

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное научное учреждение
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
«КРАСНОЯРСКИЙ НАУЧНЫЙ ЦЕНТР СИБИРСКОГО ОТДЕЛЕНИЯ
РОССИЙСКОЙ АКАДЕМИИ НАУК»
(КНЦ СО РАН, ФИЦ КНЦ СО РАН)

УДК 004.9

Рег. № НИОКТР № 124012900547-1

Рег. № ИКРБС

УТВЕРЖДАЮ

Директор ФИЦ КНЦ СО РАН

чл.-корр. РАН

_____ А.А. Шпедт

«_____» января 2025 г.

ОТЧЕТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

НОВЫЕ МЕТОДЫ И ТЕХНОЛОГИИ КОМПЛЕКСНОГО АНАЛИЗА СЛОЖНЫХ
ПРИРОДНЫХ И АНТРОПОГЕННЫХ ЭКОСИСТЕМ НА ОСНОВЕ СОВРЕМЕННЫХ
СРЕДСТВ МОДЕЛИРОВАНИЯ И ОБРАБОТКИ ДАННЫХ, РАСПРЕДЕЛЕННЫХ
ВЫЧИСЛЕНИЙ И ЦИФРОВОГО МОНИТОРИНГА
(промежуточный)

Направление фундаментальных исследований:
1.2.1. Компьютерные, информационные науки и биоинформатика
(№ FWES-2024-0014)

Руководитель НИР,
заместитель директора
ИВМ СО РАН по научной работе,
канд. физ.-мат. наук

_____ О. Э. Якубайлик

Красноярск 2025

СПИСОК ИСПОЛНИТЕЛЕЙ

Руководитель темы

заместитель директора

ИВМ СО РАН по научной работе,

канд. физ.-мат. наук

(подпись, дата)

О.Э. Якубайлик

(Разделы 1, 2)

зав. отд., канд. биол. наук

(подпись, дата)

А.В. Андрианова

(Раздел 1)

зав. отд., канд. физ.-мат. наук

(подпись, дата)

Е.Д. Карепова

(Раздел 1)

гл. науч. сотр., д-р техн. наук

(подпись, дата)

А.В. Лапко

(Раздел 3)

гл. науч. сотр., д-р физ.-мат. наук

(подпись, дата)

Н.Я. Шапарев

(Раздел 1)

вед. науч. сотр., д-р биол. наук

(подпись, дата)

В.В. Заворуев

(Раздел 1)

вед. науч. сотр., д-р техн. наук

(подпись, дата)

В.А. Лапко

(Раздел 3)

вед. науч. сотр., д-р физ.-мат. наук

(подпись, дата)

М.Г. Садовский

(Раздел 1)

ст. науч. сотр., канд. техн. наук

(подпись, дата)

А.А. Гостева

(Раздел 2)

ст. науч. сотр., канд. техн. наук

(подпись, дата)

М.Г. Ерунова

(Раздел 1)

ст. науч. сотр., канд. техн. наук

(подпись, дата)

А.А. Кадочников

(Раздел 2)

ст. науч. сотр., канд. физ.-мат. наук

(подпись, дата)

Е.С. Кирик

(Раздел 3)

ст. науч. сотр., д-р техн. наук

(подпись, дата)

В.В. Ничепорчук

(Раздел 3)

ст. науч. сотр., канд. физ.-мат. наук	_____	М.Ю. Сенашова (Раздел 1)
	(подпись, дата)	
ст. науч. сотр., канд. техн. наук	_____	А.В. Токарев (Раздел 2)
	(подпись, дата)	
науч. сотр., канд. техн. наук	_____	Т.Б. Витова (Раздел 3)
	(подпись, дата)	
науч. сотр., канд. физ.-мат. наук	_____	Т.В. Якубайлик (Раздел 1)
	(подпись, дата)	
мл. науч. сотр., канд. физ.-мат. наук	_____	О.С. Володько (Раздел 1)
	(подпись, дата)	
мл. науч. сотр.	_____	А.К. Матушко (Раздел 1)
	(подпись, дата)	
инж.	_____	А.В. Богданов (Раздел 3)
	(подпись, дата)	
инж.	_____	А.А. Шехова (Раздел 1)
	(подпись, дата)	
Нормоконтроль	_____	А.А. Кадочников
	(подпись, дата)	

РЕФЕРАТ

Отчёт 142 с., 48 рис., 14 табл., 3 прил.

МОРФОМЕТРИЧЕСКИЙ АНАЛИЗ, ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ, ЭКОСИСТЕМА, ТЕПЛОВЫЕ АНОМАЛИИ, ГЕОПОРТАЛ, СИСТЕМА МОНИТОРИНГА ВОЗДУХА, ВЗВЕШЕННЫЕ ЧАСТИЦЫ, ТЕМПЕРАТУРНАЯ ИНВЕРСИЯ, ЭВАКУАЦИЯ, БЕЗОПАСНОСТЬ ТЕРРИТОРИЙ, НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ

Основная цель исследований заключается в разработке инновационных методов и технологий для комплексного анализа и оценки состояния природных и антропогенных экосистем. Содержание работ включает создание физико-математических и геоинформационных моделей, информационно-аналитических систем, алгоритмов распределенной обработки больших данных, а также инструментов интеллектуальной поддержки решений в сфере рационального природопользования и управления территориальной безопасностью.

Выполнение работ основано на применении современных методов математического анализа, обработки данных и проектирования информационно-вычислительных технологий, создания геопространственных веб-систем и сервисов.

Научные результаты исследований представлены в ведущих отечественных и международных научных изданиях. Разработанное прикладное программное обеспечение интегрировано в региональные информационные системы, что подтверждает его технологическую востребованность.

Практическая значимость проекта заключается в создании инструментов для экологического мониторинга, обработки пространственных данных и разработке ГИС-решений. Это обеспечивает рациональное природопользование и поддержку принятия управленческих решений.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	7
1. МЕТОДЫ И МОДЕЛИ, ТЕХНОЛОГИИ КОМПЛЕКСНОГО АНАЛИЗА И ОЦЕНКИ СОСТОЯНИЯ СЛОЖНЫХ ПРИРОДНЫХ И АНТРОПОГЕННЫХ ЭКОСИСТЕМ	9
1.1. Методы морфометрического анализа агроландшафтов и оценки агроэкологических условий сельскохозяйственных территорий	9
1.2. Образование адвективного тумана охлаждения над рекой и другие эффекты, возникающие вследствие экранирования поверхности земли дымом лесных пожаров	20
1.3. Разработка инновационных методов анализа многомерных данных для задач моделирования сложных процессов в биосистемах (в том числе – в медицине).....	26
1.4. Влияние городской среды на формирование локальных климатических зон и тепловых аномалий на примере Красноярска.....	33
1.5. Оценка экологического статуса водоемов с разной степенью антропогенной нагрузки методами биоиндикации	42
1.6. Технология построения комплексного биоиндикационного показателя качества воды.....	48
2. СИСТЕМА МОНИТОРИНГА ВОЗДУХА Г. КРАСНОЯРСКА: ИНФОРМАЦИОННО- ВЫЧИСЛИТЕЛЬНОЕ ОБЕСПЕЧЕНИЕ, ОБРАБОТКА И АНАЛИЗ ДАННЫХ.....	56
2.1. Система мониторинга воздуха Красноярского научного центра СО РАН	56
2.2. Измерение концентрации PM _{2,5} в атмосфере Красноярска в 2023 году: сравнение данных четырех сетей мониторинга.....	67
2.3. Концентрация бенз(а)пирена в атмосфере Красноярска в 2019-2023 годах: анализ данных краевой и федеральной наблюдательной сети.....	72
2.4. Сравнительный анализ данных мониторинга концентрации взвешенных частиц PM _{2.5} , регистрируемых разными датчиками	78
2.5. Формирование и последствия температурных инверсий в г. Красноярске.....	89
2.6. Прогнозирование величины концентрации загрязняющих веществ в атмосфере г. Красноярска с помощью моделей машинного обучения	97
3. ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММ ДЛЯ ПРОБЛЕМНО-ОРИЕНТИРОВАННЫХ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИХ СИСТЕМ И СЕРВИСОВ	102
3.1. Разработка организационной модели управления безопасностью территорий	102
3.2. Технология применения математического моделирования движения людей при пожаре для оценки обеспечения безопасных условий эвакуации в музеях	114

3.3. Разработка и исследование непараметрических методов проверки гипотез о распределениях многомерных случайных величин большого объёма.....	120
ЗАКЛЮЧЕНИЕ.....	129
ПРИЛОЖЕНИЕ А. Научные публикации в журналах, индексируемых в российских и международных информационно-аналитических системах научного цитирования	135
ПРИЛОЖЕНИЕ Б. Зарегистрированные результаты интеллектуальной деятельности	137
ПРИЛОЖЕНИЕ В. Прочие публикации и доклады на конференциях	138

ВВЕДЕНИЕ

Настоящий отчет является промежуточным (этап 1) по теме «Новые методы и технологии комплексного анализа сложных природных и антропогенных экосистем на основе современных средств моделирования и обработки данных, распределенных вычислений и цифрового мониторинга» (FWES-2024-0014); в нем содержатся результаты за отчетный период 1-го этапа проекта – 2024 год.

Проект посвящен исследованию и разработке математического и программно-технологического обеспечения для систем мониторинга и оценки состояния природной среды. Актуальность данного направления обусловлена стремительным увеличением объемов геопространственной информации, что связано с развитием сетей передачи данных (расширением зон покрытия и повышением скорости интернет-соединений), широким распространением мобильных устройств (смартфонов и планшетов), массовым внедрением недорогих датчиков, измеряющих различные параметры окружающей среды (метеорологические данные, уровень загрязнения воздуха, и др.), а также повышением доступности спутниковой информации с высоким пространственным разрешением и ростом популярности беспилотных летательных аппаратов (БПЛА).

Развитие нового поколения систем дистанционного зондирования Земли (ДЗЗ) и внедрение суперкомпьютерных технологий привели к появлению новых массивов климатических и экологических данных с высокой степенью детализации. Увеличение скорости интернет-соединений, массовое использование мобильных устройств и доступных датчиков/регистраторов данных, а также прогресс в области веб-технологий и быстрой разработки веб-приложений создали основу для систем ввода геопозиционированных данных различной тематической направленности и оперативного цифрового мониторинга. Согласно экспертным оценкам, объем доступной геоинформации увеличивается как минимум на порядок каждые 3-5 лет. При этом важно отметить, что данный рост носит не только количественный, но и качественный характер. Современные методы и инструменты позволяют регистрировать, хранить и обрабатывать показатели, которые ранее были недоступны для исследования, причем в режиме реального времени. Это делает задачу разработки математических моделей и инструментов для многомерного нелинейного анализа больших данных чрезвычайно актуальной и востребованной.

Эффективное решение задач анализа природных и антропогенных экосистем невозможно без соответствующих инструментов хранения, представления и интеллектуальной обработки данных. В настоящее время акцент смещается в сторону специализированных серверно-ориентированных и облачных решений. Разработка приложений осуществляется по модульному принципу на основе сервис-ориентированного подхода и веб-технологий, в

рамках распределенной архитектуры с использованием мультиагентных методов. Представленный проект полностью соответствует указанным тенденциям. В его рамках предполагается создание информационно-вычислительной среды для решения поставленных задач, включая набор сервисов различного типа, ориентированных на выполнение различных функциональных подзадач – от сложной обработки спутниковых снимков на высокопроизводительных кластерах до простой картографической веб-визуализации на мобильных устройствах.

Актуальность предлагаемого подхода обусловлена потенциальными приложениями результатов исследований. Одной из характерных особенностей и проблем современных систем социально-экономического и экологического мониторинга является повышенное внимание со стороны общественности и растущая потребность в простом и быстром доступе к исходным данным. Веб-технологии, включая социальные сети, кооперативно создаваемые картографические ресурсы и облачные сервисы, способствуют выводу информационного обеспечения данного класса задач на качественно новый уровень.

Результаты, представленные в данном отчете, являются логическим продолжением ранее начатых исследований, направленных на решение актуальных задач оценки состояния природной среды, физико-математического моделирования природных процессов и явлений, а также разработки соответствующего программно-технологического и информационно-вычислительного обеспечения на основе геоинформационных веб-систем.

В работе также предложена и исследована методика проверки гипотезы о независимости многомерных случайных величин, основанная на непараметрическом алгоритме распознавания образов. Ее применение продемонстрировано на примере анализа данных дистанционного зондирования Земли (ДЗЗ) для антропогенных территорий.

Основные результаты исследований сгруппированы по трем направлениям:

- Методы и модели, технологии комплексного анализа и оценки состояния сложных природных и антропогенных экосистем;
- Система мониторинга воздуха г. Красноярска: информационно-вычислительное обеспечение, обработка и анализ данных
- Проектирование и разработка алгоритмов и программ для проблемно-ориентированных информационно-аналитических систем и сервисов.

Полученные за отчетный период основные результаты изложены в соответствующих разделах настоящего отчета.

1. МЕТОДЫ И МОДЕЛИ, ТЕХНОЛОГИИ КОМПЛЕКСНОГО АНАЛИЗА И ОЦЕНКИ СОСТОЯНИЯ СЛОЖНЫХ ПРИРОДНЫХ И АНТРОПОГЕННЫХ ЭКОСИСТЕМ

1.1. Методы морфометрического анализа агроландшафтов и оценки агроэкологических условий сельскохозяйственных территорий

Рациональное землепользование, сохранение почвенного плодородия, повышение производительной способности земель сельскохозяйственного назначения невозможно без комплексного ландшафтно-экологического подхода, включающего оценку широкого спектра факторов как природных, так и антропогенных.

Общепризнанной является роль рельефа в формировании агроэкологических условий агроландшафтов. Изменение рельефа во времени и пространстве оказывают воздействие на агроклиматические условия, сток поверхностных вод, эрозионные процессы и др. При оценке состояния агроландшафтов необходимо учитывать не только климатические условия и показатели почвенного плодородия, но и морфометрические характеристики рельефа, степень деградации почв, в совокупности, влияющие на продуктивность с.-х. культур.

На сегодняшний день, систематизацию, хранение, визуальное отображение и пространственный анализ информации следует осуществлять методами геоинформационного картографирования, создания баз пространственных геоданных и цифровой обработки данных дистанционного зондирования Земли (ДЗЗ). Развитие современных технологий обработки информации предоставляют возможность формализованного описания рельефа, в виде построения цифровых моделей рельефа (ЦМР). Оценка состояния агроландшафтов комплексная задача, требующая применения различных методов и технологий геоинформационного моделирования с использованием дополнительного набора данных. Климатические показатели, тип почв, рельеф, степень деградации земель входят в атрибутивный набор, методы ГИС и веб-технологий эффективно используются для хранения и обработки этих данных. Создание обширной пространственной и информационной базы, содержащей данных агроландшафтов, будет служить наиболее эффективным способом обработки и визуализации больших объемов информации, для принятия управленческих решений по рациональному использованию земельных ресурсов.

Традиционно главным источником количественной геоморфологической информации были результаты морфометрического анализа топографических карт. С развитием аэрокосмических и информационных технологий на смену традиционным методам пришел новый – цифровое моделирование. Основу морфометрического анализа с использованием данных дистанционного зондирования (ДДЗ) составляет анализ цифровых моделей рельефа; эти данные используются в качестве входных для количественной оценки характеристик земной поверхности.

Современные ЦМР основаны на ДДЗ, их точность постоянно повышается. В середине 50-х годов прошлого века в фотограмметрии возникло новое направление – цифровое моделирование рельефа (Digital Terrain Modeling) [1]. Широкое использование ЦМР началось около 20 лет назад с появлением отредактированной версии спутниковой модели Shuttle Radar Topography Mission (SRTM), которая имела пространственное разрешение 3 угловых секунды. Его улучшенная версия SRTM Plus с пространственным разрешением 1 угловая секунда стала доступна в 2014 году. За последние 15 лет был представлен ряд общедоступных открытых ЦМР, созданных на основе спутниковых данных и различных алгоритмов обработки: ASTER Global Digital Elevation Model (GDEM), Multiple Error Elimination Enhanced Terrain (MERIT), Copernicus DEM (GLO-30 & GLO-90), FABDEM (Forest and Buildings removed Copernicus DEM). Эти модели характеризуются пространственным разрешением в 3 или 1 угловую секунду – шаг сетки на экваторе составляет соответственно 90×90 м и 30×30 м, а на наших широтах (50-60° с.ш.) – не хуже 60×60 м и 20×20 м.

