

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 АААА-А19-119072590030-9

УТВЕРЖДАЮ
Директор ИКИ РАН
Член-корреспондент РАН

А.А. Петрукович
« 20 » декабря 2021 г.
М.П.

ОТЧЕТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

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

по теме :

КОСМОС-Д

0024-2021-0011

(промежуточный, этап 2021/1)

Научный руководитель

канд. техн. наук



А.А. Прошин

Москва

2021

СПИСОК ИСПОЛНИТЕЛЕЙ

СПИСОК ИСПОЛНИТЕЛЕЙ

Руководитель темы,
старший научный сотрудник,
зав. лаб., к.т.н.

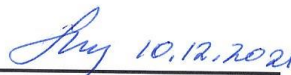


10.12.2021

Подпись, дата

Прошин А.А. (раздел
1-3, введение,
заключение)

Ответственный исполнитель,
ведущий научный сотрудник
д.ф.-м.н.



Подпись, дата

Лукьянова Р.Ю.
(раздел 1)

Ответственный исполнитель,
научный сотрудник
к.ф.-м.н.

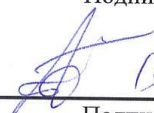


10.12.2021

Подпись, дата

Мещеряков А.В.
(раздел 2)

Ответственный исполнитель,
старший научный сотрудник,
зав. Лаб., к.т.н.

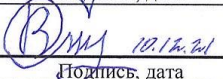
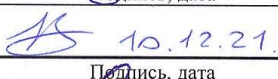
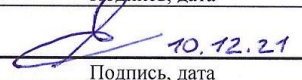


10.12.2021

Подпись, дата

Прошин А.А.
(раздел 3)

Исполнители отдела 54

старший научный сотрудник,	 Подпись, дата	А.М. Мерзлый (раздел 1)
научный сотрудник, к.н.	 Подпись, дата	В.А. Ожередов (раздел 1)
ведущий инженер	 Подпись, дата	О.В. Никифоров (раздел 1)
ведущий инженер	 Подпись, дата	А.Т. Янаков (раздел 1)

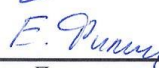
Исполнители отдела 52

старший научный
сотрудник, к.ф.-м.н.


Подпись, дата 10.12.21

Р.А. Буренин
(раздел 2)

научный сотрудник,
к.ф.-м.н.


Подпись, дата 10.12.21

Е.В. Филиппова
(раздел 2)

научный сотрудник,
к.ф.-м.н.


Подпись, дата 10.12.21

Г.А. Хорунжев
(раздел 2)

младший научный
сотрудник


Подпись, дата 10.12.21

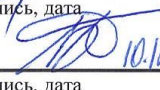
И.А. Зазнобин
(раздел 2)

младший научный
сотрудник


Подпись, дата 10.12.21

А.Р. Ляпин
(раздел 2)

инженер


Подпись, дата 10.12.21

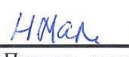
М.И. Бельведерский
(раздел 2)

старший лаборант


Подпись, дата 10.12.21

В.Д. Борисов
(раздел 2)

старший лаборант


Подпись, дата 10.12.21

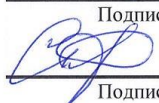
Н.В. Малышева
(раздел 2)

старший лаборант


Подпись, дата 10.12.21

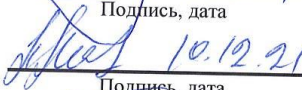
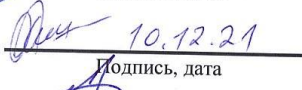
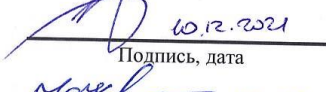
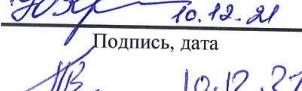
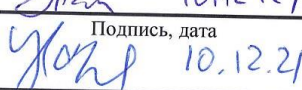
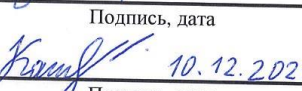
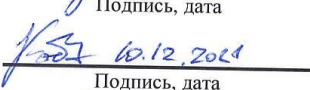
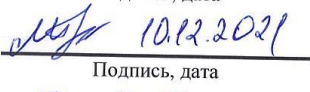
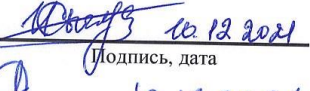
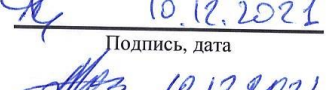
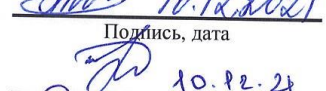
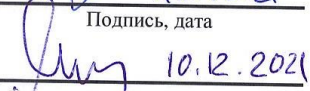
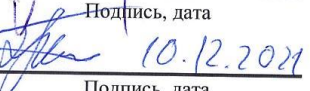
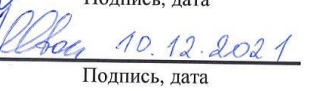
А.А. Немешаева
(раздел 2)

старший лаборант


Подпись, дата 10.12.21

С.М. Николаева
(раздел 2)

Исполнители отдела 56

зав. отдела, д.т.н.	 Подпись, дата	Е.А. Лупян (раздел 3)
ведущий научный сотрудник, к.ф.-м.н.	 Подпись, дата	А.А. Мазуров (раздел 3)
заведующий лабораторией, к.ф.-м.н.	 Подпись, дата	В.Н. Пырков (раздел 3)
старший научный сотрудник, зав. лаб., к.т.н.	 Подпись, дата	И.В. Балашов (раздел 3)
старший научный сотрудник, зав. лаб., к.т.н.	 Подпись, дата	М.А. Бурцев (раздел 3)
старший научный сотрудник, к.ф.-м.н.	 Подпись, дата	Ю.С. Крашенинникова (раздел 3)
старший научный сотрудник, к.т.н.	 Подпись, дата	В.А. Толпин (раздел 3)
старший научный сотрудник, к.т.н.	 Подпись, дата	И.А. Уваров (раздел 3)
старший научный сотрудник, к.т.н.	 Подпись, дата	А.В. Кашницкий (раздел 3)
научный сотрудник, к.т.н.	 Подпись, дата	Д.А. Кобец (раздел 3)
младший научный сотрудник	 Подпись, дата	А.М. Матвеев (раздел 3)
младший научный сотрудник	 Подпись, дата	И.Г. Сычугов (раздел 3)
главный специалист	 Подпись, дата	А.Ю. Дегай (раздел 3)
главный специалист, к.ф.- м.н.	 Подпись, дата	А.А. Козочкина (раздел 3)
ведущий инженер	 Подпись, дата	М.В. Радченко (раздел 3)
ведущий конструктор	 Подпись, дата	В.В. Марченков (раздел 3)
инженер	 Подпись, дата	А.А. Бриль (раздел 3)
инженер	 Подпись, дата	А.М. Константинова (раздел 3)

инженер



10.12.2021

Подпись, дата

К.С. Сенько
(раздел 3)

инженер



10.12.21

Подпись, дата

В.Н. Черных
(раздел 3)

старший лаборант

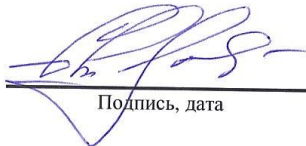


10.12.21

Подпись, дата

Д.А. Юдин
(раздел 3)

нормоконтролер



Подпись, дата

Е.А.Лупян

РЕФЕРАТ

Отчет 66 стр., 39 рис., 23 источника.

НАСТОЯЩИЙ ОТЧЕТ ПОСВЯЩЕН ОПИСАНИЮ ОСНОВНЫХ РЕЗУЛЬТАТОВ РАБОТ, ПОЛУЧЕННЫХ В 2021 ГОДУ В РАМКАХ ТЕМЫ "КОСМОС-Д" (№ 0024-2019-14), 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 2021 ГОД ПО СЛЕДУЮЩИМ НАПРАВЛЕНИЯМ:

Развитие комплекса взаимосвязанных физико-математических моделей для диагностики и прогнозирования состояния высокоширотной ионосферы. Разработки общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли.

Применение методов машинного обучения и глубокого обучения для обработки и анализа больших массивов астрономических данных в задачах наблюдательной астрофизики.

Разработка новых методов и подходов организации хранения и параллельной обработки сверхбольших объемов данных для решения задач анализа долговременных наблюдений в интересах ведения мониторинга и изучения различных процессов и явлений.

КЛЮЧЕВЫЕ СЛОВА

СИСТЕМЫ ПАРАЛЛЕЛЬНОЙ ОБРАБОТКИ СПУТНИКОВЫХ ДАННЫХ, СВЕРХБОЛЬШИЕ ДОЛГОВРЕМЕННЫЕ АРХИВЫ СПУТНИКОВЫХ ДАННЫХ, БЛОЧНЫЙ ДОСТУП К ДАННЫМ В АРХИВАХ, АНАЛИЗ ДОЛГОВРЕМЕННЫХ СЕРИЙ НАБЛЮДЕНИЙ, ТЕХНОЛОГИЯ UNISAT, МИССИЯ ЭКЗОМАРС-2016, АППАРАТ TGO, MARS EXPRESS, ИНФОРМАЦИОННЫЙ СЕРВИС ARES, ОБРАБОТКА СПУТНИКОВЫХ ДАННЫХ, ДОСТУП К АРХИВАМ СПУТНИКОВЫХ ДАННЫХ, ИНФОРМАЦИОННАЯ СИСТЕМА "АВРОРА-АРКТИКА", АНАЛИЗ СОСТОЯНИЯ ИОНОСФЕРЫ ЗЕМЛИ, ПРОГНОЗЫ СОСТОЯНИЯ АВРОРАЛЬНОГО ОВАЛА.

ИОНОСФЕРНО-МАГНИТОСФЕРНОЕ ВЗАИМОДЕЙСТВИЕ И ЭЛЕКТРОДИНАМИКА, ВЫСОКОШИРОТНАЯ ИОНОСФЕРА, ПРОДОЛЬНЫЕ ТОКИ, ЭЛЕКТРОННАЯ КОНЦЕНТРАЦИЯ, ОБРАБОТКА МАССИВОВ ДАННЫХ СПУТНИКОВЫХ И РАДАРНЫХ ИЗМЕРЕНИЙ, ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ.

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

СОДЕРЖАНИЕ

СПИСОК ИСПОЛНИТЕЛЕЙ.....	2
РЕФЕРАТ	7
СОДЕРЖАНИЕ.....	8
Введение.....	10
ГЛАВА 1 Развитие комплекса взаимосвязанных и дополняющих друг друга физико-математических моделей для диагностики и прогнозирования состояния высокоширотной ионосферы в северном и южном полушариях, которые электродинамически связаны между собой через область средних широт. Разработки общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли.	12
Введение.....	12
1.1 Развитие численной модели распределения электронной концентрации в F области высокоширотной ионосферы.....	12
1.1.1 Постановка нестационарной задачи	12
1.1.2 Оценка расчетного времени релаксации и времени реконфигурации картины конвекции.....	13
1.1.3 Отображение траекторий конвективного дрейфа при изменяющемся ММП.....	14
1.1.4 Результаты тестирования нестационарной модели.....	15
1.2 Анализ эволюции продольных токов во время магнитных бурь по данным измерений спутниками SWARM и оценки динамики границ и плотностей продольных токов, включая мелкомасштабные структуры.....	18
1.2.1 Продольные токи во время магнитных бурь	18
1.2.2 Развитие информационно-аналитической системы «Гелиогеофизика»	19
1.2.2.1 Концепция	19
1.2.2.2 Структурная схема ИАС	19
1.2.2.3 Наполнение и организация компонентов.....	20
1.3 Разработка общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли.....	21
1.3.1 Исследование структуры верхней/средней атмосферы при формировании экстремально сильного стратосферного вихря зимой 2019-2020 гг. в Арктике	21
1.3.2 Инструментарий: лимбовые измерения MLS радиометра спутника Aura, радиозондирование нижней ионосферы метеорным радаром, атмосферный реанализ	22
1.3.3 Структура теплового поля	22
1.3.4 Структура ветрового поля	24
Заключение.....	25
ГЛАВА 2 Применение методов машинного обучения и глубокого обучения для обработки и анализа больших массивов астрономических данных в задачах наблюдательной астрофизики	27
Введение.....	27
2.1 Разработка и применение алгоритмов глубокого обучения для задачи морфологической классификации галактик на оптических изображениях. Исследование и разработка алгоритмов кросс-отождествления объектов в фотометрических каталогах обзоров неба, полученных из различных спектральных диапазонов.....	27
2.2 Исследование алгоритмов машинного обучения для измерения фотометрических красных смещений галактик.....	31
Заключение.....	31
ГЛАВА 3 Разработка новых методов и подходов организации хранения и параллельной обработки сверхбольших объемов данных для решения задач анализа долговременных наблюдений в интересах ведения мониторинга и изучения различных процессов и явлений	33
Введение.....	33

3.1. Разработка подходов и методов реорганизации существующих долговременных архивов данных спутниковых наблюдений Земли для повышения эффективности организации их предоставления системам распределенного многопоточного (распараллеленного) анализа.	34
3.1.1 Модернизация программной реализации технологии динамического блочного доступа.....	34
3.1.2 Анализ эффективности процедур подготовки и предоставления требуемого набора блоков данных.....	36
3.1.2.1 Методика оценки временных затрат на подготовку блоков данных.....	36
3.1.2.2 Получение базовой зависимости T-base.....	37
3.1.2.3 Получение зависимости T-compression.....	39
3.1.2.4 Получение зависимости T-get.....	40
3.1.2.5 Вывод.....	41
3.1.3 Организация обработки данных КМСС.....	41
3.1.4 Разработка инструментов для визуального анализа наличия пространственных данных.....	43
3.2. Использование технологии динамического блочного доступа для организации доступа к данным архивов ЦКП «ИКИ-Мониторинг».....	45
3.2.1. Обновленный картографический веб-интерфейс, использующий тайловое формирование изображений.....	45
3.2.2. Новый универсальный модуль растеризации изображений.....	46
3.3. Модернизация подсистемы для анализа временных и пространственных рядов данных на основе построения графиков.....	47
3.4. Развитие методов параллельной обработки спутниковых данных.....	50
3.5. Развитие системы контроля функционирования сложного распределенного программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг».....	53
3.5.1. Общая информация.....	53
3.5.2. Контроль выполнения процессов.....	54
3.5.3. Автоматическое детектирование сбоев в работе программно-аппаратного комплекса.....	55
3.5.4. Документирование сбоев в работе комплекса. Система ведения сбоев.....	56
3.5.5. Ежедневные отчеты о функционировании отдельных подсистем комплекса.....	57
3.5.6. Инструменты анализа работы различных подсистем комплекса.....	58
3.5.7. Документирование используемых программных и аппаратных средств.....	59
3.5.8. Заключение.....	59
Заключение.....	59
ЗАКЛЮЧЕНИЕ.....	61
СПИСОК публикаций.....	62

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»
1	Пламенные процессы в геофизике и астрофизике	1.3.4.3.
2	Фундаментальные проблемы распространения радиоволн, в том числе в ионосфере	1.3.6.4.
3	Происхождение, строение и эволюция Вселенной	1.3.7.1.
4	Физика галактик и межгалактической среды	1.3.7.2.
5	Развитие методов наземной и внеатмосферной астрономии	1.3.7.6.
6	Информационно-вычислительные системы и среды в науке и образовании	1.1.8
7	Алгоритмы и программные системы в космическом мониторинге Земли и экологии	1.1.7.6
8	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 комплекса взаимосвязанных и дополняющих друг друга физико-математических моделей для диагностики и прогнозирования состояния высокоширотной ионосферы в северном и южном полушариях, которые электродинамически связаны между собой через область средних широт. Разработки общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли.
- Применение методов машинного обучения и глубокого обучения для обработки и анализа больших массивов астрономических данных в задачах наблюдательной астрофизики.
- Разработка новых методов и подходов организации хранения и параллельной обработки сверхбольших объемов данных для решения задач анализа долговременных наблюдений в интересах ведения мониторинга и изучения различных процессов и явлений.

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

ГЛАВА 1 Развитие комплекса взаимосвязанных и дополняющих друг друга физико-математических моделей для диагностики и прогнозирования состояния высокоширотной ионосферы в северном и южном полушариях, которые электродинамически связаны между собой через область средних широт. Разработки общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли.

Введение

В 2021 г. работы проводились по следующим пунктам плана: - развитие численной модели распределения электронной концентрации в F-области полярной ионосферы, параметризация и валидация модели по данным сети ионозондов и радарных измерений; - анализ эволюции продольных токов во время магнитных бурь по данным высокоточных измерений вариаций магнитного поля и параметров плазмы в верхней ионосфере, а также в направлении разработки общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли. В комплексе с исследованиями, выполненными в предыдущие два года, работы полностью отвечают общему плану раздела.

Исследования 2021 года были направлены на усовершенствование модели распределения высокоширотной ионосферы и ее применение для диагностики крупномасштабных неоднородностей ионосферной плазмы, влияющих на распространение радиоволн. Были выполнены разработка и апробация нестационарной версии модели. Модель учитывает изменяющиеся под действием нестационарного солнечного ветра условия в ионосфере в течение периода расчета конечного распределения плотности плазмы. Была проанализирована эволюция продольных токов во время магнитных бурь 24 солнечного цикла. Были продолжены работы по развитию, структурированию и наполнению информационно-аналитической системы «Гелиогеофизика». С помощью данных спутникового зондирования атмосферы и измерений метеорного радара была исследована эволюция зимнего стратосферного вихря.

Полученные в данном направлении результаты изложены в публикациях [1, 6, 8, 16].

1.1 Развитие численной модели распределения электронной концентрации в F области высокоширотной ионосферы

1.1.1 Постановка нестационарной задачи

В F области высокоширотной ионосферы существует ряд крупномасштабных неоднородностей электронной концентрации N_e , которые оказывают значительное влияние на распространение радиоволн. В связи с ограниченным количеством пунктов регулярных наблюдений в высоких широтах эмпирические модели не могут дать детального описания трехмерной структуры полярной ионосферы и её структурных особенностей для реальных гелиогеофизических условий, и актуальной задачей является разработка и использование физико-математической модели, которая позволила бы получить необходимую информацию о состоянии полярной ионосферы.

Численная модель предназначена для оперативного расчета 3-D распределения N_e на высотах 120-600 км – в F области полярной ионосферы. В высоких широтах электрические поля магнитосферного происхождения и выпадения заряженных частиц из магнитосферы оказывает большое влияние на распределение ионосферной плазмы.

Структура электрического поля определяет траектории конвекции, вдоль которых дрейфуют плазменные трубки. Конвектируя, трубка попадает в различные условия ионизации, вследствие чего вертикальное распределение плазмы в ней изменяется. Траектории конвекции моделируются в блоке расчета электрических полей в зависимости от входных параметров: значений межпланетного магнитного поля, скорости солнечного ветра, дня года, мирового времени, уровня геомагнитной и солнечной активности. В ионосферном блоке при расчете высотного распределения плазмы учитываются обобщенный молекулярный ион M^+ и ион кислорода. Для того, чтобы учесть предысторию ионизационно-рекомбинационных процессов в плазменной трубке, интегрирование уравнения фотохимического равновесия для M^+ и диффузионного уравнения для O^+ должно начинаться с некоторой начальной точки. На первом этапе осуществляется расчёт траекторий конвекции вспять по времени для определения этой стартовой точки, удалённой настолько, чтобы избежать влияния неопределённости начальных условий на высотный профиль Ne в требуемой точке. Поместив каждую трубку в ее стартовую точку, можно начинать рассчитывать вертикальное распределение плазмы в трубке по мере того, как она дрейфует вдоль линий конвекции по времени вперед.

Распределение электрического потенциала в ионосфере и, соответственно, конфигурация линий конвекции, зависят от параметров межпланетного магнитного поля (ММП). В базовой версии модели, разработанной ранее, предполагалось, что параметры ММП в течение всего периода между моментом времени, для которого производится расчет Ne , и моментом старта расчета, ММП не меняется, и соответственно не меняется и картина конвекции. Однако, для солнечного ветра типичным является изменение направления и величины вектора ММП на временных масштабах минут и часов. После изменения ММП картины конвекции постепенно перестраиваются. Несмотря на довольно большое число оценок, сделанных на основе различных методов, скорость перестройки конвекции не определена однозначно.

Согласно плану развития модели в прошедшем году была проведена ее модернизация с целью перехода к нестационарному режиму. Новая версия позволяет учесть изменение параметров ММП внутри периода времени релаксации между стартовой и расчетной точками. Это приближает модель к воспроизведению реальных условий, а также расширяет возможности постановки численных экспериментов, выявляющих зависимость крупномасштабных ионосферных неоднородностей от ММП в высоких широтах.

1.1.2 Оценка расчетного времени релаксации и времени реконфигурации картины конвекции

Входными параметрами модели являются: вертикальная (B_z) и азимутальная (B_y) компоненты ММП, скорость солнечного ветра, день года и универсальное время (UT), индексы солнечной и геомагнитной активности. Выходные параметры: высотные профили Ne в диапазоне высот 120-600 км в узлах сетки, покрывающей область широт $>50^\circ$, 2-мерные карты распределения максимума $Ne(max)$ в заданный момент времени.

Оценка влияния удаленности времени старта расчета ионосферного блока от заданного входного параметра UT (*время релаксации, τ*) сделана для отдельно для условий ММП $B_z < 0$ и $B_z > 0$. В качестве оценочного параметра выбрана сумма всех значений ΣNe в точках 3-D сетки, покрывающей расчетную область. За единицу взята сумма, полученная при $\tau = 12$ часов. Серия расчетов с уменьшением до $\tau = 2$ часа, контролем ΣNe и вычислением относительного отклонения от единицы показала (Рисунок 1.1.2.1), что сокращение τ с 12 до 6 часов практически не влияет на конечный результат. При $\tau = 4$ отклонение не превышает 10%, а при $\tau < 4$ происходит его резкий рост в отрицательную сторону. Т.е. чтобы плазма в трубке достигла равновесного состояния, и начальные условия не оказывали влияния на высотный профиль Ne , стартовая точка должна по времени отстоять от расчетной не менее, чем на 4 часа. Слишком малое τ ведет к

недооценке плотности плазмы, причем при $B_z < 0$ ошибка ожидаемо больше (т.к. скорость конвекции выше), чем при $B_z > 0$.

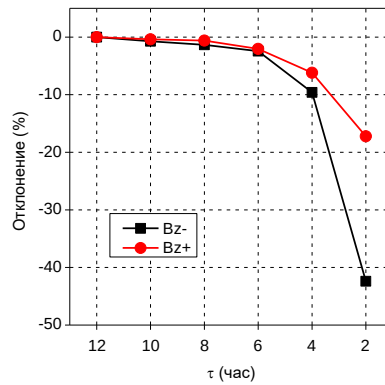


Рисунок 1.1.2.1 – Изменение (в относительных единицах) суммарного значения N_e в зависимости от времени релаксации

На изменения ММП конвекция реагирует переконфигурированием в структуру, которая соответствует новому ММП. Согласно известным оценкам, время реконфигурации варьируется от единиц до десятков минут в зависимости от сектора местного магнитного времени. Также время отклика зависит от состояния ионосферы, характера изменения ММП и других трудно учитываемых факторов. Наша статистическая оценка, проведенная по данным измерений скорости дрейфа плазмы радаром некогерентного рассеяния ESR на Шпицбергене, показала, что максимальная корреляция между ММП границе магнитосферы и скоростью дрейфа ионосферной плазмы достигается при временном лаге в 15-30 мин. Учитывая фактор τ и приблизительное время отклика конвективного дрейфа на вариации ММП, можно заменить во входных параметрах модели постоянные значения B_z и B_y на непрерывно изменяющиеся на 4-часовом интервале значения.

