



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение высшего
образования

**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ
ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»**

Кафедра информационных технологий и систем безопасности

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

(Магистерская диссертация)

На тему: Разработка интерактивной автоматизированной системы
обеспечения долгосрочной гидрометеорологической информацией
морской деятельности в Арктике: модуль обработки и анализа
геопространственных данных

Исполнитель Миронов Тимофей Алексеевич

(фамилия, имя, отчество)

Руководитель

(подпись)

кандидат географических наук

(ученая степень, ученое звание)

Шевчук Олег Игоревич

(фамилия, имя, отчество)

**«К защите допускаю»
Заведующий кафедрой**

(подпись)

доктор технических наук, профессор

(ученая степень, ученое звание)

Бурлов Вячеслав Георгиевич

(фамилия, имя, отчество)

« » 2025 г.

Санкт-Петербург

2025

Реферат

Магистерская диссертация: 146 страниц, 18 рисунков, 2 таблицы, 86 источников литературы.

Объект работы: методы и технологии обработки, анализа и визуализации геопространственных гидрометеорологических данных для Арктического региона.

Предмет работы: разработка модуля обработки и анализа геопространственных данных в составе интерактивной автоматизированной системы обеспечения долгосрочной гидрометеорологической информацией морской деятельности в Арктике.

Цель работы: разработка модуля для обработки и анализа геопространственных данных, предназначенного для использования в интерактивной автоматизированной системе, обеспечивающей долгосрочное предоставление гидрометеорологической информации для морской деятельности в Арктическом регионе.

Задачи работы:

- Провести аналитический обзор современных систем и методов сбора, обработки и предоставления гидрометеорологических данных, а также технологий и программных средств для работы с геопространственной информацией.
- Определить функциональные требования к модулю обработки и анализа геопространственных данных и его место в архитектуре общей интерактивной автоматизированной системы.
- Обосновать выбор инструментальных средств и технологических решений, включая язык программирования, библиотеки и источники данных.
- Разработать концептуальную модель алгоритмов обработки и визуализации геопространственных данных.
- Реализовать программный модуль на языке Python с использованием библиотек Pandas, NumPy и Plotly Express для обработки, анализа и визуализации данных о температуре воздуха из архива ERA-40.

- Провести апробацию разработанного модуля, проанализировать полученные результаты и производительность скриптов.
- Обсудить возможности практического применения модуля и перспективы его дальнейшего развития.

В выпускной квалификационной работе произведен анализ существующих систем и методов гидрометеорологического обеспечения морской деятельности в Арктике, выявлены их недостатки в части доступа и визуализации долгосрочных данных. Обоснован выбор Python и соответствующих библиотек для разработки модуля. Разработаны и протестированы два скрипта для обработки и визуализации данных о температуре воздуха из архива ERA-40, включая расчет среднеемноголетних значений и отображение фактических температур за выбранный период для заданных координат. Проведен анализ производительности созданных скриптов и намечены пути дальнейшего развития модуля.

Оглавление

Введение.....	6
1. Аналитический обзор и постановка задачи исследования	10
1.1 Значение гидрометеорологической информации для обеспечения безопасности и эффективности морской деятельности в Арктике.....	10
1.1.1 Особенности арктического региона	10
1.1.2 Значение Арктического региона для Российской Федерации.....	12
1.1.3 Основные гидрометеорологические параметры, значимые для арктической деятельности	14
1.2 Существующие системы и методы сбора, обработки и предоставления гидрометеорологической информации	17
1.2.1 Обзор российских и зарубежных информационных систем	18
1.2.2 Анализ недостатков существующих подходов и необходимость разработки новых решений	24
1.3 Роль геопространственных данных в гидрометеорологическом обеспечении	28
1.3.1 Типы и форматы геопространственных данных в метеорологии и океанографии	28
1.3.2 Проблемы и вызовы при работе с гидрометеорологическими геопространственными данными в Арктике	32
1.4 Обзор существующих технологий и программных средств для обработки, анализа и визуализации геопространственных данных	34
2. Методология проектирования и выбора технологических решений для модуля	48
2.1 Постановка задачи разработки модуля обработки и анализа геопространственных данных.....	48
2.1.1 Определение функциональных требований к модулю	49
2.1.2 Место модуля в архитектуре общей интерактивной автоматизированной системы.....	51
2.2 Обоснование выбора инструментальных средств и технологий разработки	53
2.2.1 Сравнительный анализ подходов и программных средств для обработки и визуализации геопространственных данных	53

2.2.2 Аргументация выбора языка программирования Python и его ключевых библиотек	61
2.3 Исследование и выбор источника гидрометеорологических данных	66
2.3.1 Обзор и сравнение доступных архивов и наборов данных о температуре воздуха для Арктического региона	66
2.3.2 Аргументация выбора ERA-40	72
2.4 Концептуальное проектирование алгоритмов обработки и визуализации ...	73
2.4.1 Выбор и обоснование методов анализа данных	73
2.4.2 Выбор и обоснование методов визуализации	76
3. Разработка модуля обработки и анализа геопространственных данных	77
3.1 Общая информация о функциях модуля и их реализации.....	77
3.2 Используемый источник данных	80
3.3 Подробное описание первого скрипта для анализа и визуализации температурных данных.....	83
3.4 Подробное описание второго скрипта для анализа и визуализации температурных данных.....	95
4. Апробация и анализ результатов работы модуля	105
4.1 Методика тестирования разработанного модуля.....	105
4.1.1 Тестовые наборы данных	105
4.1.2 Критерии оценки корректности работы	107
4.2. Результаты тестирования функциональности скриптов.	109
4.2.1 Примеры визуализаций, полученных с помощью разработанных скриптов.	109
4.2.2 Сравнение полученных результатов с ожидаемыми.....	120
4.3 Анализ производительности созданных скриптов	122
4.4 Обсуждение возможностей практического применения модуля в составе интерактивной системы.....	125
4.5 Перспективы дальнейшего развития модуля	127
Заключение	131
Список использованной литературы.....	135

Введение

Актуальность этой работы обусловлена стратегическим значением Арктического региона для Российской Федерации и возрастающей интенсивностью морской деятельности в акватории Северного Ледовитого океана. Уникальные природно-климатические условия региона, включая экстремально низкие температуры, наличие морского льда, частые и резкие изменения погодных условий, длительные периоды полярной ночи и полярного дня, а также значительную удаленность от развитой инфраструктуры, создают серьезные вызовы для обеспечения безопасности и эффективности морских операций. В условиях продолжающихся глобальных климатических изменений, оказывающих значительное воздействие на арктическую среду, необходимость в точной и долгосрочной гидрометеорологической информации приобретает критическую важность.

Существующие системы гидрометеорологического обеспечения, несмотря на их постоянное развитие, сталкиваются с проблемами доступности и удобства использования долгосрочных архивов геопространственных данных, особенно применительно к Арктике. Помимо этого, в последнее время наблюдается значительное усложнение процесса получения доступа к наборам данных для пользователей, расположенных на территории Российской Федерации. Это создает препятствия для проведения научного анализа, долгосрочного планирования и принятия обоснованных решений в сферах морской деятельности, таких как судоходство, добыча полезных ископаемых и научные исследования.

Разработка отечественной интерактивной автоматизированной системы, способной обеспечивать оперативный и удобный доступ к визуализированным архивным гидрометеорологическим данным потребителям в Российской Федерации, является важной задачей для преодоления существующих проблем и

эффективного информационного обеспечения морской деятельности в Арктическом регионе.

Цель работы заключается в разработке модуля для обработки и анализа геопространственных данных, предназначенного для использования в интерактивной автоматизированной системе, обеспечивающей долгосрочное предоставление гидрометеорологической информации для морской деятельности в Арктическом регионе.

Для достижения указанной цели были поставлены следующие задачи:

- Провести аналитический обзор современных систем и методов сбора, обработки и предоставления гидрометеорологических данных, а также технологий и программных средств, предназначенных для работы с геопространственной информацией.
- Определить функциональные требования к модулю для обработки и анализа геопространственных данных, а также его интеграцию в общую архитектуру интерактивной автоматизированной системы.
- Обосновать выбор инструментальных средств и технологических решений для разработки модуля, включая выбор языка программирования, ключевых библиотек и источников гидрометеорологических данных.
- Разработать концептуальную модель алгоритмов для обработки и визуализации геопространственных данных с обоснованием методов анализа и визуализации.
- Реализовать программный модуль для обработки и анализа геопространственных данных на основе выбранных технологий и алгоритмов.
- Провести апробацию разработанного модуля и анализ полученных результатов.
- Обсудить возможности практического применения модуля, а также перспективы его дальнейшего развития.

В рамках данной работы были применены методы системного анализа для определения функционального места разрабатываемого модуля в общей архитектуре системы. Сравнительный анализ использовался для выбора инструментов и технологий разработки, обеспечивающих оптимальное соответствие техническим требованиям и стандартам. Концептуальное проектирование алгоритмов обработки и визуализации данных базировалось на детальном анализе функциональных требований к модулю.

Для практической реализации модуля было использовано объектно-ориентированное программирование на языке Python. В процессе разработки применялись специализированные библиотеки для работы с данными (Pandas, NumPy) и их визуализации (Plotly Express).

Тестирование разработанных программных скриптов проводилось на основе реальных гидрометеорологических данных с использованием выбранных тестовых сценариев и критериев оценки качества.

Настоящая работа состоит из введения, четырёх разделов, заключения и списка использованных источников.

Глава 1 «Аналитический обзор и постановка задачи исследования» посвящена анализу значимости гидрометеорологической информации для Арктического региона, обзору существующих информационных систем, роли геопространственных данных и технологий их обработки.

Глава 2 «Методология проектирования и выбора технологических решений для модуля» рассматривает постановку задачи разработки программного модуля, обоснование выбора инструментальных средств и источников данных, а также концепцию проектирования алгоритмов.

Глава 3 «Разработка модуля обработки и анализа геопространственных данных» содержит описание функциональных возможностей модуля, использованных источников данных и детальное изложение разработанных программных скриптов.

Глава 4 «Апробация и анализ результатов работы модуля» представляет методику тестирования, результаты апробации, анализ производительности программных скриптов, а также обсуждение возможностей практического применения и перспектив развития модуля.

1. Аналитический обзор и постановка задачи исследования

1.1 Значение гидрометеорологической информации для обеспечения безопасности и эффективности морской деятельности в Арктике

1.1.1 Особенности арктического региона

Арктический регион, включающий акваторию Северного Ледовитого океана и прилегающие к ней территории Евразии и Северной Америки, характеризуется специфическими природными условиями, которые оказывают значительное влияние на все виды деятельности, включая морскую. Понимание этих факторов имеет критическое значение для разработки эффективных систем гидрометеорологического обеспечения.

Арктический регион отличается экстремально низкими температурами воздуха в зимний период, которые могут достигать критических значений. Длительные полярные ночи и короткий летний сезон формируют уникальный радиационный баланс и температурный режим.

Одной из ключевых особенностей региона является наличие морского льда, который характеризуется различной толщиной, возрастом и концентрацией. Ледовый покров существенно затрудняет навигацию и требует использования судов подходящего класса, а также применение ледокольной проводки. Состояние ледового покрова (его толщина, сплоченность, скорость и направление дрейфа) является важным фактором, определяющим безопасность и экономическую эффективность морской деятельности.

Арктический регион подвержен частым и резким изменениям погодных условий, включая сильные ветровые нагрузки, метели, туманы и возникновение обледенения. Эти явления могут возникать внезапно и представлять серьезную угрозу для морской деятельности, что может привести к задержкам, аварийным ситуациям и утрате оборудования.

Полярный день и полярная ночь создают уникальные условия для навигации и проведения различных видов работ. Длительные периоды темноты или, наоборот, постоянного дневного света оказывают влияние на биологические ритмы человека и требуют коррекции рабочих графиков и использования специализированного осветительного оборудования.

Большинство районов Арктического региона характеризуются значительной удаленностью от основных промышленных и логистических центров. Инфраструктура, включая порты, средства связи и спасательные службы, развита недостаточно, что усложняет проведение операций и повышает риски в случае чрезвычайных ситуаций.

Арктическая экосистема отличается высокой чувствительностью к внешним воздействиям и медленным восстановлением после антропогенного влияния. Любая хозяйственная деятельность в этом регионе должна осуществляться с соблюдением строгих экологических норм и стандартов, направленных на минимизацию негативного воздействия на окружающую среду.

Арктический регион является одной из наиболее подверженных глобальным климатическим изменениям областей. Наблюдается устойчивая тенденция к уменьшению площади и толщины морского льда, изменению гидрологических режимов рек и таянию вечной мерзлоты. Эти процессы, с одной стороны, создают новые возможности для судоходства и освоения природных ресурсов, но, с другой стороны, увеличивают частоту экстремальных погодных явлений и ставят новые задачи перед обеспечением безопасности.

Экстремальные климатические условия, удаленность и разреженность сети наблюдательных станций затрудняют сбор точной и своевременной гидрометеорологической информации. Это подчеркивает важность разработки автоматизированных систем, способных осуществлять сбор, обработку и предоставление данных для удаленных регионов.

Специфические особенности Арктического региона предъявляют повышенные требования к гидрометеорологическому обеспечению морской деятельности. Точная и своевременная информация о температуре воздуха,

ледовом покрове, ветровом режиме, волнении и других параметрах является ключевым фактором для обеспечения безопасности судоходства, повышения эффективности исследовательских, транспортных и добывающих операций, а также сохранения уникальной природы Арктики. В связи с этим разработка интерактивной автоматизированной системы для долгосрочного обеспечения гидрометеорологической информацией, включая модуль обработки и анализа геопространственных данных, представляет собой актуальную и практически значимую задачу [1].

1.1.2 Значение Арктического региона для Российской Федерации

Арктический регион имеет важное стратегическое значение для Российской Федерации, обусловленное комплексными экономическими, геополитическими, ресурсными, транспортными и экологическими факторами.

Арктическая зона Российской Федерации характеризуется значительными запасами природных ресурсов, включая нефть, природный газ, газовый конденсат, руды цветных и драгоценных металлов (никель, медь, платина, алмазы и др.). Освоение этих месторождений является критически важным для обеспечения энергетической безопасности страны и пополнения государственного бюджета. Арктический шельф рассматривается как перспективный регион для добычи углеводородов.

СМП представляет собой кратчайший морской маршрут между европейской частью России и Дальним Востоком, а также потенциальный транзитный коридор между Европой и Азией. В условиях изменения климата и уменьшения площади ледового покрова навигация по СМП становится более доступной, что способствует развитию транспортной инфраструктуры, снижению логистических издержек и освоению прибрежных арктических территорий. Российская Федерация активно инвестирует в развитие инфраструктуры СМП, включая строительство ледокольного флота, морских портов и систем навигационно-гидрографического обеспечения.

Арктические моря обладают значительными биоресурсами, включая ценные промысловые виды рыб и морепродуктов, что делает рыболовство важной отраслью для прибрежных регионов и национальной экономики в целом.

Арктика представляет собой обширный регион с важными военно-стратегическими характеристиками. Контроль над арктическими территориями и акваториями обеспечивает Российской Федерации стратегическое присутствие в зоне, возможность проецирования военной силы и защиты национальных интересов.

Северные границы Российской Федерации являются одними из самых протяженных в мире, и обеспечение их безопасности является приоритетной задачей государства. В Арктике расположены объекты военной инфраструктуры, такие как базы Северного флота, аэродромы и системы противовоздушной обороны, играющие ключевую роль в реализации оборонной стратегии страны.

Арктический регион становится ареной как для международного сотрудничества в области научных исследований, экологии и спасательных операций, так и для геополитического соперничества. Российская Федерация активно отстаивает свои интересы в Арктике в соответствии с международным правом и стремится к лидерству в формировании региональной повестки.

Арктика представляет собой уникальную природную зону, играющую ключевую роль в глобальных климатических процессах. Исследования арктических экосистем, ледового покрова, океанографии и атмосферы имеют важное значение для понимания климатических изменений и их последствий.

Освоение Арктического региона стимулирует развитие новых технологий и научных исследований в области судостроения (особенно ледокольного флота), добычи полезных ископаемых в экстремальных условиях, гидрометеорологии, связи и навигации в высоких широтах.

Для ряда арктических регионов Российской Федерации морские и речные пути являются основными, а иногда и единственными транспортными артериями, обеспечивающими социально-экономическое развитие территорий и логистику.

Арктические экосистемы отличаются высокой уязвимостью к антропогенному воздействию и климатическим изменениям. Сохранение уникального биоразнообразия и минимизация экологических рисков при хозяйственной деятельности являются ключевыми задачами для Российской Федерации [2].

1.1.3 Основные гидрометеорологические параметры, значимые для арктической деятельности

Морская деятельность в Арктическом регионе, включая судоходство, разведывательные и добывающие операции, научные исследования и обеспечение безопасности, осуществляется в экстремальных климатических условиях. Успешное и безопасное выполнение операций в данном регионе критически зависит от достоверной и своевременной гидрометеорологической информации. Комплексное понимание гидрометеорологических условий является ключевым фактором для разработки маршрутов, принятия стратегических решений и минимизации рисков.

К числу наиболее значимых гидрометеорологических параметров, обеспечивающих безопасность и эффективность морской деятельности в Арктике, относятся:

Температура воздуха

Данный параметр является фундаментальным для арктической морской деятельности и служит одним из основных индикаторов климатических условий региона. Температура воздуха оказывает прямое влияние на процессы льдообразования и таяния морского льда и пресноводных водоемов, определяя доступность морских путей и сложность навигации. Низкие температуры

воздуха повышают риск обледенения судов и морских платформ, что может привести к утрате устойчивости, повреждениям конструкций и оборудования, а также препятствовать работе экипажа. Экстремально низкие температуры требуют применения специализированного оборудования и материалов, адаптированных к условиям низких температур, а также строгого соблюдения протоколов безопасности. Анализ временных рядов данных о температуре воздуха в долгосрочной перспективе позволяет выявлять климатические тренды в Арктическом регионе, что имеет ключевое значение для стратегического планирования морской деятельности, оценки потенциальных рисков и адаптации инфраструктуры. Также анализ этих данных позволяет создавать прогнозы на ближайшее будущее и оценивать текущую ситуацию.

Именно анализ и визуализация временных рядов данных о температуре воздуха, относящихся к конкретным географическим точкам, являются основной задачей практической части данной работы.

Морской лед

Параметры, характеризующие морской лед (распространение, толщина, возраст, сплоченность, дрейф, наличие торосов и разводий), являются критически важными для навигации в Арктическом регионе. Ледовые условия определяют возможность прохождения судов по различным маршрутам, необходимость ледокольной проводки, скорость движения и связанные с этим временные и финансовые затраты. Долгосрочные данные о ледовых условиях и их изменчивости необходимы для оценки сезонной доступности морских путей, таких как СМП.

Температура морской воды

Температура воды в приповерхностном слое оказывает влияние на скорость льдообразования и таяния. В более глубоких слоях температурные показатели влияют на состояние подводных конструкций, работу гидроакустического оборудования, а также на биологическую активность в регионе. Долгосрочные данные о температуре воды необходимы для понимания

океанографических процессов, их влияния на ледовую обстановку и экологические системы.

Ветер

Скорость и направление ветра оказывают влияние на морское волнение, вызывают дрейф ледяных масс, определяют условия проведения грузовых операций и маневрирования судов. Сильные ветры могут представлять прямую угрозу для судов и морских сооружений. В сочетании с низкими температурами, ветер усиливает эффект охлаждения (ветровой холод), повышая риск обледенения и ухудшая условия для работы людей.

Волнение

Параметры, характеризующие морское волнение (высота, период, направление), важны для оценки нагрузки на корпус судна, определения оптимальной скорости движения, планирования шельфовых работ и обеспечения безопасности экипажа. В условиях открытой воды или в разводьях между льдами волновые характеристики могут быть значительными.

Видимость

В Арктическом регионе видимость часто ограничена из-за туманов (особенно в летний период в условиях открытой воды), снегопадов и метелей. Низкая видимость значительно повышает риски столкновений, посадки на мель и затрудняет проведение различных видов работ.

Осадки

Тип и интенсивность осадков (снег, дождь, мокрый снег) влияют на видимость, скорость обледенения судов и сооружений, а также на накопление снежного покрова на палубах и ледяных поверхностях. Снегопады могут скрывать опасные объекты (трещины, полыньи) на морском льду.

Океанические течения

Течения оказывают влияние на скорость и направление дрейфа ледяных масс, что имеет критическое значение для навигации. Они также влияют на движение судов и распределение температуры и солености в морской воде, что отражается на ледовых условиях.

Для обеспечения устойчивого функционирования морской деятельности в Арктическом регионе необходим комплексный подход к сбору, обработке, анализу и предоставлению данных по всем перечисленным параметрам. Однако, как было отмечено ранее, температура воздуха является ключевым фактором, оказывающим влияние на многие другие параметры (в первую очередь на ледовую обстановку и условия выполнения работ), и анализ ее долгосрочных изменений представляет собой важный шаг для понимания динамики арктической среды. Разработка модуля для обработки и анализа данных о температуре воздуха, основанного на многолетних временных рядах, закладывает основу для создания более комплексной системы гидрометеорологического обеспечения [3, 4].

1.2 Существующие системы и методы сбора, обработки и предоставления гидрометеорологической информации

Гидрометеорологическая информация, особенно для Арктического региона, как мы выяснили выше, является критически важной для различных отраслей экономики, обеспечения общественной безопасности, научных исследований и других сфер деятельности. Процесс сбора, обработки и распространения данных представляет собой сложную и многогранную задачу, требующую взаимодействия специалистов различных дисциплин, включая метеорологию, гидрологию, океанологию, информационные технологии и инженерию.

1.2.1 Обзор российских и зарубежных информационных систем

Федеральная служба по гидрометеорологии и мониторингу окружающей среды (Росгидромет) является ключевым национальным органом, ответственным за предоставление гидрометеорологических услуг в Российской Федерации. Организационная структура Росгидромета включает региональные управления по гидрометеорологии и мониторингу окружающей среды (УГМС), такие как Западно-Сибирское УГМС и Северное УГМС, а также специализированные центры, включая Главный авиационный метеорологический центр.

Росгидромет предоставляет широкий спектр гидрометеорологической информации, включая прогнозы погоды различной заблаговременности, штормовые предупреждения об опасных явлениях, климатические данные и информацию о состоянии окружающей среды. Для распространения оперативной информации используется мобильное приложение "ПОГОДА Гидрометцентр России".

Особая роль отводится Росгидромету в обеспечении безопасности судоходства по Северному морскому пути. Участие службы в гидрографическом и метеорологическом обеспечении этого стратегического транспортного коридора подчеркивает ее значимость для национальных экономических и инфраструктурных проектов.

В настоящее время Росгидромет активно модернизирует свою наблюдательную сеть, особенно в Арктическом регионе. Эти усилия включают обновление труднодоступных станций на побережье и островах Северного Ледовитого океана с использованием современных автоматизированных метеорологических комплексов. Модернизация направлена на повышение точности и достоверности метеорологической информации в регионе, который имеет ключевое значение в контексте изменения климата и развития хозяйственной деятельности.

Хотя термин "полярная станция" официально не используется во внутренних документах Росгидромета, в Арктике функционирует сеть

труднодоступных гидрометеорологических станций. Эти станции, расположенные в сложных физико-географических и суровых климатических условиях, играют важную роль в мониторинге природных процессов в полярных регионах. Исторически создание этой сети было связано с необходимостью информационного обеспечения судоходства по Северному морскому пути.

Росгидромет также предоставляет долгосрочные прогнозы, включая прогнозы весеннего половодья, аномалий температуры и осадков на несколько месяцев вперед. Однако следует отметить, что долгосрочные гидродинамические прогнозы находятся на начальном этапе развития и носят экспериментальный характер, что обусловлено сложностью прогнозирования на большие временные горизонты.

Активное использование геоинформационных технологий (ГИС) и геопространственных данных является важным элементом деятельности Росгидромета. Эти технологии применяются для решения широкого круга задач, включая мониторинг и прогнозирование паводков и других природных явлений, как, например, в проекте "ГИС-Амур", что свидетельствует об интеграции пространственного анализа в оперативные и исследовательские процессы службы [5].

Национальное управление океанических и атмосферных исследований (NOAA) США является одним из ведущих мировых центров в области гидрометеорологии. NOAA выполняет широкий спектр функций, включая метеорологическое прогнозирование, климатический мониторинг и океанографические исследования. Особое внимание уделяется арктическому региону, где NOAA осуществляет сбор данных, научные исследования и разработку прогнозов. Значительное внимание к Арктике обусловлено ключевой ролью региона в глобальных климатических процессах и его влиянием на метеорологические условия на глобальном уровне.

NOAA предоставляет обширный спектр метеорологических данных и прогнозов различной временной продолжительности — от краткосрочных

прогнозов на 6–10 дней до долгосрочных сезонных прогнозов. Наличие разнообразных прогнозов с различной заблаговременностью и форматами указывает на комплексный подход, направленный на удовлетворение потребностей различных потребителей.

NOAA также выпускает специализированные морские прогнозы, включая прогнозы для арктического региона. Эти прогнозы играют критически важную роль в обеспечении безопасности морской деятельности в экстремальных условиях арктической зоны. Наличие специализированных морских прогнозов для Арктики подчёркивает значимость метеорологических услуг для морских операций в арктической зоне.

NOAA располагает значительными архивами исторических метеорологических данных, предоставляя различные способы доступа к этим данным. Поддержание таких архивов является ценным ресурсом для климатологических исследований и анализа долгосрочных климатических тенденций. Платформы, такие как Climate Data Online (CDO), и наличие многолетних архивов данных свидетельствуют о том, что NOAA обеспечивают сохранение и обеспечение доступа к исторической метеорологической информации для научного анализа.

Недавние решения об ограничении поддержки ключевых арктических данных в Национальном центре данных по снегу и льду (NSIDC) со стороны NOAA вызывают беспокойство относительно долгосрочной доступности и качества важной информации для мониторинга климатических изменений в Арктике. Сокращение поддержки может ограничить возможности оперативного реагирования на запросы пользователей, решения проблем и обеспечения тщательного обслуживания этих критически важных наборов данных, необходимых для исследователей [6].

Европейский центр среднесрочных прогнозов погоды (ECMWF) представляет собой ведущую международную организацию в области метеорологии и климатологии. Центр применяет Интегрированную систему

прогнозирования (IFS) для разработки глобальных прогнозов погоды и активно занимается исследованиями, связанными с Арктическим регионом. В рамках международного сотрудничества ECMWF инициировал создание службы Copernicus Arctic Regional Reanalysis (CARRA), которая представляет собой региональный реанализ высокого разрешения, специально разработанный для анализа климатических условий Арктики. CARRA использует передовые методы ассимиляции данных и моделирования, что способствует углублению понимания климатических процессов в этом регионе.

ECMWF предоставляет долгосрочные метеорологические прогнозы, включая сезонные и расширенные прогнозы с заблаговременностью до 13 месяцев. Для учета неопределенностей в долгосрочных прогнозах используется система ансамблевого прогнозирования. SEAS5 (Система сезонного прогнозирования 5) представляет собой ансамбль из 51 модели с высоким пространственным разрешением, демонстрируя передовые возможности в области сезонного климатического прогнозирования.

Океанический реанализ ECMWF (OCEAN5) играет ключевую роль в инициализации комплексных прогнозов атмосферы и океана. OCEAN5 является ансамблевой системой реанализа, которая включает ассимиляцию различных океанических наблюдений, таких как температура, соленость и концентрация морского льда.

Набор данных CARRA, предоставляемый ECMWF, включает широкий спектр атмосферных и поверхностных переменных с высоким пространственным разрешением. Эти данные представляют собой ценный ресурс для изучения метеорологических и климатических процессов в Арктике. Переменные CARRA включают компоненты ветра, температуру, влажность, осадки, и характеристики облачности, что позволяет проводить детальный анализ различных аспектов арктической среды [7].

Сервис Яндекс.Погода, разработанный компанией Яндекс, представляет собой комплексную систему прогнозирования метеорологических условий. В

основе функционирования сервиса лежит инновационная технология "Метеум" от Яндекса, которая интегрирует методы машинного обучения и использует широкий спектр источников данных. Новая версия технологии, "Метеум 2.0", применяет гибридный подход, сочетающий традиционные метеорологические модели с передовыми методами искусственного интеллекта, включая нейронные сети.

Для обеспечения точности прогнозов Яндекс.Погода использует разнообразные источники данных, такие как метеорологические станции, спутниковые наблюдения, радиолокационные системы и пользовательские отчеты. Особое внимание уделяется функции "Зонтики", которая позволяет пользователям сообщать о фактических осадках, что значительно улучшает качество и достоверность данных.

Сервис предлагает широкий спектр функциональных возможностей, включая детальные прогнозы на уровне отдельных улиц, карты осадков и специализированные метеорологические прогнозы, предназначенные для различных категорий пользователей, включая автомобилистов. Яндекс.Погода обеспечивает высокую пространственную детализацию прогнозов, что позволяет предоставлять локализованную информацию, актуальную для конкретных районов и видов деятельности.

Для разработчиков и предприятий Яндекс.Погода предоставляет программный интерфейс приложения (API), который позволяет интегрировать метеорологические данные в собственные информационные системы и приложения. API предлагает широкий набор параметров (более 150) и возможность выбора высокого пространственного разрешения (до 250x250 метров даже в горных районах), что делает его универсальным инструментом для различных коммерческих и исследовательских целей.

Сервис также предоставляет метеорологическую информацию для высокоширотных регионов, включая Арктическую зону. Это свидетельствует о широком географическом охвате и доступности услуг Яндекс.Погода даже для населенных пунктов, расположенных за Полярным кругом [8].

Windy.com представляет собой специализированный инструмент для визуализации метеорологических прогнозов. Платформа использует данные из нескольких ведущих мировых метеорологических моделей, таких как ECMWF, GFS, ICON и другие. Windy.com осуществляет агрегацию данных из различных глобальных моделей, что позволяет пользователям проводить сравнительный анализ прогнозов и оценивать степень их неопределенности.

Отличительной особенностью платформы является её развитый функционал визуализации, включающий карты ветра, спутниковые снимки и метеорологические радары. Windy.com предоставляет высокоточные данные о ветровых условиях, что делает его востребованным среди специалистов, таких как пилоты и моряки.

Особое внимание на платформе уделяется морским метеорологическим прогнозам, включая анализ волнения и приливных колебаний. Windy.com предлагает специализированные функции, необходимые для морской навигации и рыболовства, такие как прогнозирование высоты и периода волн, а также приливов. Эти данные удовлетворяют потребности яхтсменов, серферов и рыболовов.

Для обеспечения точности прогнозов в Арктическом регионе Windy.com применяет различные глобальные и региональные метеорологические модели, включая специализированные модели для Арктики, такие как NAM Alaska и AROME Arctic. Упоминание этих моделей и обсуждение возможности интеграции дополнительных арктических моделей свидетельствуют о стремлении Windy.com предоставлять точные и детализированные метеорологические прогнозы для данного региона.

Премиум-пользователи Windy.com имеют доступ к архиву исторических данных, включающему спутниковые снимки и радарные данные за последний год. Однако, исторические данные, полученные с помощью метеорологических моделей, как правило, не сохраняются и не предоставляются пользователям. Обсуждения на специализированном форуме Windy.com подтверждают, что

платформа не хранит и не предоставляет доступ к архивным данным модельных прогнозов [9].

