре с целью выявления закономерностей появления экстремальных событий и их прогноза; - Создание деревьев принятия решений для классификации изменений состояния организма человека в связи с воздействием на него факторов космической и земной погоды на основе алгоритмов дифференциальной эволюции; - Разработка методов диагностики источников и эффектов локальной ионизации в верхних слоях атмосферы Земли на основе анализа больших массивов данных радио наблюдений; - Развитие методов диагностики метеорных потоков на основе анализа их ионосферных эффектов.

Актуальность исследований обусловлена необходимостью изучения физических механизмов солнечно-земных связей и разработкой методов прогнозирования космической погоды, от состояния которой может критически зависеть современная высокотехнологичная инфраструктура.

В текущем году исследования соответствовали плану и были направлены на усовершенствование ионосферных моделей, создание обзорного атласа средне/мелкомасштабных авроральных структур по оптическим наблюдениям, анализ межполушарной асимметрии в распределении вариаций электронной концентрации во время ионосферных бурь, оценку временных масштабов реакции геомагнитной активности на высокоскоростные потоки солнечного ветра, разработку алгоритмов для задачи обнаружения воздействия земной и космической погоды на организм человека, выявление процессов новых факторов, которые могут влиять на положение авроральных дуг полярных сияний, оценку изменчивости интегрального влагосодержания по данным спутникового микроволнового зондирования.

Полученные в рамках работ 2022 года результаты по данному направлению изложены в следующих публикациях [1,6,8,13,32-38].

1.1 Разработка подходов к построению ассимиляционных блоков для ионосферного моделирования

Изменчивость параметров высокоширотной ионосферы и ее зависимость от внешних условий – солнечного ветра и межпланетного магнитного поля, критически

усложняют задачу построения моделей для этой области. На этапе 2022 г. подготовлен обзор моделей, используемых в настоящее время для определения структуры крупномасштабных неоднородностей высокоширотной F области. Рассмотрены модели, основанные на эмпирическом и теоретических подходах. К группе эмпирических относятся международная эталонная модель IRI в ее базовой версии и модификациях IRI-RTAM (добавлен блок ассимиляции текущих данных измерений системы ионозондов) и IRI-STORM (введено масштабирование по геомагнитным индексам), эмпирическая канадская модель арктической ионосферы E-CHAIM в модификации E-CHAIM Storm (вводится корректирующий фактор, определяемый экспериментально и зависящий не от одного, как в IRI-STORM, а от нескольких интегральных прокси геомагнитных индексов Dst, ap и AE) и E-CHAIM Precipitation (введена параметризация для высыпаний энергичных заряженных частиц), полуэмпирическая модель GDMF (состоит из фоновой медианной модели для критической частоты f_oF2 и ряда динамических и аэрономических поправок к ней). Эффективность использования теоретических физико-математических моделей в значительной степени определяется точностью учёта физико-химических процессов в уравнениях этих моделей и эффективных схемах их численного решения, а также включением сторонних блоков, часть из которых часто по сути является эмпирическими. Возможности регионального моделирования для высокоширотной ионосферы рассмотрены на примере развиваемой в рамках направления численной модели PFIM, в которой прямо учитывается конвективный дрейф плазменных трубок под действием электрических полей магнитосферной конвекции. Приведены примеры результатов моделирования. Показано, что существующие на данный момент модели непротиворечиво описывают F-область высокоширотной ионосферы. Однако все они имеют те или иные проблемы с воспроизведением реалистичной структуры разномасштабных неоднородностей и их изменения во времени, что необходимо для оперативной диагностики и прогноза.

Начаты работы по реализации процедуры ассимиляции в рамках региональной численной модели ионосферы. Ассимиляция данных направлена на обновление уже существующих распределений определенных параметров с помощью ввода прямых или косвенных измерений этих параметров. На ранних этапах развития ассимиляции в фоновых ионосферных моделях использовались итерационные методы для усвоения данных измерений электронной плотности (Ne), например, метод последовательной коррекции. Эти алгоритмы могут быть реализованы ресурсосберегающим способом даже для большого количества неизвестных значений, так как они не решают систему уравнений как целое. Но они также производят артефакты. Альтернативные методы, такие как оптимальная интерполяция, применялись для модификации фоновой модели через вектор состояния в неитеративных уравнениях, тем самым позволяя использовать ковариацию измерений. Из-за представления матрицы ассимиляции данных и необходимости ковариационных вычислений, эти методы требуют повышенной вычислительной мощности. Также можно полученные модельные профили Ne использовать в качестве фона, а измеренные профили Ne (обычно ионозонда) ассимилируются с помощью гауссовой весовой функции.

Анализ имеющихся практик показал, что применительно к F2 слою ионосферы решением является минимизация разности между измерениями ионозонда и фоновой моделью с заданным временным разрешением. В результате получается обновленный вектор фоновой модели на двумерной сетке для основных характеристик: максимальной плотности плазмы $NmF2$ и высоты максимума плотности $hmF2$. Чтобы избежать резких локальных градиентов применяется следующая процедура: - вычисляются значения фоновой модели для точек всех доступных измерений; - вычисляются $\delta(hmF2)$ и $\delta(NmF2)$ так, чтобы они минимизировали уравнение (1.1.1) для измерений ионозонда $hmF2$ и $NmF2$; - используется фоновая модель с добавкой $\delta(hmF2)$ и фоновая модель с добавкой $\delta(NmF2)$ для вычисления первой оценки:

$$\min \sum_{k=1}^N (u_{mod,k} + \delta - u_{measur,k})^2 \quad (1.1.1)$$

Как показало изучение применения ассимиляции данных измерений ионозондов, входящих в систему Global Ionospheric Radio Observatory (GIRO), в модели IRI-RTAM, этот метод корректирует коэффициенты сферического (суточного) разложения, используемые ядром IRI, для получения глобального распределения $NmF2$. Набор глобальных скорректированных коэффициентов может генерироваться каждые 15 минут и распространяться с помощью одного сервера RTAM реального времени, управляемого GIRO. Однако на практике эта процедура не инкорпорирована в свободно распространяемый код (или онлайн доступ через портал) международной модели IRI.

Метод последовательной коррекции представляется наиболее легко реализуемым в рамках ионосферных моделей, включая региональные, в частности, в модели PFIM. В этом итеративном методе оценка ошибки (u_a) на первом шаге итерации задается сглаженным фоном в точке сетки. На каждой итерации n вычисляется

$$u_{a,i}^n = u_{a,i}^{n-1} + \frac{\sum_{k=1}^N W_{ik}^n (V_k - u_{a,k}^{n-1})}{\sum_{j=1}^N W_{ij}^n + \varepsilon^2} \quad (1.1.2)$$

где N - количество измерений, V_k - k -е наблюдение, $u_{a,k}$ - $(n-1)$ -я оценка в точке наблюдения k , полученная линейной интерполяцией из соседних точек сетки; ε - оценка соотношения между дисперсией ошибок измерений и дисперсией фоновой ошибки; W_{ik} - вес для очередного обновления n . Обычно вес определяется как уменьшающийся обратно пропорционально квадрату расстояния между точкой наблюдения и точкой сетки или обратно пропорционально квадрату шага сетки. Для усвоения ионозондовых данных вес определяется как

$$W_{ij}^n = \begin{cases} e^{(dist(i,j)/2cor)-2} & l(i,j) \leq Rad_n \\ 0 & l(i,j) > Rad_n \end{cases}, \quad (1.1.3)$$

где $dist(i,j)$ - расстояние между точками сетки i и j ; cor - длина между точками сетки; Rad_n - радиус влияния одной точки сетки на соседние (оценивается как $\sim 40^\circ$).

По результатам работ 2022 г. подготовлен для печати обзор «Модели F-области высокоширотной ионосферы» (38 стр.).

1.2 Параметризация и валидация моделей с использованием массивов данных спутниковых измерений и наземных измерительных комплексов

На этапе 2022 г. в рамках работ по сравнению с экспериментальными результатами было проведено исследование ионосферной асимметрии, наблюдаемой в Арктике и Антарктике, которое предсказывается моделями. Для выявления асимметрии применялись индексы, отражающие возмущения полного электронного содержания (ПЭС). Региональные (полушарные) и глобальные индексы получены по глобальным картам GIM. Карты ПЭС ионосферы, GIM-TEC, регулярно составляются Лабораторией реактивного движения с 1994 г. (https://sideshow.jpl.nasa.gov/pub/iono_daily/), так что для анализа доступны данные за период более двух солнечных циклов. Карты GIM-TEC используются для создания карт GIM-W за период с 1994 по 2021 г. путем расчета индекса W как логарифмического отклонения мгновенного ПЭС от предшествующей 15-дневной скользящей медианы для каждой ячейки в геомагнитной системе координат (всего более 5000 ячеек глобально и по 876 ячеек в полярных областях). Использовались следующие наборы ионосферных индексов. Первый набор оценивает положительное возмущение с увеличением ПЭС относительно спокойного состояния ионосферы. В набор

включены три индекса TEC: глобальный индекс - WU , арктический индекс - WU_N (магнитные широты выше $60^\circ N$) и антарктический индекс - WU_S (магнитные широты выше $60^\circ S$). Второй набор представляет отрицательное возмущение с истощением ПЭС, а соответствующие индексы — WL , WL_N , WL_S . Третий набор - индексы WE , WE_N и WE_S представляют собой разницу между положительным и отрицательным отклонением ПЭС. С помощью индексов ионосферной изменчивости анализировалась реакция ионосферы на вариации скорости солнечного ветра и рост геомагнитной активности во время магнитных бурь.

Бури идентифицировались на основе внезапного скачка скорости солнечного ветра и пиковой амплитуды индекса $Dst < -50$ нТл. Вторым критерий основан на времени начала бури t_0 (в часах), когда экваториальный индекс Dst падает ниже -30 нТл. В общей сложности с 1996 по 2021 г. было выявлено 380 бурь. На рисунке 1.2.1 показаны параметры одного из репрезентативных событий – бури июля 2000 г. Во время главной фазы Dst индекс достиг значения -300 нТл в 00 UT 16 июля. Глобальный отклик ионосферы проиллюстрирован картами GIM-W в предбуревой период (12 UT 15 июля), во время главной фазы (18 UT 15 июля, 00 и 06 UT 16 июля), и фазы восстановления (12 UT 16 июля). Красным цветом обозначены регионы, где ионосферный индекс W положительный, а синим цветом – регионы, где преобладает отрицательный эффект. Увеличение WU_S до 7.16 наблюдается в 18 UT 15 июля в Антарктиде. Отрицательных значений над антарктической зоной не фиксируется ($WL_S=0$). Преобладающее истощение ПЭС (синий цвет) наблюдается в Арктике ($WL_N = -6,36$ в 00 UT и $WL_N = -7,16$ в 06 UT 16 июля), при этом положительные значения отсутствуют ($WU_N=0$). Такое максимально полушарно-асимметричное состояние обозначается как «однознаковое». Для всех рассмотренных бурь процент образования таких условий в арктической и антарктической зонах и глобально представлен в таблице 1.2.1. Выделены предбуревой период, главная фаза, фаза восстановления и полная продолжительность бури.

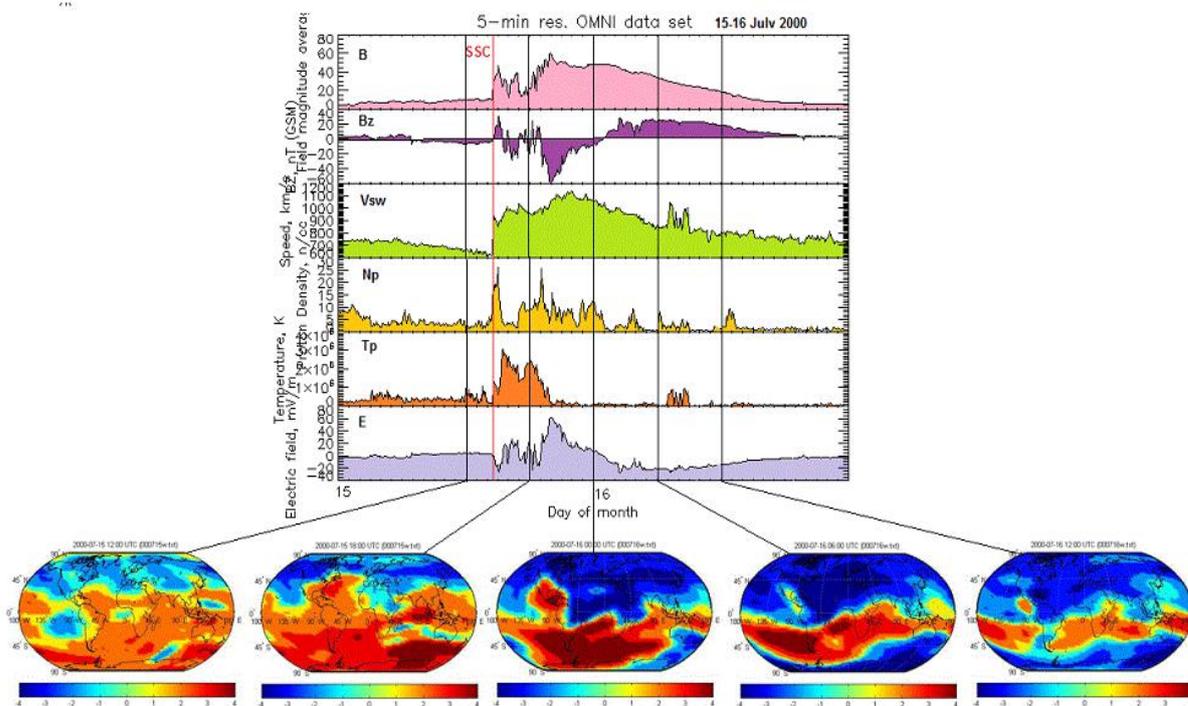


Рисунок 1.2.1 - Параметры солнечного ветра 15-16 июля 2000 г. (вверху); распределения ионосферных W индексов в 12 и 18 UT 15 июля, 00, 06 и 12 UT 16 июля

Таблица 1.2.1 - Процент образования «однозначных» условий в арктической и антарктической зонах и в глобальной ионосфере для периодов: предбури [от t_0-24 до t_0-1], главная фаза [от t_0 до t_0+23], восстановительная фаза [t_0+24 , t_0+47], полная продолжительность бури [t_0-24 , t_0+47]

Index	Arctica, %				Antarctica, %				Globe, %			
	t_1	t_2	t_3	t_{tot}	t_1	t_2	t_3	t_{tot}	t_1	t_2	t_3	t_{tot}
WU	13.68	23.11	25.57	20.79	10.43	25.40	24.52	20.12	0.71	0.37	0.59	0.56
WL	20.53	5.19	6.77	10.83	15.65	4.81	4.71	8.39	0.32	0.32	0.40	0.35
WE	0.04	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Данные, приведенные в таблице, показывают, что появление условий $WU_N=0$ варьируется от 14 до 25%, условий $WU_S=0$ - от 10 до 24%, условий $WL_N=0$ - от 5 до 20% и условий $WU_S=0$ - от 5 до 15%. Также имеется небольшое количество случаев, когда глобальный индекс $WU \approx 0$, а $WL \approx 0$. Индексы диапазона (WE) почти никогда не равны нулю, кроме нескольких исключений в предбуревой период. Доминирование во время ионосферной бури «однозначных» условий в Арктике и Антарктике (т.е. межполушарной асимметрии), которые возникают в высоких широтах и распространяются в средние и низкие широты — это результат, который требует особого внимания при разработке моделей для высоких широт. Во время бурь в арктической и антарктической зонах полярный индекс WU_N , WL_N , WU_S или WL_S равен нулю в течение 10–20 % предбуревоего времени, в течение 5–25 % времени основной фазы и 7–25 % в период восстановления. «Однозначное» состояние отражается в отсутствии корреляции индексов WU_{NS} и WL_{NS} с глобальными геомагнитными индексами.

В полярных областях обнаруживается и долговременное изменение индексов WU_{NS} и WL_{NS} . Более высокие значения суточных значений наблюдаются в высокоширотной области зимнего, а не летнего полушария. Обычно северо-южная асимметрия является результатом сезонных различий в солнечном излучении в противоположных полярных областях. Результаты модели полярной ионосферы показывают, что летом в освещенной солнцем ионосфере плотность плазмы увеличивается, а ее распределение относительно однородно. Переход от летнего к зимнему сезону приводит к резким изменениям, заключающимся не только в снижении плотности плазмы, но и в формировании крупномасштабных неоднородностей, часто с повышенными значениями N_e . Зимой, когда полярная ионосфера не освещена, образуются большие градиенты N_e , связанные с вариациями солнечного ветра и возмущенными геомагнитными условиями. Этим могут быть обусловлены наблюдаемые в W индексах полярных областей противоположных полушарий эффекты «однозначности», в частности, положительные отклонения ПЭС в зимних условиях.

Результаты данной работы опубликованы в журнале *Advances in Space Research*.

1.3 Развитие программных средств, основанных на эффективных численных схемах и пригодных для оперативного прогноза основных характеристик полярной ионосферы

Разработан входной блок для запуска численной модели высокоширотной ионосферы. Блок адаптирован для того, чтобы сделать ввод входных параметров удобным для пользователя. В дополнение к возможности задать произвольные значения входных параметров введена опция автоматизированного скачивания и ввода данных спутниковых

измерений солнечного ветра и геомагнитной активности (архивных или текущих значений). Задача поступления данных решается через скрипт (g:\scripts\update_danmp.py), который формирует входной файл указанного формата, пишет в него солнечный ветер, магнитное поле, время и константы.

Пример структуры входного файла:

Номер параметра	Название параметра	Формат ввода	Пример числового значения	Источник загрузки данных	Частота обновления
1, 2, 3	Кол-во точек разностной схемы; шаг по высоте; сезон года	I4 G8.1 I4	51 20.0D05 0	Константа для пользователя	0 - зима, 1 – лето
4	номер дня года	F6.1	8.0	online calendar	
5	поток солн. излучения на длине волны 10.7 см	F6.1	150.0	https://lasp.colorado.edu/lisird/data/cis_radio_flux_f107/	сутки
6	f10.7, усредненное за предыдущие 80 дней	F6.1	150.0	https://lasp.colorado.edu/lisird/data/cis_radio_flux_f107/	сутки
7	геомагнитный Kp индекс	F6.1	2.000	http://www-app3.gfz-potsdam.de/kp_index/Kp_ap_nowcast.txt	3 час
8	год солнечного цикла	F6.1	5.0		
9	разностный шаг по времени (сек)	F6.1	900.0	Константа для пользователя	
10	время релаксации в часах	F6.1	15.0	Константа для пользователя	
11	мировое время в часах и долях часа	F7.2	20.85	online calendar	
12	поток ионов O+ на верхней границе	F8.3	0.0	Константа для пользователя	
13	ММП Vy	F6.1	1.7	https://www.swpc.noaa.gov/products/real-time-solar-wind	1 час
14	ММП Vz	F6.1	5.9	https://www.swpc.noaa.gov/products/real-time-solar-wind	1 час
15	скорость солнечного ветра	F6.1	353.0	https://www.swpc.noaa.gov/products/real-time-solar-wind	1 час

1.4 Развитие методик глобальной и локальной оптической диагностики распределений интенсивности авроральных эмиссий в видимой и УФ-областях спектра для оценки динамики аврорального овала

Составлена первая часть аннотированного атласа, в который входит более двадцати примеров типичных авроральных изображений и измерений потоков заряженных частиц, полученных разными приборами с орбит космических аппаратов (КА) и с поверхности Земли. Отобраны события, происходившие в разных секторах магнитного местного времени (MLT) в условиях суббурь. В атласе каждый пример, а также характеристики приборов описаны и проиллюстрированы комбинированными

рисунками, адаптированными из опубликованных работ с соответствующими ссылками. Работа направлено на анализ опыта развития технологии подготовки и проведения комплексных экспериментов на орбитах КА, с помощью зондирующих ракет и с поверхности Земли, а также на расширение базы справочной информации по идентификации авроральных структур. Эксперименты были нацелены на исследования многообразных оптических и электромагнитных явлений в полярной ионосфере, которые отражаются в изменении градиентов Ne в различных слоях ионосферы. Атлас включает результаты наблюдений авроральных эмиссий, как в видимой области спектра, так и области вакуумного ультрафиолета (ВУФ). Наблюдения проводились с помощью орбитальных и наземных изображающих камер и спектрографов, а также спектрометров потоков высыпающихся частиц, работавших на борту КА на низких полярных орбитах. Все представленные примеры авроральных структур были выбраны исходя из условия, что они были возбуждены высыпающимися электронами плазменного слоя на ночной стороне овала или в дневном каспе. Энергизация частиц обусловлена процессами ускорения в ближней магнитосфере, включая альфвеновские волны, их флуктуации и резонансы, а также ускорением квазистационарными продольными электрическими полями и токами.

В атлас включены результаты измерений распределений потоков частиц и изображения авроральных эмиссий малых и средних масштабов (1-2 км и 10-15 км), полученных с орбит КА REIMEI, e-POP, TIMED и DMSP, и проанализирован опыт их получения. Некоторым из рассматриваемых примеров в конкретных событиях сопутствуют изображения авроральных эмиссий, полученные на геофизических станциях с помощью имаджеров с разными углами поля зрения и различным разрешением, а также характеристики полярной ионосферы, полученные с помощью радаров. Ниже приведены примеры изображений и комментариев по их интерпретации, представленные в атласе.

1) Квазистатические и альфвеновские структуры по данным аврорального имаджера MAC и датчика потоков частиц с орбиты КА Reimei.

На рисунке 1.4.1 показан пример наблюдений 26 декабря 2005 г. спектрограммы пич-распределения электронов и дуг полярных сияний. Наблюдалась авроральная полоса с дугами полярных сияний шириной порядка 1 км, несколько мелкомасштабных дуг, отображающих сдвиговые движения плазмы в полярной половине авроральной полосы и равномерное излучение в экваториальной части. Спектрограммы электронов показали высыпание электронов со спектрами типа «перевернутых V» шириной ~100 км, соответствующее авроральной лентообразной структуре. В полярной половине V-структуры, где появились мелкомасштабные дуги полярных сияний, есть мелкомасштабные флуктуации максимальной энергии электронов в диапазоне нескольких кэВ. При этом рассеянные во времени низкоэнергичные электроны высыпаются одновременно, свидетельствуя о влиянии механизма ускорения альфвеновскими волнами. Мелкомасштабные дуги полярных сияний генерируются не одним механизмом ускорения, а колебаниями максимальной энергии электронов в спектре типа перевернутой V, которые вызваны взаимодействием с альфвеновскими волнами.

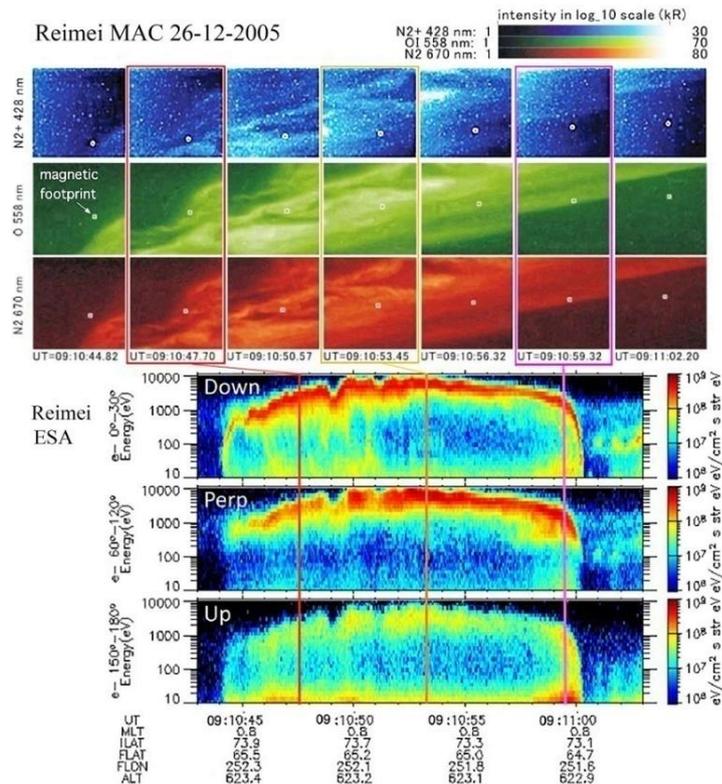


Рисунок 1.4.1 - Спектрограмма пич-распределения электронов и изображения дуг сияний (КА Reimei, 26 декабря 2005 г.) (рисунок сформирован из адаптированного материала)

2) Комплекс оптических и энергетических характеристик авроральных структур при развития суббури.

Для события 23 марта 2007 г. (развитие суббури началось в 11:20 UT) в интервале 11:20:27-11:20:39 UT с помощью наземных камер всего неба с тремя разными углами поля зрения и пространственным разрешением получены изображения в видимой части спектра на станциях Форт Юкон и Покер Флат (рисунок 1.4.2). Первый ряд рисунка состоит из трех последовательных изображений (1 кадр в сек) в 11:20:27; 11:20:33 и 11:20:39 UT на станции Форт Юкон.; второй ряд из трех изображений получен со станции Покер Флат (здесь авроральная структура находилась в магнитном зените) время экспозиции 0,5 сек (в первом ряду изображений это поле зрения ограничено трапециями); нижний ряд из трех изображений получен со станции Покер Флат с помощью изображающей камеры, имеющей полный угол поля зрения $9 \times 9^\circ$ (11° по диагонали и время экспозиции 0,033 сек). Для сравнения это поле зрения обозначено на снимках во втором ряду белыми квадратами. Снимки в третьем ряду получены с пространственным разрешением ~ 100 м.

Наблюдения показывают когерентность структур на самом мелком масштабе во время развития суббури. Мелкомасштабные дуги объединяются в пакеты, причем динамика движения самих пакетов и дискретных дуг внутри пакетов - разная. Элементарные дуги расходятся от центра дугового пакета. Ширина пакетов - в диапазоне от 2 до 8 км. На основе совокупных наблюдений полей и высыпаний частиц над активными полярными сияниями были сделаны выводы, что наиболее вероятным механизмом ускорения является корпускулярно-волновое взаимодействие в инерционной дисперсионной альфвеновской волне. Поток электронов либо модулировался, либо ускорялся волной. При этом дисперсия инерционной альфвеновской волны приводит к двум ортогональным эффектам: ускорению электронов в направлении магнитного поля B и распространению волновой энергии в плоскости B_{\perp} .