1.1.3 Отображение траекторий конвективного дрейфа при изменяющемся ММП

В конвективном блоке производится расчет пространственно-временной структуры картины конвекции плазменных трубок. Перемещение вдоль траектории описывается уравнениями

$$\begin{aligned} d\theta/dt &= -2 \cdot E_\varphi(\theta, \varphi) \cdot (R^2/M) \cdot \cos\theta / [3 \cdot \cos^2\theta + 1], \\ d\varphi/dt &= E_\theta(\theta, \varphi) \cdot (R^2/M) \cdot [2 \cdot \cos\theta \cdot \sin\theta]^{-1} + \omega, \end{aligned}$$

где E_φ и E_θ – долготная и широтная компоненты ионосферного электрического поля, M – магнитный момент Земли, ω – частота углового вращения Земли, $R = R_E + 200$ км, t – время.

При нестационарном подходе распределение E_φ и E_θ изменяется на каждом временном шаге, если при этом меняется ММП. На рисунке 1.1.3.1 представлены профили E_θ вдоль меридиана утро-вечер для различных комбинаций знаков B_y и B_z ММП. При изменении ориентации ММП появляются топологические различия в форме профилей. Так, при $B_z \leq 0$ поле внутри полярной шапки постоянно и противоположно по знаку полю в авроральной зоне, а при $B_z > 0$ в околополюсной области поле меняет знак. Переход от условий $B_z > 0$ к $B_z < 0$ ведёт к усилению поля и расширению авроральной зоны. При $B_y = 0$ картина симметрична относительно меридиана полдень-полночь. Рост интенсивности B_y ММП ведет к усилению электрического поля в вечернем или утреннем секторе полярной шапки соответственно при $B_y > 0$ или $B_y < 0$. Таким образом, геометрические и

электродинамические характеристики распределений поля определяются входными параметрами модели в виде изменяющихся во времени компонент ММП.

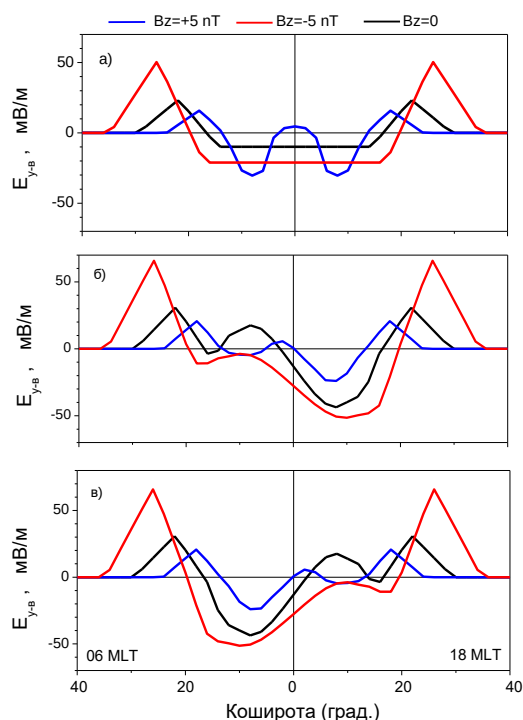


Рисунок 1.1.3.1 – Профили электрического поля вдоль меридиана утро-вечер (06-18 MLT) для различных комбинаций B_z и B_y ММП: (а) $B_y = 0$, (б) $B_y = +5$ нТл, (в) $B_y = 5$ нТл

1.1.4 Результаты тестирования нестационарной модели

Тестирование новой версии модели проводилось на синтетических входных данных. Это позволяло выявить различия в результатах при расчете идеализированных распределений N_e с использованием стационарной и нестационарной версий. Ниже приведены результаты двух серий тестовых расчетов. Синтетические входные параметры B_y и B_z меняли направление в середине 4-часового времени расчета распределения N_e (периода релаксации τ). Полученные пространственные карты выходного параметра максимума электронной концентрации $NmF2$ сравнивались с картами, полученными для постоянных B_y и B_z . Отдельно выявлялись эффекты изменение вертикальной и азимутальной компонент ММП. На рисунке 1.1.4.1 показаны графики входных параметров для первой серии расчетов: ММП B_z меняется от знака (а) с плюса на минус и (б) наоборот, B_y взято нулевым. Чтобы избежать резких скачков электрического поля и потери устойчивости расчетной схемы, B_z меняется не мгновенно, а за четыре шага по времени. Серой точкой на горизонтальной оси обозначен момент времени, который задан как входной параметр и для которого производится расчет.

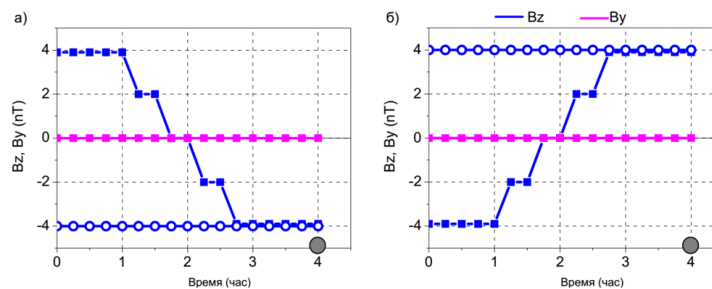


Рисунок 1.1.4.1 – Значения B_z (синий; линия с открытыми точками показывает постоянное $B_z = -4$, а линия с квадратами – переменное B_z) и B_y (сиреневый) в течение 4-часового времени расчета распределения N_e (или периода релаксации τ). Серой точкой на горизонтальной оси обозначен момент времени, для которого производится расчет. (а) B_z меняется от +4 нТ до -4 нТ; (б) B_z меняется от -4 нТ до +4 нТ.

На рисунке 1.1.4.2 показаны карты распределения $NmF2$, рассчитанные по стационарной (вверху) и нестационарной (внизу) версиям, в последней также учитывается предыстория Кр и UT. Сравнивая два типа результатов, можно видеть, что учет изменяющихся внутри периода релаксации входных параметров ММП и других связанных параметров очень существенно влияет на структуру и величину крупномасштабных ионосферных неоднородностей. Они становятся не такими ярко выраженными, с резкими границами, как в стационарном случае, и кажутся более приближенными к реально наблюдаемым формам.

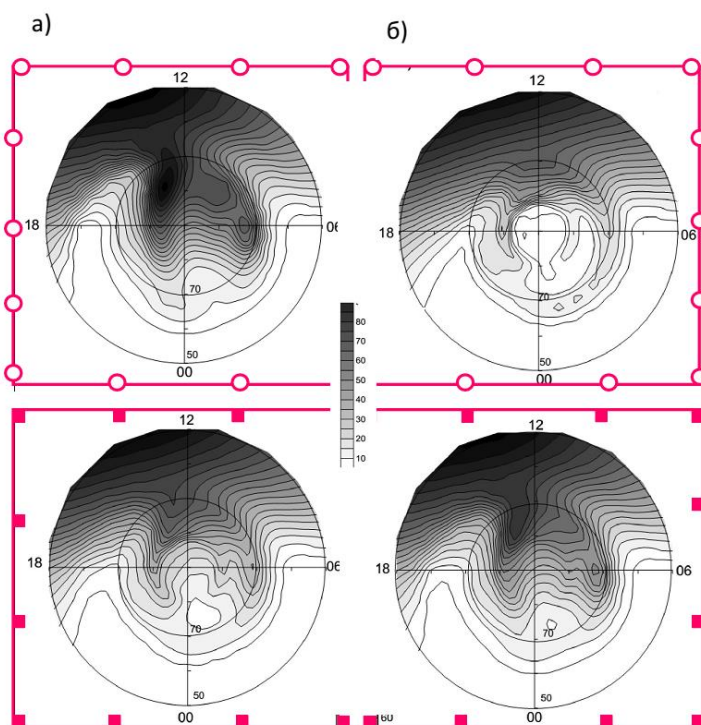


Рисунок 1.1.4.2 – Карты распределения $NmF2$. Верхний ряд соответствует расчету по стационарной модели, второй ряд – по нестационарной; колонка (а) B_z меняется от +4 нТ до -4 нТ; (б) от -4 нТ до +4 нТ. Остальные входные параметры: $V_{св} = 300$ км/с, $K_p = 2$, $F_{10.7} = 150$, $DOY = 30$, $UT = 12$

На рисунке 1.1.4.3 показаны графики входных параметров для второй серии расчетов, когда другая компонента ММП, B_y , меняется от знака (а) с плюса на минус и (б)

наоборот, при $B_z=0$. На рисунке 1.1.4.4 показаны соответствующие карты распределения $NmF2$, рассчитанные по стационарной (вверху) и нестационарной (внизу) версиям. Также, как и в предыдущем случае, можно видеть, что учет изменяющегося ММП, в данном случае – компоненты B_y , приводит к расплыванию неоднородностей. Так, язык ионизации, который ранее был очень ярко выражен, распространяясь на утреннюю или вечернюю сторону в зависимости от знака B_y , теперь становится более симметричным. Полярная полость затягивается.

Таким образом, можно заключить, что нестационарная версия модели дает более точный результат по сравнению со стационарной. Поскольку солнечный ветер очень редко бывает стабильным на временном отрезке 4-5 часов, использование нестационарной модели представляется более корректным.

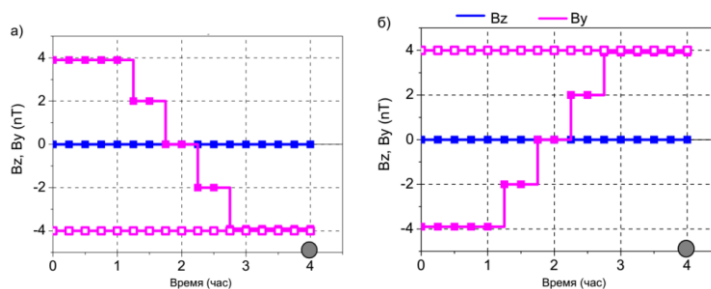


Рисунок 1.1.4.3 – Значения B_z (синий) и B_y (сиреневый; линия с открытыми точками показывает постоянное $B_z=-4$, линия с квадратиками – переменное B_y). (а) B_y меняется от +4 нТ до -4 нТ; (б) B_y меняется от -4 нТ до +4 нТ.

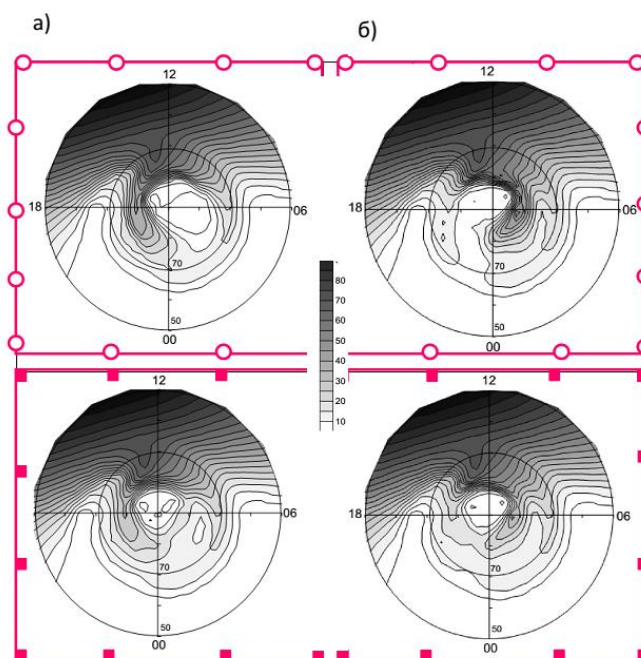


Рисунок 1.1.4.4 – То же, что и рис. 4, но для ММП B_z .

1.2 Анализ эволюции продольных токов во время магнитных бурь по данным измерений спутниками SWARM и оценки динамики границ и плотностей продольных токов, включая мелкомасштабные структуры

1.2.1 Продольные токи во время магнитных бурь

Продольные токи (ПТ) электродинамически связывают ионосферу с пограничными слоями магнитосферы и солнечным ветром. ПТ замыкаются горизонтальными ионосферными токами и являются источниками ионосферных электрических полей. В спокойных геомагнитных условиях величина ПТ составляет порядка миллиона ампер, а во время магнитных бурь увеличиваются на 1-2 порядка. Измерения векторного магнитометра на борту спутников группировки SWARM были использованы для выявления эволюции ПТ во время магнитных бурь. Три идентичных спутника SWARM находятся на низкой полярной орбите. Основным модулем является комплекс для магнитных измерений: высокочувствительный векторный и скалярный магнетометры для определения величины и направления полного вектора и вариаций геомагнитного поля с точностью до 0.1 нТл и частотой 1 Гц. По вариации горизонтальной магнитной компоненты, измеренной над ионосферой, можно получить плотность создающих эту вариацию ПТ.

Были проанализированы крупномасштабные ПТ во время магнитных бурь 24 цикла солнечной активности. Смещение области ПТ к экватору коррелирует с интенсивностью бури (Dst индекс) и связано с расширением аврорального овала. ПТ обнаруживаются вплоть до 50° MLat. На пике главной фазы бури овал в полуденном секторе смещается на несколько градусов больше, чем на ночной стороне, а во время фазы восстановления полуночный овал восстанавливается медленнее, чем дневной. Во время рассмотренных событий экваториальные границы ПТ с высокой степенью корреляции следуют изменению Dst в том случае, если суббури происходят синхронно с развитием магнитной бури и падением Dst . Суббуревая активность определяет также скорость восстановления (смещения к полюсу) границ ПТ. Мгновенные значения широтной границы области, занятой ПТ, слабо коррелируют с AL индексом. Соотношение между сглаженными значениями AL и широтой экваториальной границы ПТ остается постоянным независимо от величины Dst .

Эволюция плотностей ПТ в целом следует тем же тенденциям, что и их широтное положение. В главную фазу бури ПТ усиливаются более чем на порядок, достигая в среднем $3-4 \text{ A/m}^2$ на ночной и 2 A/m^2 на дневной стороне. Максимум плотностей тока достигается во время интенсификации суббуревой активности. Интенсивность ПТ, как крупномасштабных $R1/R2$, так и мелкомасштабных $\sim 10 \text{ км}$ наиболее эффективно контролируется динамикой суббурь, которыми сопровождается магнитная буря. Мелкомасштабные пики плотности ПТ могут быть более чем на порядок больше, чем плотности, усредненные вдоль траектории. Значения $>50 \text{ A/m}^2$ появляются в периоды общего повышения плотности в главную фазу бури во время активации суббурь. Самая большая зарегистрированная плотность 80 A/m^2 – у утекающего ПТ, находящегося в послеполуночном секторе местного магнитного времени во время бури, произошедшей в марте 2015 г. (рисунок 1.2.1.1).

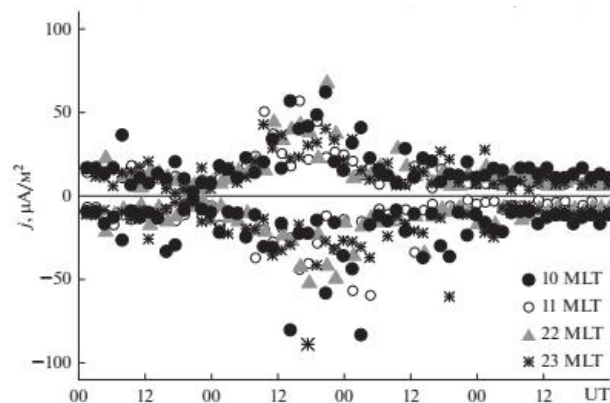


Рисунок 1.2.1.1 - Экстремумы мелкомасштабных ПТ в период магнитной бури 16–19 марта 2015 г.

1.2.2 Развитие информационно-аналитической системы «Гелиогеофизика»

1.2.2.1 Концепция

В настоящее время информационно-аналитические системы (ИАС) становятся все более востребованными в различных сферах деятельности. В научных исследованиях ИАС позволяет специалистам быстро анализировать большие объемы данных с целью мониторинга и прогноза природных процессов, составления целостной картины явления, проверки гипотез. В 2021 г. на базе ГИС «Аврора-Арктика» разработана расширенная концепция ИАС «Гелиогеофизика». ИАС предназначена для решения следующих задач:

- аккумулярование, хранение и актуализация данных из разных источников посредством автоматических сервисов;
- обеспечение исследователей гелиогеофизическими данными, полученными с наземных станций, космических аппаратов (КА), зондирующих ракет, беспилотных летающих аппаратов (БПЛА), радаров и др. инструментов;
- создание универсального стенда для тестирования и использования моделей ионосферы;
- решение научных и прикладных задач на основе эмпирических и теоретических моделей (ионосферы, распространения радиоволн, характеристик аврорального овала и др.);
- модельное сопровождение анализа экспериментальных данных и подбор эффективных моделей в конкретных областях и конкретных условиях;
- использование в образовательной деятельности (подготовка студентов и аспирантов), а также для популяризации науки.

1.2.2.2 Структурная схема ИАС

Схема ИАС (рисунок 1.2.2.2.1) включает в себя системы сбора данных, набор серверов и консоли конечных пользователей.

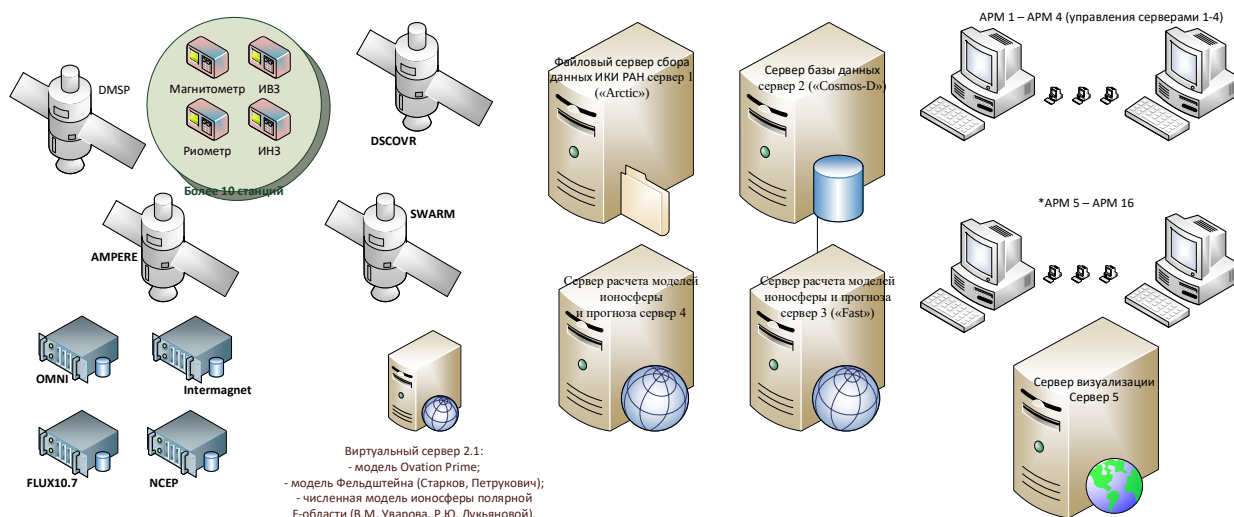


Рисунок 1.2.2.2.1 - Структурная схема ИАС «Гелиогеофизика»

К элементам АРМ 5-16 относятся:

- АРМ-ГФ (Э, И, Д) – «Эксперт», «Исследователь», «Дежурный»;
- АРМ-КП РДВ-W для расчета условий распространения декаметровых волн;
- АРМ-КП РДВ для расчета условий распространения декаметровых волн для автономной и комплексной отладки;
- АРМ-ВШ для динамической модели высокоширотной ионосферы;
- АРМ-РВ для модели распространения радиоволн КВ диапазона;
- АРМ-СШ-F2 для глобальной медианной модели F2-слоя ионосферы;
- АРМ-ВМ для программы расчета сопутствующей информации (географическая, геофизическая, астрономическая);
- АРМ-ТЧ для программы расчета взаимодействия высыпавшихся электронов и протонов в полярную верхнюю атмосферу и ионосферу;
- АРМ-ПН для модели взаимодействия электромагнитных волн КВ диапазона с воздушными и надводными объектами и ионосферными плазменными неоднородностями в арктической зоне.

Серверы, показанные на рисунке 1.2.2.2.1 работают под управлением операционной системы Ubuntu server, дистрибутив Linux, основанный на Debian GNU/Linux (в настоящее время активно развивается и поддерживается свободным сообществом, оперативная система Free Software Foundation, что является аналогом GNU GPL). АРМ работают под управлением операционной системы Linux Mint (дистрибутив Linux, развиваемый сообществом бесплатный дистрибутив Linux, основанный на Ubuntu и Debian, активно используется и поддерживается свободным сообществом, и поставляется в комплекте с обширным набором приложений с открытым исходным кодом, операционная система со свободной лицензией GNU GPL).

1.2.2.3 Наполнение и организация компонентов

В основе базы данных системы лежат пополняемые архивы данных, разрабатываемой в ИКИ РАН ГИС «Аврора-Арктика»:

- Данные по высыпаниям авроральных частиц со спутника-DMSP-18 (серия KA Defense Meteorological Satellite Program, Программа оборонных метеорологических спутников США с открытой базой данных).
- Данные наблюдений параметров солнечного ветра и межпланетного магнитного поля в точке либрации L1 KA DSCOVR (США),.
- Снимки спектро-радиометра среднего разрешения MODIS), который является ключевым прибором на борту KA Terra и Aqua (NOAA), данные радиометра высокого разрешения

AVHRR в пяти спектральных полосах и другие данные ДЗЗ, сопутствующие исследованиям околоземного пространства;

- Фактические и прогнозируемые метеорологические данные NCER (США).
- Снимки автоматических камер всего неба Полярного Геофизического Института на станциях Апатиты, Ловозеро, Баренцбург и других), камеры всего неба камеры и интерферометр Фабри-Перо Института космофизических исследований и аэронавтики им. Ю.Г. Шафера СО РАН (станции Маймага, Жиганск, Тикси);
- Российский сегмент когерентных радаров декаметрового диапазона.

За 2021 год в ИАС добавлены:

- Значения компонент межпланетного магнитного поля, скорости и плотности плазмы солнечного ветра, приведенные к границе магнитосферы, а также индексы геомагнитной активности с разрешением 1 и 5 мин. Данные доступны с 1981 г. через портал OMNI (Центр данных космической физики НАСА).
- Среднесуточные значения индекса $f_{10.7}$ потока радиоизлучения от Солнца на длине волны 10,7 см (2800 МГц) с 1999 г. через портал Канадского космического агентства.
- Минутные вариации ортогональных компонент магнитного поля Земли с 1991 г. через портал глобальной наземной сети геомагнитных обсерваторий Intermagnet и сеть магнитовариационных станций.

В перспективе будут добавляться данные орбитальных проектов ИКИ РАН, например КА «Ионосфера» и «Зонд», КА «Арктика-МП» и др., измерения наземных сетей РФ: магнитометров, риометров, ионозондов вертикального и наклонного зондирования, а также аналогичных свободно распространяемых данных, собираемых, в первую очередь, в арктической зоне, однако база данных построена без ограничений охвата, все данные собираются в полном объёме, и включают полный охват территории сбора.

В рамках создания универсального стенда для тестирования и использования моделей ионосферы, ведется работа по добавлению, тестированию и сравнению моделей ионосферы.

В системе уже работают модели границ аврорального овала Фельдштейна и OVATION Prime, реализована статистическая модель широтного положения дневного каспа, а также модель крупномасштабных ионосферных неоднородностей в F области высокоширотной ионосферы. В режиме тестирования запущена работа модели расчета и прогнозирования условий распространения радиоволн на среднеширотных и высокоширотных радиотрассах в различных геофизических условиях. Ведется разработка универсальной системы доведения специализированной информации на базе программного обеспечения с открытым исходным кодом.

1.3 Разработка общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли

1.3.1 Исследование структуры верхней/средней атмосферы при формировании экстремально сильного стратосферного вихря зимой 2019-2020 гг. в Арктике

Экстремальное климатическое событие произошло в северном полушарии зимой 2019-2020 гг. Сформировалась необычная конфигурация стратосферного полярного вихря, которая не подвергалась эрозии, как это обычно бывает, планетарными волнами, идущими вверх от тропосферы, а отражала эти волны, так что они распространялись обратно вниз. Эти уникальные условия позволили вихрю оставаться экстремально сильным и холодным (абсолютный максимум скорости стратосферного ветра по данным атмосферных реанализов начиная с 1960-х) в течение нескольких месяцев. Во время сильного стратосферного полярного вихря поверхностные температуры характеризовались аномальным теплом в Евразии (евразийское тепло было беспрецедентным за период с 1980-х) и холодом в Канаде, Гренландии и Аляске.