Яндекс Карты погоды и Windy.com представляют собой современные информационные системы, которые дополняют и в определенной степени конкурируют с традиционными гидрометеорологическими службами, такими как Росгидромет, Национальное управление океанических и атмосферных исследований (NOAA) и Европейский центр среднесрочного прогнозирования погоды (ECMWF).

Яндекс Погода отличается высокой пространственной детализацией данных и интеграцией с другими сервисами компании Яндекс. Использование технологий машинного обучения позволяет значительно повысить точность метеорологических прогнозов.

Windy.com выделяется развитой системой визуализации данных, особенно в отношении параметров ветра, а также предоставляет доступ к широкому спектру мировых прогностических моделей. Это позволяет пользователям самостоятельно проводить сравнительный анализ различных сценариев развития метеорологических процессов.

Обе платформы способствуют демократизации доступа к метеорологической информации, обеспечивая ее наглядность и удобство для широкого круга пользователей.

1.2.2 Анализ недостатков существующих подходов и необходимость разработки новых решений

Долгосрочное прогнозирование метеорологических условий сталкивается с фундаментальными трудностями в достижении абсолютной точности. Хаотический характер атмосферных процессов и недостаточное понимание механизмов взаимодействия между гидросферой и атмосферой неизбежно приводят к неопределенности в прогнозах на длительные временные интервалы.

Разреженные сети метеорологических станций, особенно в труднодоступных регионах, таких как арктическая зона, способствуют возникновению неопределенностей и ограничений в точности метеорологических прогнозов и климатических моделей. Арктическая область характеризуется суровыми климатическими условиями, низкой плотностью населения и недостаточно развитой инфраструктурой, что затрудняет организацию регулярного метеорологического мониторинга. Создание и эксплуатация сложных климатических моделей требуют значительных вычислительных мощностей, что может ограничивать частоту и детализацию прогнозов.

Ассимиляция спутниковых данных в арктической зоне, особенно в зимний период, представляет собой сложную задачу из-за потенциальных ошибок в моделях и неопределенностей, связанных с радиационным переносом над снежным и ледяным покровами.

Большинство существующих станций и пунктов наблюдения Росгидромета в арктической зоне требуют модернизации, включая замену устаревшего оборудования (в том числе энергетического), улучшение систем связи и передачи данных, оснащение станций современными измерительными приборами.

Сокращение финансовой поддержки ключевых наборов данных о состоянии морского льда и ледников, предоставляемых Национальным центром данных по снегу и льду (NSIDC), может затруднить проведение научных исследований и анализ последствий климатических изменений в арктической зоне. Уведомление NSIDC и публикации в научных изданиях подчеркивают потенциальные последствия данного сокращения, включая ограничение возможностей оперативного реагирования на запросы пользователей, что может негативно сказаться на качестве и доступности данных, критически важных для научных исследований.

Сервисы Яндекс.Погода и Windy.com, несмотря на свой продвинутый функционал в области предоставления текущих прогнозов и визуализации

метеорологических данных, имеют определенные ограничения, особенно в контексте удобного доступа к архивным данным в визуализированной форме.

Для Яндекс.Погоды основным недостатком является отсутствие интуитивно понятного пользовательского интерфейса для анализа исторических метеорологических данных за длительные периоды времени с графическим представлением. Хотя API Яндекс.Погоды предоставляет возможность получения архивных данных за более чем 70 лет, включающих различные метеорологические параметры, такие как температура, уровень осадков и скорость ветра, эти данные предоставляются в формате CSV или JSON, что требует дополнительной обработки и анализа для визуализации. В стандартном пользовательском интерфейсе отсутствует функция, позволяющая легко выбрать необходимый временной период и просмотреть динамику метеорологических изменений на графиках или картах. Пользователи могут получать доступ только к текущим метеорологическим данным и прогнозу на ближайшие дни. Для анализа исторических тенденций или сравнения метеорологических условий за различные периоды пользователям приходится использовать сторонние инструменты или самостоятельно обрабатывать данные, полученные через API.

Что касается Windy.com, данный сервис демонстрирует высокую эффективность в визуализации текущих и краткосрочных метеорологических прогнозов, особенно в отношении анализа ветровых условий. Однако при работе с архивными данными возможности платформы существенно ограничены. Windy.com предоставляет премиум-пользователям доступ к архиву данных, полученных с радаров и спутниковых снимков, за последний год. Это позволяет пользователям визуально отслеживать прошлые метеорологические явления, такие как осадки и облачность. Однако сервис не предлагает удобного инструмента для визуализированного анализа исторических прогнозов погоды или других метеорологических параметров, таких как температура или атмосферное давление, за длительные периоды времени. Пользователи могут получить доступ к данным метеостанций за последние 7 дней, однако эта информация предоставляется в виде отдельных значений, а не в виде наглядных

графиков или карт, отображающих динамику изменений. Таким образом, для получения полного представления о прошлых метеорологических условиях и их визуализации пользователям Windy.com также приходится искать альтернативные решения или использовать другие сервисы.

На основании вышеизложенного можно сделать вывод о необходимости разработки специализированной информационной системы, обеспечивающей оперативный и удобный доступ к визуализированным архивным данным гидрометеорологического характера, особенно для стратегически значимого региона Арктики.

Современные сервисы, такие как Яндекс.Погода и Windy.com, предоставляют широкие функциональные возможности для прогнозирования текущих погодных условий и их визуализации. Однако они имеют существенные ограничения в доступе к историческим данным.

Эти ограничения создают значительные препятствия для научных исследований, планирования и принятия решений в различных сферах деятельности, особенно в морской отрасли Арктического региона. Для эффективного долгосрочного планирования и анализа климатических тенденций, таких как изменения ледовой обстановки, температурных режимов или повторяемости опасных метеорологических явлений, специалистам необходимо осуществлять поиск, обработку и анализ фрагментированных данных из различных источников, что требует значительных временных и ресурсных затрат.

Разработка интерактивной автоматизированной системы для предоставления визуализированных архивных данных гидрометеорологического характера позволит:

- Упростить и ускорить процесс получения необходимой информации: пользователи смогут легко выбирать интересующий временной интервал и географическую область, а также получать данные в удобной графической форме (графики, карты и т.д.) без необходимости сложной предварительной обработки.

- Повысить эффективность анализа: визуализация данных способствует быстрому выявлению закономерностей, трендов и аномалий, что значительно облегчает анализ долгосрочных изменений и обоснованное принятие решений.
- Обеспечить информационную поддержку для долгосрочного планирования и принятия решений: удобный доступ к архивным данным позволит более точно оценивать риски и потенциальные возможности, связанные с морской деятельностью в Арктическом регионе.
- Способствовать развитию научных исследований: лёгкий доступ к историческим данным в визуализированной форме откроет новые перспективы для изучения климатических изменений в Арктике и их влияния на морскую среду.

Таким образом, ввиду ограничений существующих сервисов в предоставлении удобного доступа к визуализированным архивным данным гидрометеорологического характера, разработка специализированной информационной системы является необходимой для обеспечения эффективного планирования и принятия решений, а также для углубления научных исследований, связанных с Арктическим регионом.

1.3 Роль геопространственных данных в гидрометеорологическом обеспечении

1.3.1 Типы и форматы геопространственных данных в метеорологии и океанографии

Геопространственные данные являются ключевым элементом в области метеорологии и океанографии. Они предоставляют критически важную информацию для анализа и прогнозирования атмосферных и океанических процессов, особенно в таких динамичных и стратегически значимых регионах,

как Арктика. Разнообразие типов и форматов данных позволяет ученым и специалистам получать всестороннее представление об окружающей среде.

Одним из основных типов геопространственных данных, используемых в метеорологии, являются спутниковые снимки. Эти изображения, полученные с космических аппаратов, находящихся на геостационарных или полярных орбитах, обеспечивают непрерывный мониторинг атмосферных явлений на обширных территориях. Спутники предоставляют точную и детализированную информацию о развитии атмосферных процессов. Геостационарные спутники, расположенные на фиксированной точке над Землёй, постоянно наблюдают за одной и той же областью, что позволяет отслеживать эволюцию гидрометеорологических процессов. Полярно-орбитальные спутники, проходящие вблизи полюсов, сканируют полосу земной поверхности при каждом витке орбиты, предоставляя такие данные, как состояние озонового слоя, снежного покрова и температура поверхности океана.

В метеорологии используются три основных типа спутниковых снимков: снимки в видимом спектре, в инфракрасном и изображения водяного пара [10]. Снимки в видимом спектре формируются за счет отраженного солнечного излучения и доступны только в дневное время. На этих изображениях облачность отображается в виде белых областей, поверхность Земли имеет серый оттенок, а вода — темный. В зимний период снежный покров также выглядит белым, что затрудняет дифференциацию облаков и снега. Инфракрасные снимки показывают тепловую энергию, излучаемую поверхностью Земли, и доступны как в дневное, так и в ночное время. На таких изображениях теплые объекты, такие как поверхность Земли и вода, отображаются в темно-серых или черных тонах, в то время как более холодные облака имеют более светлые оттенки. Температура облаков может служить индикатором их высоты: более холодные облака обычно находятся на большей высоте и могут быть связаны с сильными грозовыми явлениями. Снимки водяного пара демонстрируют количество влаги в верхних слоях атмосферы (примерно от 4,5 до 9 км). Области с высокой влажностью отображаются в виде белых областей, в то время как сухие области

имеют темный оттенок. Эти снимки полезны для определения потенциальных зон сильных осадков. Комбинация различных типов спутниковых изображений позволяет метеорологам получить полное представление об атмосферных условиях.

Еще одним важным источником геопространственных данных являются данные с метеорологических станций. Метеорологические станции, расположенные на суше, на морских буях и кораблях, осуществляют сбор прямых измерений таких параметров, как температура, осадки, скорость и направление ветра, атмосферное давление и влажность. Эти данные являются основой для мониторинга и прогнозирования погодных условий, а также для изучения долгосрочных климатических тенденций. Метеорологические станции также предоставляют метаданные, включая географические координаты (широта, долгота) и высоту над уровнем моря. Информация с метеорологических станций представляется в виде станционных метеорологических сводок, содержащих данные о текущих погодных условиях, облачности, видимости, температуре, точке росы, атмосферном давлении и его изменениях за последние три часа. Однако плотность размещения метеорологических станций в Арктике значительно уступает другим регионам мира, что делает спутниковые данные и результаты моделирования особенно ценными.

Радарные данные играют ключевую роль в метеорологическом обеспечении. Метеорологические радары испускают электромагнитные волны и анализируют отраженные сигналы для идентификации характеристик атмосферных осадков, таких как их расположение, интенсивность, тип (дождь, снег, град) и направление движения. Современные метеорологические радары, преимущественно импульсно-доплеровские, способны не только определять интенсивность осадков, но и выявлять скорость движения дождевых капель. Эти данные используются для анализа структуры штормов и оценки их потенциального воздействия на формирование опасных метеорологических явлений. Радарные системы особенно эффективны для раннего выявления и мониторинга интенсивных погодных событий. Однако наземные радары имеют

ограниченную эффективность в Арктическом регионе из-за кривизны земной поверхности и специфических атмосферных условий, что делает спутниковые радары более предпочтительным источником данных для мониторинга осадков в этом регионе.

Данные моделирования представляют собой результаты численных метеорологических моделей, которые прогнозируют будущие состояния атмосферы и создают согласованные исторические наборы данных (реанализы). Моделирование основывается на численном решении системы уравнений, описывающих физические процессы в атмосфере, с использованием данных наблюдений с метеорологических станций, спутников, радаров и других источников для инициализации моделей. Примером такого реанализа является ERA-40, разработанный упоминаемым ранее Европейским центром среднесрочных прогнозов погоды (ECMWF). ERA-40 представляет собой глобальный набор данных, охватывающий атмосферные и поверхностные условия за период с сентября 1957 года по август 2002 года. Этот набор данных включает различные параметры атмосферы, поверхности и океана. Несмотря на то, что ERA-40 был заменен более новыми наборами данных, такими как ERA-Interim и ERA5, он продолжает оставаться важным ресурсом для климатических исследований и анализа атмосферных процессов в Арктике. В рамках данной работы именно ERA-40 использовался для тестирования разрабатываемых модулей.

Геопространственные данные в метеорологии и океанографии хранятся и распространяются в различных форматах. NetCDF (.nc) является широко используемым форматом для хранения многомерных научных данных, особенно сеточных данных, таких как спутниковые снимки и результаты моделирования. GRIB (Gridded Binary) представляет собой компактный и эффективный формат для обмена и хранения метеорологических данных, часто применяемый метеорологическими организациями. GeoTIFF (.tif, .tiff) является стандартным форматом для хранения геопривязанных растровых изображений, включая спутниковые и аэрофотоснимки. Shapefile (.shp, .shx, .dbf и другие) представляет

собой популярный формат для хранения векторных данных, используемый для представления геометрии географических объектов (точек, линий, полигонов) и их атрибутивной информации.

Разнообразие доступных типов геопространственных данных обеспечивает комплексное представление об арктической среде. Интеграция данных со спутников, наземных станций, радаров и моделей является критически важной для всестороннего понимания этого региона, поскольку каждый источник информации обладает своими уникальными преимуществами и ограничениями. Именно по этой причине, в рамках этой работы были использованы данные реанализа, как наиболее полного источника [11, 12].

1.3.2 Проблемы и вызовы при работе с гидрометеорологическими геопространственными данными в Арктике

Работа с гидрометеорологическими геопространственными данными в арктическом регионе связана с рядом специфических проблем и вызовов, обусловленных экстремальными климатическими условиями, ограниченностью сети наблюдений и особенностями распространения радиосигналов.

Экстремальные климатические условия Арктики, включая крайне низкие температуры, продолжительную полярную ночь, ледяной и снежный покров, а также удалённость и труднодоступность многих территорий, создают значительные логистические и технические препятствия для сбора данных и обслуживания необходимой инфраструктуры. Эти факторы могут приводить к перебоям в процессе сбора данных, снижению точности проводимых измерений и увеличению затрат на осуществление наблюдений и обработку полученной информации. Крайне важно использовать надёжные измерительные приборы и методы передачи данных, способные функционировать в условиях экстремальных нагрузок.

Ограниченность сети наземных метеорологических станций и океанографических буёв в арктическом регионе представляет собой ещё одну

серьёзную проблему. Значительные размеры региона и труднодоступность его отдельных частей приводят к тому, что плотность размещения наблюдательных пунктов здесь значительно ниже, чем в других регионах мира. Это обстоятельство увеличивает зависимость от данных, получаемых посредством дистанционного зондирования с использованием спутниковых технологий. Однако ограниченные возможности проведения прямых измерений на местах затрудняют валидацию данных дистанционного зондирования и метеорологических моделей, используемых для анализа и прогнозирования процессов в арктической зоне.

Особенности распространения сигналов в арктической среде создают значительные трудности при работе с геопространственными данными. Наличие морского льда, снежного покрова и специфические атмосферные условия оказывают влияние на распространение электромагнитных волн, используемых в технологиях дистанционного зондирования, таких как радарные системы и спутниковая связь. Это может приводить к искажениям данных и снижению их точности.

Кроме того, существуют проблемы, связанные с доступом к данным и их совместимостью. Различные национальные и международные организации осуществляют сбор данных об Арктике, которые могут различаться по форматам, стандартам метаданных и политикам доступа. Поиск, получение доступа и интеграция этих разрозненных данных представляют собой значительную сложность. Для эффективного использования гидрометеорологической информации в арктическом регионе необходимо развивать международное сотрудничество, способствовать обмену данными и внедрять стандарты, обеспечивающие их совместимость. Инициативы, такие как Арктическая инфраструктура пространственных данных (Arctic SDI), направлены на решение этих проблем путем продвижения обмена данными и стандартизации.

В условиях отсутствия унифицированной системы стандартов для гидрометеорологических данных и фрагментации информационных ресурсов, представленная в данной работе интерактивная автоматизированная система

приобретает особую значимость. Она обеспечивает пользователям оперативный доступ к обширным архивам геопространственных данных, таким как ERA-40, в визуализированном формате, оптимизированном для восприятия. Это позволяет исключить необходимость самостоятельного поиска информации в различных источниках, ее последующей обработки и анализа с целью получения практически значимых результатов. Данное преимущество особенно актуально в контексте Арктического региона, где вопросы доступа к данным и их интеграции представляют собой значительные вызовы [13, 14, 15].

1.4 Обзор существующих технологий и программных средств для обработки, анализа и визуализации геопространственных данных

Географические информационные системы (ГИС) являются фундаментальной основой геопространственных технологий, представляя собой программный инструментарий для сбора, хранения, обработки, анализа и представления различных типов пространственных и географических данных. Ключевые компоненты ГИС включают:

- Управление данными: обеспечение хранения, организации и поиска пространственных и атрибутивных данных.
- Картографирование и визуализация: создание карт и других графических представлений геопространственных данных.
- Пространственный анализ: выполнение операций для извлечения информации и выявления закономерностей из пространственных данных.
- Коммуникация: эффективное распространение и представление геопространственной информации.

ГИС выступает в качестве центральной платформы для интеграции и работы с разнообразными типами геопространственных данных, обеспечивая комплексный анализ и поддержку принятия обоснованных решений. Без использования ГИС было бы крайне затруднительно эффективно интегрировать,

анализировать и передавать сложную информацию, содержащуюся в пространственных наборах данных.

Пространственный анализ представляет собой систематическое исследование пространственного распределения различных явлений, выявление характерных закономерностей и прогнозирование на базе геопространственных данных, включающих как местоположение, так и атрибутивные характеристики. Методы пространственного анализа включают: анализ близости, направленный на идентификацию объектов, находящихся на определенном расстоянии друг от друга; моделирование пригодности, используемое для определения оптимальных местоположений на основании установленных критериев; пространственную кластеризацию, предназначенную для группировки объектов, схожих по своим пространственным характеристикам; анализ пространственной автокорреляции, оценивающий взаимосвязи между значениями одного и того же атрибута в смежных областях; сетевой анализ, анализирующий связи и потоки в сетях, таких как транспортные или коммунально-бытовые системы.

Пространственный анализ преобразует неструктурированные геопространственные данные в практически применимую информацию, выявляя скрытые закономерности, взаимосвязи и тенденции. Применение аналитических методов к пространственному компоненту данных позволяет глубже понять процессы и взаимодействия, происходящие в реальном мире, и способствует принятию более обоснованных решений.

Картография представляет собой искусство и науку создания карт, охватывающие принципы визуальной коммуникации и разработку эффективных и информативных картографических продуктов на основе геопространственных данных. Принципы картографического дизайна включают учет масштаба, выбор картографических проекций, использование символики, цветовой гаммы и композиционного решения. Картографы применяют геопространственные данные для создания визуально привлекательных и информативных карт,

способствующих лучшему пониманию пространственных закономерностей и взаимосвязей.

Картография играет ключевую роль в эффективной передаче информации, полученной в результате пространственного анализа, широкой аудитории. Она обеспечивает представление сложной пространственной информации в доступной и понятной форме. Карты являются основным визуальным результатом пространственного анализа, и принципы картографии определяют создание картографических продуктов, которые не только эстетически привлекательны, но и точно передают заложенную информацию.

Первый этап обработки геопространственных данных включает сбор первичных материалов, полученных непосредственно в процессе геодезических измерений, GPS-съемки, аэрофотосъемки, беспилотной или спутниковой съемки. В процессе используются разнообразные датчики, включая оптические, LiDAR и SAR. Также осуществляется сбор вторичных данных, извлекаемых из существующих геопространственных ресурсов, географических и наземных информационных систем или других источников.

Геопространственные данные могут быть представлены в различных форматах, таких как векторные данные (точки, линии, полигоны, отображающие дискретные объекты, например, дороги, реки, здания), растровые данные (пикселизированные или сеточные структуры, представляющие непрерывные поля, такие как температура, высота, спутниковые изображения) и облака точек (наборы трехмерных координат с потенциальной информацией RGB, часто полученные с помощью LiDAR).

Выбор соответствующих источников данных и анализ их характеристик, таких как точность, разрешение и формат, имеют решающее значение для последующих этапов обработки и анализа. Основой любого геопространственного проекта является корректное получение необходимых данных, что включает определение требуемой информации, выбор оптимальных источников и понимание свойств собираемых данных.

Следующий этап — предварительная обработка собранных геопространственных данных. Этот процесс включает геометрическую коррекцию (привязку изображений к их правильному географическому положению и выравнивание различных слоев данных), радиометрическую коррекцию (настройку значений пикселей для представления физических величин, таких как отражательная способность), атмосферную коррекцию (удаление атмосферных помех из изображений дистанционного зондирования), шумоподавление (устранение артефактов датчиков или нежелательных вариаций), ортотрансформирование (коррекцию смещения рельефа на изображениях), паншарпенинг (повышение пространственного разрешения мультиспектральных изображений) и объединение каналов (интеграцию нескольких спектральных слоев изображения) [16].

Интеграция данных включает объединение информации из различных источников в единый формат. Тщательная предварительная обработка является необходимым условием для обеспечения точности и надежности результатов анализа, так как позволяет устранить искажения и несоответствия в необработанных данных. Перед проведением значимого анализа геопространственные данные часто подвергаются ряду этапов предварительной обработки для исправления ошибок и приведения данных в формат, соответствующий предполагаемым аналитическим методам.

Организация и управление геопространственными данными в структурированном виде включают использование ГИС, которые могут быть реализованы в виде отдельных файлов или реляционных баз данных с пространственной поддержкой. Основные методы включают индексирование баз данных (для повышения скорости поиска данных), сжатие данных (для уменьшения объема хранения) и использование геоинформационных баз данных (для организации и управления географической информацией) [17].

Также важны эффективные протоколы и платформы для обмена данными, такие как ArcGIS Server, GeoServer и Mapbox. Эффективное управление данными является критически важным для обработки больших объемов

геопространственной информации и обеспечения их доступности, хорошей организации и защиты от утраты или повреждения. В условиях растущих объемов геопространственных данных разработка надежных стратегий и систем управления данными становится все более значимой для обеспечения эффективности рабочих процессов и долгосрочного использования данных.

Применение различных методов для извлечения значимой информации из геопространственных данных включает анализ векторной информации, который охватывает геометрическую обработку, такую как буферизация для формирования зон вокруг объектов, и оверлейный анализ для интеграции пространственных наборов данных через операции объединения, пересечения и разности. Анализ близости определяет расстояния и взаимосвязи между объектами, а сетевой анализ оценивает связность и потоки в сетях, таких как транспортные или гидрографические системы.

Анализ растровых данных включает выполнение математических операций над растровыми слоями с использованием алгебры карт, пространственную фильтрацию, классификацию изображений на основе спектральных характеристик, обнаружение изменений между различными временными периодами, спектральный анализ для изучения спектральных сигнатур объектов и анализ временных рядов для оценки изменений растровых данных во времени.

Анализ рельефа основывается на цифровых моделях рельефа (ЦМР, или DEM) для определения характеристик поверхности, таких как уклон, градиент, экспозиция и кривизна. Методы анализа видимости идентифицируют области, видимые из заданной точки. Геостатистический анализ применяет статистические методы для анализа пространственно привязанных данных, включая пространственную интерполяцию (например, кригинг для прогнозирования значений в ненаблюдаемых местоположениях), анализ пространственной автокорреляции для измерения степени кластеризации или дисперсии объектов или значений в пространстве и анализ точечных паттернов для изучения пространственного распределения точечных объектов.

Широкий спектр методов геопространственного анализа позволяет исследователям и практикам решать большой круг задач и извлекать ценную информацию из пространственных данных, адаптированную к конкретным потребностям и целям. Выбор аналитического метода определяется исследовательским вопросом, типом доступных геопространственных данных и ожидаемыми результатами. Понимание принципов, лежащих в основе каждого метода, критически важно для их правильного применения и интерпретации результатов.

Интеграция геопространственных данных с негеопространственными наборами данных, такими как данные переписи населения, метеорологические данные, данные социальных сетей или чертежи (САПР), способствует более глубокому пониманию явлений. Эта интеграция выявляет взаимосвязи между переменными и позволяет обнаружить закономерности и тенденции, которые могут оставаться незамеченными при анализе данных в отдельности. Методы интеграции включают пространственные соединения для объединения атрибутов на основе пространственных связей, атрибутивные соединения для объединения данных на основе общих атрибутов и наложение различных наборов данных в ГИС.

Интеграция геопространственных данных с релевантной информацией значительно увеличивает аналитическую мощь и предоставляет более широкий контекст для принятия решений. Реальные явления часто зависят от комбинации пространственных и непространственных факторов. Объединение различных типов данных позволяет сформировать более глубокое и детализированное понимание окружающего мира.

ГИС Esri ArcGIS, включающая продукты ArcGIS Pro, ArcGIS Online и ArcGIS Enterprise, является ведущим мировым программным обеспечением в области геоинформатики. Она представляет собой комплексную корпоративную геопространственную платформу с полным спектром функций для SaaS-картографирования. Ключевые функциональные возможности включают

продвинутые инструменты пространственного анализа, такие как обширные инструменты геообработки и специализированные наборы инструментов, а также впечатляющие возможности визуализации в 2D, 3D и 4D. ArcGIS обеспечивает бесшовную интеграцию данных из различных источников, предлагает комплексные инструменты управления данными и широкий спектр расширений (таких как ArcGIS Spatial Analyst, ArcGIS Network Analyst и ArcGIS 3D Analyst) для решения специализированных задач.

ArcGIS является отраслевым стандартом и широко используется профессионалами и организациями в различных секторах благодаря своей надежности и обширной экосистеме. Однако его проприетарный характер и высокая стоимость могут быть ограничивающими факторами для индивидуальных пользователей и небольших организаций. Доминирование ArcGIS на рынке обеспечивает широкий спектр инструментов и интеграционных возможностей, что делает его оптимальным решением для сложных проектов и крупных организаций.

Global Mapper часто описывается как универсальное и простое в использовании решение для геопространственного анализа. Он предоставляет комплексный набор инструментов для обработки пространственных данных и поддерживает широкий спектр форматов данных. Ключевые функции включают обширную поддержку более чем 300 форматов файлов, надежное отображение и обработку данных LiDAR, расширенные инструменты для работы с высотными данными и облаками точек, удобный пользовательский интерфейс и поддержку скриптов для автоматизации рабочих процессов. Global Mapper выделяется высокой совместимостью данных и мощными возможностями обработки высотных данных, часто по более доступной цене в сравнении с ArcGIS. Однако его картографический и веб-картографический функционал может быть менее обширным по сравнению с ArcGIS. Преимущество Global Mapper заключается в способности эффективно обрабатывать разнообразные форматы данных, что делает его ценным инструментом для пользователей, ориентированных на

преобразование и анализ данных, а не на создание расширенных картографических продуктов.

ERDAS IMAGINE представляет собой комплексное решение для обработки и анализа данных дистанционного зондирования Земли (ДЗЗ). Он объединяет функциональные возможности классификации изображений, фотограмметрии, анализа данных LiDAR и обработки радиолокационных данных в единую платформу. Ключевые функции включают расширенные инструменты для извлечения значимой информации из данных ДЗЗ, инструменты для сегментации изображений, ортотрансформирования, мозаичного объединения и извлечения высотных данных, а также интеграцию технологий искусственного интеллекта для задач, таких как обнаружение объектов и семантическая сегментация. ERDAS IMAGINE является мощным инструментом, разработанным для профессионалов, работающих с растровыми данными и изображениями ДЗЗ, предлагая расширенные возможности обработки изображений. Для пользователей, основной задачей которых является анализ спутниковых, авиационных или дроновых изображений, ERDAS IMAGINE предоставляет специализированную среду с полным набором инструментов для извлечения ценной информации.

ENVI представляет собой признанный стандарт в области программного обеспечения для обработки и анализа изображений. Оно предназначено для извлечения актуальной, достоверной и точной информации из широкого спектра геопространственных данных, включая гиперспектральные, мультиспектральные, инфракрасные, LiDAR и SAR. Программное обеспечение включает в себя интуитивно понятные инструменты визуализации, обработки и анализа данных, а также автоматизированные рабочие процессы для обнаружения аномалий и изменений, классификации объектов и интеграции с ArcGIS. ENVI также предоставляет модуль глубокого обучения для расширенного анализа изображений. Это программное обеспечение широко

используется в сфере ДЗЗ благодаря своей научно обоснованной аналитике и удобному пользовательскому интерфейсу.

QGIS (Quantum GIS) является одной из наиболее популярных бесплатных и открытых геоинформационных систем, предлагающей широкий спектр функций для решения различных задач в области ГИС. Это программное обеспечение с открытым исходным кодом поддерживает визуализацию данных, редактирование, анализ и создание картографических композиций. QGIS отличается высокой степенью расширяемости благодаря обширному набору плагинов (более 2000), активному сообществу разработчиков и регулярной поддержке. Оно обеспечивает кроссплатформенную совместимость (Windows, macOS, Linux, Android) и поддерживает множество форматов данных, включая векторные, растровые, сетчатые и базы данных. QGIS предоставляет мощные инструменты для пространственного анализа, геообработки и картографического проектирования, что делает его надёжной и универсальной альтернативой коммерческому программному обеспечению.

GRASS GIS (Geographic Resources Analysis Support System) представляет собой мощную геоинформационную систему, как и QGIS, с открытым исходным кодом, предназначенную для управления и анализа геопространственных данных, обработки изображений и пространственного моделирования. Программное обеспечение известно своими передовыми аналитическими возможностями, включая расширенный геостатистический анализ, надёжную обработку растровых и векторных данных, поддержку временных рядов и обширный набор инструментов геообработки (более 350). GRASS GIS способен обрабатывать большие объёмы данных и может функционировать как автономное приложение или интегрироваться с другими программными продуктами, такими как QGIS. Эта система особенно эффективна в научных и экологических исследованиях благодаря своим мощным инструментам для растрового анализа и моделирования рельефа. Однако её интерфейс может быть

менее интуитивным по сравнению с QGIS, что требует более долгого обучения. GRASS GIS предоставляет широкий спектр аналитических возможностей, что делает её ценным инструментом для исследователей и специалистов в области геоинформатики [18].

GeoDa представляет собой ещё одно программное обеспечение с открытым исходным кодом и бесплатным доступом. Оно предназначено для разведочного анализа пространственных данных (ESDA). GeoDa эффективно выявляет пространственные закономерности и выполняет статистический анализ. Основные функции включают специализированные инструменты и методы пространственной статистики, такие как локальные индикаторы пространственной ассоциации (LISA), которые позволяют идентифицировать статистически значимые горячие и холодные точки. GeoDa поддерживает широкий спектр векторных форматов данных, включая GeoJSON, shapefile, MapInfo, GML, geodatabase, KML и другие форматы, поддерживаемые библиотекой GDAL.

GeoDa является ценным инструментом для исследователей и аналитиков, занимающихся изучением и пониманием пространственного распределения данных, а также выявлением статистически значимых пространственных закономерностей. Его ориентация на ESDA делает его особенно полезным в таких областях, как здравоохранение и экологическое исследование. Основные преимущества GeoDa заключаются в его специализированных инструментах для пространственной статистики, которые позволяют пользователям проводить строгий анализ пространственных взаимосвязей в своих данных и выявлять значимые закономерности.