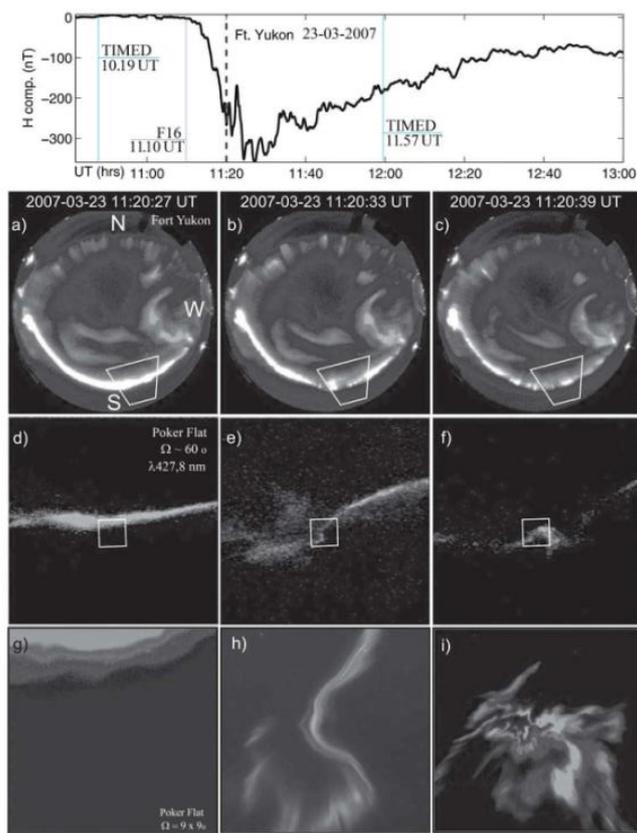


Рисунок 1.4.2 - Мультимасштабные изображения авроральных структур, полученные с трех камер на станциях на Аляске 23 марта 2007 г.: Форт Юкон (первый ряд); Покер Флат (второй и третий ряды). На верхней панели показана геомагнитная Н-компонента на станции Форт Юкон, пунктиром отмечен момент получения изображений (рисунок сформирован из адаптированного материала)

Одновременно с наземными измерениями проводились спутниковые оптические наблюдения и измерения энергетических спектров высыпающихся частиц. Для интервала 10:55-11:25 UT (начальная стадия суббури) доступны спектрограмма электронов и ионов, полученная с помощью электростатического анализатора SSJ5 с орбиты КА DMSP F16 и изображения четырех ВУФ-эмиссий, полученные в полосе строчной пространственной развертки с помощью спектрографа SSUSI. Эти данные представлены на рисунке 1.4.3. Во время развития суббури над Аляской также проходила траектория КА TIMED. На рисунке 1.4.4 представлены измерения ВУФ-эмиссий и рассчитанные из них характеристики (поток энергии высыпающихся электронов, средняя энергия электронов, электронная концентрация в максимуме Е-области и высота максимума N_e в Е-области). Данные получены с помощью изображающего спектрографа GUVI перед началом и в конце суббури. На панели в середине рисунка показан график ВУФ спектра, а ниже - фрагмент контурной карты Аляски с указанием расположения станций Форт Юкон и Покер Флат. Внизу левой и правой колонок панелей показаны распределения средней энергии вдоль всей пространственной полосы, отображаемой GUVI.

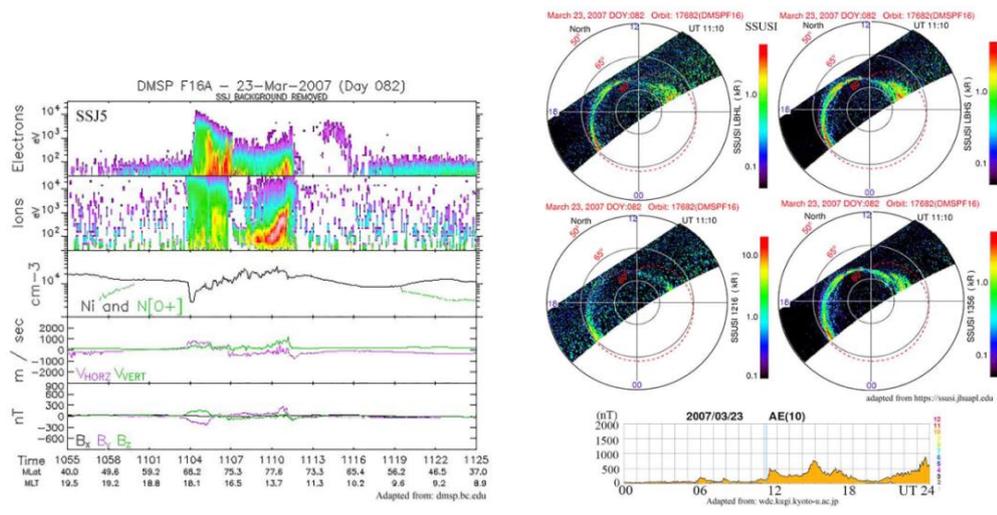


Рисунок 1.4.3 - (слева) Спектрограмма электронов и ионов, полученная электростатическим анализатором SSJ5 с орбиты КА DMSP F16; (справа) изображения ВУФ эмиссий LBH_L, LBH_S, λ 135,6 нм, λ 121 и 6+121,8 нм Ly- α в координатах InvLat-MLT, полученные с помощью спектрографа SSUSI на КА DMSP F16

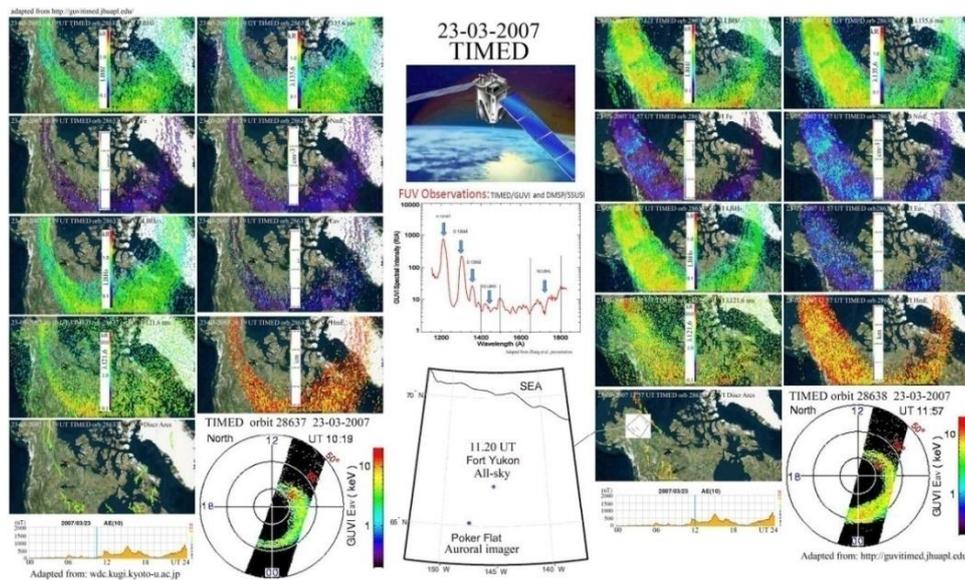


Рисунок 1.4.4 - Среднемасштабные измерения структур в авроральном овале над Аляской, ВУФ-спектрографом GUVI с орбиты КА TIMED в 10:19 UT (слева) и в 11:57 UT (справа)

Составленный атлас примеров мелко- и среднемасштабных наблюдений авроральных эмиссий и характеристик плазмы дает информацию о возможностях и трудностях методической координации одновременных наблюдений с разных орбит КА, траекторий зондирующих ракет и поверхности Земли, особенно в условиях динамики развития авроральных суббурь. Представленные в атласе примеры, а также методический опыт, полученный при их получении, могут быть использованы в дальнейшем при развитии технологии подготовки и проведения готовящихся экспериментов ИКИ РАН, нацеленных на исследования механизмов образования и динамики градиентов N_e в различных секторах полярной ионосферы. Примеры отображают базовые элементы развиваемой технологии получения изображений авроральных эмиссий с низких полярных орбит, которая в перспективе позволит приблизиться к лучшему пониманию

деталей процессов, приводящих к градиентам Ne разных масштабов в полярной зоне ионосферы, влияющих на условия распространения трансполярных сигналов.

По материалам данной работы подготовлена статья, принятая к публикации в журнале «Геофизические исследования».

1.5 Исследование статистических связей глобальных и локальных геомагнитных вариаций с процессами в ионосфере, магнитосфере и солнечном ветре с целью выявления закономерностей появления экстремальных событий и их прогноза

В рамках тематики данного раздела были сделаны оценки временных масштабов, на которых происходит отклик высокоширотной геомагнитной активности на воздействие на магнитосферу высокоскоростных потоков (ВСП) солнечного ветра (СВ). Известно, что основными источниками возбуждения геомагнитной активности являются скорость СВ и межпланетное магнитное поле (ММП) южного направления. В фазу спада солнечного цикла ВСП СВ, исходящие из долгоживущих корональных дыр, вызывают рекуррентные геомагнитные возмущения повышенной интенсивности. Ранее было показано, что среднемесячные значения скорости СВ могут быть с достаточной точностью реконструированы по рядам геомагнитных вариаций на высокоширотных обсерваториях. Однако, вопрос о том, насколько надежны реконструкции на более коротких временных масштабах, оставался открытым.

Для оценки корреляции между СВ и параметром суббуравой геомагнитной активности на изменяющемся временном масштабе используются среднечасовые значения горизонтальной северной (H) составляющей геомагнитного поля на обсерватории Соданкюля (географическая широта и долгота $67.4^{\circ}N$, $26.6^{\circ}E$; геомагнитная широта и долгота 64° , 119°). Магнитометр расположен вблизи экваториальной границы аврорального овала и подходит для мониторинга суббуравого западного аврорального электроджета. При суббуре наибольшие отрицательные отклонения H от спокойного уровня наблюдаются в секторе 22-01 MLT, что для Соданкюля соответствует 20-23 UT. Положительные отклонения наблюдаются в послеполуденном секторе и связаны с восточным электроджетом. Отрицательные – в ночном и утреннем секторах.

Для каждого месяца года определялся параметр ΔH как разность между H_q (спокойный уровень, рассчитанный по часовым значениям H в интервале 20-23 UT для 5 локально наиболее спокойных дней каждого месяца) и H_d (минимальное значение H , усредненное по часовым данным). Значения скорости СВ (V_{sw}) получены из базы данных НАСА (<http://omniweb.gsfc.nasa.gov/>).

Рассчитывался коэффициент корреляции между ΔH и V_{sw} при изменении усреднения от 1 до 30 дней. Статистика набиралась за 33-летний период с 1998 по 2021 гг. На рисунке 1.5.1 показано изменение коэффициента корреляции CC между ΔH и V_{sw} для сезона солнцестояния (январь), и двух сезонов равноденствия (апрель, октябрь). На масштабах от 30 до 8 дней значение CC уменьшается примерно с постоянной скоростью. После этой точки корреляция начинает резко падать. Также наблюдаются сезонные различия так, что значения CC в осеннее равноденствие превышают значения в другие месяцы и оказываются на $\sim 20\%$ больше, чем в апреле. Это соотношение исчезает при усреднении на масштабе менее 7 дней. Таким образом проявляется сезонная асимметрия геоэффективности ВСП СВ, которая обусловлена как геометрическим (координатным) фактором, так и асимметрией полушарий Солнца. Потеря корреляции на масштабе менее 7 дней, (что сравнимо с продолжительностью вызываемой ВСП СВ магнитной бурей) может быть связана с тем, что в магнитосфере начинают преобладать внутренние «loading-unloading» процессы.

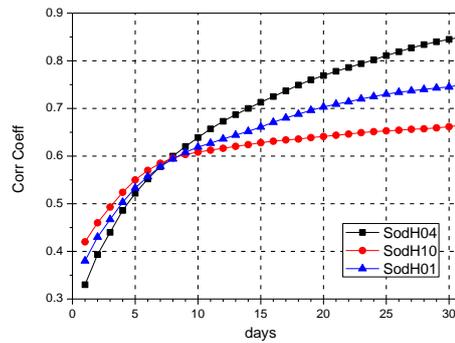


Рисунок 1.5.1 – Изменение коэффициента корреляции между ΔH и V_{sw} для января (синий цвет), апреля (черный) и октября (красный) при изменении временного масштаба усреднения от 1 до 30 дней

1.6 Создание деревьев принятия решений для классификации изменений состояния организма человека в связи с воздействием на него факторов космической и земной погоды на основе алгоритмов дифференциальной эволюции

Решением задачи обнаружения воздействия земной и космической погоды на организм человека является предсказание значений физиологических параметров, в частности, длины кардиоинтервала в ответ на изменения погодных факторов. Однако, отклик на изменение одного фактора зависит от совокупности всех остальных, т.е. искомая зависимость является многопараметрической. Существует два инструмента, эффективно решающих поставленную проблему – это решающее дерево и решающий лес. Второй инструмент немного выигрывает в точности решения по сравнению с решающим деревом, однако значительно проигрывает ему в плане возможности интерпретации решения, скажем, медицинскими работниками. Поэтому мы остановили свой выбор на рассмотрении именно решающих деревьев.

На текущем этапе исследований был разработан эволюционный алгоритм, который (а) в случае произвольным образом сконструированного принципа отбора способен решать задачу многокритериальной оптимизации и (б) благодаря стохастичности имеет шансы найти решение, близкое к глобальному оптимуму. В данном случае оптимизация деревьев идет по трем критериям: компактности представления, проценту допущенных ошибок при классификации и вероятности отказа от принятия решений. Предложенные технологии, обеспечивающие возможность приблизиться к глобальному максимуму сильнее, чем детерминированные деревья – это особая конструкция алгоритмов скрещивания-мутации и основанный на методе туннельной конкуренции и операторе многокритериального приоритета принцип отбора.

Основной структурной единицей рассматриваемого эволюционного алгоритма является дерево из предикторных узлов (фрейм). Для заданного фрейма существует полностью детерминированный алгоритм, превращающий его в конечное решающее дерево, оптимальность которого была доказана. Общая схема эволюционирования ансамбля фреймов (популяции) такова: на основе созданной таблицы узлов случайно инициализируется исходная популяция фреймов. Сущностью эволюционной методики создания оптимального объекта является комбинирование свойств случайных пар объектов и незначительное изменение свойств каждого объекта в случайную сторону. За первое отвечает операция скрещивания, за второе – мутации. Для создания случайных пар объектов исходная популяция делится на две части. Образующиеся после скрещивания потомки мутируют и превратившись в решающие деревья, вступают в конкуренцию с так же превращенной в решающие деревья исходной популяцией по всем трем перечисленным критериям, за что отвечает оператор отбора. В результате половина

суммарной популяции потомков и предков уничтожается, и получившееся новое поколение фреймов поступает на вход следующей итерации алгоритма. На последнем этапе выбирается наилучший фрейм.

Оптимальным в плане быстрой генерации случайных деревьев, их скрещивания и мутации оказался афферентный код. Афферентным кодом бинарного дерева из N вершин со списком узлов I , $|I| = N$ называем пару $A \equiv [I, J]$, где J – подстановка длиной N без повторений из множества $\{-N; -(N-1); \dots; -1; 0; 1; \dots; N-1; N\}$, в которой обязательно присутствие нуля. Следующее свойство подстановки J препятствует появлению разрывов и петель в графе. Оно состоит в том, что ни для какого элемента J его модуль не может совпадать с его номером, т.е. для любого $i = 1, 2, \dots, N$: $|J(i)| \neq i$.

В искусственно запрограммированной эволюции аналогами изменчивости и наследственности являются операторы скрещивания \times и мутации \Rightarrow , определенные на множестве графов Ω (в данном случае это класс бинарных деревьев), представляющем собой всевозможные деревья с принадлежащем некоторому заранее заданному списку узлов вершинами. При этом, если A и B – деревья из Ω , то $A \times B \in \Omega$ и $A \Rightarrow \in \Omega$ (во избежание двусмысленности мы ставим унарный оператор мутации позади своего аргумента, т.е. если дерево A мутирует в дерево A' , то $A \Rightarrow A'$). Иными словами, операторы скрещивания и мутации должны быть эндоморфны. Эндоморфность позволяет не «вылетать» за пределы Ω .

Базовыми действиями над деревьями в контексте скрещивания и мутации являются слияние ($+_o$) и расщепление ($-_o$). Нижний индекс o обозначает валентность, т.е. потенциальную или реализованную возможность прикрепления ребра определенного цвета к вершине. Поскольку рассматриваются только бинарные деревья, валентностей у любой вершины i ровно две: i и $-i$. Процедура слияния деревьев проиллюстрирована на рисунке 1.6.1. Валентности делятся на свободные (к которым существует возможность присоединения ребра без выхода полученного объекта за пределы класса бинарных деревьев) и занятые, или связанные (к которым ребро уже присоединено). Процедурой слияния деревьев A и B по свободной валентности o является образование графа C из вершин A и B путем присоединения всего дерева B своей корневой вершиной к вершине с номером $|o|$ ($|\cdot|$ обозначает модуль числа) дерева A . Валентность o обозначается в индексе оператора слияния: $C = A +_o B$. Процедурой расщепления дерева Z по (связанной) валентности o называем разрыв восходящего к вершине с номером $|o|$ ребра определенного цвета, соответствующего знаку валентности o : $[A, B, \dots] = -_o Z$. На основе операторов слияния и расщепления строятся алгоритмы скрещивания и мутации.

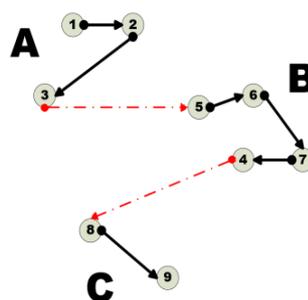


Рисунок 1.6.1 – Схема преобразования графа, первоначально содержащего три компоненты связности, в однокомпонентный путем добавления двух ребер (красные пунктиры)

Затем осуществляется поиск референсной точки. Процедура проиллюстрирована на рисунке 1.6.2. Пусть узел с фактором ζ и референсной точкой x (обозначаемый далее как $\{\zeta, x\}$) порождает два условия – A и B , соответствующие красной и черной нисходящей

ветви. Обозначим за Θ совокупность условий на восходящем маршруте до корневой вершины дерева, обозначенном на рисунке 1.6.2 синим цветом. Часть целевого функционала, зависящая от рассматриваемой вершины (на рисунке 1.6.2 обозначена зеленым), запишется как

$$\Phi = 1 - \sum_{i \in I_A} q_{\alpha(i)}^i - \sum_{j \in I_B} q_{\alpha(j)}^j \quad (1.6.1)$$

где i, j – обвешанные валентности, относящиеся к нисходящим маршрутам, ассоциированным с черной ($i \in I_A$) и красной ($j \in I_B$) ветвями вершины $\{\zeta, x\}$. Распишем фидуциальные вероятности в суммах (1.6.1):

$$\begin{cases} q_{\alpha(i)}^i = \lambda_{\alpha(i)} P_{\alpha(i)}^i = \lambda_{\alpha(i)} \frac{P[G_i A \Theta \alpha(i)]}{P[\alpha(i)]} \\ q_{\alpha(j)}^j = \lambda_{\alpha(j)} P_{\alpha(j)}^j = \lambda_{\alpha(j)} \frac{P[H_j B \Theta \alpha(j)]}{P[\alpha(j)]} \end{cases} \quad (1.6.2)$$

Где G_i и H_j – условия на нисходящих маршрутах от узла $\{\zeta, x\}$, отвечающих черной и красной ветви, $P[*]$ – вероятности пересечения условий в квадратных скобках (во избежание излишне перегруженных формул пересечение событий записывается в виде их «произведения»), $P[\alpha(i)]$ и $P[\alpha(j)]$ – априорные вероятности появления соответствующих классов. Тогда:

$$\begin{cases} q_{\alpha(i)}^i = \lambda_{\alpha(i)} \frac{P[G_i | A \Theta \alpha(i)] P[A \Theta | \alpha(i)] P[\alpha(i)]}{P[\alpha(i)]} = \lambda_{\alpha(i)} P[G_i | A \Theta \alpha(i)] P[A \Theta | \alpha(i)] \\ q_{\alpha(j)}^j = \lambda_{\alpha(j)} \frac{P[H_j | B \Theta \alpha(j)] P[B \Theta | \alpha(j)] P[\alpha(j)]}{P[\alpha(j)]} = \lambda_{\alpha(j)} P[H_j | B \Theta \alpha(j)] P[B \Theta | \alpha(j)] \end{cases} \quad (1.6.3)$$

Вероятности $\lambda_{\alpha(i)} P[A \Theta | \alpha(i)] \equiv \pi_{\alpha(i)}^A(x)$ и $\lambda_{\alpha(j)} P[B \Theta | \alpha(j)] \equiv \pi_{\alpha(j)}^B(x)$ измеримы для каждой подстановки $\alpha(i)$, $\alpha(j)$ и очевидным образом зависят от референсной точки x , поскольку последняя порождает разделение прецедентов по ветвям A и B . Подставим выражения (1.6.2-1.6.3) в (1.6.1):

$$\Phi = 1 - \sum_{i \in I_A} \pi_{\alpha(i)}^A(x) P[G_i | A \Theta \alpha(i)] - \sum_{j \in I_B} \pi_{\alpha(j)}^B(x) P[H_j | B \Theta \alpha(j)] \quad (1.6.4)$$

Минимум функционала Φ достигается тогда, когда обе суммы максимальны. Тогда Φ представляет собой разность единицы и двух выпуклых комбинаций переменных $\pi_{\alpha(i)}^A(x)$ и $\pi_{\alpha(j)}^B(x)$. Требуется дополнительное ограничение для того, чтобы сформулировать условие минимума для Φ , а именно приоритет решений одного класса на свободных валентностях, отвечающих маршрутам G_i и H_j . Обозначим эти классы за $\bar{\alpha}$ (для маршрутов G_i) и $\bar{\beta}$ (для маршрутов H_j) Тогда

$$\begin{cases} P[G_i | A \Theta \alpha(i)] \approx 0, \alpha(i) \neq \bar{\alpha} \\ P[H_j | B \Theta \alpha(j)] \approx 0, \alpha(j) \neq \bar{\beta} \end{cases}$$

и:

$$\begin{aligned} \Phi &\approx 1 - \pi_{\bar{\alpha}}^A(x) \sum_{i \in I_A} P[G_i | A \Theta \bar{\alpha}] - \pi_{\bar{\beta}}^B(x) \sum_{j \in I_B} P[H_j | B \Theta \bar{\beta}] = \\ &= 1 - \pi_{\bar{\alpha}}^A(x) - \pi_{\bar{\beta}}^B(x) \geq 1 - \max_x \max_{\bar{\alpha}, \bar{\beta}} (\pi_{\bar{\alpha}}^A(x) + \pi_{\bar{\beta}}^B(x)) \end{aligned} \quad (1.6.5)$$

Итак, оптимальной референсной точкой для заданного узла $\{\xi, x\}$ является та, которая максимизирует выражение (1.6.5) для некоторой пары классов $\bar{\alpha}$ и $\bar{\beta}$. Этот критерий требует знания конфигурации узлов на восходящем к корню маршруте, поэтому расстановка референсных точек по фрейму происходит при движении от корневого узла.

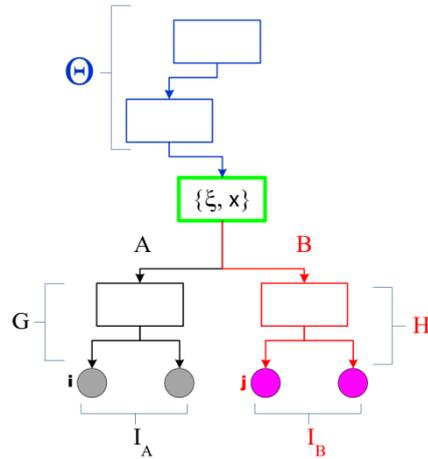


Рисунок 1.6.2 – Фрагмент решающего дерева с входящими и исходящими маршрутами для поиска референсной точки заданной вершины

Для предотвращения застревания алгоритма в локальных минимумах был разработан метод туннельной конкуренции. Его суть в приписывании меньшей вероятности уничтожения для особей дочернего поколения. Работу данного принципа в простейшей задаче оптимизации функции двух переменных иллюстрирует рисунок 1.6.3.

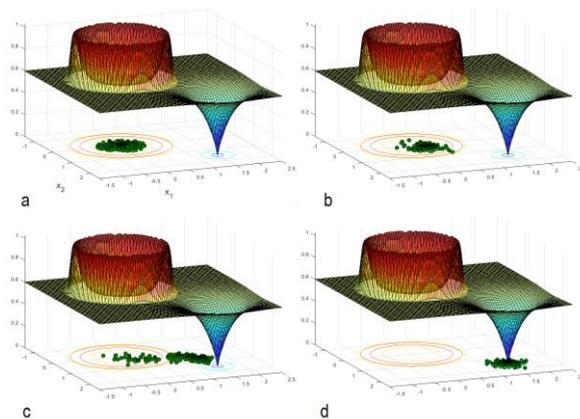


Рисунок 1.6.3 - Результат применения туннельной конкуренции и сходимости эволюционного алгоритма к глобальному минимуму: на плоскости $[x_1, x_2]$ показаны контуры равного уровня, зеленые точки обозначают популяцию; (а) старт алгоритма – исходная популяция равномерно распределена по дну бассейна; (б) появление мутационной цепочки, туннелирующей сквозь высокий барьер; (с) большая часть популяции преодолела барьер; (д) сходимость алгоритма к глобальному максимуму без вырождения популяции

Для сравнения разработанного нами метода построения деревьев с общепринятыми алгоритмами использовалась библиотека *scikit-learn* на языке Python. Пятиуровневое решающее дерево на Python показало интегральную ошибку классификации 16.2% против ошибки «эволюционного» дерева в 14.3%. В обоих случаях вероятность отказа от принятия решений была равна нулю. Затем мы снизили ошибку до 5%, отказавшись от принятия решений в 60% случаев. На рисунке 1.6.4 показано выращенное при помощи эволюционных алгоритмов решающее дерево с пятью уровнями предикторных узлов в приложении к иерархии факторов. В корневой вершине находится диагноз – это главный предиктор. Следующие два уровня занимают погодные факторы, и наконец последние два уровня занимает магнитное поле (участвуют минимальный-максимальный уровень и отношение сигнал/шум; от индекса монотонности поля, как видно, решения не зависят). Таким образом, факторы физиологии, земной и космической погоды упорядочены по убыванию своего влияния следующим образом:

- ✓ диагноз;
- ✓ погодные условия;
- ✓ магнитное поле.

По материалам данной работы готовится публикация в журнале LIFE.