Данные с размером ячейки 1 угловая секунда представляют особый интерес для специалистов аграрного сектора, т. к. позволяют оценить морфометрические характеристики рельефа уже на уровне отдельных полей. Благодаря глобальному покрытию, хорошему качеству исходной информации подобные общедоступные открытые ЦМР, основанные на спутниковых данных, имеют большие перспективы использования в решении практических задач. Стоит также отметить, что сегодня также существует значительный интерес к данным ЦМР, полученным на основе аэрофотосъемки с беспилотных летательных аппаратов (БПЛА) – их точность может достигать нескольких сантиметров. Целесообразность, необходимость создания и обработки данных ЦМР средствами БПЛА определяется решаемыми задачами – с одной стороны качественно другой уровень детализации данных, но и совершенно другие затраты (финансовые, технические, организационные).

Качество ЦМР является мерой того, насколько точно указана высота каждого пикселя (абсолютная точность) и насколько точно представлена морфология (относительная точность) [2]. По совокупности факторов использование некоммерческой открытой ЦМР FABDEM представляется сегодня наиболее целесообразным для рассматриваемого класса задач оценки агроэкологических условий сельскохозяйственных территорий.

Цифровое моделирование рельефа позволяет проводить анализ рельефа изучаемой территории с целью получения морфометрических параметров. Применение таких параметров используется в исследованиях сельскохозяйственных земель относительно недавно [3], например, топографический индекс влажности (TWI) применяют в качестве показателя влажности почвы, показатель экспозиции склонов применяют для выявления сезонного изменения NDVI посевных площадей. Крутизна и экспозиция склонов применяется для

совершенствования методики определения природно-ресурсного потенциала земель, предназначенных для сельскохозяйственного производства, путем введения поправочных коэффициентов на рельеф, и т.д. Морфометрические параметры рельефа могут использоваться в качестве источника для прогнозирования целого ряда экзогенных процессов, позволяют оценивать вероятность и интенсивность их развития. Моделирование временных водотоков применяется для оценки поверхностного стока, степени увлажнения почвы. Цифровое моделирование рельефа дает возможность создавать тематические карты важнейших геоморфометрических параметров рельефа.

Целью данной работы является формирование геопространственной базы данных агроландшафтов Красноярского края морфометрическими показателями на примере сельскохозяйственной территории ОПХ «Михайловское» на основе ЦМР FABDEM.

В качестве объекта исследования выбрано опытно-производственное хозяйство «Михайловское», которое является базовым хозяйством Федерального исследовательского центра «Красноярский научный центр СО РАН» (ФИЦ КНЦ СО РАН). Данное хозяйство расположено в северо-западной части Ужурского района в 30 км от районного центра г. Ужур, 250 км от г. Красноярска. Общая площадь ОПХ составляет 15209 га, в том числе площадь сельскохозяйственных угодий – 11294 га. ОПХ «Михайловское» расположено в лесостепной зоне Причулымья.

В качестве базового источника для определения морфометрических характеристик поверхности использовали ЦМР FABDEM [4] с исходным пространственным разрешением менее 20 м. Для обеспечения корректности расчетов оригинальную FABDEM перепроецировали из географической системы координат в метрическую проекцию (универсальная поперечного Меркатора – UTM) в открытой ГИС QGIS. Затем полученную ЦМР обрезают по границе хозяйства и по границам отдельных полей.

Геоморфометрический анализ ЦМР выполнялся с использованием функциональных возможностей набора инструментов «Basic terrain analysis» SAGA GIS [5]. В каждой ячейке раstra рассчитывали следующие морфометрические параметры рельефа: крутизна, экспозиция, кривизна (плановая и профиля), индекс расчлененности рельефа (TRI), фактор риска развития эрозии (LS-фактор), топографический индекс влажности (TWI). Выбор этих параметров связан с их применением при исследовании сельскохозяйственных земель, в частности для прогнозирования ряда экзогенных процессов. Так, параметр крутизна склонов используется для оценки потенциального стока и эрозионной опасности. индекс потенциала плоскостной эрозии (Slope Length and Steepness factor, LS-factor) – один из комплексных морфометрических показателей, характеризующих эрозионный потенциал склонов. Он определяет влияние длины и крутизны склона на эрозию почвы, то есть процессы смыва и

размыва почвы и пород поверхностным стоком временных водных потоков, возникающих главным образом в результате таяния снега или дождевых осадков. LS-factor используют в моделях универсальных уравнений потерь почвы USLE и RUSLE, которые наиболее часто применяют для оценки риска эрозии и потерь почв [6]. Чем больше величина этого показателя, тем выше влияние рельефа на процессы водной эрозии. Для выделения форм рельефа использовали инструмент классификации SAGA «TPI Based Landform Classification».

Для исследуемой территории ОПХ была сформирована база данных ключевых морфологических показателей рельефа. Для всей территории хозяйства и каждого поля в отдельности были получены статистические данные морфометрических параметров. По каждому полю рассчитывали средние, максимальные и минимальные величины всех показателей. Выявление диапазона влияния каждого показателя осуществляли путем извлечения зональной статистики по ячейкам регулярной сети с использованием рабочего модуля QGIS. По полученным показателям была построена серия крупномасштабных карт ключевых морфологических показателей рельефа: крутизны и экспозиций склонов, кривизны (плановой и профиля), индекса расчлененности рельефа (TRI), топографического индекса влажности (TWI), LS-фактора, и др.

Морфометрический анализ территории ОПХ «Михайловское» показал, что, несмотря на небольшую площадь, устройство его поверхности неоднородно. Сельскохозяйственные поля находятся на возвышенной равнине с преобладающими высотами порядка 400 м. Морфометрический параметр «крутизна склонов» используется для оценки потенциального стока и эрозионной опасности. На рисунке 1 представлена тематическая карта крутизны склонов полей хозяйства, построенная по классификации для равнинных территорий.

Поля территории, находящиеся на склонах с крутизной до 3° , занимают 92% территории (плоские поверхности – 51%, очень пологие – 30%, пологие 11%). Участки с крутизной склонов $3...5^\circ$ составляют 6%. Площади с крутизной склонов выше 5° расположены в юго-западной части хозяйства и составляют 2% территории всех полей. Полученная карта крутизны склона представляет собой растровый слой с шагом 10 м, что позволяет проводить более детальный анализ на уровне поля. Так, например, поле 22, которое находится севернее села Михайловка, в основном расположено на склонах до 1° (65% от общей площади поля), 26% площади поля расположено на склонах $1...2^\circ$, 8% – $2-3^\circ$, и 1% склоны от $3...4^\circ$.

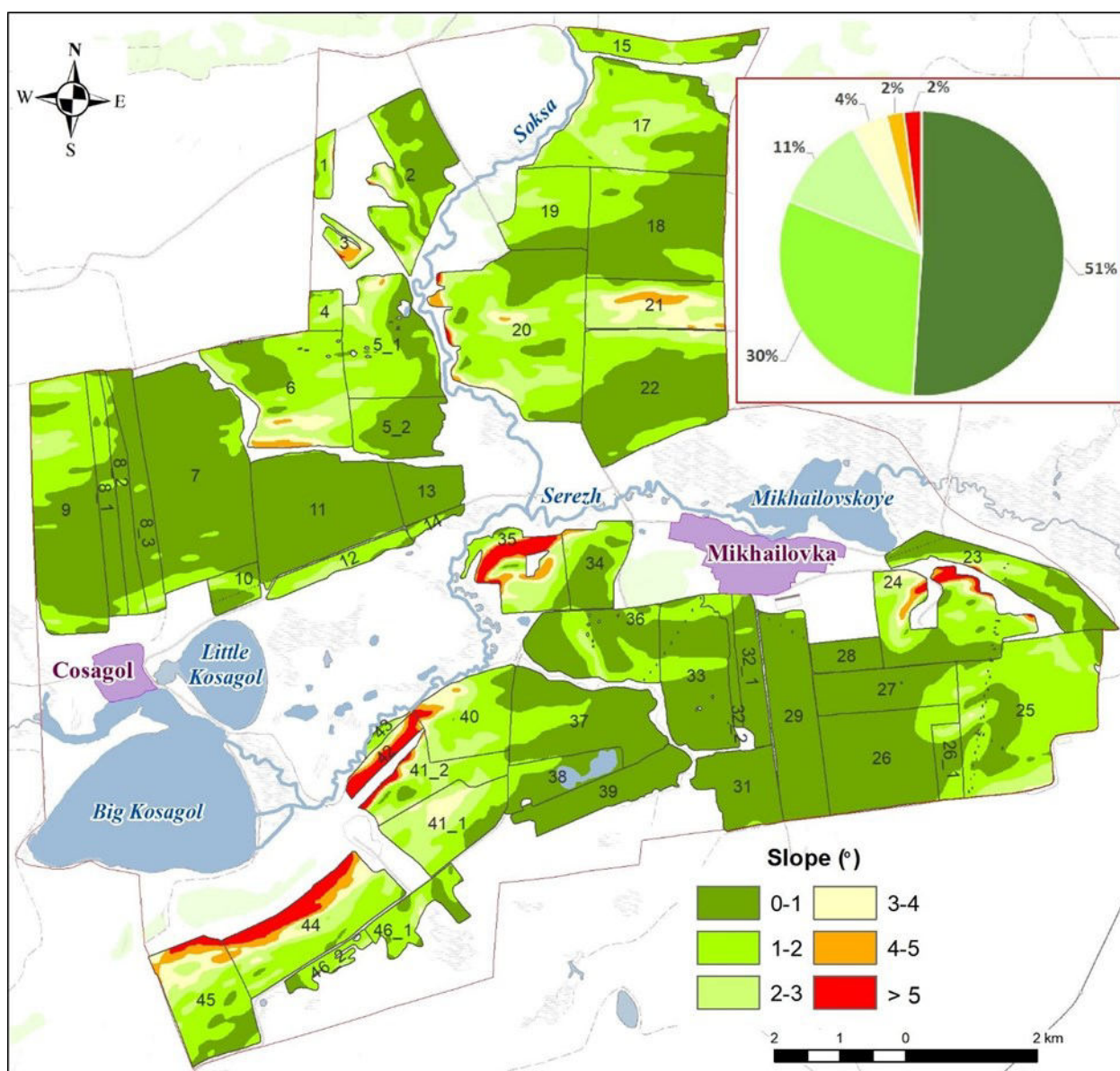


Рисунок 1 – Тематическая карта по параметру крутизна склона с диаграммой занимаемых площадей, %.

Аналогично были проанализированы другие морфологические показатели. На рисунке 2 представлен параметр «элементов рельефа». Анализ карты показал, что 85% территории полей ОПХ «Михайловское» представляют собой ровные участки (плоскую поверхность). Остальная территория (15%) представлена локальными преимущественно приподнятыми формами рельефа.

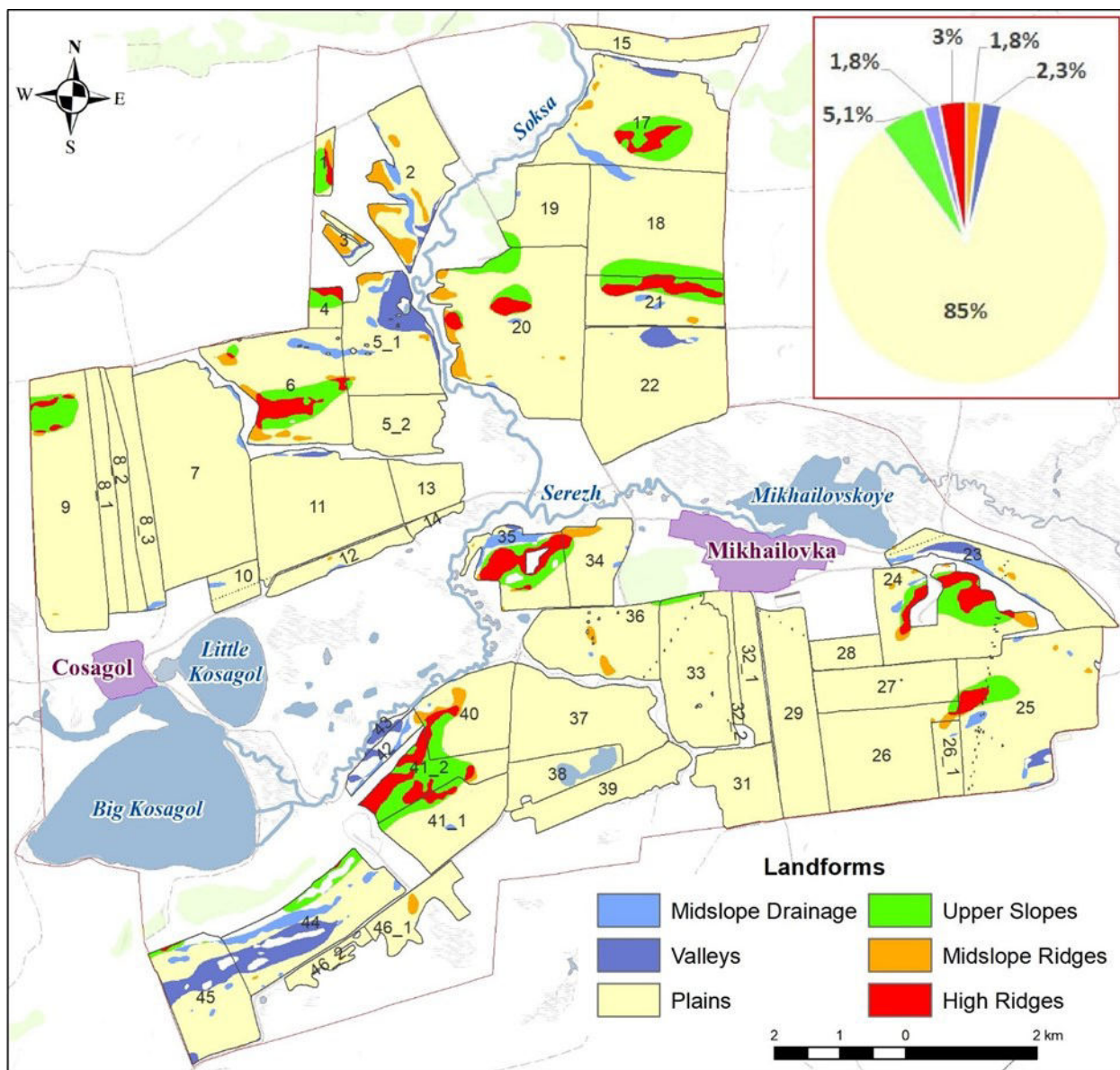


Рисунок 2 – Тематическая карта элементов рельефа с диаграммой занимаемых площадей, %.

На рисунке 3 представлен параметр «экспозиция склона», который используется для оценки теплообеспеченности и влагообеспеченности. На полях ОПХ распределены участки следующим образом: южные склоны – 36% от общей площади, северные склоны составляют 22 %, западные склоны составляют 15 %, а восточные склоны составляют 27 %. По этой карте можно определить характер освещения и прогревания поверхности, перераспределения света и тепла на каждом отдельном поле, выделить участки с более благоприятным световым, термическим режимом.