Google Earth Engine (GEE) представляет собой облачную платформу для крупномасштабного анализа геопространственных данных. GEE предоставляет доступ к обширному архиву спутниковых изображений и геопространственных наборов данных. Основные функции включают каталог данных петабайтного

масштаба, мощную вычислительную инфраструктуру и API на Python и JavaScript для эффективной разработки рабочих процессов. GEE идеально подходит для исследователей и организаций, которым необходимо обрабатывать и анализировать огромные наборы геопространственных данных в облаке. Облачный характер платформы и обширный репозиторий данных делают её мощным инструментом для крупномасштабных исследований окружающей среды и землепользования.

CARTO представляет собой облачную платформу для определения местоположения, предназначенную для пространственного анализа, визуализации и разработки приложений. Она интегрируется с облачными хранилищами данных, такими как BigQuery, Snowflake, Databricks и Redshift. Основные функции включают мощную аналитику с инструментарием визуального проектирования рабочих процессов, выдающиеся возможности визуализации данных, встроенные возможности генеративного искусственного интеллекта (GenAI) и доступ к тысячам глобальных наборов данных через Data Observatory. CARTO предоставляет комплексное облачное решение для организаций, стремящихся интегрировать информацию о местоположении в свои существующие экосистемы данных. Интеграция пространственного анализа непосредственно в облачные хранилища данных обеспечивает масштабируемость и безопасность для геопространственных приложений корпоративного уровня.

Существует широкий спектр методов визуализации геопространственных данных, каждый из которых обладает своими преимуществами и недостатками, зависящими от специфики данных и цели представления информации. Точечные карты применяют точечные символы для обозначения данных в конкретных географических локациях, эффективно демонстрируя распределение и плотность объектов. Карты пропорциональных символов варьируют размер и/или окраску символов (например, круговых маркеров) для представления различных

значений или нескольких переменных в определенных местах. Кластерные карты, являющиеся разновидностью карт пропорциональных символов, агрегируют близлежащие точки в кластеры, что особенно полезно при работе с большими массивами данных.

Картограммы используют цветовую градацию для заполнения определенных географических или административных единиц, отображая значения данных в этих областях. Важно отметить, что картограммы могут искажать форму и размеры географических объектов, чтобы они соответствовали пропорциям значений конкретной переменной. Тепловые карты применяют цветовую шкалу для визуализации плотности или интенсивности распределения точек данных на географической плоскости. Карты плотности используют контурные линии для демонстрации распределения и плотности точечных данных.

Трехмерные модели предоставляют возможность визуализации геопространственных данных в трехмерной перспективе, что особенно полезно для анализа рельефа, архитектурных объектов и других пространственных элементов. Статические карты представляют собой фиксированное визуальное отображение данных, в то время как динамические карты позволяют пользователям интерактивно исследовать данные и выполнять различные запросы. Информационные панели предоставляют комплексное представление нескольких визуализаций и ключевых метрик в едином интерфейсе.

Эффективная визуализация играет ключевую роль в обеспечении понятного и доступного представления пространственных закономерностей, тенденций и взаимосвязей. Выбор подходящего метода визуализации является критически важным для достижения цели передачи информации. Различные типы геопространственных данных и аналитические задачи требуют применения соответствующих методов визуализации. Понимание сильных и слабых сторон каждого метода визуализации является необходимым условием для создания карт и моделей, точно и ясно передающих требуемую информацию.

Для создания геопространственных визуализаций существует широкий спектр программных инструментов. ГИС, такие как ArcGIS, QGIS и Global Mapper, предоставляют надежные возможности для создания различных типов карт и трехмерных моделей. Например, ArcGIS Pro предоставляет широкий набор инструментов и возможностей для создания высококачественных визуализаций в двухмерном, трехмерном и четырехмерном пространствах. QGIS предлагает исключительные функции для картографического дизайна, позволяя создавать выдающиеся карты для настольных компьютеров, мобильных устройств, облачных сред и печати. Global Mapper обеспечивает гибкую визуализацию двумерных и трехмерных данных с возможностью настройки шейдеров.

Веб-платформы для картографирования, такие как ArcGIS Online, Google Maps Platform, Mapbox, CARTO, kepler.gl и Atlas.co, разработаны специально для создания и публикации интерактивных веб-карт. ArcGIS Online позволяет пользователям создавать интерактивные веб-карты с интеллектуальными стилями картографирования, управляемыми данными. Google Maps Platform предоставляет возможность визуализации геопространственных и бизнес-данных на настраиваемых картах Google. Mapbox предлагает инструменты для создания пользовательских карт и визуализации данных для веб- и мобильных приложений. CARTO — это облачная платформа для создания насыщенных и динамических визуализаций данных, работающих с облачными хранилищами данных. kepler.gl представляет собой высокопроизводительное веб-приложение для визуального исследования крупномасштабных геоданных. Atlas.co — это современная браузерная ГИС-платформа с функциями визуализации данных и создания интерактивных карт для совместного использования.

Инструменты визуализации данных с геопространственными возможностями, такие как Tableau и Microsoft Power BI, позволяют интегрировать и визуализировать геопространственные данные вместе с другими типами данных, создавая информативные панели и отчеты. Tableau фокусируется на взаимосвязи между данными и их физическим расположением, позволяя

создавать аналитические выводы с помощью различных геопространственных визуализаций, таких как карты пропорциональных символов и картограммы.

Программные библиотеки, такие как Leaflet, OpenLayers и пространственные пакеты R, предоставляют разработчикам гибкость для создания высоконастраиваемых и интерактивных геопространственных визуализаций. Leaflet — это JavaScript-библиотека для создания интерактивных веб-карт, оптимизированных для мобильных устройств. OpenLayers — это еще одна бесплатная JavaScript-библиотека с открытым исходным кодом для создания динамических карт. R, с такими пакетами, как ggplot2, leaflet и tmap, предлагает мощные инструменты для создания статических и интерактивных карт.

Специализированные инструменты визуализации, такие как EO Browser, NASA Worldview и GloVis, предназначены для визуализации конкретных типов геопространственных данных, особенно спутниковых изображений. EO Browser позволяет пользователям быстро визуализировать спутниковые данные со множества спутников и коллекций данных, предоставляя возможности настройки визуализации и комбинаций каналов. NASA Worldview предоставляет интерактивный доступ к более чем 900 глобальным спутниковым изображениям с полным разрешением. GloVis (USGS Global Visualization Viewer) — это простая в использовании платформа для просмотра и загрузки данных ДЗЗ [19].

Среди множества инструментов и библиотек для анализа и визуализации геопространственных данных особое место занимает язык программирования Python. Его популярность обусловлена рядом существенных преимуществ, включая обширную коллекцию библиотек, специально разработанных для работы с геопространственными данными. Библиотека GeoPandas предоставляет мощные инструменты для обработки и анализа векторных данных, включая чтение и запись различных форматов файлов, а также выполнение сложных пространственных операций. Для работы с растровыми данными используется библиотека Rasterio, обеспечивающая эффективное управление и анализ данных в пиксельной форме.

Для визуализации геопространственных данных Python предлагает несколько решений. Библиотека Matplotlib предоставляет базовые возможности для создания статических карт и графиков, в то время как Plotly позволяет генерировать интерактивные и динамические визуализации, которые могут быть интегрированы в веб-приложения.

Python является языком с открытым исходным кодом, что способствует активному развитию библиотек и наличию обширной документации и системы поддержки. Его гибкость позволяет интегрировать язык с ГИС, такими как QGIS и ArcGIS, расширяя возможности анализа и визуализации данных.

Благодаря этим преимуществам и богатой экосистеме библиотек Python стал предпочтительным инструментом для специалистов в области геопространственного анализа, обработки и визуализации данных. Именно этот язык был выбран для разработки модуля обработки и анализа геопространственной информации в рамках данного исследования.

2. Методология проектирования и выбора технологических решений для модуля

2.1 Постановка задачи разработки модуля обработки и анализа геопространственных данных

Разработка интерактивной автоматизированной системы (ИАС) для долгосрочного обеспечения гидрометеорологической информацией, необходимой для морской деятельности в Арктическом регионе, требует создания специализированного модуля, предназначенного для работы с геопространственными данными. Данный модуль, названный в рамках этой работы модулем обработки и анализа геопространственных данных, является ключевым компонентом системы, ответственным за извлечение, преобразование

и визуализацию информации в удобной для пользователя форме. Настоящий раздел посвящен формулированию функциональных требований к этому модулю и его интеграции в общую архитектуру ИАС.

2.1.1 Определение функциональных требований к модулю

Модуль обработки и анализа геопространственных данных предназначен для предоставления пользователям ИАС актуальной и исторической информации о гидрометеорологических условиях в заданной точке арктического региона. На основе анализа задач, которые необходимо решить разрабатываемыми алгоритмами, можно выделить следующие ключевые функциональные требования.

Прием входных параметров от основной системы:

- Географические координаты: Модуль должен принимать от ИАС точные географические координаты (широту и долготу), указанные пользователем на интерактивной карте.
- Тип запрашиваемой информации: Модуль должен быть способен обрабатывать запросы на конкретные гидрометеорологические параметры. В рамках текущей версии основным параметром является температура воздуха. В перспективе планируется расширение функциональности для обработки других параметров, таких как скорость и направление ветра, толщина льда, высота волн и т.д.
- Временной диапазон: Модуль должен принимать от ИАС спецификации временного периода, за который пользователь желает получить данные. Это может включать:
 - Запрос на отображение фактических данных за определенный исторический период (например, с 01-01-1970 по 31-12-1990).

- Запрос на отображение средних многолетних значений для конкретного диапазона дат в году (например, с 01-05 по 01-10 для всех доступных лет).

Доступ и загрузка исходных данных:

- Модуль должен обеспечивать загрузку гидрометеорологических данных из определенных источников. В текущей реализации предусмотрено использование архивов данных в формате CSV, содержащих информацию о температуре воздуха за многолетний период (1957-2002 гг.).

- Необходимо предусмотреть возможность объединения данных из нескольких файлов (например, по годам) в единый набор данных для последующего анализа.

Предварительная обработка данных:

- Фильтрация по координатам: Модуль должен осуществлять отбор данных, соответствующих указанным пользователем географическим координатам (широте и долготе).

- Конвертация единиц измерения: Данные должны быть преобразованы к стандартизированным и понятным для пользователя единицам измерения (например, перевод температуры из Кельвинов в градусы Цельсия).

- Обработка формата дат: Даты должны быть приведены к единому формату, необходимому для анализа и визуализации, включая локализацию (например, перевод названий месяцев на русский язык).

Анализ данных:

- Агрегация данных: Модуль должен обеспечивать вычисление статистических показателей по временным рядам. Основной функцией является расчет средних многолетних значений параметра (температуры воздуха) для каждой календарной даты на основе исторических данных, указанных пользователем.

- Фильтрация по временному диапазону: Модуль должен осуществлять отбор данных, соответствующих заданному пользователем

временному интервалу (как для фактических данных, так и для средних многолетних).

Визуализация данных:

- Модуль должен генерировать графические представления обработанных данных. Основным типом визуализации является линейный график, отображающий динамику изменения температуры воздуха во времени (либо фактические значения, либо средние многолетние) для выбранной точки и периода.

- Оформление графика должно включать информативные подписи осей, заголовок с указанием координат, временного периода и типа отображаемых данных.

Вывод результатов:

- Модуль должен возвращать сгенерированный график (или данные для его построения) основной информационно-аналитической системе для последующего отображения пользователю через веб-интерфейс.

2.1.2 Место модуля в архитектуре общей интерактивной автоматизированной системы

Модуль обработки и анализа геопространственных данных, разрабатываемый в рамках проекта, является ключевым компонентом серверной части ИАС. Его функционирование тесно интегрировано с другими элементами системы, включая пользовательский интерфейс (фронтенд) и, возможно, базу данных или хранилище исходных геопространственных данных.

Схема взаимодействия модуля с ИАС включает следующие этапы:

Инициация запроса посредством пользовательского интерфейса (фронтенд):

- Пользователь взаимодействует с веб-сайтом ИАС, содержащим интерактивную карту Арктического региона.

- При клике на конкретную точку карты пользователь выбирает географические координаты (широта и долгота).

- Через элементы управления на сайте (например, выпадающие списки, календари) пользователь указывает тип гидрометеорологического параметра (в данном случае — температура воздуха) и выбирает один из двух режимов отображения:

- Отображение фактических температур воздуха за конкретный исторический период (например, с 01 января 1980 года по 31 декабря 1980 года).

- Отображение графика средних многолетних температур воздуха для определённого сезонного периода (например, с 01 июня по 31 августа, осреднённое по данным за 1960–2000 годы).

- Сформированный запрос, содержащий выбранные координаты, тип параметра и временные рамки, передаётся с фронтенда на серверную часть ИАС.

Обработка запроса и вызов модуля на серверной стороне:

- Серверная часть ИАС принимает запрос от пользователя.

- На основании анализа запроса серверная часть определяет необходимость использования модуля обработки и анализа геопространственных данных.

- Полученные от пользователя параметры (координаты, тип данных, выбранный период и режим отображения) передаются в качестве входных данных соответствующему скрипту модуля.

Работа модуля (исполнение скриптов):

- Модуль, получив входные данные, выполняет следующие операции:

- Загрузка необходимых файлов с историческими данными о температуре воздуха.

- Фильтрация данных по указанным координатам.

- В зависимости от выбранного пользователем режима отображения:

- Для фактических данных: отбор данных за указанный исторический период, преобразование дат и подготовка к визуализации.

- Для средних многолетних данных: агрегация данных по всем годам для каждой календарной даты, фильтрация по выбранному сезонному периоду, преобразование дат и подготовка к визуализации.

- После выполнения необходимых операций модуль генерирует линейный график, который отображает динамику температуры воздуха.

Возврат результата и отображение пользователю:

- Сгенерированный модулем график (например, в виде HTML-кода для встраивания или изображения) возвращается на серверную часть ИАС.

- Серверная часть передаёт полученный результат на фронтенд.

- Веб-сайт ИАС отображает сгенерированный график пользователю в специально отведённой области интерфейса, предоставляя визуализированную гидрометеорологическую информацию по выбранной точке и периоду.

2.2 Обоснование выбора инструментальных средств и технологий разработки

2.2.1 Сравнительный анализ подходов и программных средств для обработки и визуализации геопространственных данных

В предыдущей главе был представлен обзор существующих технологий и программных средств для обработки, анализа и визуализации геопространственных данных, с фокусом на детальное рассмотрение различных ГИС. Настоящий раздел посвящен сравнительному анализу альтернативных языков программирования, пригодных для решения аналогичных задач без использования традиционного ГИС-интерфейса.

Цель данного анализа — сравнить Python с другими языками программирования с точки зрения их функциональности, эффективности и применимости для обработки и визуализации геопространственных данных. Сравнение проведено по таким критериям, как простота использования,

вычислительная производительность, доступность специализированных программных библиотек, уровень поддержки со стороны сообщества разработчиков, а также стоимость (в случае коммерческих решений). Кроме того, рассмотрена пригодность каждого из анализируемых языков для выполнения различных типов геопространственных задач, включая обработку как векторных, так и растровых данных, создание интерактивных картографических приложений и проведение статистического анализа.

Многообразие языков программирования, упомянутых в различных источниках, свидетельствует о наличии широкого спектра инструментов для геопространственного анализа. Несмотря на лидирующие позиции Python в данной области, другие языки программирования обладают уникальными характеристиками и функциональными возможностями, которые было необходимо учесть при выборе языка для написания разрабатываемого модуля.

R

Язык программирования R был разработан как инструмент для статистических вычислений и визуального анализа данных. Он обладает широким спектром возможностей для статистического анализа и находит применение в области геопространственных данных. В R реализованы специализированные пакеты для работы с геопространственной информацией, включая `sf` (Simple Features) для обработки векторных данных, `raster` для анализа растровых данных, представленных в виде сеток, а также `ggplot2` для создания сложных визуальных представлений данных на основе грамматики графики. Дополнительно, пакет `leaflet` позволяет разрабатывать интерактивные веб-карты, а пакет `tmap` предоставляет инструменты для статического и интерактивного тематического картографирования.

Сравнительный анализ синтаксиса и уровня сложности изучения R и Python показывает, что синтаксис R, несмотря на свою выразительность и оптимизацию для статистических операций, может быть менее интуитивно понятным для пользователей без опыта в области статистики. В то время как

синтаксис Python отличается прямолинейностью и универсальностью, что делает его более доступным для широкого круга разработчиков.

В контексте производительности в геопространственных задачах наблюдается, что пакет `sf` в R может демонстрировать более высокую скорость выполнения определенных операций с векторными данными по сравнению с `GeoPandas` в Python. Однако Python показывает высокую эффективность при работе с растровыми данными благодаря библиотекам `NumPy` и `Rasterio`.

Сообщество пользователей R активно поддерживает развитие и использование языка для геопространственного анализа. Существует значительное количество пакетов, разработанных специально для работы с геопространственными данными. Одной из сильных сторон R является его способность создавать привлекательные и информативные карты, что делает его предпочтительным выбором в академической и исследовательской среде.

Таким образом, основное преимущество R заключается в его специализации на статистическом анализе и визуализации данных. Это делает его мощным инструментом для проектов, требующих глубокого статистического анализа и визуализации результатов. Однако, сложность синтаксиса и кривой обучения R может представлять определенные трудности для разработчиков, не имеющих опыта работы с языками программирования, ориентированными на статистический анализ. В таких случаях Python может быть более предпочтительным выбором благодаря своей универсальности и доступности [20].

MATLAB

Язык программирования MATLAB является признанным решением для выполнения численных вычислений, работы с матрицами и обработки изображений. Эти функциональные возможности делают его особенно полезным в области геопространственного анализа, включая растровый анализ и анализ данных ДЗЗ. MATLAB предоставляет специализированные наборы инструментов, такие как `Mapping Toolbox` и `Image Processing Toolbox`, которые

обеспечивают комплексную среду для импорта, экспорта, визуализации и анализа геопространственных данных в растровом и векторном форматах, включая анализ рельефа и геоморфологию.

Синтаксис MATLAB был разработан с акцентом на математические вычисления и тесно интегрирован с математической нотацией, что отличает его от более общего и, как правило, более многословного синтаксиса Python, используемого для аналогичных задач. Стоимость MATLAB является значительным фактором, так как это проприетарное программное обеспечение, требующее лицензирования, в отличие от бесплатного и открытого кода Python. Это может существенно повлиять на бюджет проекта, особенно такого небольшого, как модуль, реализованный в этой работе.

Анализ поддержки сообщества и экосистемы библиотек для MATLAB и Python показывает, что Python имеет значительно более обширное и разнообразное сообщество, особенно в быстро развивающихся областях, таких как наука о данных (Data science) и машинное обучение, которые все чаще применяются в геопространственном анализе.

Сравнение производительности MATLAB и Python в вычислительно интенсивных задачах геопространственного анализа показывает, что оптимизированный движок численных вычислений MATLAB может превосходить Python в некоторых матричных операциях. Однако Python, с библиотеками, такими как NumPy, которые часто поддерживаются языками программирования C или Fortran, способен достигать сопоставимой или даже более высокой производительности в различных сценариях.

Таким образом, MATLAB выделяется своей интерактивной средой и специализированными инструментами для численных вычислений и обработки изображений, что делает его привлекательным для проектов, ориентированных на эти виды геопространственного анализа. Однако высокая стоимость лицензирования и относительно небольшое сообщество могут быть существенными недостатками, ограничивающими доступность сторонних библиотек и решений, разработанных сообществом [21].

Julia

Язык программирования Julia представляет собой современное высокопроизводительное решение, сочетающее динамическую типизацию и функциональные возможности, характерные для языков высокого уровня, с эффективностью и скоростью, присущими низкоуровневым языкам. Эта уникальная комбинация делает Julia особенно подходящим для решения вычислительно интенсивных задач, включая геопространственный анализ.

Экосистема Julia для геопространственной обработки данных активно развивается, что подтверждается созданием и активным развитием таких библиотек, как JuliaEarth, GeoStats.jl и GeoMakie.jl. Эти инструменты предоставляют обширные возможности для геостатистического моделирования, анализа пространственных данных и визуализации, включая поддержку сферических координат.

Синтаксис Julia отличается интуитивностью и удобством для пользователей, имеющих опыт работы с MATLAB или R. Однако, несмотря на эти преимущества, основное применение языка пока сосредоточено в области научных и технических вычислений. Вопрос о его доступности для начинающих пользователей по сравнению с Python остается дискуссионным.

Анализ производительности Julia в геопространственных задачах демонстрирует, что язык может значительно превосходить Python в выполнении численных и массивно-ориентированных операций, особенно при обработке растровых данных и реализации специализированных алгоритмов. Это делает Julia привлекательным выбором для проектов, требующих высокой производительности и оптимизации.

Тем не менее, сообщество Julia, хотя и активно развивается, пока не достигло масштабов сообщества Python. Это отражается на разнообразии и зрелости геопространственных библиотек. В настоящее время спектр специализированных инструментов для геопространственного анализа в Julia ограничен по сравнению с Python, что может повлиять на скорость разработки и доступность функциональных возможностей.

Таким образом, Julia обладает значительным потенциалом для повышения эффективности геопространственной обработки данных, особенно в задачах, требующих высокой вычислительной мощности и оптимизации. Однако текущая стадия развития экосистемы Julia для геопространственного анализа предполагает необходимость тщательного планирования и адаптации к возможным ограничениям в доступе к специализированным инструментам [22].

JavaScript

Язык программирования JavaScript является ключевым инструментом в области веб-ориентированной геопространственной визуализации и разработки интерактивных картографических приложений. В этой сфере широко применяются такие мощные библиотеки, как Leaflet, OpenLayers, Cesium для создания трехмерных глобусов и Mapbox GL JS, предоставляющая возможности для настройки пользовательских карт.

Сравнение JavaScript и Python в контексте серверной геопространственной обработки и анализа данных выявляет определенные различия. Хотя Node.js позволяет выполнять JavaScript на сервере, Python располагает более развитой и зрелой экосистемой библиотек, предназначенных для научных вычислений, обработки данных и сложных геопространственных анализов.

Анализ производительности этих языков в геопространственных задачах демонстрирует, что JavaScript может быть более эффективным для операций рендеринга на стороне клиента и выполнения интерактивных задач. В то же время, Python часто использует оптимизированные бэкенды на C/C++ для решения вычислительно интенсивных аналитических задач.

JavaScript обладает значительной поддержкой со стороны сообщества и широким спектром доступных ресурсов, особенно в области веб-разработки. Это упрощает поиск решений и библиотек для задач веб-картографирования. Важно отметить, что существует возможность интеграции JavaScript с Python через библиотеки, такие как Folium и Plotly, которые позволяют разработчикам,

работающим с Python, создавать интерактивные веб-карты, используя картографические библиотеки JavaScript.

Таким образом, JavaScript является доминирующим языком для разработки интерактивных и веб-ориентированных геопространственных визуализаций. Его применение особенно целесообразно для проектов, требующих отображения карт через веб-интерфейс и обеспечения интерактивного взаимодействия с пользователем. Однако, несмотря на его эффективность в визуализации на стороне клиента, серверные возможности Python и обширный набор аналитических библиотек делают его более предпочтительным выбором для основной обработки геопространственных данных, проведения анализа и выполнения задач сложной манипуляции, которые часто предшествуют процессу визуализации [23, 24].

Java

Язык программирования Java широко применяется для разработки надежных, масштабируемых и кроссплатформенных геопространственных приложений, особенно в сфере корпоративных ГИС, серверов веб-карт и мобильных ГИС-приложений. Ключевыми геопространственными библиотеками для Java являются GeoTools и JTS Topology Suite. Они предоставляют обширные функциональные возможности для создания, обработки и анализа векторной геометрии, а также поддерживают различные форматы геопространственных данных.

Сравнивая более многословный и статически типизированный синтаксис Java с лаконичным и динамически типизированным синтаксисом Python, можно отметить, что многословность Java способствует созданию более надежного и поддерживаемого кода в масштабных проектах. В то же время, лаконичность Python часто ускоряет процесс разработки и облегчает прототипирование, что было критически важно для данной работы.

Производительность компилируемого языка Java обычно превосходит производительность интерпретируемых языков, таких как Python, в задачах,

требующих значительных вычислительных ресурсов, и в корпоративных приложениях. Экосистема библиотек и уровень поддержки сообщества также играют важную роль в выборе языка программирования. Хотя Java имеет сильные позиции в корпоративной разработке, Python обладает более широким и активным сообществом, особенно в области науки о данных, машинного обучения и геопространственного анализа.

Таким образом, Java является предпочтительным выбором для разработки крупных, масштабируемых и кроссплатформенных геопространственных приложений, что делает его подходящим для корпоративных ГИС-решений и серверных систем, требующих высокой производительности и надежности. С другой стороны, Python с его простым синтаксисом и обширной библиотечной поддержкой в области науки о данных и геопространственного анализа может обеспечить более эффективное решение для задач обработки и визуализации геопространственных данных, особенно при необходимости быстрого прототипирования и использования широкого спектра инструментов [25].

C++

Язык программирования C++ является ключевым инструментом для разработки высокопроизводительных и оптимизированных по использованию памяти геопространственных библиотек и ГИС. Среди наиболее значимых проектов, реализованных на C++, можно выделить такие фундаментальные библиотеки, как GDAL, GEOS, которая используется в Shapely и PostGIS, а также Mapnik.

C++ предоставляет разработчикам возможность детального контроля над системными ресурсами, что позволяет достигать оптимальной производительности. В отличие от Python, который характеризуется более высоким уровнем абстракции и относительной простотой разработки, изучение и освоение C++ требует более глубокого понимания низкоуровневых аспектов программирования, что влечет за собой более длительную кривую обучения и

увеличение временных затрат на разработку, которые в рамках этой работы играли бы критичную роль.

Анализ производительности при использовании C++ для вычислительно интенсивных геопространственных задач показывает, что данный язык способен обеспечить значительное увеличение скорости, особенно в контексте пользовательских алгоритмов и операций, требующих интенсивного использования памяти. Однако библиотеки Python, такие как NumPy и SciPy, часто нивелируют этот разрыв, предоставляя высокооптимизированный инкапсулированный код на C/C++.

Сообщество разработчиков C++ в области геопространственной аналитики и ГИС в основном сосредоточено на создании базовых библиотек и инфраструктуры программного обеспечения, а не на разработке пользовательских скриптов или аналитических инструментов.

Таким образом, C++ остается предпочтительным языком для разработки высокопроизводительных геопространственных библиотек и программного обеспечения благодаря своей способности напрямую взаимодействовать с аппаратными ресурсами и оптимизировать использование системных ресурсов, таких как процессорное время и память.

Тем не менее, для большинства задач обработки и визуализации геопространственных данных Python предоставляет более продуктивную и удобную среду разработки, используя свои привязки к базовым библиотекам C/C++. Это позволяет разработчикам достигать баланса между производительностью и простотой использования, не прибегая к сложному низкоуровневому программированию в большинстве случаев [26].

2.2.2 Аргументация выбора языка программирования Python и его ключевых библиотек

Выбор языка программирования Python в качестве инструмента для обработки и визуализации геопространственных данных обусловлен его

значительными преимуществами, выявленными в ходе сравнительного анализа. Python демонстрирует оптимальное сочетание простоты использования, высокой производительности, обширной экосистемы специализированных библиотек, активного сообщества разработчиков и открытого исходного кода, что делает его максимально подходящим для реализации проекта в рамках этой работы.

Синтаксис языка, ориентированный на лёгкость чтения и поддерживаемость кода, способствует сокращению времени разработки и снижению вероятности возникновения ошибок. Интеграция Python с существующими ГИС, такими как ArcGIS (через модуль ArcPy) и QGIS (через PyQGIS), обеспечивает беспрепятственную автоматизацию задач по геообработке и настройку рабочих процессов в рамках уже существующих ГИС-инфраструктур.

Универсальность Python как языка общего назначения позволяет использовать его для решения широкого спектра задач, выходящих за пределы геопространственного анализа. Это включает в себя очистку данных, веб-разработку для визуализации и обмена данными, а также интеграцию с другими инструментами и методами, применяемыми в науке о данных и машинном обучении. Это расширяет возможности по дальнейшему развитию разрабатываемого модуля.

Обширная документация, онлайн-руководства и активное сообщество разработчиков, занимающихся поддержкой Python и его экосистемы, облегчают процесс обучения, устранения неполадок и поиска решений для задач, связанных с геопространственными данными.

Выбор библиотек для модуля:

NumPy

Библиотека NumPy является ключевой для работы с многомерными массивами и матрицами, что делает её незаменимым инструментом для анализа и обработки геопространственных данных, таких как спутниковые изображения

и цифровые модели рельефа. Эта библиотека предоставляет мощные средства для выполнения численных операций с векторными данными, включая преобразование координат и геометрические вычисления.

Одной из ключевых особенностей NumPy является поддержка векторизованных операций, которые позволяют осуществлять математические вычисления над целыми массивами без использования явных циклических конструкций. Это значительно повышает производительность и эффективность обработки больших объемов геопространственных данных, что особенно важно для научных и прикладных исследований.

NumPy служит основой для множества других научных библиотек Python, таких как Pandas, а также специализированных геопространственных инструментов, включая Rasterio и Shapely. Благодаря своей архитектуре и принципам совместимости, NumPy обеспечивает интеграцию и обмен данными между различными программными компонентами, что способствует созданию более сложных и масштабируемых аналитических систем.

Таким образом, NumPy играет фундаментальную роль в геопространственном анализе благодаря своей способности эффективно обрабатывать числовые данные, особенно в контексте растровых изображений. Его функциональность и производительность делают его незаменимым инструментом для выполнения широкого спектра задач, связанных с обработкой и анализом геопространственной информации [23].

Pandas

Библиотека Pandas представляет собой высокопроизводительный инструмент для работы с данными, включающий такие структуры, как DataFrame и Series. Эти структуры предназначены для обработки и анализа табличных данных, часто связанных с геопространственной информацией в виде атрибутивных таблиц. Примерами таких данных могут служить сведения о населении, привязанные к административным границам, или показания датчиков, ассоциированные с координатами местоположения.

Pandas предоставляет широкий спектр функций для очистки данных, таких как обработка пропущенных значений и устранение несоответствий. Кроме того, инструмент включает возможности для манипуляции данными, включая фильтрацию, сортировку, слияние и объединение, а также агрегирование. Эти функции играют ключевую роль в подготовке геопространственных данных для последующего анализа и визуализации.

Важно отметить, что Pandas интегрируется с геопространственными расширениями, такими как GeoPandas. Это расширение расширяет функциональность Pandas, добавляя поддержку геометрических типов данных. Благодаря этой интеграции становится возможным проводить комплексный анализ как пространственной, так и атрибутивной информации в рамках единой структуры DataFrame [23].

Plotly

Библиотека представляет собой инструмент для создания интерактивных визуализаций геопространственных данных. Она поддерживает широкий спектр типов карт, включая точечные карты, фоновые картограммы, карты плотности, а также диаграммы рассеяния, линейные графики, столбчатые диаграммы и прочие статистические графики, которые играют ключевую роль в изучении и представлении геопространственных закономерностей и связей.

Plotly предоставляет возможность создания веб-ориентированных визуализаций, которые могут быть интегрированы в отчеты, информационные панели или опубликованы в интернете. Это способствует эффективной передаче результатов широкому кругу пользователей, включая лиц, не обладающих техническими знаниями.

Богатый функционал Plotly, включающий масштабирование, панорамирование, всплывающие подсказки, информацию при наведении курсора и анимацию, является важным инструментом для разведочного анализа данных. Эти возможности позволяют пользователям детально исследовать

пространственные данные и выявлять скрытые закономерности, которые могут быть упущены при использовании статических визуализаций.