Большое количество работ, в том числе опубликованных в тематическом выпуске рейтинговых научных журналов *Geophysical Research Letters* и *J. Geophysical Research Atmosphere*, посвящено исследованию метеорологических эффектов арктической зимы 2019-2020 гг. В то же время, сам механизм и условия формирования экстремально сильного стратосферного вихря не были выявлены. Необычная структура вихря позволяет предположить, что область верхней атмосферы, выше стратопавзы могла быть вовлечена в данный климатический экстремум. Для доказательства необходимо было привлечь массив данных высотных измерений. В проведенном исследовании на основе анализа глобальных и локальных наблюдений была определена структура и эволюция термодинамических полей в верхней стратосфере и мезосфере зимой 2019/2020 гг. Следует отметить, что текущий период низкой солнечной активности благоприятен для более четкого выявления внутриатмосферной изменчивости на больших высотах.

1.3.2 Инструментарий: лимбовые измерения MLS радиометра спутника Aura, радиозондирование нижней ионосферы метеорным радаром, атмосферный реанализ

Спутник Aura имеет солнечно-синхронную полярную орбиту на высоте около 700 км и делает в сутки 13 витков вокруг Земли. Зонд MLS на борту спутника детектирует слабое микроволновое излучение от лимба Земли по направлению полета с вертикальным разрешением 8 км и горизонтальным 170×12 км. Вертикальный профиль температуры строится на 54 уровнях до 0,0001 гПа с точностью 1–2,5 К.

Метеорный радар (МР) - это серийно выпускаемый интерферометрический радар всего неба SKiYMET со стандартизированным программным обеспечением для обработки данных. Он состоит из одной передающей антенны на частоте 36,9 МГц и пяти антенн, принимающих отражения от ионизованных метеорных следов. Наблюдения за положением и скоростью нескольких тысяч радио эхо в сутки позволяют определять зональную и меридиональную составляющие скорости нейтрального ветра на интервалах высот от 82 до 98 км с вертикальным разрешением 3 км и временным разрешением 1 час. Суточные данные о температуре на высоте максимального обнаружения метеоров ~90 км рассчитываются по времени распада метеорных следов, обусловленного амбиполярной диффузией. В работе использовались измерения МР геофизической обсерватории Соданкюля (СГО, 67° с.ш., 23° в.д.) Зимой 2019/2020 СГО МР обеспечил измерения вблизи края стратосферного полярного вихря, где наиболее сильная аномалия геопотенциальной высоты при уровне давления 10 гПа составляла примерно $-5 \cdot 10^4$ м. Траектория спутника Aura проходит над локацией СГО МР в 3 и 12 часов местного времени.

Данные климатического реанализа с MERRA-2 до высоты 0,1 гПа с 3-часовым временным разрешением и пространственным разрешением $0,5^\circ \times 0,625^\circ$ использовались для построения стратосферных ветров в декабре 2019 г. – феврале 2020 г. Данные доступны на основе системы NASA Geospatial Interactive Online Visualization And ANalysis Infrastructure (GIOVANNI), разработанной Годдардовским Центром данных и информационных услуг по наукам о Земле (GES DISC).

1.3.3 Структура теплового поля

Температурные поля MERRA показывают проявление планетарной волны 1 в высотном диапазоне от мезосферы до тропосферы. На рисунке 1.3.3.1 показана диаграмма температур, усредненных по 60–90° с.ш., в формате «время-долгота» на уровне 1 гПа (~47 км) за три зимних месяца. Диаграмма позволяет визуализировать динамику процесса по статическим распределениям. На рисунке 1.3.3.1а показана эволюция теплового поля в верхней стратосфере, а на рисунке 1.3.3.1б – карта приземной температуры. Можно

видеть, что в первой декаде января наблюдается усиление зональной асимметрии волны 1 в восточном полушарии по всей высоте атмосферы. При этом наблюдения метеорного радара проводятся в непосредственной близости от гребня волны.

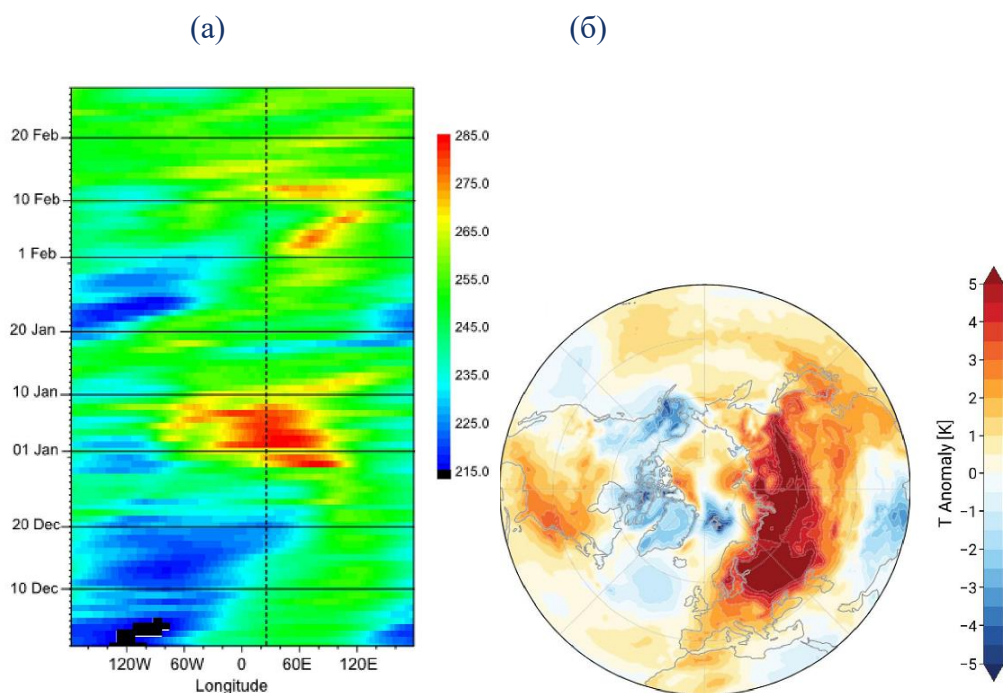


Рисунок 1.3.3.1 – (а) эволюция поля температуры на высоте 1 гПа, усреднение по 60–90° с.ш. Вертикальной линией отмечена долгота расположения СГО МР; (б) карта наблюдаемых аномалий приземных температур в январе–феврале 2020 г.

Спутниковые измерения позволяют восстановить глобальное тепловое поле до высот 100 км. На рисунке 1.3.3.2 показана эволюция температур MLS для локации СГО МР (рисунок 1.3.3.2а) и симметричной локации в западном полушарии (рисунок 1.3.3.2б). Данные MLS показывают: (1) наличие глобальной долготной асимметрии, (2) формирование вертикальной инверсии в конце декабря: потепление в верхней части стратосферы до 280К и похолодание мезосферы до 180К, (3) Резкий срыв процесса по всей толще атмосферы в середине января: внезапное похолодание с обратным эффектом в мезосфере.

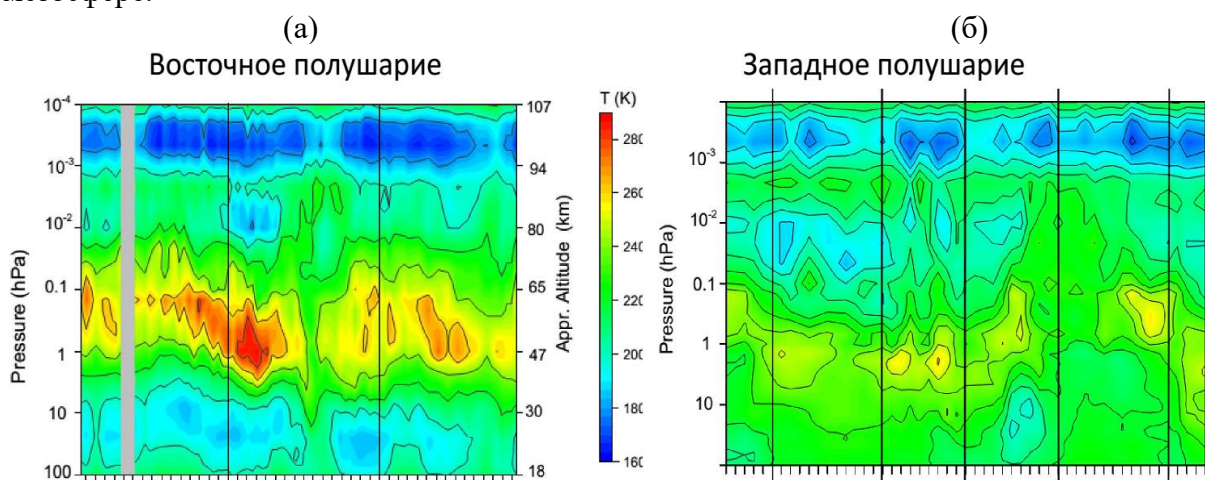


Рисунок 1.3.3.2 - Вертикальный разрез поля температуры по данным Aura MLS декабрь 2019-февраль 2020 гг.: (а) для локации СГО МР в восточном полушарии и (б) для симметричной локации в западном полушарии

1.3.4 Структура ветрового поля

Вертикальное сечение циркумполярного потока до высот 0.1 гПа (~45 км) может быть построено по данным реанализа. На больших высотах, в мезосфере-нижней термосфере, метеорный радар - единственный инструмент, способный регулярно измерять ветер. СГО МР находится в непосредственной близости от основных термических и циркуляционных возмущений. Вертикальные сечения ветра на высотах 82–98 км над локацией СГО МР в декабре-феврале показаны на рисунке 1.3.4.1. В зональном потоке (рисунок 1.3.4.1а) выделяются волновые структуры, состоящие из четырех короткоживущих импульсов усиления зонального потока до скоростей >50 м/с. Ядро зонального ветра находится экстремально высоко в термосфере, выше 100 км, и в начале января резко опускается вниз в верхнюю стратосферу, образуя при этом несколько сдвиговых слоев. В меридиональной составляющей потока (рисунок 1.3.4.1б) доминируют 16-дневные планетарные волны.

Стена опускающегося зонального ветра продолжается в стратосфере. Вертикальный разрез стратосферных ветров на высотах 100 – 0.1 гПа может быть получен по данным реанализа MERRA. Региональный зональный ветер над локацией СГО МР (рисунок 1.3.4.2) в начале января демонстрирует резкое опускание ядра с 60 до 35 км, а также спектр планетарных волн.

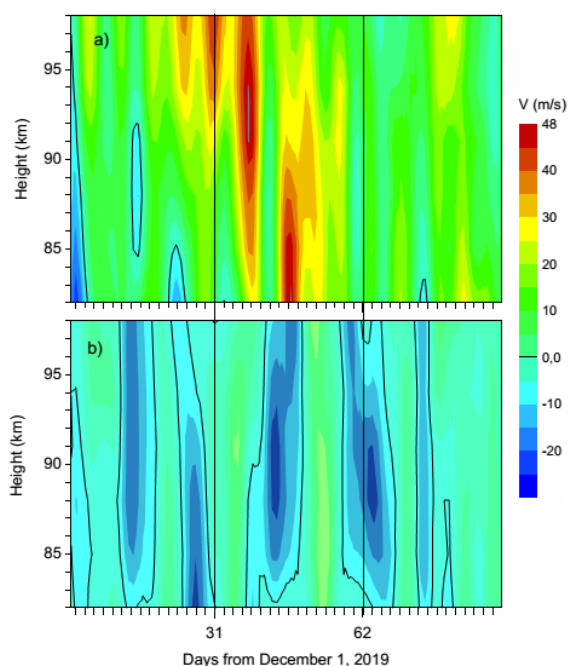


Рисунок 1.3.4.1 - Вертикальный разрез поля (а) зонального и (б) меридионального ветра по данным СГО МР. Положительные значения - направление на восток

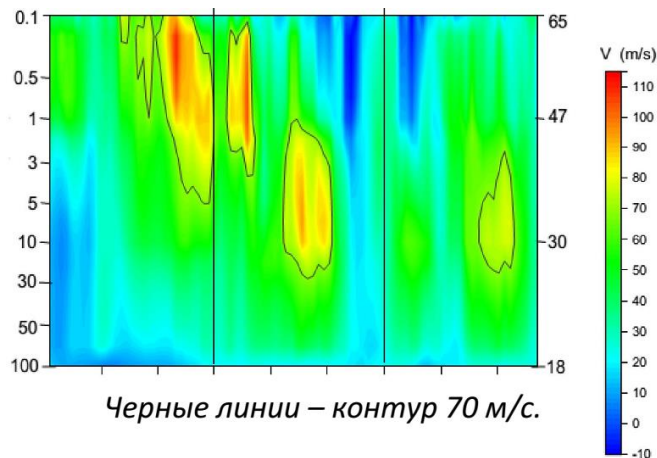


Рисунок 1.3.4.2 - Вертикальный разрез поля зонального ветра в полярной стратосфере

Таким образом, анализ данных показывает, что формирование стены горизонтального сдвига ветра и расслоение зонального потока происходит в широком диапазоне высот от нижней термосферы до стратосферы. Сдвиг зонального ветра формирует высокую и прочную «стену» или «шапку», окружающую полярный вихрь и отражающую волны, идущие из низких широт и тропосферы, что способствует устойчивости вихря. Подобные условия могут реализовываться в южном полушарии в Антарктике, но уникальны для северного, в котором захват полярным вихрем воздуха на термосферных высотах ранее не наблюдался. Сам механизм и причины появления зимнего циркумполярного потока на высотах мезосферы-нижней термосферы в сезоне 2019-2020 гг. остаются пока невыясненными. Возможно, это связано с усилением высотной межполушарной циркуляции или формированием субтропического высотного джета. Выяснение этих вопросов требует дальнейших исследований с привлечением дополнительных данных.

Заключение

Представленные в настоящей главе результаты показывают, что по направлению «Развитие комплекса взаимосвязанных и дополняющих друг друга физико-математических моделей для диагностики и прогнозирования состояния высокоширотной ионосферы в северном и южном полушариях, которые электродинамически связаны между собой через область средних широт. Разработки общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли» проведенные работы соответствуют установленному плану работ на 2021 год и выполнены полностью в соответствии с планом.

При этом получены следующие основные результаты:

- Проведена модернизация численной модели F области высокоширотной ионосферы с целью перехода к нестационарному режиму. Разработана и протестирована новая версия в нестационарном приближении, которая позволяет учесть изменение входных параметров внутри периода времени релаксации между стартовой и расчетной точками. Это приближает модель к воспроизведению реальных условий, а также расширяет возможности постановки численных экспериментов, выявляющих зависимость крупномасштабных ионосферных неоднородностей от ММП в высоких широтах. Можно заключить, что нестационарная версия модели дает более точный результат по сравнению со стационарной. Поскольку солнечный ветер очень редко бывает стабильным на временном отрезке 4-5 часов, использование нестационарной модели представляется более корректным и повышает точность моделирования.

- Получены оценки мелкомасштабных продольных токов во время магнитных бурь 24 солнечного цикла. Самая большая зарегистрированная плотность $80 \text{ } \mu\text{A}/\text{m}^2$ – у втекающего ПТ, находящегося в послеполуночном секторе местного магнитного времени. Мелкомасштабные пики плотности ПТ могут более чем на порядок превышать средние крупномасштабные плотности токов зон R1/R2.
- На базе ГИС «Аврора-Арктика» разработана расширенная концепция ИАС «Гелиогеофизика». Функционал ИАС расширен с целью комплексирования с другими ГИС, базами данных и моделями, разрабатываемыми в ИКИ РАН, увеличения информационного наполнения и более широким спектром инструментов и возможностей, предоставляемым пользователям для решения задач исследования солнечного ветра, магнитосферы и ионосферы.
- Выявлены механизмы термических и динамических возмущений на высотах от верхней стратосферы до мезосферы-нижней термосферы в период формирования аномально сильного арктического стратосферного вихря зимой 2019-2020 гг. В мезосферных ветрах преобладают несколько импульсных возрастаний зональной составляющей, событие резкого опускания зонального ветра и чередование северных и южных течений с периодом длинных планетарных волн. Понижение температуры на высоте ~ 90 км, сопровождалось сильными термическими инверсиями и сдвигами ветра. Сформировались высокие «стены», окружающие полярный вихрь, что способствовало его необычной устойчивости. Такие условия могут реализовываться в южном полушарии, но уникальны для северного.

ГЛАВА 2 Применение методов машинного обучения и глубокого обучения для обработки и анализа больших массивов астрономических данных в задачах наблюдательной астрофизики

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 настоящей главе отчета представлены основные результаты работ по данному направлению, полученные по следующим разделам плана работ ИКИ РАН на 2021 год:

- Разработка и применение алгоритмов глубокого обучения для задачи морфологической классификации галактик на оптических изображениях. Исследование и разработка алгоритмов кросс-отождествления объектов в фотометрических каталогах обзоров неба, полученных из различных спектральных диапазонов
- Исследование алгоритмов машинного обучения для измерения фотометрических красных смещений галактик

Результаты, полученные по данным направлениям, описаны в разделах 2.1, 2.2, 2.3 настоящей главы. Подробно полученные в данном направлении результаты изложены в научной публикации [2], а также представлены в рамках 10-и устных докладов и тезисах на научной конференции: «Ломоносовские чтения 2021» [17-19], «Тихоновские чтения 2021» [20-23], и международной конференции «Астрофизика высоких энергий сегодня и завтра 2021 (HEA-2021)» [24,25].

2.1 Разработка и применение алгоритмов глубокого обучения для задачи морфологической классификации галактик на оптических изображениях. Исследование и разработка алгоритмов кросс-отождествления объектов в фотометрических каталогах обзоров неба, полученных из различных спектральных диапазонов

13 июля 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). Главной задачей обсерватории «Спектр-РГ» является 4-х летний рентгеновский обзор всего неба с беспрецедентной чувствительностью. 8 декабря 2019 года обсерватория начала сканирование неба в режиме обзора, в июне 2021 года выполнила третий обзор неба, а в середине декабря 2021 года закончила четвертый обзор неба в рентгеновских лучах (то есть половину всех запланированных наблюдений обзора всего неба). Разработка и применение алгоритмов интеллектуального анализа больших данных на основе машинного и глубокого обучения к многоволновым данным по более чем миллиону рентгеновских источников СРГ, ключевым образом влияет на качество научных результатов в ходе данной космической миссии. По направлению «Космос-Д-Астрофизика», основные усилия были направлены на совершенствование системы SRGz (создаваемой в отделе Астрофизики высоких энергий ИКИ РАН) для интеллектуального анализа данных обсерватории Спектр-РГ (в первую очередь - телескопа eROSITA). Скопления галактик и сверхмассивные черные дыры (квазары) являются основными целями рентгеновского телескопа eROSITA при проведении обзора неба. В 2021 году наши усилия были сконцентрированы, во-первых, на автоматическом извлечении информации по скоплениям галактик из многоволновых данных спутника Planck, и, во-вторых, на построении вероятностных моделей оптического отождествления точечных рентгеновских источников для обзора всего неба СРГ/eROSITA.

Информация по скоплениям галактик во Вселенной может быть получена из трех основных диапазонов: рентгеновского, микроволнового и оптического. Основная цель работы в 2021 году — создание нейросетевого алгоритма сегментации и детекции для автоматического извлечения информации по скоплениям галактик eRosita в микроволновом диапазоне, основанном на анализе искажений производимых гало темной материи на карте реликтового излучения (эффект Сюняева-Зельдовича (CЗ)). Мы построили алгоритмы глубокого обучения на базе адаптированной архитектуры U-Net с применением техники активного обучения для сегментации скоплений галактик в данных спутника Planck, выполнившим наиболее глубокий обзор всего неба в микроволновом диапазоне. Ниже на Рисунке 2.1.1 показана архитектура системы для нейросетевой сегментации и построения каталога кандидатов в скопления галактик по данным спутника Planck.

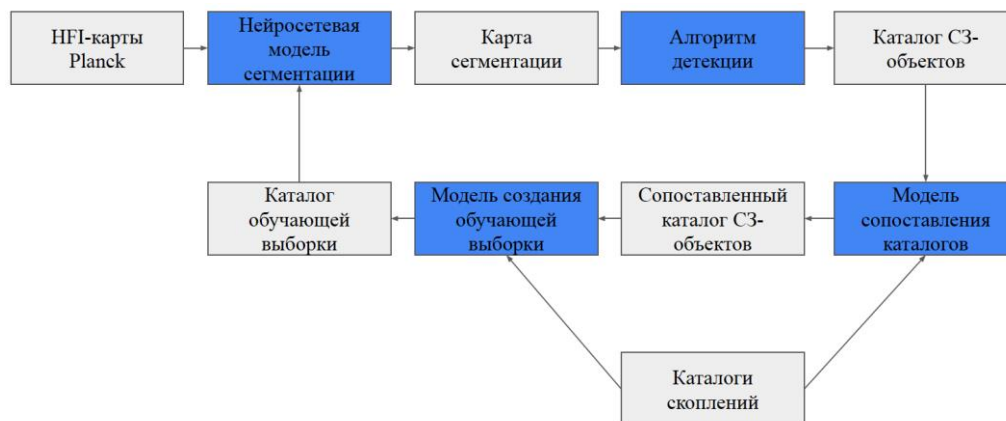


Рис.2.1.1 Архитектура системы для нейросетевой сегментации и построения каталога кандидатов в скопления галактик (CЗ-объектов) по данным спутника Planck

В качестве обучающей выборки выбраны каталоги скоплений в литературе полученные классическими методами детектирования по микроволновым (PSZ2, ACT) данным и рентгеновским данным спутника ROSAT (MCXC). Изображения для нейросети были предобработаны в стандартной пикселизации неба HEALPix. Наша нейросетевая модель позволила извлечь из данных Planck наиболее полный и точный, на данный момент, каталог кандидатов в скопления галактик в области неба обзора eROSITA. Далее мы сравнили полученный каталог кандидатов в скопления с каталогами рентгеновских скоплений eRosita в восточной галактической полусфере. На графике Рис.2.1.2 показана зависимость Полнота—Точность для предложенной нами нейросетевой модели (SZcat(AL)) и для моделей сегментации на микроволновых данных Planck построенных в литературе - нейросетевой (Bonjean 2019, полностью воспроизведенной нами с нуля - SZcat(base)) и полученной из готовых карт параметра комптонизации (SZcat(y-map)). Как можно видеть из графика, предложенная нами нейросетевая модель заметно лучше моделей из литературы.

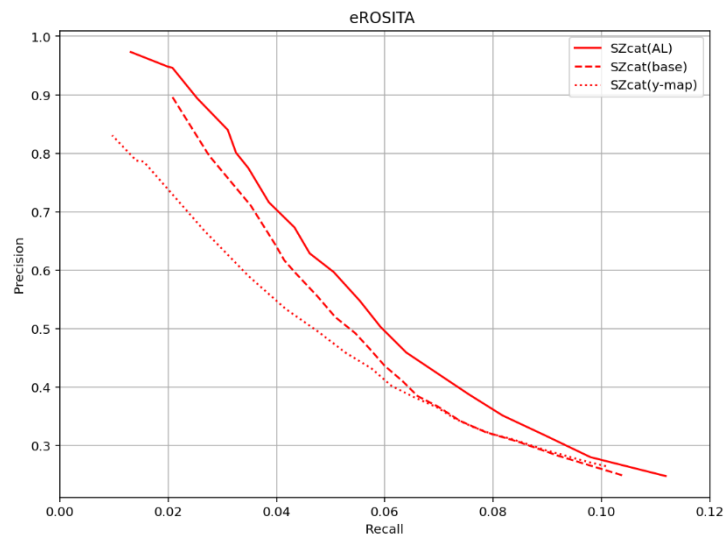


Рис. 2.1.2 Зависимость Полнота—Точность для предложенной нами нейросетевой модели (SZcat(AL)) и для моделей в литературе - нейросетевой (SZcat(base)) и полученной из готовых карт параметра комптонизации (SZcat(y-map)).