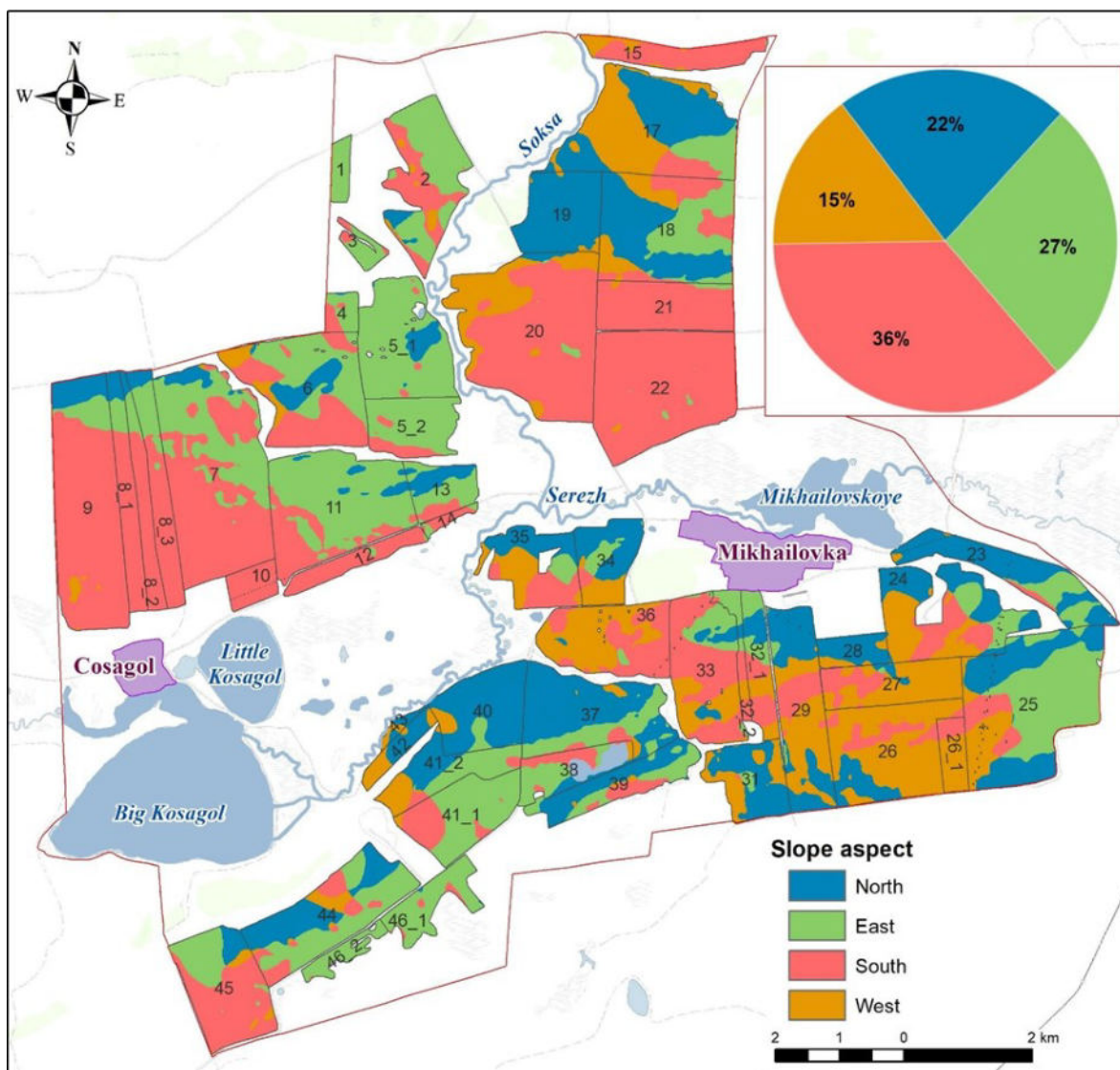


Рисунок 3 – Тематическая карта экспозиции склонов с диаграммой занимаемых площадей, %.

Топографический индекс влажности (Topographic Wetness Index, TWI) – это параметр, описывающий склонность конкретной ячейки раstra к накоплению в ней воды, т.е. склонность к аккумуляции осадков [7]. Топографический индекс влажности (Рис. 4) – показатель гидроморфности почвенного покрова, который во многом определяется особенностями рельефа территории. Он позволяет оценить предпосылки к развитию переувлажненных земель и развитию процессов заболачивания, а также учесть данный фактор при планировании оптимизационных (мелиоративных) мероприятий. Поэтому топографический индекс используют в качестве показателя влажности почвы. Анализ карты показал, что наибольшие значения TWI (более 17) представляют собой дренажные впадины, т.е. эти участки с повышенным содержанием (накоплением) влаги в почве. Участки с наибольшие значения TWI расположены в юго-восточной части хозяйства. Более низкие значения (тенденция к

красному цвету) представляют собой возвышенность. Низкие значения индекса TWI на полях 2, 3, 44, 42, 43 – это значения до 7 и эти участки соответствуют участкам с самым крутым склоном. Анализ топографического индекса влажности ОПХ в целом показал незначительную эрозионную опасность: только 0,5 % земель хозяйства имеют дренажные понижения, а 3 % расположены на холмах.

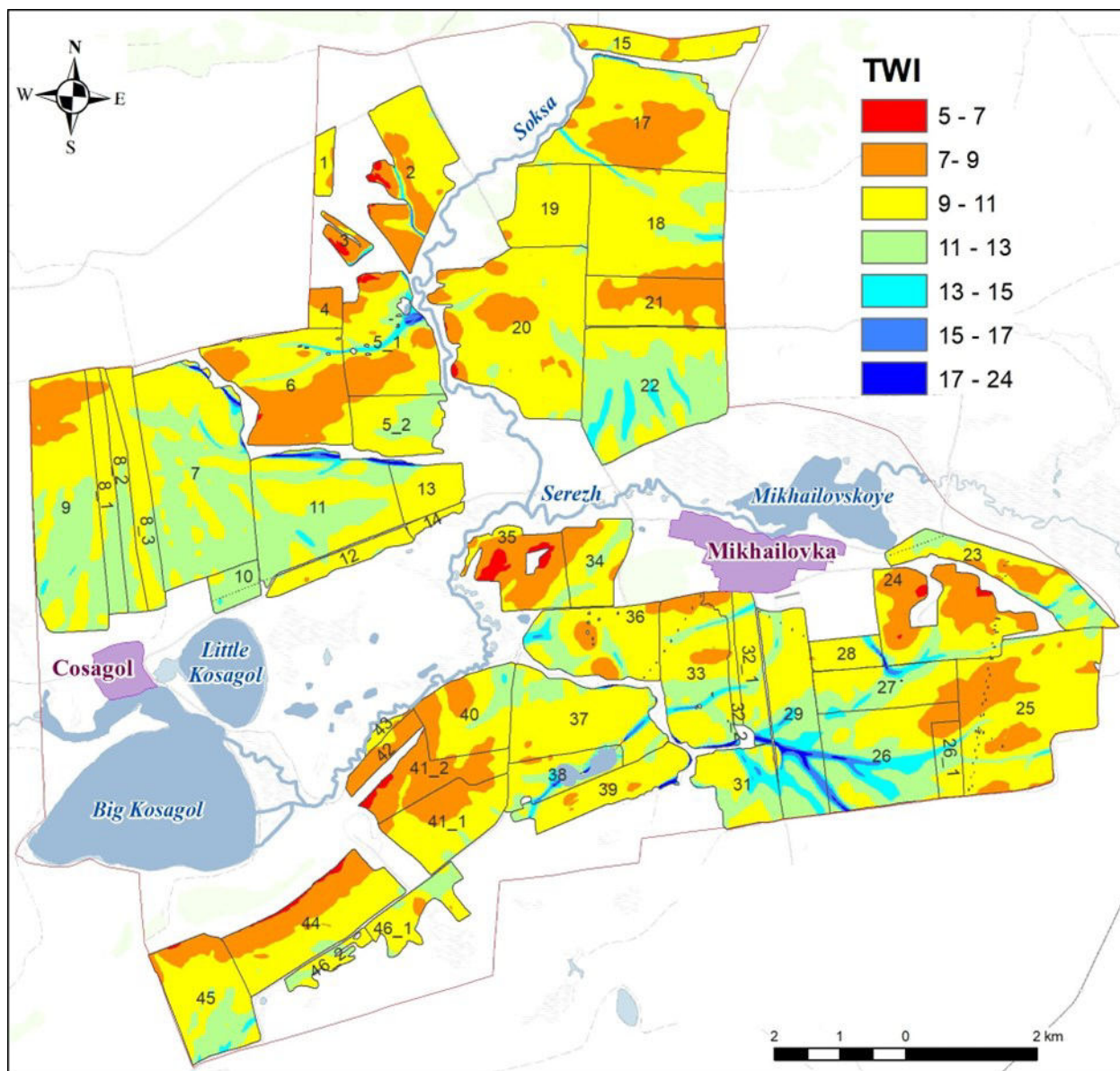


Рисунок 4 – Тематическая карта топографического индекса влажности (TWI).

Одним из комплексных морфометрических показателей, используемым для оценки эрозионного потенциала склонов, является индекс потенциала плоскостной эрозии (Slope Length and Steepness factor, LS-factor). Он определяет влияние длины и крутизны склона на эрозию почвы, т. е. процесс смыва и размыва почвы и пород поверхностным стоком временных водных потоков, возникающих главным образом в результате таяния снега или дождевых осадков. LS-factor используется в моделях универсальных уравнений потерь почвы

USLE и RUSLE, которые являются наиболее часто используемыми моделями для оценки риска эрозии и потерь почв [8].

Тематическая карта потенциала плоскостной эрозии LS-factor представлена на рисунке 5. Чем больше значение LS-factor, тем выше влияние рельефа на процессы водной эрозии, т.е. тем более ярко выражены процессы смыва почвы. Наибольшие значения данного параметра (темно-розовый цвет) для данной территории характерны для участков на полях с уклоном более 4° и с увеличением расчленённости рельефа. Данные участки расположены на полях 3, 21, 35, 42, 44, 45. В этот класс также попали участки с наибольшими значениями топографического индекса влажности TWI, т.е. дренажные впадины. В целом территория ОПХ «Михайловское» расположена на ровном участке и лишь небольшая доля территории подвержена эрозионным процессам.

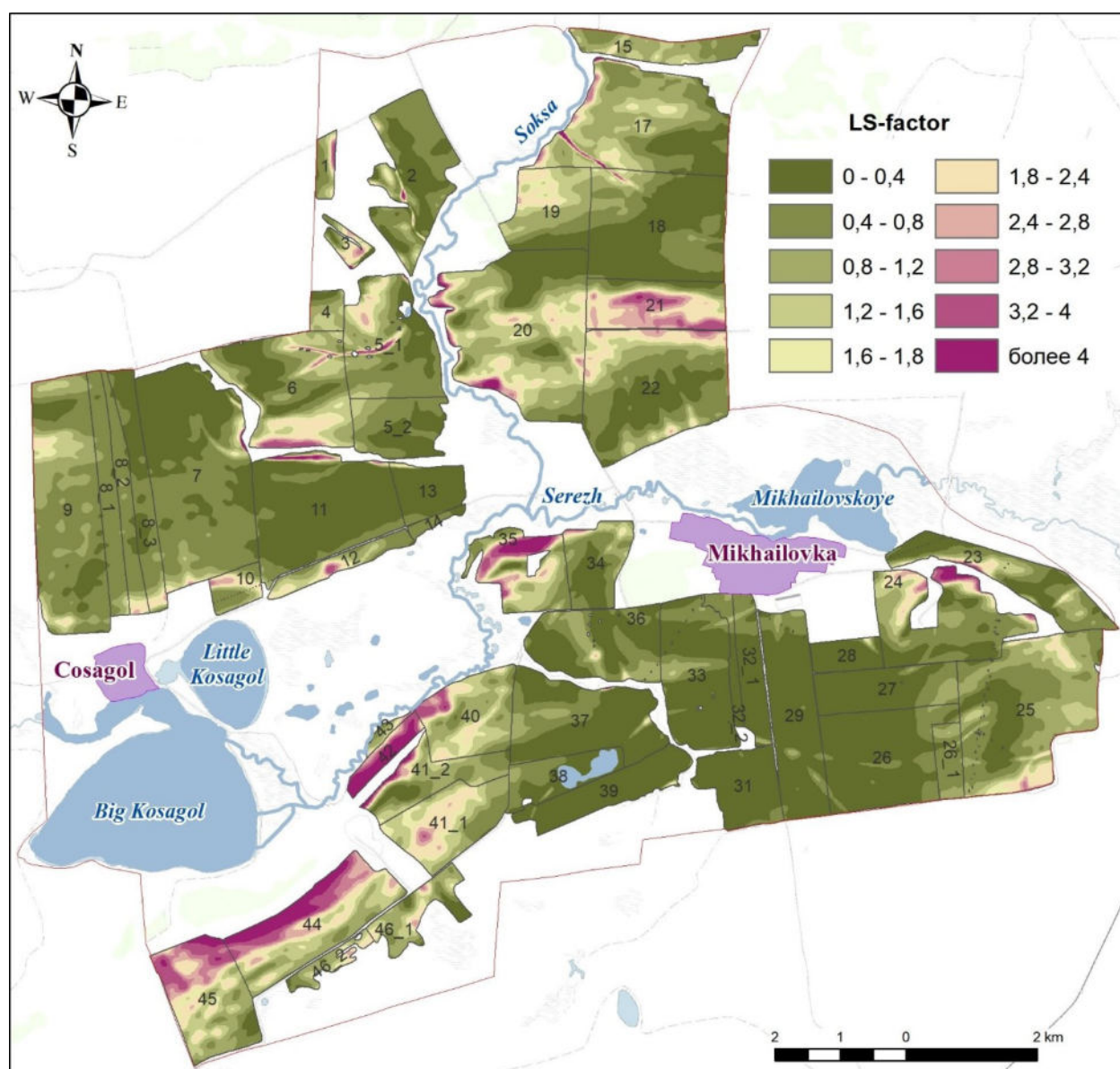


Рисунок 5 – Тематическая карта индекса потенциала плоскостной эрозии.

Для территории ОПХ «Михайловское» проведено цифровое моделирование рельефа на основе ЦМР FABDEM и функциональных возможностей набора инструментов SAGA GIS. Проведен геоморфометрический анализ территории, получены ключевые параметры: крутизны и экспозиций склонов, кривизны (плановой и профиля), индекса расчлененности рельефа (TRI), LS-фактора, топографического индекса влажности (TWI), и др., которые хранятся в виде геопространственной базы данных хозяйства, а также в виде набора тематических карт ключевых параметров рельефа данной территории. Для всей территории хозяйства и каждого поля в отдельности были получены статистические данные морфометрических параметров. По каждому полю получены средние, максимальные и минимальные величины всех показателей.

Выполненный в работе геоморфометрический анализ территории ОПХ «Михайловское» показал, что 85 % территории полей хозяйства занимают ровные участки (плоская поверхность). На остальной площади (15 %) расположены локальные преимущественно приподнятые формы рельефа. Несмотря на небольшую площадь, устройство поверхности территории ОПХ «Михайловское» неоднородно, в том числе 85 % приходится на ровные участки (плоская поверхность), остальные 15 % на локальные, преимущественно приподнятые, формы рельефа. На склонах крутизной до 3° приходится 92 % от общей площади, участки с крутизной склонов более 3° занимают всего 8 %. На полях хозяйства преобладают территории с западными и восточными склонами – 42 % от общей площади, на южные приходится 36 %, на северные – 22 %. Результаты анализа топографического индекса влажности территории ОПХ в целом свидетельствуют о незначительной эрозионной опасности: 0,5 % от общей площади хозяйства представляют собой участки с дренажными впадинами и 3 % – возвышенности. По величине индекса потенциала плоскостной эрозии участки на возвышенностях расположены на склонах крутизной более 4°. Представляется целесообразным вывести эти участки из пашни и перевести в сенокосы или засеять многолетними травами. На остальной (ровной) территории необходимо провести агрохимические исследования для рационального внесения минеральных удобрений под сельскохозяйственные культуры.

Полученные результаты показали, что современная ЦМР FABDEM пригодна для получения информации об общих топографических особенностях рельефа для формирования агроландшафтов как на уровне хозяйства, так и на уровне Красноярского края. Использование ЦМР и ДДЗ открытого пользования позволяет осуществить полноценный сбор и анализ рельефа любой сельскохозяйственной территории.