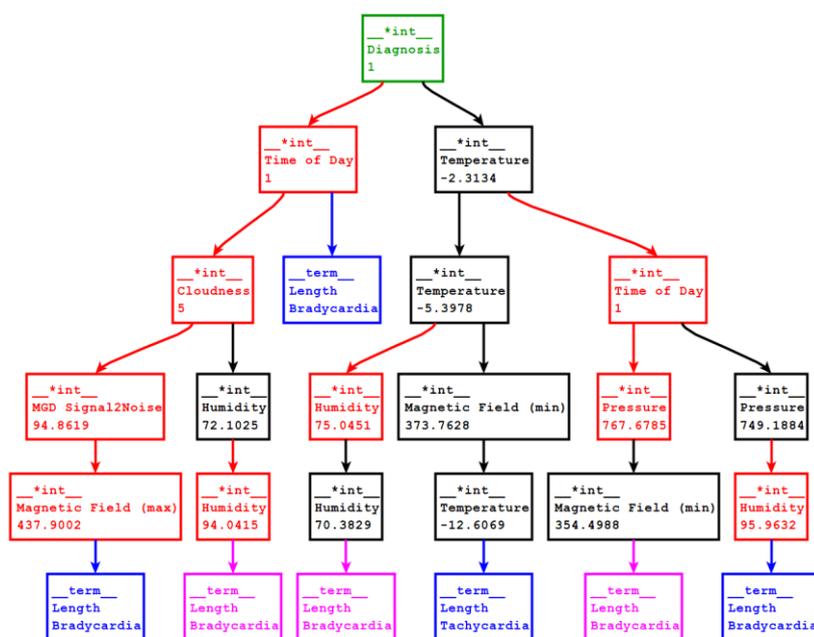


Рисунок 1.6.4 - Решающее дерево для предсказания длины кардиоинтервала в зависимости от условий земной и космической погоды

1.7 Разработка методов диагностики источников и эффектов локальной ионизации в верхних слоях атмосферы Земли на основе анализа больших массивов данных радио наблюдений

На этапе 2022 г. тематика данного раздела была посвящена анализу факторов, влияющих на ориентацию авроральных дуг. Традиционно считается, что дуги полярных сияний вытянуты вдоль аврорального овала под углом к географической параллели. С ростом геомагнитной активности овал расширяется, а угол растет. Характеристики авроральных дуг контролируются параметрами солнечного ветра и магнитосферной динамикой. Однако статистический анализ большого массива данных наблюдений

неожиданно показал наличие определенного влияния на авроральные дуги со стороны динамики нижележащих слоев атмосферы.

Статистически исследована ориентация порядка 10000 авроральных дуг, наблюдавшихся в течение 2016-2021 гг. камерой всего неба на ст. Абиско, Швеция (АВК, 68,4° с.ш., 18,8° в.д.) в экваториальной части ночного аврорального овала. Наблюдения проводились между 19 и 06 часами MLT. В среднем ориентация дуг, т. е. угол между дугой и направлением запад-восток, линейно менялся в зависимости от MLT со скоростью 2,2°/час. В большинстве случаев средняя ориентация дуг около полуночи ожидаемо была параллельна геомагнитной широте, на которой находится ст. Абиско. Исключением явились только два месяца в 2018 г., когда дуги отклонились от той линии, которую им предписывает магнитное поле, параметры солнечного ветра и геомагнитные возмущения. Оказалось, что аномалии ориентации авроральных дуг совпадают с аномальным весенним переходом атмосферной циркуляции от зимнего к летнему типу. Перестройка циркуляции проявлялись в данных метеорологического радара в районе наблюдений как поворот зонального ветра на высотах 90—100 км с востока на запад. Результаты предполагают, что эти весенние переходные изменения в глобальной атмосферной циркуляции проникают в вышележащую термосферу на высоту 100-150 км, где нейтральные ветры могут воздействовать на ионосферу, а затем и на всю ионосферно-магнитосферную систему. Эти инициированные снизу модуляции ионосферно-магнитосферной системы проявляются в изменении ориентации авроральных дуг.

Для анализа использовались изображения, полученные на длине волны 557,7 нм в условиях темного неба. Из-за этого между апрелем и сентябрем происходит летний перерыв. Пример изображения приведен на рисунке 1.7.1. Вручную было установлено не менее 5 точек вдоль нижнего края дуги, наиболее удаленной от зенитного пикселя. Координаты точек были сохранены вместе с датой изображения. Дуги аппроксимировались с применением ортогональной регрессии

$$Y = D + X \tan \alpha \quad (1.7.1)$$

где D — расстояние вдоль географического меридиана, α — угол между дугой и направлением на географический восток, положительный с востока на север.

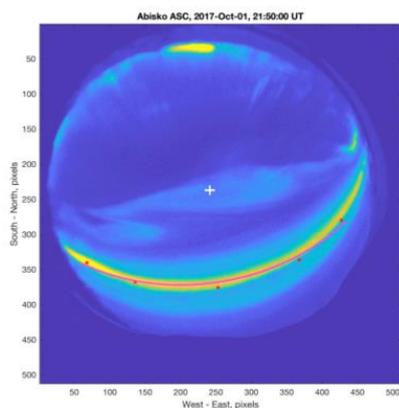


Рисунок 1.7.1 - Пример оцифровки изображения камеры всего неба. Красные точки определяют положение дуги. Красная кривая показывает проекцию на изображение линии регрессии, представляющей дугу. Зенитный пиксель обозначен белым крестом.

Для расчета углов между дугами и геомагнитной широтой использовалась модель главного геомагнитного поля и преобразование из географических координат в исправленные геомагнитные (CGM) (<https://omniweb.gsfc.nasa.gov/vitmo/cgm.html>). Расчетные параметры для ст. АВК на высоте 110 км приведены в таблице 1.7.1. Из-за

векового хода главного поля угол между географической и геомагнитной широтами (угол овала) изменялся с 6,2° в 2016 г. до 5,3° в 2021 г.

Таблица 1.7.1. - Исправленные геомагнитные координаты и параметры модели IGRF/DGRF

Год	ABK 68.4°N, 18.8°E		
	CGM широта (°)	Угол овала (°)	Полночь UT
2016	65.78	6.19	21.62
2017	65.81	6.00	21.63
2018	65.84	5.81	21.64
2019	65.87	5.64	21.66
2020	65.90	5.48	21.68
2021	65.94	5.33	21.69
Среднее	65.86	5.74	21.65

На рисунке 1.7.2 показаны углы дуг с географической широтой (α) в зависимости от MLT. График демонстрирует линейную тенденцию, так что угол увеличивается с MLT примерно от -5° в 19 MLT до примерно 20° в 06 MLT. Линейная подгонка методом наименьших квадратов дает выражение:

$$\alpha = A_0 + A_1 \cdot \text{MLT} \quad (1.7.2)$$

Коэффициенты регрессии $A_0=(5,8\pm 0,1)^\circ$ и $A_1=(2,21\pm 0,03)^\circ/\text{час}$ получены путем минимизации квадрата расстояния по вертикали от точек данных до линии регрессии. При $\text{MLT} = 0$ угол α близок к углу овала в ABK, и это означает, что в полночь дуги в среднем ориентированы по геомагнитной широте $65,9^\circ$. Если дуги ориентированы вдоль овала, то смысл коэффициента A_1 легко понять в рамках простой модели в геомагнитных координатах, где центральная линия овала аппроксимируется окружностью радиуса r , центр которой расположен на полуночном меридиане на расстоянии d от полюса. Тогда угол α_o между овалом и геомагнитной широтой при изменении MLT определяется как

$$\alpha_o = \text{atan} \left(\frac{\sin \varphi_M}{\cos \varphi_M - d/C_0} \right) - \varphi_M \quad (1.7.3)$$

где угол $\varphi_M = 15^\circ \cdot \text{MLT}$, а геомагнитная широта центральной линии овала:

$$C_0 = \sqrt{r^2 - d^2 \sin^2 \varphi_M} + d \cos \varphi_M \quad (1.7.4)$$

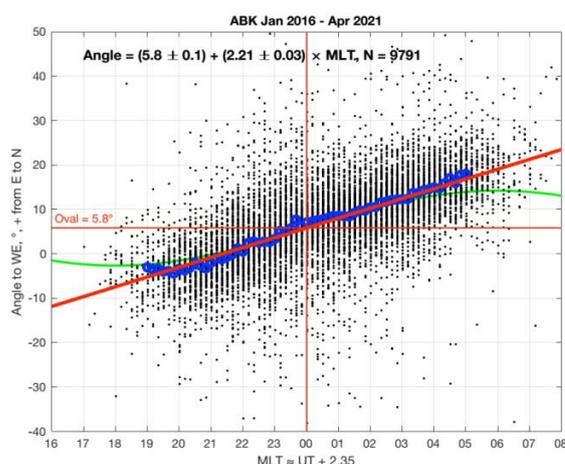


Рисунок 1.7.2 – Зависимость угла между дугой и географической широтой по MLT: черные точки - для каждой из оцифрованных дуг, синие кружки - усреднение по 10-минутным интервалам, красная линия - линия регрессии согласно (1.7.1), зеленая - модельная зависимость α_o от MLT, рассчитанная по (1.7.3-1.7.4)

Значения углов наклона дуг к географической параллели были усреднены помесячно. Корреляция среднемесячных значений углов с локальным параметром геомагнитной активности ΔH (отклонение от спокойного уровня северной компоненты геомагнитного поля, измеренной магнитометром ст. Абиско,) представлена на рисунке 1.7.3. В целом, зависимость линейная, т.е. с ростом геомагнитной активности овал расширяется, угол растет. Однако имеются две выпадающие точки в феврале и марте 2018 г., когда угол очень мал, т.е. дуги ориентированы не вдоль геомагнитной, а больше вдоль географической параллели.

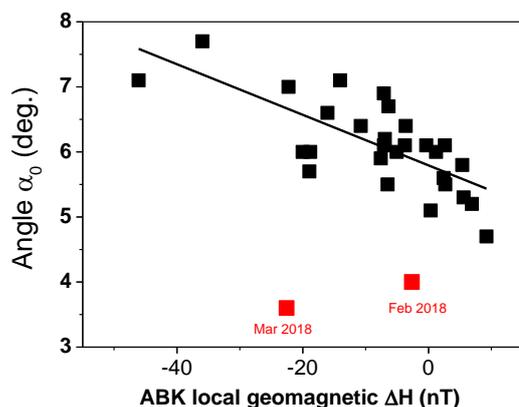


Рисунок 1.7.3 – Корреляция и линия линейной регрессии для зависимости между среднемесячными значениями угла и параметром локальной геомагнитной активности ΔH (черный цвет) и две выпадающие точки в 2018 г. (красный цвет)

Конец зимы 2018 г. отличался необычно ранней и активной перестройкой атмосферной циркуляции. Это показывают измерения нейтрального ветра на высотах мезосферы (80-100 км) с помощью метеорного радара, расположенного в обсерватории Соданкюля (67,2°N, 26,4°E) вблизи ст. Абиско. На рисунке 1.7.4 представлены усредненные помесячно (период январь-август) значения скорости зонального ветра. В конце зимы 2018 г. преобладало противоположное обычному направлению ветра, что свидетельствует об аномалии циркуляционного режима.

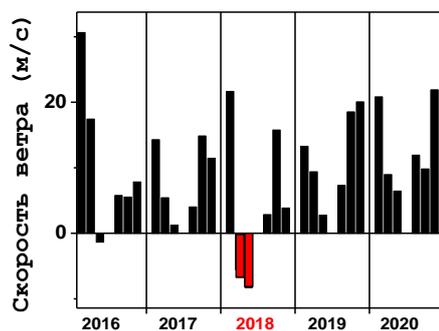


Рисунок 1.7.4 – Усредненные за месяц (январь-июль) значения скорости зонального ветра на высоте ~100 км по данным метеорного радара обс. Соданкюля. Положительное направление на восток. Красным цветом выделены февраль и март 2018 г., когда преобладающий ветер имел противоположное направление

Маловероятно, что нейтральные ветры на мезосферных высотах могут напрямую влиять на ориентацию дуг. Однако, если весенняя перестройка охватывает термосферу и

на больших высотах (до 150 км), где эффективен механизм динамо генерации электрического поля, то это может воздействовать на электродинамику ионосферно-магнитосферной системы. Для моделирования ориентации дуг при весенних перестройках системы нейтральных ветров нужны данные о скорости ветра выше 100 км, но таких данных пока нет. И в целом нейтральный ветер на этих высотах изучен слабо, хотя очевидно, что ветер в ионосфере может кардинально отличаться от того, что измеряется на 90-100 км. Поэтому можно только качественно предположить, что нейтральный ветер может повлиять на ионосферно-магнитосферную систему и изменить ориентацию дуг посредством проникновения электрических полей ионосферного динамо в область плазменного слоя. Сами условия плазменного слоя контролируются внутренней динамикой хвоста магнитосферы, и обычно именно это определяет ориентацию авроральных дуг. Однако, в сезон перестройки циркуляции при больших ее аномалиях эффекты нейтральной атмосферы могут распространяться вплоть до высот магнитосферы. Если в нижней термосфере появляется сильный восточный ветер (это всегда связано с внезапным стратосферным потеплением в полярной стратосфере в конце зимы), то дуги разворачиваются больше вдоль географической параллели, и меньше вдоль аврорального овала. Во время сезонной перестройки атмосферной циркуляции нейтральный ветер в средней атмосфере может оказывать влияние на ориентацию дуг, т.е. воздействовать на магнитосферно-ионосферную систему.

По результатам исследований подготовлена статья к публикации в журнале «AGU Advances».

1.8 Развитие методов диагностики метеорных потоков на основе анализа их ионосферных эффектов

Работы, первоначально запланированные по данному разделу как «Развитие методов диагностики метеорных потоков на основе анализа их ионосферных эффектов», оказались невыполнимыми вследствие закрытия доступа к данным из иностранного источника, ранее бывшим открытыми в рамках международного сотрудничества. Существующий отечественный метеорный радар (Обнинск) не обладает достаточной чувствительностью, а также временным и пространственным разрешением, позволяющим в настоящее время выполнить намеченное исследование. В связи с этим тематика работ была модифицирована. В данном разделе описываются полученные в 2022 г. результаты по исследованию изменчивости интегрального влагосодержания атмосферы по данным спутникового микроволнового зондирования. В работе использовались данные, находившиеся в открытом доступе.

На базе коллекции цифровых карт среднемесячных полей интегрального влагосодержания (ИВС) атмосферы по данным радиотеплового зондирования спутников серии DMSР за период 1988-2020 гг. исследована широтная и межгодовая изменчивость интегрального содержания водяного пара в тропосфере над северными частями Атлантического и Тихого океанов. Показано, что в Тихоокеанском регионе влагосодержание росло линейно и за рассматриваемый период увеличилось на 1.5 мм. В северной Атлантике до начала 2000-х количество пара также довольно быстро увеличивалось, и его общее повышение составило около 2.5 мм. Однако в последующие декады тренд существенно замедлился. В Атлантическом секторе в межгодовой изменчивости выделяются восемь зим, которые характеризовались ослаблением переноса атмосферного тепла и влаги из низких широт в высокие; все они соответствуют годам фазы спада солнечной активности.

Регулярные радиотепловые измерения с борта низкоорбитальных космических аппаратов начались более 30 лет назад, когда НАСА были запущены первые аппараты серии DMSР (Defense Meteorological Satellite Program, <http://dmsp.ngdc.noaa.gov/dmsp.html>). К настоящему времени калиброванная база данных

глобального распределения интегрального влагосодержания атмосферы охватывает период, отвечающий критерию достаточности для оценки климатических норм, что позволяет поставить задачу выявления долговременных изменений. Источником данных для анализа изменчивости ИВС служит портал Remote Sensing Systems (RSS, <http://www.remss.com>). Электронная коллекция продуктов RSS содержит карты полей ИВС, а также ряд других параметров атмосферы (водозапас облачных структур, зоны интенсивного выпадения осадков, скорость ветра и др.) над океанами. Глобальные карты, представленные на портале, строятся по калиброванным данным сканирующих многоканальных микроволновых радиометров серии SSM/I (Special Sensor Microwave/Imager) в спектральном диапазоне от 19 до 85 ГГц, которые установлены на борту спутников долговременной миссии DMSP. В рамках программы DMSP на орбиту последовательно выводились космические аппараты F08-F18 таким образом, что одновременно и непрерывно функционируют два (с 1995 г. – три) типовых радиометра. Данные архива RSS охватывает интервал радиотепловых спутниковых наблюдений с 1988 года по настоящее время. Информация является открытой для использования в научных целях.

Временной ряд среднемесячных и среднегодовых аномалий, усредненных по широтному диапазону $\pm 70^\circ$, показан на рисунке 1.8.1. Оба параметра имеют восходящий тренд, который описывается линейным регрессионным уравнением $y = +0.04 \cdot x - 85.65$. Это дает скорость увеличения влагозаписа атмосферы 0.45 мм/10 лет. Коэффициент детерминации равен $r^2 = 0.58$, т.е. линейный тренд описывает более половины изменчивости временного ряда.

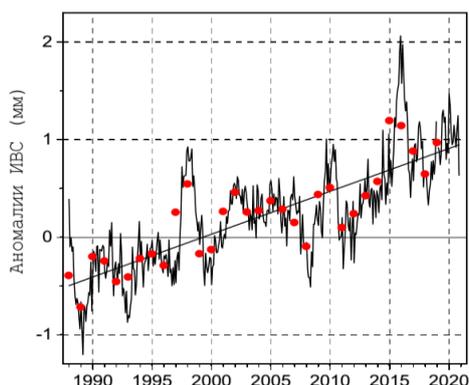


Рисунок 1.8.1 - Среднемесячные и среднегодовые (красные точки) аномалии ИВС в широтном диапазоне $\pm 70^\circ$ и линейный тренд

Хотя в планетарном масштабе ИВС демонстрирует восходящий линейный тренд, в региональном масштабе тренд менее равномерный. Среднегодовые аномалии ИВС, рассчитанные для северной части Атлантического и Тихого океанов в широтном диапазоне $40-70^\circ$ с.ш., представлены на рисунке 1.8.2. Можно видеть, что в Атлантике довольно быстрое увеличение ИВС наблюдалось до начала 2000-х, после чего рост практически прекратился. Трендовая составляющая изменчивости явно нелинейна и описывается полиномом 2-го порядка $y = 0.22 \cdot x^2 - 0.04 \cdot x - 1.92$. Тренд объясняет большую часть изменчивости ($r^2 = 0.86$), общее увеличение ИВС составляет примерно 2.5 мм, т.е. более 10%. В тихоокеанском секторе влагозапас увеличивался более равномерно, следуя линейному тренду $y = 0.04 \cdot x - 0.71$, и за тридцатилетний период вырос на 1.5 мм.

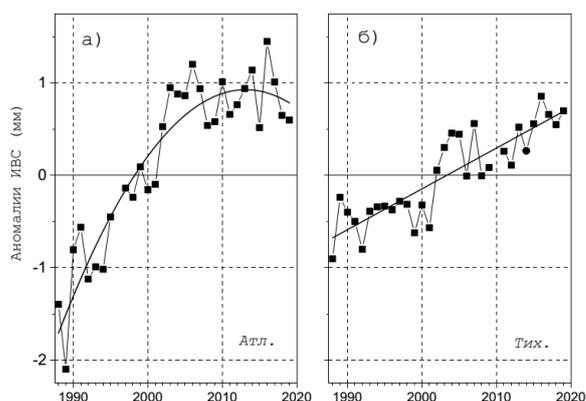


Рисунок 1.8.2 - Среднегодовые аномалии ИВС для (а) Атлантического и (б) Тихого океанов в широтном диапазоне 40-70° с.ш.

Для выявления межгодовой изменчивости ИВС были построены диаграммы Ховмёллера, которые помогают визуализировать эволюцию параметра на основе его статических изображений. Диаграммы в осях “время-широта” показывают, как водяной пар распространяется с юга на север от месяца к месяцу в течение всего периода имеющихся спутниковых наблюдений. Для построения диаграммы с каждой карты среднемесячных распределений ИВС на определенной долготе выбирался ряд пикселей (ширина 1°), и последовательно помещался на ось у. В качестве репрезентативных были выбраны меридианы 15° в.д. и 165° в.д. (для Атлантического и Тихого океанов, соответственно). Широтный диапазон ограничен 40° и 70° с.ш.

На рисунке 1.8.3 представлена широтно-временная диаграмма ИВС для северной Атлантики. В сезонном ходе параметр изменяется от 10 мм зимой до 20 мм летом. При смещении к полюсу от 40° до 70° с.ш. количество пара уменьшается примерно на 30%. На фоне систематической сезонной вариации наблюдается значительная межгодовая изменчивость как в общем количестве пара, так и в его распространении от южной границы области к северной. На рисунке визуально заметно, что в зимние сезоны в 1993-1995, 2004-2006 и 2015-2016 гг. распространение пара от экватора к северу было менее интенсивным, чем в другие годы. В эти же периоды летнего пара также было меньше.

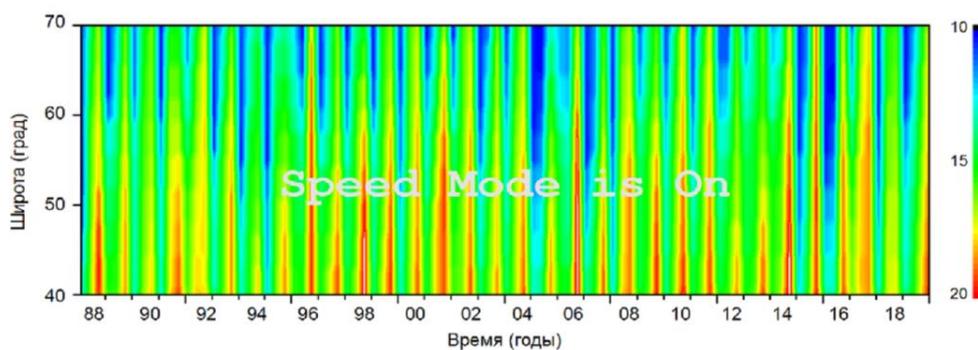


Рисунок 1.8.3 - Диаграмма «время-широта» для среднемесячных значений ИВС в северной Атлантике на меридиане 15° в.д.

Для того, чтобы определить, насколько систематически проявляются периоды слабого меридионального тепловлагопереноса и с чем они могут быть связаны, для каждой зимы была выделена самая низкая широта, которой достигает изолиния 10 мм. Были отобраны годы, когда эта широта была меньше порогового значения, в качестве которого была выбрана широта 65°. Для северной Атлантики таких зим оказалось восемь.

Для тихоокеанского сектора, где значения ИВС в целом выше, была выбрана изолиния 13 мм. Годы пониженного влагосодержания были соотнесены с 11-летним циклом солнечной активности (рисунок 1.8.4). Оказалось, что в атлантическом регионе все зимы, характеризующиеся минимальными ИВС, происходили в годы спада солнечного цикла, когда обычно к Земле приходят максимальные высокоскоростные потоки солнечного ветра (ВСП). Это согласуется с полученными ранее результатами, что ВСП влияют на озон и стратосферную циркуляцию в высоких широтах, усиливая зональный перенос и ослабляя меридиональный. В тихоокеанском регионе соотношение с солнечным циклом не однозначно: три зимы минимальных ИВС были с годы солнечного минимума, тогда как другие две – в периоды солнечного максимума в 2000 и 2012 гг. При этом оба года совпали с периодами холодной фазы южной осцилляции (эпизод Эль-Нино), что обусловило общее уменьшение влагосодержания и, как следствие, его пониженные значения в высоких широтах.

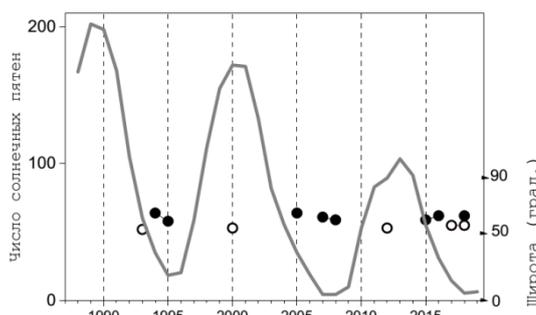


Рисунок 1.8.4 - Солнечный цикл (серая линия) и годы пониженного влагосодержания в северной Атлантике (черные точки) и северной части Тихого океана (открытые точки)

Таким образом, в североатлантическом регионе в интенсивности зимнего распространения водяного пара из приэкваториальных широт на север выделяется когерентность, совпадающая с солнечным циклом. Зимы, характеризующиеся минимальным количеством водяного пара, и как следствие, ослаблением полярного переноса тепла и влаги, соответствуют годам фазы спада солнечной активности и усиления ВСП. Это может быть еще одним подтверждением влияния солнечной активности на глобальный климат.

Результаты опубликованы в журнале «Метеорология и гидрология».

Заключение

Изложенные выше результаты показывают, что по направлению «Развитие методов моделирования высокоширотной ионосферы; построение численных и гибридных (включающих элементы ассимиляции данных измерений ионосферно-магнитосферных параметров) моделей распределения электронной концентрации в полярной ионосфере и зависимости от изменяющихся гелио-геофизических условий» проведенные работы соответствуют перечню, установленному в плане работ ИКИ РАН на 2022 год. Сложности возникали в связи с закрытием доступа к ряду прежде открытых источников данных, нарушением международного сотрудничества, отсутствием возможности обеспечить публикации в высокорейтинговых международных журналах. Тем не менее, работы проводились по каждому из заявленных разделов и выполнены в соответствии с планом.

Основные полученные результаты состоят в следующем:

На большой статистике оптических наблюдений определены зависимости ориентации авроральных дуг от параметров солнечного ветра и геомагнитной активности.

Дуги непосредственно связаны с магнитосферной электродинамикой. Но при определенных условиях во время сезонной перестройки атмосферной циркуляции нейтральный ветер в средней атмосфере может оказывать влияние на ориентацию дуг, т.е. воздействовать на магнитосферно-ионосферную систему.

С помощью набора ионосферных индексов показано, что во время магнитных бурь в полном электронном содержании регулярно наблюдается межполушарная асимметрия, которая особенно явно выражена в Арктике и Антарктике. Причем, увеличение ПЭС связано не столько с ростом освещенности в летний сезон, сколько с неоднородностями, формирующимися в зимней ионосфере.

Оценки временных масштабов, на которых происходит отклик суббулевой активности на воздействие на магнитосферу высокоскоростных потоков солнечного ветра, показали, что корреляция понижается сравнительно медленно и линейно на масштабах 8-30 дней. На более коротком временном масштабе происходит резкая потеря корреляции так как, в магнитосфере начинают преобладать внутренние диссипативные процессы.

Подготовлен обзор современных моделей, используемых для определения структуры крупномасштабных неоднородностей высокоширотной F-области, и подходов к введению блока ассимиляции данных. Разработан пользовательский входной блок для запуска численной модели высокоширотной ионосферы.

Подготовлен аннотированный атлас изображений эмиссий в авроральных структурах, обусловленных определенными проявлениями магнитосферной динамики.

По данным спутникового микроволнового зондирования выявлены тренды меридионального переноса водяного пара в северном полушарии. Проявляется когерентность, совпадающая с солнечным циклом. Зимы, характеризующиеся минимальным ослабленным полярным переносом, соответствуют годам фазы спада солнечной активности и усиления высокоскоростных потоков солнечного ветра.

Разработан алгоритм для построения деревьев принятия решений при классификации изменений состояния организма человека в связи с воздействием на него факторов космической погоды.

Материалы, изложенные в разделе 1, позволяют считать, что все работы по направлению «Космос-Д-Плазма» выполнены. По результатам подготовлены публикации в реферируемых журналах и материалах конференций.

ГЛАВА 2 Исследование и разработка алгоритмов машинного обучения для измерения расстояний (красных смещений) астрономических объектов по данным многоволновых фотометрических обзоров неба

Введение

Направление 2 ("Космос-Д-Астрофизика") темы "Космос-Д" посвящено вопросам, связанным с разработкой и совершенствованием научных основ, методов, технологий и систем работы с большими данными для решения научных задач интеллектуального анализа больших данных в наблюдательной астрофизике.