Нами были построены вероятностные модели оптического отождествления точечных рентгеновских источников для обзора всего неба СРГ/еРОЗИТА в области Lockman Hole по данным фотометрического обзора DESI Legacy Imaging Surveys, которая учитывает как позиционную, так и фотометрическую информацию об оптических объектах в окрестности рентгеновских источников и позволяет производить отбор оптических компаньонов с точностью и полнотой на уровне $\approx 94\%$ (для рентгеновских источников с потоком в диапазоне 0.5-2кэВ выше $1e-14$ эрг/с/см², что соответствует ожидаемой характерной чувствительности после 4-х лет обзора СРГ/еРОЗИТА). Предложенная модель оптического отождествления в дальнейшем будет использована в составе системы SRGz анализа данных рентгеновского обзора всего неба.

В 2021 году в рамках направления «Космос-Д-астрофизика» мы рассмотрели несколько нейросетевых подходов, представляющих потенциальный интерес для астрофизических исследований: (i) комбинаторная оптимизации в задачах построения маршрутов транспортных средств с ограничениями, (ii) применение визуальных трансформеров в задачах распознавания действий в видеоряде в условиях ограниченной разметки.

Первая задача довольно тесно связана с задачей массового кросс-отождествления большого числа источников из многоволновых каталогов и в дальнейшем мы планируем применить полученные нами наработки в области обучения с подкреплением для задачи комбинаторной оптимизации для астрофизических задач.

Проблема построения и оптимизации маршрутов представляет собой класс задач комбинаторной оптимизации с многочисленными практическими применениями: курьерская доставка, работа такси, логистика товаров между складами и точками продаж. Существуют программные инструменты, позволяющие получить субоптимальное решение для классической постановки задачи ограниченного размера (например “Google OR-Tools”). Однако, для больших задач (>500 точек) и динамических задач, в которых новые точки, требующие быстрого пересчета пути, появляются по ходу выполнения маршрута, не было предложено эффективных полиномиальных алгоритмов. В прикладных областях возникают дополнительные ограничения, такие как: временные окна доступности клиентов, ограничение на последовательность посещения, ограничения вместимости грузовых средств. Классические подходы для решения подобных задач делятся на два вида: поиск точных решений при помощи целочисленного

программирования и эвристические подходы. Точные решения часто разрабатываются на основе метода ветвей и границ, который имеет теоретическую гарантию нахождения оптимального решения, но неприменим для реальных задач из-за экспоненциальной сложности. На практике эвристика лучший выбор: несмотря на отсутствие теоретической гарантии качества, субоптимальное решение может быть получено за адекватное время. Однако разработка эвристики требует большого количества проб и ошибок, а производительность с точки зрения качества решения во многом зависит от интуиции и опыта людей-экспертов. Решение для задачи построения маршрута представляет собой последовательность узлов (местоположение клиентов и складов). Глубокие нейросетевые модели обучения для обработки естественного языка, которые имеют дело с последовательностями, вероятно, являются идеальным выбором для обработки решений построения маршрутов. Именно поэтому в последнее время стали популярны нейронные подходы, основанные алгоритмах внимания для решения задачи оптимизации маршрутов с ограничениями. В данной работе исследованы нейросетевые алгоритмы внимания при обучении с подкреплением в задачах оптимизации маршрутов. Рассматриваемые задачи содержат ограничения на вместимость транспортных средств и временные окна доступности клиентов. В качестве базового подхода используется эвристический алгоритм “Google OR-Tools”.

Вторая задача, по которой была нами проведена исследовательская работа в 2021 году также важна для астрофизики тем, что видеоданные постепенно входят в круг астрофизических исследований с запуском проектов многократного сканирования неба, таких как ZTF, LSST (в скором будущем). Многократные изображения неба позволяют изучать астрономические объекты в динамике. Таким образом, далее мы планируем перенести наши достижения в распознавании действий на данных видеоряда при помощи моделей трансформеров в астрофизические исследования транзиентных и переменных объектов.

Распознавание действий людей по видео имеет также множество применений в реальной жизни. Современные решения в данной области задействуют методы машинного обучения, при использовании которых необходимы новые данные для каждой прикладной задачи, а также для каждого нового набора действий людей. Для обработки больших наборов данных требуются краудсорсинговые решения, применение которых требует значительных денежных затрат и контроля качества работы. Минимизация разметки данных позволяет сэкономить время и деньги разработчиков моделей машинного обучения. Классическим решением задачи распознавания действий по видео является применение свёрточных нейронных сетей, таких как I3D и SlowFast. Однако в последнее время архитектура трансформера на основе механизма внимания, пришедшая из обработки естественных языков, показывает хорошие результаты и для видео данных. Её преимуществом является сокращение времени обучения относительно свёрточных подходов при сохранении качества. Также модель трансформера отличает возможность выбора соседей блоков кадра, что увеличивает гибкость модели и позволяет подстраивать её под особенности данных. В данной работе предложен подход с обучением модели трансформера с двумя различными функциями потерь, в том числе с Triplet Loss, для применения которой реализован подход Сиамских сетей. Такое решение позволяет обучить модель получать признаки, удаленные для различных классов и сгруппированные для примеров одного класса. Это также позволяет не осуществлять обучение большой модели в процессе разметки, а обучать лишь легковесный классификатор, который работает с признаками действий. Точность модели оценивалась на выборке UCF-101, на котором она превзошла по точности классификации рассмотренные свёрточные нейронные сети. Помимо классификации тестовой выборки, были проведены эксперименты с классификацией новых для модели трансформера классов действий по различному числу имеющихся размеченных экземпляров.

2.2 Исследование алгоритмов машинного обучения для измерения фотометрических красных смещений галактик

Классификация космических объектов имеет фундаментальное значение для многих областей астрономических исследований, в особенности для разметки данных, получаемых из рентгеновского обзора SRG/eROSITA. Целью данного направления работы в 2021 году являлось создание эффективных моделей машинного обучения для классификации рентгеновских объектов на основе многоволновых данных по схеме звезда-галактика-квазар. Наши модели импользуют высококачественные однородные фотометрические обзоры значительных площадей неба такие, как SDSS, Pan-STARRS, DESI Legacy Imaging Survey и WISE. В индустрии и в астрономии для задачи классификации по табличным данным чаще используются древовидные архитектуры такие, как случайный лес и градиентный бустинг. Также была предложена модель глубокой нейронной сети TabNet для табличных данных. В данной работе использовались табличные оптические признаки из всех ранее представленных обзоров, где разметка получена из обзора SDSS. В качестве тестовой выборки были взяты рентгеновские объекты из 3-го обзора SRG/eROSITA с величиной потока $> 1e-14$. Было произведено сравнение упомянутых ранее моделей на представленных данных. Как превосходящий остальные алгоритмы использовался градиентный бустинг реализации LightGBM. Глобальные гиперпараметры модели бустинга настраивались при помощи байесовской оптимизации (с применением библиотеки HyperOpt). Для оценки качества алгоритма использовалась двукратная кросс-валидация, оценивались следующие метрики: Recall — полнота классификации, ROC AUC — площадь под графиком операционной характеристики приемника (ROC) является стандартной оценкой качества ранжирования.

Полученный нами результат на выборке рентгеновских источников после 1.5 лет обзора СРГ/еРОЗИТА имеющих спектроскопический класс (STAR/GALAXY/QSO по данным обзора SDSS) показан в таблице на Рис.2.2.1.

STAR			GALAXY			QSO		
recall	roc-auc	Кол-во объектов	recall	roc-auc	Кол-во объектов	recall	roc-auc	Кол-во объектов
0.853	0.981	2284	0.950	0.997	1714	0.985	0.988	81386

Рис.2.2.1 Полнота классификации (recall) и качество ранжирования (roc-auc) по результатам модели градиентного бустинга на тестовой выборке рентгеновских источников СРГ/еРОЗИТА со спектральной классификацией SDSS.

Таким образом, полученные нами модели классификации достигают высокой полноты и качества ранжирования для всех классов объектов и вошли в стандартный пайплайн SRGz анализа данных обзора СРГ/еRosita.

В 2021 нами были также построены новые модели вероятностной регрессии на основе квантильного случайного леса для прогнозирования фотометрических красных смещений точечных и протяженных рентгеновских источников. Модели были протестированы путем прямых измерений спектроскопических красных смещений у внегалактических объектов (квазаров и массивных далеких скоплений галактик), которые показали их высокую точность и эффективность.

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 2021 г. тематике. Таким образом, работы выполнены полностью в соответствии с планом.

При этом получены следующие основные результаты:

- Разработана нейросетевая модель для сегментации скоплений галактик в многоволновых данных телескопа Planck. На ее основе был получен наиболее полный каталог кандидатов в скопления в микроволновом диапазоне и измерена его полнота и точность.
- Разработаны модели (древовидные и нейросетевые) машинного обучения для классификация (по схеме звезда/галактика/квазар) оптических объектов по фотометрическим данным в полях рентгеновских источников и показана высокие характеристики полноты и точности ранжирования для модели градиентного бустинга на тестовых данных 3-го обзора неба СРГ/ePOЗИТА. Модель включена в состав системы SRGz.
- Разработана модель оптического отождествления СРГ/ePOЗИТА и измерены ее основные характеристики (точность и полнота в задаче отбора оптических компаньонов рентгеновских источников). Модель включена в состав системы SRGz.
- Исследованы и разработаны модели на базе визуальных трансформеров для распознавания действий в данных видеоряда в условиях ограниченной разметки, полученные модели превзошли лучшие результаты в литературе на выборке UCF-101, изучены возможности их применения в астрономических исследованиях.

Таким образом, работы по направлению «Астрофизика» темы Космос-Д в 2021 году были выполнены в полном объеме.

ГЛАВА 3 Разработка новых методов и подходов организации хранения и параллельной обработки сверхбольших объемов данных для решения задач анализа долговременных наблюдений в интересах ведения мониторинга и изучения различных процессов и явлений

Введение

Настоящая глава посвящена описанию работ по разработке новых методов к организации хранения и параллельной обработки данных сверхбольших архивов ДЗЗ, позволяющих эффективно решать задачи анализа долговременных наблюдений по большим территориям в интересах ведения мониторинга и изучения различных процессов и явлений.

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 блочного (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 наборов данных ресурсы. Такое 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, когда для решения конкретных задач по анализу и обработке спутниковых данных необходимо использовать различные пространственные разбиения для широкого спектра разных типов данных. В частности, такая задача 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. В этих случаях подготовка и хранение одних и тех же типов спутниковых наблюдений в разных пространственных разбиениях оказывается чересчур затратной.

Для того чтобы использовать все основные преимущества блочного доступа к данным и избежать излишних накладных расходов, в 2019 году в рамках работ по данному направлению 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 полномасштабно внедрена в эксплуатацию для задач построения безоблачных атмосферно скорректированных ежедневных композитов по

данным прибора КМСС, установленного на спутниках серии Метеор-М. В частности, такие ежедневные композиты были получены за весь 2020 год по большой территории, соответствующей зерновому поясу России. В процессе реального массового применения разработанной технологии был выявлен и исправлен ряд недостатков в ее реализации. Также был добавлен дополнительный функционал, в первую очередь, связанный с решением задачами контроля выполнения обработки и анализа получаемых на ее основе результатов. Также в 2021 году были проведены исследовательские работы по анализу эффективности процедур предоставления блоков данных под обработку, которые привели к разработке методики оценки временных затрат на подготовку требуемого набора данных с использованием имеющейся аппаратной инфраструктуры.

Другим важным применением разработанной в ИКИ РАН технологии динамического блочного доступа к данным при создании обновленного картографического веб-интерфейса, использующего тайловую схему подготовки изображений. В 2021 году новый картографический интерфейс был существенно доработан, в частности, существенно увеличена скорость его работы и устранены основные проблемы со стабильностью его функционирования под большой нагрузкой. Это, в свою очередь, позволило внедрить его в эксплуатацию в рамках системы «Вега-Science», а также в ряде других систем, входящих в проект «Созвездие-Вега». Новый интерфейс предлагает пользователям более удобный и распространенный в настоящее время стиль работы с пространственными данными. В рамках создания обновленного картографического веб-интерфейса был также реализован новый универсальный модуль растеризации изображений, благодаря которому удалось достичь высокой скорости работы интерфейса.

В 2021 году были также продолжены работы по развитию подсистемы анализа временных и пространственных рядов данных на основе графиков. Благодаря использованию единого реестра показателей, отображаемых в виде графиков, удалось существенно повысить гибкость реализованной подсистемы, что позволяет создавать новые тематические интерфейсы анализа рядов данных с минимальными трудозатратами.

Также в 2021 году были продолжены работы по развитию методов параллельной обработки данных, в частности, были проанализированы и внедрены различные варианты распараллеливания обработки на каждом из вычислительных узлов.

Задача организации хранения и обработки сверхбольших объемов спутниковых данных для решения задач мониторинга и изучения различных процессов и явлений неразрывно связана с необходимостью обеспечения бесперебойного функционирования сложного программно-аппаратного комплекса. В 2021 году активно велись работы по совершенствованию системы контроля функционирования сложного программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг», на базе которого реализованы и используются описанные выше технологии.

Полученные в рамках работ 2021 года результаты по данному направлению изложены в следующих публикациях [3-5, 7, 9-15].

3.1. Разработка подходов и методов реорганизации существующих долговременных архивов данных спутниковых наблюдений Земли для повышения эффективности организации их предоставления системам распределенного многопоточного (распараллеленного) анализа.

3.1.1 Модернизация программной реализации технологии динамического блочного доступа

В 2021 году были продолжены работы по совершенствованию программной реализации технологии динамического блочного доступа к данным для проведения их эффективной параллельной обработки. В процессе внедрения представленной технологии

возникла необходимость в поддержке более широкого спектра параметров, описывающих задания на предоставление данных под обработку, касающихся как описания исходных данных в архивах, так и формата их предоставления. Была улучшена эффективность работы ключевых процедур формирования блоков данных и передачи их на обработку. Также были исправлены ошибки, связанные с округлением координат ячеек разбиения, которые проявились только при наборе очень большой статистики выполнения процедур.

Также в процессе внедрения технологии возникла необходимость в более полном протоколировании всех операций по подготовке и предоставлению требуемых наборов данных для обработки, а также фиксации статусов обработки каждого из предоставленных блоков данных. С этой целью была разработана специальная служебная база данных, структура основных таблиц которой приведена на рисунке 3.1.1.1. При этом в базу данных помещается информация не только о подготовленных блоках данных и параметрах их подготовки, но и об исходных данных, на основе которых они были получены. Это позволяет полностью контролировать как процесс подготовки, так и процессы их обработки и архивации. Особенно это актуально для сложных процедур обработки спутниковых данных, в рамках которых результаты обработки подготовленных блоков данных могут быть различными, в частности, данные могут быть отбракованы по степени облачности, недостаточному для используемого алгоритма обработки комплекта соседних блоков данных и качеству привязки. Таблица jobs содержит информацию об отдельных запусках процедур подготовки данных для указанного задания, таблица job_scenes – информацию об используемых для подготовки каждого из блока данных исходных сценах спутниковой информации. В таблицу job_bfiles протоколируется информация о всех подготовленных блоках данных и статусах их обработки, а таблица bfiles_scenes является связующей, и отражает тот факт, что для каждый из блоков может быть подготовлен на основе набора исходных сцен.



Рисунок 3.1.1.1 - Структура базы данных для протоколирования операций с блоками данных

3.1.2 Анализ эффективности процедур подготовки и предоставления требуемого набора блоков данных

В рамках работ, проводимых в 2020 году, была разработана методика выбора оптимального размера блоков данных для проведения конкретной процедуры по их обработке. Однако в рамках этой методики не был еще детально проработан вопрос анализа эффективности самих процедур подготовки и предоставления данных. Поэтому в 2021 году основным направлением изучения было выбрано исследование зависимостей времени подготовки блоков данных от множества влияющих на него факторов. К таким факторам можно отнести производительность серверов обработки и систем хранения данных, скорость сетевых соединений, размер и тип предоставляемых спутниковых данных и др. В статье представлена методика, позволяющая достаточно оперативно оценить ресурсы, необходимые для предоставления заданного набора блоков данных с нужными характеристиками. Она позволяет на этапе реализации нового типа обработки оценить временные затраты на предоставление блоков данных, оценить требуемые для этой задачи ресурсы и выработать оптимальное решение, исходя из реальной производительности имеющейся инфраструктуры.

3.1.2.1 Методика оценки временных затрат на подготовку блоков данных

В рамках разработанной технологии динамического блочного доступа к данным в архивах спутниковых данных подготовка блоков данных производится в параллельном режиме на кластере специализированных серверов, на которых реализован прямой доступ к файлам данных в архиве. При этом в каждом из информационных узлов распределенного архива устанавливается группа таких серверов, обслуживающих запросы на получение данных из соответствующего раздела архива. За формирование к ним HTTP запросов на получение требуемых для проведения обработки файлов данных отвечает сервера диспетчеризации, который реализуют гибкий механизм балансировки нагрузки, учитывающий физическое расположение исходных файлов данных в архиве. Процедура формирования блока данных реализована на основе свободно распространяемого ПО GDAL (Geospatial Data Abstraction Library). В рамках ее выполнения производится чтение только необходимых фрагментов исходных файлов данных и формируется файл в формате GeoTIFF со следующими заданными характеристиками: размер, проекция, разрешение, набор каналов и алгоритм сжатия. После этого файл пересылается по сети на сервер, выполняющий функцию дискового буфера, который обеспечивает кэширование подготовленных блоков данных и является источником исходной информации для проведения одной или нескольких обработок. Для того чтобы оценить суммарное время подготовки данных, необходимых для проведения конкретного типа обработки необходимо получить среднее время выполнения запроса на получение одного блока данных.

Время выполнения запроса на подготовку блока данных зависит от множества факторов, включая вышеперечисленные характеристики подготавливаемых файлов, а также типа исходных данных, схемы организации их хранения, скорости доступа к конкретным файловым серверам и производительности сетевых каналов. Установление всех зависимостей без использования упрощенной модели требует очень большого времени и ресурсов, что не позволяет оперативно оценить требуемое время для нового типа обработки. Для построения такой модели было проанализировано влияние наиболее значимых из них, и был установлен целый ряд упрощенных зависимостей, которые можно использовать для оценки времени подготовки разных выборок исходных данных. В частности, время передачи данных по сети, также как и время, необходимое для их сжатия не зависят существенно от типа исходных данных и организации их хранения. Также в рамках выработанной модели используются зависимости времени подготовки блоков

данных не от их размера в географических координатах, а от размера в пикселах изображения. Важно отметить, что описанная ниже модель является приближенной, а оценки для выбранного на ее использования варианта могут быть при необходимости уточнены. В качестве основного времени, которое необходимо смоделировать было выбрано время подготовки данных по одному типу (продукту) исходных данных. Ниже приводится формула для оценки этого времени как сумма разных зависимостей, которые могут быть установлены экспериментально:

T-product(n, arch_type, dst_proj, compression) =

$$\begin{aligned} & \mathbf{T\text{-}base(n, arch_type, dst_proj)} \\ & + \mathbf{T\text{-}compression(n, compression)} \\ & + \mathbf{T\text{-}get(size(n, compression))} \end{aligned}$$

Где **n** – линейный размер получаемого блока данных в пикселах в предположении, что блоки имеют квадратную форму,

arch_type – совокупность характеристик исходных файлов, ключевыми из которых являются их проекция и характерный размер фрагментов,

dst_proj – проекция получаемых данных,

compression – используемый алгоритм сжатия,

size(n, compression) – размер блока данных после применения заданного алгоритма сжатия

Первое слагаемое в формуле **T-base** описывает базовое (основное) время формирования блока данных в зависимости от его размера, проекции и фрагментации исходных данных и проекции, в которой должны быть получены результаты. Существенно, что полученная для конкретного типа данных зависимость может быть применена и для похожих по организации хранения других типов спутниковых данных. Второе слагаемое **T-compression** соответствует времени, необходимому для сжатия получаемых данных по блоку и зависит только от размера блока и используемого алгоритма сжатия. Третье слагаемое **T-get** определяет время, необходимое для передачи результирующего файла по сети, и зависит только от размера данных. Далее в статье кратко приводятся основные методы для оценки каждой из рассматриваемых выше зависимостей с примерами полученных зависимостей.

3.1.2.2 Получение базовой зависимости T-base

Для получения базовой зависимости времени подготовки блоков данных от их размера в пикселах необходимо получить это время для различных значений такого размера. Важно отметить, что время подготовки блока данных существенно зависит от скорости доступа к конкретной системе хранения файлов, на которой находятся необходимые для его формирования исходные данные. Разница в производительности разных систем хранения может быть связана с особенностями их аппаратной реализации, степени загруженности и пропускной способностью сетевых каналов связи. Также время формирования блоков данных может зависеть и от ряда других факторов, таких как географические координаты запрашиваемой области и степень наполненности ее данным. Поэтому необходимо для каждого размера блока получить среднее значение времени по всем блокам данных, которые должны быть предоставлены под обработку. В рамках представленной методики для оценки этого времени берется среднее значение по набору случайно выбранных блоков данных, относящихся к искомой выборке. При этом для каждого размера блоков эта выборка генерируется заново, что позволяет избежать

кэширования чтения исходных файлов данных, которое может существенно исказить результаты.

Ниже на рисунке 3.1.2.2.1 приводятся примеры базовых зависимостей, полученных при формировании блоков по данным прибора OLI_TIRS, установленного на спутниках серии Landsat, за период времени с 2013 по 2021 год и диапазон координат, примерно соответствующий территории России. Приведены зависимости для случая с сохранением проекции UTM и для случая с получением результата в географической проекции. Проведенные исследования показали, что при наборе достаточной статистики получаемые зависимости хорошо аппроксимируются полиномами второй степени, что позволяет на основе небольшого числа замеров для различных размеров блока в пикселах, достаточно точно оценить время подготовки для любого такого размера. То обстоятельство, что при уменьшении размера блока время подготовки не стремится к нулю связано с наличием накладных расходов на каждый запуск процедуры формирования блока данных. Важно отметить, что существенное падение эффективности дисковых операций чтения и записи данных при уменьшении объемов данных, с которыми они оперируют, вносит существенный вклад практически во все исследуемые зависимости.

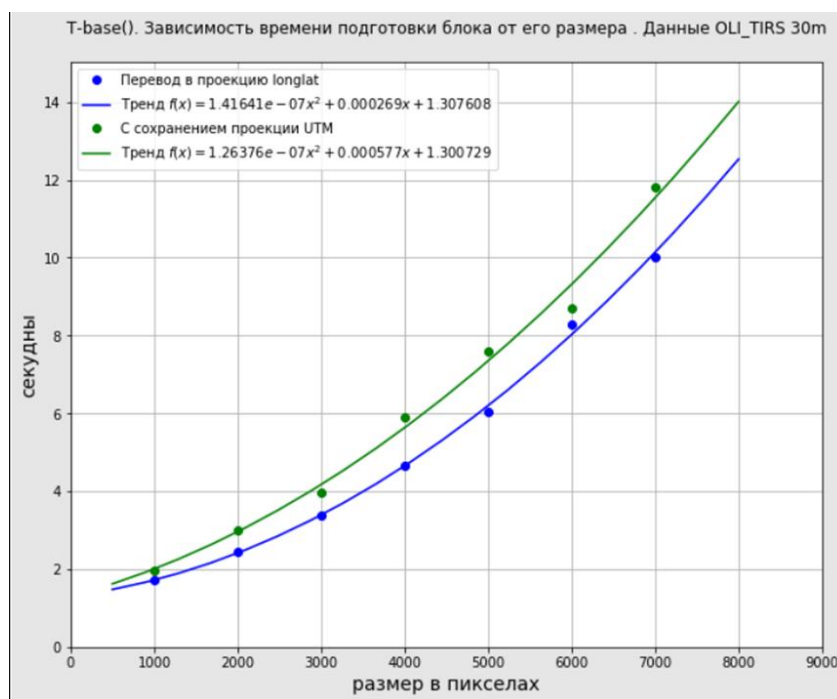


Рисунок 3.1.2.2.1 - Примеры полученных базовых зависимостей

Для того чтобы оценить эффективность процедур формирования данных для различных размеров блоков, можно получить зависимость удельного времени подготовки фиксированной области данных от размера блока в пикселах. Пример такой зависимости приведен на рисунке 3.1.2.2.2. Из нее видно, что удельное время падает при увеличении размера блока, но для больших значений этого размера этот эффект практически исчезает.