Разработанная база геоданных морфологических показателей может стать основой для изучения и прогнозирования проявления водно-эрозионных процессов на сельскохозяйственном угодье, позволят своевременно разработать и организовать противоэрозионные

мероприятия. Предложенная методика геоинформационного картографирования (получение тематических карт морфометрических показателей) рельефа агроландшафтов на местном (локальном) уровне может быть экстраполирована для дальнейших агроландшафтных исследований в свете цифровизации АПК Красноярского края.

Список цитируемой в разделе литературы:

1. Modeling topographic potential for erosion and depositing using GIS / H. Mitasova, J. Hofierka, M. Zlocha, et al. // Int. J. Geogr. Inf. Syst. 1996. Vol. 10. No. 5. P. 629–641. DOI: 10.1080/0269379960890210..
2. Vertical Accuracy of Freely Available Global Digital Elevation Models (ASTER, AW3D30, MERIT, TanDEM-X, SRTM, and NASADEM) / E. Uuemaa, S. Ahi, B. Montibeller, et al. // Remote Sensing. 2020. Vol. 12(21). Article 3482. DOI: 10.3390/rs12213482.
3. Wilson J. P. Digital Terrain Modeling // Geomorphology. 2012. Vol. 137. P. 107–121. DOI: 10.1016/j.geomorph.2011.03.012.
4. A 30 m global map of elevation with forests and buildings removed / L. Hawker, P. Uhe, L. Paulo, et al. // Environmental Research Letters. 2022. Vol. 17(2).
5. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4 / O. Conrad, B. Bechtel, M. Bock, et al. // Geosci. Model Dev. 2015. Vol. 8. P. 1991–2007. DOI: 10.5194/gmd-8-1991-2015.
6. Гопп Н. В. Использование почвенно-геоморфологической базы данных для изучения пространственной изменчивости содержания гумуса, физической глины и ила в почвах Кузнецко-салаирской геоморфологической провинции // Почвоведение. 2021. № 7. С. 783–796. DOI: 10.31857/S0032180X21070054.
7. Мальцев К. А., Голосов В. Н., Гафуров А. М. Цифровые модели рельефа и их использование в расчётах темпов смыва почв на пахотных землях // Учен. зап. Казан. ун-та. Сер. Естеств. науки. 2018. Т. 160. кн. 3. С. 514–530.
8. Гопп Н. В. Использование почвенно-геоморфологической базы данных для изучения пространственной изменчивости содержания гумуса, физической глины и ила в почвах Кузнецко-салаирской геоморфологической провинции // Почвоведение. 2021. № 7. С. 783–796. DOI: 10.31857/S0032180X21070054.

Основные публикации по разделу:

1. Геоморфометрический анализ сельскохозяйственных территорий на основе новой цифровой модели рельефа / М. Г. Ерунова, А. С. Кузнецова, А. А. Шпедт и др. // Достижения науки и техники АПК. 2024. Т. 38. № 4. С. 4–9. DOI: 10.53859/02352451_2024_38_4_0 [БАК, Ядро РИНЦ, RSCI]

2. Erunova M.G., Yakubailik O.E. Geomorphometric analysis of agricultural areas based on the FABDEM digital elevation model // ИнтерКарто. ИнтерГИС, 2024. Т. 30. Ч. 2. С. 252–262. DOI: 10.35595/2414-9179-2024-2-30-252-262 [Scopus, Ядро РИНЦ, RSCI]
3. Геоморфометрический анализ рельефа сельскохозяйственных территорий на основе данных дистанционного зондирования / Ерунова М.Г., Кузнецова А.С., Шпедт А.А., Якубайлик О.Э. // Проблемы плодородия почв в современном земледелии. Материалы Международной научно-практической конференции, посвященной 70-летию освоения целинных и залежных земель. Красноярск, 2024. С. 413-416. DOI: 10.52686/9785605087878_413
4. Показатели состояния сельскохозяйственных угодий опытно-производственного хозяйства "Михайловское" Ужурского района Красноярского края / Ерунова М.Г., Кузнецова А.С., Якубайлик О.Э. // Свидетельство о регистрации базы данных RU 2024623107, 15.07.2024. Заявка от 27.06.2024.

1.2. Образование адвективного тумана охлаждения над рекой и другие эффекты, возникающие вследствие экранирования поверхности земли дымом лесных пожаров

В отчетный период были продолжены ранее начатые исследования по оценке воздействия задымления территории от лесных пожаров на состояние природной среды.

Рассматривались периоды возникновения за несколько последних лет значительных по интенсивности лесных пожаров, дым от которых достигал г. Красноярска. Как показывают данные наблюдений, в ряде случаев дым лесных пожаров может распространяться на расстояние в несколько тысяч километров (Рис. 6).

Лесные пожары представляют собой экстремальное явление, которое нарушает экологический баланс и влияет на изменение климатических условий. Исследователи уже более 40 лет назад обратили внимание на то, что блокирование солнечного света смогом от лесных, городских и промышленных пожаров может приводить к значительному похолоданию природной среды [1]. Это явление позднее получило название «ядерная зима» [2].

В конце 1980-х были опубликованы первые результаты исследований, посвященные моделированию и оценке массовых природных и «ядерных» пожаров [3, 4]. Дымовые аэрозоли, образующиеся при сгорании биомассы, поглощают и рассеивают солнечное излучение, что приводит к снижению количества солнечной радиации, достигающей поверхности Земли [5, 6]. Важно отметить, что атмосферный аэрозоль является одним из ключевых факторов, влияющих на климатические характеристики [7, 8].

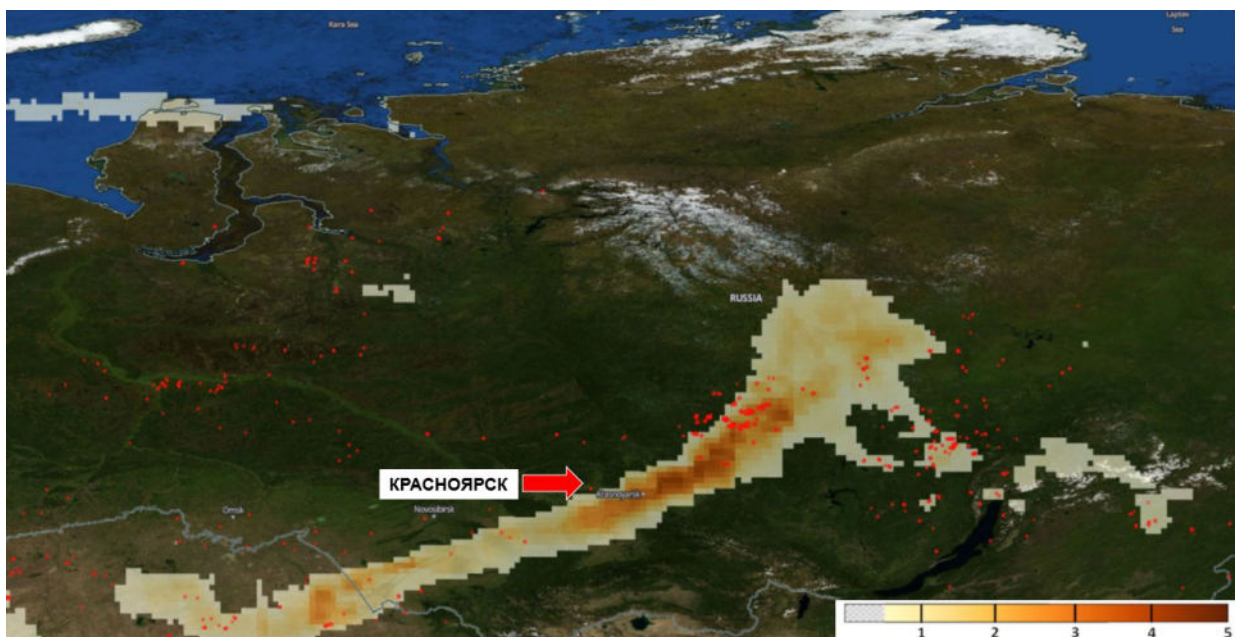


Рисунок 6 – Аэрозольный индекс UVAI по данным спутника NASA Suomi-NPP на 19 июля 2019 г. иллюстрирует распространение дымового шлейфа от лесных пожаров Якутии через Красноярск в направлении Казахстана на расстояние более 2500 км.

Объектом нашего исследования являлось поведение окружающей природной среды в г. Красноярске и его пригороде в августе 2021 года, когда в Якутии (Россия) возникли масштабные лесные пожары, дым от которых достиг г. Красноярска 7 августа и сохранялся до 10 августа. Для указанного периода времени анализировались данные метеостанций по температуре воздуха и почвы в пригороде Красноярска государственной наблюдательной сети ФГБУ «Среднесибирское УГМС», данные датчиков «Системы мониторинга воздуха КНЦ СО РАН», спутниковые данные по ультрафиолетовому аэрозольному индексу UVAI, формируемые с помощью прибора OMPS (Ozone Mapping Profiler Suite), установленного на метеорологическом спутнике Suomi-NPP, данные видеорегистрации таманов над рекой Енисей в Красноярске, полученные с помощью IP-видеокамеры ИВМ СО РАН, размещенной на городской телебашне Красноярска.

Анализ данных позволил сформулировать следующие результаты. Солнечное излучение, достигая земной поверхности, частично поглощается в ней. В результате верхний деятельный слой почвы, испытывает суточные и годовые колебания температуры. На рисунке 7 представлено поведение температуры деятельного слоя почвы, температуры воздуха и концентрации взвешенных частиц PM_{2.5} в приземном слое атмосферы в пригороде Красноярска с 6 по 9 августа 2021 г., а также значение индекса UVAI. На рисунке видно, что концентрация частиц PM_{2.5} начала увеличиваться после 22 часов 6 августа, имела два максимума 7-го и 8-го августа, и далее уменьшалась после 22 часов 8 августа. Фоновые и максимальные

значения концентрации PM2.5 составляли соответственно 25 и 300 мкг/м³. В отсутствии дымов 6 августа значение индекса UVAI равно нулю, что соответствует оптически тонкой атмосфере [9].

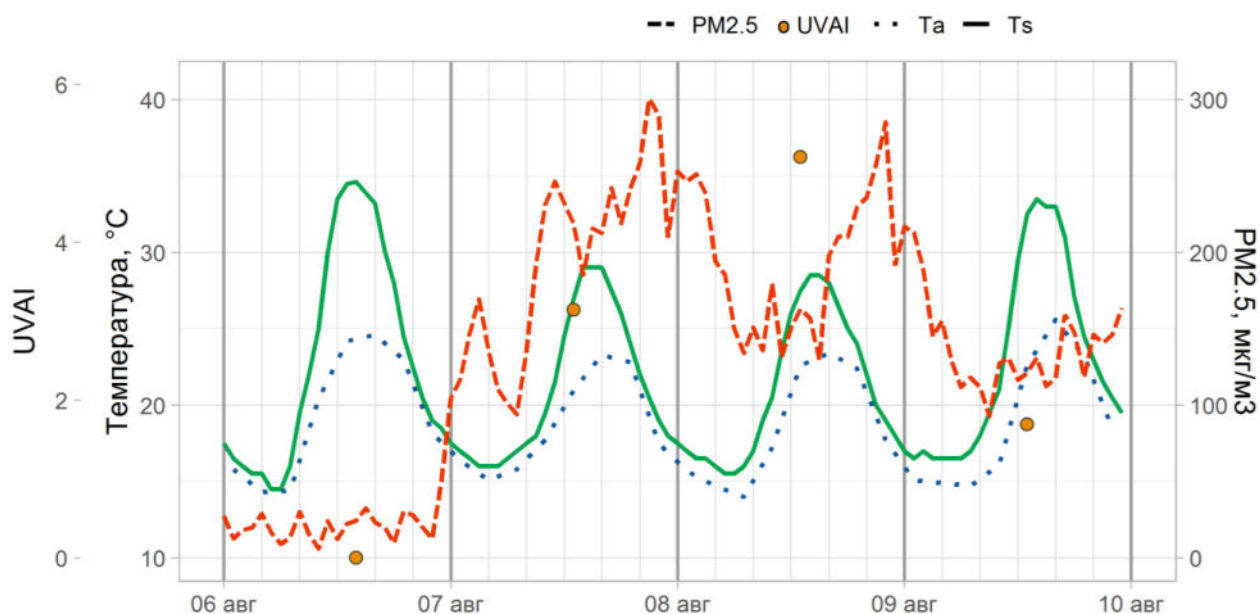


Рисунок 7 – Температура деятельного слоя почвы Ts (на глубине 0.5 см), температура воздуха Ta, концентрация PM2.5 (на высоте 2 м) и значение аэрозольного индекса UVAI в пригороде Красноярска.

По нашим оценкам поступление энергии в деятельный слой почвы в летнее время обусловлено в основном поглощением прямого и рассеянного солнечного излучения (W_R), а потери энергии связаны с излучением почвой тепловой инфракрасной радиации (ТИР) (W_S). В начале суток 6-го августа идет охлаждение почвы, обусловленное излучением ТИР (W_S), далее, после восхода Солнца (5 часов) начинается поглощение солнечного излучения W_R , и температура почвы достигает минимума. Положение минимума обусловлено равенством прихода и ухода энергии в деятельном слое почвы ($W_R = W_S$). В последующем за счет увеличения поглощаемого солнечного излучения температура почвы достигает максимума в полдень. После 13:00 интенсивность солнечного излучения падает, температура почвы уменьшается и после заката вновь определяющей является W_S . Повышение концентрации частиц PM2.5 вызванное дымами пожаров увеличивает оптическую толщину атмосферы, соответственно, уменьшает величину падающей солнечной радиации. Это приводит к уменьшению температуры деятельного слоя почвы 7 и 8 августа 2021 г.

Теплообмен между верхним слоем почвы и атмосферой осуществляется за счёт турбулентного процесса, а изменения температуры в нижнем приземном слое атмосферы в значительной степени определяются изменениями температуры поверхности земли. Темпера-

тура атмосферы в обоих случаях «следит» за ходом температуры почвы за счет турбулентного теплообмена.

На рисунке 8 представлены максимальные, минимальные и среднесуточные значения температуры почвы и атмосферы с 6 по 9 августа. Видно, что изменение температуры почвы, вызванное дымами, гораздо сильнее, чем изменение температуры атмосферы. Снижение максимальной дневной температуры T_s составляет 5.6°C , а T_a – 3.3°C .

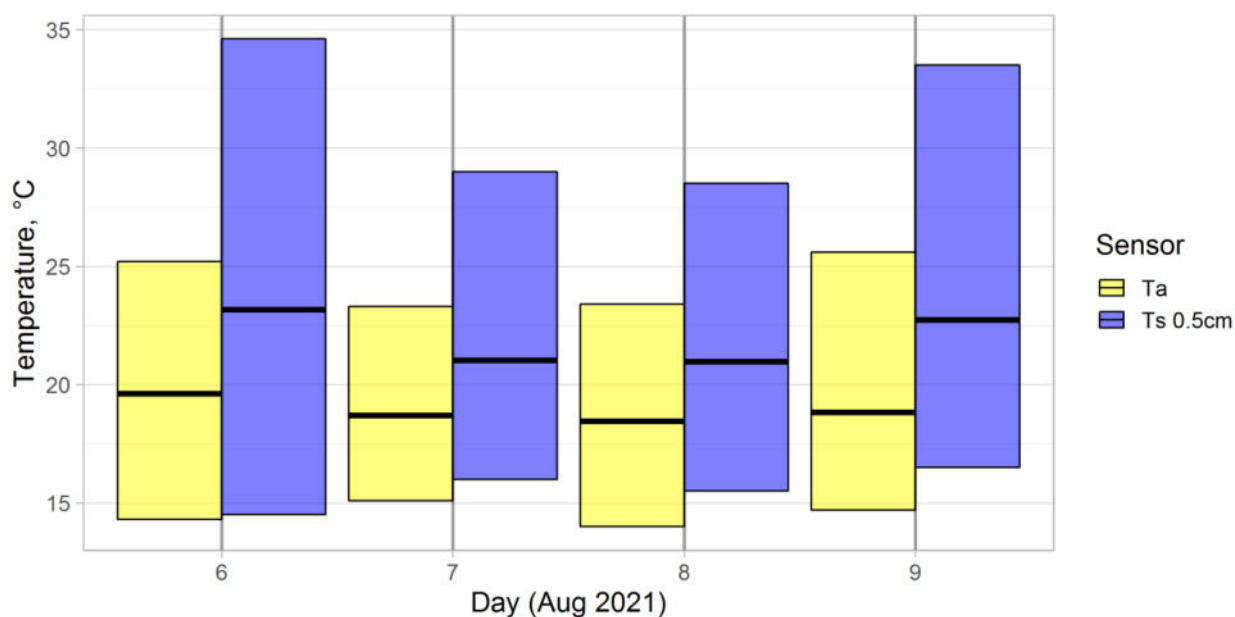


Рисунок 8 – Диапазоны изменения и среднесуточные значения температуры воздуха T_a и температуры почвы T_s .