В 2019 года была запущена российская национальная космическая обсерватория «Спектр-РГ» («Спектр-Рентген-Гамма») с двумя рентгеновскими телескопами на борту: eROSITA и телескопом ART-XC им.М.В.Павлинского). В середине декабря 2021 года телескоп eROSITA закончил четвертый обзор неба в рентгеновских лучах. Создание моделей и систем для интеллектуального анализа астрономических данных на основе алгоритмов машинного и глубокого обучения, применительно к уникальному рентгеновскому обзору спутника СРГ, представляет огромный интерес. Скопления галактик и рентгеновские транзитные источники являются одними из основных целей рентгеновского обзора неба eROSITA. По направлению «Космос-Д-Астрофизика», основные усилия в 2022 были сконцентрированы, во-первых, на создании моделей автоматического распознавания скоплений галактик в сырых данных изображения с микроволновых телескопов - Atacama Cosmology Telescope (ACT) и PLANCK. Во-вторых, строились модели машинного обучения для уточнения космологических расстояний далеких квазаров в спектральном каталоге SDSS DR16Q. Далее, мы приступили к созданию системы анализа космических транзиентов eРОЗИТА - был разработан ее первый рабочий прототип для полуавтоматической разметки транзиентов и разработаны алгоритмы измерения космологических расстояний (красных смещений) до данного класса источников. Последним направлением работы в 2022 году в рамках направления «Космос-Д-Астрофизика» стало совершенствование алгоритмов обучения с подкреплением для задач комбинаторной оптимизации с перспективой использования данных подходов в астрономической предметной области.

В настоящей главе отчета представлены основные результаты работ по направлению «Космос-Д-Астрофизика», полученные по следующим разделам плана работ ИКИ РАН на 2022 год:

- Исследование и разработка алгоритмов глубокого обучения для сегментации и детектирования объектов, а также для измерения их физических характеристик в данных космических и наземных обсерваторий в радио, рентгеновском и оптическом диапазонах спектра.
- Исследование и развитие алгоритмов машинного обучения для классификации астрономических объектов по комбинированным данным астрономических каталогов и фотометрическим изображениям неба.
- Исследование и разработка вероятностных моделей кросс-отождествления астрономических объектов в многоволновых обзорах неба на основе пространственной и фотометрической информации.
- Разработка технологий и систем для обработки и интеллектуального анализа больших астрономических данных в различных спектральных диапазонах.
- Исследование и разработка алгоритмов инкрементального моделирования с применением в области наблюдательной астрофизики. Исследование и

разработка моделей обучения с подкреплением для решения задач комбинаторной оптимизации.

Результаты полученные по данным направлениям описаны в разделах 2.1, 2.2, 2.3, 2.4, 2.5 настоящей главы. Подробно полученные в данном направлении результаты изложены в следующих научных публикациях [3,4,9], а также представлены в рамках 10-и устных докладов и тезисов на научных конференциях «Тихоновские чтения 2022» [21-25], «Ломоносовские чтения 2022» [26–29], и международной конференции «Астрофизика высоких энергий сегодня и завтра 2022 (HEA-2022)» [30,31].

2.1 Исследование и разработка алгоритмов глубокого обучения для сегментации и детектирования объектов, а также для измерения их физических характеристик в данных космических и наземных обсерваторий в радио, рентгеновском и оптическом диапазонах спектра

Информация по скоплениям галактик во Вселенной может быть получена из трех основных диапазонов: рентгеновского, микроволнового (радио) и оптического. Одной из целей исследования в 2022 году в рамках направления «Космос-Д-Астрофизика» стало создание нейросетевых алгоритмов распознавания скоплений галактик в микроволновом диапазоне, (основанных на анализе искажений производимых гало темной материи на карте реликтового излучения за счет эффекта Сюняева-Зельдовича (СЗ)).

В работе [9,21,30] представлен расширенный каталог источников Сюняева-Зельдовича (СЗ) по данным спутника PLANCK, основанный на соединении выборок, полученных двумя независимыми подходами к детектированию объектов в микроволновых данных: сегментация источников моделью глубокого обучения (семейства U-Net) на картах интенсивности в HFI-каналах 100-850 GHz дополнена поиском СЗ-источников на готовых картах комптоновского у-параметра (NILC). Путем сравнения полученной выборки СЗ-объектов PLANCK с наиболее полными каталогами скоплений галактик полученных по микроволновым (PSZ2, ACT) и рентгеновским (SRG/eROZITA предварительный каталог скоплений 1 года на восточном галактическом небе; MCXC по наблюдениям спутника ROSAT) обзорам неба, мы проанализировали функцию отбора скоплений галактик от массы (M_{500}), красного смещения (z) и рентгеновского потока в диапазоне 0.5-2 кэВ (F_x). Мы видим два сценария использования полученных нами расширенных каталогов СЗ-источников по данным PLANCK: (а) выборка только надежных СЗ-объектов из каталога кандидатов может использоваться непосредственно, например для поиска в ней уникальных объектов или исследования популяции скоплений, (б) весь каталог СЗ-объектов содержит значительное число ложных источников, но при этом располагает практически всеми возможными детектированными скоплениями галактик в данных PLANCK (он может использоваться совместно с другими каталогами кандидатов в скопления галактик, полученных независимо на данных в других диапазонах или другими инструментами).

В работе [31] мы продемонстрировали возможности автоматической классификации массивных рентгеновских скоплений галактик по эффекту Сюняева-Зельдовича в микроволновых данных телескопа ACT при помощи искусственных нейросетей. Для тестирования моделей использовались каталоги рентгеновских скоплений галактик обсерватории XMM-Newton в области XMM-XXL-N. Мы показали, что предложенная нейросетевая модель хорошо идентифицирует скопления на карте ACT и позволит отождествить ~70% рентгеновских скоплений двухлетнего обзора eROZITA (в области покрытия обзора ACT) с массой $M_{500} > 3 \cdot 10^{14} M_{\odot}$, и ~90% с массой $M_{500} > 4.4 \cdot 10^{14} M_{\odot}$.

2.2 Исследование и развитие алгоритмов машинного обучения для классификации астрономических объектов по комбинированным данным астрономических каталогов и фотометрическим изображениям неба

В декабре 2021 года орбитальная обсерватория «Спектр-Рентген-Гамма» завершила двухлетний обзор всего неба в рентгеновских лучах. Обзор SRG/eRosita обнаружил несколько десятков тысяч рентгеновских транзиентов (астрономических источников с сильно меняющейся яркостью) – переменных квазаров, событий приливного разрушения звезд сверхмассивными черными дыр, гамма-всплесков и др. Эти источники представляют большой интерес для науки.

Для физического анализа транзиенты необходимо поместить на трехмерную карту Вселенной, для чего требуется измерять расстояние до них (космологическое красное смещение). Методы измерения космологического красного смещения разделяются на спектральные, где красное смещение измеряется напрямую из спектра источника, и фотометрические (photo-z), где красное смещение может быть опосредованно оценено для любого источника, имеющего детектирование на изображениях неба. Последние являются на порядок менее точными, однако представляют единственную возможность измерения красного смещения для десятков тысяч транзиентов, которые найдены в рамках обзора. Наиболее точными алгоритмами photo-z являются методы, основанные на моделях машинного обучения.

В рамках работы [4,23] были построены модели photo-z рентгеновских источников на основе алгоритма случайного леса, предложенные в работе, и исследовался вопрос о применимости этих моделей для измерения красного смещения рентгеновских транзиентов (перенос обучения). Построенные модели photo-z были протестированы на выборке ~200 рентгеновских транзиентов, для которых измерено спектральное красное смещение. Прогнозы показали высокую точность по метрикам NMAD – 0,027 – и доля выбросов – 0,055. Таким образом, был получен метод, позволяющий построить трехмерную карту рентгеновских транзиентов, найденных в обзоре SRG/eRosita.

2.3 Исследование и разработка вероятностных моделей кросс-отождествления астрономических объектов в многоволновых обзорах неба на основе пространственной и фотометрической информации

Крупномасштабная структура Вселенной на больших расстояниях (красных смещениях) состоит из квазаров (активных ядер далеких галактик), поэтому для построения карты структуры Вселенной необходимо иметь точные оценки красного смещения данных объектов. Однако, существующие астрономические каталоги квазаров, такие как каталог DR16Q, содержат значительное количество объектов, имеющих ошибочные отождествление с далекими квазарами. Наибольший интерес представляют квазары, находящиеся на больших красных смещениях ($Z > 4$), поскольку именно в этой области сосредоточено большинство объектов с ошибочными спектроскопическими отождествлениями. Целью нашей работы [24] является построение и оценка модели, позволяющей отождествить квазары с ошибочными спектроскопическими оценками красного смещения в каталоге SDSS DR16Q.

Спектроскопическая оценка красного смещения вычисляется по линиям в снятом на телескоп спектре объекта. Фотометрическая оценка красного смещения производится на основе данных с фотографий объекта в нескольких цветовых фильтрах. Спектроскопические и фотометрические данные - независимы. В работе Wu & Shen, 2022, arXiv:2209.03987 ранее было выявлено по крайней мере 1900 квазаров из каталога DR16Q с ошибками в оценках красного смещения. Для классификации авторы статьи использовали спектры объектов. В свою очередь, в данной работе мы используем

фотометрические признаки и проведенную на их основе с помощью модели, представленной в нашей работе [4], фотометрическую оценку красного смещения квазаров. В качестве модели классификации был выбран градиентный бустинг, обученный на квазарах из каталогов DR14Q, VHzQ, и отдельной выборке Superset DR14Q, состоящей из объектов, не попавших в каталог квазаров DR16Q.

Качество описанной модели оценивалось на двух тестовых наборах данных. Первый — является отложенной подвыборкой тренировочного набора данных. Второй — состоит из объектов DR14Q/VHzQ и найденных в работе Wu & Shen, 2022, arXiv:2209.03987 выбросов. Метрики построенной модели на описанных тестовых наборах данных равны соответственно: (i) Recall: 0.925, Precision: 0.952, 1 - FPR: 0.994.; (ii) Recall: 0.817, Precision: 0.993, 1 - FPR: 0.993. В результате проделанной работы удалось выявить 1260 квазаров из каталога DR16Q на $Z > 3$, являющихся кандидатами на проверку существующей оценки красного смещения. Можно оценить, что ~1177 (93%) из них действительно содержат ошибки в спектральной оценке красного смещения. В работе Wu & Shen, 2022, arXiv:2209.03987 было выявлено 967 подобных объектов. При этом только 528 квазаров попали в обе выборки. Мы делаем вывод, что предложенная нами модель машинного обучения для определения ошибочных спектральных красных смещений превосходит подобные модели в литературе и позволит существенно уточнить известную карту крупномасштабной структуры далеких квазаров.

2.4 Разработка технологий и систем для обработки и интеллектуального анализа больших астрономических данных в различных спектральных диапазонах

В процессе анализа данных выполняется предварительная классификация транзиентов путем ручного сопоставления данных из многоволновых обзоров неба. Ручная разметка транзиентов, в настоящее время, позволяет классифицировать только малую часть (несколько сотен объектов) из десятков тысяч потенциально интересных событий. Целью данной работы было исследование и разработка системы для полуавтоматической классификации рентгеновских транзиентов «СРГ/еРОЗИТА» с использованием многоволновых данных обзоров неба. Данная система позволит значительно повысить эффективность работы астрофизиков и увеличить объем анализируемых данных.

Система должна соответствовать определенным требованиям: надежность хранения данных исследования, а также результатов анализа данных экспертами, соблюдение проприетарности данных исследования «СРГ/еРОЗИТА», возможность загрузки данных определенного формата в систему и сохранение взаимосвязей в данных, соответствие интерфейса требованиям экспертов. В результате обзора инструментов разметки данных было принято решение создать собственное решение, которое будет полностью отвечать вышеописанным требованиям. В процессе работы мы столкнулись со сложностями в создании схемы базы данных отражающей связи между анализируемыми источниками, а также в визуализации мультидоменных данных рентгеновских и оптических обзоров. Прототип системы разрабатывался на базе веб-фреймворка Django и СУБД SQLite, в итоговой версии используется более мощная СУБД PostgreSQL. Для автоматизации классификации используются predefined критерии, которые отражают знания астрофизиков о классах транзиентов «СРГ/еРОЗИТА». На основе этих критериев осуществляется предварительная разметка объектов.

В 2022 году в рамках направления «Космос-Д-Астрофизика» была создана и протестирована астрофизиками система [28,3] для полуавтоматической классификации рентгеновских транзиентов «СРГ/еРОЗИТА». На данный момент система активно используется астрофизиками ИКИ РАН, занимающимися исследованием рентгеновских транзиентов «СРГ/еРОЗИТА». В дальнейшем планируется интеграция моделей машинного обучения в целях усовершенствования автоматической разметки.

2.5 Исследование и разработка алгоритмов инкрементального моделирования с применением в области наблюдательной астрофизики. Исследование и разработка моделей обучения с подкреплением для решения задач комбинаторной оптимизации

В 2022 году в рамках направления «Космос-Д-астрофизика» мы исследовали нейросетевые алгоритмы обучения с подкреплением (RL) в рамках решения NP-трудных задач комбинаторной оптимизации с ограничениями. Подобные подходы находят себе применение и в астрофизической предметной области, но требуют для создания точных RL-алгоритмов хорошего моделирования среды, в которой обучаются данный тип нейросетевых моделей. В 2022 мы поставили эксперименты по сравнению точных, эвристических моделей комбинаторной оптимизации (Gurobi, Or-Tools), с одной стороны, и нейросетевых алгоритмов обучения с подкреплением (JAMPR), с другой, на хорошо известных задачах исследования операций - разновидности задач коммивояжера (комбинаторной оптимизации с ограничениями из области логистики), где создание среды для обучения алгоритма не представляет сложности. В дальнейшем мы планируем перенести полученные нами знания о сравнении различных подходов комбинаторной оптимизации на задачи возникающие в области астрофизики (например, в задаче массового кросс-отождествления большого числа источников из многоволновых каталогов, или в других астрофизических приложениях RL-подходов).

Сама по себе задача построения оптимальных маршрутов вывоза и доставки товаров является одной из наиболее перспективных задач в условиях глобального роста городского населения. Хотя такого рода проблемы с небольшим размером могут быть решены с помощью различных классических подходов, быстрый (или работающий в реальном времени) оптимизатор маршрута в условиях ограничений реального мира (таких как ограничения пропускной способности и временных окон) для задач среднего и большого размера все еще остается сложной задачей. В нашей работе [22,25,26] мы впервые успешно применили подход глубокого обучения с подкреплением (модифицированная модель JAMPR) для решения логистической задачи забора и доставки с ограничениями по вместимости транспортных средств и по временным окнам (CPDPTW). Мы получили робастную нейросетевую модель, которая дает быстрое оптимальное решение для задач малого и среднего размера и дает быстрое субоптимальное решение для задач большого (>200) размера.

Заключение

Представленные в настоящей главе результаты показывают, что по направлению «Большие данные в космических исследованиях: астрофизика» проведенные работы соответствуют установленной в плане работ ИКИ РАН на 2022 г. тематике. Таким образом, работы выполнены полностью в соответствии с планом.

При этом получены следующие основные результаты:

- Разработана нейросетевая модель для детектирования рентгеновских скоплений галактик в данных телескопа АСТ. Разработана нейросетевая модель для детектирования скоплений галактик в данных HFI-инструмента телескопа PLANCK. Путем комбинации результатов нейросети и стандартного детектирования СЗ-источников на у-картах PLANCK получен наиболее полный комбинированный каталог кандидатов в скопления галактик в микроволновом диапазоне и измерена его полнота и точность относительно рентгеновского

каталога скоплений ePOZITA. Мы продемонстрировали возможности нейросетевой классификации массивных рентгеновских скоплений галактик по эффекту Сюняева-Зельдовича в микроволновых данных телескопа АСТ.

- Созданы модели машинного обучения для уточнения карты далеких $z > 4$ квазаров по данным каталога SDSS DR16Q и показана их точность при отборе объектов с ошибочными красными смещениями далеких квазаров.
- Создана и протестирована астрофизиками система для полуавтоматической разметки рентгеновских транзиентов «СРГ/ePOZITA» и разработаны модели машинного обучения для измерения фотометрических красных смещений для внегалактических рентгеновских транзиентов
- Созданы модели глубокого обучения с подкреплением (на основе модификации JAMPR) для решения задач комбинаторной оптимизации с ограничениями (логистическая проблема забора и доставки с ограничениями мощности и временных окон, CPDPTW). Полученная модель дает быстрое оптимальное решение для графов малого и среднего размера и дает быстрое субоптимальное решение для графов задач большого (> 200) размера. В дальнейшем мы планируем применить данные подходы комбинаторной оптимизации в астрофизической предметной области.

Таким образом, работы по направлению «Астрофизика» темы Космос-Д в 2022 году были выполнены в полном объеме.

ГЛАВА 3 Разработка новых подходов и методов организации сверхбольших архивов данных спутниковых наблюдений Земли и доступа к ним для обеспечения эффективной обработки и анализа данных ДЗЗ для решения научных и прикладных задач

Введение

Настоящая глава посвящена описанию основных результатов, полученных в рамках проведения работ по направлению «Разработка новых подходов и методов организации сверхбольших архивов данных спутниковых наблюдений Земли и доступа к ним для обеспечения эффективной обработки и анализа данных ДЗЗ для решения научных и прикладных задач». В 2022 году по нему велись работы по следующим задачам плана работ по теме «Космос-Д»:

- Разработка новых методов и алгоритмов организации динамического блочного доступа к сверхбольшим разнородным архивам данных спутниковых наблюдений Земли для повышения эффективности распределенной многопоточковой обработки информации, поступающей от различных систем ДЗЗ.
- Разработка подходов и методов организации динамического блочного доступа к сверхбольшим разнородным архивам данных спутниковых наблюдений Земли для повышения эффективности инструментов онлайн доступа и анализа спутниковых данных и результатов их обработки.
- Развитие с использованием созданных методов и подходов различных специализированных научных информационных систем, в том числе. ИС ВЕГА-Лес (комплексный научный мониторинг лесных ресурсов России), ИС STS (комплексный научный мониторинг различных явлений в пограничных морях России), ИС VSV (комплексный научный мониторинг вулканической активности Камчатки и Курил)

Для проведения эффективной параллельной обработки и анализа долговременных рядов спутниковых наблюдений по большим территориям оптимальным является использование данных, представленных в виде регулярного пространственного разбиения, т.е. использование «тайлового» подхода. Это позволяет эффективно распараллелить процесс обработки данных, реализовать кэширование фрагментов данных, что, в свою очередь, позволяет существенно снизить нагрузку на централизованные вычислительные ресурсы и каналы связи. В настоящее время такой подход используют практически все системы, обеспечивающие массовый доступ пользователей к различным пространственным данным. Для реализации этого подхода, как правило, используются заранее подготовленные данные в требуемых пространственных разбиениях. Такой вариант достаточно хорошо подходит в тех случаях, когда необходимо реализовать работу с ограниченным набором различных типов пространственных данных и имеются достаточные для подготовки и хранения требуемых данных ресурсы. Это решение может быть эффективно использовано также тогда, когда хранение того или иного типа спутниковых данных уже реализовано в виде регулярного пространственного разбиения, которое удовлетворяет требованиям конкретных процедур по их анализу и обработке.

В то же время использование заранее подготовленных данных в фиксированных пространственных разбиениях оказывается неприемлемым решением в тех случаях, когда одни и те же типы спутниковых данных необходимы в разных пространственных разбиениях и проекциях. При этом такая ситуация часто возникает как при реализации процедур обработки спутниковых данных, так и при обеспечении к ним доступа пользователей. Спутниковые данные, как правило, характеризуются достаточно большими объемами, поэтому подготовка и хранение одних и тех же типов спутниковых

наблюдений одновременно в разных пространственных разбиениях требует неоправданно много как вычислительных ресурсов, так и ресурсов хранения данных. В частности, описанная ситуация возникает при необходимости совместной обработки разных типов спутниковой информации, хранящихся в разных пространственных разбиениях, и в тех случаях, когда для разных задач по анализу и обработке какого-то типа данных оптимальные для их проведения разбиения существенно отличаются.

Для того чтобы использовать все основные преимущества блочного доступа к данным и избежать излишних накладных расходов, в 2019 году в рамках работ по теме «Космос-Д» был предложен новый подход, основанный на динамическом формировании блоков данных для их обработки и визуализации. Такой подход позволяет формировать требуемые наборы данных различных типов спутниковой информации в таком пространственном разбиении и формате представления, которые были бы оптимальны для проведения конкретных задач по их обработке и анализу. В 2020 году активно велись работы по созданию эффективной программной реализации для этой технологии, пробному ее применению для решения реальных задач и разработке методики выбора оптимальных параметров пространственных разбиений. В 2021 году разработанная технология была успешно внедрена в эксплуатацию для задач построения безоблачных атмосферно скорректированных ежедневных композитов по данным прибора КМСС, установленного на спутниках серии «Метеор-М».

В 2022 году были продолжены работы по совершенствованию технологии динамического блочного доступа и применению ее для эффективной реализации задач обработки долговременных рядов спутниковых наблюдений по большим территориям. Существенные улучшения были внесены в базовое программное обеспечение, в частности, была реализована новая подсистема для предоставления в обработку данных, уже находящихся в архивах в требуемом пространственном разбиении. Также был реализован ряд новых подходов к организации системы обработки данных прибора КМСС, использующей технологию динамического блочного доступа. В частности, для детектирования облачности и теней на монохроматических изображениях приборов серии КМСС была применена сверточная нейронная сеть U-Net, что в перспективе может существенно ускорить проведение обработки. Также технология динамического блочного доступа стала применяться для предоставления данных целого ряда других спутниковых приборов, установленных на российских спутниках метеорологического и природно-ресурсного назначения.

Другим важным применением разработанной в ИКИ РАН технологии динамического блочного доступа к данным стало создание картографического веб-интерфейса, использующего тайловую схему формирования изображений. С одной стороны это позволило предложить пользователям более удобный и распространенный в настоящее время стиль работы с пространственными данными, а с другой, благодаря динамическому формированию тайлов изображений интерфейс сохранил свои уникальные функциональные возможности, которыми не обладают системы доступа, работающие с данными в фиксированных пространственных разбиениях. Это относится как к поддержке широкого спектра «виртуальных» информационных продуктов, т.е. продуктов, динамически формируемых на основе базовых продуктов, так и к реализации большого набора интерактивных инструментов для анализа и обработки спутниковых данных непосредственно в интерфейсе. Работы по созданию обновленного картографического интерфейса были начаты в 2020 году, а в 2021 году он был внедрен в эксплуатацию в рамках системы «Вега-Science», а также в ряде других систем, входящих в проект «Созвездие-Вега».

В 2022 году были продолжены работы по развитию, как самого картографического интерфейса, так и модуля растеризации, непосредственно отвечающего за формирование тайлов изображений. В частности, была реализована поддержка большого числа различных картографических проекций, что является уникальной возможностью

описываемого интерфейса. Также была реализована новая подсистема профилирования времени выполнения запросов на получение тайлов и времен выполнения отдельных их этапов. Это позволило лучше выявлять причины задержек с формированием изображений, а в некоторых случаях, и вмешиваться в выполнение соответствующих процессов.

В текущем году также активно велись работы по разработке новых подходов к созданию различных программных компонент, используемых для построения специализированных научных информационных систем, в том числе: ИС ВЕГА-Лес (комплексный научный мониторинг лесных ресурсов России), ИС STS (комплексный научный мониторинг различных явлений в пограничных морях России), ИС VSV (комплексный научный мониторинг вулканической активности Камчатки и Курил). В частности, была разработана новая система управления комплексными задачами по обработке спутниковых данных. И хотя эта система позиционируется как универсальная, основным ее применением на текущий момент является подготовка данных для предоставления их в рамках вышеперечисленных информационных систем. С использованием этой системы были разработаны сервисы для эффективного предоставления доступа к сериям данных сверхбольших распределенных архивов. Одним из наиболее актуальных применений на текущий момент является формирование рядов значений, полученных по данным геостационарного спутника Himawari, для мониторинга вулканической активности Камчатки и Курил. Также была разработана и полностью внедрена технология создания безоблачных композитных изображений по данным серии Sentinel-2, позволяющая на постоянной основе создавать качественные композитные изображения в высоком разрешении на всю территорию России. Такие продукты являются очень востребованными в большинстве разрабатываемых в отделе «Технологии спутникового мониторинга» информационных систем, а также используются в качестве исходных данных для получения производных тематических продуктов.

Полученные в рамках работ 2022 года результаты по данному направлению изложены в следующих публикациях [5,10-12,14-20].

3.1 Разработка новых методов и алгоритмов организации динамического блочного доступа к сверхбольшим разнородным архивам данных спутниковых наблюдений Земли для повышения эффективности распределенной многопоточковой обработки информации, поступающей от различных систем ДЗЗ

3.1.1 Модернизация программной реализации технологии динамического блочного доступа

В 2022 году продолжались работы по совершенствованию программной реализации технологии динамического блочного доступа к данным. В частности, был поддержан механизм контекстной фильтрации списка предоставляемых под обработку блоков данных, который позволяет выбрать только минимально необходимый набор данных и тем самым снизить нагрузку на вычислительные ресурсы. Такой механизм актуален в тех случаях, когда в архиве содержатся частично дублирующие друг друга данные, или когда для получения требуемого результата достаточно обработать только подмножество имеющихся в архивах данных. При этом логика выбора минимального набора требуемых блоков данных зависит от конкретной задачи по их обработке. Для реализации этого механизма после получения полного списка блоков и соответствующих им метаданных может быть вызвана процедура из указанной в конфигурации задания библиотеки. Также в 2022 году были внесены изменения в программное обеспечение, предназначенное для протоколирования всех операций с данными и оперативное детектирование сбоев при их выполнении. В специализированную базу данных теперь заносится не только информация о результатах операций, но также и информация на

каком из серверов распределенного кластера они выполнялись, что позволяет быстро локализовать сервер, на котором возникли ошибки.

В рамках рассматриваемой технологии после формирования списка требуемых блоков данных запросы на их получение поступают на выделенный сервер балансировки. На этом сервере располагается разработанный в ИКИ РАН сервис, позволяющий перенаправить эти запросы на один из серверов соответствующего центра данных распределенного архива. До недавнего времени каждому из этих серверов мог быть задан весовой коэффициент, соответствующий его производительности, в соответствии с которым запросы распределялись по имеющимся в том или ином центре серверам. На практике на используемых для этой цели серверах могут выполняться и другие задачи, требующие вычислительных ресурсов, что приводит к неравномерной загрузке серверов и, как следствие, к снижению скорости подготовки блоков данных. В обновленной реализации сервиса балансировки на периодической основе производится запрос информации о текущей загрузке каждого из серверов, что позволяет более равномерно распределить на них нагрузку и быстрее получить заданный набор блоков данных.