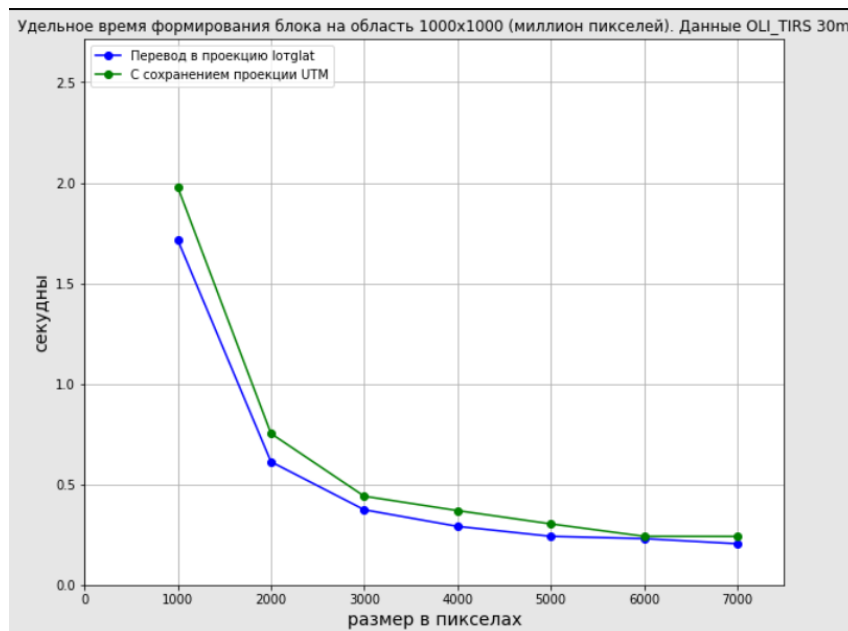


Рисунок 3.1.2.2.2 - Пример зависимости удельного времени подготовки области данных от размера блока

3.1.2.3 Получение зависимости T-compression

Зависимость времени сжатия от размера блока может быть получена вычитанием из зависимости времени подготовки блоков с использованием заданного алгоритма сжатия зависимости, полученной для этой процедуры без сжатия. Иллюстрация такого подхода приведена на рисунке 3.1.2.3.1. На нем приведены кривые, соответствующие времени подготовки данных без сжатия, с использованием разных алгоритмов сжатия, поддерживаемых ПО GDAL, и кривые, соответствующие искомой зависимости T-compression. Для иллюстрации выбран наиболее быстрый из используемых на практике алгоритм сжатия LZW и более медленный и эффективный алгоритм DEFLATE. Отметим, что получаемые зависимости времени, затрачиваемого на сжатие блоков, от их размера также хорошо аппроксимируются полиномами второй степени.

Также была оценена степень сжатия данных при использовании каждого из вышеперечисленных алгоритмов. При сжатии полностью заполненных данными блоков процент сжатия для алгоритма DEFLATE стабильно держится в районе 22%, а для алгоритма LZW – размер файлов даже возрастает примерно на 3%. Однако использование алгоритма LZW все же может быть целесообразным когда значительная часть блоков данных заполнена данными не полностью и используются предыдущие версии ПО GDAL, в которых хранение таких данных реализовано неэффективно.

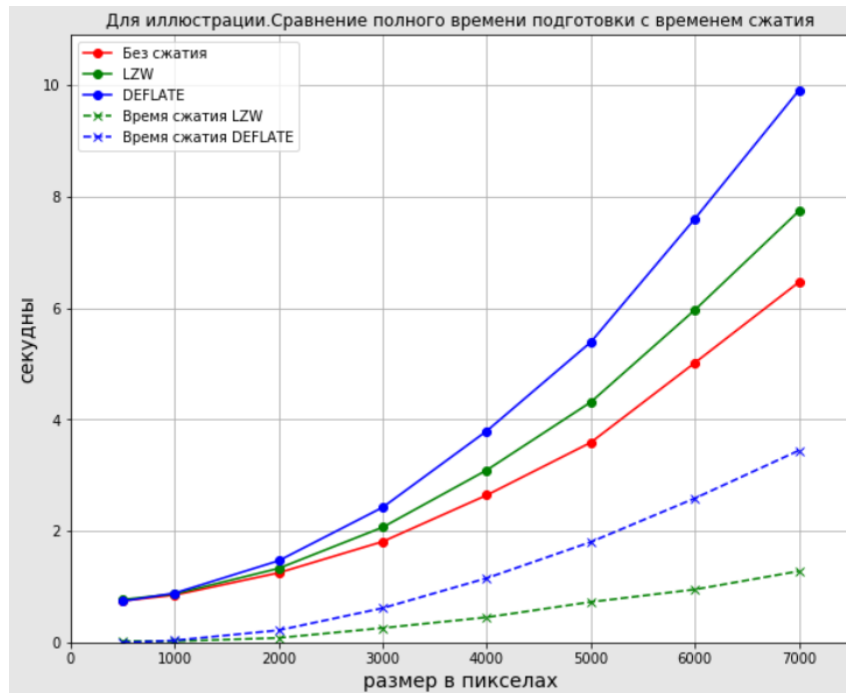


Рисунок 3.1.2.3.1. Пример выявления зависимостей времени сжатия блоков данных от их размера.

3.1.2.4 Получение зависимости T-get

Так как время скачивания файлов данных при использовании стандартного протокола HTTP практически не зависит от их содержимого, то для начала можно установить зависимость времени скачивания файлов от их размера в мегабайтах, характерную для имеющейся инфраструктуры. Пример такой зависимости, хорошо аппроксимируемой линейной функцией, приведен на рисунке 3.1.2.4.1.

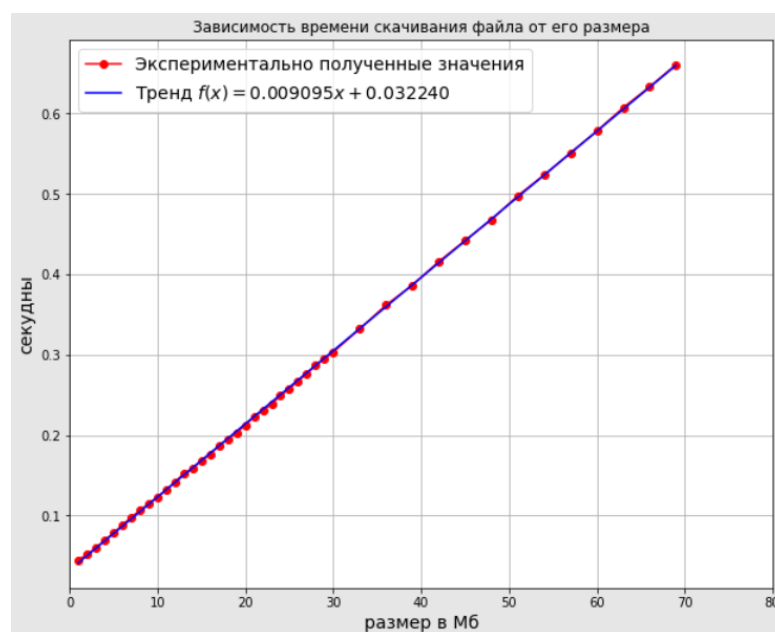


Рисунок 3.1.2.4.1 - Пример зависимости времени получения файлов от их размера в мегабайтах

Далее можно определить искомые зависимости времени скачивания блоков данных от их размера в пикселах при использовании различных алгоритмов сжатия, которые аппроксимируются полиномами второй степени с вычисляемыми на основе установленной линейной зависимости коэффициентами, приведенной на рисунке 3.1.2.4.2.

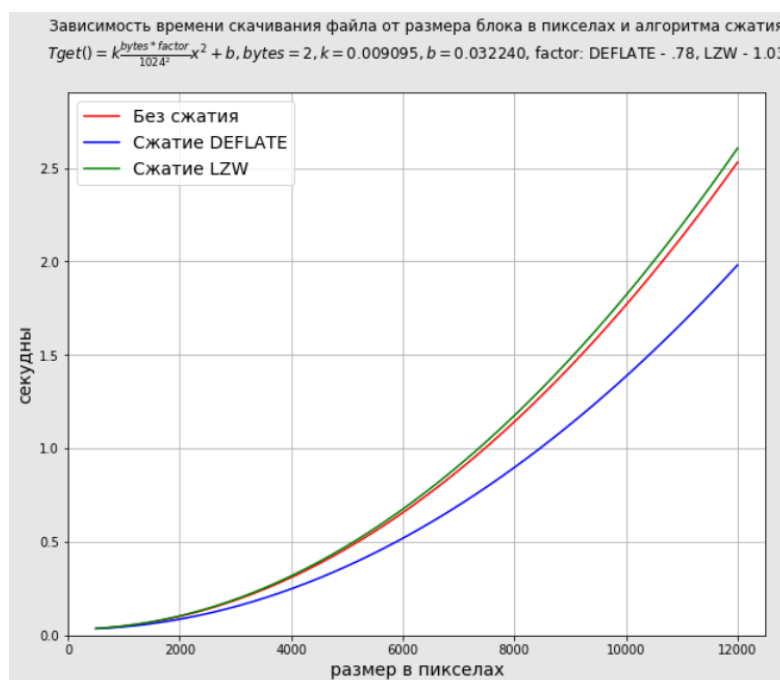


Рисунок 3.1.2.4.2 - Примеры зависимостей времени скачивания блоков данных от их размера в пикселах при использовании разных алгоритмов сжатия.

3.1.2.5 Вывод

Представленная методика позволяет оперативно оценить временные затраты на подготовку заданного набора требуемых для обработки данных на базе имеющейся инфраструктуры хранения и обработки данных. Такие оценки могут быть использованы для более точного выбора оптимальных характеристик предоставляемых в обработку блоков данных, а также для того, чтобы определить, какие вычислительные ресурсы необходимо задействовать для ее проведения в заданные сроки. Последняя из задач особенно актуальна при проведении массовой обработки спутниковых данных по большим территориям.

3.1.3 Организация обработки данных 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. На основе использования рассматриваемой технологии была реализована комплексная процедура обработки для получения ежедневных безоблачных композитных изображений КМСС, которые могут быть использованы для количественной оценки состояния земной поверхности. Первые результаты были получены в 2020 году, а в 2021 году технология была существенно доработана, что позволило запустить разработанную процедуру обработки в эксплуатацию.

Реализованная процедура получения безоблачных композитных изображений по данным приборов серии КМСС базируется на совместной обработке этих данных с данными прибора MODIS. Эти типы данных существенно отличаются как пространственным разрешением, так и схемой фрагментации данных в архивах. В частности, хранение данных КМСС реализовано в виде сцен в географической проекции, на которые разбиты сеансы данных, а хранение используемых данных прибора MODIS в регулярном разбиении на фрагменты в синусоидальной проекции. Для того, чтобы реализовать их эффективную совместную обработку на кластере вычислительных узлов они должны быть приведены к единому пространственному разбиению. Поэтому для предоставления всех необходимых для обработки данных этих приборов в едином пространственном разрешении и проекции используется разработанная в ИКИ РАН технология динамического блочного доступа. При этом получаемые в результате проведения двухэтапной процедуры обработки данные также формируются в этом фиксированном пространственном разрешении, что упрощает их дальнейшее использование. Первый вариант подсистемы организации рассматриваемой процедуры обработки был реализован в 2020 году. В 2021 году были устранены выявленные в ее работе недостатки, а также был добавлен новый функционал, отвечающий за контроль выполнения всех этапов выполнения процедуры. В частности, было реализовано формирования ежедневного отчета о выполнении оперативных и ретроспективных задач по обработке данных, предоставляемых на основе механизма динамического блочного доступа с автоматической диагностикой их успешности.

Принципиальная схема организации обработки для получения безоблачных композитов КМСС приведена на рисунке 3.1.3.1. Как показано на этой схеме данные КМСС поступают на обработку из трех центров «НИЦ «Планета», в каждом из которых функционируют сервисы предоставления блоков данных, а данные прибора MODIS поступают из архива, находящегося в ИКИ РАН. В рамках выполнения задания на подготовку данных под обработку сначала производится запрос к каталогам распределенных архивов, а затем на основе полученной информации формируются запросы на получение требуемых блоков данных к соответствующим сервисам их предоставления. После этого с использованием специального сервера балансировки нагрузки производится многопоточное скачивание необходимых блоков данных. В рамках первого этап рассматриваемой процедуры обработки производится радиометрическая и геометрическая коррекция данных. После чего полученные калиброванные данные по сеансам КМСС заносятся в архив. После того как необходимый набор таких данных поступает в БД, выполняется второй этап процедуры, а именно, построение интерполированных безоблачных ежедневных композитов.

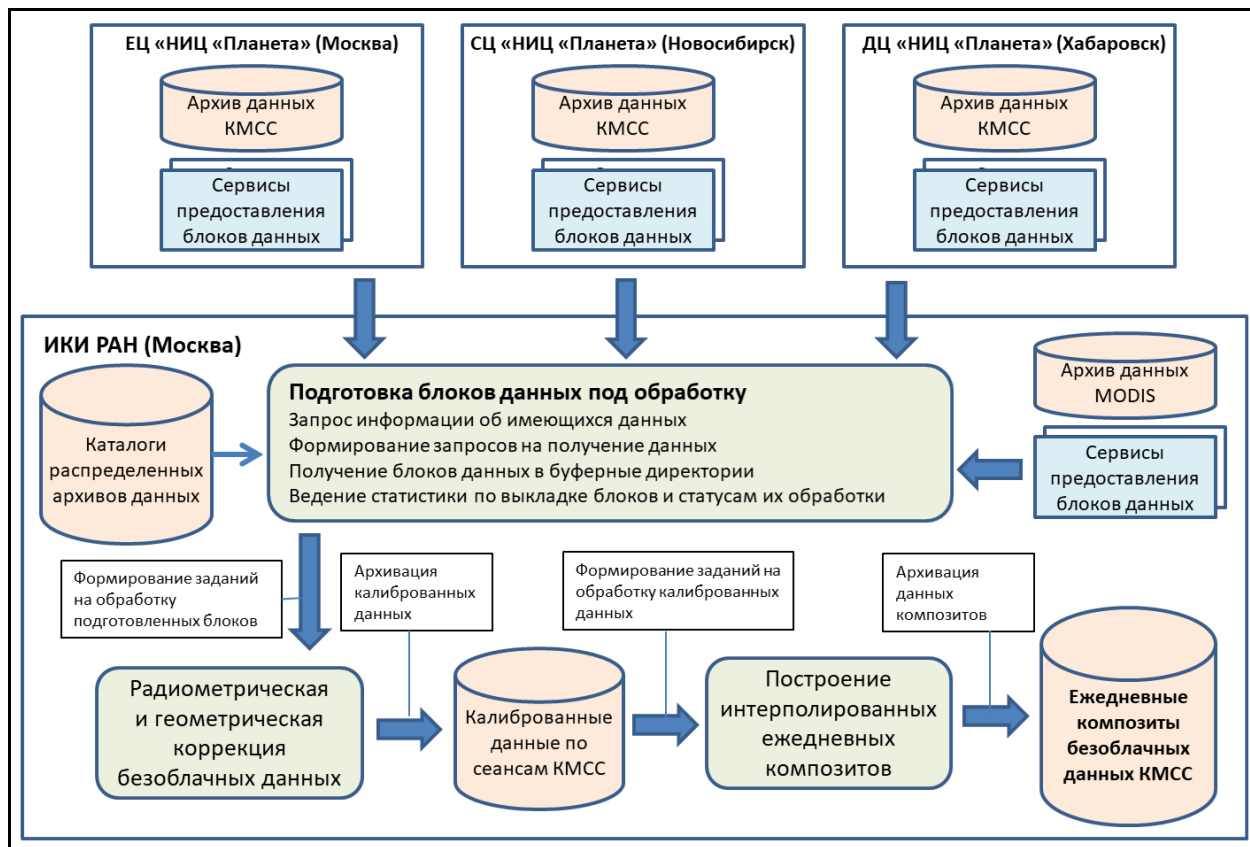


Рисунок 3.1.3.1 - Принципиальная схема организации обработки для получения безоблачных композитов КМСС

3.1.4 Разработка инструментов для визуального анализа наличия пространственных данных

Многолетний опыт получения и обработки спутниковых данных по большим территориям показывает, что для оценки ситуации с наличием требуемых наборов данных по разным территориям бывают очень полезны интерактивные интерфейсы, которые позволяют отобразить требуемую информацию в виде «карты». Они, конечно, не могут заменить автоматически получаемую диагностику о поступлении требуемых данных, но для оценки сложившейся ситуации с наличием данных по разным территориям могут быть очень полезны. В частности, на основе такой информации может быть принято решение о повторном получении недостающих наборов данных из центров их распространения или о необходимости получения их на основе обработки других имеющихся в архивах данных. При этом, хотя такая задача актуальна для любых данных, хранящихся в фиксированных разбиениях, особую актуальность она приобретает для задач по обработке данных, полученных на основе динамического блочного доступа.

Как уже было отмечено выше, в 2021 году в рамках реализации технологии динамического блочного доступа было поддержано детальное протоколирование всех операций по предоставлению и обработке блоков данных, включая информацию о результатах их обработки. Наличие такой информации позволяет реализовать интерфейс для визуального анализа числа блоков с различным статусом обработки, а также числа блоков полученных информационных продуктов по каждой из ячеек разбиения. Пример такого интерфейса, реализованного для контроля проведения обработки данных прибора КМСС, приведен на рисунке 3.1.4.1. Кроме отображения количества блоков в нем также реализован ряд вычисляемых показателей, в частности: процент блоков заданного типа от числа всех поступивших на обработку блоков, максимальный интервал в днях между соседними данными в ячейках разбиения, влияющий на качество интерполяции данных и другие. В получаемой в результате запроса с заданными параметрами карте в каждой ячейке разбиения указывается требуемый численный показатель, а использование

градаций серого цвета для фона ячеек позволяет визуально оценить значения этих показателей для различных областей карты.

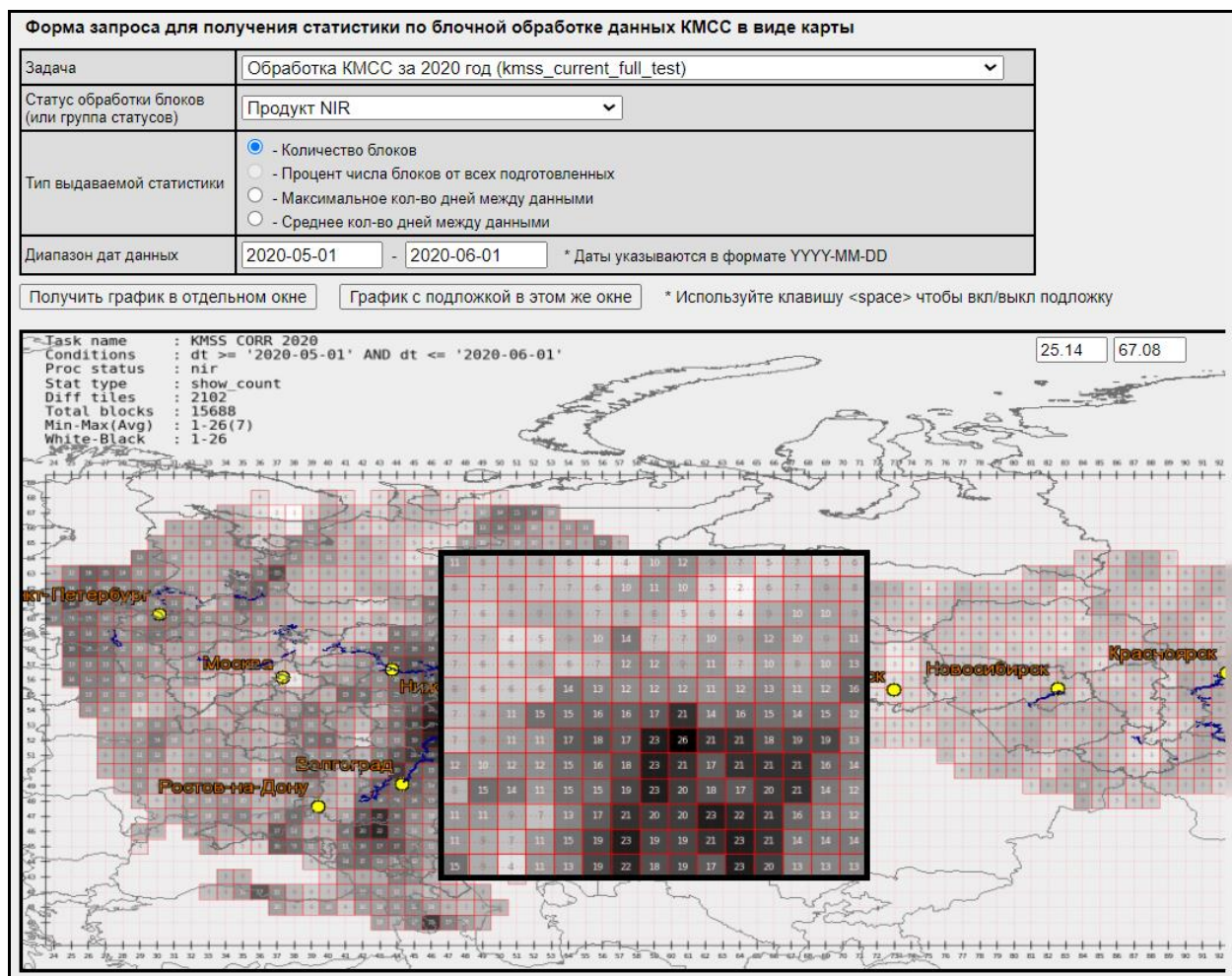


Рис. 3.1.4.1 - Интерфейс для визуального анализа наличия заданного типа блоков данных КМСС

На рисунке 3.1.4.2 также представлен пример еще одного интерфейса для визуального анализа наличия пространственных данных, который был также разработан в 2021 году и предназначен для контроля поступления в архивы данных спутников серии Sentinel 2. Особенностью схемы хранения этого типа данных является использование сложного пространственного разбиения на тайлы. В рамках этого разбиения каждая UTM зона делится на набор пересекающихся тайлов 110 на 110 километров, причем соседние тайлы перекрываются на 10 километров, при этом количество получающихся для разных зон тайлов может существенно различаться. В результате отображение контуров ячеек этого разбиения на карте оказывается малоинформативным, особенно в крупном масштабе, поэтому в интерфейсе была реализована «таблица» с информацией по каждой из зон, которая попадает в зону интереса. По клику на ячейку этой таблицы в текстовом поле выдается информация по наличию экземпляров данных для каждого из относящихся к ней тайлов. Использование различных цветовых палитр для фона ячеек таблицы позволяет более наглядно оценить разницу в наличии данных по разным зонам UTM.