Особое внимание в исследованиях было уделено формированию адвективного тумана охлаждения. После строительства Красноярской ГЭС в черте города в летнее время образуется «холодная» подстилающая поверхность на р. Енисей, а в зимнее время «тёплая» поверхность. В результате над поверхностью реки стали формироваться адвективные туманы охлаждения летом и туманы парения зимой.

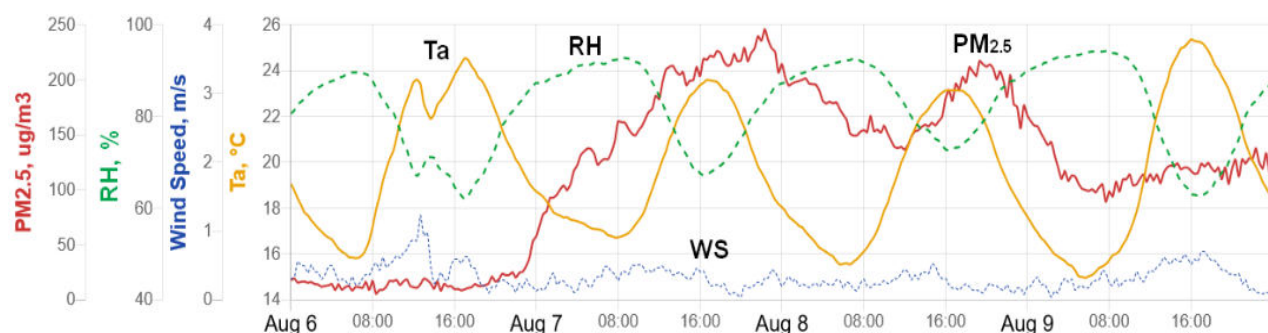


Рисунок 9 – Усредненные по г. Красноярску метеоданные за период с 6 по 9 августа 2021 г.

На рисунке 9 для рассматриваемого периода времени с 6 по 9 августа 2021 года представлены данные с нашего геопортала ИВМ СО РАН, усредненные по постам наблюдений в черте города Красноярска, включающие: концентрация взвешенных частиц в атмосфере (PM2.5), температура воздуха (Ta) относительная влажность воздуха (RH), скорость ветра (WS).

Согласно критерию, полученному нами ранее [10], адвективные туманы охлаждения на р. Енисей в естественных условиях формируются при условиях, когда относительная влажность $RH > 80\%$, а разность между температурой атмосферы Ta и температурой воды Tw определяются неравенством $0 < (Ta - Tw) < 6^\circ\text{C}$. При $(Ta - Tw) = 0^\circ\text{C}$ охлаждение атмосферы невозможно, а при $(Ta - Tw) \geq 6^\circ\text{C}$ влажный атмосферный воздух не успевает охладиться до точки росы. Эти условия, согласно данным на рисунке 9, выполнялись 9 августа 2021 г. в 5:10 ($Tw = 9.8^\circ\text{C}$, $RH = 94\%$, $Ta - Tw = 5.2^\circ\text{C}$).

На рисунке 10 представлено состояние атмосферы в виде изображений с городской веб-камеры № 3 за 6 и 9 августа 2021 г. в 05:10. 6 августа нет загрязнений в атмосфере, воздух прозрачный. 9 августа фрагментарная пелена обусловлена формированием тумана, это позволяет сделать вывод, что на изображении виден туман с примесью дыма от пожаров.



Рисунок 10 – Состояние атмосферы в Красноярске по данным городской веб-камеры № 3: а) 6 августа 05:10, б) 9 августа 05:10.

Основные итоги проведенного исследования состоят в следующем. Экспериментально установлено, что в период с 7 по 9 августа 2021 г. пожары обуславливали изменение концентрации PM2.5 от 25 до 300 мкг/м^3 . Наличие аэрозолей PM2.5 вызвало увеличение оптической толщины атмосферы и уменьшение величины поглощаемой солнечной радиации почвой. В период наблюдения мы зафиксировали холодный след якутских пожаров: уменьшение температуры верхнего деятельного слоя почвы ($\Delta T = 5.6^\circ\text{C}$), обусловленное уменьшением величины солнечной радиации за счет аэрозоли дымов; уменьшение температуры нижней части приземного слоя атмосферы, вызванное турбулентным теплообменом с почвой, которое составило ($\Delta T = 3.3^\circ\text{C}$); и формирование адвективных туманов охлаждения над поверхностью реки Енисей в г. Красноярске.

Список цитируемой в разделе литературы:

1. Crutzen P., Birks J. Twilight at noon: The atmosphere after a nuclear war // *Ambio*. 1982. Т. 11, № 2–3. С. 114–125.
2. Turco R.P., Toon O.B., Ackerman T.P., Pollack J.B., Sagan C. Nuclear Winter: Global Consequences of Multiple Nuclear Explosions // *Science*. 1983. Т. 222, № 4630. С. 1283–1292.
3. Seitz R. Siberian fire as “nuclear winter” guide // *Nature*. 1986. Т. 323, № 6084. С. 116–117.
4. Вельтищев Н.Н., Гинзбург А.С., Голицын Г.С. Климатические эффекты массовых пожаров // *Изв. АН СССР. Физика атмосферы и океана*. 1988. Т. 24, № 3. С. 296–305.
5. Ackerman A.S., Toon O.B., Stevens D.E., Heymsfield A.J., Ramanathan V., Welton E.J. Reduction of tropical cloudiness by soot // *Science*. 2000. Т. 288, № 5468. С. 1042–1047.
6. Панченко М.В., Журавлева Т.Б., Козлов В.С., Насртдинов, И.М., Полькин В.В., Терпугова С.А., Чернов Д.Г. Оценка радиационных эффектов аэрозоля в фоновых и задымленных условиях атмосферы Сибири на основе эмпирических данных // *Метеорология и гидрология*. 2016. № 2. С. 45–54.
7. Кондратьев К.Я. Аэрозоль и климат: современное состояние и перспективы разработок. 3. Аэрозольное радиационное возмущающее воздействие // *Оптика атмосферы и океана*. 2006. Т. 19, № 7. С. 565–575.
8. Ramanathan V., Carmichael G. Global and regional climate changes due to black carbon // *Nat. Geosci.* 2008. Т. 1, № 4. С. 221–227.
9. Herman, J. R. et al. Global distribution of UV-absorbing aerosols from Nimbus 7/TOMS data. // *J. Geophys. Res. Atmos.* 1997. 102, 16911–16922.
10. Shaparev N., Tokarev A. & Yakubailik O. Fogs on the Yenisei River (Krasnoyarsk, Russia) // *Thermal Science*. 2022. V. 26, No. 5B. P. 4447–4458.

Основные публикации по разделу:

1. Shaparev N., Tokarev A., Yakubailik O. Cold traces of smoke from wildfires in the environment // *Environmental Processes*. 2024. Т. 11. № 1. С. 5.
DOI: 10.1007/s40710-024-00687-6 [WoS Q3, Scopus Q1, БС-2]
2. Шапарев Н.Я., Токарев А.В., Якубайлик О.Э. Влияние дымов якутских пожаров на состояние природной среды в г. Красноярске в августе 2021 г // *Оптика атмосферы и океана*. 2023. Т. 36. № 6 (413). С. 487–493. DOI: 10.15372/AOO20230609. [RSCI, БС-2]
3. Якубайлик Т.В., Токарев А.В., Якубайлик О.Э. Почасовые данные о наличии туманов над рекой Енисей в г. Красноярске по данным видеонаблюдений за 2022 – 2023 гг. // Свидетельство о регистрации базы данных RU 2024623202, 19.07.2024.

1.3. Разработка инновационных методов анализа многомерных данных для задач моделирования сложных процессов в биосистемах (в том числе – в медицине)

Одной из важных задач данного проекта в области анализа больших данных сложных природных и антропогенных экосистем являлось исследование внутренней структуры данных и обнаружение скрытых взаимосвязей в них. Хотя методы такого анализа во многом определяются спецификой данных, эта зависимость не носит абсолютного характера. Подходы, разработанные для изучения параметров окружающей среды, могут быть эффективно адаптированы для других областей, таких как обработка медицинских данных или биоинформатика. Аналогично, успешные методики анализа медицинской информации могут быть использованы для моделирования сложных природных экосистем. В связи с этим в рамках проекта особое внимание было уделено моделированию сложных процессов в биосистемах, включая медицинские исследования.

Исследования отчетного периода по данному направлению были посвящены изучению особенностей и деталей структуры нуклеотидных последовательностей – одной из важнейших задач биологии в настоящее время. Исследования проводились в двух аспектах – структурно-функциональном и эволюционном. Выявление связи между структурными компонентами и соответствующим им функциями представляет собой классическую проблему молекулярной и системной биологии и несмотря на обширный поток публикаций и исследований в этом направлении, она всё ещё далека от завершения. Более того, исследователи выявляют всё новые и новые структурные элементы либо новые виды и формы взаимодействий и взаимоотношений между структурными элементами биологических макромолекул, этому способствует развитие техники и инструментов исследований. Кроме того, затруднения в исследованиях такого рода всегда вызывают выбор и качество того биологического материала, который берётся в рассмотрение. Это связано с большой сложностью таких объектов, как геномы либо отдельные хромосомы. Рассматривая эти объекты, приходится анализировать набор характеристик: структуру, функцию и филогению. Эти характеристики очень сильно взаимодействуют и сильно влияют друг на друга. Причем это влияние далеко не всегда удаётся выделить в качестве отдельного и независимого фактора.

Прокариотические организмы с этой точки зрения являются более удобными объектами для исследования, чем эукариотические; геном бактерий заметно короче генома эукариот и всегда представлен одной хромосомой. Однако ещё удобнее использовать геномы органелл, в нашем случае – митохондрий и хлоропластов, поскольку для них полностью исключается влияние различий в кодируемых функциях: в пределах одной группы органелл функциональные различия отсутствуют.

Такой параметр, как GC -состав очень часто используется в исследованиях структуры геномов и функций отдельных участков. Под GC -составом понимают долю гуанина и цитозина по отношению к общему числу нуклеотидов в рассматриваемой нуклеотидной последовательности.

Было изучено распределение значений GC -состава выделенных фрагментов геномов хлоропластов, митохондрий и бактерий в пространственной структуре генома, полученной на основе частотных словарей этих фрагментов. Рассматривался вопрос – имеется ли упорядоченность значений GC -состава в нуклеотидных последовательностях, либо они расположены случайным образом. Естественно было бы ожидать случайное распределение значений GC -состава, но как показывают результаты работы, наблюдается достаточно высокая упорядоченность такого распределения.

Будем рассматривать генетическую последовательность длины L , состоящую из символов алфавита $\mathfrak{K} = \{A, C, G, T\}$. Для этой последовательности будем составлять частотный словарь толщины 3. Частотный словарь W_3 толщины 3 символьной последовательности, соответствующей ДНК – это список всех троек $v_1 v_2 v_3$ идущих подряд нуклеотидов с указанием частот этих троек; всего может быть 64 триплета. Заметим, что приведенное определение частотного словаря триплетов является частным случаем: при подсчете числа триплетов окно считывания может перемещаться на один, два, три, четыре и вообще произвольное число нуклеотидов, тем самым порождая разные частотные словари. Как правило используется частотный словарь вида W_3^k , где $k = 1$.

Использовался частотный словарь вида W_3^3 , тем самым триплеты в нашей работе подсчитывались таким образом, что они полностью покрывают последовательность и при этом не пересекаются. Частота f_ω – это отношение числа копий n_ω данного триплета к общему числу всех триплетов N , где N – сумма всех n_ω :

$$f_\omega = \frac{n_\omega}{N} \quad (1)$$

Всякий частотный словарь W_3^3 отображает геном в 64-мерное метрическое (евклидово) пространство. То есть каждому фрагменту генома, для которого подсчитывается частотный словарь, ставится в соответствие точка в 64-мерном пространстве. Осям этого пространства соответствуют триплеты. Координатами точки являются значения частот триплетов. В результате разбиения генома на фрагменты он отображается во множество точек в пространстве триплетов.

Генетическая последовательность сканировалась окном длины Δ с шагом t . Для каждого положения i окна определялся участок генетической последовательности, для которого вычислялся частотный словарь $W_3^{(i)}$ соответствующий i -ой точке в 64-мерном пространстве. Кроме того, с каждой точкой в 64-мерном пространстве связывался номер центрального символа рассматриваемого участка, который совпадает с номером этого символа в последовательности. Для всех последовательностей длина окна $\Delta = 603$, шаг $t = 11$.

Для визуализации полученного множества точек использовалось программное обеспечение для кластерного анализа VidaExpert (<http://bioinfo-out.curie.fr/projects/vidaexpert/>), в котором строилась проекция данных в пространстве первых трех главных компонент, вычисленных для данного 64-мерного пространства. Тем самым сокращалось 64-мерное пространство до трехмерного; данные становились обозримыми. При обработке геномов было обнаружено, что точки располагаются достаточно однотипно в проекции на плоскость, определяемую 1-ой и 2-ой главными компонентами. Для более полной картины рассматривалась также плоскость, определяемая 2-ой и 3-ей главными компонентами.

Было рассмотрено 570 геномов хлоропластов, 60 геномов бактерий и 418 митохондрий растений и животных, в среднем по 25 видов на тип, класс или группу из базы EMBL. Для всех геномов, кроме геномов митохондрий наземных растений и печеночных мхов, наблюдается трехлучевая структура. Под такой структурой понималось распределение точек, представляющих частотные словари фрагментов в виде трех эллипсоидов (со значительным эксцентриситетом), пересекающихся одним из своих концов. В месте пересечения концентрируются точки, соответствующие некодирующим областям генома. При этом большие оси обнаруженных эллипсоидов отстоят на равные углы. Отметим, что для геномов митохондрий наземных растений и печеночных мхов все три эллипсоида совпадают и образуют единый эллипсоид.

Для всех геномов здесь и далее интервал имеющихся значений GC-состава разбивался на 7 подинтервалов равной длины. Интервалу с минимальными значениями соответствует фиолетовый цвет, с максимальными – красный (Рис. 11).

Обнаруженная упорядоченность имеет типичный вид для отдельных групп геномов. Наибольшую однородность распределения величины GC-состава показали геномы хлоропластов наземных растений. Для них характерно градиентное распределение. Тип распределения для геномов бактерий обусловлен величиной GC-состава генома в целом. Для геномов, величина GC-состава которых менее 50%, наблюдается центральносимметричное распределение. Для геномов с величиной GC-состава больше 50% распределение носит градиентный характер.