В отличие от статических библиотек для построения графиков, таких как Matplotlib, часто применяемой в качестве бэкенда для базовых графиков в других геопространственных инструментах, Plotly предлагает дополнительные преимущества в виде интерактивности, необходимой для динамического исследования данных и создания привлекательных отчетов [23].

Таблица 2.2.2.1 – Сравнение программных средств для обработки и визуализации геопространственных данных

Характеристика	Простота использования	Производительность (геопространство)	Геопространственные библиотеки	Поддержка сообщества	Стоимость	Основные геопространственные задачи
Python	Высокая	Хорошая (с оптимизированными библиотеками)	Обширные (GeoPandas, Rasterio и др.)	Очень большая и активная	Бесплатный и открытый исходный код	Общая обработка, визуализация, автоматизация, веб-интеграция
R	Средняя (для не-статистиков)	Хорошая (векторная геообработка)	Сильные (sf, raster, ggplot2, leaflet)	Большая и активная	Бесплатный и открытый исходный код	Статистический анализ, визуализация, тематическое картографирование
MATLAB	Средняя (ориентирован на математику)	Хорошая (численные расчеты и обработка изображений)	Средние (Mapping & Image Processing Toolboxes)	Средняя	Проприетарный (может быть дорогостоящим)	Численные расчеты, обработка и анализ изображений
Julia	Средняя	Отличная (потенциал высокой производительности)	Растущие (JuliaEarth, GeoMakie)	Растущая	Бесплатный и открытый исходный код	Высокопроизводительные вычисления, научное моделирование
JavaScript	Средняя	Средняя (бэкенд может быть оптимизирован)	Средние (Leaflet, OpenLayers, CesiumJS)	Очень большая и активная	Бесплатный и открытый исходный код	Веб-картографирование, интерактивная визуализация
Java	Средняя - Высокая (многословный синтаксис)	Хорошая - Отличная (компилируемый язык)	Средние (GeoTools, JTS)	Средняя	Бесплатный и открытый исходный код	Корпоративные приложения, масштабируемые системы
C++	Низкая (крутая кривая обучения)	Отличная (прямой доступ к оборудованию)	Фундаментальные (GDAL, GEOS, Mapnik)	Средняя (ориентирована на разработку ядра)	Бесплатный и открытый исходный код	Разработка высокопроизводительного ПО, создание библиотек

2.3 Исследование и выбор источника гидрометеорологических данных

2.3.1 Обзор и сравнение доступных архивов и наборов данных о температуре воздуха для Арктического региона

Существует достаточное количество источников данных, которые могли быть использованы для тестирования разрабатываемого модуля. Ниже приведена краткая информация по ним.

Наборы данных реанализа:

ERA-40

- Период охвата: Сентябрь 1957 года – август 2002 года. Начало данных совпадает с Международным геофизическим годом, что обеспечивает значительное количество наблюдений.
- Пространственное разрешение: Около 1,125 градусов по широте и долготы (гауссова сетка T159). Соответствует сетке N80.
- Временное разрешение: 6-часовое и месячное. Доступны также суточные данные.
- Формат данных: GRIB. Возможность доступа в формате NetCDF через некоторые архивы.
- Доступность: Данные доступны через архив NCAR RDA после регистрации. Также доступны через другие архивы, такие как BADC (требуется регистрация).
- Известные ограничения и погрешности: Обнаружен разрыв в трендах арктической температуры около 1997 года, вызванный изменениями в обработке спутниковых данных. Возможны проблемы с анализом долгосрочных трендов из-за неоднородности наблюдательных данных. Наблюдаются

нереалистичные тренды для морского льда. Холодный сдвиг (bias) в нижней тропосфере был скорректирован в 1997 году. Сдвиги в стратосфере. Проблемы с гидрологическим циклом. Завышенное количество осадков над тропическими океанами. Слишком сильная циркуляция Брюера-Добсона в поздние годы [27].

ERA5

- Период охвата: январь 1940 года – настоящее время (данные продолжают обновляться).
- Пространственное разрешение: 0,25 градуса по широте и долготе (31 километр).
- Временное разрешение: часовое, суточное, месячное.
- Формат данных: GRIB, NetCDF, Zarr (оптимизированный для облачных хранилищ).
- Доступность: данные предоставляются на безвозмездной основе через Copernicus Climate Data Store (CDS) после регистрации пользователя. Также доступны в Google Cloud (безвозмездно, при условии указания авторства).
- Известные ограничения и погрешности: наблюдается теплый сдвиг температуры воздуха на высоте 2 метра и температуры поверхности над арктическим морским льдом, особенно в холодные зимние и ясные периоды. Отмечено высокий сдвиг солнечной энергии и низкий сдвиг ветровой энергии по сравнению с наблюдениями. В период с 2000 по 2006 годы наблюдался холодный сдвиг в нижней стратосфере, который был исправлен в версии ERA5.1 [28].

ERA-Interim

- Период охвата: 1979 год – август 2019 года (производство прекращено).
- Пространственное разрешение: приблизительно $0,7^\circ \times 0,7^\circ$ (гауссова сетка N128).
- Временное разрешение: 6-часовое и месячное.

- Формат данных: GRIB (оригинальный), NetCDF (версия ICDC).
- Доступность: производство прекращено в 2019 году. Данные, вероятно, останутся доступными через CDS после регистрации.
- Известные ограничения и погрешности: проблемы с моделированием гидрологического цикла, качеством стратосферной циркуляции. Улучшенная временная согласованность по сравнению с ERA-40. Сдвиги в верхней и нижней стратосфере (аналогично ERA5) [29].

NCEP/NCAR Reanalysis 1 (R1)

- Период охвата: с 1948 года по настоящее время.
- Пространственное разрешение: $2,5^\circ \times 2,5^\circ$ на гауссовой сетке T42, спектральное разрешение T62.
- Временное разрешение: ежесуточно, ежемесячно, четыре раза в сутки.
- Формат данных: GRIB.
- Доступность: предоставляется бесплатно через NOAA PSL.
- Известные ограничения и погрешности: данный набор данных является более старым по сравнению с более новыми версиями реанализа и характеризуется более низким пространственным разрешением. Также следует учитывать возможные искусственные тренды в полях среднего уровня моря (MSLP) в районе Антарктиды [30].

MERRA-2 (Modern-Era Retrospective Analysis for Research and Applications, Version 2)

- Период охвата: с 1980 года по настоящее время.
- Пространственное разрешение: $0,5^\circ$ по долготе и $0,625^\circ$ по широте.
- Временное разрешение: часовое, суточное и месячное.
- Формат данных: NetCDF.

- Доступность: бесплатный доступ через NASA Goddard Earth Sciences Data and Information Services Center (GMAO).
- Известные ограничения и погрешности: данные по большей части полностью соответствуют наземным наблюдениям в Арктическом регионе [31].

JRA-55 (Japanese 55-year Reanalysis)

- Период охвата: 1958 – 2023 гг.
- Пространственное разрешение: около 55 км (TL319, редуцированная гауссова сетка).
- Временное разрешение: 6-часовое, 3-часовое (для некоторых переменных), суточное, месячное.
- Формат данных: GRIB.
- Доступность: предоставляется бесплатно через NCAR RDA после прохождения процедуры регистрации и принятия условий использования.
- Известные ограничения и погрешности: значительное уменьшение сдвига значений температуры в нижней стратосфере в сравнении с предшествующей версией (JRA-25); наличие влажного сдвига над морской поверхностью [32].

Arctic System Reanalysis (ASR) (ASRv1, ASRv2)

- Период охвата: ASRv2: период с 2000 по 2016 год; ASRv1: период с 2000 по 2012 год.
- Пространственное разрешение: ASRv1: 30 километров; ASRv2: 15 километров (полярные стереографические сетки).
- Временное разрешение: 3 часа и среднемесячные данные.
- Формат данных: файлы в формате NetCDF (формат выходных данных модели WRF, полярные стереографические сетки).
- Доступность: открытый доступ через NCAR RDA после процедуры регистрации.

- Известные ограничения и погрешности: обнаружены систематические ошибки, такие как сухой сдвиг в холодные месяцы, а также избыток коротковолновой и дефицит длинноволновой радиации, что указывает на необходимость улучшения расчёта физики облаков [33].

Copernicus Arctic Regional Reanalysis (CARRA) (CARRA-West, CARRA-East)

- Период охвата: сентябрь 1990 года – настоящее время (ежемесячные обновления с задержкой в 2–3 месяца).
- Пространственное разрешение: 2,5 километра x 2,5 километра (ламбертова конформная проекция).
- Временное разрешение: трехчасовые анализы и часовые прогнозы.
- Формат данных: GRIB2 и NetCDF.
- Доступность: бесплатные данные предоставляются через Copernicus Climate Data Store (CDS) после прохождения процедуры регистрации.
- Известные ограничения и погрешности: CARRA демонстрирует улучшенное представление температуры поверхности морского льда по сравнению с ERA5. Тем не менее, сохраняются сдвиги, особенно в отношении альбедо морского льда в весенний период [34].

Наборы наблюдательных данных:

GISS Surface Temperature Analysis (GISTEMP)

- Период охвата: с 1880 года по настоящее время, данные обновляются ежемесячно.
- Пространственное разрешение: 2 градуса по широте на 2 градуса по долготы.
- Временное разрешение: ежемесячное.
- Формат данных: формат NetCDF, также доступны текстовые файлы.

- Доступность: данные предоставляются на бесплатной основе через веб-сайт NASA GISS.
- Известные ограничения и погрешности: присутствие сглаживания температурных аномалий вследствие применения методов гомогенизации и реконструкции данных [35].

Berkeley Earth Surface Temperature (Berkeley)

- Период охвата: 1850 год – настоящее время (ежемесячные данные); экспериментальная версия с ежедневными данными с 1880 года.
- Пространственное разрешение: $1^{\circ} \times 1^{\circ}$; в некоторых регионах доступны более высокие разрешения.
- Временное разрешение: Ежемесячное, ежедневное (экспериментально).
- Формат данных: NetCDF; также доступны текстовые файлы в формате CSV.
- Доступность: Бесплатные данные доступны через веб-сайт Berkeley Earth.
- Известные ограничения и погрешности: Аномалии данных подвержены сильному сглаживанию [35, 36].

Глобальная сеть исторических метеорологических наблюдений NOAA (GHCN-Daily)

- Период охвата: 1763 год – настоящее время (ежедневные данные, регулярно обновляемые).
- Пространственное разрешение: Данные метеостанций, характеризуются неравномерным покрытием.
- Временное разрешение: Ежедневно.
- Формат данных: Текстовый формат ASCII (текстовые файлы).

- **Доступность:** Бесплатный доступ предоставляется через NOAA NCEI и AWS (Amazon Web Services).
- **Известные ограничения и погрешности:** Неравномерное пространственное покрытие, неоднородности в данных, обусловленные изменениями времени наблюдения, местоположения станций и типов используемых термометров [37].

2.3.2 Аргументация выбора ERA-40

Набор данных ERA-40 является подходящим выбором для тестирования работоспособности разрабатываемого программного модуля по следующим причинам:

- **Наличие исторических данных:** ERA-40 охватывает период с 1957 по 2002 годы, что позволяет проводить тестирование модуля на различных временных масштабах и выявлять потенциальные проблемы, связанные с обработкой данных за длительный временной интервал.
- **Глобальный охват:** данные ERA-40 предоставляют информацию для всех географических координат, включая Арктический регион, что обеспечивает универсальность и возможность тестирования модуля в различных климатических условиях.
- **Широкая распространенность:** несмотря на то, что ERA-40 был заменен более новыми наборами данных реанализа, данный набор продолжает широко использоваться в климатических исследованиях, а также имеет обширную документацию и инструменты для работы с ним.
- **Известные недостатки:** наличие известных ограничений и сдвигов в данных ERA-40 может быть использовано для тестирования способности модуля обрабатывать и корректировать потенциальные неточности в исходных данных. Например, разрыв в температурных трендах около 1997 года может служить проверкой устойчивости модуля к таким аномалиям.

- Совместимость с инструментами разработки: данные ERA-40 предоставляются в форматах GRIB и NetCDF, которые могут быть легко интегрированы с библиотеками Python, планируемыми к использованию для разработки модуля.
- К моменту начала разработки модуля уже был доступен необходимый массив данных из архива ERA-40, что стало решающим фактором для выбора именно этого источника.

Хотя современные наборы данных, такие как ERA5, обладают более высокой точностью и детализацией, ERA-40 предоставляет достаточный уровень информации и временной охват для первоначального тестирования функциональности модуля обработки и анализа геопространственных данных. Это особенно важно на этапе разработки и отладки. После успешного тестирования на ERA-40 модуль может быть проверен с использованием более современных наборов данных, таких как ERA5, что позволит оценить его производительность с применением самых актуальных данных.

2.4 Концептуальное проектирование алгоритмов обработки и визуализации

2.4.1 Выбор и обоснование методов анализа данных

Выбор методов анализа данных для разрабатываемого модуля был продиктован необходимостью извлечения значимой гидрометеорологической информации из исходного массива данных ERA-40 для решения практических задач морской деятельности в Арктике.

Фильтрация данных (пространственная и временная)

Описание: данный метод предполагает отбор подмножества данных в соответствии с заданными критериями. В разрабатываемых скриптах реализованы фильтрация по географическим координатам (широта и долгота) и по временному диапазону (конкретные даты или периоды).

Обоснование: морская деятельность в Арктическом регионе часто связана с определенными географическими районами (например, трассы Северного морского пути, зоны буровых платформ) и конкретными временными периодами или операционными окнами. Фильтрация обеспечивает получение гидрометеорологической информации, релевантной для конкретных задач пользователя, исключая избыточные данные. Это способствует упрощению анализа и повышению оперативности получения необходимых сведений. Использование целочисленных координат для фильтрации обусловлено дискретной сеткой исходных данных ERA-40.

Агрегация данных (расчет среднемесячных значений)

Описание: разрабатываемый модуль применяет метод агрегации путем вычисления среднего арифметического значения температуры для каждого календарного дня (например, средняя температура для 1 января, рассчитанная на основе данных за период с 1957 по 2002 год).

Обоснование: среднемесячные значения (климатические нормы) играют ключевую роль в долгосрочном планировании морских операций. Они позволяют оценить типичные температурные условия для любого дня года в заданной точке, выявить сезонные закономерности и определить периоды с наиболее экстремальными температурами. Данная информация имеет критическое значение для оценки ледовых условий, планирования сроков навигации и выбора оборудования и материалов, устойчивых к арктическим климатическим условиям. Метод позволяет сгладить межгодовую изменчивость и выделить устойчивые климатические характеристики.

Анализ временных рядов (фактических значений)

Описание: разрабатываемый модуль направлен на визуализацию фактических значений температуры за выбранный пользователем непрерывный временной период с указанием дня, месяца и года.

Обоснование: анализ временных рядов фактических значений температуры необходим для изучения конкретных гидрометеорологических условий в прошлом, оценки изменчивости температуры в заданном интервале времени и выявления экстремальных значений или аномалий. Это может быть полезно для ретроспективного анализа условий, при которых происходили различные события, а также для верификации прогнозных моделей и понимания динамики температурных изменений в определенные годы.

Преобразование единиц измерения

Описание: перевод температуры из шкалы Кельвина в градусы Цельсия.

Обоснование: хотя шкала Кельвина является стандартной единицей измерения в научных расчетах и ряде наборов данных, градусы Цельсия более привычны и интуитивно понятны для большинства пользователей, занимающихся морской деятельностью. Это повышает практическую применимость и удобство использования предоставляемой информации.

Выбранные методы анализа являются базовыми, но эффективными для первичной обработки и представления долгосрочной гидрометеорологической информации. Они позволяют получить как обобщенные климатические характеристики, так и детализированные данные о фактических условиях, что соответствует целям создания системы обеспечения морской деятельности в Арктике.

2.4.2 Выбор и обоснование методов визуализации

Для наглядного представления результатов обработки и анализа гидрометеорологических данных в разработанном модуле был выбран метод построения линейных графиков. Этот выбор обусловлен следующими факторами.

Основным анализируемым параметром является температура воздуха, которая представляет собой непрерывную величину, изменяющуюся во времени. Задача заключается в визуализации данной динамики.

Линейные графики (или графики временных рядов) являются классическим и наиболее подходящим инструментом для визуализации изменений непрерывной переменной, такой как температура, в зависимости от времени. Они позволяют легко анализировать тенденции (повышение, понижение), сезонные колебания, экстремальные значения и общую картину изменчивости параметра.

Линейный график с временной осью (ось X) и осью значений температуры (ось Y) является интуитивно понятным для широкого круга пользователей, включая специалистов в области морской деятельности.

Пользователь может быстро оценить температурный режим в выбранной точке и за указанный период. Разрабатываемый модуль создаёт как график среднесезонных температур, который демонстрирует годовой ход температуры, характерный для данной локации, так и график, иллюстрирующий фактические значения температуры.

Библиотека `plotly.express` позволяет создавать интерактивные и визуально привлекательные графики с минимальным количеством программного кода.

Интерактивность (даже базовая, такая как отображение точных значений при наведении курсора или возможность масштабирования) значительно повышает удобство анализа данных. Библиотека `Plotly` ориентирована на создание веб-интегрируемых визуализаций, что соответствует общей задаче этой

работы: разработки модуля для "интерактивной автоматизированной системы". Plotly.express упрощает процесс создания стандартных типов графиков, таких как линейные, делая код более лаконичным и читаемым.

Графики сопровождаются информативными заголовками, содержащими указание на отображаемый параметр (температура), географические координаты и временной период. Оси подписаны, а названия месяцев локализованы (переведены на русский язык).

Такая адаптация делает визуализацию максимально понятной и готовой к использованию конечным пользователем системы без необходимости дополнительных пояснений.

3. Разработка модуля обработки и анализа геопространственных данных

3.1 Общая информация о функциях модуля и их реализации

В рамках данной работы была реализована программная система для визуализации многолетних данных температуры на высоте 2 м с целью анализа климатических изменений и сезонных колебаний температуры воздуха на определённых географических координатах. Основной задачей системы является преобразование и последующая обработка метеорологических данных, представленных в формате txt-файлов, с последующей визуализацией температурных кривых. Обработка и анализ данных производились с использованием языка программирования Python и специализированных библиотек, таких как pandas, numpy и plotly.express.

Ввиду значительного объема исходных данных — ежегодных архивов температурных измерений, охватывающих 45 лет, — была реализована процедура автоматического объединения отдельных годовых наборов в единый

датафрейм. Сначала файлы txt были преобразованы в CSV формат. Каждый CSV-файл содержал значения температуры, указанные в Кельвинах, а также координаты (широта и долгота), временные метки и прочую вспомогательную информацию. Для обеспечения удобства дальнейшего анализа температура была конвертирована в шкалу Цельсия путём вычитания 273.15 из каждого значения.

Одной из ключевых задач данной работы стало обеспечение возможности интерактивного выбора пользователем конкретной точки на глобусе — путём ввода широты и долготы — для которой осуществляется дальнейшая обработка данных. Таким образом, пользователь может проанализировать температурные характеристики интересующей его области. Фильтрация по координатам позволяет значительно уменьшить размер обрабатываемого массива данных и повысить эффективность вычислений.

Для анализа температурных трендов и климатических закономерностей были разработаны два скрипта. Первый предназначен для расчёта и построения графика средних многолетних температур для каждого дня года. Это позволяет исследовать характерные сезонные изменения в долгосрочной перспективе. В этом случае для каждого дня (например, 15 марта) рассчитывается среднее значение температуры по всем годам наблюдений. Таким образом формируется обобщённый температурный профиль года, усреднённый по всей выборке. Пользователь может задать интересующий период в пределах года — например, с 1 января по 31 апреля, либо только с мая по август — и получить графическое представление среднемноголетней температуры в указанный интервал. Также несложно изменить период, данные за который осредняются, для этого достаточно лишь убрать те исходные данные, которые не интересуют пользователя. Однако для большей наглядности в этой работе все расчёты делались по всему доступному периоду.

Второй скрипт разработан для отображения фактических температурных значений во времени для заданных координат. В отличие от предыдущего, данный инструмент предоставляет возможность отследить изменения температуры во времени в конкретной точке за полный период с 1957 по 2002

год. Это особенно полезно для анализа климатических трендов, выявления экстремальных температурных значений, а также для построения регрессионных моделей изменения климата в долгосрочной перспективе. В отличие от первого скрипта, пользователь может задать конкретную начальную и конечную дату, между которыми будет производиться анализ данных. При этом датой начала может быть, например, 01 сентября 1980, а окончанием — 30 августа 2000.

Для повышения наглядности и удобства восприятия визуализации результатов обоих скриптов реализованы с использованием библиотеки `plotly.express`, обеспечивающей интерактивные графики с возможностью масштабирования, навигации по временной шкале и отображения точных значений. Кроме того, все англоязычные обозначения дат и месяцев автоматически переводятся на русский язык, что делает интерфейс работы с программой более интуитивным и удобным для русскоязычного пользователя. Например, обозначения месяцев вида «JAN», «FEB» и т.д. были приведены к форме «Январь», «Февраль» и т.д., а также реализована локализация форматов дат.

Таким образом, реализованная программная система предоставляет пользователю мощный инструмент для климатического анализа, позволяющий гибко взаимодействовать с большими историческими массивами температурных данных, визуализировать их в разных разрезах и формировать наглядные графики, отображающие как усреднённые, так и фактические значения. Полученные результаты затем были встроены в интерфейс разрабатываемой веб-страницы, позволяя пользователю быстро и без труда получать интересующие его данные в удобно визуализированном виде.

3.2 Используемый источник данных

Созданный в ходе работы модуль обработки и анализа геопространственных данных сделан таким образом, чтобы обрабатывать любые данные, представленные в подходящем формате. Однако при разработке и тестировании модуля для удобства использовался конкретный набор данных. В данной главе подробно описывается источник, структура и особенности этих данных, а также обоснование их выбора в рамках данной работы.

Данные, использованные в рамках настоящего исследования, взяты из открытого доступа, предоставленного Азиатско-Тихоокеанским центром дистанционного зондирования климата (Asia-Pacific Data Research Center) при Гавайском университете в Маноа. Конкретный адрес используемого набора данных: http://apdrc.soest.hawaii.edu/dods/public_data/ERA-40/daily-surface. Этот ресурс предоставляет доступ к климатическим данным, собранным в рамках проекта ERA-40 — одного из наиболее обширных реанализов метеорологических наблюдений, осуществлённого Европейским центром среднесрочных прогнозов погоды (ECMWF, European Centre for Medium-Range Weather Forecasts).

Проект ERA-40 представляет собой глобальный набор данных реанализа, охватывающий период с сентября 1957 по август 2002 года. Реанализ — это методология, при которой исторические данные наблюдений объединяются с результатами численного моделирования атмосферы с целью получения наиболее полной и согласованной картины климатических условий прошлого. Таким образом, данные ERA-40 обладают высокой степенью достоверности и применимы для различных климатологических, метеорологических и геоинформационных исследований.

Таблица 3.2.1 – Основные характеристики проекта ERA-40

Характеристика	Детали
Временной период	Сентябрь 1957 г. — август 2002 г.
Пространственное разрешение	Редуцированная гауссова сетка N80, приблизительно 125 км
Вертикальные уровни	60 уровней
Техника ассимиляции данных	3D-вариационная
Используемая модель	IFS CY23r4
Частота анализа	Четыре раза в день (00, 06, 12, 18 UTC)
Формат данных	GRIB

В рамках данной работы использовались 46 текстовых файлов формата .txt, каждый из которых содержит данные за один календарный год, начиная с 1957 и заканчивая 2002 годом включительно. Эти файлы имеют идентичную структуру и состоят из двух частей: метаданных и непосредственно табличных данных.

В начале каждого файла содержится блок описательной информации, содержащий основные сведения о наборе данных. Пример типичного заголовка:

«Total Number of Records 315810.

DATA SET: http://apdrc.soest.hawaii.edu/dods/public_data/ERA-40/daily-surface

ERA40 surface daily averaged

X: 0.5 to 315810.5

Column 1: DATETIME is Time String

Column 2: TIME is Time Coordinates (days since 01-JAN-0001 00:00:00)

BAD FLAG : -1.E+34

Column 3: LON is Longitude (degrees_east) BAD FLAG : -1.E+34

Column 4: LAT is Latitude (degrees_north) BAD FLAG : -1.E+34

Column 5: PARAM24 is * 2 metre temperature [k] (k) BAD FLAG : -1.E+34»

В этих метаданных содержится информация о:

- общем количестве записей в файле;
- ссылке на источник данных;
- описании измеряемого параметра;
- диапазонах координат;
- форматах столбцов и их значениях по умолчанию при отсутствии

данных (BAD FLAG).

После блока метаданных следует строка с заголовками столбцов и непосредственно табличные записи. Все данные хранятся в текстовом виде, разделены запятыми и представлены в следующем формате (на примере 2002 года):

```
«DATETIME,TIME,LON,LAT,PARAM24
```

```
"01-JAN-2002",730853,0,70,274.73
```

```
"01-JAN-2002",730853,2.5,70,274.896
```

```
...
```

```
"30-AUG-2002",731094,357.5,90,270.656
```

```
"30-AUG-2002",731094,360,90,270.656»
```

Здесь:

- DATETIME — строковое представление даты в формате ДД-МММ-ГГГГ (например, "01-JAN-2002");
- TIME — числовой индекс времени, отражающий количество дней, прошедших с 1 января 0001 года;
- LON — долгота в градусах восточной долготы (от 0 до 360);

- LAT — широта в градусах северной широты (от -90 до 90, но в данном исследовании рассматривались значения от 70 до 90);
- PARAM24 — значение температуры воздуха на высоте 2 метра над уровнем поверхности в кельвинах.

Таким образом, каждая строка представляет собой измерение температуры в определённой точке геопространственной сетки на конкретную дату.

Данные имеют ежедневную дискретность, что позволяет отслеживать как долгосрочные тренды, так и сезонные, а также суточные колебания при необходимости агрегирования.

Каждый файл содержит сотни тысяч записей, в зависимости от количества дней в году и плотности геопространственной сетки. Например, в среднем для одного дня фиксируется около 720 измерений. Сетка данных имеет равномерное распределение точек по широте и долготе. Долгота варьируется от 0° до 360° с шагом 2.5°, широта — от -90° до 90°, также с шагом 2.5°. В случае данных за 2002 год общее количество записей равно 315 810.

Как уже было упомянуто выше, поскольку полученный в результате этой работы модуль создавался для использования в Арктическом регионе, из исходных данных сразу же были убраны все данные, полученные южнее 70-й северной широты.

3.3 Подробное описание первого скрипта для анализа и визуализации температурных данных

Этот скрипт является инструментом для анализа и визуализации температурных данных на основе файлов CSV, содержащих информацию о температуре воздуха в разные годы. Его целью является извлечение, обработка и визуализация этих данных для конкретных географических координат (широта и долгота) с использованием различных библиотек Python, таких как pandas для обработки данных и plotly для построения графиков.

Импорт необходимых библиотек

В начале скрипта импортируются несколько библиотек, которые будут использоваться для обработки данных и построения графиков:

```
«import numpy as np
import pandas as pd
import plotly.express as px
from datetime import datetime, timedelta»
```

- NumPy (np) — библиотека для работы с многомерными массивами и матрицами, а также для выполнения различных математических операций.
- Pandas (pd) — это одна из самых популярных библиотек для работы с данными в Python. Она позволяет легко манипулировать таблицами данных, такими как DataFrame, выполнять различные операции с данными, например, фильтрацию, агрегацию и группировку.
- Plotly Express (px) — библиотека для визуализации данных. В частности, она используется для создания графиков и диаграмм. В нашем случае она применяется для построения графиков изменения температуры.
- Datetime и timedelta — эти два компонента модуля datetime позволяют работать с датами и временем. В данном случае они будут использоваться для манипуляций с датами в датасетах, таких как создание нового диапазона дат для будущих анализов.

Загрузка данных

Для начала работы с данными скрипт загружает данные из нескольких CSV файлов. Каждый файл содержит информацию о температуре воздуха для конкретного года. Все файлы имеют схожую структуру, и скрипт последовательно загружает каждый из них с помощью функции `pd.read_csv()`. Каждое чтение файла загружает его в DataFrame, объект, используемый в библиотеке pandas для представления таблиц данных.

```
«df_2002 = pd.read_csv('C:/Users/miron/Desktop/startap/y2002y.csv')
df_2001 = pd.read_csv('C:/Users/miron/Desktop/startap/y2001y.csv')
```

```
df_2000 = pd.read_csv('C:/Users/miron/Desktop/startap/y2000y.csv')
```

и так далее...»

Каждый файл содержит данные о температуре воздуха за определенный год, начиная с 1957 года и заканчивая 2002 годом. Все эти файлы находятся на одном уровне в файловой системе, и каждый из них имеет название в формате y<год>.csv, где <год> — это год, для которого собраны данные. Например, файл y2002y.csv содержит данные за 2002 год.

Компоновка DataFrame из нескольких файлов

После того как все данные загружены, они должны быть объединены в один общий DataFrame для упрощения дальнейшей обработки. Для этого создается список, содержащий все загруженные DataFrame:

```
«df_1957_2002 = [df_1957, df_1958, df_1959, df_1960, df_1961, df_1962,
df_1963, df_1964, df_1965, df_1966, df_1967, df_1968, df_1969, df_1970, df_1971,
df_1972, df_1973, df_1974,
df_1975, df_1976, df_1977, df_1978, df_1979, df_1980, df_1981,
df_1982, df_1983, df_1984, df_1985, df_1986, df_1987, df_1988, df_1989, df_1990,
df_1991, df_1992,
df_1993, df_1994, df_1995, df_1996, df_1997, df_1998, df_1999,
df_2000, df_2001, df_2002]»
```

Каждый из этих DataFrame представляет собой таблицу данных, где строки содержат записи о температуре в разные дни года, а столбцы включают такие данные, как дата, температура, широта и долгота.

Далее эти DataFrame объединяются в один с помощью функции `pd.concat()`, которая объединяет список DataFrame в один большой DataFrame. В результате получается единая таблица, содержащая данные за все годы с 1957 по 2002.

```
«df_1957_2002_ct = pd.concat(df_1957_2002)»
```

Перевод температуры в градусы Цельсия

Температура, записана в Кельвинах, так как это стандартная единица измерения в научных наблюдениях и экспериментах. Однако для удобства работы и визуализации данные необходимо перевести в более привычные градусы Цельсия. Для этого из значений температуры вычитается 273.15 (разница между шкалами Кельвина и Цельсия):

```
«df_1957_2002_ct['PARAM24'] -= 273.15»
```

Здесь столбец PARAM24 содержит данные о температуре, и операция - 273.15 переводит температуру из Кельвинов в Цельсий. Это важно для дальнейшего анализа и визуализации данных.

Фильтрация данных по заданной широте

После того как данные загружены и преобразованы, скрипт позволяет пользователю задать географическую широту, для которой будет проводиться анализ. Пользователь вводит это значение через стандартный ввод:

```
«lat_in = int(input("Введите широту: "))»
```

Затем скрипт фильтрует исходный DataFrame, оставляя только те строки, которые соответствуют введенной широте:

```
«df_1957_2002_ct_prntd_lat = df_1957_2002_ct[df_1957_2002_ct['LAT'] == lat_in]»
```

Здесь столбец LAT соответствует широте в данных, и фильтрация возвращает только те строки, где широта совпадает с введенной пользователем.