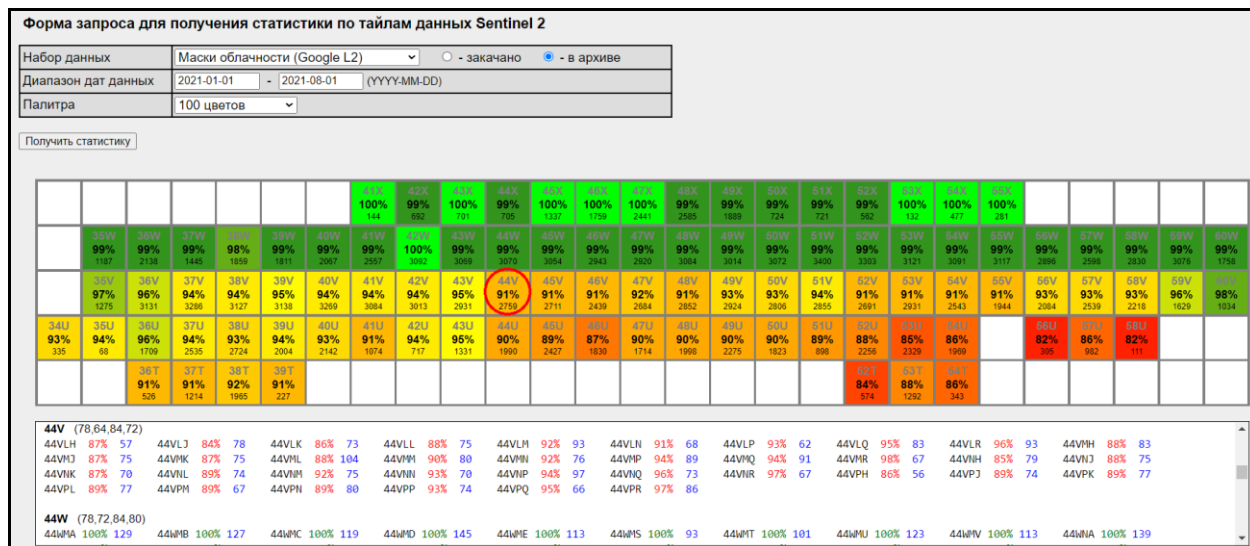


Рисунок 3.1.4.2 - Интерфейс для визуального контроля наличия в архиве данных прибора MSI спутников серии Sentinel 2

3.2. Использование технологии динамического блочного доступа для организации доступа к данным архивов ЦКП «ИКИ-Мониторинг»

3.2.1. Обновленный картографический веб-интерфейс, использующий тайловое формирование 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».

В 2020 году были начаты работы по использованию динамического блочного доступа для реализации тайлового доступа к спутниковым и другим изображениям в картографическом веб-интерфейсе системы «Вега-Science», являющемся основным инструментом для работы с данными архивов ЦКП «ИКИ-Мониторинг». В рамках этого варианта формирования данных реализуется асинхронный запрос требуемого изображения небольшими фрагментами (по умолчанию 256x256 пикселей) с использованием стандартных пространственных разбиений для разных масштабов. Такой подход к предоставлению изображений 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 различных информационных продуктов.

В рамках системы «Вега-Science» решается задача предоставления доступа пользователям к широкому спектру различных информационных продуктов, причем с возможностью использованию различных картографических проекций. Кроме того, пользователю предлагается большой набор интерактивных инструментов для анализа и обработки спутниковых изображений и других пространственных данных. Такая высокая вариативность достигается благодаря использованию динамического формирования требуемых информационных продуктов на базе имеющейся в архиве информации. Поэтому вариант с предварительной подготовкой всех вариантов информационных продуктов в виде тайлов в данном случае оказывается абсолютно неприемлемым. Для того чтобы сохранить имеющийся в системе «Вега-Science» функционал, но при этом

обеспечить более комфортный для пользователя тайловый вариант доступа к изображениям в интерфейсе, было решено использовать разработанную в ИКИ РАН технологию динамического блочного доступа. При этом по результатам макетирования различных вариантов нового интерфейса для этого было решено использовать открытую библиотеку «leaflet.js», которая обладает большим набором дополнительных плагинов, а также легко наследуемых и расширяемых компонент.

В 2021 году новый картографический интерфейс был существенно доработан, в частности, была существенно увеличена скорость его работы и устранены основные проблемы со стабильностью его функционирования под большой нагрузкой. Это, в свою очередь, позволило внедрить его в эксплуатацию в рамках системы «Вега-Science», а также в ряде других систем, входящих в проект «Созвездие-Вега». Новый интерфейс предлагает пользователям более удобный и распространенный в настоящее время стиль работы с пространственными данными. При этом благодаря эффективному кэшированию тайлов изображений не только в браузере пользователя, но и на предназначенных для этого специализированных серверах, удалось существенно повысить максимальное число пользователей, которые могут одновременно работать с пересекающимися наборами данных. Внешний вид нового тайлового веб-интерфейса приведен на рисунке 3.2.1.1.

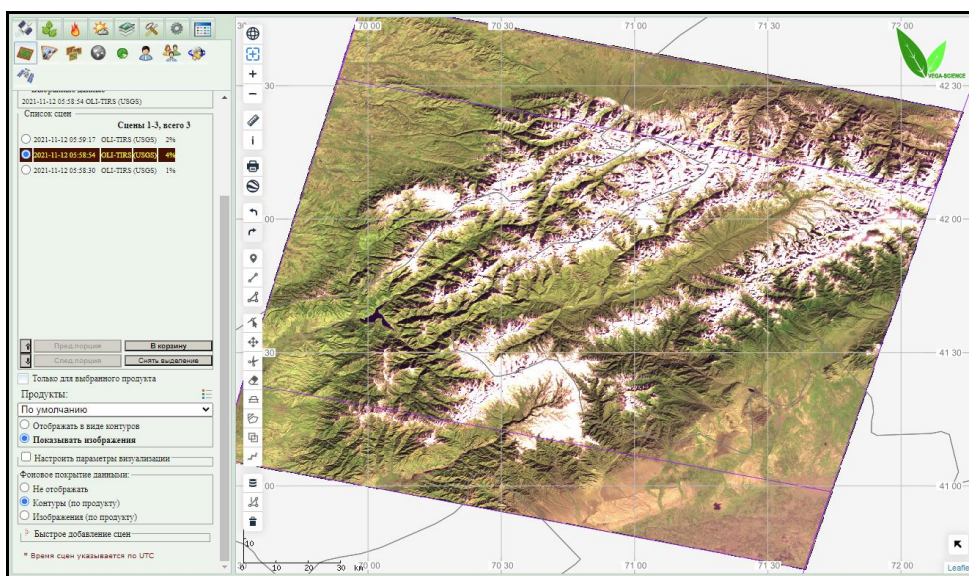


Рис. 3.2.1.1 - Внешний вид обновленного картографического веб-интерфейса

3.2.2. Новый универсальный модуль 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. Это, в частности, связано с тем, что в отличие от прежнего интерфейса при формировании одного изображения для области просмотра в новом интерфейсе необходимо получить целую группу отдельных тайлов. Прежде всего, были проанализированы основные недостатки существовавших ранее программных компонент, предназначенных для решения этой задачи. К числу основных из них можно отнести использование разных процедур и библиотек для решения схожих задач по формированию изображений и недостаточную их гибкость, что существенно усложняло поддержку и модификацию этих компонент. Также существенным фактором, снижающим их эффективность, было использование временных файлов, которого не удавалось избежать в прежней реализации. На основе анализа имеющихся недостатков было решено разработать новый универсальный модуль растеризации изображений.

Новый модуль растеризации был 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 то же время программный код, реализованный на языке Python существенно проще поддерживать и не нужно компилировать при установке на сервера с различными операционными системами.

Ниже приводится базовая последовательность операций, выполняемых при вызове модуля растеризации:

1. Определение параметров вызова
 - a. Определение параметров карты: диапазон координат, проекция, размер и др.
 - b. Определение правил построения продуктов: идентификаторы набора исходных данных, правила формирования на их основе требуемого информационного продукта, получаемые из справочных баз данных или от пользователя
2. Формирование изображения в памяти
 - a. Чтение данных из файлов
 - b. Обработка данных при помощи программного пакета GDAL
 - c. Получение мозаики фрагментов
 - d. Получение вычисляемых продуктов
3. Предоставление изображения
 - a. Формирование и отдача в выходной поток (или в файл, если нужно) изображения в формате PNG
 - b. Формирование и отдача изображения в формате GeoTIFF

В случае формирования изображения в формате PNG реализуются следующие основные варианты данных:

- трехканальный RGB продукт;
- одноканальный продукт с палитрой;
- одноканальный продукт без палитры;
- пустое прозрачное изображение в случае отсутствия данных в области интереса.

На основе обработки массивов NumPy производятся следующие основные операции: сшивка изображений, вычисления в соответствии с правилами построения продукта, нормировка данных, формирование маски прозрачности по каждому каналу изображения и др. Существенно, что для повышения эффективности работы модуля обработка каждого из входных изображений производится в параллельном режиме (до 8 одновременных потоков).

Создание универсального модуля растеризации изображений позволило существенно повысить эффективность работы нового тайлового веб-интерфейса. Благодаря его внедрению также существенно упростилась задача расширения имеющегося функционала по формированию спутниковых изображений.

3.3. Модернизация подсистемы для анализа временных и пространственных рядов данных на основе построения графиков

В 2021 году одним из основных направлений работ по реализации новых подходов к анализу долговременных рядов спутниковых наблюдений было развитие подсистемы анализа временных и пространственных рядов данных на основе построения графиков,

реализованной в виде специализированного веб-приложения. Примеры реализованных интерфейсов приведены на рисунке 3.3.1. В настоящее время, реализованные на базе этого приложения интерфейсы позволяют решать широкий круг различных задач, в том числе:

- анализ динамики различных показателей по объектам мониторинга, в частности спектральных характеристик, вегетационных индексов, метеорологических показателей и многих других характеристик, вычисленных на основе обработки спутниковых данных;
- исследование спектральных характеристик поверхности на заданном изображении;
- определение высотной изменчивости метеорологических величин;
- анализ пространственной изменчивости спектральных свойств, рассчитанных для объектов, а также метеорологических величин.

Необходимую гибкость при внедрении новых подходов обеспечило использование фреймворка Svelte, реализующего гибкий механизм модульного проектирования приложения. Также этот пакет позволяет существенно повысить скорость работы создаваемых на его основе интерфейсов по сравнению с интерфейсами, разработанными на основе стандартного подхода с включением в проект большого числа различных библиотек и модулей. Достигается это благодаря использованию специального этапа «компиляции», в рамках которого создаются оптимизированные файлы на языках программирования JavaScript и CSS, включающие в себя только тот функционал используемых библиотек и модулей, который необходим для работы конкретного интерфейса. К достоинствам фреймворка Svelte также относится хорошо продуманный механизм разработки и использования виджетов, т.е. компонент интерфейса, предназначенных для решения таких стандартных задач, как работа со списком и календарем, использование вкладок и т.п.

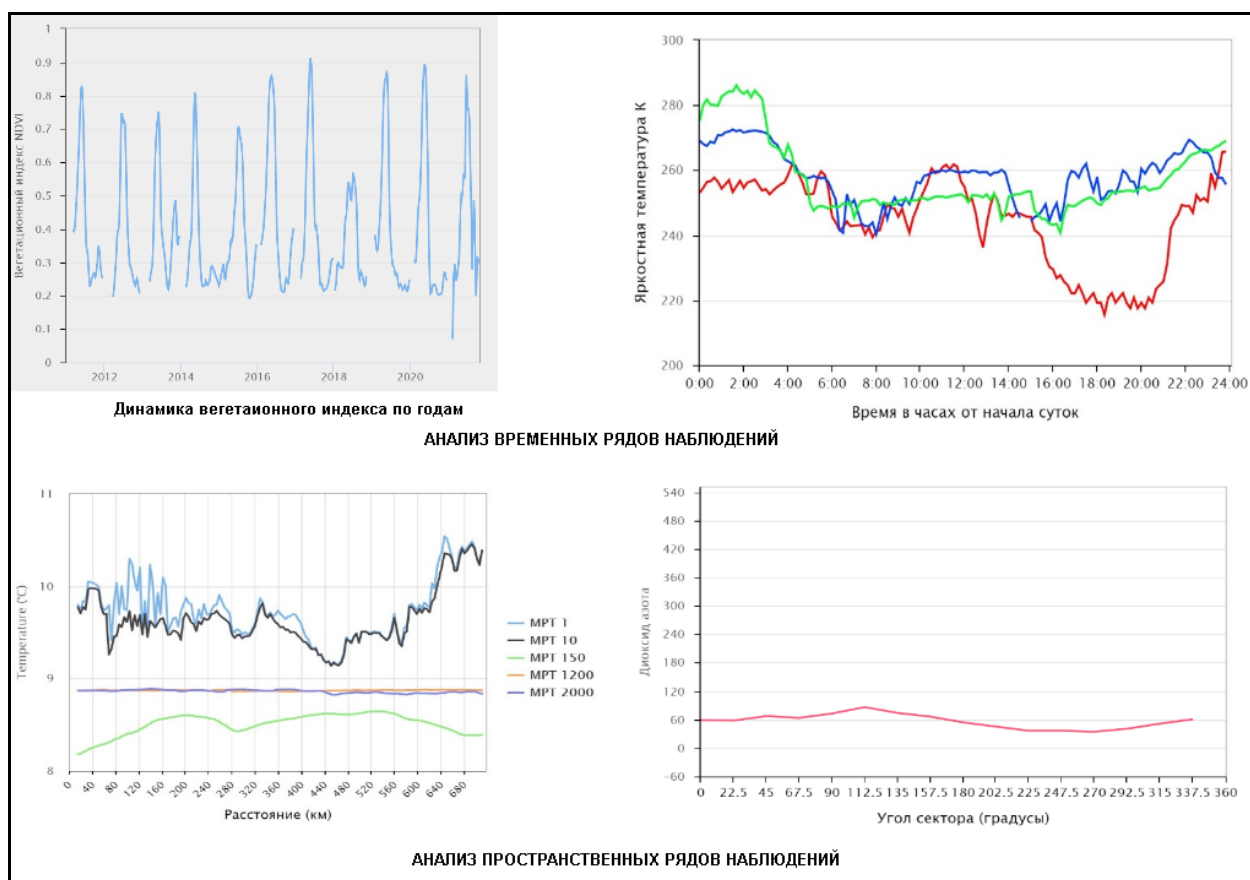


Рисунок 3.3.1 - Примеры интерфейсов для анализа временных и пространственных рядов наблюдений

С ростом числа используемых в подсистеме различных показателей по объектам возникла необходимость в унификации их описания и создания для этой цели специализированной базы данных. Структура основных таблиц базы данных object_indicators приведена на рисунке 3.3.2.1. Использование этой базы данных позволило вести единый реестр всех поддерживаемых в системе показателей, для каждого из которых может быть указан необходимый набор параметров, отвечающих за получение требуемых рядов данных и отображение их в интерфейсе. При этом благодаря заложенной в структуру базы данных гибкости один и тот же показатель может по-разному отображаться даже в рамках одного тематического интерфейса для разных пользователей и режимов работы. Для того чтобы сократить список показателей по объектам, которые необходимо рассчитывать в постоянном режиме на основе обработки спутниковых изображений, а также существенно расширить их состав в интерфейсе, значительная часть показателей теперь вычисляется непосредственно в клиентском интерфейсе на основе правил, задаваемых в БД object_indicators. В этом случае необходимые для их получения исходные ряды данных, например яркостные температуры по отдельным каналам, предварительно загружаются на клиентское приложение по стандартному протоколу. Для редактирования информации о показателях был реализован специализированный служебный интерфейс, пример которого приведен на рисунке 3.3.2.2. В нем также реализован функционал, позволяющий сформировать запросы, которые затем могут быть использованы в интерфейсе.

Использование единой базы данных показателей по объектам позволило централизованно управлять как их составом в рамках того или иного интерфейса, так и особенностями отображения соответствующих им временных рядов в виде графиков. В частности, это позволило поддерживать различный дизайн отображения одних и тех же показателей, используемых в различных информационных системах. Благодаря работам по структурированию используемого программного кода, проведенным в 2021 году, новые интерфейсы для анализа рядов в виде графиков во многих случаях могут быть реализованы с минимальными трудозатратами на основе компоновки уже готовых компонент. Также в этом году была поддержана возможность формирования ссылки на состояние интерфейса, включая все используемые настройки, которыми могут обмениваться между собой пользователи интерфейса.

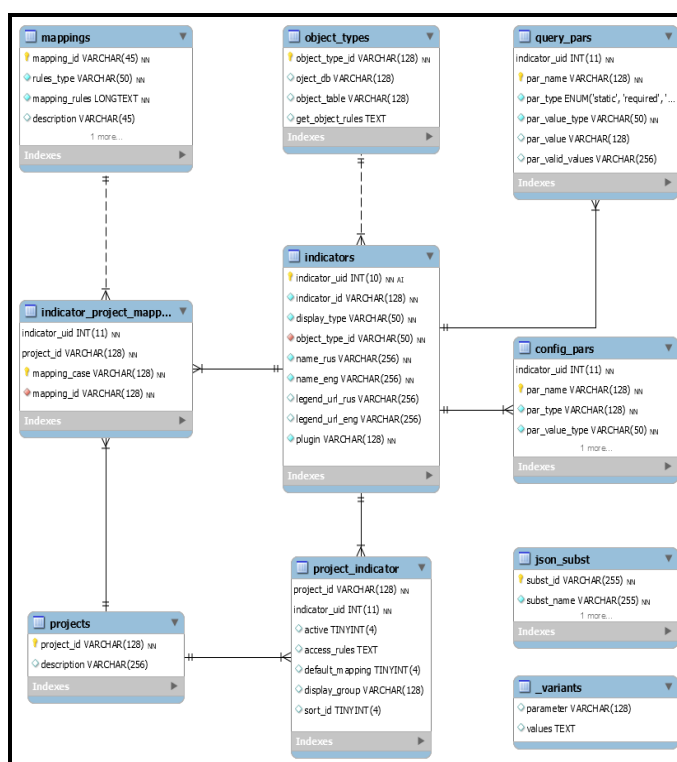


Рисунок 3.3.2 - Структура базы данных object_indicators

Каталог показателей по объектам: [Показатели](#) [Отображение в проектах](#) [Правила визуализации](#) [JSON подстановки](#) [Проекты](#) [Форма запроса метаданных](#)

Авторизованный пользователь: Прошкин Андрей База данных: [brsatsdb] object_indicators

Тип отображения: chart Тип объектов: любой

Выбор варианта показателя: [prec (113) | chart | point | Количество осадков, кг/м2] [Описать новый вариант показателя](#)

Общая информация (indicators)

Идентификатор	Показатель	Тип отображения	Тип объектов	Русское название	Английское название	Русская легенда	Английская легенда	Плагин	Изменить	Удалить	Сохранить как новый
indicator_uid	indicator_id	display_type	object_type_id	name_ru	name_eng	legend_ru	legend_eng	plugin			
113	prec	chart	point	Количество осадков, кг/м ²	Accumulated precipitation, kg/m ²	-	-	meteo_direct			

Параметры запроса показателя (query_pars)

Название параметров	Тип параметра	Тип значений	Значение параметра	Возможные значения параметра	Изменить	Удалить
par_name	par_type	par_value_type	par_value	par_valid_values		
year	required	string	п.п.	-		

Добавить параметр типа: static required optional

Параметры конфигурации показателя (config_pars)

Название параметров	Тип параметра	Тип значений	Значение параметра	Изменить	Удалить
par_name	par_type	par_value_type	par_value		
temporal_resolution	data	string	0.25		
axis	data	json	"chart_v_prec"		

Добавить параметр

Рисунок 3.3.3 - Интерфейс редактирования информации о показателях по объектам мониторинга

3.4. Развитие методов параллельной обработки спутниковых данных

Стремительное развитие национальных и международных систем ДЗЗ привело к доступности большого количества спутниковой информации для потребителя, в том числе с российских приборов и спутников. Это привело к существенному расширению области применения спутниковых данных, и, соответственно, к резкому возрастанию числа информационных продуктов, которые необходимо получать на их основе. В результате, для обработки как «сырых» данных, так и информационных продуктов требуется все больше вычислительных ресурсов.

Рассмотрим структуру вычислительного комплекса, существующего в ИКИ РАН для обработки и хранения данных ДЗЗ, которая проиллюстрирована на рисунке 3.4.1.

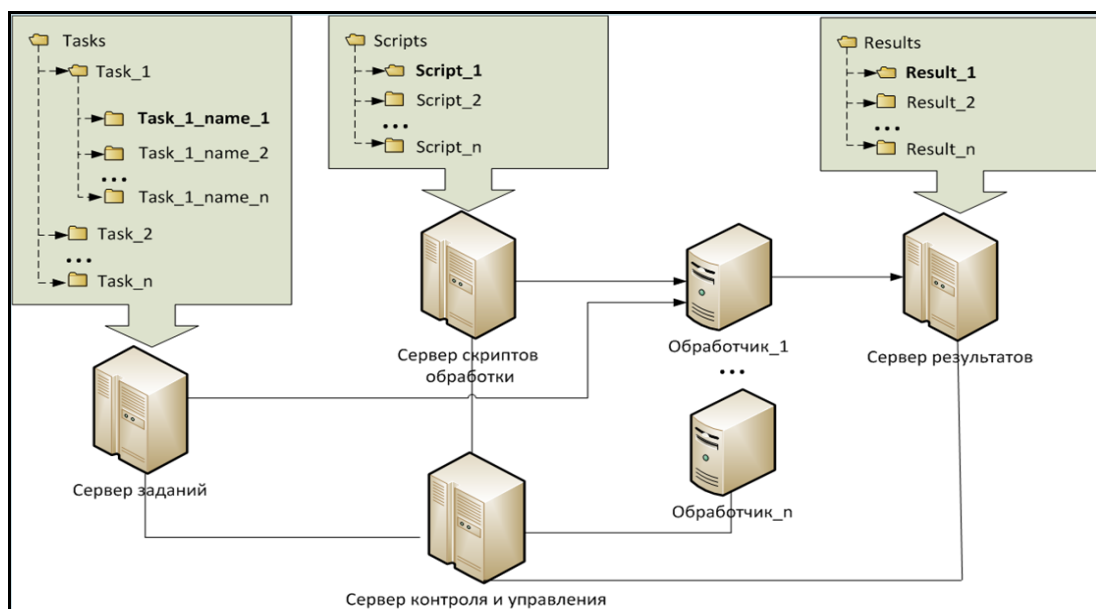


Рисунок 3.4.1 - Схема комплекса обработки данных.

Кластер серверов (станций) обработки, используемых в ИКИ РАН, состоит на данный момент из множества выделенных компьютеров (преимущественно 1-юнитовых). Каждая машина имеет в наличии от 8 до 16 процессорных ядер и от 8 до 32 Гигабайт оперативной памяти. В рамках разработанной в ИКИ РАН архитектуры построения системы обработки спутниковых данных используется понятие «задание», как

элементарной единицы выполнения обработки, которой соответствует совокупность требуемого набора данных, необходимого программного обеспечения и инструкций по его применению. До недавнего времени на каждой станции обработки в каждый момент времени выполнялось только одно задание на обработку. Однако такой вариант оказывается недостаточно эффективным, особенно если учесть, что некоторые средства обработки 32-разрядные и/или 1-процессорные. Такие приложения имеют ограничение в 3 Гб занимаемой оперативной памяти и, таким образом, могут быть запущены одновременно в несколько потоков в рамках одной сессии обработки, без особого ущерба для производительности каждого экземпляра процесса. Это позволяет создать высокую нагрузку и, как следствие, увеличить скорость обработки исходных данных.

Рассмотрим время, необходимое на обработку небольших легких заданий при параллельном запуске от 1 до 10 потоков на одном обработчике. Всего 20 заданий 500 Мб общим весом. Характеристики обработчика – 16 Гб оперативной памяти, 8 ядер, скорость диска - 100 мб/с. Время обработки полного комплекта данных для различного числа потоков приведено на рисунке 3.4.2.