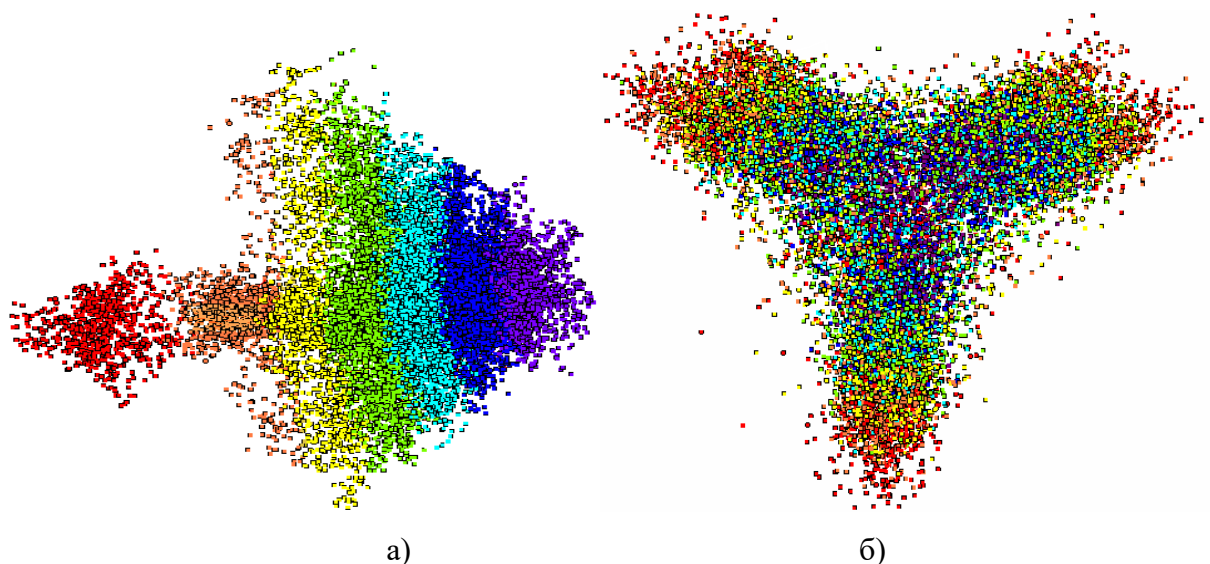


Рисунок 11 – Градиентное распределение на примере *Anthoceros angustus* в пространстве 1 и 2 главных компонент (а) и центральносимметричное распределение на примере *Clostridium baratii* str. Sullivan (б).

Наибольшим разнообразием в типах распределений GC-состава фрагментов обладают геномы митохондрий. Тем не менее, можно проследить некоторые закономерности и для этих геномов. Так, для геномов митохондрий растений (наземных растений, одноклеточных и многоклеточных водорослей, обычных и печеночных мхов), а также для высших и низших грибов характерно градиентное распределение. Для геномов митохондрий насекомых, паукообразных и ракообразных характерно центральносимметричное распределение. Для всех остальных групп геномов митохондрий животных наблюдались оба типа распределений в каждой группе. Возможно, это связано с более сложным строением организмов и разнообразием среды их обитания.

Градиентное или центральносимметричное распределение GC-состава определяется скалярным произведением вектора, компонентами которого будут значения GC-состава всех триплетов, на вектора главных компонент. Построим 64-мерный вектор GC-состава триплетов следующим образом. Триплету AAA будет соответствовать значение GC-состава 0, триплету AAC будет соответствовать значение GC-состава 1, триплету ACC – значение 2 и т.д. Нормируем этот вектор, чтобы сумма всех компонент была равна 1. Анализ скалярного произведения показал, что при градиентном распределении значений GC-состава максимальное по абсолютной величине значение скалярного произведения наблюдается при умножении вектора GC-состава триплетов на первую или вторую главную компоненты. Для центральносимметричного распределения максимальное по абсолютной величине значение скалярного произведения наблюдается при умножении вектора GC-состава триплетов на третью или четвертую главные компоненты.

Чтобы выяснить, насколько устойчивым является обнаруженное пространственное распределение GC-состава фрагментов геномов, были рассмотрены такие распределения для различной длины окна Δ . Распределения были рассмотрены в интервале сохранения тех структур геномов, которые представлены в статье. Для геномов хлоропластов и митохондрий этот интервал Δ равен от 300 до 1500, для бактерий от 600 до 3000. Поскольку выше мы рассматривали распределение величины GC-состава фрагментов геномов для длины окна $\Delta = 600$, то для выяснения устойчивости распределения были взяты длины окна Δ равные 300, 900, 1200 и 1500 для хлоропластов и митохондрий, и 900, 1200, 1500 и 3000 для бактерий. На рисунке 12 показаны полученные распределения. В качестве примера для хлоропластов взят геном *Triticum aestivum* (ID AB042240), для бактерий *Clostridium baratii* str. *Sullivan* (ID CP006905) и для митохондрий *Cladonia petrophila* (ID MG941021).

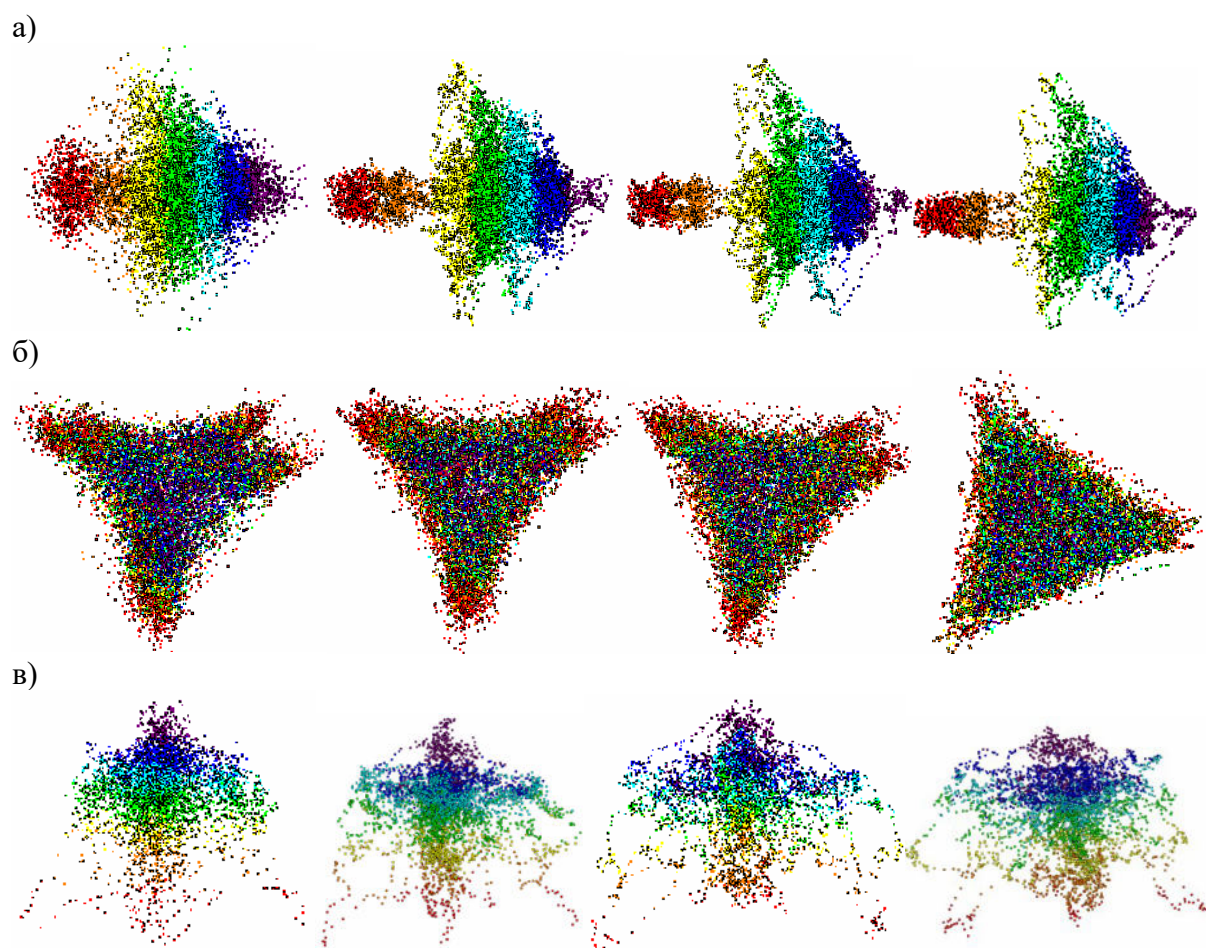


Рисунок 12 – Распределение величины GC-состава фрагментов геномов хлоропластов а), бактерий б) и митохондрий в) для различной длины окна Δ .

Структуру биологических макромолекул можно определить многими различными способами. В работе под структурой понимается частотный словарь (частотный профиль)

триплетов – список всех триплетов, встречающихся в исследуемой последовательности, с указанием частот этих триплетов. Построение частотного словаря преобразует нуклеотидную (символическую) последовательность в точки в 64-мерном метрическом пространстве, делая их математическими объектами, позволяющими использовать весь арсенал соответствующих инструментов для вашего исследования. Изучены особенности кластеризации частотных словарей триплетов генов бактериальной 5S РНК с точки зрения связи между такой кластеризацией и таксономическим составом кластеров. Отметим, что бактериальная 5S РНК не является наиболее широко используемым объектом для такого рода исследований, и сравнение полученных результатов с другими, например, полученными на основе кластеризации по генам бактериальной 16S РНК, определяет актуальность данной работы. Проанализирована устойчивость наблюдаемой кластеризации (если таковая имеется) к случайному выбору исключенных перепредставленных таксонов.

База данных генов SILVA содержит всего 182697 записей бактериальных генов 5S РНК. Для анализа были выбраны гены 5S РНК бактерий, относящихся к 23 типам. Количество генов для разных типов сильно отличалось. Было создано 6 баз данных, в которых записи генов в таксонах с низкой представленностью (т.е. с небольшим числом видов в них, до 30 записей) были исключены из анализа, т.к. такие данные не способны сформировать сигнал должной силы, однако формируют шум, существенно искажающий картину кластеризации. Гены в перепредставленных таксонах были частично удалены случайным образом. Это нужно для того, чтобы подавить слишком сильный сигнал от такой группы организмов, который также может существенно исказить результаты кластеризации.

Было сформировано 6 баз данных, полученных из исходных данных удалением генов в перепредставленных таксонах. Каждая из шести баз данных была обработана упругой картой с одинаковыми значениями параметров. Сравнительный анализ распределений (Рис. 13) позволяет утверждать, что влияние состава базы, на которой строится кластеризация, имеет место, однако это влияние невелико. Действительно, если взять любую пару карт, представленных на рисунке 4, и сравнить их между собой, то можно увидеть, что буквального, по-пиксельного соответствия нет. Другими словами, если две такие карты положить одну на другую, они не совпадут. С другой стороны, видимые различия между этими картами очень малы. Все эти карты, очевидно, визуальны очень близки друг к другу. Фактически, это наблюдение дает ответ на главный вопрос данной работы: правда ли, что разные базы, полученные случайным образом, дают близкий или даже идентичный результат.

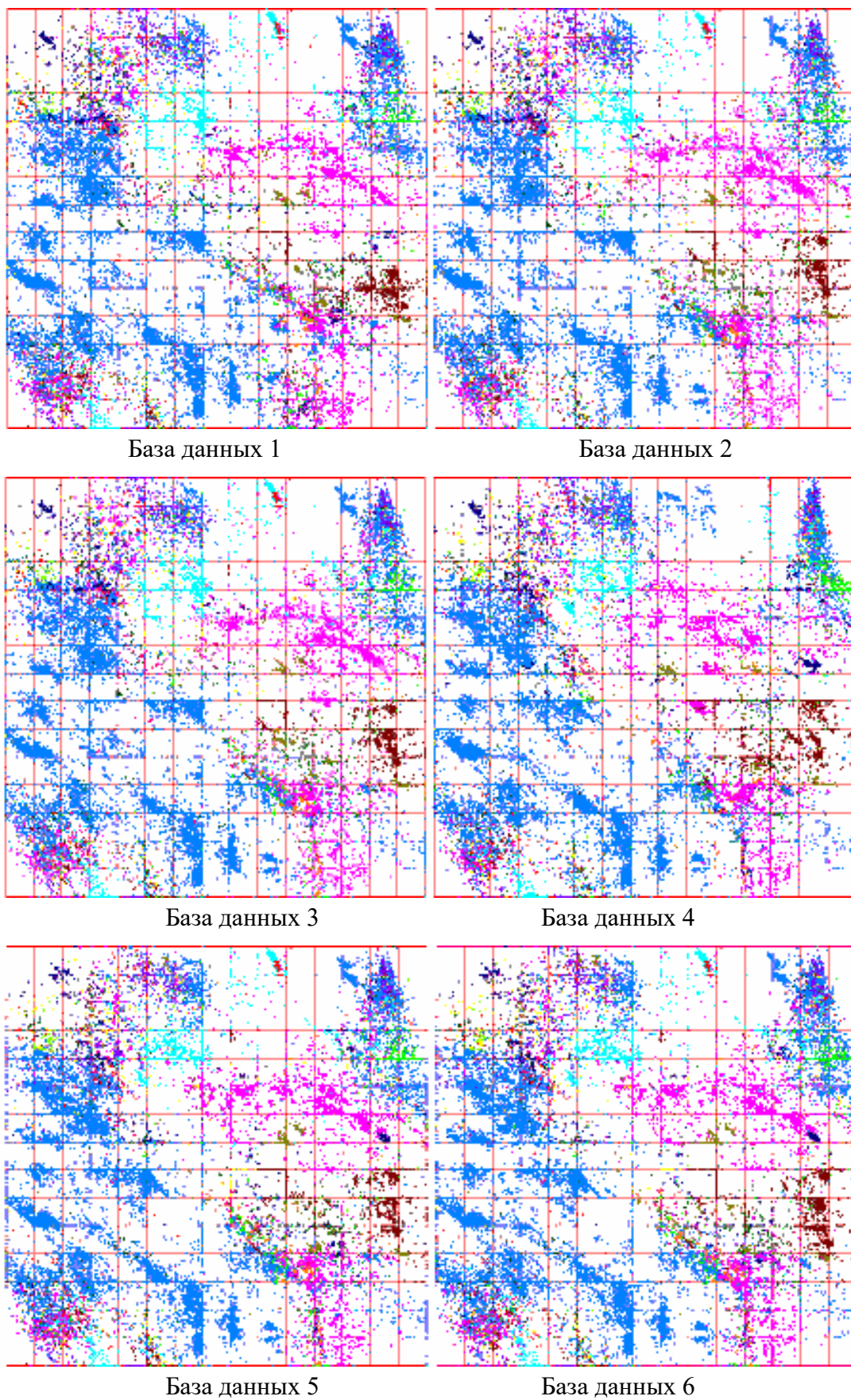


Рисунок 13 – Распределение всех двадцати трех типов на упругой карте в шести различных версиях бактериальной геномной базы 5S РНК.

Ответ на вопрос об отсутствии сильного влияния случайной индексации на картину кластеризации с точки зрения таксономического состава выявленных кластеров положительный: такое влияние следует считать весьма слабым и незначительным для дальнейших исследований. Результаты, полученные в данной работе, доказывают, что состав базы данных, определяемый случайным исключением перепредставленных записей, не оказывает существенного влияния на результаты кластеризации. Это наблюдение позволяет в дальнейшем использовать любую из индексированных баз, полученных случайным исключением перепредставленных таксонов, для исследования взаимосвязи структуры, функции и таксономии.

Основные публикации по разделу:

1. Senashova M.Yu., Sadovsky M.G. The Distribution of the GC Content in the Spatial Structure of the Fragments of Mitochondrial, Chloroplast, and Bacterial Genomes // Biophysics, V. 69(2), p. 209-218. DOI: 10.1134/S0006350924700234 [Scopus, BAK]
2. Сенашова М.Ю. Упорядоченность значений GC-состава фрагментов в пространственной структуре геномов органелл // Информационные и математические технологии в науке и управлении, 2024 № 2 (34), с. 33-40.
DOI: 10.25729/ESI.2024.34.2.003 [BAK]
3. Ovchinnikova Iu., Sadovsky M., Sokolov V. Clustering of bacteria by the triplet composition of 5S RNA // ITM Web of Conferences. 2024. V. 59, 04014.
DOI: 10.1051/itmconf/20245904014

1.4. Влияние городской среды на формирование локальных климатических зон и тепловых аномалий на примере Красноярска

Город с промышленными объектами, асфальтированными дорогами, пробками на дорогах и железобетонными зданиями, отличается климатом от окружающей местности [1]. В городе другой климат, называемый городским, который отличается от окружающей местности температурой, ветром, влажностью и другими метеорологическими характеристиками. Сегодня эти различия еще более заметны, поскольку климат меняется [2].