3.1.2 Разработка новых подходов к организации обработки для получения ежедневных безоблачных композитов по данным прибора КМСС

В рамках работ, нацеленных на обеспечение возможности автоматической тематической обработки данных с российских спутниковых систем, в ИКИ РАН была разработана автоматическая технология потоковой обработки изображений с приборов серии КМСС, находящихся на борту спутников серии «Метеор-М». Данные КМСС характеризуются хорошо востребованным набором показателей пространственного разрешения и повторяемости наблюдений, благодаря чему территория России полностью покрывается результатами съёмки с пространственным разрешением 60 метров в течение 1-2 дней при работе двух приборов. Однако до недавнего времени потенциал этих данных не мог быть по-настоящему использован ввиду ограничений совместимости со стандартными методами, используемыми для предварительной и тематической обработки данных имеющихся спутниковых систем для оптико-электронной съёмки. Разработанные в рамках упомянутой выше технологии методы позволили автоматически устранять локальные и систематические искажения географической привязки изображений, детектировать облачность и тени от нее, а также выполнять радиометрическую коррекцию данных. Это стало возможным благодаря проведению совместной обработки данных КМСС с хорошо калиброванными данными прибора MODIS, установленного на спутниках Terra и Aqua. Для реализации совместной обработки данных этих приборов, отличающихся как по пространственному разрешению, так и по схемам хранения в архивах, для выборки данных была использована созданная в ИКИ РАН технология динамического блочного доступа.

В 2021 году разработанная по представленной технологии комплексная процедура обработки была внедрена в эксплуатацию. В настоящее время она позволяет получать ежедневные интерполированные безоблачные композитные изображения по данным КМСС по территории сельскохозяйственного пояса России. Благодаря высокому качеству получаемых информационных продуктов, они могут быть использованы для решения задач мониторинга и количественной оценки характеристик растительного покрова, в том числе, и в оперативном режиме. Пример ежедневного безоблачного изображения по данным прибора КМСС приведен на рисунке 3.1.2.1.

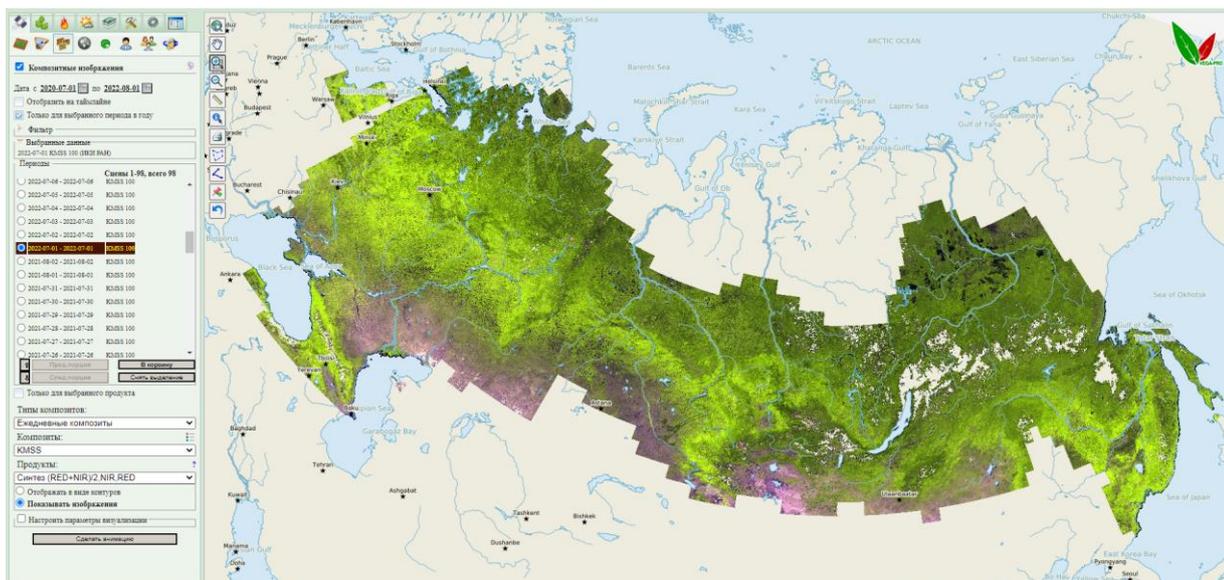


Рисунок 3.1.2.1 - Пример ежедневного безоблачного изображения КМСС за 1 июля 2022

В 2022 году активно велись работы по повышению качества получаемых продуктов, а также по оптимизации самих процедур их обработки, необходимость которой связана с большими вычислительными затратами, необходимыми для проведения комплексной процедуры обработки. Так, в настоящее время для получения одного оперативного безоблачного интерполированного композита необходимы вычислительные ресурсы как минимум 6 выделенных серверов обработки данных. В тех же случаях, когда возникает задача повторной обработки данных за предшествующие периоды времени с использованием более совершенных алгоритмов, нагрузка на сервера может возрастать в разы. Ниже рассматриваются два новых подхода к организации рассматриваемой комплексной обработки, которые в перспективе могут позволить радикально снизить требования на вычислительные ресурсы, необходимые для ее проведения.

Данные прибора КМСС, установленного на российских спутниках серии «Метеор-М», используемые для проведения рассматриваемой обработки, принимаются и архивируются в трех центрах ФГБУ НИЦ «Планета», расположенных в Москве, Новосибирске и Хабаровске, что позволяет оперативно получать данные по всей территории России. Существенно, что этих в центрах могут приниматься и поступать в архивы одни и те же сеансы (витки) спутниковых данных, а часть сеансов также поступает в центр в Москве в режиме записи. При этом данные одного и того же сеанса, поступившего в архивы из разных источников зачастую не являются идентичными как по пространственному покрытию, так и по качеству их географической привязки и калибровки. Это приводит к тому, что в рамках рассматриваемой процедуры обработки блоки данных каждой из версий таких сеансов обрабатываются отдельно. Важно отметить, что в результате проведения обработки примерно только половина исходных блоков данных используется для получения итогового продукта. Основной причиной этого является высокая облачность данных, но и иногда это связано и с плохой географической привязкой исходных данных. Поэтому для получения максимально полных данных было решено использовать блоки по всем вариантам одних и тех же сеансов, что приводит к тому, что около трети блоков данных обрабатываются два и более раз.

На рисунке 3.1.2.2 представлена схема, на которой градациями серого цвета представлена степень дублирования блоков данных в текущем году по обрабатываемой территории. При этом черному цвету соответствует ситуация, когда блок данных всегда дублировался, а белому то, что он был в единственном экземпляре. Наибольшее

дублирование блоков имеет место на востоке России, что связано с тем, что почти все принятые там сеансы поступают в архивы также и в режиме записи. Дублирование данных в европейской части России также связано и с тем, что иногда одни и те же сеансы принимаются сразу на двух станциях приема в Европейском центре НИЦ «Планета» в Москве. Для снижения требуемых для проведения обработки данных прибора КМСС вычислительных ресурсов было решено рассмотреть вопрос исключения из обработки наименее подходящих из дублирующих сеансов. Для этого в рамках использования реализованного в этом году механизма контекстной фильтрации была реализован метод, позволяющий на основе сравнения пространственного покрытия таких сеансов и источника их поступления выбрать из них наиболее полный. В настоящее время производится анализ того, насколько исключение дублирующих сеансов данных скажется на качестве получаемых композитных изображений. В случае, если негативный эффект окажется несущественным, контекстная фильтрация будет внедрена в эксплуатацию, что позволит сократить требования к вычислительным ресурсам примерно на треть.

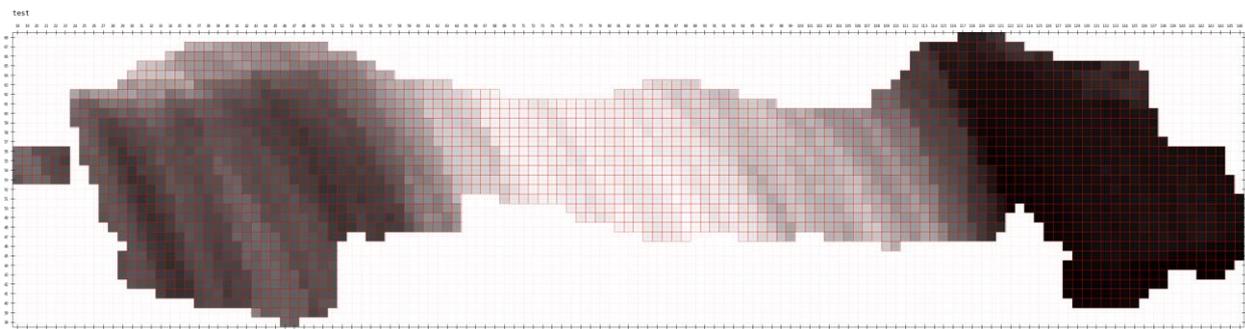


Рисунок 3.1.2.2 – Степень дублирования блоков данных по одним и тем же сеансам данных прибора КМСС

В качестве еще одного перспективного подхода, который потенциально может существенно ускорить выполнение обработки, рассматривается использование технологий Deep Learning для детектирования облачности и теней на изображениях прибора КМСС. В текущей реализации для выполнения этой операции используется достаточно сложный алгоритм, требующий весьма значительных вычислительных ресурсов. Для реализации этого подхода была выбрана сверточная нейронная сеть U-Net, которая считается одной из стандартных архитектур для задач сегментации изображений, когда нужно не только определить класс изображения целиком, но и сегментировать его области по классам. Одним из основных ее преимуществ является возможность использования сравнительно небольшого количества обучающих данных для достижения хороших результатов. Архитектура этой сети включает в себя 23 сверточных слоя и состоит из стягивающего пути для захвата контекста и симметричного расширяющегося пути, который позволят осуществить точную локализацию. Пример топологии сети U-Net представлен на рисунке 3.1.2.3. На нем каждый синий квадрат соответствует многоканальной карте свойств, для которого сверху приведено количество каналов, а в нижнем левом краю – размер признаков, белые квадраты представляют собой копии карты свойств, а стрелки обозначают различные операции.

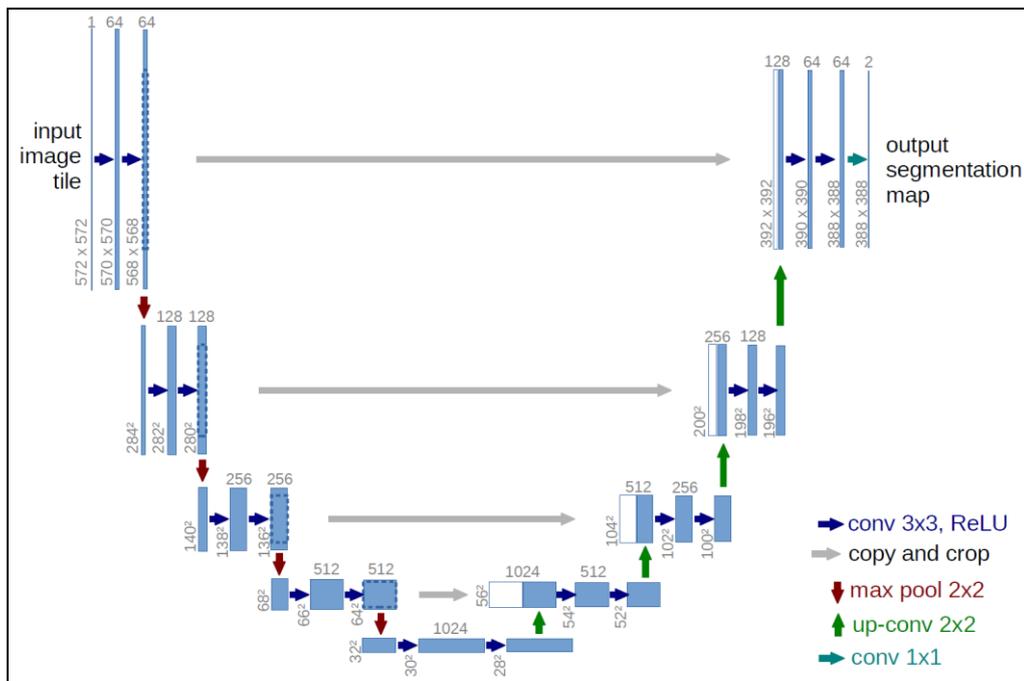


Рисунок 3.1.2.3 Архитектура U-net (пример изображения с разрешением 32×32 пикселя)

Для реализации сверточной нейронной сети U-net был выбран язык программирования Python, библиотека Keras, обеспечивающая взаимодействие с искусственными нейронными сетями, и нейросетевая библиотека TensorFlow. В качестве обучающей выборки были использованы фрагменты данных прибора КМСС, полученные при помощи технологии динамического блочного доступа, и результаты детектирования облачности и теней, полученные при помощи уже реализованного алгоритма. Отдельно производится обучение по красному и ближнему инфракрасному каналам прибора КМСС. На рисунке 3.1.3.4 приведен график обучения реализованной нейронной сети.

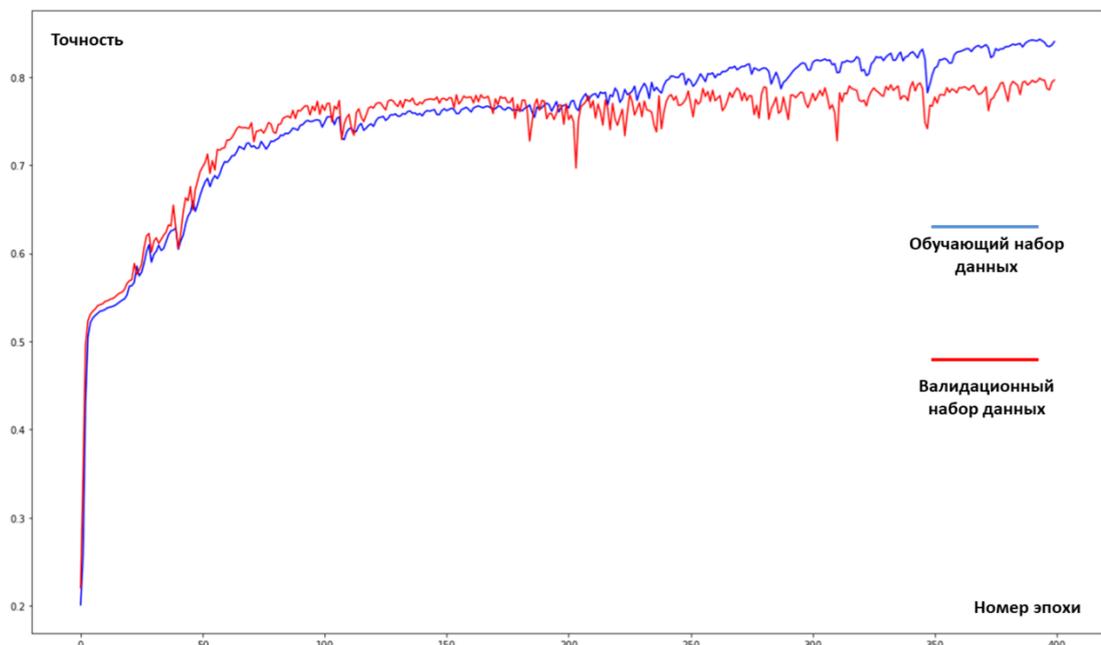


Рисунок 3.1.2.4 - График обучения нейронной сети

В настоящее время работы по совершенствованию и обучению сети еще не завершены. Тем не менее, обученная нейронная сеть уже сейчас используется в работе комплексной процедуры обработки в качестве вспомогательного инструмента, используемого для грубой экспресс оценки облачности на изображениях прибора КМСС. В том случае, если в результате дальнейшего обучения сети точность ее результатов станет приемлемой, она будет использоваться в качестве основного инструмента для детектирования облачности и теней, что позволит в разы ускорить выполнение процедуры атмосферной и геометрической коррекции данных, которая вносит основной вклад во время выполнения всей комплексной процедуры обработки.

3.1.3 Подсистема подготовки блоков данных для проведения дальнейшей их обработки

Для предоставления данных для проведения их дальнейшей обработки в 2022 году была разработана новая программная подсистема. Первоначально она была реализована для выгрузки блоков данных, полученных в результате проведения обработки данных прибора КМСС, но затем ее функционал был расширен, и в настоящее время она позволяет подготавливать для обработки любые типы файлов данных, находящихся в архивах ЦКП «ИКИ-Мониторинг». Основным ее отличием от используемых ранее аналогов является гибкость настроек и поддержка большого числа различных вариантов именования получаемых файлов данных, включая использование требуемых структур директорий. Ниже сначала приводятся общие сведения об организации архивов данных «ЦКП ИКИ-Мониторинг», а затем описывается разработанная подсистема для предоставления хранящихся в них файлов данных для проведения их обработки.

Система ведения архивов спутниковых данных ЦКП «ИКИ-Мониторинг» реализована на базе унифицированной технологии ведения сверхбольших распределенных архивов разнородных спутниковых данных UNISAT. В рамках использования этой технологии метаданные, т.е. информация об имеющихся в архивах данных и их атрибуты, сохраняются в базу данных unisat, а сами файлы данных помещаются в файловое хранилище. Структура БД unisat включает в себя две основные таблицы:

- fragments – описание фрагментов спутниковых продуктов;
- fragment_products – описание информационных продуктов, относящихся к фрагментам, которым соответствуют файлы данных, которые в основном представлены в виде спутниковых изображений в формате GeoTIFF.

Хранение файлов спутниковых данных в архивах ЦКП «ИКИ-Мониторинг» в каждом из центров реализуется на основе использования распределенных файловых хранилищ, состоящих из UNIX серверов и специализированных систем хранения (NAS), объединенных в единое логическое пространство при помощи протокола NFS. На текущий момент в ИКИ РАН задействовано более 30 серверов хранения. Файлы архива хранятся поддиректориях, в каждой из которых по умолчанию может быть до ста тысяч файлов. Имена поддиректорий и самих файлов формируются на основании идентификатора соответствующей записи в таблице product_fragments. К примеру, записи с идентификатором 12345678, описывающей GeoTIFF файл, соответствует относительный путь к файлу “00123/45678.tif”.

За пополнение и изменение данных каждого из архивов отвечает выделенный сервер архивации. На нем располагается базовая директория хранилища файлов архива, в которую линкуются поддиректории с данными, находящимися на различных сетевых разделах. При этом сами разделы должны быть присоединены к этому серверу в режиме чтения и записи по протоколу NFS.

На серверах, которым необходим только доступ к файлам в архиве, присоединяется по NFS базовая директория архива, а требуемые сетевые разделы присоединяются по протоколу NFS в режиме только чтения. На серверах, которые обеспечивают формирование данных для занесения их в архив, присоединяется по NFS специальная директория с сервера архивации, которая всегда указывает на текущий раздел пополнения данными, а также сам этот раздел и следующие за ним. Это позволяет избежать лишнего копирования файлов и помещать их сразу на требуемый сетевой раздел, после чего в рамках процедур архивации они перемещаются внутри этого раздела в поддиректории с данными, что является нетребовательной к ресурсам операцией. Схема организации распределенного файлового хранилища спутниковой информации приведена на рисунке 3.1.3.1.

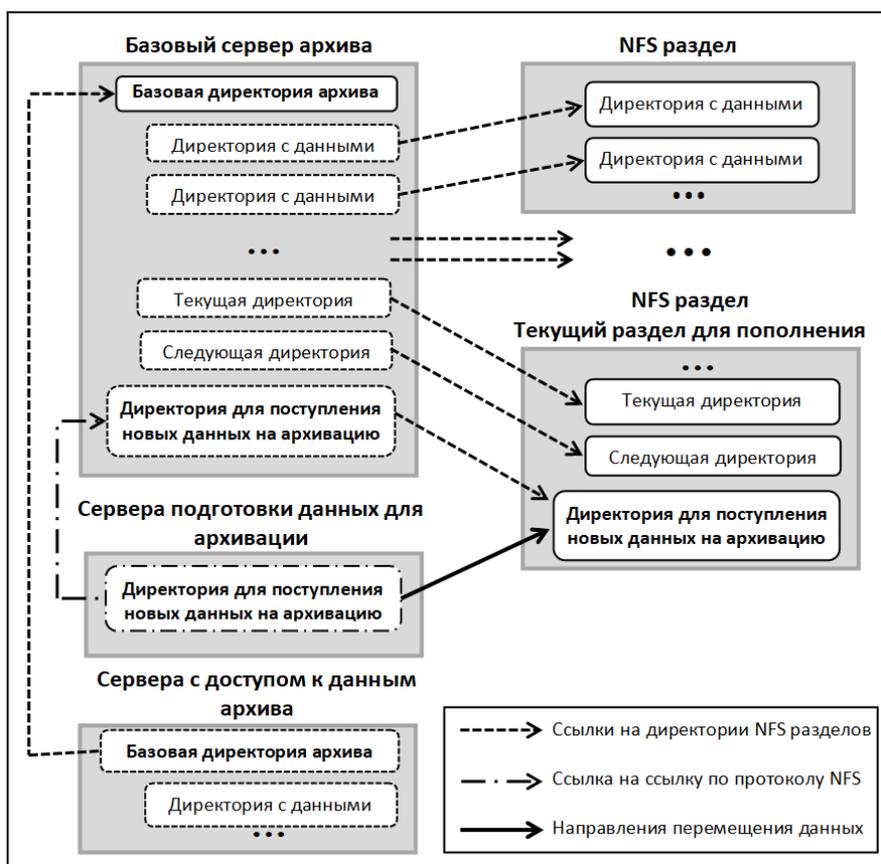


Рисунок 3.1.3.1 - Схема организации распределенного файлового хранилища

Процедура получения файлов данных из заданного распределенного архива может выполняться на серверах, к которым по протоколу NFS подсоединены базовая директория соответствующего архива и все его локальные (находящиеся в пределах одного центра данных) разделы. Функциональная схема подсистемы подготовки данных приведена на рисунке 3.1.3.2. На вход процедуре подготовки данных поступает конфигурационный файл, содержащий все необходимые параметры требуемых данных, включая идентификатор распределенного архива данных, центры данных, спутники, схему представления результатов и т.п. На первом этапе выполнения процедуры производится запрос к каталогам архивов указанных в конфигурации центров данных для получения списка требуемых файлов. Необходимость в этом связана с тем, что в каждом из узлов распределенного архива содержится информация обо всех имеющихся в нем данных, но информация об именах файлов данных находится локально. После этого производится непосредственно получение данных из файловых хранилищ. В том случае, когда это

хранилище является локальным, производится операция копирования или линкования файла по протоколу NFS, а данные из удаленных центров запрашиваются в параллельном режиме по протоколу http при помощи программы aria2c. Помимо самих файлов данных, в директорию для выкладки файлов могут также помещаться файлы, содержащие их метаданные. Информация о полученных файлах или ошибках при их получении сохраняется в специализированную базу данных.

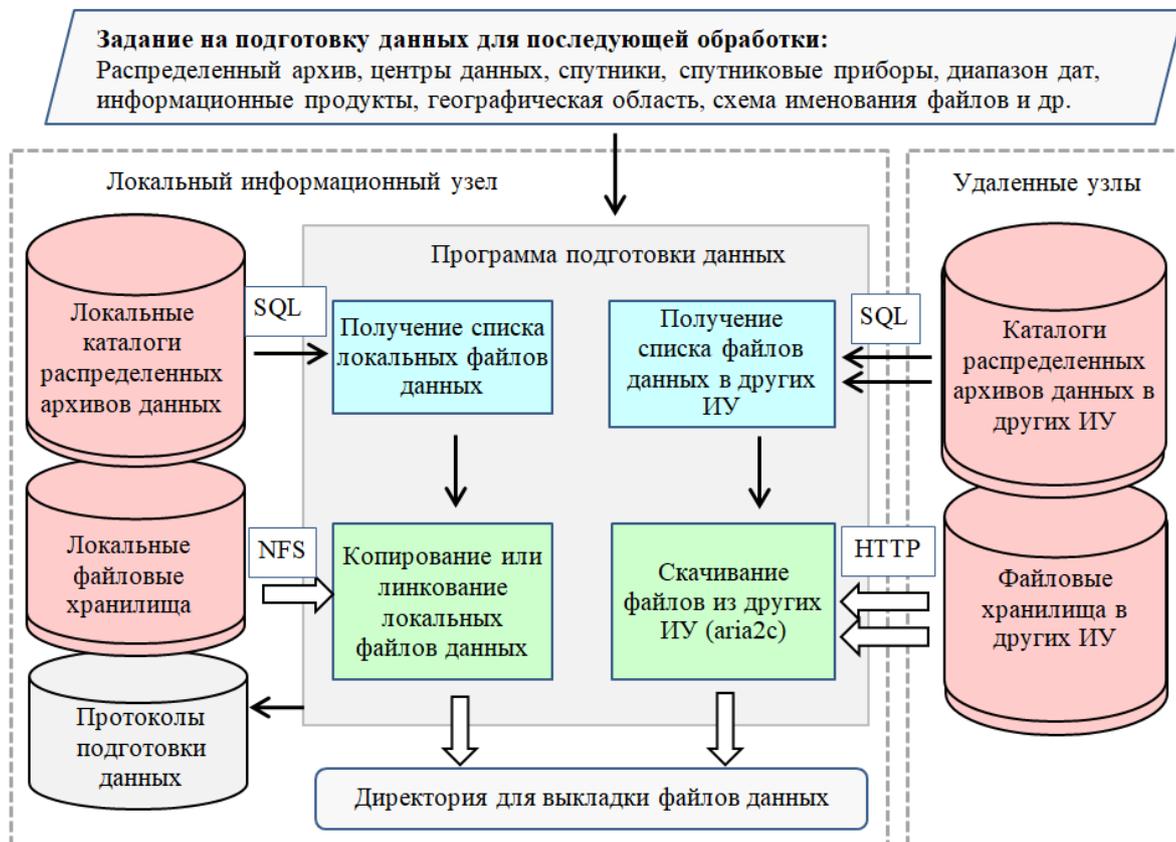


Рисунок 3.1.3.2 - Функциональная схема подсистемы подготовки данных

3.2 Разработка подходов и методов организации динамического блочного доступа к сверхбольшим разнородным архивам данных спутниковых наблюдений Земли для повышения эффективности инструментов онлайн доступа и анализа спутниковых данных и результатов их обработки

3.2.1 Модернизация картографического веб-интерфейса, использующего динамическое формирование тайлов изображений

За последние годы значительно изменились представления пользователей о комфортной работе с картографическими данными. При этом также существенно возросли требования на максимальное количество пользователей, одновременно работающих с картографическими веб-интерфейсам. Все это привело к необходимости обновления интерфейсов, предназначенных для обеспечения доступа к данным архивов ЦКП «ИКИ-Мониторинг». По этим причинам в 2021 году была внедрена в эксплуатацию новая версия картографического веб-интерфейса, в котором спутниковые изображения в интерфейсе формируются в виде тайлов. Получение изображений в виде набора тайлов в

настоящее время используется в большинстве систем визуализации пространственных данных, так как благодаря эффективному кэшированию отдельных тайлов изображения достигается более комфортный для пользователя режим работы с пространственной информацией. Однако, как правило, для его реализации используется предварительно насчитанный набор тайлов, соответствующих ограниченному набору типов информационных продуктов.