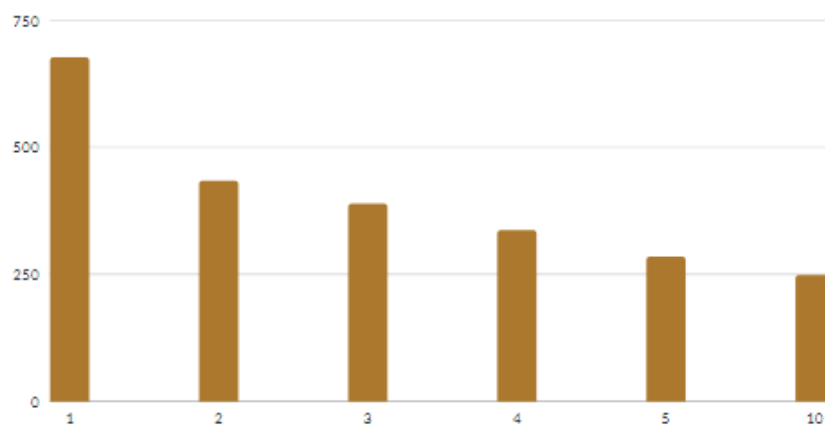


Рисунок 3.4.2 – Время обработки полного комплекта данных при параллельном выполнении заданий.

Видно, что при использовании такой схемы, можно получить выигрыш почти в 3 раза по скорости уже на 5 потоках. Дальнейшее увеличение числа одновременно обрабатываемых заданий уже не приносит существенной пользы. Нужно отметить, что расчет количества потоков может быть произведен как статически, так и динамически, в зависимости от количества ядер и оперативной памяти конкретного обработчика.

Другой вариант распараллеливания можно применить, если требуется обработать большое количество файлов в рамках одного задания. При этом суммарный объем обрабатываемых данных не позволяет одновременно запускать несколько заданий, из-за ограниченного объема жесткого диска обработчика. Тогда файлы можно разделить на несколько частей, к примеру, положив их в разные директории. Каждая часть может быть обработана независимо (и одновременно). В примере представлена предобработка данных для создания композита данных спутника Landsat 8, а именно построение масок облачности. На входе 400 сцен, 200 Гб общим весом, 12 ядерный процессор с 24 Гб оперативной памяти.

Нужно отметить, что в данной обработке некоторые части также распараллелены внутри кода через CreateThread, что, по сути, приводит к некоторой избыточности, так что применять данный подход следует с осторожностью, как видно из графика, можно получить даже замедление суммарной скорости обработки. На данном графике (рисунок 3.4.3) по вертикали отмечены часы, затраченные на обработку, по горизонтали – число потоков (порций данных).

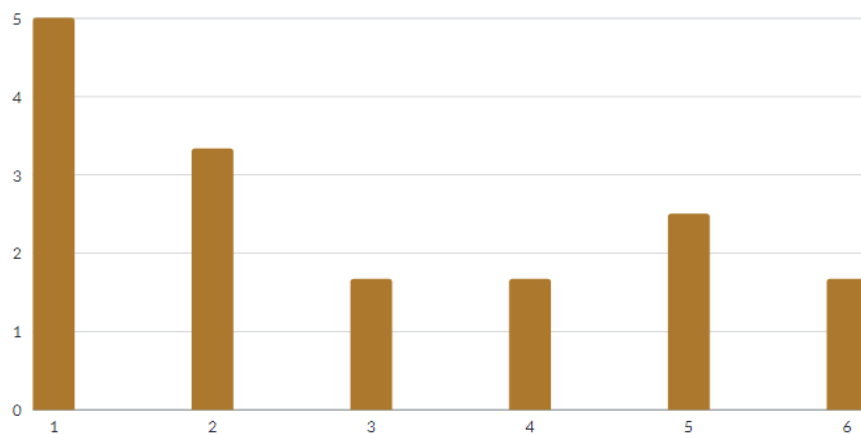


Рисунок 3.4.3 – Время для случая с параллельной обработкой комплектов

Также встречаются ситуации, когда файлов много, но разделение их на части невозможно. В качестве примера можно привести задачу интерполяции временных серий данных, полученных с использованием прибора MODIS. Интерполяция каждой точки временного ряда может быть выполнена независимо друг от друга. В этом случае внутри операции можно с легкостью использовать thread-ы. На следующем графике, приведенном на рисунке 3.4.4, показаны времена при использовании такого распараллеливания кода. По вертикали отложено время в часах, а по горизонтали количество потоков (threads). На входе 194 файла MODIS со спутников Terra и Aqua в синусоидальной проекции, на выходе 97 файлов интерполированных значений индексов NDVI. Обработчик имеет 16 ядер и 24 Гб. оперативной памяти.

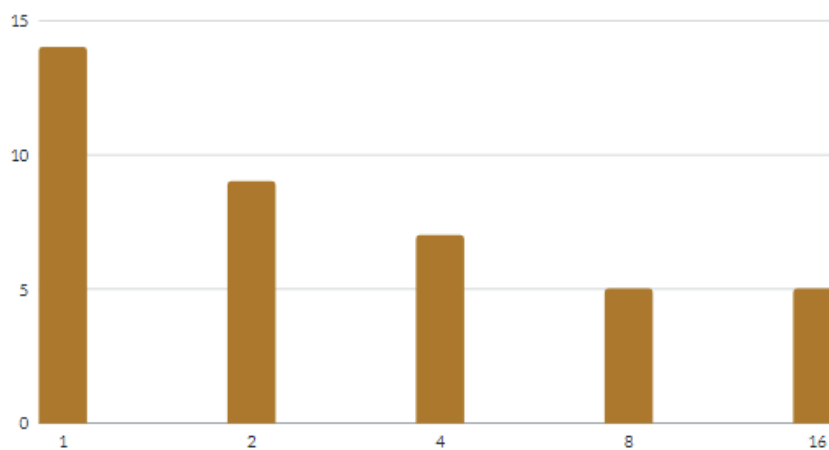


Рисунок 3.4.4 - Вариант с проведением параллельной обработка при помощи метода CreateThread

Отсутствие выигрыша при увеличении количества потоков до 16 объясняется ограничением дискового массива, он не может отдавать данные с требуемой скоростью. Как показывает практика это основное препятствие (наряду со скоростью получения данных с сервера по сети) на пути к ускорению вычислений.

На основании проведенного исследования была показана эффективность применения рассматриваемых подходов к распараллеливанию обработки в рамках одного вычислительного узла. Все представленные варианты в настоящее время внедрены и используются для повышения эффективности выполнения различных задач по обработке спутниковых данных. При этом они дают максимальный выигрыш практически в три раза.

Избыточное распараллеливание также можно использовать, но требуется предварительное тестирование, так как оно может привести и к падению эффективности выполнения обработки, как показано на рисунке 3.4.3.

3.5. Развитие системы контроля функционирования сложного распределенного программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг»

3.5.1. Общая информация

Современные системы доступа к данным ДЗЗ и результатам их обработки представляют из себя сложные территориально распределенные программно-аппаратные комплексы, в работе которых участвуют многие десятки компьютеров, функционирующих в автоматизированном режиме. К таким системам, безусловно, относится реализованный в ИКИ РАН центр коллективного пользования системами архивации, обработки и анализа данных спутниковых наблюдений ЦКП «ИКИ-Мониторинг». В настоящее время в его работе задействовано более 50 серверов архивации и специализированных систем хранения данных (NAS), свыше 60 серверов, предназначенных для обработки спутниковых данных, а также более десяти специализированных серверов, предназначенных для формирования спутниковых изображений для картографических WEB интерфейсов. В настоящее время пользователям ЦКП «ИКИ-Мониторинг» обеспечивается доступ к архивам данных ДЗЗ, суммарный объем которых превышает 5 петабайт, при этом ежедневно в архивы центра поступает более 3 Тб новых данных.

Для обеспечения бесперебойного функционирования программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг» был разработан и развивается в настоящее время целый ряд программных решений, предназначенный для решения следующих основных задач:

1. Контроль выполнения процессов

Эта задача контроля является наиболее очевидной. Для ее качественного решения необходимо вести протоколы выполнения всех процессов по работе с данными в системе, а также их коды возврата, которые при этом должны быть стандартизированы.

2. Автоматическое детектирование сбоев в работе программно-аппаратного комплекса

Как показывает многолетний опыт разработки и эксплуатации сложных распределенных систем, задача автоматического обнаружения различных типов критических неполадок в работе системы является ключевой для достижения своевременного оперативного их устранения и обеспечения бесперебойной работы программно-аппаратного комплекса

3. Документирование сбоев в работе комплекса

Для того чтобы эффективно устранять неполадки в работе сложного программно-аппаратного комплекса необходимо вести систематизированную информацию о возникающих ошибках, оперативно оповещать о них ответственных специалистов и отслеживать процесс их устранения.

4. Ежедневные отчеты о функционировании отдельных подсистем комплекса

Ошибки при выполнении отдельных процессов не обязательно являются критическими и требующими немедленного исправления, при этом даже отсутствие ошибок в работе процессов еще не означает, что в работе системы отсутствуют неполадки, в частности, возникшие по внешним для нее причинам. Для качественной оценки работы целых подсистем комплекса используются ежедневные отчеты, рассылаемые ответственным специалистам по электронной почте.

5. Инструменты для анализа работы различных подсистем комплекса

Инструменты анализа позволяют разобраться в причинах возникших неполадок, выявить скрытые ошибки, найти «узкие» места в работе подсистемы, получить необходимую статистику о наличии тех или иных данных в архивах и т.д.

6. Документирование используемых программных и аппаратных средств

Хотя сама по себе информация об аппаратной и программной конфигурациях комплекса не позволяет детектировать или устранять ошибки в его работе, без нее практически невозможно качественно решить все вышеописанные задачи.

3.5.2. Контроль выполнения процессов

Контроль выполнения всех процессов, а прежде всего, автоматически запускаемых, является базовым элементом всей системы контроля. На UNIX серверах он реализуется на основе использования программного пакета PMS (Process Monitoring System). Для этого все процессы запускаются при помощи специальной программы, которая заносит протоколы их выполнения, а также коды возврата в специализированную базу данных. При этом используется стандартный набор кодов возврата процессов, основные из которых сигнализируют о следующих результатах выполнения процесса: успешное выполнение, ошибка при выполнении, холостой запуск, блокировка запуска другим процессом. Специализированный веб-интерфейс обеспечивает доступ к статистике выполнения заданной группы процессов за указанный интервал времени с использованием целого ряда фильтров, позволяющих показывать только: ошибочные запуски, активные запуски, результативные запуски и другие. Также для каждого процесса могут быть заданы различные критерии обнаружения неполадок в его работе, например, может быть указано время, в течение которого должен быть хотя бы один результативный успешный запуск процесса. Специальный программный интерфейс позволяет предоставлять контрольную информацию внешним программным средствам. Примеры базовых инструментов, реализованных в специализированном веб-интерфейсе, приведены на рисунках 3.5.2.1, 3.5.2.2.

Process	5	4	3	2	1	Last Success	Last Error
600 [unlabeled] Add launcher Developer: Tikhonov, Pavel	13:35	13:39	13:43	13:46	14:10	13:46 9 Nov	15:13 ago
601 [unlabeled] Import planets_united metadata	13:01	13:10	13:27	13:33	13:35	13:35 9 Nov	104 days ago
602 [unlabeled] Import zscipod_united metadata	12:47	13:01	13:10	13:25	14:10	14:10 9 Nov	26 Sep 2016
603 [unlabeled] Import divcpod_united metadata	13:32	13:35	13:38	13:43	14:10	14:10 9 Nov	15:13 ago
604 [unlabeled] Import ktweb metadata	13:35	13:39	13:43	13:46	14:10	14:10 9 Nov	15:12 ago
622 [unlabeled] Add source_panc0 landstatpsh_landstat	13:25	13:35	13:40	13:43	14:10	14:10 9 Nov	16:22 30 Oct
623 [unlabeled] Add source_panc1 landstatpsh_landstat	13:25	13:35	13:37	13:41	13:47	13:47 9 Nov	17:07 30 Oct
624 [unlabeled] Add source_panc2 landstatpsh_landstat	13:12	13:28	13:33	13:47	14:10	14:10 9 Nov	17:07 30 Oct
627 [unlabeled] Add source_panc3 landstatpsh_landstat	13:12	13:28	13:34	13:50	14:16	14:16 9 Nov	17:07 30 Oct
628 [unlabeled] Add source_panc4 landstatpsh_landstat	13:25	13:27	13:41	13:43	13:51	13:51 9 Nov	17:07 30 Oct
629 [unlabeled] Add source_panc5 landstatpsh_landstat	13:26	13:34	13:38	13:42	13:51	13:51 9 Nov	17:08 30 Oct

Рисунок 3.5.2.1 - Просмотр информации о последних запусках процессов

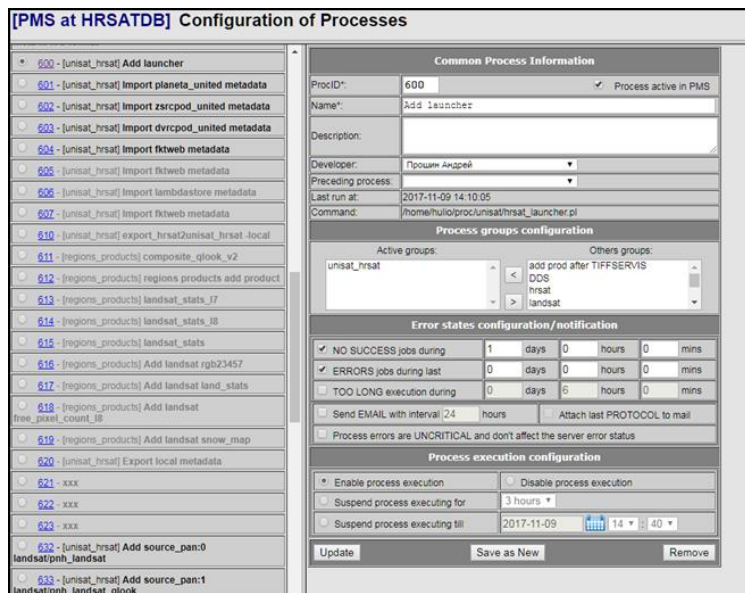


Рисунок 3.5.2.2 - Конфигурация процессов в системе PMS

Контроль за выполнением процедур обработки данных на Windows серверах обработки данных реализован на основе использования системы управления многопоточковой обработкой спутниковых данных. В рамках этой системы реализуется централизованное протоколирование выполнения всех процедур обработки данных и их кодов возврата. Доступ к контрольной информации и ее анализ реализованы на базе использования интерактивных отчетных форм, построенных по технологии Business Intelligence (BI). На рисунке 3.5.2.3. приведен пример одной из многочисленных интерактивных отчетных форм, реализованных для контроля за выполнением процессов по обработке спутниковых данных.

Группы заданий	Задания	ldy	Удачных	Процесс	Ошибки	Crash	Пустых
регулярные-приоритетные	AM_STN_1DC daily composite oper (session)		46		3		49
	benchmark composite Landsat 1		1	1	1	1	
	benchmark composite Landsat 2		1	2	1	1	1
	benchmark composite Landsat 3	1		1	1	1	
	benchmark composite Landsat 4		1	1	1	2	
	benchmark composite Landsat 5	1		1	1	1	
	benchmark composite Landsat 6		1	1	2		
	benchmark_limires_interpol-1		1		1		
	benchmark_limires_interpol-2		1		1		
	benchmark_limires_interpol-4		1		1		
	benchmark_limires_interpol-8		1		1		
	Burns 250 TV product (session)	3					

Рисунок 3.5.2.3 - Просмотр статистики выполнения группы процессов по обработке данных

3.5.3. Автоматическое детектирование сбоев в работе программно-аппаратного комплекса

В настоящее время на серверах программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг» реализовано выполнение многих сотен процессов, выполняющих задачи по сбору, обработке и архивации данных, а также специальных процессов, отвечающих за контроль работы различных системных и аппаратных компонент. Естественно, что не все возникающие в их работе ошибки являются одинаково важными и требуют немедленного устранения. В то же время, даже отсутствие детектированных ошибок в работе тех или иных процессов в рамках такой сложно многокомпонентной системы еще не означает отсутствие существенных неполадок в ее работе. Поэтому возникает необходимость в автоматическом детектировании наиболее критичных для функционирования комплекса

неполадок в его работе, о которых следует оперативно оповестить специалистов, занимающихся технической поддержкой работы комплекса.

Ниже приводятся основные типы проверок, предназначенных для их автоматического выявления:

- Контроль функционирования аппаратных компонент:
 - Проверка доступности всех серверов и хранилищ данных
 - Проверка состояния дисковых массивов (RAID)
- Контроль системных программных компонент:
 - Диагностика и прогнозирования переполнения дисковых разделов
 - Проверка работы основных служб на серверах
 - Проверка доступности по NFS всех томов архивов данных
 - Проверка синхронизации времени на серверах
- Контроль оперативного поступления в архивы всех типов информационных продуктов (много десятков)
 - Для каждого информационного продукта проверяется давность экземпляров данных и если она превышает заданную в конфигурации, то диагностируется сбой
- Контроль выполнения заданий обработки данных
 - Контроль процента успешных выполнений
 - на каждом из серверов обработки
 - Контроль допустимого времени простоя серверов обработки
 - Обнаружение «зависших» процессов
- Контроль целостности тематического программного обеспечения
 - На каждом из серверов для каждого программного пакета проверяется комплектность файлов и их контрольные суммы

3.5.4. Документирование сбоев в работе комплекса. Система ведения сбоев

Для документирования всех неполадок в работе системе, оперативного оповещения о них ответственных специалистов и отслеживания процесса их устранения была реализована специализированная система ведения сбоев. При этом реализуется два основных источника поступления в нее сведений об ошибках: получение информации от программных компонент автоматического контроля различных типов сбоев и «ручное» занесение этой информации пользователями информационных систем и специалистами, осуществляющими их разработку и поддержку. Ниже перечислены основные функциональные блоки разработанной системы

- Интерфейс для поиска и просмотра информации об обнаруженных неполадках в работе различных информационных систем, реализованных на базе возможностей ЦКП «ИКИ-Мониторинг». Отдельный режим работы этого интерфейса реализован для дежурного на текущий день специалиста, основной задачей которого является контроль устранения всех актуальных сбоев в работе комплекса.
- Блок редактируемой справочной информации о типах работ и соответствующих им ответственных специалистах
- Интерфейс занесения информации об ошибках с возможностями классификации сбоев по проектам, организациям и направлениями работ, указания ответственного, степени критичности сбоя, ведения истории изменений и комментариев и др.
- API для заведения сбоев системами автоматического детектирования ошибок
- Подсистема оперативного оповещения ответственных специалистов по электронной почте
- Подсистема формирования ежедневных отчетов о неисправленных на текущий момент критических неполадках в работе комплекса с указанием ответственных, длительности сбоя и другой информации

[СДКП] Система ведения сбоев

Общая информация

 анду
Прошин

Дежурный: Бурцев ☐ Режим дежурного

Сегодня: чт, 11 ноября 2021, 18:41:31

Управление выводом информации:

Показывать за период:
 Сортировать по:
 Всегда показывать:
 Фильтр по описанию

☒ крит. ☐ новые ☐ оповещ.
☒ ☒

Легенда

- Новый сбй
- Сбой направлен
- Сбой принят
- Сбой исправлен
- Критический сбй

Новый сбй
 Статистика

Найдено 14 записей

ID	Проекты	Центр	Обнаружил	Ответственный	Направляемая работ	Статус	Описание и комментарии	Время обнаружения сбоя
---	---	---	---	Прошин	---	Не исправлен		
46211	56 отдел	ИКИ РАН	RMS	Дежурный (Бурцев)	Система контроля за серверами в стационных образах RMS	Новый	Сервер MRSAUTOB: Процесс: 5410 - NFS mount ERRORS during 11 minutes Last non-error job at: 11 Nov 17 25 MSK 03 Alert timeout parameter: 10 minutesPLUS	11 ноя 17:26 11 ноя 17:26
46210	56 отдел	ИКИ РАН	RMS	Дежурный (Бурцев)	Система контроля за серверами в стационных образах RMS	Новый	Сервер SBD: Процесс: 10018 - www_forms+ru_reports ERRORS during 1 minute 50 seconds Last non-error job at: 11 Nov 17 03 MSK 03 PLUS	11 ноя 17:05 11 ноя 17:05
46178	56 отдел	ИКИ РАН	RMS	Прошин	Система контроля за серверами в стационных образах RMS	Направлен	Сервер MRSAUTOB: Процесс: 1008 - Check NFS mounts ERRORS during 19 days 9 hours Last non-error job at: 22 Oct 18 50 MSK 03 Alert timeout parameter: 1 hour Показать скриншот	10 ноя 18:20 10 ноя 18:20
46106	56 отдел	ИКИ РАН	RMS	Прошин	Система контроля за серверами в стационных образах RMS	Направлен	Сервер DEV: Процесс: 6410 - NFS mount NG SUCCESS job during 1 hour 5 minutes Last success job at: 10 Nov 11 10 MSK 03 Alert timeout parameter: 1 hourPLUS Показать скриншот	10 ноя 12:15 10 ноя 12:15
45984	Вега 56 отдел	ИКИ РАН	SMISPKG	Прошин	SMISPKG	Направлен	Сервер zeltastore check_smispg_integrity AUMPS-Probin dnf_update - 11119 State Invalid, Master dev info aviales ru, AUMPS-Probin Changed files. Показать скриншот	5 ноя 17:06 11 ноя 15:36
45926	56 отдел	ИКИ РАН	RMS	Сенько	Система контроля за серверами в стационных образах RMS	Направлен	Сервер SBD: Процесс: 418 - id_upd_sks_snaps (thread 1) ERRORS during 34 minutes Last non-error job at: 4 Nov 03 16 MSK 03 PLUS Показать скриншот	4 ноя 03:20 4 ноя 03:21
45901	ИСМД Вуланы	ИКИ РАН	SMISPKG	Прошин	SMISPKG	Направлен	Сервер firemaps check_smispg_integrity i_sys_object_indicators - 1 State Invalid, Master dev info aviales ru, AUMPS-Probin	3 ноя 14:06

Рисунок 3.5.4.1 - Просмотр информации о сбоях в работе комплекса

[СДП] Редактирование информации о сбое


Логин: andry (Прошин)
Должность: Бурцев

Общая информация

Идентификатор сбоя	31499
Кто обнаружил	Смирнов Артем Валерьевич
Отвечает дежурный	Прошин
Текущий статус	Принят
Время обнаружения	2020-04-21 15:00:51

Проблемы

Удалено: Проспект
 Планета - Геоизмиса-Центры
 Сервис: BEFA
 BEFA-Лос
 Тайповый интерфейс
 Буланы
 BEFA ЕФИС ЗСН

Центры

ЦЭПЛ РАН
 ЦСМС
 див РГЦПОД
 ВЦ ДВО РАН
 ГАЗКОМ Щетково
 ИКИЗ
 ииш РАН

Проинформировать

Елизия
 Аргаманова
 Байбаков
 Барталиев
 Барталиев(мл)
 Белоконь
 Бочарова

Определение типа работ и ответственного

Направленные работы	Система доступа
тип работ	Формы отчетности по пожарам и др.
Ответственный	Фишер
Критический сбой:	☒
Изменить статус:	Принят

Описание и изменения:

Сбой от:
 Организация: МЦ Сибирского и Дальневосточного федеральных округов ФБУ "Авиалесоохрана"
 e-mail: smirnyagorodovaa@mail.ru
 По адресу: Ссылка в описании
 Проблема описана в вложенном файле: [ссылка \[Файл\]](#)

Новые изменения:

Сбой был занесен пользователем системы, который не может просматривать журнал сбоя.
 Чтобы послать ему письмо с новыми изменениями поставить галочку ☐

Вынести изменения

Рисунок 3.5.4.2 - Редактирование информации о сбое

3.5.5. Ежедневные отчеты о функционировании отдельных подсистем комплекса

Ежедневные отчеты о функционировании, как правило, формируются для таких подсистем, от функционирования которых зависит реализация основного функционала программно-аппаратного комплекса, а также для подсистем, находящихся в процессе разработки и отладки. К первому случаю относится, в частности, формирование ежедневных отчетов о работе подсистемы сбора, предварительной обработки и архивации данных, поступающих в архивы ЦКП «ИКИ-Мониторинг» из различных центров распространения спутниковых данных. Это связано с тем, что от работы этой подсистемы зависит наличие в архивах требуемых данных, в том числе оперативных. Сложность задачи сбора данных связана, в частности, с тем, что во многих случаях задержки с поступлением данных бывают вызваны различными внешними причинами, например, профилактическими работами в центре распространения данных или изменениями в протоколах их взаимодействия с пользователями и даже в форматах предоставления

самых данных. В рамках ежедневного отчета для каждого типа получаемых спутниковых данных указывается количество скачанных за последние сутки данных, количество обработанных и усвоенных в архивы данных, текущая очередь их обработки, а также оценка того, насколько эти значения отличаются от средних значений за заданный интервал времени. Также под особым контролем находится система получения на основе собранных данных различных тематических информационных продуктов. Ко второй группе ежедневных отчетов относится, к примеру, контроль функционирования новых подсистем обработки спутниковых данных и контроль формирования различных наборов данных, наиболее актуальных для выполнения текущих проектов.