Фильтрация данных по заданной долготе

Аналогично фильтрации по широте, скрипт предлагает пользователю ввести долготу, для которой будет выполняться анализ:

```
«lon_in = int(input("Введите долготу: "))»
```

После этого происходит еще одна фильтрация данных, чтобы оставить только те строки, которые соответствуют заданной долготе:

```
df_1957_2002_ct_prntd_lat_lon =  
df_1957_2002_ct_prntd_lat_lon[df_1957_2002_ct_prntd_lat_lon['LON']  
lon_in].reset_index(drop=True)»
```

Здесь столбец LON отвечает за долготу, и фильтрация оставляет только те строки, где долгота совпадает с заданным значением.

Переименование столбцов

После того как данные были отфильтрованы по широте и долготе, скрипт переименовывает столбцы с английских на русские. Это необходимо для удобства дальнейшего анализа и визуализации, чтобы сделать выводы более понятными для пользователей, которые могут не знать английского языка. Это делается с помощью метода `rename()`, который изменяет имена столбцов в `DataFrame`.

```
df_1957_2002_ct_prntd_lat_lon =  
df_1957_2002_ct_prntd_lat_lon.rename(columns={'LON': 'Долгота', 'LAT':  
'Широта', 'PARAM24': 'Температура °C', 'DATETIME': 'Дата'})»
```

- Столбец LON переименовывается в Долгота.
- Столбец LAT переименовывается в Широта.
- Столбец PARAM24 (который содержит значения температуры) переименовывается в Температура °C.
- Столбец DATETIME, который представляет собой даты, переименовывается в Дата.

Теперь все данные удобны для восприятия, так как пользователи, говорящие на русском языке, могут легко интерпретировать эти столбцы.

Удаление части даты (год)

После того как данные были переименованы, скрипт продолжает обработку дат. В исходных данных даты могут быть представлены в формате,

который включает в себя год, месяц и день, но для анализа достаточно оставить только месяц и день. Для этого из столбца с датами (Дата) убирается часть, связанная с годом.

```
«df_1957_2002_ct_prntd_lat_lon_minyr =  
df_1957_2002_ct_prntd_lat_lon.copy()  
df_1957_2002_ct_prntd_lat_lon_minyr['Дата'] =  
df_1957_2002_ct_prntd_lat_lon_minyr['Дата'].str[:5]»
```

Метод `str[:5]` используется для обрезки последних 5 символов в строках столбца Дата. Это позволяет оставить только месяц и день, избавившись от информации о годе. Например, если дата была представлена как 01-Jan-2024, то результатом будет только 01-Jan. Это делает столбец с датами универсальным, подходящим для анализа многолетних данных.

Удаление лишних столбцов

Далее скрипт очищает данные от лишних столбцов, которые больше не нужны для анализа. Мы видим, что столбцы TIME, Долгота и Широта могут быть исключены, так как они не несут дополнительной информации после фильтрации данных по этим параметрам.

```
«df_1957_2002_ct_prntd_lat_lon_minyr_mincln =  
df_1957_2002_ct_prntd_lat_lon_minyr.copy()  
df_1957_2002_ct_prntd_lat_lon_minyr_mincln =  
df_1957_2002_ct_prntd_lat_lon_minyr_mincln.drop(['TIME', 'Долгота', 'Широта'],  
axis=1)»
```

- Столбец TIME убирается, поскольку он не используется для анализа в дальнейшем.
- Столбцы Долгота и Широта уже были использованы для фильтрации данных, и теперь они становятся лишними, так как данные уже конкретизированы по координатам.

После выполнения этой операции остаются только два столбца: Дата и Темпераутра °C, которые будут использоваться для дальнейших расчетов и визуализации.

Осреднение данных по датам

Важным шагом в обработке данных является вычисление средней температуры за каждый день. Скрипт группирует данные по столбцу Дата и вычисляет среднюю температуру для каждого дня. Это нужно для того, чтобы получить более сглаженную картину температурных колебаний и уменьшить возможные шумы, вызванные отдельными аномальными значениями.

```
«df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg =  
df_1957_2002_ct_prntd_lat_lon_minyr_mincln.groupby('Дата',  
as_index=False)['Темпераутра °C'].mean()»
```

- Метод `groupby('Дата')` группирует все строки по уникальным значениям в столбце Дата (каждая дата представлена только один раз).
- После этого метод `.mean()` вычисляет среднее значение температуры по каждой дате.

В результате мы получаем новый DataFrame с двумя столбцами:

1. Дата — уникальные даты, для которых была вычислена средняя температура.
2. Темпераутра °C — средняя температура для каждой даты.

Упорядочивание дат

Поскольку данные о температуре за каждый день представлены в виде строковых значений, скрипт также устанавливает правильный порядок месяцев. Для этого создается новый список всех возможных дат в году (с 1 января по 31 декабря), и порядок этих дат задается как категориальный тип:

```
«start_date = datetime(2024, 1, 1)  
end_date = datetime(2024, 12, 31)»
```

```

new_order = [(start_date + timedelta(days=i)).strftime('%d-%b').upper() for i in
range((end_date - start_date).days + 1)]
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg['Дата'] =
pd.Categorical(df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg['Дата'],
categories=new_order, ordered=True)
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg =
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg.sort_values('Дата')
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg =
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg.reset_index(drop=True)»

```

- Сначала создается список всех дат в году, начиная с 1 января (start_date) и заканчивая 31 декабря (end_date), с шагом в один день.
- Этот список преобразуется в категориальный тип с использованием pd.Categorical(), что гарантирует правильный порядок месяцев.
- Затем DataFrame сортируется по столбцу Дата, чтобы даты шли в хронологическом порядке.

После выполнения этих шагов даты в DataFrame будут расположены в правильном порядке, что имеет большое значение для дальнейшей визуализации.

Перевод месяцев на русский язык

Следующим шагом является замена англоязычных сокращений месяцев на их русскоязычные аналоги. Это делается с помощью словаря month_mapping, где каждому англоязычному сокращению месяца соответствует русскоязычное:

```

«month_mapping = {
    'JAN': 'Января',
    'FEB': 'Февраля',
    'MAR': 'Марта',
    'APR': 'Апреля',
    'MAY': 'Мая',
    'JUN': 'Июня',
    'JUL': 'Июля',

```

```
'AUG': 'Августа',  
'SEP': 'Сентября',  
'OCT': 'Октября',  
'NOV': 'Ноября',  
'DEC': 'Декабря'
```

```
}»
```

Затем с помощью списка `rus_month_dates` происходит замена каждого месяца на его русский аналог:

```
«rus_month_dates = [date.replace(date.split('-')[1], month_mapping[date.split('-')[1]]) for date in new_order]  
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg[df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg.columns[0]] = rus_month_dates»
```

Этот шаг позволяет создать наглядный и понятный график с названиями месяцев на русском языке.

Выбор периода для анализа

Далее скрипт предоставляет пользователю возможность ввести диапазон дат, для которого будет построен график. Это позволяет более гибко подойти к анализу данных:

```
«date_start_in = input("Введите дату начала периода в формате 01-Января:  
")  
date_end_in = input("Введите дату начала периода в формате 31-Декабря:  
")»
```

Пользователь вводит дату начала и окончания периода в формате день-месяц. Скрипт находит индексы этих дат в `DataFrame` и выбирает подмножество данных, которое соответствует введенному диапазону:

```
«start_index =  
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg[df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg['Дата'] == date_start_in].index[0]
```

```

end_index =
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg[df_1957_2002_ct_prntd_lat_lon
_minyr_mincln_avg['Дата'] == date_end_in].index[0]
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg_dates =
df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg.loc[start_index:end_index].reset
_index(drop=True)»

```

Это подмножество данных будет использоваться для дальнейшего построения графика.

```

Введите широту: 85
Введите долготу: 55
Введите дату начала периода в формате 01-Января: 01-Марта
Введите дату начала периода в формате 31-Декабря: 31-Мая

```

	Дата	Средняя многолетняя температура °C
0	01-Марта	-28.597681
1	02-Марта	-28.976043
2	03-Марта	-27.802660
3	04-Марта	-26.740213
4	05-Марта	-26.794936
..
87	27-Мая	-3.360191
88	28-Мая	-2.903979
89	29-Мая	-2.636085
90	30-Мая	-2.691128
91	31-Мая	-2.112660

```

[92 rows x 2 columns]

```

Рисунок 3.3.1 – Интерфейс первого скрипта с выводением данных, по которым строится график

Построение графика

После того как данные были отфильтрованы по выбранному диапазону дат, скрипт строит график средней многолетней температуры для выбранного периода. Для этого используется библиотека `plotly.express`:

```
«fig = px.line(df_1957_2002_ct_prntd_lat_lon_minyr_mincln_avg_dates,
x="Дата", y="Средняя многолетняя температура °C", title=f"График средней
многолетней температуры с {date_start_in} по {date_end_in} для точки с широтой
{lat_in}°с.ш. и долготой {lon_in}°')
fig.show()»
```

Этот код создает линейный график, который отображает изменение средней температуры за каждый день в указанном пользователем интервале. График будет отображаться в интерактивном режиме, что позволяет пользователю исследовать данные более детально.

График средней многолетней температуры с 01-Марта по 31-Мая для точки с широтой 85°с.ш. и долготой 55°

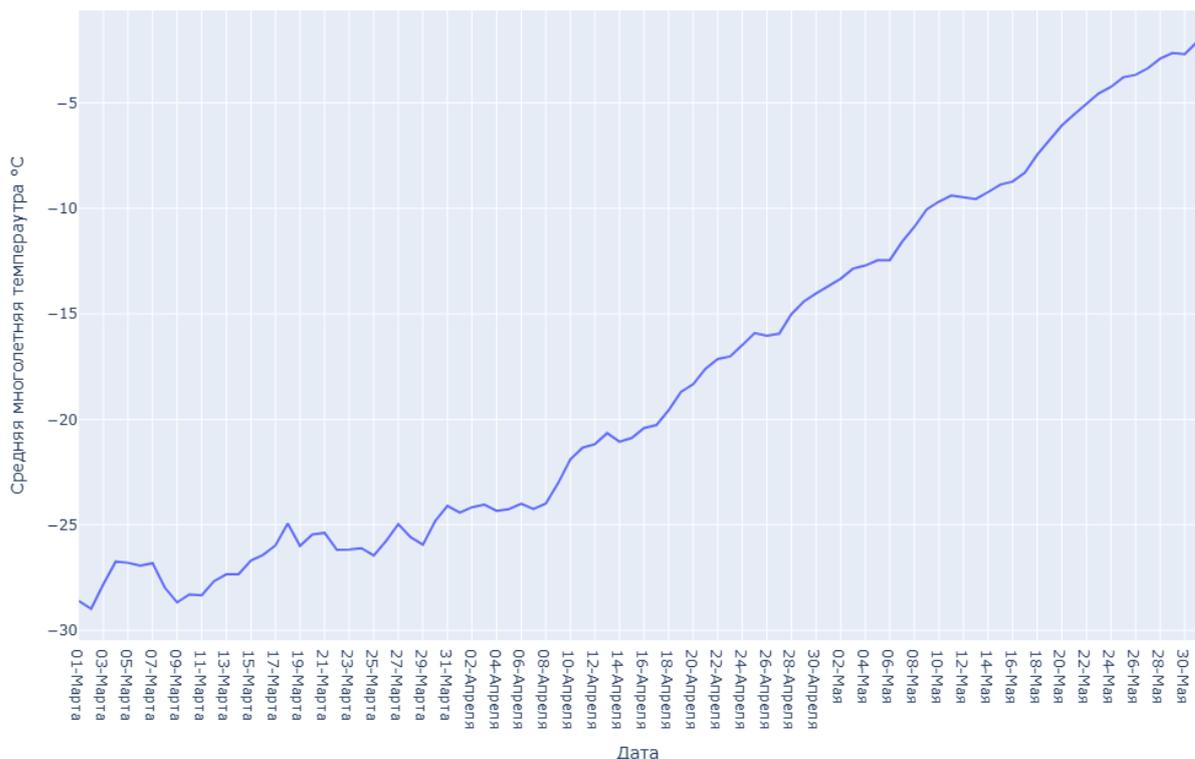


Рисунок 3.3.2 – Полученный в результате работы первого скрипта график

График средней многолетней температуры с 01-Марта по 31-Мая для точки с широтой 85°с.ш. и долготой 55°



Рисунок 3.3.3 – Демонстрация возможностей графика первого скрипта: приближение и просмотр точных значений в любой точке

Итог

Первая часть скрипта включает в себя загрузку данных из файлов, их объединение в один DataFrame, преобразование температурных данных в градусы Цельсия и фильтрацию данных по широте и долготе. Вся эта информация подготовлена для дальнейшего анализа, а также для визуализации в следующей части скрипта. Важно отметить, что именно в этой части скрипта создается основа для дальнейшего анализа, который будет направлен на построение графиков и получение полезных выводов о температурных изменениях в зависимости от географических координат.

Во второй части скрипта производится детальная подготовка данных для визуализации, включая осреднение температур, изменение формата дат, фильтрацию по выбранному периоду и создание графика. Это позволяет пользователю исследовать и анализировать изменения температуры в конкретной

географической точке за выбранный период, делая выводы о климатических изменениях и их сезонных колебаниях.

3.4 Подробное описание второго скрипта для анализа и визуализации температурных данных

В этом пункте подробно рассматривается структура второго скрипта, направленного на визуализацию температурных данных на основе координатных данных (широта и долгота). Основное назначение этого скрипта — загрузка многолетних климатических данных, их объединение в единую структуру, последующая фильтрация по заданным географическим координатам и построение визуализации динамики температуры воздуха во времени.

Как и первый, этот скрипт использует данные из файлов CSV, содержащих информацию о температуре воздуха в разные годы. И создан с использованием различных библиотек Python, о чем ниже.

Импорт необходимых библиотек

В начале скрипта производится импорт стандартных и внешних библиотек, необходимых для работы:

```
«import numpy as np
import pandas as pd
import plotly.express as px»
```

На первом этапе скрипт подключает три ключевые библиотеки:

- NumPy (numpy) — одна из самых популярных библиотек в языке Python, ориентированная на выполнение высокопроизводительных математических операций и работу с многомерными массивами.
- Pandas — основная библиотека для анализа и манипуляции данными в табличном формате. Все температурные данные в скрипте представлены в виде таблиц (dataframe), которые загружаются, обрабатываются и объединяются с помощью инструментов, предоставляемых pandas.

- Plotly Express (`plotly.express`) — библиотека для построения интерактивных графиков. В скрипте она используется для финальной визуализации температурных данных, что делает результаты более наглядными и доступными для анализа.

Загрузка температурных данных за период 1957–2002 гг.

Далее скрипт осуществляет последовательную загрузку CSV-файлов, каждый из которых содержит данные за один календарный год. Данные охватывают временной диапазон с 1957 по 2002 год включительно, что составляет 46 лет наблюдений.

```
«df_2002 = pd.read_csv('C:/Users/iron/Desktop/startap/y2002y.csv')
df_2001 = pd.read_csv('C:/Users/iron/Desktop/startap/y2001y.csv')
...
df_1957 = pd.read_csv('C:/Users/iron/Desktop/startap/y1957y.csv')»
```

Каждое обращение к `pd.read_csv()` создаёт отдельный объект `DataFrame`, содержащий табличные данные за конкретный год. Каждый файл содержит данные в структурированном табличном виде с параметрами, среди которых ключевыми являются:

- LAT — широта точки измерения;
- LON — долгота точки измерения;
- PARAM24 — температурное значение, представленное в градусах Кельвина;
- DATETIME — дата измерения.

Формат CSV (`Comma Separated Values`) является широко используемым для хранения табличных данных, и библиотека `pandas` автоматически определяет структуру файла, используя первую строку в качестве заголовка. При этом скрипт предполагает, что все файлы находятся в одной директории на локальном диске.

Путь к файлам прописан в абсолютном формате. Это означает, что для запуска скрипта на другом устройстве потребуется адаптировать путь к месту хранения файлов. Загрузка данных осуществляется через метод `pd.read_csv()`,

который позволяет загружать табличные данные из текстовых файлов с разделителем — запятой. Каждый загруженный файл сохраняется в отдельную переменную (`df_2002`, `df_2001` и т.д.), что позволяет сохранить привязку данных к соответствующему году.

Загрузка данных вручную, построчно и по одному файлу, пусть и кажется избыточной, выполняет важную функцию: она обеспечивает явную читаемость и предсказуемость работы скрипта, так как список годов не предполагается изменять динамически. В более универсальной реализации это можно было бы заменить автоматизированным циклом, однако для целей этой работы приоритет отдается прозрачности и гибкости получаемого решения.

Формирование списка с датасетами

```
«df_1957_2002 = [df_1957, df_1958, df_1959, ..., df_2002]»
```

На следующем этапе происходит компоновка всех отдельных датафреймов в один список `df_1957_2002`. Это необходимо для последующего объединения данных в один общий датафрейм. Порядок добавления элементов в список строго соответствует хронологии (от 1957 до 2002 года), что важно для корректной визуализации временных рядов на финальном этапе.

Формирование единого списка позволяет воспользоваться преимуществом метода `pd.concat()`, который принимает список объектов `DataFrame` и объединяет их в одну общую таблицу, что значительно упрощает манипуляции с большим числом однотипных таблиц.

Благодаря этому подходу удалось сохранить чистую и последовательную структуру, а также при необходимости производить повторную загрузку данных без изменений в коде.

Объединение данных в один датафрейм

```
«df_1957_2002_ct = pd.concat(df_1957_2002)»
```

После того как все годовые таблицы собраны в список, с помощью функции `pandas.concat()` происходит их вертикальное объединение в единый

датафрейм `df_1957_2002_ct`. Эта операция объединяет все таблицы по строкам, предполагая, что структура столбцов в каждой из таблиц одинакова. Это справедливо, поскольку все файлы являются результатом автоматизированной выгрузки данных с одинаковой структурой.

Результирующий датафрейм содержит все доступные значения температуры за 46-летний период по множеству координат. Такой объем данных позволяет проводить как локальный, так и глобальный анализ климатических изменений в заданный временной промежуток.

Перевод температуры из Кельвинов в градусы Цельсия

```
«df_1957_2002_ct['PARAM24'] -= 273.15»
```

После объединения данных скрипт выполняет одну из ключевых преобразующих операций — перевод температурных значений из шкалы Кельвина в шкалу Цельсия. Это делается с целью повышения читаемости данных для конечного пользователя, так как температура в градусах Цельсия является общепринятой единицей измерения в метеорологии и повседневной жизни.

Математически перевод осуществляется путём вычитания значения 273.15 из каждого значения в столбце `PARAM24`. Стоит отметить, что операция выполняется векторизовано, то есть сразу ко всему столбцу, без использования циклов, что делает процесс быстродействующим и эффективным.

Таким образом, к этому моменту в датафрейме находятся очищенные и объединённые данные по температуре в Цельсиях, подготовленные для дальнейшей фильтрации по координатам и визуализации.

Ввод координат и фильтрация данных по широте

```
«lat_in = int(input("Введите широту: "))  
df_1957_2002_ct_prntd_lat = df_1957_2002_ct[df_1957_2002_ct['LAT'] ==  
lat_in]»
```

На этом этапе начинается интерактивная часть скрипта, в которой пользователю предлагается ввести значение широты исследуемой точки. Широта

— это одна из двух координат, необходимых для точного указания географического положения. Введённое значение сохраняется в переменной `lat_in`.

С помощью метода фильтрации `DataFrame` по условию `df['LAT'] == lat_in` создаётся новый датафрейм `df_1957_2002_ct_prntd_lat`, в который попадают только те строки, где значение широты соответствует введённому пользователем. Это позволяет сузить массив данных, оставив только интересующую географическую полосу — горизонтальный срез глобальных данных.

Важно отметить, что данный фильтр предполагает точное соответствие значения, то есть пользователь должен вводить целое число, полностью совпадающее с тем, что содержится в колонке `LAT`. В дальнейшем, при модернизации скрипта, может быть реализована функция автоматического определения ближайшего значения при отсутствии точного совпадения.

Ввод долготы и фильтрация данных по координатной паре

```
«lon_in = int(input("Введите долготу: "))
df_1957_2002_ct_prntd_lat_lon =
df_1957_2002_ct_prntd_lat[df_1957_2002_ct_prntd_lat['LON']
lon_in].reset_index(drop=True)»
```

Аналогично предыдущему этапу, пользователь вводит значение долготы. Оно сохраняется в переменной `lon_in`. Далее данные из предыдущего отфильтрованного датафрейма (уже с заданной широтой) дополнительно фильтруются по значению долготы. Это означает, что в датафрейме `df_1957_2002_ct_prntd_lat_lon` остаются только те строки, где значения широты и долготы одновременно соответствуют введённым пользователем.

Метод `reset_index(drop=True)` применяется для сброса старых индексов и создания нового непрерывного индексного ряда. Это повышает читаемость и корректность отображения данных в последующих шагах.

Таким образом, формируется финальный набор данных, привязанный к строго определённой географической точке. Все дальнейшие операции и

визуализация будут выполняться только для этой конкретной координатной пары, что делает анализ строго локализованным.

Локализация формата даты: перевод названий месяцев на русский язык

```
«month_mapping = {
    'JAN': 'Января',
    'FEB': 'Февраля',
    'MAR': 'Марта',
    'APR': 'Апреля',
    'MAY': 'Мая',
    'JUN': 'Июня',
    'JUL': 'Июля',
    'AUG': 'Августа',
    'SEP': 'Сентября',
    'OCT': 'Октября',
    'NOV': 'Ноября',
    'DEC': 'Декабря'
}

def replace_months(date_str):
    day, month, year = date_str.split('-')
    full_month = month_mapping.get(month.upper(), month)
    return f'{day}-{full_month}-{year}'
```

```
df_1957_2002_ct_prntd_lat_lon['DATETIME'] =
df_1957_2002_ct_prntd_lat_lon['DATETIME'].apply(replace_months)»
```

Следующим важным этапом является локализация данных — перевод англоязычных аббревиатур месяцев на русский язык. Это необходимо, во-первых, для повышения читаемости и удобства восприятия графиков русскоязычными пользователями; во-вторых, чтобы сделать итоговую

визуализацию более презентабельной в рамках отечественного научного сообщества.

Для реализации перевода создаётся словарь `month_mapping`, в котором каждому трёхбуквенному обозначению месяца на английском языке соответствует его аналог на русском в родительном падеже.

Функция `replace_months()` принимает строку в формате день-месяц-год, разбивает её на составляющие с помощью метода `.split('-')`, заменяет название месяца на соответствующее русское и возвращает обратно отформатированную строку. Эта функция затем применяется к столбцу `DATETIME` с помощью метода `apply()`, что позволяет обработать весь столбец без использования циклов.

Ввод временного диапазона для отображения

```
«date_start_in = input("Введите дату начала периода в формате 01-Сентября-1957: ")
```

```
date_end_in = input("Введите дату начала периода в формате 30-Августа-2002: ")
```

```
start_index = df_1957_2002_ct_prntd_lat_lon[df_1957_2002_ct_prntd_lat_lon['DATETIME'] == date_start_in].index[0]
```

```
end_index = df_1957_2002_ct_prntd_lat_lon[df_1957_2002_ct_prntd_lat_lon['DATETIME'] == date_end_in].index[0]
```

```
df_1957_2002_ct_prntd_lat_lon_dates = df_1957_2002_ct_prntd_lat_lon.loc[start_index:end_index].reset_index(drop=True)»
```

Пользователь вводит даты начала и окончания интересующего временного отрезка в формате день-месяц-год, причём названия месяцев должны соответствовать русским, полученным на предыдущем этапе. Эти даты сохраняются в переменных `date_start_in` и `date_end_in`.

Для извлечения диапазона строк, соответствующего временному отрезку, скрипт определяет индексы строк, где значения в колонке `DATETIME` совпадают

с введёнными датами. С использованием метода `.index[0]` извлекается первый найденный индекс соответствующей строки (предполагается, что такие строки в выборке есть).

С помощью `.loc[start_index:end_index]` производится срез датафрейма по заданному диапазону. Результатом является датафрейм `df_1957_2002_ct_prntd_lat_lon_dates`, содержащий все строки между начальной и конечной датой. И снова используется `reset_index(drop=True)` для приведения индексов к последовательному виду.

Переименование столбцов на русский язык

```
df_1957_2002_ct_prntd_lat_lon_dates =  
df_1957_2002_ct_prntd_lat_lon_dates.rename(columns={  
    'LON': 'Долгота',  
    'LAT': 'Широта',  
    'PARAM24': 'Температура °C',  
    'DATETIME': 'Дата'  
})»
```

На завершающем этапе перед построением графика производится переименование столбцов на русский язык для повышения удобства восприятия графика. Колонки получают следующие названия:

- LON → Долгота
- LAT → Широта
- PARAM24 → Температура °C
- DATETIME → Дата

Это улучшает читаемость как самой таблицы, так и подписей к осям на графике, особенно при использовании графических библиотек, которые автоматически берут названия столбцов для построения осей.

```

Введите широту: 85
Введите долготу: 55
Введите дату начала периода в формате 01-Сентября-1957: 01-Января-1980
Введите дату начала периода в формате 30-Августа-2002: 01-Января-1984
      Дата      TIME  Долгота  Широта  Темпераутра °С
0      01-Января-1980  722817    55.0    85.0    -27.230
1      02-Января-1980  722818    55.0    85.0    -25.001
2      03-Января-1980  722819    55.0    85.0    -29.094
3      04-Января-1980  722820    55.0    85.0    -33.095
4      05-Января-1980  722821    55.0    85.0    -34.081
...
1453   27-Декабря-1983  724273    55.0    85.0    -37.067
1454   28-Декабря-1983  724274    55.0    85.0    -34.593
1455   29-Декабря-1983  724275    55.0    85.0    -28.304
1456   30-Декабря-1983  724276    55.0    85.0    -22.776
1457   01-Января-1984  724278    55.0    85.0    -20.634

[1458 rows x 5 columns]

```

Рисунок 3.4.1 – Интерфейс второго скрипта с выводением данных, по которым строится график

Визуализация температурных данных

```

«fig = px.line(
    df_1957_2002_ct_prntd_lat_lon_dates,
    x="Дата",
    y="Темпераутра °С",
    title=f'График температуры с {date_start_in} по {date_end_in} для точки с
широтой {lat_in}°с.ш. и долготой {lon_in}°'
)
fig.show()»

```

Финальный этап — построение линейного графика температурных данных с помощью библиотеки `plotly.express`. Визуализация осуществляется методом `px.line`, которому передаётся датафрейм, координата по оси X (дата), координата по оси Y (температура), а также заголовок графика.

Заголовок автоматически формируется на основе введённых пользователем координат и временного диапазона. Он даёт полную информацию о

местоположении и исследуемом периоде, что удобно как для анализа, так и для включения изображения в отчёты или публикации.

Вызов `fig.show()` запускает интерактивное окно отображения графика, где пользователь может масштабировать изображение, просматривать значения в точках и сохранять результат в виде изображения.

График температуры с 01-Января-1980 по 01-Января-1984 для точки с широтой 85°с.ш. и долготой 55°

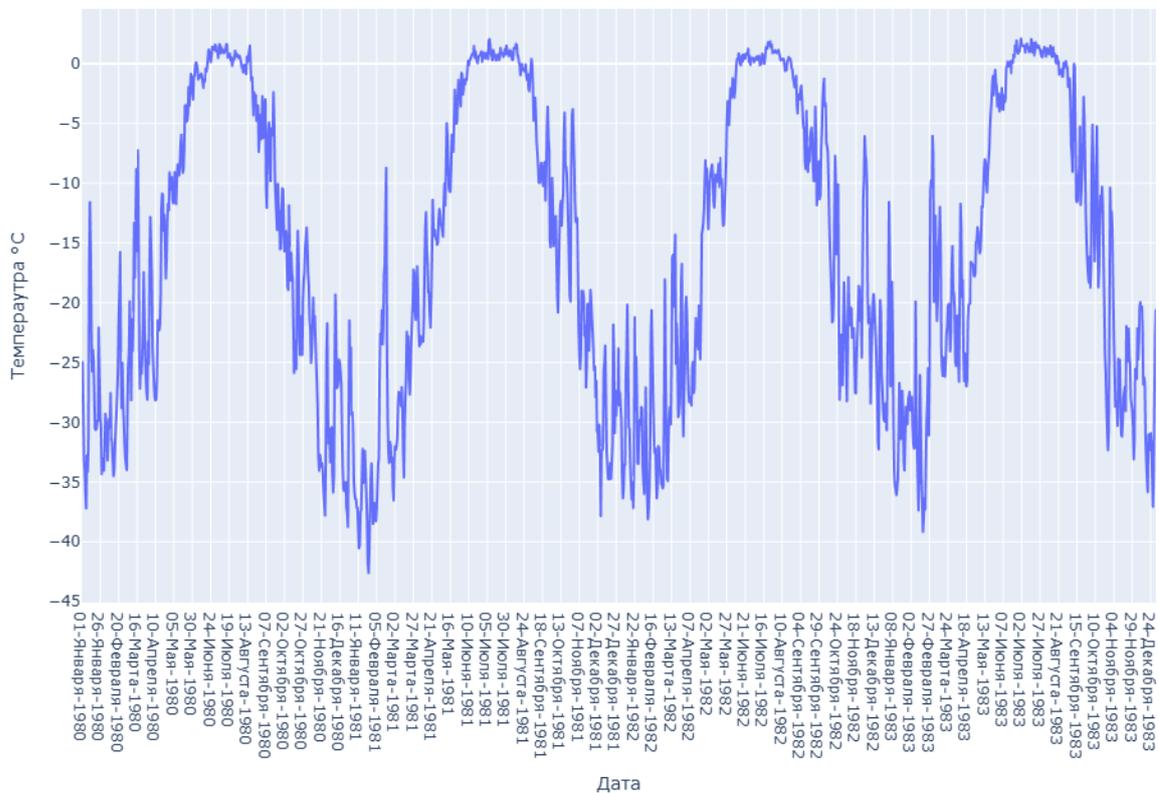


Рисунок 3.4.2 – Полученный в результате работы второго скрипта график

График температуры с 01-Января-1980 по 01-Января-1984 для точки с широтой 85°с.ш. и долготой 55°

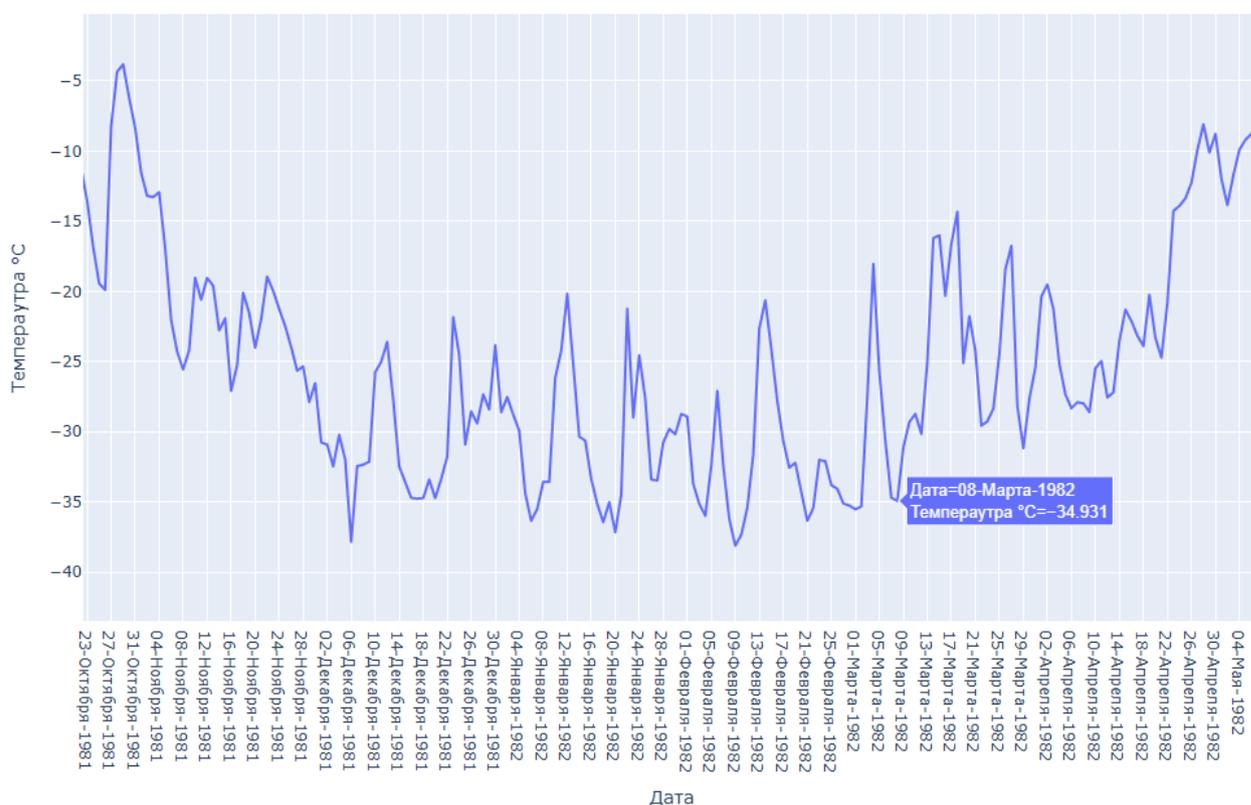


Рисунок 3.4.3 – Демонстрация возможностей графика второго скрипта: приближение и просмотр точных значений в любой точке

4. Аprobация и анализ результатов работы модуля

4.1 Методика тестирования разработанного модуля

4.1.1 Тестовые наборы данных

В качестве основного массива данных для тестирования функциональности разработанных скриптов был использован архив гидрометеорологических данных ERA-40, предоставляемый Европейским центром среднесрочных прогнозов погоды (ECMWF). Обоснование выбора

данного источника подробно изложено в разделах 2.3.2 и 3.2 настоящего исследования. Для тестирования были задействованы данные о температуре воздуха на высоте 2 метров (параметр 'PARAM24' в исходных файлах) за период с 1957 по 2002 годы. Данные предварительно были подготовлены и сохранены в формате CSV-файлов, каждый из которых содержал информацию за весь год для арктического региона.