При помощи реализованных в ИКИ РАН картографических веб-интерфейсов пользователи получают доступ к широкому спектру различных информационных продуктов, включая «виртуальные», которые формируются по запросу пользователя на основе использования базовых информационных продуктов и предоставляются ему в заданной проекции. Кроме этого реализованы различные интерактивные инструменты, позволяющие пользователю проводить анализ и обработку данных непосредственно в интерфейсе. Поэтому при реализации новой версии интерфейса для формирования тайлов изображений была использована технология динамического блочного доступа.

В 2022 году были проведены работы по поддержке большего числа различных проекций в обновленном картографическом веб-интерфейсе. Сложность заключалась в том, что используемая в программной реализации интерфейса открытая библиотека «leaflet.js» по умолчанию поддерживает лишь небольшой набор проекций, основными среди которых являются географическая проекция (EPSG4326) и проекция Меркатора (EPSG3857). При этом для решения целого ряда исследовательских задач пользователям интерфейсов могут быть необходимы и другие проекции. В частности, для исследования Арктики наиболее подходят полярные проекции, а для сохранения площадей анализируемых объектов – равновеликая коническая проекция Альберса. Для поддержки требуемого функционала была использована библиотека преобразования картографических проекций «proj4leaflet.js». Данная библиотека позволяет организовать работу с картой в разных проекциях, если новая система координат будет описана строкой в формате PROJ.4 и для неё будет задан список разрешений для каждого масштабного уровня карты. Однако для поддержки ряда проекций, в частности, проекции Альберса, этого оказалось недостаточно, и потребовалась существенная доработка модулей самой библиотеки «leaflet.js». На рисунке 3.2.1.1 приведен пример использования картографического веб-интерфейса в проекции Альберса.



Рисунок 3.2.1.1 - Пример использования в интерфейсе проекции Альберса

3.2.2 Модернизация модуля растеризации изображений

Создание тайлового веб-интерфейса потребовало разработки нового, более эффективного программного решения для формирования отображаемых в нем тайлов изображений, которое позволило бы при сохранении имеющегося функционала существенно повысить скорость выполнения операций. С этой целью был разработан новый универсальный модуль растеризации изображений. Он полностью реализован на языке программирования Python и использует для работы с изображениями свободно распространяемый программный пакет GDAL. Необходимые вычисления в процессе формирования изображений производятся с использованием высокоэффективной библиотеки NumPy, являющейся свободно распространяемой альтернативой пакета MATLAB. В результате, несмотря на использование интерпретируемого языка программирования, основные операции по обработке данных производятся с высокой эффективностью, характерной для компилируемых языков программирования.

В 2022 году велись работы по устранению выявленных в процессе эксплуатации недостатков в работе модуля растеризации, а также по повышению его функционала и быстродействия. В частности, были устранены возникавшие при работе с некоторыми проекциями ошибки формирования изображений, пересекающих меридиан 180 градусов. Также была устранена выявленная в ряде случаев пикселизация изображений, которая была связана с неверным выбором уровня в пирамиде разрешений. На основе метода `ndimage.correlate`, реализованного в библиотеке с открытым исходным кодом SciPy, были реализованы матричные фильтры для повышения резкости изображений или их размытия. Кроме этого был проведен детальный анализ времени выполнения различных операций, используемых для формирования разных типов изображений, и выявлены фрагменты программного кода, которые нуждались в доработке. Модернизация программного кода позволила в ряде случаев существенно повысить скорость формирования требуемых изображений.

3.2.3 Профилирование времени выполнения запросов на получение тайлов изображений

Внедрение в эксплуатацию тайлового веб-интерфейса привело к резкому увеличению числа запросов на формирование изображений, которые выполняются на выделенном кластере серверов. При этом большая задержка с получением хотя бы одного из тайлов изображения в картографическом интерфейсе визуально очень заметна для пользователей. Это привело к необходимости реализации контроля выполнения запросов на получение изображений на качественно более высоком уровне. Для анализа причин задержек при выполнении таких запросов необходимо протоколировать не только время их выполнения, но и вклад, вносимый в это время отдельными операциями, такими как запрос метаданных в БД, копирование файлов, перевод изображений в требуемую проекцию и т.п. Наиболее критичными требованиями к программной организации профилирования запросов является то, что она не должна замедлять их выполнение и тем более быть фактором, увеличивающим вероятность сбоев при их выполнении.

В 2022 году была реализована новая подсистема профилирования выполнения запросов на получение изображений. Для того, чтобы избежать зависимости от состояния и загруженности базы данных, используемой для этой цели, информация о времени выполнения запроса и отдельных его операций сначала заносится в локальный временный файл. Уникальность имен таких файлов достигается использованием при формировании его имени не только времени, но и идентификатора процесса. При этом запись совсем небольшого файла (как правило, размером в пределах одного килобайта) на локальный диск не приводит к существенным накладным расходам. На каждом сервере формирования изображений периодически запускается процедура, которая вычитывает

информацию из появившихся файлов и заносит ее в базу данных. И даже если база данных на какое-то время окажется недоступной, это никак не повлияет на выполнение самих запросов.

Получаемая информация о выполнении запроса состоит из его параметров и наборов показателей для отдельных замеров времени (таймеров). В программном коде процедуры формирования изображений перед началом выполнения той или иной операции производится вызов, осуществляющий запуск таймера с указанным текстовым идентификатором, а после – вызов для его остановки. При этом одни и те же таймеры могут запускаться многократно, а также могут быть вложенными. По окончании выполнения процедуры (включая ошибочные) для каждого из таймеров формируется следующий набор показателей: суммарное время, количество повторений, среднее время и максимальное время из всех его запусков. Наряду с этими показателями в базу данных также помещается следующая информация: время начала выполнения, идентификатор сервера, параметры формирования изображения, логин пользователя, идентификатор проекта, а также IP адрес, с которого был произведен запрос. Для анализа информации о профилировании времени выполнения запросов реализован специализированный служебный веб-интерфейс.

3.3 Развитие с использованием созданных методов и подходов различных специализированных научных информационных систем, в том числе. ИС ВЕГА-Лес (комплексный научный мониторинг лесных ресурсов России), ИС STS (комплексный научный мониторинг различных явлений в пограничных морях России), ИС VSV (комплексный научный мониторинг вулканической активности Камчатки и Курил)

3.3.1 Разработка новой системы управления комплексными задачами по обработке данных

Эффективность любых систем обработки данных во многом зависит от того, насколько оптимально реализовано управление запусками процедур обработки и контроль их выполнения. Естественно, это справедливо и для систем обработки данных ДЗЗ, которые, как правило, оперируют большими объемами информации, что требует весьма значительных вычислительных ресурсов. Конкретная реализация системы управления заданиями по обработке спутниковых данных во многом зависит от специфики выполняемых задач, используемого программного обеспечения и операционных систем, установленных на серверах обработки. В рамках функционирования ЦКП «ИКИ-Мониторинг» в настоящее время используется ряд систем управления обработкой. Основной из них является разработанная в отделе «Технологии спутникового мониторинга» система управления многопоточной обработкой спутниковых данных на кластере windows серверов, который в настоящее время включает в себя более 100 компьютеров и виртуальных машин. Отдельная система управления заданиями реализована для поддержки работы интерактивных инструментов обработки и анализа спутниковых информации, реализованных в рамках веб-интерфейсов для доступа к данным ЦКП «ИКИ-Мониторинг». Кроме этого в настоящее время используется ряд частных решений для управления специфическими задачами по обработке данных, в частности, для получения информации по объектам мониторинга, выгрузке данных из больших архивов и др. Наличие ряда различных программных компонент, предназначенных для решения во многом схожих задач, безусловно, усложняет их поддержку и развитие. Поэтому актуальной для отдела «Технологии спутникового мониторинга» ИКИ РАН задачей является разработка более универсальных решений, позволяющих управлять широким спектром различных типов обработки спутниковой информации.

В 2022 году была разработана и внедрена новая универсальная система управления комплексными задачами обработки данных TASK JOBS. Ключевым ее преимуществом является возможность управления составными задачами по обработке как спутниковых данных, так и другой информации. Типичным примером такого рода задач является построение композитных спутниковых изображений по большим территориям, фрагменты которых формируются в параллельном режиме на кластере серверов обработки. Также к таким задачам можно отнести и любые другие обработки, которые могут быть выполнены в параллельном режиме и для которых важно получить итоговый результат в заданное время. Таким образом, основной особенностью представленной системы управления является возможность отслеживания не только отдельных заданий по обработке данных, но и их комплектов.

Также при разработке системы TASK JOBS большое внимание было уделено гибкости ее настроек, что во многом обеспечивается использованием JSON-структур для описания как отдельных заданий по обработке, так и задач, в которые они входят. Это позволяет использовать эту систему для широкого спектра различных задач. В частности, наряду с выполнением комплексных (составных) задач обработки данных, поддерживается также режим с потоковым поступлением новых заданий по обработке спутниковых данных. А в тех случаях, когда результатом обработки отдельного задания является относительно компактная текстовая информация, он может быть сохранен в самой системе управления обработкой. Таким образом, при создании рассматриваемой системы ставилась задача предоставления разработчикам программной среды для организации самых разных вариантов управления задачами обработки.

В рамках реализации системы TASK JOBS используются три основных типа объектов: тип задач, задача и задание.

Тип задачи по логике построения системы соответствует типу обработки или конкретному программному пакету, однако разработчик процедур обработки сам вправе решать каким образом ему удобнее структурировать управление разрабатываемого им программного обеспечения. Например, тип задач может соответствовать совокупности задач по построению различных композитных изображений, использующих для этого один и тот же программный пакет.

Задача описывает набор конкретных процедур. Этот набор может быть фиксированным, например, в случае построения композитного спутникового изображения за указанную дату из заданных фрагментов, а может быть постоянно пополняемым, например, когда задача соответствует потоковой обработке запросов, поступающих из интерфейсов.

Задания соответствуют выполнению «элементарных» процедур обработки на отдельных серверах обработки и являются элементом задачи. Так, при построении композитного изображения, задания соответствуют процедурам получения отдельных его фрагментов, описания которых определяются в его параметрах.

Принципиальная схема функционирования системы TASK JOBS приведена на рисунке 3.3.1.1. На схеме представлены три типа серверов, участвующих в работе системы: базовый сервер системы, сервера управления конкретными типами задач по обработке данных и непосредственно сервера обработки данных. На базовом сервере устанавливается основное программное обеспечение:

- базовый API, предназначенный для выполнения всего спектра операций в системе,
- REST API, предназначенный для получения заданий на обработку серверами обработки и отправку в систему информации о ходе и результатах их выполнения,
- базовое ПО, предназначенное для решения таких задач, как детектирование ошибок при выполнении задач и оповещение о них специалистов, удаление

неактуальных данных, создание регулярных отчетов о выполнении задач обработки и др.

- служебный веб-интерфейс, предназначенный для конфигурации системы и визуального контроля выполнения задач обработки

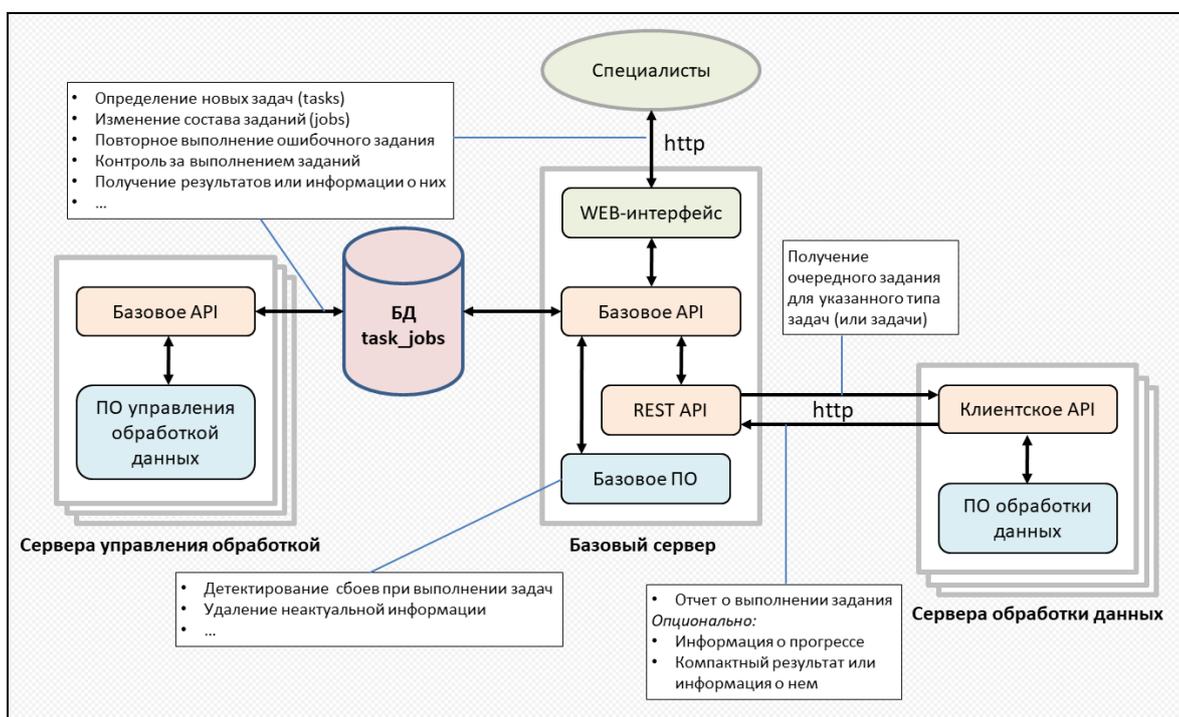


Рисунок 3.3.1.1 - Пример использования служебного интерфейса системы TASK JOBS

На серверах управления обработкой устанавливается только базовый API, взаимодействующий со специфическим программным обеспечением, предназначенным для управления выполнением конкретных типов задач. На серверах обработки устанавливается «легковесный» клиентский API, позволяющий им получать очередные задания на обработку и «отчитываться» о ходе их выполнения. Существенно, что клиентский API является достаточно простым, что позволяет легко адаптировать его под любую программную платформу.

Конфигурация новых типов задач, самих задач и соответствующих им заданий может быть произведена разработчиками при помощи служебного веб-интерфейса. Пример использования этого интерфейса приведен на рисунке 3.3.1.2. При этом основным вариантом является автоматическая их генерация специфическим программным обеспечением, установленным на серверах управления обработкой.

[TASK_JOBS] Интерфейс системы управления выполнением комплексных задач обработки данных

Выбор типа задач: web_data_queue | Очередь заданий для получения данных в WEB интерфейсы | Прочии | Описать новый тип задач

Тип задач (task_type)	Описание (task_desc)	Ответственный (task_type_user)	Максимальное время выполнения задания (max_job_runtime_hours)	Задания сохраняют результат в БД (jobs_store_results)	Вариант добавления заданий (add_job_case)	Изменить	Удалить
web_data_queue	Очередь заданий для получения данных в WEB интерфейсы	Прочии	1	1	queue		

Выбор задачи: 16 | get_unisat_point_iki | failed | активна | Описать новую задачу

ID (task_id)	Название (task_name)	Описание (task_desc)	Параметры (task_pars_json)	Время создания (task_create_dt)	Статистика по заданиям					Приоритет (task_priority)	Статус (task_status)	Активность (task_active)	Изменить	Удалить
					jobs_count	jobs_suc	jobs_inproc	jobs_done	jobs_failed					
16	get_unisat_point_iki	Получение рядов данных по точке (IKU1)		2022-09-20 15:15:53	290	0	1	182	107	1	failed	активна		

Номер (job_number)	Параметры (job_pars_json)	Добавление (job_dt)	Начало (job_start_dt)	Окончание (job_finish_dt)	Сервер (job_server)	Протокол (ошибка) (job_log)	Перезапусков (restart_count)	Статус (job_status)	Процент вып. (done_percent)	Результат (done_percent)	Время выполнения	Отметить
293	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-11-02 16:16:57", "satellite": "TERRA", "AQIA": "AQIA" }	2022-11-02 16:16:57	2022-11-02 16:17:01	2022-11-02 16:17:07	192.168.200.61	-	0	done	-	2.72b	6 sec	<input type="checkbox"/>
292	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-11-02 14:12:24", "satellite": "TERRA", "AQIA": "AQIA" }	2022-11-02 14:12:24	2022-11-02 16:16:48	2022-11-02 16:16:50	192.168.200.61	-	1	inproc	4	-	2 sec	<input type="checkbox"/>
291	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-11-02 10:43:02	2022-11-02 16:04:37	2022-11-02 16:04:33	192.168.200.61	-	24	done	-	2.72b	4 sec	<input type="checkbox"/>
290	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-25", "satellite": "TERRA", "AQIA": "AQIA" }	2022-11-02 10:43:02	2022-11-02 16:04:37	2022-11-02 16:04:33	192.168.200.61	-	1	done	-	9b	12 sec	<input type="checkbox"/>
289	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-25", "satellite": "TERRA", "AQIA": "AQIA" }	2022-11-02 10:43:02	2022-11-02 10:43:06	2022-11-02 10:43:06	192.168.200.61	-	0	done	-	2.56b	6 sec	<input type="checkbox"/>
288	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-11-02 10:41:32	2022-11-02 18:18:02	2022-11-02 18:19:04	192.168.200.61	-	1	done	-	2.56b	1 min 2 sec	<input type="checkbox"/>
287	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-10-31 09:35:45	2022-11-02 18:18:02	2022-11-02 18:18:58	192.168.200.61	-	2	done	-	3.36b	56 sec	<input type="checkbox"/>
286	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-10-28 13:39:53	2022-11-02 18:18:02	2022-11-02 18:18:50	192.168.200.61	-	1	done	-	3.36b	48 sec	<input type="checkbox"/>
285	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-10-28 13:39:48	2022-11-02 18:18:01	2022-11-02 18:18:44	192.168.200.61	-	3	done	-	2.84b	45 sec	<input type="checkbox"/>
284	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-10-28 13:39:48	2022-11-02 18:18:01	2022-11-02 18:18:39	192.168.200.61	-	11	done	-	2.72b	38 sec	<input type="checkbox"/>
283	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-10-28 13:39:09	2022-11-02 18:18:01	2022-11-02 18:18:34	192.168.200.61	-	5	done	-	126b	30 sec	<input type="checkbox"/>
282	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-10-27 13:33:04	2022-11-02 13:11:00	2022-11-02 13:11:04	192.168.200.61	-	4	done	-	781b	4 sec	<input type="checkbox"/>
281	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-10-27 13:32:29	2022-11-02 18:18:01	2022-11-02 18:18:08	192.168.200.61	-	1	done	-	781b	7 sec	<input type="checkbox"/>
280	{ "DB_PKID_MODE": "comp", "device": "MODIS", "dt": "2022-10-27 10:20:00", "satellite": "TERRA", "AQIA": "AQIA" }	2022-10-27 13:32:08	2022-11-02 18:18:00	2022-11-02 18:18:02	192.168.200.61	-	1	done	-	781b	6 sec	<input type="checkbox"/>

Рисунок 3.3.1.2 - Пример использования служебного интерфейса системы TASK JOBS

Централизованное хранение информации в системе TASK JOBS реализовано на основе использования БД task_jobs, в структуру которой входят четыре основных таблицы:

- таблица task_types, описывающая типы задач и общие для их выполнения параметры;
- таблица tasks содержит параметры задач, а также целый ряд показателей, описывающих ход их выполнения;
- таблица task_jobs содержит конфигурацию отдельных заданий, а также информацию о об их выполнении;
- таблица task_jobs_results позволяет сохранять в БД текстовые результаты выполнения заданий.

В настоящее время система TASK JOBS внедрена для решения ряда задач, связанных с получением рядов данных по долговременным спутниковым наблюдениям, которые используются в подсистеме анализа данных в виде графиков. В перспективе планируется использовать ее для всех специфических задач по обработке данных.

3.3.2 Разработка сервисов для эффективного предоставления доступа к временным сериям данных сверхбольших распределенных архивов

В настоящее время для исследований Земли из космоса и мониторинга различных процессов и явлений доступны глобальные и протяженные во времени архивы спутниковых данных, такие как, например, архивы ЦКП «ИКИ-Мониторинг». Существующие инструменты позволяют «на лету» получать различные информационные продукты на основе данных архивов, а так же, рассчитывать историю необходимых для мониторинга характеристик для отдельных объектов наблюдений в «офлайн» режиме. Кроме них, однако, востребованными являются сервисы, позволяющие «на лету» получать данные по большому количеству разновременных наблюдений в произвольных точках Земной поверхности при решении целого ряда отдельных задач спутникового мониторинга. Опыт эксплуатации таких сервисов спутникового мониторинга показал востребованность инструмента анализа данных, представляющих протяженные временные серии в виде графиков. При этом реализованные до недавнего времени сервисы в основном были ориентированы на работу с заранее сформированными рядами данных для фиксированных объектов наблюдения. В то же время, существует

потребность в эффективных инструментах, позволяющих «на лету» получать и анализировать временные серии большого числа наблюдений в произвольных точках Земной поверхности. Такие серии данных могут быть сформированы как на основе частых наблюдений (с периодичностью в несколько минут), так и более редких наблюдений (с периодичностью в недели, месяцы или годы), но при этом за много лет.

До недавнего времени реализованные в ИКИ РАН подходы для формирования больших временных серий данных были недостаточно эффективны, что приводило к большому времени ожидания результатов, неприемлемому для использования этих данных для онлайн-инструментов работы с данными. Для реализации таких инструментов необходима разработка специализированных сервисов и инструментов, ориентированных на быструю обработку запросов пользователей на формирование рядов данных и их визуализации, позволяющих «на лету» получать и анализировать значения необходимых продуктов по любой точке поверхности в границах покрытия архивов. Для создания этих инструментов потребовалась разработка новой архитектуры и программного обеспечения, позволяющих максимально быстро получать временные серии данных, строить по ним информационные продукты и осуществлять их визуализацию.

Новые сервисы для более быстрого получения временных серий данных интегрированы в существующую схему работы с данными распределенных архивов спутниковых данных ЦКП «ИКИ-Мониторинг», включая программные интерфейсы и сервисы анализа данных. Одной из ключевых особенностей технологии UNISAT, на базе которой реализованы распределенные архивы, является поддержка механизма «виртуальных» информационных продуктов, а именно, продуктов, которые формируются по запросу пользователя на основе имеющихся в архивах данных и заданных правил их формирования. Этот механизм позволяет реализовать доступа пользователей к широкому спектру самых разных информационных продуктов, при том что в архивах в основном находятся только данные базового уровня обработки.

Основными задачами новых сервисов для получения временных серий данных являются:

- нормализация, калибровка и вычисление виртуальных продуктов по большим рядам исходных данных на «лету» как в указанной точке, так и в ее окрестности;
- быстрое получение необходимого комплекта метаданных для всей серии данных, включая информацию о файлах данных и правилах построения виртуальных продуктов;
- постановка заданий на формирование временных рядов, запуск процедур получения временных серий и получение результатов (диспетчеризация) на серверах информационных узлов распределенных архивов;
- формирование единого ряда, состоящего из данных, расположенных в разных узлах распределенного архива;
- кэширование результатов чтения временных рядов из исходных файлов для минимизации повторных обращений к данным.

Для реализации вышеперечисленных задач с учетом общей асинхронной «онлайн» схемы взаимодействия пользовательских интерфейсов с сервисами данных были реализованы архитектура системы и необходимые программные блоки. Схема основных блоков и их взаимодействия представлена на рисунке 3.3.2.1. Синим цветом показаны блоки, не требующие доработки, зелеными – блоки, требующие расширения функционала, а оранжевым – вновь создаваемые блоки.

Для реализации сервисов были созданы следующие новые блоки и компоненты, необходимые для решения задач эффективного доступа к временным сериям:

- блок диспетчеризации заданий чтения серий TASK JOBS;
- блок быстрого получения наборов метаданных временных серий;

- блок быстрого получения значений исходных данных и вычисления продуктов;
- блоки кэширования данных в составе как блоков чтения, так и в блоках, получающие «виртуальные» продукты на лету;
- специализированные интерфейсы и плагины для построения графиков по данным временных серий данных, включая модули формирования заданий на получение временных серий данных в программной инфраструктуре графиков.

Существенной доработке подверглись следующие уже реализованные программные решения:

- модули блоков построения интерфейсов работы с графиками для интеграции данных UNISAT (включая необходимые описания) и создаваемой инфраструктурой получения рядов данных;
- инструмент работы с «показателями» для построения по ним временных серий;
- блоки картографических интерфейсов сервисов спутникового мониторинга.

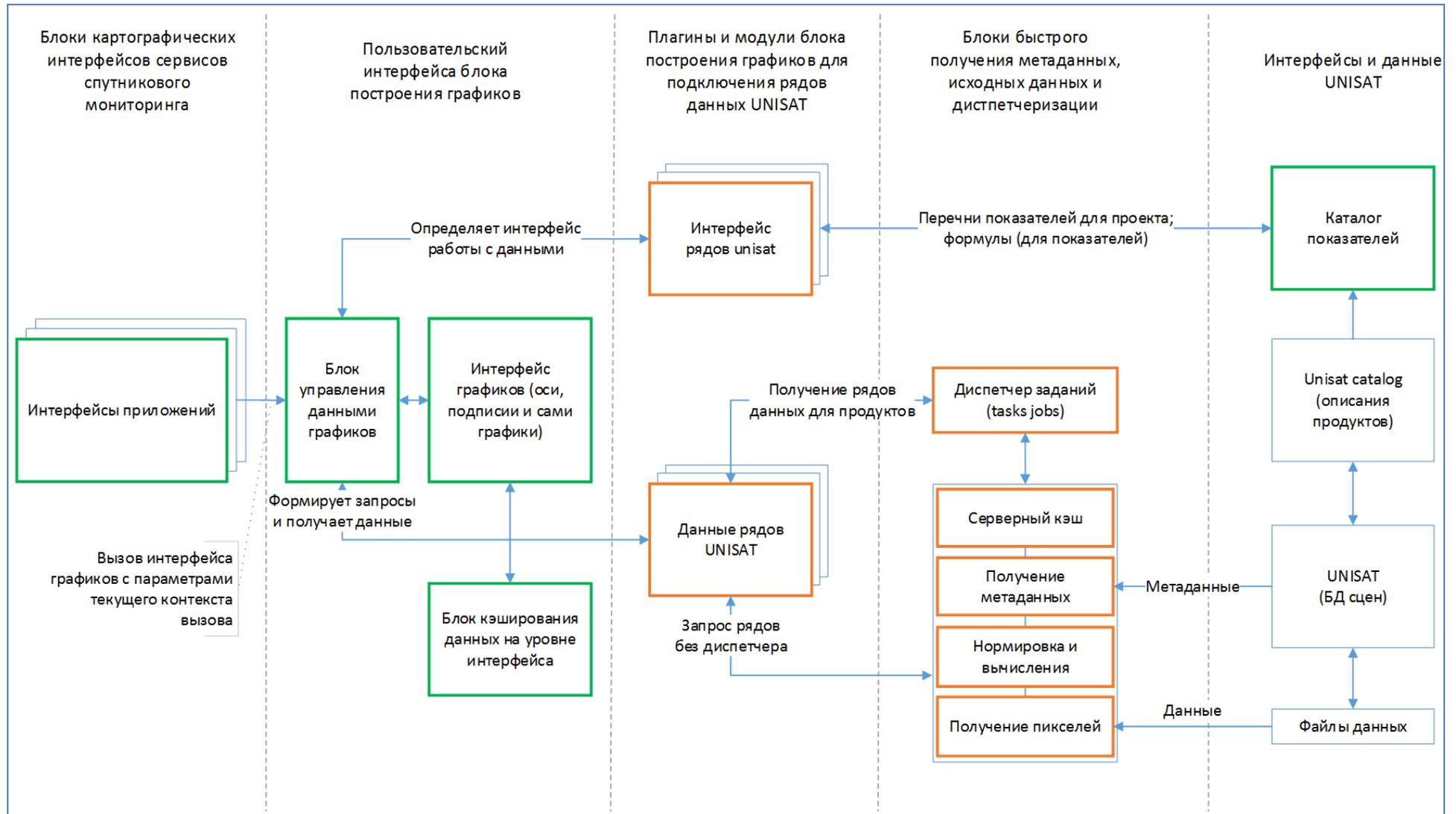


Рисунок 3.3.2.1 - Общая схема структуры и блоков сервисов доступа к временным сериям данных

Реализация асинхронной «онлайн»-схемы работы с сервисом требует универсального инструмента, позволяющего по запросу пользователя сформировать задания на получение требуемых рядов данных в узлах распределенного архива. Для решения этой задачи была использована новая система управления комплексными задачами обработки данных TASK JOBS.