3.5.6. Инструменты анализа работы различных подсистем комплекса

Для анализа функционирования отдельных подсистем комплекса было разработано большое количество специализированных инструментов, позволяющих в интерактивном режиме провести исследование различные показатели работы таких подсистем. Некоторые из реализованных на текущий момент инструментов перечислены ниже:

- Комплексная информации о состоянии и степени наполнения дисковых хранилищ данных
- Информация о задержках при обработке принимаемых в информационных центрах сеансов спутниковых данных
- Использование ПО Zabbix для мониторинга аппаратной нагрузки серверов
- Использование технологий BI (Business Intelligent) для анализа информации о выполнении заданий по обработке спутниковых данных (время, продолжительность, статус, ошибки, сервера обработки и др.)
- Формы запроса на наличие разных типов данных в архивах по целому ряду различных критериев, таких как диапазон дат, географическая область, станции приема и т.п.
- Статистика по данным, поступившим в архивы за указанный диапазон времени
- Статистика по аппаратным сбоям серверов
- Инструменты для визуального анализа пространственного наличия различных типов спутниковой информации

На рисунке 3.5.6.1 приведен пример интерфейса, построенного по технологии BI и предназначенного для анализа информации о выполнении заданий по обработке спутниковых данных.

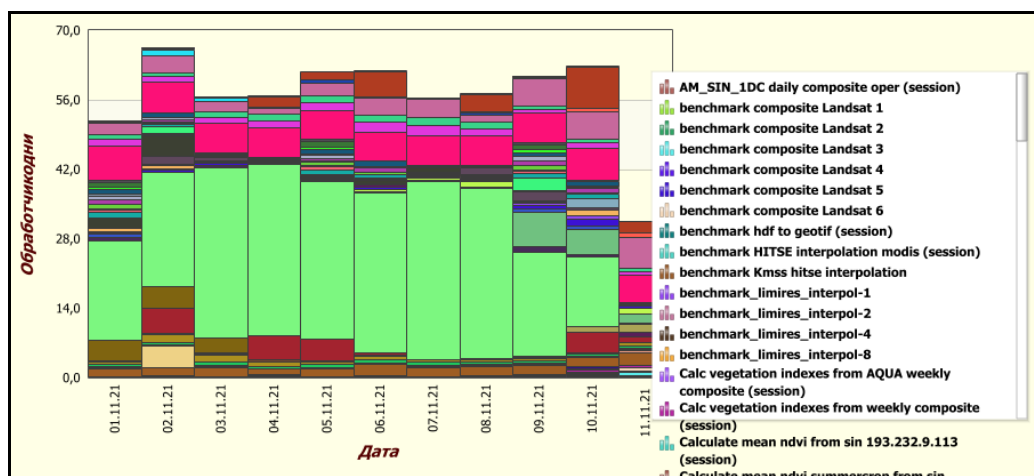


Рисунок 3.5.6.1 - Пример интерактивной отчетной формы, используемой для анализа информации о выполнении заданий по обработке спутниковых данных

3.5.7. Документирование используемых программных и аппаратных средств

Описанные выше инструменты для обеспечения бесперебойной работы программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг» концептуально являются часть разработанной в ИКИ системы документирования и контроля проектов (СДКП). В состав этой системы входят функциональные блоки, отвечающие за описание конфигурации программных и аппаратных компонент, таких как: используемые компьютеры, базы данных, реализованные пользовательские и служебные интерфейсы и т.п. Пример заглавной страницы интерфейса системы СДКП приведен на рисунке 3.4.7.1

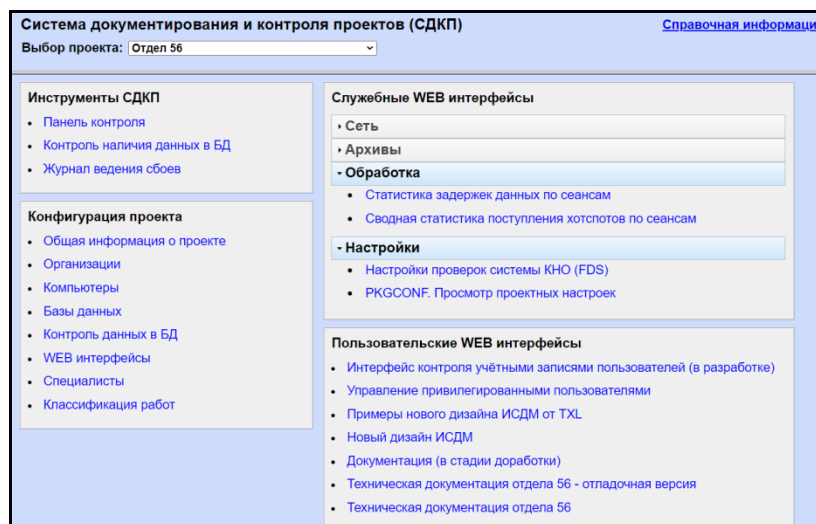


Рисунок 3.5.7.1 - Заглавная страница интерфейса системы СДКП

3.5.8. Заключение

В 2021 году продолжались работы по совершенствованию имеющихся инструментов для контроля функционирования программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг» и созданию новых программных решений, предназначенных для решения этой задачи. В частности, был существенно расширен набор инструментов для анализа работы различных подсистем комплекса, добавлены новые интерактивные формы для контроля за работой системы обработки данных, а также был переработан блок контроля за системой получения спутниковых данных из центров их распространения. Выработанные в рамках построения представленной системы контроля решения могут быть использованы и для обеспечения бесперебойного функционирования схожих по назначению программно-аппаратных комплексов.

Заключение

Представленные в настоящей главе результаты показывают, что по направлению «Разработка новых методов и подходов организации хранения и параллельной обработки сверхбольших объемов данных для решения задач анализа долговременных наблюдений в интересах ведения мониторинга и изучения различных процессов и явлений» проведенные работы соответствуют установленной в плане работ ИКИ РАН на 2021 год тематике. Таким образом, работы выполнены полностью в соответствии с планом.

При этом получены следующие основные результаты:

- Проведены модернизация и внедрение технологии динамического блочного доступа к данным сверхбольших архивов спутниковых данных для проведения их эффективной параллельной обработки
- Доработан и внедрен в систему «Вега-Science» и ряд других информационных систем новый тайловый веб-интерфейс на базе использования механизма динамического блочного доступа к данным в архивах

- Модернизирована подсистема анализа временных и пространственных рядов данных на основе графиков
- Были проведены работы по развитию системы контроля функционирования сверхбольшого программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг»

Таким образом, можно считать, что работы по данному направлению темы в 2021 году полностью выполнены.

ЗАКЛЮЧЕНИЕ

Таким образом, представленный отчет по теме «Большие данные в космических исследованиях: астрофизика, солнечная система, геосфера» (№ 0024-2019-0014) показывает, что в соответствии с планом ИКИ РАН на 2021 год они велись по следующим основным направлениям:

- Развитие комплекса взаимосвязанных и дополняющих друг друга физико-математических моделей для диагностики и прогнозирования состояния высокоширотной ионосферы в северном и южном полушариях, которые электродинамически связаны между собой через область средних широт. Разработки общей модели взаимодействия в системе солнечный ветер-магнитосфера-ионосфера-нейтральная атмосфера Земли.
- Применение методов машинного обучения и глубокого обучения для обработки и анализа больших массивов астрономических данных в задачах наблюдательной астрофизики.
- Разработка новых методов и подходов организации хранения и параллельной обработки сверхбольших объемов данных для решения задач анализа долговременных наблюдений в интересах ведения мониторинга и изучения различных процессов и явлений.

В рамках данной тематики опубликовано в 2021 году 25 работ, из них 7 в рецензируемых изданиях, входящих в Scopus и WoS (плановый показатель 3).

Таким образом, можно считать, что поставленные в плане задачи на 2021 год в рамках данной темы полностью решены.

ПУБЛИКАЦИИ

Опубликовано – 25 работ, из них:

- в зарубежных изданиях – 5
- книг – 0
- статьи в отечественных научных рецензируемых журналах – 2
- статьи в сборниках материалов конференций – 0
- материалы конференции - 8
- доклады, тезисы – 10
- публикаций по грантам РФФИ – 0
- в соавторстве с зарубежными авторами – 1
- статьи в WOS и Scopus – 7
- статьи в WOS и Scopus (Q1, Q2) – 2
- статьи в WOS и Scopus (Q3, Q4) - 3

Публикации в зарубежных изданиях

1. *Lukianova, R., Kozlovsky, A., Lester, M.* Upper stratosphere mesosphere-lower thermosphere perturbations during the formation of the Arctic polar night jet in 2019–2020. *Geophysical Research Letters*, 48, e2021GL094926. 2021. <https://doi.org/10.1029/2021GL094926>. (**WoS, Scopus Q1**)
2. *Bikmaev I. F., Meshcheryakov A. V., Irtuganov E. N., Nikolaeva E. A., Sakhibullin N. A., Gumerov, R. I., Sklyanov A. S., Glushkov M. V., Khamitov I. M., Borisov V. D., Burenin R. A., Zaznobin I. A., Krivonos R. A., Lyapin A. R., Medvedev P. S., Sazonov S. Yu., Sunyaev R. A., Khorunzhev G. A., and Gilfanov M. R.* Spectroscopic Redshift Determination for a Sample of Distant Quasars Detected by the SRG Observatory Based on RTT-150 Observations. II // *Astronomy Letters*, 2021, Vol. 47, No. 5, pp. 277–290, DOI: 10.1134/S1063773721050029 (**WoS, Scopus Q3**)
3. *Proshin A.A., Loupian E.A., Bartalev S.A.* Metodology for evaluation the effectiveness of the system of dynamic block access to data of ultra-large distributed remote sensing archives // *Ceur Workshop Proceedings. ITHPC 2021 - Short Paper Proceedings of the 6th International Conference on Information Technologies and High-Performance Computing*, 2021. Vol. 2930. P. 62-69. (**Scopus**)
4. *Proshin A.A., Loupian E.A.* Comprehensive analysis of the efficiency of the system of dynamic block access to data of ultra-large distributed remote sensing archives // *CEUR Workshop Proceedings. Proceedings of the All-Russian Conference With International Participation "Spatial Data Processing for Monitoring of Natural and Anthropogenic Processes" (SDM-2021)*, 2021. Vol. 3006. P. 222-230. DOI: 10.25743/SDM.2021.28.16.027. (**Scopus**)
5. *Loupian E., Burtsev M., Proshin A., Kashnitskii A., Balashov I., Bartalev S., Konstantinova A., Kobets D., Radchenko M., Tolpin V., Uvarov I.* Usage Experience and Capabilities of the VEGA-Science System // *Remote Sensing*. 2021. 14. – в печати (. (**Web of Sciences - Q1, Scopus**)

Публикации в отечественных научных рецензируемых изданиях

6. *Мингалева И. В., Суворова З.В., Шубин В. Н., Мёрзлый А.М., Тихонов В.В., Талалаев А.Б., Мингалева В.С.* Отличия прогнозов условий КВ-радиосвязи между передатчиком на средних широтах и приемником в арктическом регионе при использовании различных эмпирических моделей ионосферы, «Геомагнетизм и аэрономия», Т. 61, № 4, с.506-519, 2021.DOI: 10.31857/S001679402104009X. (**WoS, Scopus Q3**)

7. *Лузян Е.А., Прошин А.А., Бурцев М.А., Кашицкий А.В., Балашов И.В., Барталев С.А., Егоров В.А., Жарко В.О., Константинова А.М., Кобец Д.А., Мазуров А.А., Марченков В.В., Матвеев А.М., Миклашевич Т.С., Плотников Д.Е., Радченко М.В., Стыценко Ф.В., Сычуглов И.Г., Толтин В.А., Уваров И.А., Хвостиков С.А., Ховратович Т.С.* Система "Вега-Science": особенности построения, основные возможности и опыт использования // Современные проблемы дистанционного зондирования Земли из космоса. 2021. Т. 18. № 6. (Scopus, Q3)– в печати.

Материалы конференции

8. *Лукьянова Р.Ю.* Структура средней атмосферы при формировании арктического стратосферного вихря зимой 2019-2020 гг. // Материалы 19-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. 2021. ИКИ РАН, 2021.
9. *Прошин А.А., Бурцев М.А., Сычуглов И.Г., Кобец Д.А.* Система контроля функционирования сложных распределенных программно-аппаратных комплексов, обеспечивающих доступ к данным ДЗЗ и результатам их обработки // Материалы 19-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2021. С. 51. DOI: 10.21046/19DZZconf-2021a.
10. *Константинова А.М., Руткевич Б.П., Балашов И.В.* Обновленный картографический web-интерфейс для работы с данными в системах семейства Созвездие-Вега // Материалы 19-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2021. С. 99. DOI: 10.21046/19DZZconf-2021a.
11. *Марченков В.В., Уваров И.А.* Архитектура web-приложения графического анализа рядов данных в системах спутникового мониторинга // Материалы 19-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2021. С. 103. DOI: 10.21046/19DZZconf-2021a.
12. *Прошин А.А., Лузян Е.А.* Анализ эффективности системы динамического блочного доступа к данным для предоставления их системам обработки // Материалы 19-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2021. С. 106. DOI: 10.21046/19DZZconf-2021a.
13. *Прошин А.А., Лузян Е.А., Матвеев А.М., Плотников Д.Е., Колбудаев П.А.* Организация обработки данных КМСС на основе использования системы динамического блочного доступа к данным // Материалы 19-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2021. С. 107. DOI: 10.21046/19DZZconf-2021a.
14. *Руткевич Б.П., Балашов И.В., Кашицкий А.В., Прошин А.А.* Универсальный модуль растеризации продуктов спутниковых данных UNISAT // Материалы 19-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований Российской академии наук. Москва, 2021. С. 109. DOI: 10.21046/19DZZconf-2021a.
15. *Матвеев А.М., Кобец Д.А., Радченко М.В.* Опыт использования различных подходов к распараллеленной обработке спутниковых данных // Материалы 19-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Электронный сборник материалов конференции. Институт космических исследований

Российской академии наук. Москва, 2021. С. 104. DOI: 10.21046/19DZZconf-2021a.
«Мониторинг» это должно быть в Космос Д

Тезисы, доклады

16. Лукьянова Р.Ю., Козловский А.Е. Эффекты асимметрии конвективного дрейфа ионосферной плазмы на границе полярной шапки // Сборник тезисов докладов 16-й ежегодной конференции «Физика плазмы в солнечной системе». 8-12 февраля 2021, Москва. ИКИ РАН. С. 208. (<https://plasma2021.cosmos.ru>)
17. Мещеряков А.В., Герасимов С.В., Селякин А.С. «Применение обучения с подкреплением в задаче построения маршрутов» // Ломоносовские чтения 2021. Секция вычислительная математика и кибернетика, 20-29 апреля 2021, Москва, Россия, тезисы докладов.
18. Немешаева А.А., Мещеряков А.В., Герасимов С.В. «Нейросетевые модели поиска скоплений галактик в микроволновом диапазоне по данным спутника Planck» // Ломоносовские чтения 2021. Секция вычислительная математика и кибернетика, 20-29 апреля 2021, Москва, Россия, тезисы докладов, устный доклад.
19. Мещеряков А.В., Герасимов С.В., Полевой А.В. «Анализ методов инкрементального моделирования в области директ-маркетинга» // Ломоносовские чтения 2021. Секция вычислительная математика и кибернетика, 20-29 апреля 2021, Москва, Россия, тезисы докладов.
20. Мещеряков А.В., Герасимов С.В., Сорока А.Г. Применение нейросетевых моделей внимания при обучении с подкреплением в задачах оптимизации маршрутов // Научная конференция "Тихоновские чтения 2021", 25-30 октября 2021, Москва, Россия, тезисы докладов.
21. Немешаева А.А., Мещеряков А.В., Герасимов С.В. Нейросетевая сегментация скоплений галактик в микроволновом и рентгеновском диапазонах // Научная конференция "Тихоновские чтения 2021", 25-30 октября 2021, Москва, Россия, тезисы докладов, устный доклад.
22. Малышева Н.В., Мещеряков А.В., Герасимов С.В. Модели машинного обучения для базовой классификации рентгеновских объектов СРГ/eRosita с использованием данных многоволновых обзоров неба // Научная конференция "Тихоновские чтения 2021", 25-30 октября 2021, Москва, Россия, тезисы докладов, устный доклад.
23. Мещеряков А.В., Герасимов С.В., Димитриенко А.Ю. Исследование методов распознавания действий людей в видео с минимальным использованием разметки // Научная конференция "Тихоновские чтения 2021", 25-30 октября 2021, Москва, Россия, тезисы докладов.
24. Мещеряков А.В. «» // Астрофизика высоких энергий сегодня и завтра 2021 (HEA-2021), 21-24 декабря 2021, Москва, Россия, тезисы докладов, устный доклад.
25. Бельведерский М., Мещеряков А.В., Гильфанов М.Р., Медведев П.С. Построение модели оптического отождествления рентгеновских источников СРГ/eРОЗИТА на примере данных области Дыры Локмана // Астрофизика высоких энергий сегодня и завтра 2021 (HEA-2021), 21-24 декабря 2021, Москва, Россия, тезисы докладов, устный доклад.

Публикации, подготовленные в соавторстве с зарубежными учеными

1. Lukianova, R., Kozlovsky, A., Lester, M. Upper stratosphere mesosphere-lower thermosphere perturbations during the formation of the Arctic polar night jet in 2019–2020. Geophysical Research Letters, 48, e2021GL094926. <https://doi.org/10.1029/2021GL094926>. (WoS, Scopus Q1)

Публикации по грантам РФФИ

1. Khorunzhev G. A., Meshcheryakov A. V., Medvedev P. S., Borisov V. D., Burenin R. A., Krivonos R. A., Uklein R. I., Shablovinskaya E. S., Afanasiev V. L., Dodonov S. N., Sunyaev R. A., Sazonov S. Yu., Gilfanov M. R. Discovery of the Most X-ray Luminous Quasar SRGE J170245.3+130104 at Redshift $z \approx 5.5$ // *Astronomy Letters*, 2021, volume 47, pages 123–140, DOI: 10.1134/S1063773721030026, **WoS/Scopus Q3**
2. Додин А.В., Шатский Н.И., Белинский А.А., Атапин К.Е., Бурлак М.А., Желтоухов С.Г., Татарников А.М., Постнов К.А., Черепашук А.М., Бельведерский М.И., Борисов В.Д., Буренин Р.А., Гильфанов М.Р., Кривonos Р.А., Медведев П.С., Мещеряков А.В., Сазонов С.Ю., Сюняев Р.А., Хорунжев Г.А. Оптическая спектроскопия квазаров, открытых телескопом SRG/eROSITA, на 2.5-м телескопе кавказской горной обсерватории ГАИШ МГУ // *Письма в Астрономический журнал*, издательство Наука (М.), том 47, № 10, с. 683–696, DOI: 10.31857/S0320010821100028, **WoS/Scopus Q3**
3. Burenin R. A., Bikmaev I. F., Gilfanov M. R., Grokhovskaya A. A., Dodonov S. N., Eiselevich M. V., Zaznobin I. A., Irtuganov E. N., Lyskova N. S., Medvedev P. S., Meshcheryakov A. V., Moiseev A. V., Sazonov S. Yu., Starobinsky A. A., Sunyaev R. A., Uklein R. I., Khabibullin I. I., Khamitov I. M., Churazov E. M. Observation of a Very Massive Galaxy Cluster at $z = 0.76$ in the SRG/eROSITA All-Sky Survey // *Astronomy Letters*, 2021, Vol. 47, No. 7, pp.443–453, DOI: 10.1134/S1063773721070045, **WoS/Scopus Q3**

Список публикаций по теме «Космос-Д» в рецензируемых изданиях

1. Lukianova, R., Kozlovsky, A., Lester, M. Upper stratosphere mesosphere-lower thermosphere perturbations during the formation of the Arctic polar night jet in 2019–2020. *Geophysical Research Letters*, 48, e2021GL094926. 2021. <https://doi.org/10.1029/2021GL094926>. (**WoS, Scopus Q1**)
2. Bikmaev I. F., Meshcheryakov A. V., Irtuganov E. N., Nikolaeva E. A., Sakhibullin N. A. Gumerov, R. I., Sklyanov A. S., Glushkov M. V., Khamitov I. M., Borisov V. D., Burenin R. A., Zaznobin I. A., Krivonos R. A., Lyapin A. R., Medvedev P. S., Sazonov S. Yu., Sunyaev R. A., Khorunzhev G. A., and Gilfanov M. R. Spectroscopic Redshift Determination for a Sample of Distant Quasars Detected by the SRG Observatory Based on RTT-150 Observations. II // *Astronomy Letters*, 2021, Vol. 47, No. 5, pp. 277–290, DOI: 10.1134/S1063773721050029 (**WoS, Scopus Q3**)
3. Proshin A.A., Loupian E.A., Bartalev S.A. Metodology for evaluation the effectiveness of the system of dynamic block access to data of ultra-large distributed remote sensing archives // *Ceur Workshop Proceedings. ITHPC 2021 - Short Paper Proceedings of the 6th International Conference on Information Technologies and High-Performance Computing*, 2021. Vol. 2930. P. 62–69. (**Scopus**)
4. Proshin A.A., Loupian E.A. Comprehensive analysis of the efficiency of the system of dynamic block access to data of ultra-large distributed remote sensing archives // *CEUR Workshop Proceedings. Proceedings of the All-Russian Conference With International Participation "Spatial Data Processing for Monitoring of Natural and Anthropogenic Processes" (SDM-2021)*, 2021. Vol. 3006. P. 222–230. DOI: 10.25743/SDM.2021.28.16.027. (**Scopus**)
5. Loupian E., Burtsev M., Proshin A., Kashnitskii A., Balashov I., Bartalev S., Konstantinova A., Kobets D., Radchenko M., Tolpin V., Uvarov I. Usage Experience and Capabilities of the VEGA-Science System // *Remote Sensing*. 2021. 14. – в печати (. (**Web of Sciences - Q1, Scopus**)
6. Мингалева И. В., Суворова З.В., Шубин В. Н., Мёрзлый А.М., Тихонов В.В., Талалаев А.Б., Мингалева В.С. Отличия прогнозов условий КВ-радиосвязи между передатчиком на средних широтах и приемником в арктическом регионе при использовании различных эмпирических моделей ионосферы, «Геомагнетизм и аэрономия», Т. 61, № 4, с.506–519,

2021.DOI: 10.31857/S001679402104009X. (**WoS, Scopus Q3**)

7. *Лупян Е.А., Прошин А.А., Бурцев М.А., Кашицкий А.В., Балашов И.В., Барталев С.А., Константинова А.М., Кобец Д.А., Мазуров А.А., Марченко В.В., Матвеев А.М., Радченко М.В., Сычугов И.Г., Толпин В.А., Уваров И.А.* Система "Вега-Science": особенности построения, основные возможности и опыт использования // Современные проблемы дистанционного зондирования Земли из космоса. 2021. Т. 18. № 6. (**Scopus**) – в печати.