Тестирование проводилось на полных годовых наборах данных и различных подмножествах, имитирующих пользовательские запросы на конкретные пространственно-временные срезы:

- Анализ обработки полных временных рядов: использование данных за весь период (1957–2002) для оценки агрегирующих функций, таких как расчет средней многолетней температуры, в первом скрипте.
- Анализ обработки по конкретным координатам: ввод различных значений широты и долготы в пределах арктического региона для проверки корректности пространственной фильтрации данных.
- Анализ обработки по конкретным временным интервалам: задание различных дат начала и окончания временного периода для визуализации, включая короткие и длительные интервалы, а также периоды, охватывающие смену месяцев и лет.
- Тестирование на данных для различных точек Арктики: были выбраны тестовые координаты, соответствующие различным физико-географическим условиям в Арктике, например, прибрежные зоны и центральная часть Северного Ледовитого океана, для обеспечения адекватного представления температурных рядов.

Использование данных ERA-40 обеспечило репрезентативность тестовых сценариев благодаря длительному периоду наблюдений, обширному пространственному охвату и высокому качеству данных, признанному в научном сообществе.

4.1.2 Критерии оценки корректности работы

Оценка корректности работы разработанных скриптов проводилась на основании следующих ключевых критериев.

Отсутствие ошибок при выполнении

Скрипты должны функционировать без возникновения исключений Python (runtime errors) при корректном вводе пользовательских данных, таких как координаты и даты. Осуществлялась проверка стабильности работы используемых библиотек (Pandas, Plotly Express, NumPy) в процессе обработки данных.

Корректная обработка входных данных

Точность фильтрации по координатам: необходимо убедиться, что скрипты корректно отбирают данные в соответствии с заданными пользователем значениями широты (lat_in) и долготы (lon_in). Проверка осуществлялась путем сопоставления выходных данных DataFrame с контрольными значениями, полученными вручную или с использованием стандартных инструментов анализа данных для выбранных координат.

Правильность выборки данных по временному диапазону: скрипты должны корректно идентифицировать и извлекать данные в заданном пользователем периоде (между date_start_in и date_end_in). Особое внимание уделялось корректной обработке граничных дат и формированию временных рядов.

Корректность агрегации данных (для первого скрипта): необходимо проверить правильность расчета средних многолетних значений температуры для каждой даты года. Результаты сравнивались с выборочными ручными расчетами для соответствующих координат и дат.

Точность перевода единиц измерения: контролировалась корректность перевода температуры из Кельвинов в градусы Цельсия.

Корректная визуализация данных

Соответствие типа графика поставленной задаче: необходимо убедиться, что генерируется линейный график (px.line), как и было предусмотрено для отображения временных рядов температуры.

Точность отображения данных на графике: ось X (Дата) и ось Y (Температура) должны корректно отображать значения из обработанного DataFrame. Линия графика должна достоверно отражать динамику температуры за выбранный период.

Информативность заголовков и подписей осей: заголовки графиков должны формироваться динамически на основе введенных координат и временного диапазона, обеспечивая однозначную интерпретацию данных. Подписи осей должны быть четкими и отражать величину и единицы измерения отображаемых данных.

Корректность локализации: необходимо проверить корректность отображения названий месяцев в датах, заголовков графиков и подписей осей на русском языке в соответствии с установленными требованиями.

Соответствие выходных DataFrame

Структура и содержание промежуточных и конечных DataFrame (выводимых через функцию print) должны соответствовать ожидаемым результатам на каждом этапе обработки данных. Это включает корректное переименование столбцов и удаление лишней информации.

Оценка по указанным критериям проводилась на основании анализа выводимых данных (DataFrame) на различных этапах работы скриптов. Также осуществлялся визуальный анализ сгенерированных графиков.

4.2. Результаты тестирования функциональности скриптов.

4.2.1 Примеры визуализаций, полученных с помощью разработанных скриптов.

В данном разделе, ради экономии места, представлены исключительно общие изображения графиков, генерируемых посредством разрабатываемого программного модуля. Все графики обладают интерактивными функциями, которые были подробно описаны и продемонстрированы в третьей главе.

График температуры с 01-Сентября-1957 по 30-Августа-2002 для точки с широтой 80°с.ш. и долготой 100°

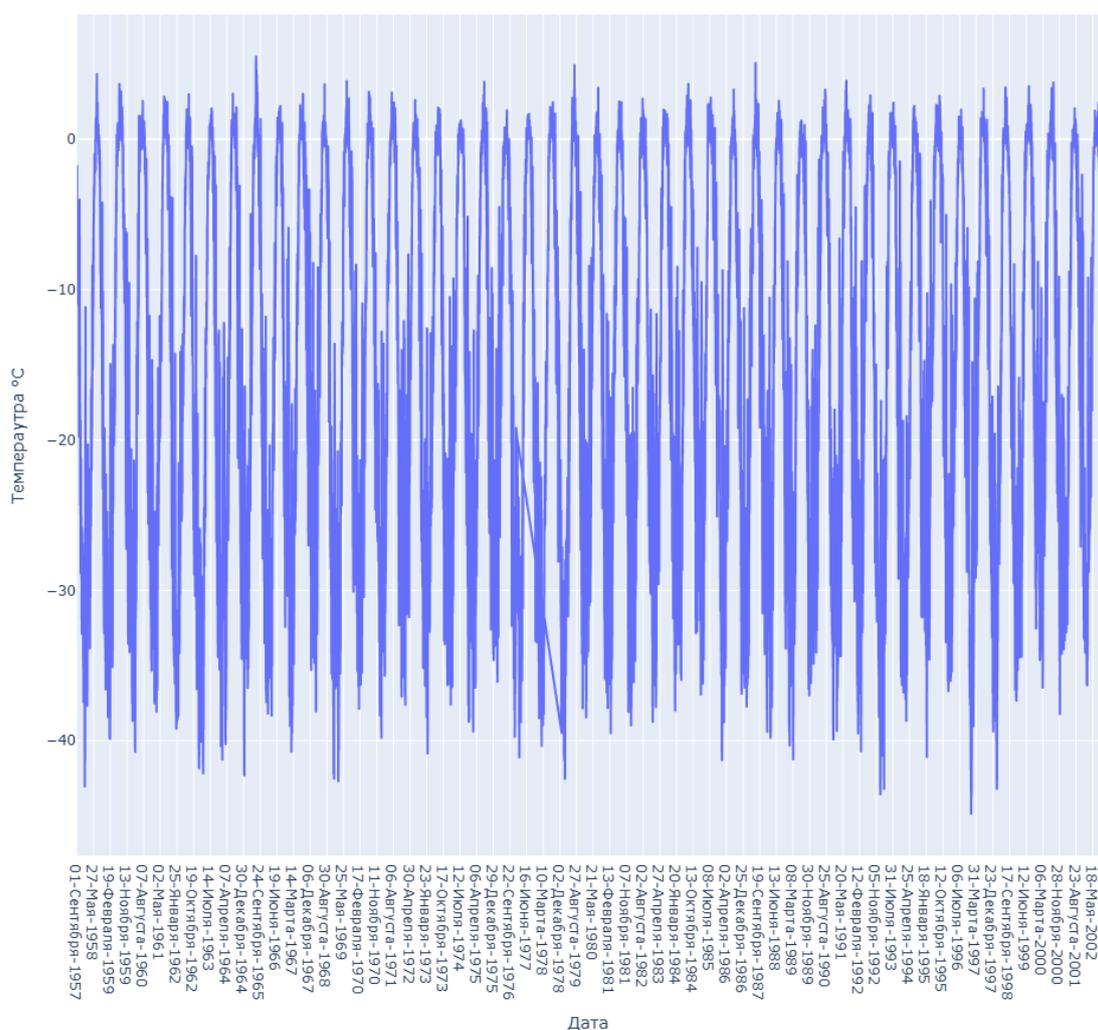


Рисунок 4.2.1.1 – Демонстрация возможности второго скрипта визуализировать весь ряд данных для выбранной точки

График температуры с 01-Сентября-1960 по 01-Сентября-1970 для точки с широтой 90°с.ш. и долготой 10°

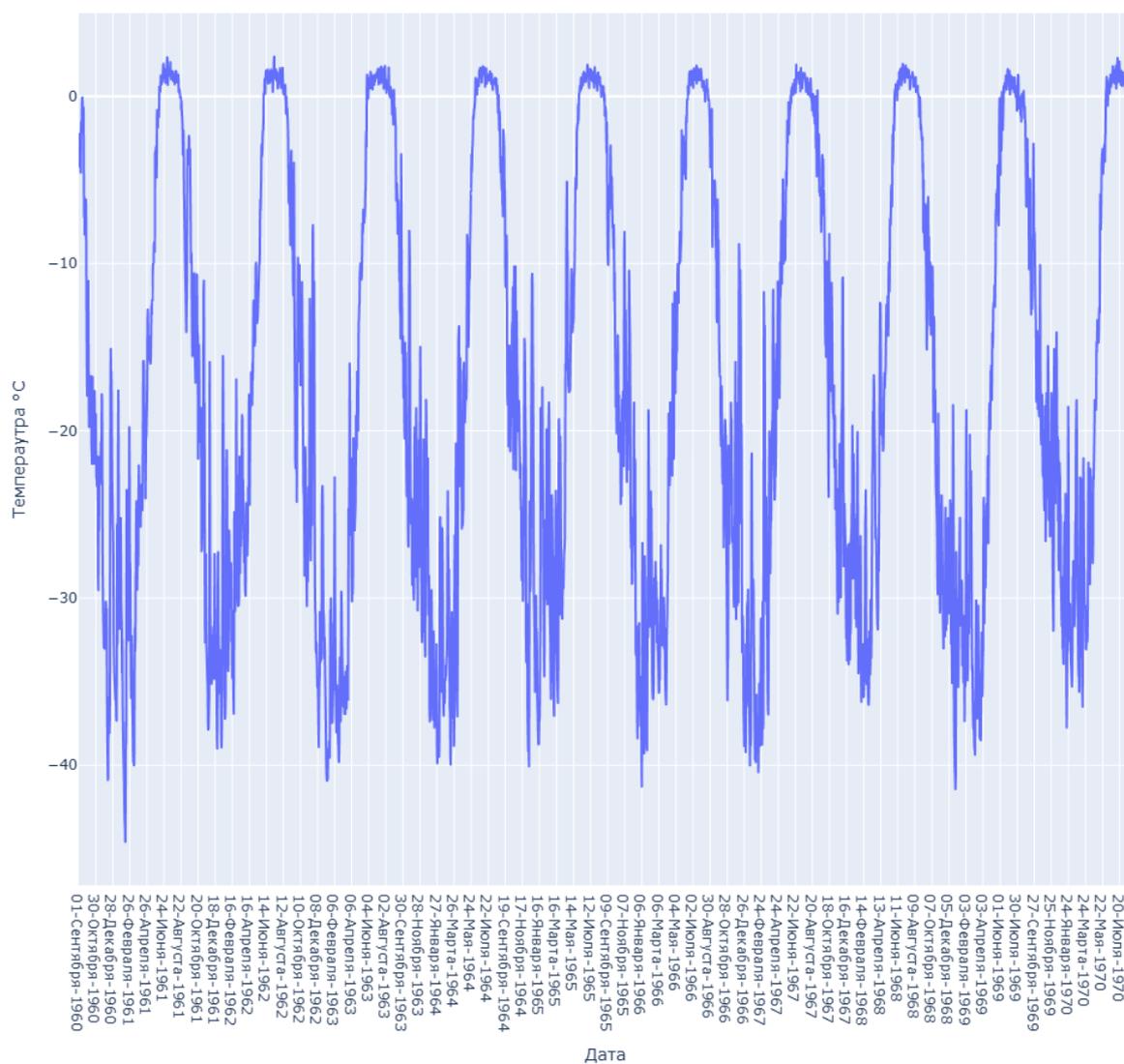


Рисунок 4.2.1.2 – Пример визуализации, полученной с помощью второго скрипта для фиксированной долготы, фиксированных дат и 90° северной широты

График температуры с 01-Сентября-1960 по 01-Сентября-1970 для точки с широтой 85°с.ш. и долготой 10°

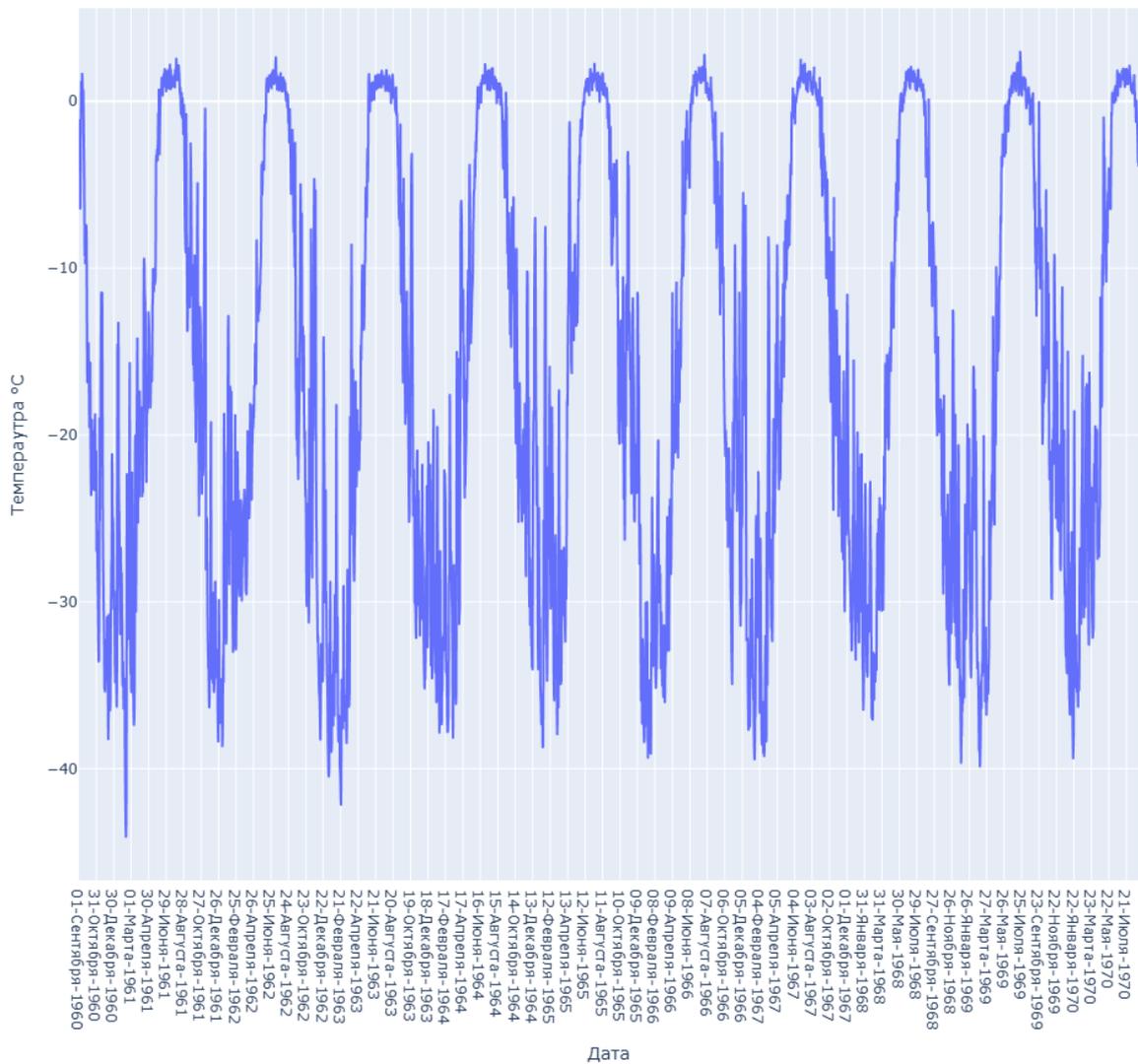


Рисунок 4.2.1.3 – Пример визуализации, полученной с помощью второго скрипта для фиксированной долготы, фиксированных дат и 85° северной широты

График температуры с 01-Сентября-1960 по 01-Сентября-1970 для точки с широтой 80°с.ш. и долготой 10°

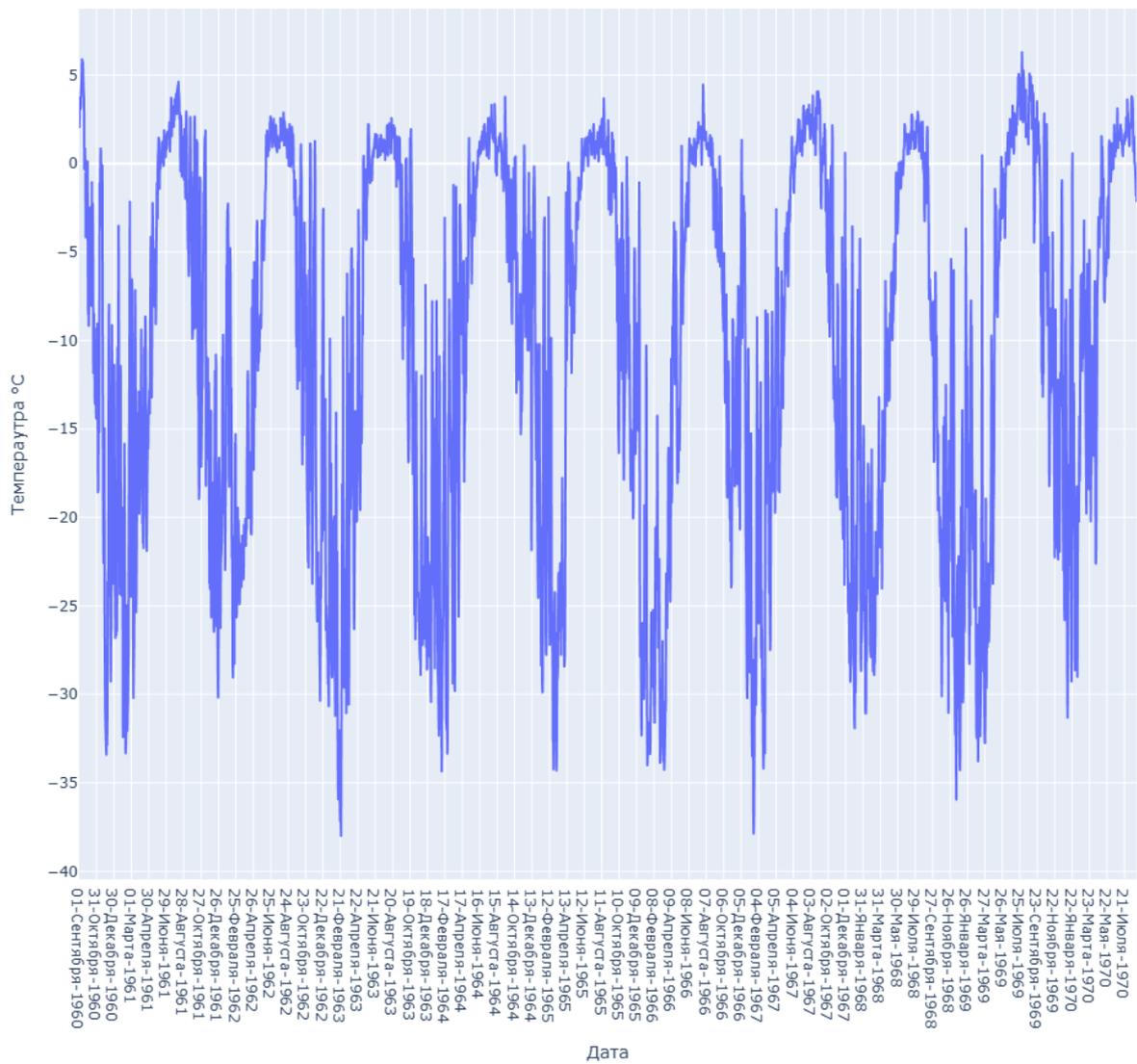


Рисунок 4.2.1.4 – Пример визуализации, полученной с помощью второго скрипта для фиксированной долготы, фиксированных дат и 80° северной широты

График температуры с 01-Сентября-1960 по 01-Сентября-1970 для точки с широтой 75°с.ш. и долготой 10°

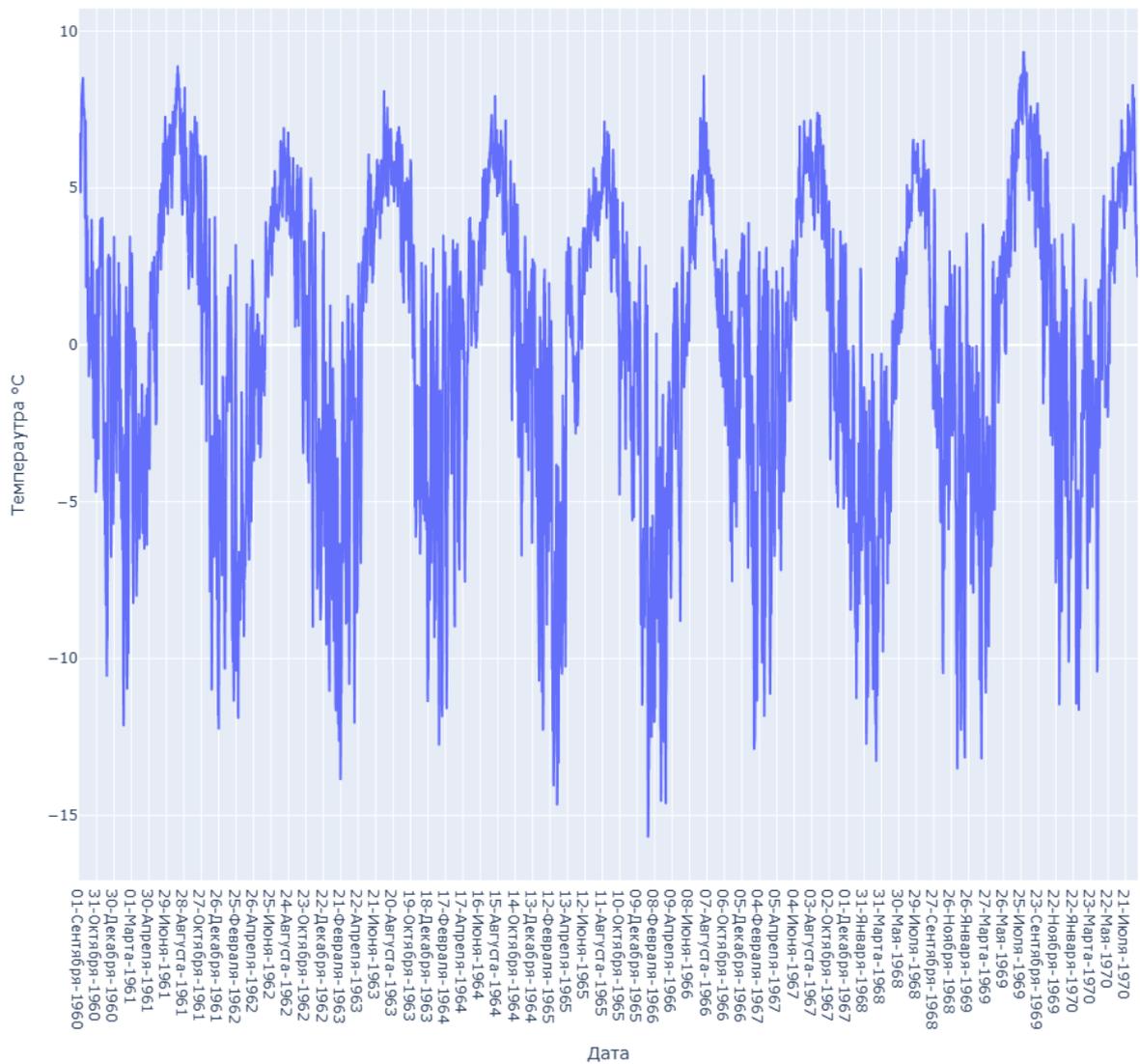


Рисунок 4.2.1.5 – Пример визуализации, полученной с помощью второго скрипта для фиксированной долготы, фиксированных дат и 75° северной широты

График температуры с 01-Сентября-1960 по 01-Сентября-1970 для точки с широтой 70°с.ш. и долготой 10°

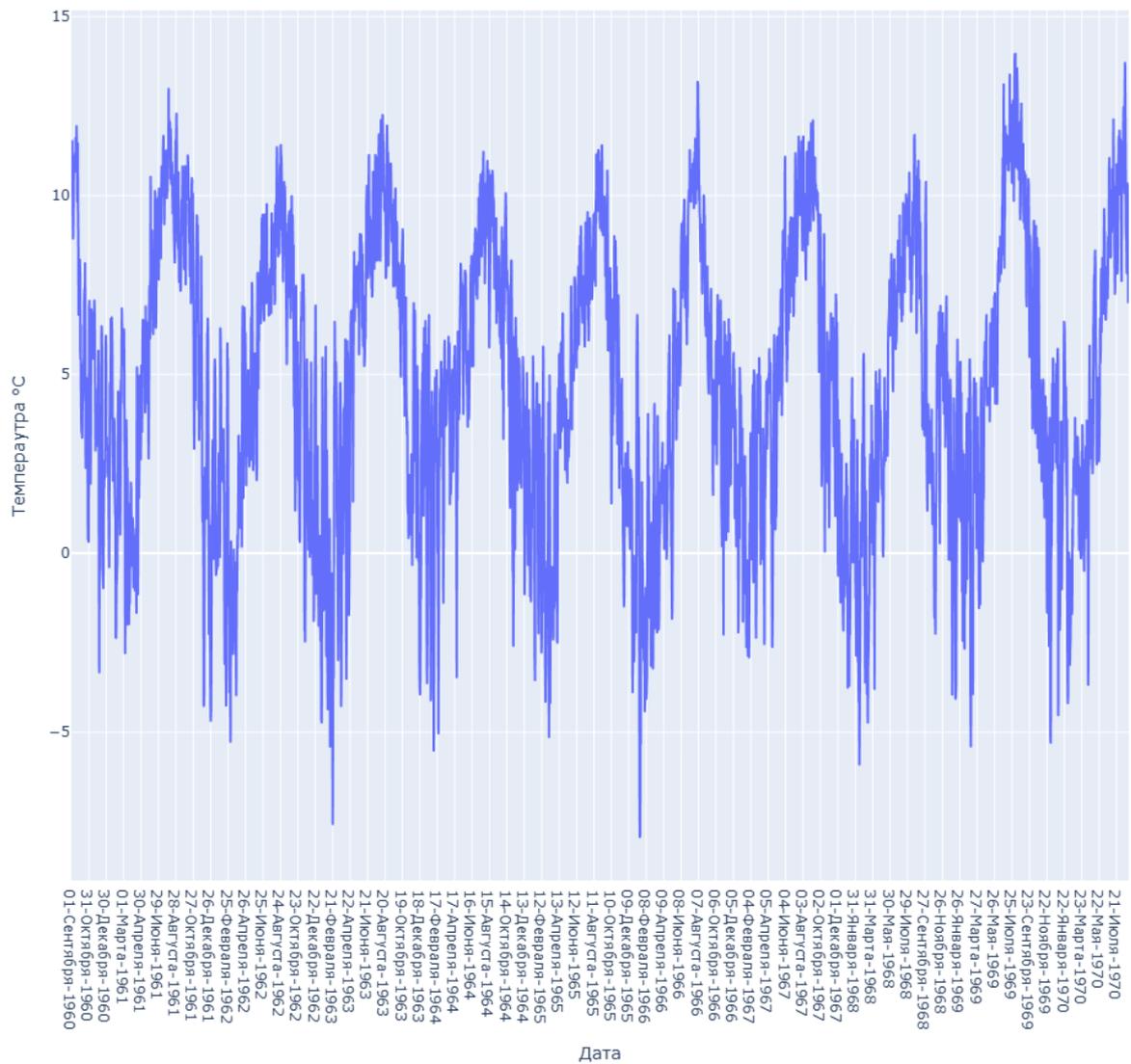


Рисунок 4.2.1.6 – Пример визуализации, полученной с помощью второго скрипта для фиксированной долготы, фиксированных дат и 70° северной широты

График средней многолетней температуры с 01-Января по 31-Декабря для точки с широтой 80°с.ш. и долготой 100°