Быстрое получение наборов метаданных временных серий необходимо, в первую очередь, при работе с протяженными рядами данных, насчитывающими более десятков тысяч отдельных элементов, с которыми существующие «поэлементные» схемы становятся крайне неэффективными и затратными по времени выполнения. Метаданные элементов описывают не только коэффициенты калибровки и нормализации данных, но и варианты построения продукта и правила (формулы) его вычисления из базовых составляющих. Следует отметить, что не для всех имеющихся в архивах ЦКП «ИКИ-Мониторинг» рядов продуктов может быть применен одинаковый набор метаданных – в этом случае необходимо использовать оптимизированные варианты поэлементного запроса. В рамках создания блока для быстрого получения наборов метаданных временных серий были реализованы программные модули, работающие в инфраструктуре архивов, построенных по технологии UNISAT, обеспечивающие формирование комплектов метаданных, для каждого задания в очереди системы TASK_JOBS.

Ядром системы сервисов для эффективного предоставления доступа к временным сериям является блок быстрого получения значений продуктов в указанных точках или заданных областях. Блок поддерживает необходимый для работы функционал, такой как:

- получение задания из очереди заданий блока диспетчеризации;
- запрос и получение метаданных и информации о доступе к исходным файлам;
- пере-проектирование координат запроса в конкретные проекции исходных данных;
- быстрое извлечение значений в пикселях файлов исходных данных в точке и ее окрестности;
- применение коэффициентов коррекции, нормировки в соответствии с формулами продукта;
- применение вычисления значений для вычисляемых по исходным данным продуктов.

Кроме этого, модули блока получения значений предусматривают средства оптимизации выполнения запросов, такие как параллельное чтение с учетом параметров вычислительных средств узла, а также оптимизации, связанные с получением данных разных каналов в общих исходных файлах. Информация о ходе обработки и ее результатах поступает в систему диспетчеризации обработки TASK JOBS.

Кэширование (сохранение результатов выполнения заданий) данных для их повторного использования позволяет использовать уже полученные временные серии данных или их подмножества. Кэширование данных реализовано для эффективного формирования временных серий как на серверном уровне (уровне исходных данных) – в блоке диспетчеризации, так и на уровне «клиента» - базовыми средствами интерфейса построения графиков.

В спутниковых сервисах, использующих возможности ЦКП «ИКИ-Мониторинг», для работы с временными сериями, сформированными в «оффлайн»-сценарии подготовки данных, используются специализированные инструменты анализа рядов данных в виде графиков. Архитектура построения данных инструментов обеспечивает их гибкость при решении различных задач и возможность реализации на их основе сервисов для эффективной работы с временными сериями данных сверхбольших распределенных архивов спутниковых данных. Для создания инструмента для работы с временными сериями по данным архивов ЦКП были реализованы специальные плагины для работы с рядами данных UNISAT и отображения сведений о продуктах, включая средства формирования вычисляемых продуктов «на лету». В сценарии асинхронной «онлайн»-работы с данными инструменты построения графиков вызываются из картографических

интерфейсов с указанием необходимых для исследования координат и продуктов обработки спутниковых данных. Инструмент графиков, взаимодействуя с блоком диспетчеризации, каталогом показателей запрашивает и получает необходимые для вычисления продуктов и отображения данные временных серий, позволяя пользователю оперативно управлять как диапазонами отображения данных, так и производными статистическими характеристиками.

В качестве примера использования созданного сервиса можно привести инструмент для работы с данными Himawari-8 в целях ретроспективного анализа и мониторинга активности вулканов в информационной системе дистанционного мониторинга активности вулканов Камчатки и Курил «VolSatView». Данные геостационарного аппарата Himawari-8 характеризуются высокой частотой наблюдений – каждые 10 минут. Созданные сервисы работы с временными сериями позволяют пользователям системы «VolSatView» в настоящее время «на лету» получать графики значений в произвольной точке в зонах мониторинга вулканической активности. Пример графика по временной серии наблюдений инфракрасным каналом прибора АН1 зоны вулкана Карымский в приведен на рисунке 3.3.2.2. График, содержащий более 150 тысяч наблюдений за последние три года, был сформирован за 24 минуты на существующем оборудовании. По значениям температуры пользователи системы могут выявлять как не исследованные ранее события (извержения), так и исследовать подтвержденные другими спутниковыми системами.

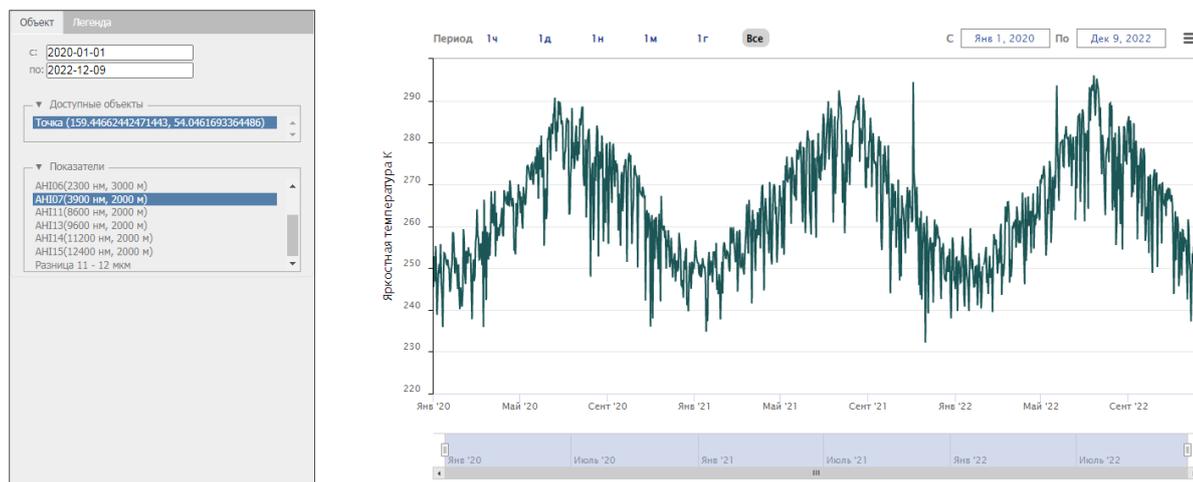


Рисунок 3.3.2.2 - Временная серия наблюдений зоны извержения вулкана Карымский

3.3.3 Технология создания безоблачных композитных изображений по данным спутников серии Sentinel-2

Для изучения изменений, происходящих на поверхности Земли, во многих случаях необходимо получать регулярные и сопоставимые спутниковые данные дистанционного зондирования Земли (ДЗЗ). Особенно остро стоит вопрос получения периодической съемки в задачах изучения растительности. Например, для получения карт растительного покрова или анализа его изменений необходима информация о динамике отражательной способности в различных диапазонах электромагнитного спектра каждой точки картируемой территории в течение сезона вегетации. Надежное регулярное получение такой информации по спутниковым данным может быть затруднено в основном из-за влияния облачности. При этом для разработки автоматических методов анализа данных по большим территориям необходимы полностью ее покрывающие однородные бесшовные данные, в которых полностью отсутствует облачность и связанные с ней мешающие факторы. Одним из основных путей получения таких данных является создание безоблачных композитных изображений.

В 2022 году была доработана и внедрена в эксплуатацию технология получения безоблачных бесшовных композитных изображений по данным спутников серии Sentinel-

2. Исходные данные прибора MSI спутников серии Sentinel-2 распространяются в рамках программы Европейского космического агентства (ESA) Copernicus. В настоящее время в архивах ЦКП «ИКИ-Мониторинг» имеется информация спутников серии Sentinel-2 по территории Северной Евразии и ряду локальных участков по всему миру за период с 2015 года по настоящее время.

В основе рассматриваемой технологии лежит маскирование мешающих факторов с использованием стандартного продукта классификации сцены, полученного при помощи алгоритма Sen2Cor. В этом продукте (далее SC) для каждого пикселя данных указывается один из 11 классов, включая три класса облаков, тени от облаков, темные объекты, растительность, не растительность, воду, неклассифицированные пиксели, значение “нет данных” и сбои. Для качественной оценки точности этого продукта было в экспертном режиме проанализировано несколько сотен сцен, что позволило выявить ряд его недостатков:

- Наличие класса «неклассифицированные пиксели». Особенно часто встречается на урбанизированной территории. Также он присутствует на облачности или тенях от облаков;
- Класс «темные объекты» часто перепутан с тенями от облаков;
- Много проблем на границах классов, когда пограничные или соседние с границей классов пиксели классифицированы неверно. Проблема особенно актуальна для классов облачности;
- Существуют одиночные ошибочные пиксели или линии пикселей в разных классах. Некоторые мелкие объекты, например, в низинах, или на урбанизированной территории могут находиться в классах облачности непрерывно;
- Проблема с классом «дымка». Часто дымка незначительно искажает значения яркостей, однако под ней хорошо просматривается чистая поверхность;
- Классы снега и облаков часто перепутаны, особенно это актуально зимой для снежных облачных фронтов.

Для устранения некоторых указанных проблем к каждой сцене дополнительно применяется ряд операций. Схема построения композитного изображения для бесснежного периода приведена на рисунке 3.3.3.1.

Исходные данные хранятся в фиксированном разбиении в тайлах размером около 100 км на 100 км в проекции UTM/WGS84. В процессе проведения обработки композитное изображение создается независимо по каждому тайлов и по каждому из его спектральных каналов. При этом для каждого тайла отбирается не более определенного числа измерений за период с общим процентом облачности по тайлу меньше заданной. Предельный процент облачности и максимальное число измерений задается опционально в зависимости от периода. В случае если за период есть более указанного числа измерений выбираются те, у которых общий процент облачности по тайлу меньше.

В каждой сцене на первом этапе строится маска на основе продукта SC. Для бесснежного периода считаются чистыми классы «растительность», «не растительность», «вода», «темные объекты» и «неклассифицированные пиксели», а для снежного периода также класс «снег». Остальные классы считаются маской. Для снежного (зимнего) композитного изображения операция улучшения маски проводится в один этап, выполняется только стягивание одиночных групп пикселей. Для бесснежного периода к каждой маске применяются независимо морфологические операции расширения (“Уточнение dilation”) и сужения (“Уточнение erosion”).

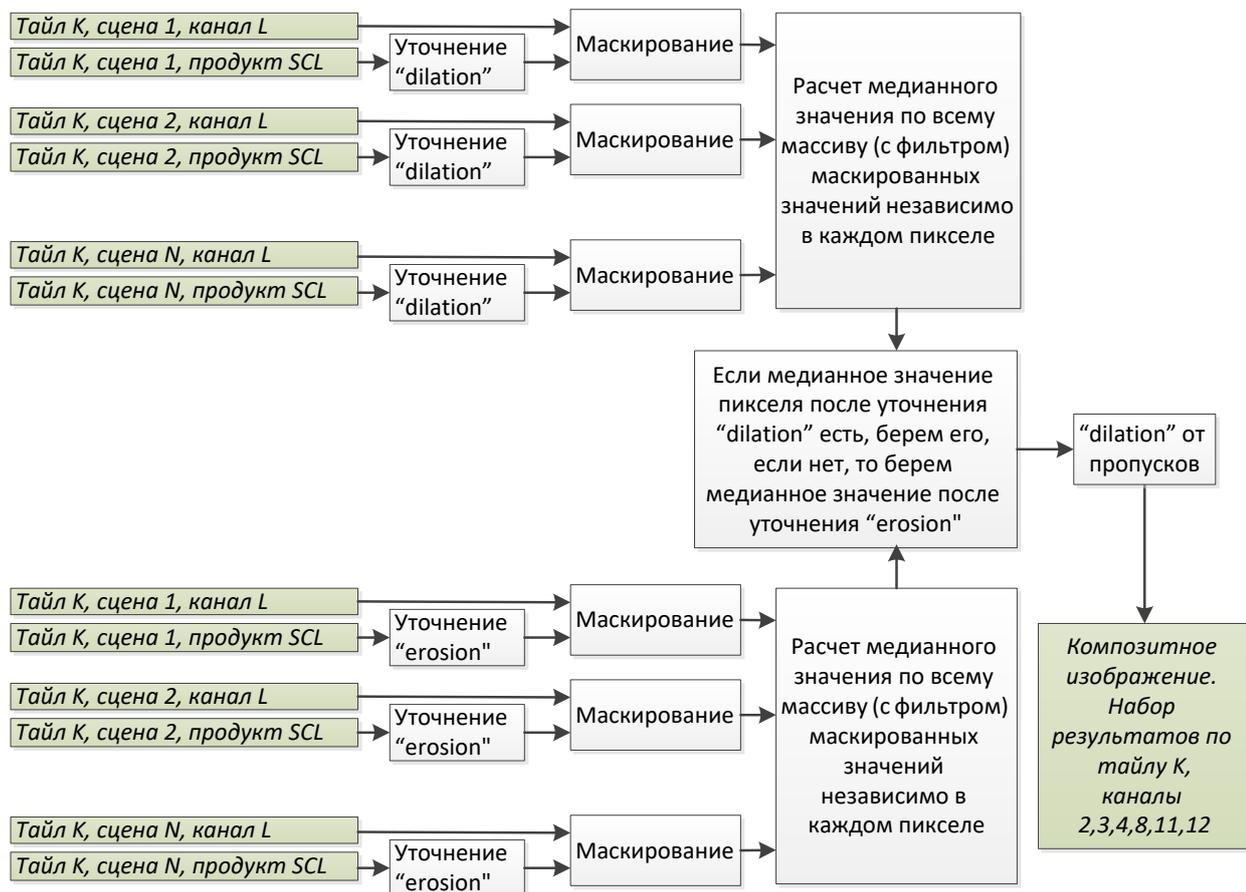


Рисунок 3.3.3.1 - Схема построения композитного изображения для бесснежного периода

После применения маски независимо для улучшений путем расширения и сужения считается медианное значение по всему массиву измерений. Если чистое измерение не было получено в ветке с применением операции расширения, то оно заполняется значением из ветки с применением операции сужения. Применение операции расширения в первую очередь необходимо для устранения проблем на границе класса облачности. Применение операции сужения необходимо для устранения одиночных ошибочных пикселей или ошибочных линий пикселей в разных классах. При отсутствии этого этапа некоторые мелкие объекты были бы полностью исключены из композитного изображения. Размер расширения (10 пикселей) и сужения (1 пиксель) был подобран эмпирически при анализе сотен различных сцен за 2021 год. Результат оценивался в экспертном режиме на качественном уровне, без точной количественной оценки. Дополнительно исследовалась возможность учета иной комбинации классов из продукта SC, например, принятие темных объектов или класса неклассифицированных пикселей в качестве маски. Однако, учет иного набора классов в качестве чистых приводил к существенно худшему результату, особенно при числе безоблачных измерений за период менее 5. На последнем этапе от оставшихся без данных пикселей дополнительно проводится отступление операцией расширения. Эта операция необходима, так как на границе мест отсутствия безоблачных измерений качество результирующего композитного изображения, полученного всего по нескольким сценам, получается низким.

Отметим также особенности технологии и временные затраты на создание композитных изображений по описываемой технологии. При создании композитного изображения данные за отобранные даты внутри периода выгружаются с сервера хранения непосредственно в оперативную память сервера, выполняющего обработку. Такой подход позволяет избежать каких либо дополнительных временных затрат на копирование данных. В результате, при расчете в 8 потоков с использованием одного

высокопроизводительного сервера получено среднее время на создание композитного изображения за месячный период на территорию России около 3 дней.

С помощью описанной технологии был создан ряд безоблачных композитных изображений на территорию России по данным 2019, 2020 и 2021 года. При этом были построены следующие варианты этих продуктов: ежегодные за сезон вегетации (с 1 мая по 31 сентября), ежемесячные, сезонные и «снежные» (с 1 января по 31 марта). На рисунке 3.3.3.2 приведены примеры созданных композитных изображений.

В настоящее время реализована сезонная и месячная периодичности создания композитных изображений. Месячный период обусловлен с одной стороны достаточным обеспечением безоблачными данными каждой точки за этот период, а с другой стороны, тем, что скважность в один месяц позволяет выявить большинство особенностей развития растительного покрова в течение сезона вегетации. Выбор времени построения снежного (зимнего) композитного изображения обусловлен более вероятным наличием в этот период снежного покрова. Для создания композитных изображений использованы стандартные продукты уровня обработки L2A, то есть после проведения процедуры атмосферной коррекции. Отметим, что с помощью описанной технологии возможно создание композитных изображений по информации в других спектральных каналах, а также по данным уровня обработки L1C.

Доступ к композитным изображениям предоставляется в рамках информационной системы «ВЕГА-Science» (<http://sci-vega.ru/>). Кроме данных в отдельных спектральных каналах также доступен ряд тематических информационных продуктов, являющихся синтезом информации в отдельных каналах композитного изображения или результатом вычислений над ними. Такие информационные продукты получают в режиме реального времени в момент запроса на основе ранее разработанной технологии предоставления виртуальных информационных продуктов. Также доступны разновременные продукты на основе композитных изображений, например, синтез за три месяца, в котором для каждого из RGB каналов цветного изображения используются данные одного из месяцев. Такие продукты хорошо показывает динамику изменений земной поверхности. К примеру, они могут быть использованы для задачи картографирования лесных гарей. Различные цвета в синтезированном изображении явно показывают период появления поврежденных участков леса. Также они находят применения для решения различных задач анализа сельскохозяйственной растительности, таких как: определение границ полей, выявление неиспользуемых земель, выявление пахотных земель, определение различных видов угодий и типов выращиваемых культур. При этом благодаря непрерывному покрытию территории созданными композитными изображениями они могут быть использованы для построения различных тематических карт по большим территориям, в частности, карт различных типов растительности.

Представленная технология позволяет достаточно оперативно создавать безоблачные композитные изображения на большую территорию. В перспективе планируется реализовать возможность восстановления данных за длительные облачные периоды, улучшить качество масок облачности, а также применить созданные композитные изображения для задач автоматического картографирования Земной поверхности.

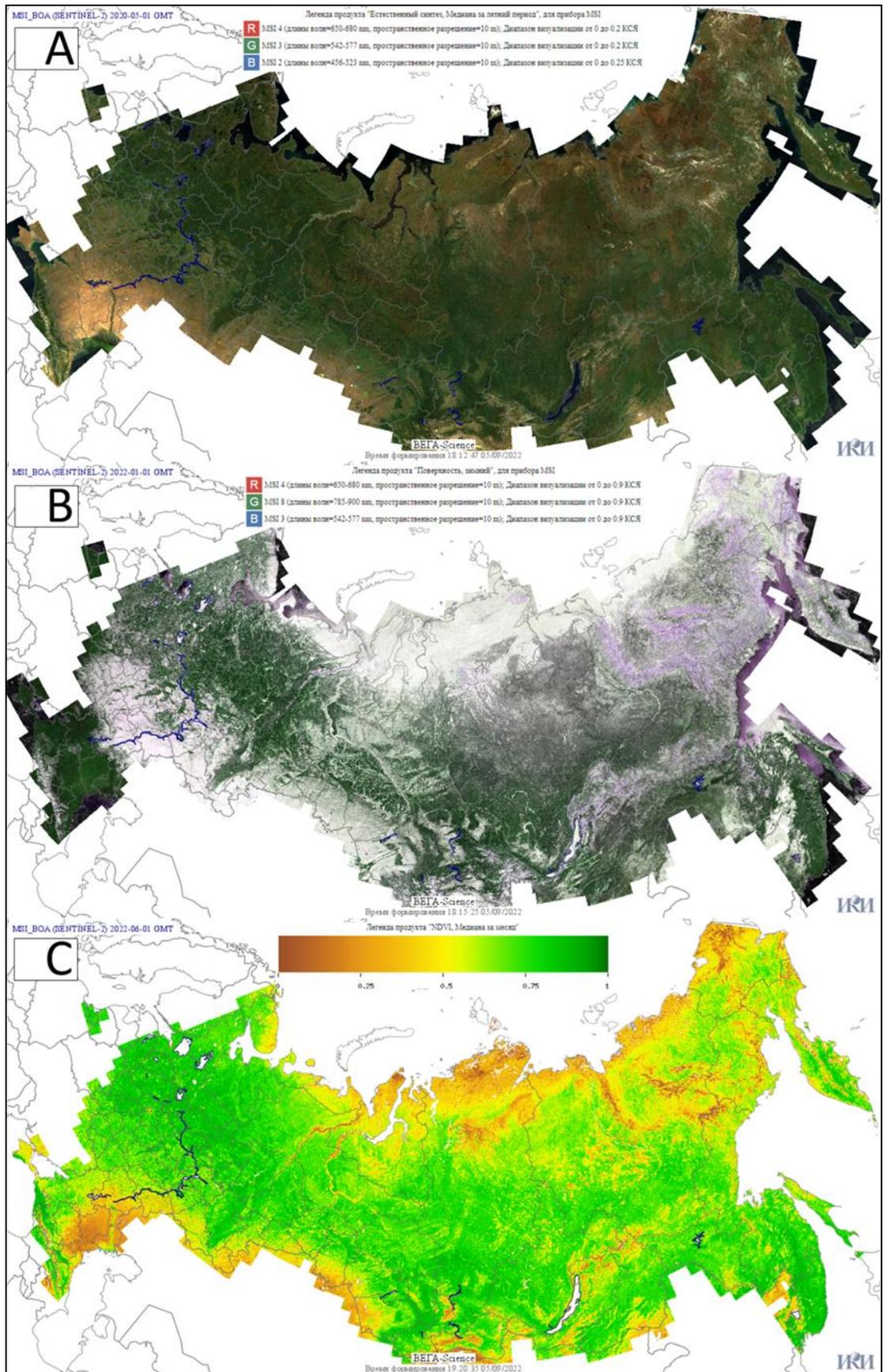


Рисунок 3.3.3.2 - Примеры созданных композитных изображений. А - за сезон вегетации 2020 года (синтез в естественных цветах, каналы Red-Green-Blue). В - за снежный период 2022 года (синтез в ложных цветах, каналы Red-Nir-Green). С - за июнь месяц 2022 года (продукт индекс NDVI)

Заключение

Представленные в настоящей главе результаты показывают, что по направлению «Разработка новых подходов и методов организации сверхбольших архивов данных спутниковых наблюдений Земли и доступа к ним для обеспечения эффективной обработки и анализа данных ДЗЗ для решения научных и прикладных задач» проведенные работы соответствуют установленной в плане работ ИКИ РАН на 2022 год тематике. Таким образом, работы выполнены полностью в соответствии с планом.

При этом получены следующие основные результаты:

- Проведена модернизация и расширение функционала базового программного обеспечения, реализующего технологию динамического блочного доступа к данным для проведения их обработки;
- Реализованы новые подходы к повышению эффективности обработки данных сверхбольших архивов спутниковых наблюдений;
- Проведена модернизация картографического веб-интерфейса, реализованного на базе технологии динамического блочного доступа;
- Разработана новая подсистема управления комплексными задачами по обработке данных, используемая при реализации функционала создаваемых в ИКИ РАН информационных систем;
- Разработаны новые, более эффективные сервисы для предоставления доступа к большим временным сериям данных, получаемых по данным сверхбольших распределенных архивов;
- Разработана и внедрена технология построения безоблачных композитных изображений по данным спутников серии Sentinel-2, которые востребованы для решения целого ряда исследовательских задач рамках создаваемых в ИКИ РАН информационных систем.

Таким образом, можно считать, что работы по данному направлению темы в 2022 году полностью выполнены.

ЗАКЛЮЧЕНИЕ

Таким образом, представленный отчет по теме «Большие данные в космических исследованиях: астрофизика, солнечная система, геосфера» (№ FFWG-2022-0012) показывает, что в соответствии с планом ИКИ РАН на 2022 год они велись по следующим основным направлениям:

- Развитие методов моделирования высокоширотной ионосферы; построение численных и гибридных (включающих элементы ассимиляции данных измерений ионосферно-магнитосферных параметров) моделей распределения электронной концентрации в полярной ионосфере и зависимости от изменяющихся гелио-геофизических условий.
- Исследование и разработка алгоритмов машинного обучения для измерения расстояний (красных смещений) астрономических объектов по данным многоволновых фотометрических обзоров неба
- Разработка новых подходов и методов организации сверхбольших архивов данных спутниковых наблюдений Земли и доступа к ним для обеспечения эффективной обработки и анализа данных ДЗЗ для решения научных и прикладных задач

В рамках данной тематики опубликовано в 2022 году 36 работ, из них 6 в рецензируемых изданиях, входящих в Scopus и WoS (плановый показатель 3).

Таким образом, можно считать, что поставленные в плане задачи на 2022 год в рамках данной темы полностью решены.