Исследование по настоящему проекту было направлено на изучение связи распределения тепловых аномалий и городской среды Красноярска. Проблема качества воздуха весьма актуальна для г. Красноярска – это промышленно развитый город со множеством источников выбросов, окруженный горами и расположенный глубоко в Евразии. Котловинный характер рельефа местности города способствует ограничению циркуляции воздушных масс, что приводит к накоплению холодного воздуха, загрязняющих веществ и туманов в

приземном слое атмосферы. В сочетании со штилевыми условиями и температурными инверсиями, особенно в зимний период, это создает предпосылки для формирования неблагоприятных метеорологических условий (НМУ). В результате в приземном слое атмосферного воздуха происходит интенсивное накопление загрязняющих веществ, что ухудшает качество воздуха и способствует возникновению смога. Экологи часто фиксируют превышение показателей загрязняющих веществ, таких как бенз(а)пирен, гидрохлорид, взвешенные частицы РМ, и проч. – по данным Минприроды Красноярск считается одним из самых загрязненных городов России. Тепловое загрязнение также является неблагоприятным фактором для экологического состояния и комфорта городской среды.

В качестве данных ДЗЗ использованы спутниковые снимки Landsat. Эти спутники эксплуатируются с 1972 года, и в течение десятилетий создавался архив спутниковых снимков [3]. Программа, разработанная для изучения суши, внесла значительный вклад в ряд научных областей, начиная от естественных и заканчивая социальными [Zhou, 2019]. Landsat 8, запущенный на орбиту 11 февраля 2013 года, получает данные с помощью двух разных датчиков – Operational Land Imager (OLI) и теплового инфракрасного датчика (TIRS). Снимки Landsat 8 состоят из 11 спектральных диапазонов, из которых 10-й и 11-й – дальние инфракрасные диапазоны с пространственным разрешением 100 м [4]. В марте 2021 года стали доступны новые уровни обработки Landsat 8 Collection 2 Level 2. С 10 февраля 2022 года появились данные нового аппарата Landsat 9 – это последний спутник, он был запущен в сентябре 2021 года и сейчас полностью работоспособен после прохождения некоторых первоначальных испытаний [5]. Landsat 9 имеет два основных прибора: OLI-2 и TIRS-2. После запуска нового спутника Landsat 9 количество доступных изображений увеличилось примерно вдвое.

В работе авторов [6] подробно описан алгоритм вычисления температуры поверхности (LST) по данным Landsat 8/9 Collection 2 Level 2 для территории г. Красноярска с использованием собственного ПО. В этом наборе данных уже проведена атмосферная, геометрическая и радиометрическая коррекция и рассчитана температура LST. Необходимость создания собственного ПО для расчета LST вызвана тем, что в архиве Landsat данные 2-го уровня обработки отсутствуют на часть территории города (> 25%).

Наблюдение за температурой LST поверхности по спутниковым данным с 2013 года позволяет выделить характерные для территории тепловые аномалии, которые возникают в границах городского острова тепла (ГОТ). Температура воздуха так же имеет свой архив по данным сети наблюдений «Системы мониторинга воздуха КНЦ СО РАН» с 2020 г., который необходимо сопоставить со спутниковыми данными. По определению ГОТ – это различие температуры, наблюдаемое в городе от температуры окружающей его местности. В данной

работе рассматривалась температура поверхности в границах города по данным Landsat 8/9 с разрешением теплового канала 100 м. А температура воздуха – по данным 20-ти станций мониторинга КНЦ СО РАН. Такая детализация позволяет исследовать неоднородность распределения температуры поверхности и воздуха внутри ГОТ.

Для изучения ГОТ из каталога данных были выбраны спутниковые снимки Landsat 8 за период с 2013 по 2022 гг. и Landsat 9 за 2022 г. Выбирались снимки в бесснежный период с отсутствием облаков над городом, по этим критериям было собрано 35 снимков. Для каждого изображения были выделены тепловые аномалии с использованием методики, ранее предложенной авторами [7].

На 1-м этапе производится вычисление LST по данным Landsat 8/9. Следующим 2-м этапом определяются повышенные значения температур – это верхний квартиль диапазона (от 75 до 100%). Далее растровые значения преобразуются в векторный формат, все полигоны со значениями температуры меньше, чем квартиль удаляются. Из оставшихся полигонов необходимо удалить все контуры размер, которых менее трех пикселей. Это объясняется тем, что пиковые значения такого размера могут быть или ошибкой, или точечным объектом (например, металлическая крыша), но при этом тепловая аномалия не распространяется за пределы самого объекта.

Оставшиеся полигоны авторами работы определены как контуры тепловых аномалий. Накопленный архив контуров тепловых аномалий внутри ГОТ Красноярска позволяет проанализировать динамику их изменения с 2013 по 2022 годы. Все полученные контуры тепловых аномалий были сгруппированы по трем временам года: весна, лето и осень, т. к. было замечено, что сезонность влияет на распределение городского острова тепла. А также были выделены две группы постоянных и переменных контуров тепловых аномалий. В группу постоянных вошли те контуры, которые максимально сохранили свои границы в период с 2013 по 2022 год вне зависимости от сезона. В работе анализируется именно постоянные контуры тепловых аномалий.

При изучении городской среды в качестве минимального участка местности выбраны локальные климатические зоны (ЛКЗ). ЛКЗ представляют собой территории с однородным покрытием поверхности, структурой, материалами и особым характером деятельности человека и имеют специфический тип взаимодействия с приземным слоем атмосферы [8]. Впервые метод выявления ЛКЗ, предложен Стюартом И.Д. и Оке Т.Р., он представляет собой классификацию, состоящую из 17 типов зон, основанную главным образом на свойствах подстилающей поверхности, этажности зданий и плотности застройки. Классификация является универсальной, для исследования были выбраны только классы городской среды, некоторые классы были исключены из-за отсутствия, а некоторые были объединены. Границы

классов векторизовались вручную по открытым данным Open Street Map. В итоге для анализа распределения тепловых аномалий территории г. Красноярска были выбраны только классы с различными типами застройки: малоэтажная жилая застройка; многоэтажная жилая застройка; крупногабаритная застройка; промышленность. Два последних класса были объединены в один, потому что их поверхности идентичны, на обеих территориях есть строения большой площади, которые окружены парковками или другими асфальтобетонными покрытиями, они не имеют зеленых насаждений рядом.

На рисунке 14 показано расположение постов наблюдений «Системы мониторинга воздуха КНЦ СО РАН», соотношение постоянных контуров тепловых аномалий и ЛКЗ в границах г. Красноярска.

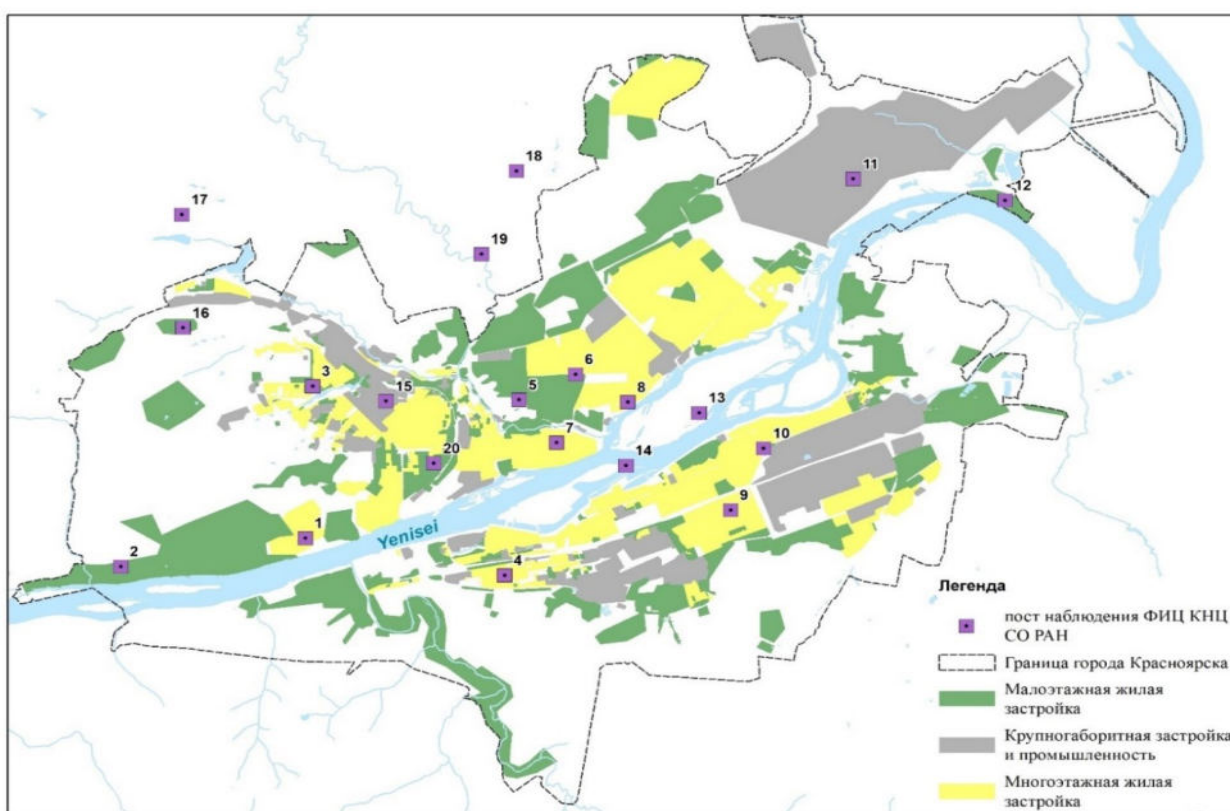


Рисунок 14 – Распределение станций мониторинга «Системы мониторинга воздуха КНЦ СО РАН» и сформированных локальных климатических зон (ЛКЗ).

Для изучения тепловых аномалий в ГОТ, необходимо учитывать дополнительные метеорологические параметры такие параметры, как: атмосферное давление, влажность, температура воздуха, индекс качества воздуха (AQI). В каталог данных исследования добавлены эти параметры за 2020, 2021 и 2022 годы в дни съемки территории исследования Landsat 8/9. Из-за облачности выбор спутниковых снимков сильно ограничен: были выбраны дни, максимально подходящие к разным временам года. В 2020 году было очень мало доступных сцен, и они всем больше относятся к летнему сезону. В 2021 и 2022 году данные

достаточно хорошо отображают сезонные изменения. В таблице 1 представлены метеорологические среднесуточные параметры в соответствующие даты исследования, полученные на посту «Покровка» «Системы мониторинга воздуха КНЦ СО РАН».

Таблица 1. Метеорологические условия на даты исследования.

Дата	Атмосферное давление	Влажность, %	Температура воздуха, °С	Индекс качества воздуха, РМ _{2.5}
15 августа 2020	733.79	75.09	27.53	36
12 июня 2020	736.18	51.43	22.69	8
20 мая 2020	740.35	27.18	26.3	15
21 апреля 2021	744.83	36.72	1.5	14
26 июля 2021	729.44	92.11	30.12	60
5 октября 2021	746.46	100	6.5	78
17 мая 2022	735.87	60.1	22.14	38
12 июня 2022	732.61	51.59	20.32	11
14 сентября 2022	747.43	44.89	12.21	57

Результаты исследований.

Контуры тепловых аномалий определены для трех времен года (весна, лето, осень). Весенние острова тепла составляют 81 полигон, осенние острова тепла – 21 полигон, летние острова тепла – 82 полигона. Зимний период не рассматривался из-за отсутствия бесснежного периода в это время года. Высокие значения температур поверхности более заметны летом и весной, именно в это время преобладает количество контуров тепловых аномалий. Стоит так же отметить, что тепловые аномалии находятся не только в местах городской застройки промышленными объектами, но и на природных возвышенностях, учитывая специфику сформированных ЛКЗ в данной работе не будем рассматривать природные тепловые аномалии, а только промышленные и городские. Для проведения анализа был сформирован один общий векторный слой контуров тепловых аномалий, который авторы формировали по принципу максимального пересечения контуров в разные сезоны. Таким образом, было выделено наиболее значимых 24 контура. Стоит отметить, что преобладающее большинство тепловых аномалий расположено в ЛКЗ крупногабаритная застройка и промышленность (рис. 15).

При сравнении карт температуры поверхности 2020 и 2022 годов были обнаружены изменения городской среды. На месте Комбайнового завода были построены жилые дома, следовательно классы ЛКЗ необходимо изменить с «крупногабаритная застройка и промышленность» на «многоэтажная жилая застройка».

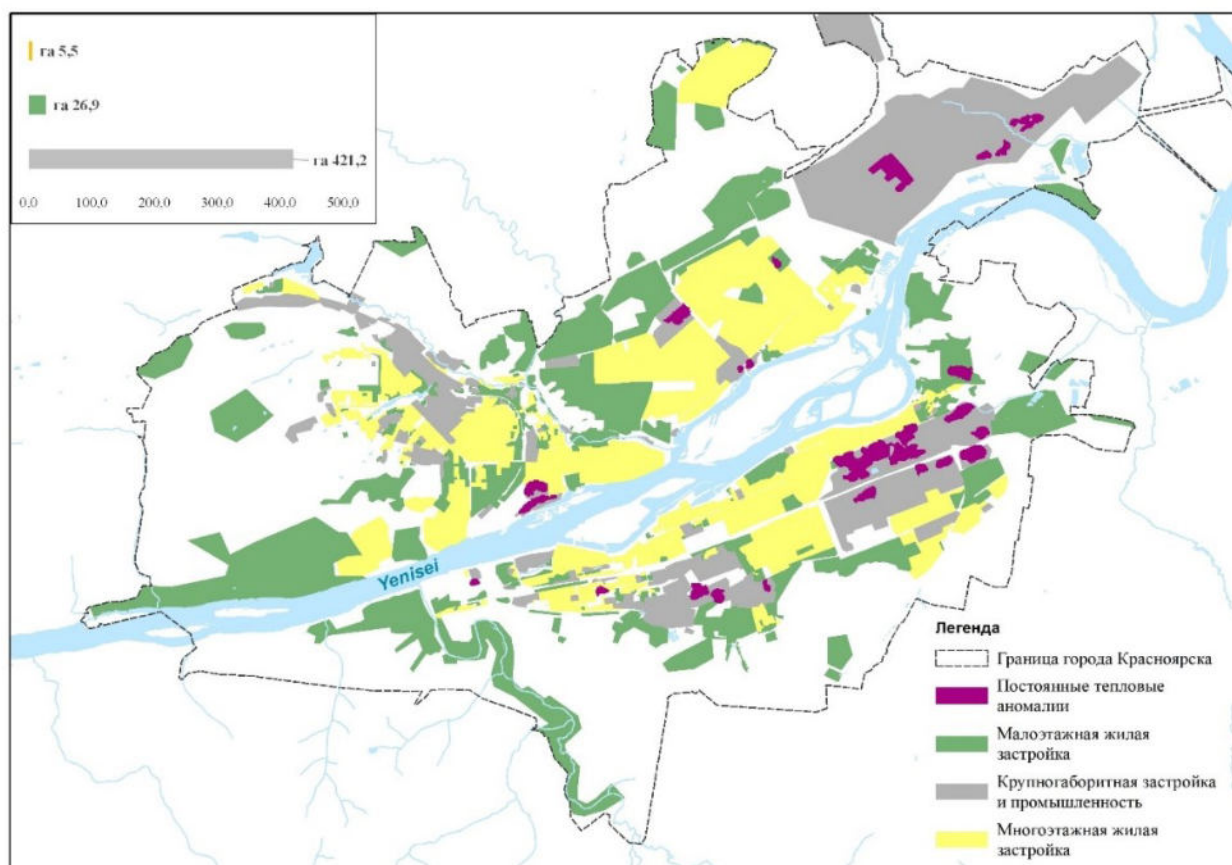


Рисунок 15 – Распределение локальных климатических зон (ЛКЗ) в постоянных контурах тепловых аномалий.