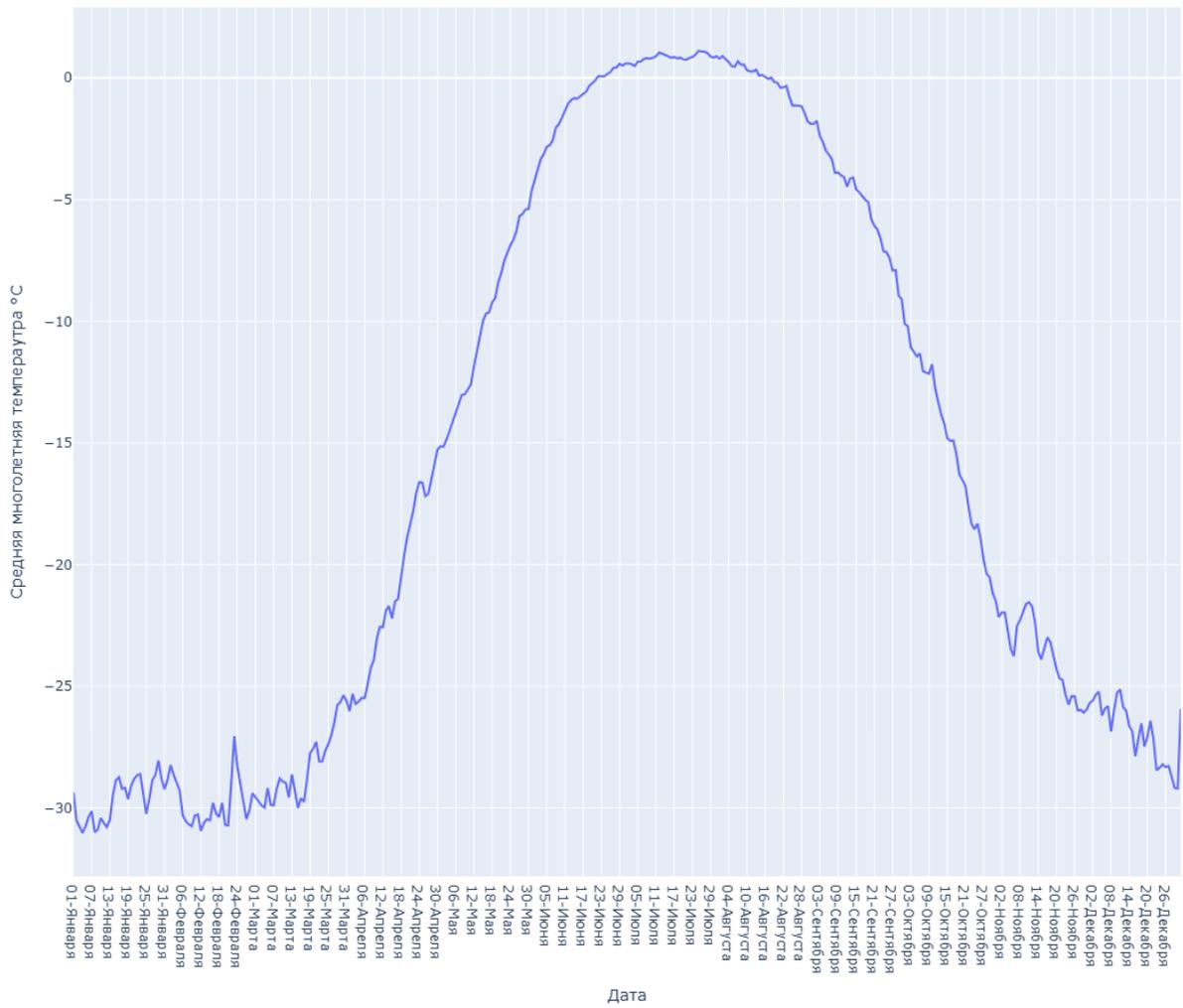


Рисунок 4.2.1.7 – Демонстрация возможности первого скрипта визуализировать весь ряд данных для выбранной точки

График средней многолетней температуры с 01-Сентября по 31-Декабря для точки с широтой 90°с.ш. и долготой 10°

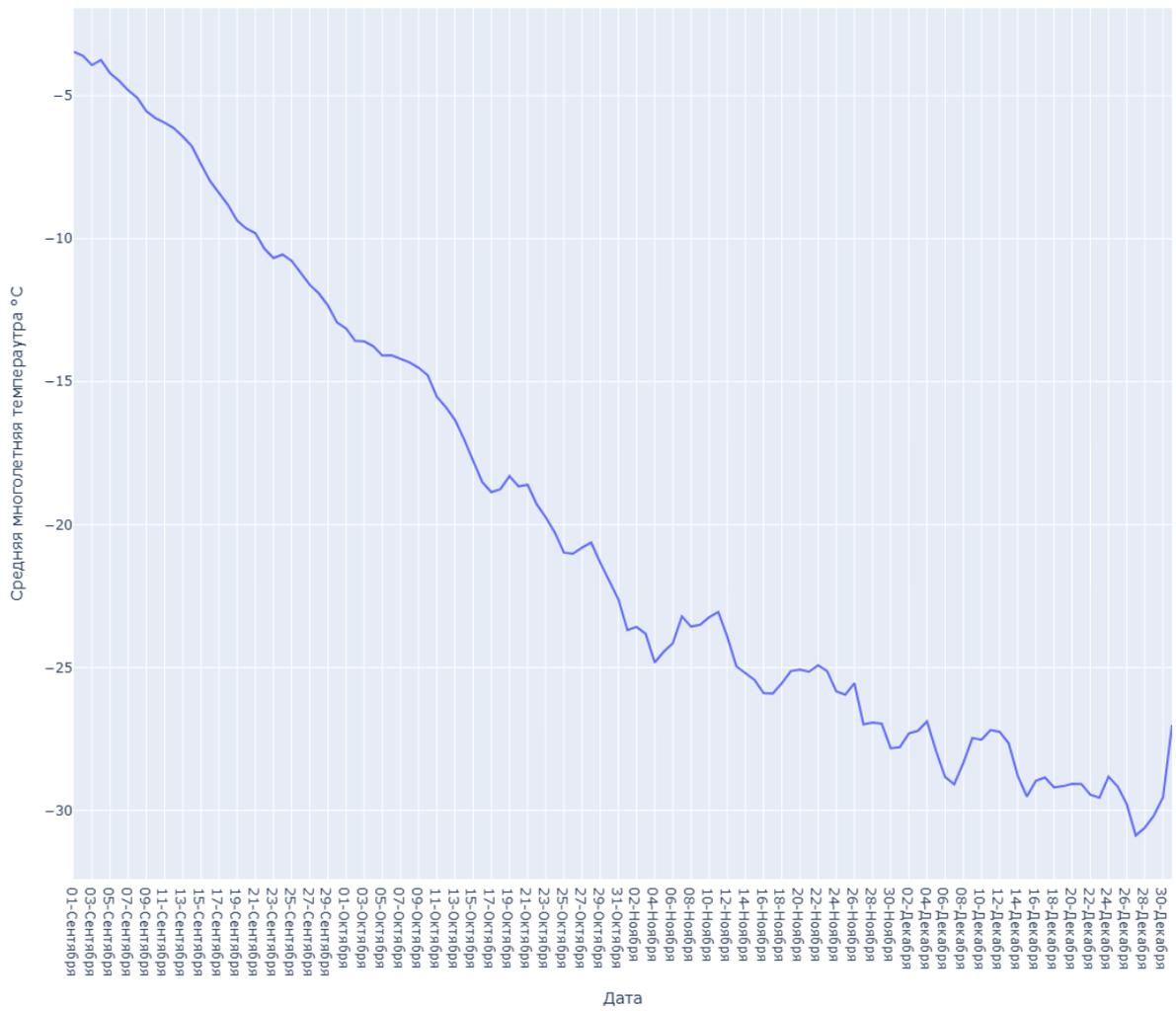


Рисунок 4.2.1.8 – Пример визуализации, полученной с помощью первого скрипта для фиксированной долготы, фиксированных дат и 90° северной широты

График средней многолетней температуры с 01-Сентября по 31-Декабря для точки с широтой 85°с.ш. и долготой 10°

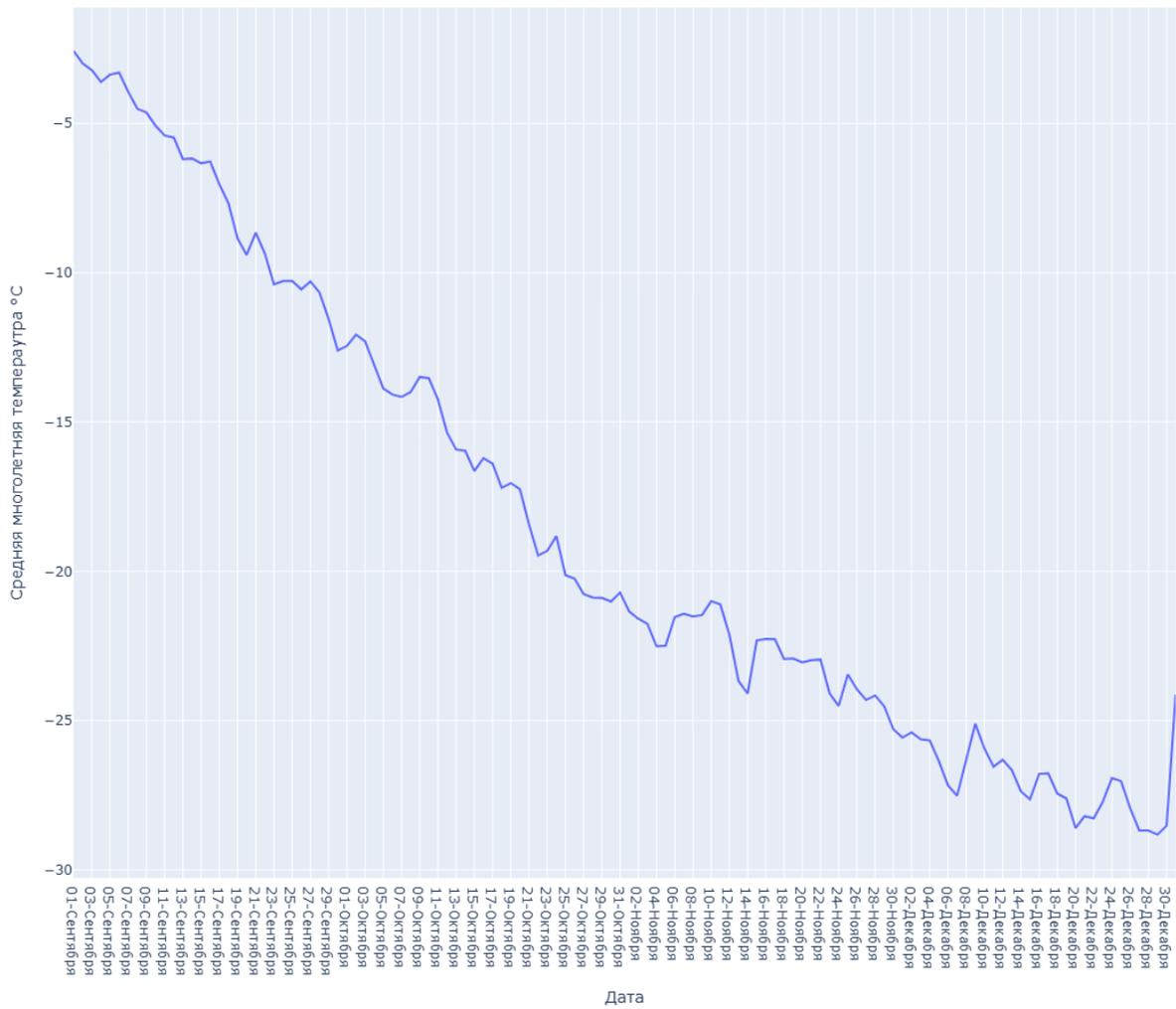


Рисунок 4.2.1.9 – Пример визуализации, полученной с помощью первого скрипта для фиксированной долготы, фиксированных дат и 85° северной широты

График средней многолетней температуры с 01-Сентября по 31-Декабря для точки с широтой 80°с.ш. и долготой 10°

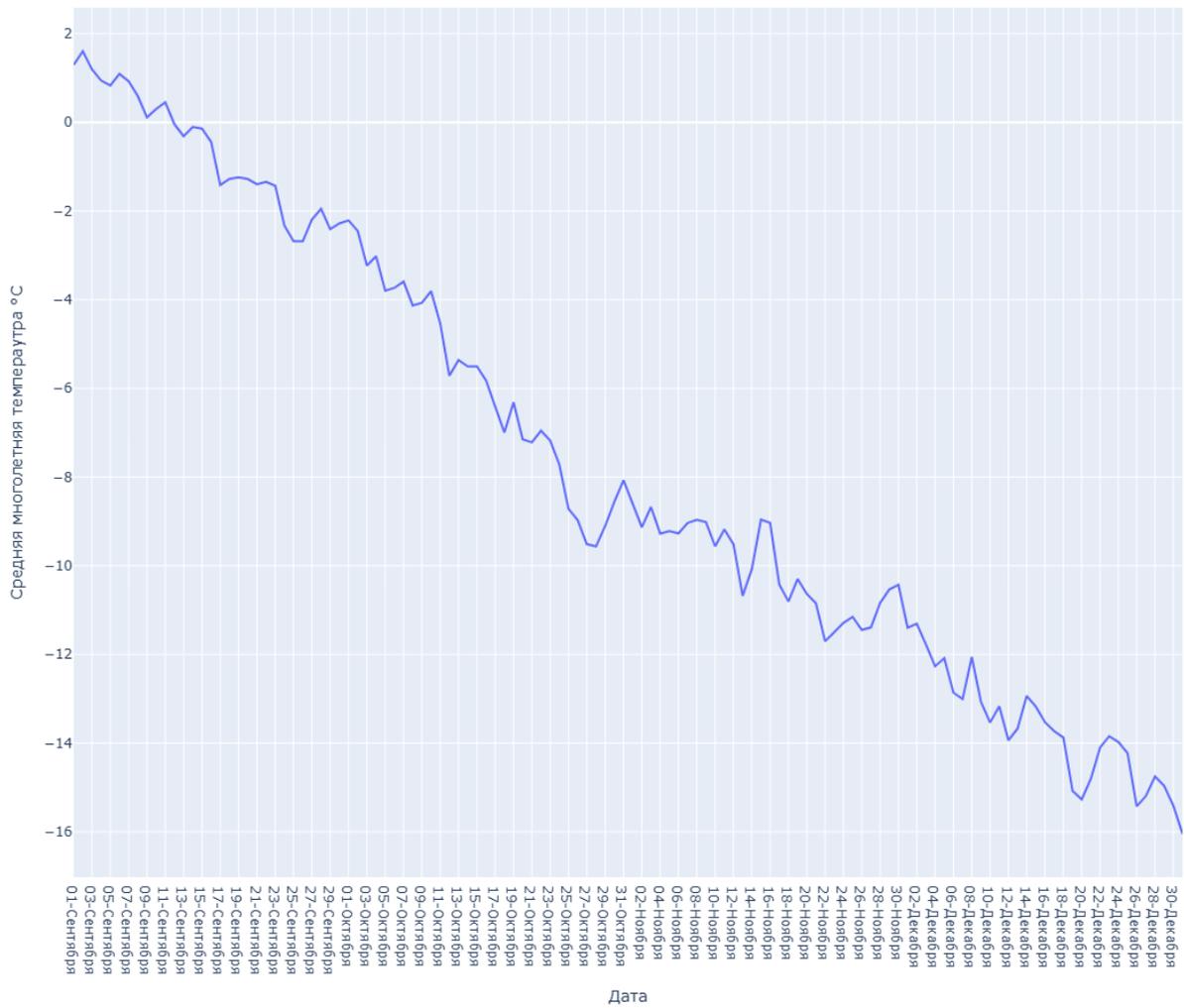


Рисунок 4.2.1.10 – Пример визуализации, полученной с помощью первого скрипта для фиксированной долготы, фиксированных дат и 80° северной широты

График средней многолетней температуры с 01-Сентября по 31-Декабря для точки с широтой 75°с.ш. и долготой 10°

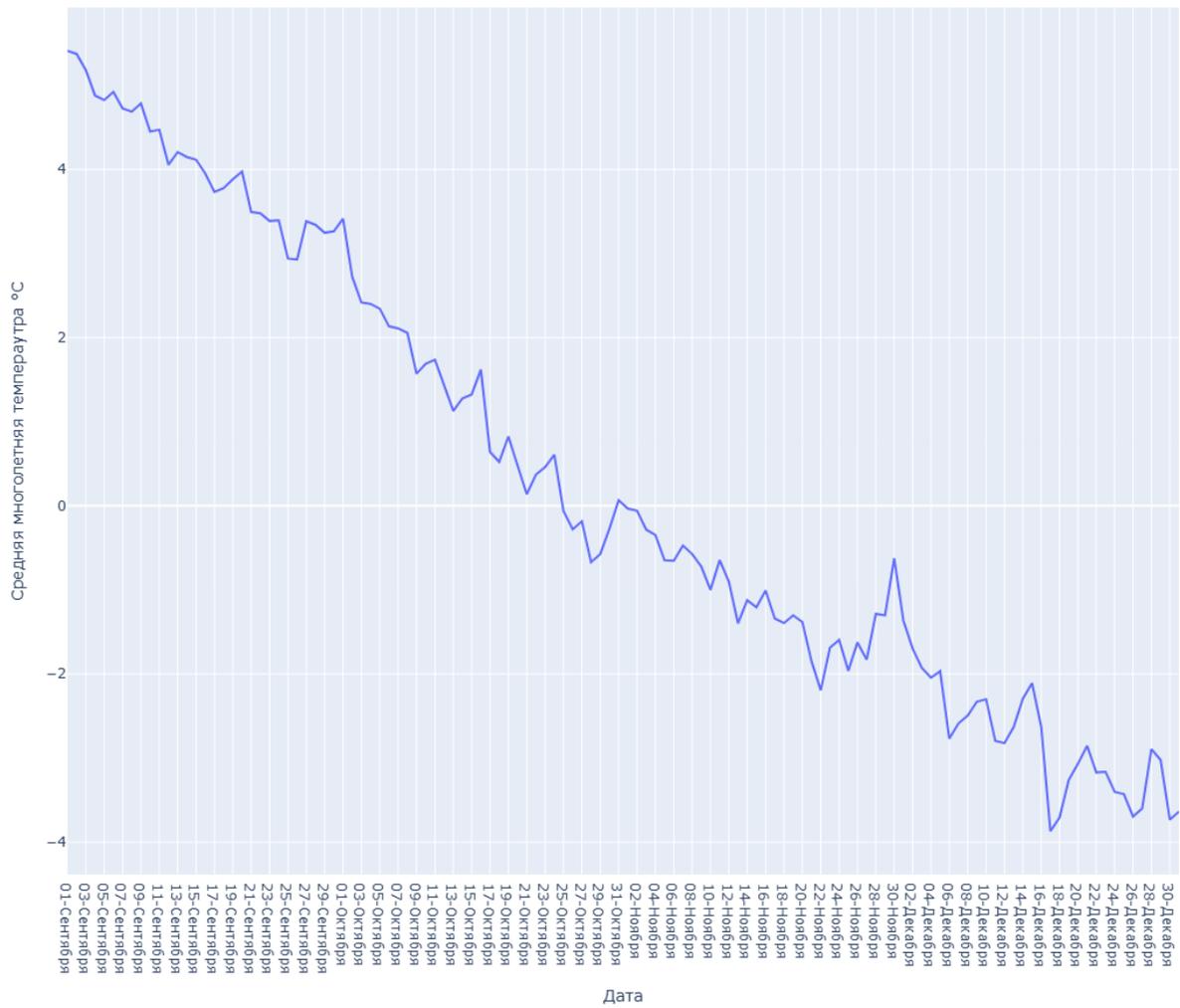


Рисунок 4.2.1.11 – Пример визуализации, полученной с помощью первого скрипта для фиксированной долготы, фиксированных дат и 75° северной широты

График средней многолетней температуры с 01-Сентября по 31-Декабря для точки с широтой 70°с.ш. и долготой 10°

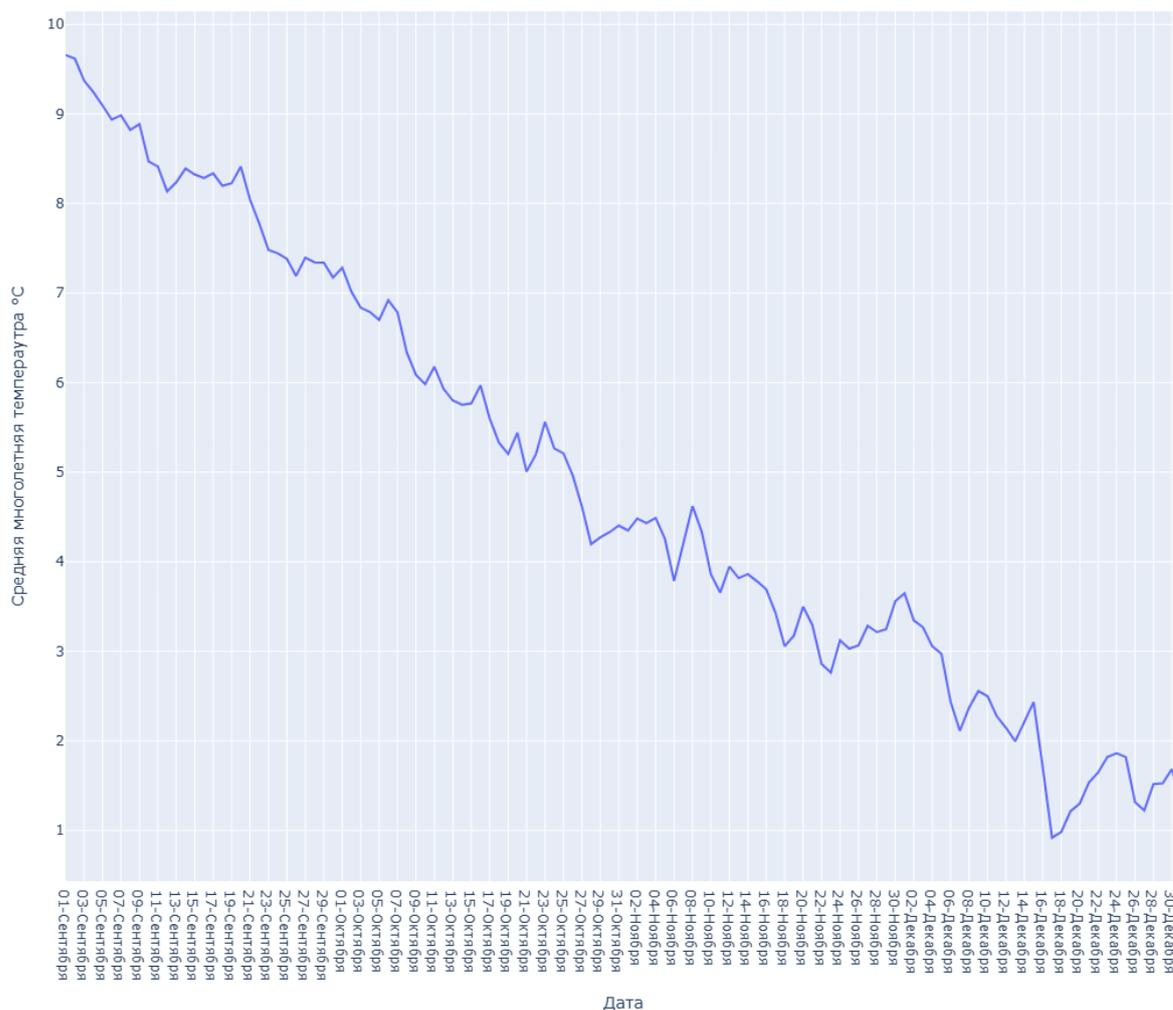


Рисунок 4.2.1.12 – Пример визуализации, полученной с помощью первого скрипта для фиксированной долготы, фиксированных дат и 70° северной широты

4.2.2 Сравнение полученных результатов с ожидаемыми

В процессе тестирования функциональности разработанных алгоритмов для обработки и анализа геопространственных данных о температуре воздуха в Арктическом регионе было установлено, что полученные результаты в целом соответствуют первоначальным требованиям и ожиданиям. Основная задача алгоритмов заключалась в корректной обработке данных из выбранного источника и их последующей визуализации в виде графиков изменения

температуры по заданным координатам и временным периодам. Данная задача была успешно выполнена.

Графические представления, представленные выше, демонстрируют логичные и предсказуемые климатические тренды, характерные для арктического региона, включая сезонные колебания температуры. Это подтверждает корректность разработанных алгоритмов, включающих фильтрацию данных по географическим координатам, агрегацию и осреднение температурных значений, а также правильное отображение временных рядов.

Однако следует отметить, что проверка работоспособности и апробация модуля были ограничены использованием единственного набора данных – ERA-40. Как было подробно рассмотрено в пунктах 2.3 «Исследование и выбор источника гидрометеорологических данных» и 3.2 «Использованный источник данных», а также упомянуто в методике тестирования (пункт 4.1 «Методика тестирования разработанного модуля»), выбор данного источника был обусловлен рядом факторов, включая его доступность и уровень полноты.

В рамках исследования планировалось провести сравнительный анализ и тестирование алгоритма на нескольких различных архивах гидрометеорологической информации для более комплексной оценки его универсальности и адаптивности. Однако в связи с техническими трудностями и проблемами, возникшими при получении стабильного и оперативного доступа к альтернативным наборам геопространственных данных для Арктического региона, реализация данной части тестирования не была завершена в полном объеме.

Таким образом, хотя работа алгоритма на данных ERA-40 продемонстрировала ожидаемые и корректные результаты, его поведение и эффективность при использовании с другими источниками данных требуют дополнительного изучения в рамках будущих работ.

4.3 Анализ производительности созданных скриптов

Первый скрипт предназначен для визуализации среднесуточной температуры воздуха на основе заданных координат и временного интервала.

Загрузка и интеграция данных: наиболее ресурсоемкой операцией на начальном этапе является последовательное чтение 46 CSV-файлов (от y1957.csv до y2002.csv) с использованием функции `pd.read_csv()`. Каждый вызов данной функции включает операции ввода-вывода файлов, что приводит к накоплению задержек.

Последующее объединение полученных `DataFrame` в один большой `DataFrame` (`df_1957_2002_ct = pd.concat(df_1957_2002)`) также является ресурсоемким процессом, особенно при таком значительном объеме данных, так как производительность операции `pd.concat()` зависит от количества и объема объединяемых `DataFrame`.

Трансформация и фильтрация: конвертация температуры из шкалы Кельвина в градусы Цельсия (`df_1957_2002_ct['PARAM24'] -= 273.15`) осуществляется как векторизованная операция над столбцом `DataFrame` и характеризуется высокой скоростью выполнения.

Фильтрация данных по географическим координатам (широта и долгота) осуществляется эффективно для `DataFrame` среднего размера.

Обработка временных данных и агрегация: удаление года из временного столбца (`df_1957_2002_ct_prntd_lat_lon_minyr['Дата'].str[:-5]`) осуществляется как строковая операция, применяемая к каждой ячейке столбца, что происходит относительно медленно при таком большом количестве строк.

Осреднение данных по датам с использованием метода `groupby` (`groupby('Дата', as_index=False)['Температура °C'].mean()`) является ключевой операцией. Эффективность этого метода в `Pandas` высока, хотя время выполнения прямо пропорционально объему группируемого `DataFrame`, эта операция занимает немного времени.

Упорядочивание дат с использованием `pd.Categorical` и `sort_values()` необходимо для корректного отображения временного ряда. Данная операция также вносит небольшой вклад во время обработки из-за небольшого количества уникальных дат (в данном случае не более 366).

Перевод месяцев на русский язык через списковое включение и замену строк также не сильно увеличивает временные затраты, пропорциональные количеству дат.

Визуализация: построение графика с использованием `plotly.express.line()` выполняется быстро для умеренного количества точек (до нескольких сотен или тысяч). В данном скрипте количество точек на графике определяется временным интервалом, выбранным пользователем (максимум 366 точек), что обеспечивает высокую скорость визуализации.

Второй скрипт осуществляет визуализацию фактических значений температуры воздуха для заданных географических координат и временного интервала, выбранного пользователем (с указанием года).

Загрузка и объединение данных: как и в первом скрипте, загрузка и объединение 46 CSV-файлов являются наиболее ресурсоемкими операциями на начальном этапе.

Преобразование и фильтрация: процесс преобразования данных в формат температуры и фильтрации по координатам идентичен первому скрипту.

Обработка дат: для локализации формата даты (перевода названий месяцев на русский язык) используется метод `df['DATETIME'].apply(replace_months)`. Данный подход менее эффективен по сравнению с векторизованными строковыми операциями библиотеки `Pandas`, поскольку метод по умолчанию выполняет итерацию по строкам. В случае работы с большим объемом данных это станет узким местом производительности.

Выбор периода и визуализация: фильтрация `DataFrame` по выбранному временному диапазону осуществляется с использованием метода `loc[start_index:end_index]`. Данный метод эффективен, так как индексы `DataFrame`

(после соответствующей обработки) находятся в упорядоченном состоянии. Время построения графика с использованием функции `plotly.express.line()` зависит от длительности выбранного временного периода. При выборе продолжительного временного интервала (например, несколько лет или десятилетий) количество отображаемых точек на графике может значительно увеличиться, что приводит к небольшому увеличению времени рендеринга графика.

Разработанные программные скрипты демонстрируют удовлетворительную производительность при анализе данных, извлечённых из архива ERA-40 за период с 1957 по 2002 годы. Скрипты могут обрабатывать информацию как по одной, так и по нескольким точкам данных. Выполнение анализа на современном персональном компьютере, на котором тестировался разрабатываемый модуль, осуществляется в приемлемые сроки.

Основные временные затраты связаны с процессами загрузки и начальной интеграции большого объёма данных, представленных в формате CSV. Операции фильтрации, агрегации и визуализации данных для ограниченного числа точек и временных периодов выполняются с достаточной скоростью.

В случае интеграции скриптов в интерактивные системы, где требуется высокая скорость отклика, особенно при частом выполнении запросов, будет необходимо рассмотреть оптимизацию процедур загрузки данных. Потребуется внедрение механизмов кэширования данных и предварительной обработки информации на стороне сервера или в системе хранения данных, с которой взаимодействует данный модуль.

Тем не менее, для задач, требующих однократного или эпизодического анализа данных по конкретным точкам, текущая производительность программных скриптов является достаточной.

4.4 Обсуждение возможностей практического применения модуля в составе интерактивной системы

Возможности разработанного модуля в области визуализации данных играют ключевую роль в преобразовании первичных исторических данных о температуре в информацию, имеющую практическое значение. Система, предоставляющая интуитивно понятные визуальные представления, позволяет пользователям выявлять критические закономерности, аномалии и тенденции, которые сложно идентифицировать при анализе табличных данных. Это способствует более обоснованному и эффективному принятию решений.

Возможные сценарии применения модуля

- Историческая информация о температуре может быть использована для анализа типичных моделей формирования и таяния льда в различные сезоны. Это позволяет оптимизировать и повысить безопасность судоходных маршрутов, исключая районы с исторически низкими температурами. Учет исторических температурных трендов способствует долгосрочному стратегическому планированию в судоходстве.
- Несмотря на то, что модуль ориентирован на анализ температуры воздуха, исторические данные могут коррелировать с состоянием морского льда. Более низкие температуры в определенные периоды и в определённых регионах могут указывать на повышенную вероятность наличия льда. Визуализация исторических температурных трендов позволяет прогнозировать потенциальные изменения ледовой обстановки.
- Модуль предоставляет исследователям ценные исторические данные о температуре, которые могут быть использованы для анализа тенденций потепления в Арктическом регионе. Долгосрочный анализ температурных данных способствует пониманию воздействия климатических изменений на арктическую экосистему. Модуль может быть применен для верификации климатических моделей и улучшения точности будущих прогнозов.