ПУБЛИКАЦИИ

опубликовано – 38 работ, из них опубликовано:

- в зарубежных изданиях – 5 (1- в печати)
- книг – 0
- статьи в отечественных научных рецензируемых журналах – 5 (-1 в печати)
- статьи в сборниках материалов конференций - 2
- материалы конференций - 8
- доклады, тезисы, циркуляры – 18
- Патенты – 0
- публикаций по грантам РФФ – 3
- в соавторстве с зарубежными учёными - 2
- статьи в WoS и Scopus – 7 (из них 1- в печати)
- статьи в WOS и Scopus (Q1, Q2) – 3 (из них 1 - в печати)

Публикации в зарубежных изданиях

1. *Gulyaeva T., Stanislavska I., Lukianova R.* Arctic–Antarctic asymmetry of the ionospheric weather // *Advances in Space Research*, Available online 11 May 2022, <https://doi.org/10.1016/j.asr.2022.05.008>. (РИНЦ, Scopus, Q2)
2. *P. Hartley, A. Bonaldi, R. Braun, D. Cornu, B. Semelin, X. Lu, S. Aicardi, P. Salomé, A. Marchal, J. Freundlich, F. Combes, C. Tasse, C. Heneka, M. Delli Veneri, A. Soroka, F. Gubanov, A. Meshcheryakov, B. Fraga, C.R. Bom, M. Brügger, A. K. Shaw, N. Patra, A. Chakraborty, R. Mondal, S. Choudhuri, A. Mazumder, M. Jagannath, M. J. Hardcastle, J. Forbrich, L. Smith, V. Stolyarov, M. Ashdown, J. Coles, H. Håkansson, A. Sjöberg, M. C. Toribio, M. Önnheim, M. Olberg, E. Gustavsson, M. Lindqvist, M. Jirstrand, J. Conway, K. M. Hess, R. J. Jurek, S. Kitaeff, P. Serra, A. X. Shen, J. M. van der Hulst, T. Westmeier, A. Alberdi, J. Cannon, L. Darriba, J. Garrido, J. Góza, D. Herranz, M. G. Jones, P. Kamphuis, D. Kleiner, I. Márquez, J. Moldón, M. Pandey-Pommier, M. Parra, J. Sabater, S. Sánchez, A. Sorgho, L. Verdes-Montenegro, G. Fourestey, A. Galan, C. Gheller, D. Korber, A. Peel, M. Sargent, E. Tolley, B. Liu, R. Chen, B. Peng, L. Yu, H. Xi, K. Yu, Q. Guo, W. Pei, Y. Liu, Y. Wang, X. Chen, X. Zhang, S. Ni, J. Zhang, L. Gao, M. Zhao, L. Zhang, H. Zhang, X. Wang, J. Ding, S. Zuo, Y. Mao, A. Vafaei Sadr, M. Kunz, B. Bassett, V. Nistane, N. Oozeer, S. Jaiswal, B. Lao, J. N. H. S. Aditya, Y. Zhang, A. Wang, X. Yang* SKA Science Data Challenge 2: analysis and results // *Monthly Notices of the Royal Astronomical Society*. 2023. (WoS/Scopus Q1)– в печати
3. *Soroka Andrey, Meshcheryakov Alex, Gerasimov Sergey* Morphological Classification of Astronomical Images with Limited Labelling // *Astronomical Data Analysis Software and Systems XXX*. ASP Conference Series. Vol. 532. Proceedings of a virtual conference held 8-12 November 2020. Edited by Jose Enrique Ruiz, Francesco Pierfederici, and Peter Teuben. San Francisco: Astronomical Society of the Pacific. 2022. p.307
4. *Borisov, Viktor; Meshcheryakov, Alex; Gerasimov, Sergey* Probabilistic Photo-z Machine Learning Models for X-ray Sky Surveys // *Astronomical Data Analysis Software and Systems XXX*. ASP Conference Series. Vol. 532. Proceedings of a virtual conference held 8-12 November 2020. Edited by Jose Enrique Ruiz, Francesco Pierfederici, and Peter Teuben. San Francisco: Astronomical Society of the Pacific. 2022. p.231
5. *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](#)

Публикации в отечественных научных рецензируемых изданиях

6. Лукьянова Р.Ю., Алексеев А.В., Сериков М.В. Изменчивость интегрального влагосодержания атмосферы на севере Атлантического и Тихого океанов по данным спутникового микроволнового зондирования // Метеорология и гидрология. № 4, с. 46–58, 2022, doi: 10.52002/0130-2906-2022-4-46-58. (РИНЦ) Переводная версия: Lukianova, R.Y., Alekseev, G.V. & Serikov, M.V. Variability of Total Precipitable Water over the North Atlantic and North Pacific According to Satellite Microwave Sounding. Russ. Meteorol. Hydrol. 47, 281–289 (2022). <https://doi.org/10.3103/S1068373922040045/> (WOS, Scopus)
7. Мингалев О.В., Сецко П.В., Мельник М.Н., Мингалев И.В., Малова Х.В., Артемьев А.В., Мерзлый А.М., Зеленый Л.М. Роль ионов кислорода в структуре токового слоя ближнего хвоста магнитосферы Земли // Физика плазмы. 2022. Т. 48, № 3, с. 237–258, doi: 10.31857/S036729212203009X. (РИНЦ) Переводная версия: Mingalev O.V., Setsko P.V., Mel'nik M.N., Mingalev I.V., Malova K.V., Artem'ev A.V., Merzlyi A.M., Zelenyi L.M. Role of Oxygen Ions in the Structure of the Current Sheet of the near-Earth Magnetotail // Plasma Physics Reports. 2022. Т. 48. № 3. С. 242–262. DOI: 10.1134/S1063780X22030096. (WOS, Scopus)
8. Кузьмин А.К., Мерзлый А.М., Никифоров О.В., Петрукович А.А., Потанин Ю.Н., Садовский А.М., Соколов А.Д., Янаков А.Т. Аннотированный атлас примеров изображений эмиссий в авроральных структурах, зарегистрированных имаджерами и изображающими спектрографами с разных орбит и поверхности земли. Часть 1. Авроральные структуры, возбужденные природными источниками, включая Альфвеновские волны // Гелиогеофизические Исследования. 2022. - Принято в печать (РИНЦ)
9. Мещеряков А. В., Немешаева А., Буренин Р. А., Гильфанов М. Р., Сюняев Р. А. Расширенный каталог объектов Сюняева-Зельдовича по данным спутника PLANCK с использованием глубокого обучения // Письма в астрономический журнал. 2022. Т. 48. No 9. с. 603–621 DOI: 10.31857/S0320010822090066 (РИНЦ) Переводная версия: Meshcheryakov A.V. , Nemeshaeva A. , Burenin R.A., Gilfanov M.R., Sunyaev R.A. An Extended Catalog of Sunyaev–Zeldovich Objects from Planck Data with Deep Learning // Astronomy Letters, 2022, Vol. 48, No. 9, pp. 479–496. DOI: 10.1134/S1063773722090055 (WoS, Scopus, Q3)
10. Кашиницкий А.В., Бурцев М.А., Прошин А.А. [Технология создания безоблачных композитных изображений по данным спутников серии Sentinel-2](#) // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 5. С. 76–85. DOI: 10.21046/2070-7401-2022-19-5-76-85. (РИНЦ, Scopus)

Статьи в сборниках материалов конференций

11. Прошин А.А., Лупян Е.А. [Технология ведения сверхбольших архивов спутниковых данных с поддержкой динамического формирования информационных продуктов и инструментов для их анализа и обработки](#) // Материалы XVIII научно-технической конференции "Системы наблюдения, мониторинга и дистанционного зондирования Земли". Москва, Зеленоград. (г. Сочи, 12-17 сентября 2022 г.), 2022. С. 260–266.
12. Лупян Е.А., Прошин А.А. [Основные возможности центра коллективного пользования системами архивации, обработки и анализа спутниковой информации \(ЦКП "ИКИ-мониторинг"\)](#) // Материалы XVIII научно-технической конференции "Системы

наблюдения, мониторинга и дистанционного зондирования Земли". Москва, Зеленоград. (г. Сочи, 12-17 сентября 2022 г.), 2022. С. 252-259.

Материалы конференций

13. *Лукьянова Р.Ю.* Эволюция стратосферного аэрозоля над регионом Казахстана по данным спутника Suomi NPP // Материалы 20-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Москва: ИКИ РАН, 2022. С. 429. DOI 10.21046/20DZZconf-2022a (РИНЦ)
14. *Прошин А.А., Лупян Е.А., Бурцев М.А.* [Особенности построения распределенных архивов спутниковых данных ЦКП "ИКИ-Мониторинг"](#) // Восьмой Белорусский космический конгресс. 25-27 октября 2022. Минск, 2022. С. 282-285.
15. *Бурцев М.А., Прошин А.А., Матвеев А.М., Егоров В.А.* Организация работы с данными прибора VIIRS для обеспечения преемственности работ на базе прибора MODIS // Материалы 20-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2022. С. 82. DOI: 10.21046/20DZZconf-2022a. (РИНЦ)
16. *Колбудаев П.А., Плотников Д.Е., Матвеев А.М.* Детектирование облачности и теней на монохроматических изображениях приборов серии КМСС с использованием свёрточной нейронной сети U-Net // Материалы 20-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2022. С. 40. DOI: 10.21046/20DZZconf-2022a. (РИНЦ)
17. *Плотников Д.Е., Колбудаев П.А., Матвеев А.М., Прошин А.А., Лупян Е.А.* Технология оперативного построения и результаты оценки продуктов КМСС уровня L2A для решения задач оперативного мониторинга растительного покрова России // Материалы 20-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2022. С. 108. DOI: 10.21046/20DZZconf-2022a. (РИНЦ)
18. *Балашов И.В., Лупян Е.А., Марченко В.В., Прошин А.А., Руткевич Б.П., Уваров И.А.* Построение инструментов онлайн-анализа временных серий данных сверхбольших распределенных архивов // Материалы 20-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2022. С. 77. DOI: 10.21046/20DZZconf-2022a. (РИНЦ)
19. *Прошин А.А., Бурцев М.А., Радченко М.В.* Организация распределенных файловых хранилищ для ведения сверхбольших архивов спутниковых данных ДЗЗ // Материалы 20-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2022. С. 110. DOI: 10.21046/20DZZconf-2022a. (РИНЦ)
20. *Прошин А.А., Марченко В.В., Руткевич Б.П.* Универсальная система управления комплексными задачами обработки спутниковых данных TASK JOBS // Материалы 20-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции.

Тезисы, доклады

21. *Немешаева А.А., Мещеряков А.В., Герасимов С.В.* Нейросетевые методы поиска скоплений галактик в данных космического обзора Planck // Научная конференция "Тихоновские Чтения - 2022". Москва. 24-29 октября 2022. тезисы докладов. устный доклад.
22. *Мещеряков А.В., Сорока А.Г., Герасимов С.В.* Применение алгоритмов обучения с подкреплением в системе оптимизации маршрутов с ограничениями вида временных окон и пропускной способности // Научная конференция "Тихоновские Чтения - 2022". Москва. 24-29 октября 2022. тезисы докладов. устный доклад.
23. *Борисов В.Д., Мещеряков А.В., Герасимов С.В.* Применение машинного обучения для интеллектуального анализа данных о космических транзиентах // Научная конференция "Тихоновские Чтения - 2022". Москва. 24-29 октября 2022 тезисы докладов. устный доклад.
24. *Васильев С.М., Мещеряков А.В., Герасимов С.В.* Применение машинного обучения для построения карты крупномасштабной структуры Вселенной по данным оптических обзоров неба // Научная конференция "Тихоновские Чтения - 2022" Москва. 24-29 октября 2022. тезисы докладов. устный доклад.
25. *Мещеряков А.В., Михельсон Г.В., Герасимов С.В., Сорока А.Г.* Сравнение классических эвристических и точных методов оптимизации маршрутов в статической задаче VRP с ограничениями на грузоподъемность и временные окна доставки товаров // Научная конференция "Тихоновские Чтения - 2022". Москва. 24-29 октября 2022. тезисы докладов. устный доклад.
26. *Мещеряков А.В., Герасимов С.В., Сорока А.Г.* Достоинства и недостатки моделей обучения с подкреплением при оптимизации маршрутов в задачах большой размерности с ограничениями вида временные окна, объем грузов и мультидепо // Ломоносовские чтения – 2022. Секция вычислительная математика и кибернетика. 14-22 апреля 2022. Москва. МГУ имени М.В.Ломоносова. факультет ВМК. тезисы докладов. устный доклад.
27. *Мещеряков А.В., Герасимов С.В., Васильев С.М.* Классификация скоплений галактик по картам волокон крупномасштабной структуры Вселенной // Ломоносовские чтения – 2022. Секция вычислительная математика и кибернетика. 14-22 апреля 2022. Москва. МГУ имени М.В.Ломоносова. факультет ВМК. тезисы докладов. устный доклад
28. *Мещеряков А.В., Герасимов С. В., Кривонос Р.А., Губанов Ф.Д.* Разработка технологии для полуавтоматической классификации транзиентов SRG/eРОЗИТА. Ломоносовские чтения - 2022, Секция вычислительная математика и кибернетика, 14-22 апреля 2022. Москва. МГУ имени М.В.Ломоносова. факультет ВМК. тезисы докладов. устный доклад
29. *Мещеряков А.В., Герасимов С.В., Дмитриенко А.Ю.* Технология разметки действий людей на видео по нескольким примерам с применением глубокого обучения. Ломоносовские чтения – 2022. Секция вычислительная математика и кибернетика. 14-22 апреля 2022. Москва. МГУ имени М.В.Ломоносова. факультет ВМК. тезисы докладов. устный доклад
30. *Мещеряков А.В.* Расширенный каталог объектов Сюняева-Зельдовича по данным спутника PLANCK с использованием глубокого обучения // Астрофизика высоких

- энергий сегодня и завтра 2022 (HEA-2022). 19-22 декабря 2022. Москва. тезисы докладов. устный доклад.
31. *Воскресенская С., Мещеряков А.В.* Поиск массивных рентгеновских скоплений галактик в данных АСТ с помощью глубокого обучения // *Астрофизика высоких энергий сегодня и завтра 2022 (HEA-2022)*. 19-22 декабря 2022. Москва. тезисы докладов.
 32. *Lukianova R.* Modeling the large scale irregularities in the F region polar ionosphere // *The SCOSTEP 15th Quadrennial Solar-Terrestrial Physics Symposium (STP-15) 21-25 February 2022*. Alibag, India.
 33. *Сахаров Я.А., Золотой С.А., Мёрзлый А.М., Садовский А.М., Петрукович А.А., Янаков А.Т., Никифоров О.В., Селиванов В.Н.* Оценка воздействия магнитосферных возмущений на энергетические системы в средних широтах // *Физика авроральных явлений*. 45 ежегодный семинар. 14 – 18 March 2022. ПГИ. Апатиты. Сборник тезисов. С. 18.
 34. *Сахаров Я.А., Золотой С.А., Мёрзлый А.М., Моисеев И.А., Никифоров О.В., Петрукович А.А., Садовский А.М., Селиванов В.Н., Янаков А.Т.* Проявление воздействий магнитосферных возмущений в средних широтах // *ITES 2022, VI Всероссийская конференция с международным участием «Информационные технологии для наук о Земле и цифровизация в геологии и горнодобывающей промышленности»*. Владивосток, 3-8 октября 2022. Сборник тезисов. Изд ДВФУ. С.31.
 35. *Мёрзлый А.М., Калишин А.С., Мингалев И.В., Моисеев И.А., Никифоров О.В., Сахаров Я.А., Суворова З.В., Черняков С.М., Шубин В.Н., Янаков А.Т.* Оценка характеристик зон засветки односкачковыми лучевыми траекториями в высоких широтах в 24-м цикле солнечной активности // 17 Конференция "Физика плазмы в солнечной системе". 7-11 февраля 2022. ИКИ РАН. Москва. Сборник тезисов. https://plasma2023.cosmos.ru/docs/2022/Plasma-2022-AbstractBook_v4.pdf
 36. *Замогильный Д., Мёрзлый А.М., Никифоров О.В., Петрукович А.А., Сахаров Я.А., Чернышов А.А., Янаков А.Т., Куприянов А.О., Алибин Д.Ю.* Основные результаты применения приемника глобальных навигационных спутниковых систем для оценки параметров высокоширотной ионосферы на архипелаге Земля Франца-Иосифа в марте 2021 г. // 17 Конференция "Физика плазмы в солнечной системе". 7-11 февраля 2022 г. ИКИ РАН. Москва. Сборник тезисов. https://plasma2023.cosmos.ru/docs/2022/Plasma-2022-AbstractBook_v4.pdf
 37. *Мингалев И.В., Суворова З.В., Мингалев О.В., Ахметов О.И., Мёрзлый А.М., Попов В.Ю.* Численное моделирование прохождения метровых радиоволн через ионосферные неоднородности // 17 Конференция "Физика плазмы в солнечной системе". 7-11 февраля 2022 г.. ИКИ РАН. Москва. Сборник тезисов. https://plasma2023.cosmos.ru/docs/2022/Plasma-2022-AbstractBook_v4.pdf
 38. *Сахаров Я.А., Золотой С.А., Мёрзлый А.М., Садовский А.М., Петрукович А.А., Янаков А.Т., Никифоров О.В., Селиванов В.Н.* Особенности воздействия возмущений магнитосферы и ионосферы на энергетические системы в средних широтах // 17 Конференция "Физика плазмы в солнечной системе", 7-11 февраля 2022 г. ИКИ РАН. Москва. Сборник тезисов. https://plasma2023.cosmos.ru/docs/2022/Plasma-2022-AbstractBook_v4.pdf

РНФ

1. *Belvedersky, M. I. ; Meshcheryakov, A. V. ; Medvedev, P. S. ; Gilfanov, M. R.* SRGz: Building an Optical Cross-Match Model for the X-ray SRG/eROSITA Sources Using the Lockman Hole Data // *Astronomy Letters*, 20026 Volume 48, Issue 2, p.109-125 DOI: 10.1134/S106377372202001 (**Scopus, Q2**)

2. Бикмаев И.Ф., Колбин А.И., Шиманский В.В., Хамитов И.М., Иртуганов Э.Н., Николаева Е.А., Сахибуллин Н.А., Гумеров Р.И., Буренин Р.А., Гильфанов М.Р., Зазнобин И.А., Кривонос Р.А., Медведев П.С., Мещеряков А.В., Сазонов С.Ю., Сюняев Р.А., Хорунжеев Г.А., Моисеев А.В., Малыгин Е.А., Шабловинская Е.С., Желтоухов С.Г. SRGe2149+6736 - кандидат в переменные типа AM Her, обнаруженный телескопом eРозита орбитальной обсерватории ``Спектр-Рентген-Гамма`` // Письма в астрономический журнал. 2022. Т. 48. № 9. С. 645-656 (перевод статьи выйдет в журнале Astronomy Letters. 2022. Vol.48. №9 (**WoS/Scopus Q3**))
3. Мещеряков А.В., Борисов В.Д., Хорунжеев Г.А., Васильев С.М., Бельведерский М.И., Медведев П.А., Гильфанов М.Р., Сазонов С.Ю., Буренин Р.А., Бикмаев И.Ф., Хамитов И.М., Машечкин И.В., Сюняев Р.А. SRGz: методы машинного обучения и свойства каталога оптических компаньонов точечных рентгеновских источников СРГ/eРОЗИТА в области покрытия DESI Legacy Imaging Surveys // принята к публикации в Письма в астрономический журнал, 2023, том 49 (перевод статьи выйдет в журнале Astronomy Letters, 2022, Vol.49 (**WoS, Scopus Q3**))

Публикации в соавторстве с зарубежными авторами

1. P. Hartley, A. Bonaldi, R. Braun, D. Cornu, B. Semelin, X. Lu, S. Aicardi, P. Salomé, A. Marchal, J. Freundlich, F. Combes, C. Tasse, C. Heneka, M. Delli Veneri, A. Soroka, F. Gubanov, A. Meshcheryakov, B. Fraga, C.R. Bom, M. Brüggem, A. K. Shaw, N. Patra, A. Chakraborty, R. Mondal, S. Choudhuri, A. Mazumder, M. Jagannath, M. J. Hardcastle, J. Forbrich, L. Smith, V. Stolyarov, M. Ashdown, J. Coles, H. Håkansson, A. Sjöberg, M. C. Toribio, M. Önnheim, M. Olberg, E. Gustavsson, M. Lindqvist, M. Jirstrand, J. Conway, K. M. Hess, R. J. Jurek, S. Kitaeff, P. Serra, A. X. Shen, J. M. van der Hulst, T. Westmeier, A. Alberdi, J. Cannon, L. Darriba, J. Garrido, J. Góza, D. Herranz, M. G. Jones, P. Kamphuis, D. Kleiner, I. Márquez, J. Moldón, M. Pandey-Pommier, M. Parra, J. Sabater, S. Sánchez, A. Sorgho, L. Verdes-Montenegro, G. Fourestey, A. Galan, C. Gheller, D. Korber, A. Peel, M. Sargent, E. Tolley, B. Liu, R. Chen, B. Peng, L. Yu, H. Xi, K. Yu, Q. Guo, W. Pei, Y. Liu, Y. Wang, X. Chen, X. Zhang, S. Ni, J. Zhang, L. Gao, M. Zhao, L. Zhang, H. Zhang, X. Wang, J. Ding, S. Zuo, Y. Mao, A. Vafaei Sadr, M. Kunz, B. Bassett, V. Nistane, N. Oozeer, S. Jaiswal, B. Lao, J. N. H. S. Aditya, Y. Zhang, A. Wang, X. Yang SKA Science Data Challenge 2: analysis and results // Monthly Notices of the Royal Astronomical Society. 2023. (**WoS, Scopus Q1**)– в печати
2. Gulyaeva T., Stanislavska I., Lukianova R. Arctic–Antarctic asymmetry of the ionospheric weather // Advances in Space Research, Available online 11 May 2022, <https://doi.org/10.1016/j.asr.2022.05.008>. (**РИНЦ, Scopus, Q2**)

Публикации в рецензируемых изданиях по теме Космос-Д

1. Gulyaeva T., Stanislavska I., Lukianova R. Arctic–Antarctic asymmetry of the ionospheric weather // Advances in Space Research, Available online 11 May 2022, <https://doi.org/10.1016/j.asr.2022.05.008>. (**РИНЦ, Scopus, Q2**)
2. P. Hartley, A. Bonaldi, R. Braun, D. Cornu, B. Semelin, X. Lu, S. Aicardi, P. Salomé, A. Marchal, J. Freundlich, F. Combes, C. Tasse, C. Heneka, M. Delli Veneri, A. Soroka, F. Gubanov, A. Meshcheryakov, B. Fraga, C.R. Bom, M. Brüggem, A. K. Shaw, N. Patra, A. Chakraborty, R. Mondal, S. Choudhuri, A. Mazumder, M. Jagannath, M. J. Hardcastle, J. Forbrich, L. Smith, V. Stolyarov, M. Ashdown, J. Coles, H. Håkansson, A. Sjöberg, M. C. Toribio, M. Önnheim, M. Olberg, E. Gustavsson, M. Lindqvist, M. Jirstrand, J. Conway, K. M. Hess, R. J. Jurek, S. Kitaeff, P. Serra, A. X. Shen, J. M. van der Hulst, T. Westmeier, A.

- Alberdi, J. Cannon, L. Darriba, J. Garrido, J. Gósz, D. Herranz, M. G. Jones, P. Kamphuis, D. Kleiner, I. Márquez, J. Moldón, M. Pandey-Pommier, M. Parra, J. Sabater, S. Sánchez, A. Sorgho, L. Verdes-Montenegro, G. Fourestey, A. Galan, C. Gheller, D. Korber, A. Peel, M. Sargent, E. Tolley, B. Liu, R. Chen, B. Peng, L. Yu, H. Xi, K. Yu, Q. Guo, W. Pei, Y. Liu, Y. Wang, X. Chen, X. Zhang, S. Ni, J. Zhang, L. Gao, M. Zhao, L. Zhang, H. Zhang, X. Wang, J. Ding, S. Zuo, Y. Mao, A. Vafaei Sadr, M. Kunz, B. Bassett, V. Nistane, N. Oozeer, S. Jaiswal, B. Lao, J. N. H. S. Aditya, Y. Zhang, A. Wang, X. Yang* SKA Science Data Challenge 2: analysis and results // Monthly Notices of the Royal Astronomical Society. 2023. (**WoS, Scopus Q1**)– в печати
3. *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. Vol. 14. №. 1. P. 77. DOI: doi.org/10.3390/rs14010077. (**РИНЦ, WOS, Scopus, Q1**)
 4. *Лукьянова Р.Ю., Алексеев А.В., Сериков М.В.* Изменчивость интегрального влагосодержания атмосферы на севере Атлантического и Тихого океанов по данным спутникового микроволнового зондирования // Метеорология и гидрология. № 4, с. 46-58, 2022, doi: 10.52002/0130-2906-2022-4-46-58. (**РИНЦ**) Переводная версия: Lukianova, R.Y., Alekseev, G.V. & Serikov, M.V. Variability of Total Precipitable Water over the North Atlantic and North Pacific According to Satellite Microwave Sounding. Russ. Meteorol. Hydrol. 47, 281–289 (2022). <https://doi.org/10.3103/S1068373922040045/> (**WOS, Scopus**)
 5. *Мингалев О.В., Сецко П.В., Мельник М.Н., Мингалев И.В., Малова Х.В., Артемьев А.В., Мерзлый А.М., Зеленый Л.М.* Роль ионов кислорода в структуре токового слоя ближнего хвоста магнитосферы Земли // Физика плазмы. 2022. Т. 48, № 3, с. 237– 258, doi: 10.31857/S036729212203009X. (**РИНЦ**) Переводная версия: Mingalev O.V., Setsko P.V., Mel'nik M.N., Mingalev I.V., Malova K.V., Artem'ev A.V., Merzlyi A.M., Zelenyi L.M. Role of Oxygen Ions in the Structure of the Current Sheet of the near-Earth Magnetotail // Plasma Physics Reports. 2022. Т. 48. № 3. С. 242-262. DOI: 10.1134/S1063780X22030096. (**WOS, Scopus**)
 6. *Кузьмин А.К., Мерзлый А.М., Никифоров О.В., Петрукович А.А., Потанин Ю.Н., Садовский А.М., Соколов А.Д., Янаков А.Т.* Аннотированный атлас примеров изображений эмиссий в авроральных структурах, зарегистрированных имаджерами и изображающими спектрографами с разных орбит и поверхности земли. Часть 1. Авроральные структуры, возбужденные природными источниками, включая Альфвеновские волны // Гелиогеофизические Исследования. 2022. - Принято в печать (**РИНЦ**)
 7. *Мещеряков А. В., Немешаева А., Буренин Р. А., Гильфанов М. Р., Сюняев Р. А.* Расширенный каталог объектов Сюняева-Зельдовича по данным спутника PLANCK с использованием глубокого обучения // Письма в астрономический журнал. 2022. Т. 48. No 9. с. 603–621 DOI: 10.31857/S0320010822090066 (**РИНЦ**) Переводная версия: Meshcheryakov A.V. , Nemeshaeva A. , Burenin R.A., Gilfanov M.R., Sunyaev R.A. An Extended Catalog of Sunyaev–Zeldovich Objects from Planck Data with Deep Learning // Astronomy Letters, 2022, Vol. 48, No. 9, pp. 479–496. DOI: 10.1134/S1063773722090055 (**WoS, Scopus, Q3**)
 8. *Кашицкий А.В., Бурцев М.А., Прошин А.А.* [Технология создания безоблачных композитных изображений по данным спутников серии Sentinel-2](#) // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 5. С. 76-85. DOI: 10.21046/2070-7401-2022-19-5-76-85. (**РИНЦ, Scopus**)