Помимо температуры воздуха, модуль может обрабатывать и визуализировать данные о: температуре поверхности моря; концентрации и протяженности морского льда; скорости и направлении ветра; осадках (снегопаде и дожде); влажности; атмосферном давлении; океанских течениях; высоте и периоде волн; видимости; толщине снежного покрова; солнечной радиации; солености; приливах и уровнях воды.

Разработанный модуль может быть адаптирован для удовлетворения потребностей широкого спектра пользователей, имеющих различные требования. Судоходные компании акцентируют внимание на обеспечении безопасности и эффективности навигации, в то время как рыболовные суда заинтересованы в данных, влияющих на состояние рыбных запасов и эксплуатационные риски. Научные экспедиции требуют предоставления детализированной информации для проведения исследований. Понимание этих разнообразных потребностей является ключевым фактором для проработки удобства и функциональности интерактивной системы. Различные группы пользователей имеют различные приоритеты и контексты применения. Например, рыболовные суда могут быть заинтересованы в краткосрочных прогнозах температуры поверхности моря, в то время как судоходные компании, планирующие маршруты на несколько месяцев вперед, будут сосредоточены на долгосрочных тенденциях изменения ледовой обстановки. Дизайн модуля может учитывать эти различные требования.

Разработанный модуль, интегрированный в интерактивную систему, предоставляет пользователям удобный интерфейс для доступа к историческим гидрометеорологическим данным. Система позволяет осуществлять выбор интересующих координат и временных периодов, и получать визуализированные данные, удобные для интерпретации исторических тенденций и закономерностей.

Система может быть адаптирована к специфическим потребностям различных групп пользователей, предлагая возможность выбора соответствующих слоев данных и параметров визуализации. А интеграция с внешними информационными ресурсами, включая навигационные и ледовые карты, расширит её функциональные возможности.

4.5 Перспективы дальнейшего развития модуля

Модуль обработки и анализа геопространственных данных, разработанный в рамках данной работы, является ключевым компонентом интерактивной автоматизированной системы, предназначенной для обеспечения долгосрочной гидрометеорологической информацией, необходимой для морской деятельности в Арктическом регионе.

Текущая реализация модуля, включающая два скрипта для анализа и визуализации температурных характеристик атмосферного воздуха, сформировала базовую основу для проведения определенных видов анализа. Однако функционал модуля обладает значительным потенциалом для дальнейшего развития и улучшения в нескольких ключевых аспектах.

Расширение спектра обрабатываемых гидрометеорологических параметров

В настоящее время функционал модуля ограничен обработкой данных о температуре воздуха. Для обеспечения полноценного информационного сопровождения морской деятельности в Арктическом регионе необходимо расширить спектр обрабатываемых геопространственных гидрометеорологических параметров.

Конкретный перечень возможных дополнительных данных представлен в пункте 4.4.

Интеграция данных из различных источников

Текущая версия модуля осуществляет обработку данных, полученных из архива ERA-40 в формате CSV. В перспективе целесообразно рассмотреть вопрос о расширении источников данных за счет включения информации из других глобальных и региональных архивов, таких как ERA5, спутниковых систем, а также данных наблюдений, проводимых на полярных станциях. Кроме того, возможна интеграция данных из оперативных информационных источников.

Это повысит актуальность и полноту данных, доступных для анализа. Для реализации этой задачи потребуется разработка универсальных инструментов для обработки и преобразования данных, включая парсеры и адаптеры, способные работать с различными форматами данных, такими как NetCDF, GRIB, стандартные форматы метеорологических сводок и другие.

Оптимизация и рефакторинг программного кода

Существующая методика последовательной загрузки всех CSV-файлов в начале выполнения каждого скрипта демонстрирует низкую производительность, особенно при увеличении объема обрабатываемых данных и количества файлов. Необходимо будет рассмотреть переход к более эффективным методам хранения и доступа к данным. Они могут включать использование реляционных баз данных, таких как PostgreSQL с расширением PostGIS для работы с геопространственными данными, а также применение специализированных форматов файлов, оптимизированных для работы с большими объемами данных, например, NetCDF или Zarr.

В текущих реализациях скриптов наблюдаются дублирующиеся блоки кода. Для улучшения читаемости, поддерживаемости и возможности повторного использования кода необходимо будет провести рефакторинг. Это предполагает выделение отдельных модулей для функций загрузки, предобработки и фильтрации данных.

Необходимо будет внедрить надежные механизмы обработки ошибок, связанных с некорректным вводом данных пользователем, отсутствием необходимых данных, проблемами с файлами и другими подобными ситуациями.

Разработка расширенных аналитических функций

Помимо вычисления средних многолетних значений, функциональные возможности модуля могут быть усовершенствованы для проведения более детального анализа, востребованного в сфере морской деятельности: оценка долгосрочных изменений исследуемых параметров в Арктическом регионе; определение отклонений текущих или усредненных значений от стандартных показателей; проведение расчетов дисперсии, стандартного отклонения, минимальных и максимальных значений, а также перцентилей для оценки вариабельности параметров; реализация функций пространственного анализа, включая построение карт распределения параметров по акватории на заданный момент времени или за определенный период; обеспечение возможности сопоставления параметров для различных точек, регионов или временных периодов на единой графической платформе.

Оптимизация и расширение функционала визуализации

Графические интерфейсы, созданные с использованием Plotly, уже обеспечивают базовую интерактивность. Планируется дальнейшее развитие пользовательского взаимодействия с графиками, включая улучшение возможностей масштабирования, панорамирования и отображения детализированной информации при наведении курсора.

В дополнение к существующим графикам, планируется внедрение новых типов визуализации, оптимизированных для анализа метеорологических данных. К ним относятся тепловые карты для отображения суточной или годовой динамики параметров, диаграммы «роза ветров» и картографические представления распределения метеорологических характеристик.

Также планируется расширение возможностей механизмов экспорта визуализированных данных и результатов анализа в различные форматы, включая графические (PNG, JPEG, SVG) и табличные (CSV, Excel).

Интеграция в общую архитектуру системы

Для обеспечения бесперебойной работы в составе единой интерактивной автоматизированной системы модуль должен быть адаптирован для взаимодействия с веб-интерфейсом и другими элементами системы. Это может включать выполнение следующих задач: разработка программного интерфейса приложения (API), который позволит веб-серверу осуществлять передачу запросов к модулю (например, координат, дат, выбранных параметров) и получать результаты обработки в стандартизованном формате (например, JSON); обеспечение асинхронной обработки запросов для предотвращения блокирования веб-интерфейса при выполнении ресурсоемких вычислений.

Заключение

Настоящая работа посвящена актуальной задаче разработки модуля обработки и анализа геопространственных данных для интерактивной автоматизированной системы обеспечения долгосрочной гидрометеорологической информацией морской деятельности в Арктике.

Стратегическое значение Арктической зоны для Российской Федерации и увеличение интенсивности морских операций в условиях сложных и изменяющихся природно-климатических условий обуславливают необходимость в точной и своевременной гидрометеорологической информации. Существующие системы характеризуются ограниченным удобством доступа и визуализации долгосрочных геопространственных данных, особенно для пользователей в Российской Федерации, что существенно затрудняет проведение научного анализа, долгосрочного планирования и принятия обоснованных решений.

Проведен анализ значимости гидрометеорологической информации для Арктики, выявлены особенности региона и ключевые гидрометеорологические параметры. Рассмотрены существующие российские и зарубежные информационные системы, выявлены их недостатки, связанные с доступом и визуализацией долгосрочных данных. Показана роль геопространственных данных и обзор современных технологий их обработки и визуализации.

Определены функциональные требования к модулю, включая прием входных параметров, доступ к данным, предварительную обработку, анализ и визуализацию. Обосновано место модуля в архитектуре ИАС как ключевого компонента серверной части. Проведен сравнительный анализ языков программирования и обоснован выбор Python и его библиотек для геопространственного анализа и визуализации. Исследованы доступные архивы гидрометеорологических данных для Арктики и аргументирован выбор ERA-40

для тестирования. Сформулирована концепция алгоритмов обработки и визуализации на основе выбранных методов.

Реализована программная система на Python, включающая функции загрузки, объединения, преобразования и фильтрации температурных данных. Созданы два скрипта: один для расчета и визуализации средних многолетних температур, другой — для отображения фактических значений температуры за заданный период. Реализована локализация дат и названий месяцев на русский язык. Использован источник данных ERA-40.

Разработана методика тестирования модуля с использованием данных ERA-40 и определены критерии оценки корректности работы. Проведена апробация, показавшая, что полученные результаты соответствуют ожиданиям и демонстрируют логичные климатические тренды. Проанализирована производительность скриптов, выявлены наиболее ресурсоемкие операции, связанные с загрузкой данных. Обсуждены возможности практического применения модуля для обеспечения безопасности судоходства, анализа ледовой обстановки, климатических исследований и верификации моделей.

Разработанный модуль может быть успешно интегрирован в интерактивную автоматизированную систему для предоставления пользователям визуализированной исторической гидрометеорологической информации по Арктическому региону. Он может использоваться для:

- Анализа среднемноголетних температурных условий для планирования морских операций и оценки сезонной доступности маршрутов.
- Исследования фактической динамики температуры в конкретной точке за интересующий исторический период.
- Выявления климатических трендов и аномалий температуры в Арктике.
- Обеспечения информационной поддержки при принятии решений, связанных с безопасностью и эффективностью морской деятельности в условиях Арктики.

Для дальнейшего развития модуля и повышения его функциональности и эффективности предлагаются следующие направления:

- Расширение спектра обрабатываемых гидрометеорологических параметров, включив данные о морском льде, ветре, волнении, осадках и других критически важных для Арктики параметрах.
- Интеграция данных из различных источников, помимо ERA-40, включая более современные реанализы (например, ERA5), спутниковые данные и данные наблюдений с метеорологических станций, а также оперативные данные.
- Оптимизация процедур загрузки и хранения данных, возможно, с использованием баз данных (например, PostgreSQL с PostGIS) или форматов, оптимизированных для больших данных (NetCDF, Zarr), для повышения производительности.
- Рефакторинг программного кода с выделением модулей и внедрением механизмов обработки ошибок для улучшения структуры и надежности.
- Разработка расширенных аналитических функций, таких как расчет отклонений от нормы, пространственный анализ распределения параметров и сопоставление данных для различных регионов.
- Оптимизация и расширение функционала визуализации, включая добавление новых типов графиков (тепловые карты, розы ветров) и улучшение интерактивности.
- Разработка API для бесшовной интеграции модуля в общую архитектуру интерактивной автоматизированной системы и обеспечения асинхронной обработки запросов.

Таким образом, цель работы, заключающаяся в разработке модуля для обработки и анализа геопространственных данных для интерактивной

автоматизированной системы обеспечения долгосрочной гидрометеорологической информацией морской деятельности в Арктике, достигнута. Созданный модуль способен загружать, обрабатывать, фильтровать и визуализировать долгосрочные температурные данные для заданных географических координат и временных периодов, предоставляя пользователю наглядные графические представления.

Список использованной литературы

1. Зеленина Лариса Ивановна, Федькушова Светлана Ивановна Прогнозирование и последствия изменения климата Арктического региона // АИС. 2012. №5. URL: <https://cyberleninka.ru/article/n/prognozirovanie-i-posledstviya-izmeneniya-klimata-arkticheskogo-regiona> (дата обращения: 15.04.2025).
2. Быковский Вадим Кириллович Арктическая зона Российской Федерации: правовой режим развития, обеспечения безопасности, охраны природы, защиты национальных интересов // Международное сотрудничество евразийских государств: политика, экономика, право. 2019. №1. URL: <https://cyberleninka.ru/article/n/arkticheskaya-zona-rossiyskoy-federatsii-pravovoy-rezhim-razvitiya-obespecheniya-bezopasnosti-ohrany-prirody-zaschity-natsionalnyh> (дата обращения: 15.04.2025).
3. К В. Фильчук, В Б. Коробов, А В. Юлин, Т В. Шевелева ВЛИЯНИЕ НАБЛЮДАЕМЫХ ИЗМЕНЕНИЙ КЛИМАТИЧЕСКИХ УСЛОВИЙ НА ХОЗЯЙСТВЕННУЮ ДЕЯТЕЛЬНОСТЬ В МОРЯХ РОССИЙСКОЙ АРКТИКИ // Российская Арктика. 2022. №2 (17). URL: <https://cyberleninka.ru/article/n/vliyanie-nablyudaemyh-izmeneniy-klimaticheskikh-usloviy-na-hozyaystvennuyu-deyatelnost-v-moryah-rossiyskoy-arktiki> (дата обращения: 15.04.2025).
4. Шипко Ю. В., Шувакин Е. В., Шуваев М. А. Регрессионные модели оценки безопасности работ персонала на открытой территории в жестких погодных условиях // Воздушно-космические силы. Теория и практика. 2017. №1 (1). URL: <https://cyberleninka.ru/article/n/regressionnyye-modeli-otsenki-bezopasnosti-rabot-personala-na-otkrytoy-territorii-v-zhestkih-pogodnyh-usloviyah> (дата обращения: 15.04.2025).
5. О службе / [Электронный ресурс] // РОСГИДРОМЕТ : [сайт]. — URL: <https://www.meteorf.gov.ru/about/service> (дата обращения: 16.04.2025).

6. About our agency / [Электронный ресурс] // National Oceanic and Atmospheric Administration : [сайт]. — URL: <https://www.noaa.gov/about-our-agency> (дата обращения: 16.04.2025).
7. About us / [Электронный ресурс] // ECMWF 50 : [сайт]. — URL: <https://www.ecmwf.int/en/about> (дата обращения: 16.04.2025).
8. Метеум — собственная технология прогноза Яндекс Погоды / [Электронный ресурс] // Я.Погода : [сайт]. — URL: https://yandex.ru/pogoda/technology?utm_source=main_page (дата обращения: 16.04.2025).
9. Меню / [Электронный ресурс] // Windy : [сайт]. — URL: <https://www.windy.com/ru/-Меню/menu> (дата обращения: 16.04.2025).
10. Three types of satellite imagery / [Электронный ресурс] // National Weather Service : [сайт]. — URL: <https://www.weather.gov/mrx/sattype> (дата обращения: 16.04.2025).
11. ЖУКОВ Ю. Н., БАХМУТОВ В. Ю. СИСТЕМЫ ХРАНЕНИЯ ГЛОБАЛЬНОЙ ГЕОПРОСТРАНСТВЕННОЙ ИНФОРМАЦИИ //Навигация и гидрография. – 2018. – С. 52.
12. Жукова Н. А., Смирнова О. В. Методы адаптивной обработки и анализа метеорологических и океанографических данных в Арктическом регионе //Информатика и автоматизация. – 2013. – Т. 7. – №. 30. – С. 258-272.
13. Матишов Г. Г. и др. Совместные подходы ММБИ, ЮНЦ и Лаборатории климата океана НОАА (США) к организации океанографических и гидробиологических баз данных арктических и южных морей России //Труды Кольского научного центра РАН. – 2013. – №. 1 (14). – С. 135-152.
14. Абрамов Н. С. и др. Современные методы интеллектуальной обработки данных ДЗЗ //Программные системы: теория и приложения. – 2018. – Т. 9. – №. 4 (39). – С. 417-442.
15. Наумова Т. Е., Михайлова Д. А. Международная научно-практическая конференция «Проблемы предупреждения и ликвидации чрезвычайных ситуаций и создание комплексных аварийно-спасательных

центров в Арктике»: Материалы конференции/МЧС России. М.: ФГБУ ВНИИ ГОЧС (ФЦ), 2012. 196 с.: ил. – 2012.

16. Remote Sensing / [Электронный ресурс] // EARTHDATA : [сайт]. — URL: <https://www.earthdata.nasa.gov/learn/earth-observation-data-basics/remote-sensing> (дата обращения: 17.04.2025).

17. Prof. Dr. Peter Kogut Spatial Analysis: Data Processing And Use Cases / Prof. Dr. Peter Kogut [Электронный ресурс] // EOS DATA ANALYTICS : [сайт]. — URL: <https://eos.com/blog/spatial-analysis/> (дата обращения: 17.04.2025).

18. GIS Tools and Resources / [Электронный ресурс] // GISDegree : [сайт]. — URL: <https://www.gisdegree.org/resources/tools/> (дата обращения: 17.04.2025).

19. Spatial without Compromise / [Электронный ресурс] // QGIS : [сайт]. — URL: <https://qgis.org/> (дата обращения: 17.04.2025).

20. Comparing Top R and Python Libraries for Geospatial Data Visualization / [Электронный ресурс] // MoldStud : [сайт]. — URL: <https://moldstud.com/articles/p-comparing-top-r-and-python-libraries-for-geospatial-data-visualization> (дата обращения: 17.05.2025).

21. MATLAB vs Python: Ultimate Showdown of Programming Titans [2024] / [Электронный ресурс] // Simplilearn : [сайт]. — URL: <https://www.simplilearn.com/matlab-vs-python-article> (дата обращения: 17.05.2025).

22. Why Choose Julia Over Python (or Vice Versa)? Full Breakdown / [Электронный ресурс] // eLuminous : [сайт]. — URL: <https://eluminoustechnologies.com/blog/julia-vs-python/> (дата обращения: 17.05.2025).

23. Python mapping libraries / [Электронный ресурс] // HEX : [сайт]. — URL: <https://hex.tech/templates/data-visualization/python-mapping-libraries/> (дата обращения: 17.05.2025).

24. JavaScript Geo Library: Definition and Top 4 Geo Libraries / [Электронный ресурс] // GeoPlugin : [сайт]. — URL:

<https://www.geoplugin.com/resources/javascript-geo-library/> (дата обращения: 17.05.2025).

25. Python vs Java: Making the Right Choice for Your Software Development Project / [Электронный ресурс] // MOBILUNITY : [сайт]. — URL: <https://mobilunity.com/blog/python-vs-java/> (дата обращения: 17.05.2025).

26. c,c++,java and python in gis development / [Электронный ресурс] // slideshare : [сайт]. — URL: <https://www.slideshare.net/slideshow/ccjava-and-python-in-gis-development/135429443> (дата обращения: 17.05.2025).

27. Arctic Temperatures / [Электронный ресурс] // Zachary Labe : [сайт]. — URL: <https://zacklabe.com/arctic-temperatures/> (дата обращения: 17.05.2025).

28. Complete ERA5 global atmospheric reanalysis 1940 to present / [Электронный ресурс] // Arctic Data Center : [сайт]. — URL: <https://www.arcticdata.io/catalog/view/doi%3A10.18739%2FA2639K70N> (дата обращения: 17.05.2025).

29. ECMWF ERA-Interim. European Reanalysis / [Электронный ресурс] // Universität Hamburg : [сайт]. — URL: <https://www.cen.uni-hamburg.de/en/icdc/data/atmosphere/reanalysis-atmosphere/era-interim.html> (дата обращения: 17.05.2025).

30. NCEP-NCAR Reanalysis 1 / [Электронный ресурс] // Physical Sciences Laboratory : [сайт]. — URL: <https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html> (дата обращения: 17.05.2025).

31. Near-Surface Air Temperatures / [Электронный ресурс] // National Snow and Ice Data Center : [сайт]. — URL: <https://nsidc.org/data/soac/near-surface-air-temperatures> (дата обращения: 17.05.2025).

32. JRA-55: Japanese 55-year Reanalysis, Daily 3-Hourly and 6-Hourly Data / [Электронный ресурс] // NSF NCAR : [сайт]. — URL: <https://rda.ucar.edu/datasets/d628000/> (дата обращения: 17.05.2025).

33. Arctic System Reanalysis (ASR) / [Электронный ресурс] // NSF NCAR : [сайт]. — URL: <https://climatedataguide.ucar.edu/climate-data/arctic-system-reanalysis-asr> (дата обращения: 17.05.2025).
34. Arctic regional reanalysis on single levels from 1991 to present / [Электронный ресурс] // Climate Data Store : [сайт]. — URL: <https://cds.climate.copernicus.eu/datasets/reanalysis-carra-single-levels?tab=overview> (дата обращения: 17.05.2025).
35. GISS Surface Temperature Analysis (v4) / [Электронный ресурс] // National Aeronautics and Space Administration : [сайт]. — URL: <https://data.giss.nasa.gov/gistemp/maps/> (дата обращения: 17.05.2025).
36. Data Overview / [Электронный ресурс] // Berkeley Earth : [сайт]. — URL: <https://berkeleyearth.org/data/> (дата обращения: 17.05.2025).
37. Climate Change: Earth Surface Temperature Data / [Электронный ресурс] // Kaggle : [сайт]. — URL: <https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data/data> (дата обращения: 17.05.2025).
38. Александрова Л. В. Геоинформационная модель и концепция комплексного мониторинга прибрежных регионов на примере Финского залива : дис. – Автореферат диссертации на соискание ученой степени кандидата технических наук, 2014.–20 с, 2014.
39. Гвишиани А. Д. и др. Геоинформатика и системный анализ в геофизике и геодинамике //Физика Земли. – 2019. – №. 1. – С. 42-60.
40. Кондратов Н. А. Стратегические направления развития гидрометеорологии и смежных наук в России //География: развитие науки и образования. – 2021. – С. 112-116.
41. Макоско А. А. Гидрометеорологическое обеспечение плавания по трассам Северного морского пути //Арктика: экология и экономика. – 2013. – Т. 11. – №. 3. – С. 40.
42. Михайлов Н. Н. и др. Единая государственная система информации об обстановке в Мировом океане и ее применение для информационной

поддержки морской деятельности Российской Федерации //Труды Всероссийского научно-исследовательского института гидрометеорологической информации-Мирового центра данных. – 2014. – №. 177. – С. 95-118.

43. Кутинов Ю. Г., Копосов С. Г. Возможности и перспективы развития Центра космического мониторинга Арктики САФУ для решения задач рационального природопользования //Электронное научное издание Альманах Пространство и Время. – 2013. – Т. 4. – №. 1. – С. 19.

44. Вязилов Е. Д., Пузова Н. В. Перспективы использования средств искусственного интеллекта в гидрометеорологии.

45. Кулаков М. Ю., Демчев Д. М. Моделирование дрейфа айсбергов как часть ледового мониторинга в Западной Арктике //Метеорология и гидрология. – 2015. – №. 12. – С. 47-55.

46. Шокин Ю. И. и др. Распределенная информационно-аналитическая система для поиска, обработки и анализа пространственных данных //Вычислительные технологии. – 2007. – Т. 12. – №. S3. – С. 108-115.

47. Брежнев Р. В. и др. Развитие технологии интерактивного формирования процессов обработки и анализа данных ДЗЗ //Региональные проблемы дистанционного зондирования Земли. – 2023. – С. 9-13.

48. Ямашкин С. А., Ямашкин А. А. Интеграция, хранение и обработка больших массивов пространственно-временной информации в цифровых инфраструктурах пространственных данных //Современные наукоемкие технологии. – 2021. – №. 5. – С. 108-113.

49. Басаргин А. А., Бугаков П. Ю., Кацко С. Ю. Решение логистических задач на основе обработки геопространственных данных с помощью модуля TRACKING ANALYST в составе ARCGIS //Вестник СГУГиТ (Сибирского государственного университета геосистем и технологий). – 2019. – Т. 24. – №. 1. – С. 83-96.

50. Григорьев А. Н. и др. Современные средства автоматизации обработки геопространственных данных //Труды Военно-космической академии имени АФ Можайского. – 2012. – №. 635. – С. 66-70.

51. Горбачева Е. Н. Программный комплекс ENVI-профессиональное решение для комплексной обработки мультиспектральных, гиперспектральных и радарных данных //Геоматика. – 2013. – №. 2. – С. 50-55.
52. Ямашкин С. А., Ямашкин А. А., Федосин С. А. Разработка проектно-ориентированной инфраструктуры пространственных данных с применением облачных технологий //Радиопромышленность. – 2019. – №. 3. – С. 79-90.
53. Бондур В. Г. Современные подходы к обработке больших потоков гиперспектральной и многоспектральной аэрокосмической информации //Исследование Земли из космоса. – 2014. – №. 1. – С. 4-17.
54. Кадочников А. А., Якубайлик О. Э. Сервис-ориентированные веб-системы для обработки геопро пространственных данных //Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2015. – Т. 13. – №. 1. – С. 37-45.
55. Шокин Ю. И., Потапов В. П. ГИС сегодня: состояние, перспективы, решения //Вычислительные технологии. – 2015. – Т. 20. – №. 5. – С. 175-213.
56. Истомин Е. П. и др. Анализ моделей и систем обработки гетерогенных данных для использования в прикладных ГИС //Научный вестник. – 2015. – №. 4. – С. 53.
57. Матин С. В., Круглов А. Н. Разработка клиентского модуля для анализа и визуализации пространственных данных в картографических веб-подсистемах сервис-ориентированных корпоративных систем управления //Энергетика. Инновационные направления в энергетике. CALS-технологии в энергетике. – 2014. – №. 1. – С. 316-324.
58. Алешко Р. А. и др. Технологии непрерывной обработки пространственных данных //Состояние арктических морей и территорий в условиях из. – 2014. – С. 84.
59. Кошкарев А. В. и др. Виртуальная ГИС-лаборатория как инструмент анализа пространственных данных //Труды Института системного программирования РАН. – 2012. – Т. 23. – С. 245-256.

60. Шокин Ю. И. и др. Информационная система приема, обработки и доступа к спутниковым данным и ее применение для решения задач мониторинга окружающей среды //Вычислительные технологии. – 2015. – Т. 20. – №. 5. – С. 157-174.
61. Кадочников А. А. Особенности построения геопространственных веб-приложений и сервисов для систем мониторинга состояния окружающей природной среды //Журнал Сибирского федерального университета. Техника и технологии. – 2015. – Т. 8. – №. 7. – С. 908-916.
62. Ямашкин А. А., Ямашкин С. А. Геопорталы как инструмент доступа к облачным хранилищам данных о метагеосистемах //Известия Саратовского университета. Новая серия. Серия Науки о Земле. – 2022. – Т. 22. – №. 4. – С. 251-261.
63. Гурьев А. Т. и др. Разработка геоинформационной системы на базе программного обеспечения с открытым исходным кодом //Транспортное машиностроение. – 2014. – №. 3 (43). – С. 114-118.
64. Карин С. А. Построение предметно-ориентированной онтологии в системах обработки пространственных данных //Информационно-управляющие системы. – 2014. – №. 4 (71). – С. 78-84.
65. Воробьева Г. Р., Фарваев Э. Ф. Высокореактивное веб-приложение для интеграции разнородных пространственных данных из распределённых источников на основе REST API //Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2023. – №. 4. – С. 152-166.
66. Матвеев А. Г., Якубайлик О. Э. Разработка веб-приложения для обработки и представления пространственных метаданных геопортала //Сибирский аэрокосмический журнал. – 2012. – №. 2 (42). – С. 48-54.
67. Гараева А. Р. и др. Программный модуль интеллектуального анализа пространственно-временных данных для систем экстренных служб //Вестник Технологического университета. – 2018. – Т. 21. – №. 11. – С. 138-142.
68. Гордов Е. П., Окладников И. Г., Титов А. Г. Использование веб-ГИС-технологий для разработки информационно-вычислительных систем для анализа

пространственно-привязанных данных //Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2011. – Т. 9. – №. 4. – С. 94-102.

69. Вагизов М. Р. и др. Разработка интеллектуальной геоинформационной системы для отрасли лесного хозяйства //Геоинформатика. – 2021. – №. 3. – С. 4.

70. Мальцев Е. А. Разработка алгоритмов обработки данных дистанционного зондирования Земли для распознавания пространственных объектов сельскохозяйственного назначения с линейной геометрической структурой : дис. – Сибирский федеральный университет, 2016.

71. Бобков В. А. и др. Программный комплекс визуализации синоптических данных //Информатика и системы управления. – 2014. – №. 1. – С. 141-150.

72. Алешко Р. А. и др. Разработка методики визуализации и обработки геопространственных данных //Научная визуализация. – 2015. – Т. 7. – №. 1. – С. 20-29.

73. Васильев В. Р. и др. Контекстная визуализация пространственных данных //Информационные технологии и вычислительные системы. – 2004. – №. 4. – С. 25-34.

74. Беляков С. Л., Белякова М. Л., Савельева М. Н. Адаптивная к изменению структуры базы данных визуализация пространственных данных //Приборы и системы. Управление, контроль, диагностика. – 2016. – №. 1. – С. 25-32.

75. Молокина Т. С., Колесников А. А. Анализ состояния и перспективы развития визуализации пространственных данных //Вестник СГУГиТ (Сибирского государственного университета геосистем и технологий). – 2021. – Т. 26. – №. 4. – С. 73-82.

76. Кошкарев А. В. Картографическая визуализация на геопорталах ИПД //Геодезия, картография и маркшейдерия. – 2014. – С. 18-25.

77. Казьмина И. Г., Усков В. М. Системы управления качеством атмосферного воздуха на основе визуализации и трансформации пространственных данных //Медико-биологические и социально-психологические проблемы безопасности в чрезвычайных ситуациях. – 2016. – №. 3. – С. 88-92.
78. Базина М. А., Побединский Г. Г. Проблема оптимальной визуализации геопространственных данных в процессе управления территориями и объектами //Системы и средства информатики. – 2008. – С. 204-223.
79. Титов А. Г., Окладников И. Г., Гордов Е. П. Разработка Веб-ГИС на основе сервисов обработки и визуализации пространственных данных для анализа и прогнозирования региональных климатических изменений //Информационные и математические технологии в науке и управлении. – 2016. – №. 4-2. – С. 96-109.
80. Майоров А. А., Гвоздев О. Г. СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТЕХНОЛОГИЙ ЛОКАЛЬНОЙ ОПТИМИЗАЦИИ ПРОГРАММ НА ЯЗЫКЕ PYTHON В ЗАДАЧАХ ОБРАБОТКИ ПРОСТРАНСТВЕННЫХ ДАННЫХ //Известия высших учебных заведений. Геодезия и аэрофотосъемка. – 2014. – №. 5. – С. 61-69.
81. Чечин А. В. Обработка геоданных с использованием языка Python //Великие реки'2018. – 2018. – С. 379-380.
82. Вестра Э. Разработка геоприложений на языке Python. – Litres, 2022.
83. Алексеева К. С., Кириллова А. В. Программное обеспечение для обработки больших геопространственных данных //Путь в науку. Прикладные науки. Строительство. – 2022. – №. 44. – С. 3-5.
84. Полякова А. С. Программное приложение для визуализации пространственных данных на языке Python //Colloquium-journal. – Голопристанський міськрайонний центр зайнятості= Голопристанский районный центр занятости, 2019. – №. 14-2. – С. 177-180.

85. Бородинский Р. Е. и др. Создание инструментария для геопространственного анализа речной сети реки Оки на языке программирования Python //Известия высших учебных заведений. Геодезия и аэрофотосъемка.

86. Галимзянов О. А. Библиотеки Python для анализа геоданных //Актуальные проблемы авиации и космонавтики. – 2019. – Т. 2. – С. 112-114.