Расширение городских границ тоже изменяет классы ЛКЗ. В городе ведется активная застройка по нескольким направлениям (в северо-западном, северном и юго-восточном), в основном новые территории классифицируются, как многоэтажная жилая застройка. Смена класса ЛКЗ с жилой на крупногабаритную застройку и промышленность, достаточно редкая, и в период с 2020 до 2022 не обнаружена.

Анализ выбранных данных за 2020 – 2022 гг. показывает, что в целом контуры тепловых аномалий, полученных по LST, совпадают с высокими значениями температуры воздуха со станций наземного мониторинга. Следует отметить, что 20 станций на территорию большого города представляется недостаточным для выявления неоднородностей температурного поля. Для заполнения пропуска значений проведена интерполяция температуры с использованием метода обратно взвешенных расстояний.

На рисунке 16 показаны карты интерполяции по температуре воздуха на даты исследования, температура в зависимости от сезона меняется от -2°C до $+33^{\circ}\text{C}$. При разных значениях температуры воздуха наблюдается увеличение значений в точках наблюдения.

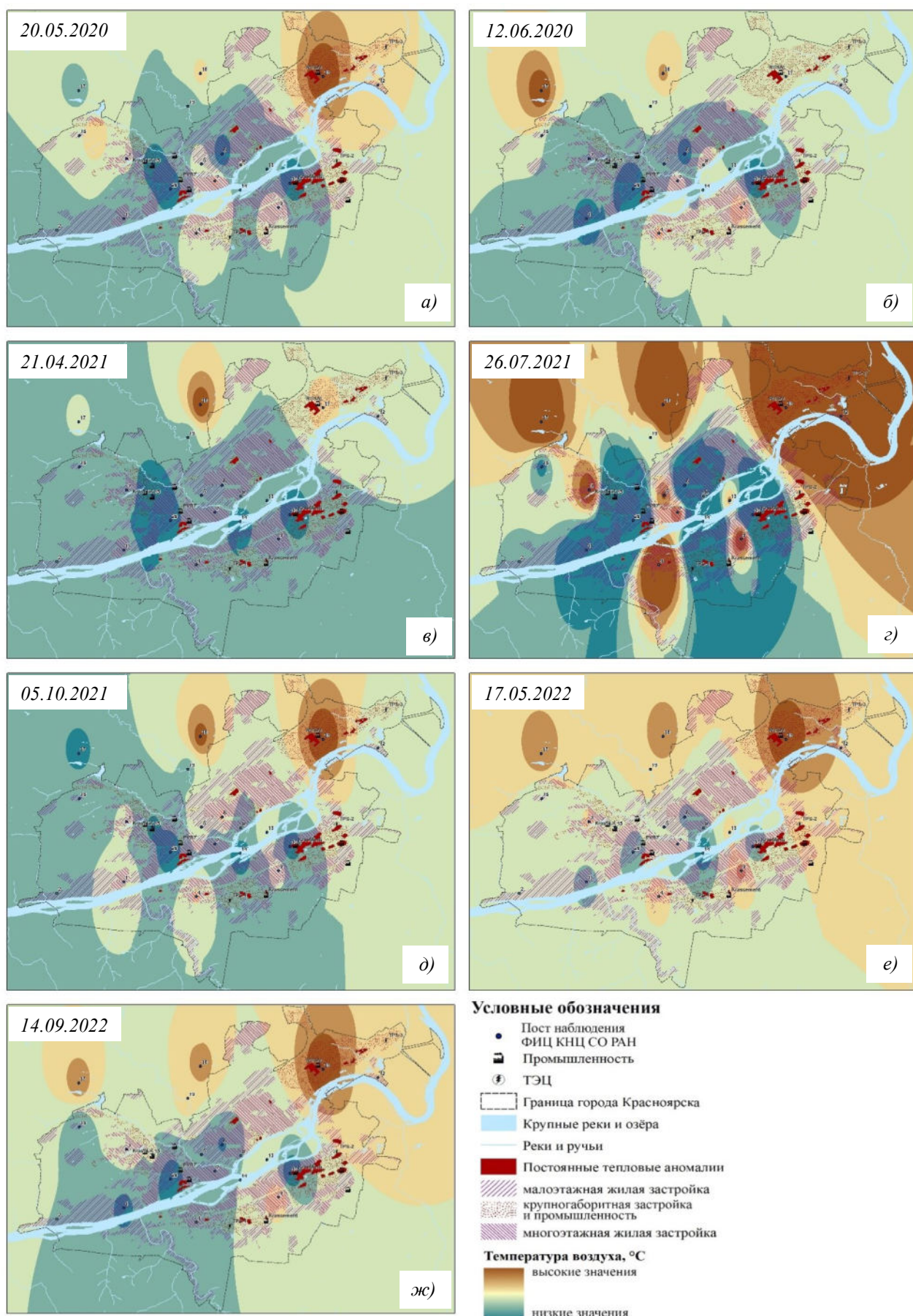


Рисунок 16 – Карта интерполяции температуры воздуха: а) 20 мая 2020, б) 12 июня 2020, в) 21 апреля 2021, г) 26 июля 2021, д) 5 октября 2021, е) 17 мая 2022 г. ж) 14 сентября 2022.

На 6-ти из 7-ми изображений, максимальные значения температуры зафиксированы в пункте наблюдений 11, в этом месте находится крупный промышленный объект – алюминиевый завод (КрАЗ).

На рис. 16а, средняя температура воздуха за 20.05.2020 составляет 26,3 °С (по табл. 1), пиковые значения в районе КрАЗа = 29,66 °С, остальная температура воздуха распределена более равномерно. Похожая ситуация повторяется 14.09.22 (рис. 16ж).

В пунктах наблюдений 17 и 18 на большинстве изображений тоже зафиксированы максимальные температуры. Точки мониторинга 17 и 18 не относятся к промышленным и находятся за пределами города.

Основные выводы состоят в следующем. Исследование тепловых аномалий за несколько лет по данным наземного мониторинга и тепловых спутниковых снимков позволило выделить на территории города локальные климатические зоны (ЛКЗ), которые были сгруппированы в три класса. Оценивая взаимное расположение постоянных контуров тепловых аномалий и ЛКЗ, можно отметить, что расположение контуров тепловых аномалий совпадает с классом крупногабаритной застройки и промышленных предприятий.

Комплексный подход к изучению городской среды позволяет расширить представление о взаимодействии различных факторов, например, промышленных объектов, вырубки лесов, сложного ландшафта и состояния окружающей среды. Учитывая неравномерность получения доступных тепловых данных ДЗЗ Landsat – обычно 3-4 безоблачных изображения за сезон – важно определить постоянные контуры тепловых аномалий и найти их взаимное расположение с ЛКЗ. Общие тенденции распределения температуры сохраняются как при изучении LST, так и при интерполяции температуры воздуха. Вклад промышленных объектов отображается на картах интерполяции температуры воздуха.

Геопространственные инструменты и инструменты дистанционного зондирования для городских исследований дают возможность планировать этапы городского развития, чтобы смягчить эффект воздействия тепловых аномалий на городскую среду.

Список цитируемой в разделе литературы:

1. Cheval S., Micu D., Dumitrescu A., Irimescu A., Frighenciu M., Iojă C., Tudose N.C., Davidescu Ș., Antonescu B. Meteorological and Ancillary Data Resources for Climate Research in Urban Areas // Climate. 2020, 8, 37. DOI: 10.3390/cli8030037.
2. Borzino N.; Chng S.; Mughal M.O.; Schubert R. Willingness to Pay for Urban Heat Island Mitigation: A Case Study of Singapore // Climate. 2020, 8, 82. DOI: 10.3390/cli8070082
3. Liu L., Zhang Y. Urban heat island analysis using the landsat TM data and ASTER Data: A case study in Hong Kong // Remote Sens. 2011, 3, 1535-1552.

4. Malakar N.K., Hulley G.C., Hook S.J., Laraby K., Cook M., Schott J.R. An operational land surface temperature product for Landsat thermal data: Methodology and validation // IEEE Trans. Geosci. Remote Sens. 2018, 56, 5717-5735.
5. Choate M.J., Rengarajan R., Storey, J.C., Lubke M. Landsat 9 Geometric Characteristics Using Underfly Data // Remote Sens. 2022, 14, 3781. DOI: 10.3390/rs14153781
6. Гостева А. А., Матузко А. К., Якубайлик О.Э. Алгоритм вычисления температуры поверхности для восстановления потери данных Landsat 8-9 Collection 2 Level 2. // ИнтерКарто. ИнтерГИС. 2023. Т. 29, № 1. С. 318-329. DOI 10.35595/2414-9179-2023-1-29-318-329.
7. Matuzko A.K., Yakubailik O.E. Urban heat island effects over Krasnoyarsk obtained on the basis of Landsat 8 remote sensing data // IOP Conference series: Earth and Environmental Science. 2018. V. 211. article 012010. DOI: 10.1088/1755-1315/211/1/012010.
8. Самсонов Т.Е., Тригуб К.С. Картографирование локальных климатических зон Москвы по космическим снимкам // Геодезия и картография, 2018. Т. 79. № 6. С. 14-25. DOI: 10.22389/0016-7126-2018-936-6-14-25.

Основные публикации по разделу:

1. Гостева А.А., Матузко А.К. Влияние городской среды на формирование локальных климатических зон и тепловых аномалий // ИнтерКарто. ИнтерГИС. 2024. Т. 30. Ч. 2. С. 543–555. DOI: 10.35595/2414-9179-2024-2-30-543-555. [RSCI, ядро РИНЦ, ВАК]
2. Гостева А.А., Матузко А.К., Таныгин К.Е. Разделение территории города Красноярска на локальные климатические зоны по спутниковым данным / Региональные проблемы дистанционного зондирования Земли : материалы XI Междунар. науч. конф. Красноярск, 10–13 сентября 2024 г. / отв. ред. Р. В. Брежнев – Красноярск : Сиб.федер. ун-т, 2024. С. 204-207.
3. Гусев Б.А., Гостева А.А., Тамаровская А.Н. Особенности пространственного распределения тепловых островов в городе / Региональные проблемы дистанционного зондирования Земли : материалы XI Междунар. науч. конф. Красноярск, 10–13 сентября 2024 г. / отв. ред. Р. В. Брежнев – Красноярск : Сиб.федер. ун-т, 2024. С. 226-229.

1.5. Оценка экологического статуса водоемов с разной степенью антропогенной нагрузки методами биоиндикации

Оценка качества вод, в том числе и поверхностных, является одной из наиболее актуальных и приоритетных задач устойчивого развития водного хозяйства России. Биологический контроль в оценке экологического состояния водных объектов зачастую оказывается приоритетным [1, 2], и наиболее перспективными индикаторами общепризнаны организмы зообентоса [2, 3]. Рекомендуется использовать комплексный подход, включающий набор разнообразных биоиндикационных показателей [1–4]. Однако отсутствие единой общепринятой балльной градации снижает чувствительность использованных показателей в градиенте загрязнения, что приводит к искажению полученных значений категорий качества воды. В качестве обобщающей характеристики качества воды предлагается рассматривать различные интегральные показатели, основанные на совокупности нескольких биотических индексов [1–4].

Экологический статус реки или речного бассейна возможно оценивать по отношению к эталонным створам, расположенным на участках, не подверженных антропогенным воздействиям [1, 5]. Однако многообразие различных водных экосистем на территории нашей страны, связанное с природными условиями, затрудняет выбор эталонных биологических сообществ. Усугубляет положение недостаток отечественных гидробиологических исследований на удаленных от антропогенного влияния фоновых территориях в разных климатических зонах. Как правило, такие территории являются труднодоступными, а их исследование сопряжено с большими финансовыми затратами. Плотность гидрологической сети Росгидромета в Красноярском крае составляет 6 тыс. км² на одну гидрологическую станцию, что явно недостаточно, особенно в условиях освоения северных территорий края.

Альтернативным способом обобщения множества факторов и показателей является использование многомерной статистики, а именно подходов к классификации на основе различных метрик сходства. Данные методы позволяют проводить типизацию сообществ зообентоса и их биотопов, выявлять множественные факторы формирования и пространственного распределения донной фауны, осуществлять зонирование водных объектов, в том числе по степени загрязненности вод [3, 4, 6, 7]. Целью данного исследования является оценка экологического состояния р. Енисей по комплексному биоиндикационному индексу, разработанному с привлечением методов многомерной статистики.

Для достижения поставленной цели привлечены результаты собственных многолетних гидробиологических исследований р. Енисей. На всем протяжении реки исследованы 5 участков, из них 3 расположены в Верхнем Енисее: 1 – Республика Тыва, между гг. Кызыл и Шагонар (2 станции), 2 – ниже Саяно-Шушенского водохранилища, между гг. Саяногорск

и Минусинск (3 станции), 3 – от плотины Красноярской ГЭС до устья р. Ангара (10 станций). Участок 4 находится в Среднем Енисее – от устья Ангары до д. Сургутиха (21 станция); участок 5 – в Нижнем Енисее от г. Дудинка до дельты, включая Бреховские острова (12 станций).

Для оценки качества воды по донным организмам использовали комплексный подход из 13 биоиндикационных индексов и метрик [5]: количество видов в пробе (n), индекс видового разнообразия Шеннона (H), хирономидный индекс Балуткиной (K), олигохетный индекс Гуднайта (G), EPT -индексы (число видов в пробе EPT_n , численность EPT_N (г/м²) и относительная численность $EPT_{N\%}$ поденок, веснянок и ручейников), FBI (Family Biotic Index), $IBGN$ (Indice Biologique Global Normalized), $BMWP$ (Biological Monitoring Working Party Index) и его производное $ASPT$ (Average Score Per Taxon Index), EPT_n/Ch_n и EPT_N/Ch_N – соотношение числа видов и численности группы EPT к хирономидам.

Статистический анализ данных проводился в среде MS Excel и с помощью программ на языке Python. При корреляционном анализе вычислялся коэффициент корреляции Пирсона, для построения корреляционного графа принимали во внимание достоверные связи и коэффициенты корреляции (r) не ниже 0,5. Для оценки влияния типа грунта на значения показателей качества воды проводился дисперсионный анализ с использованием параметрического критерия Фишера с поправками Уэлча для степеней свободы для учета разности дисперсии в группах, а также вычислялся коэффициент детерминации R^2 . Нулевые гипотезы отвергались при $p < 0,05$. Для построения и обоснования интегрального индекса качества воды использовался метод главных компонент. Предварительно все исходные данные были усреднены по станциям и стандартизированы. Для разбиения станций на классы по качеству воды применялась иерархическая кластеризация с использованием внутри и между кластерами расстояния Евклида и Варде соответственно. Кластеризация производилась в пространстве трех главных компонент.

Основными обитателями грунтов на большей части Енисея являлись амфиподы, хирономиды и олигохеты (Рис. 17). Поденки, ручейники и веснянки наибольшего развития достигли лишь в верховье Енисея (участок 1). Ниже Саяно-Шушенского водохранилища (участок 2) широко распространены фитофильные биоценозы, массово заселенные амфиподами. От плотины Красноярской ГЭС до устья р. Ангары (участок 3) доминировали хирономиды и амфиподы. Ниже устья Ангары (участок 4) увеличилась доля олигохет, и в низовье (участок 5) они составляли уже половину бентосных сообществ. Видовой состав беспозвоночных увеличивался вниз по течению: в верховье выявлено 48 видов, далее – 53, выше устья р. Ангары – 69. После впадения р. Ангары (участок 4) видовой состав расширился до

120 видов, но в низовье обеднел до 51. Видовое богатство комплекса ЕРТ сокращалось по мере продвижения к низовью реки, а хирономид и олигохет – напротив увеличивалось.

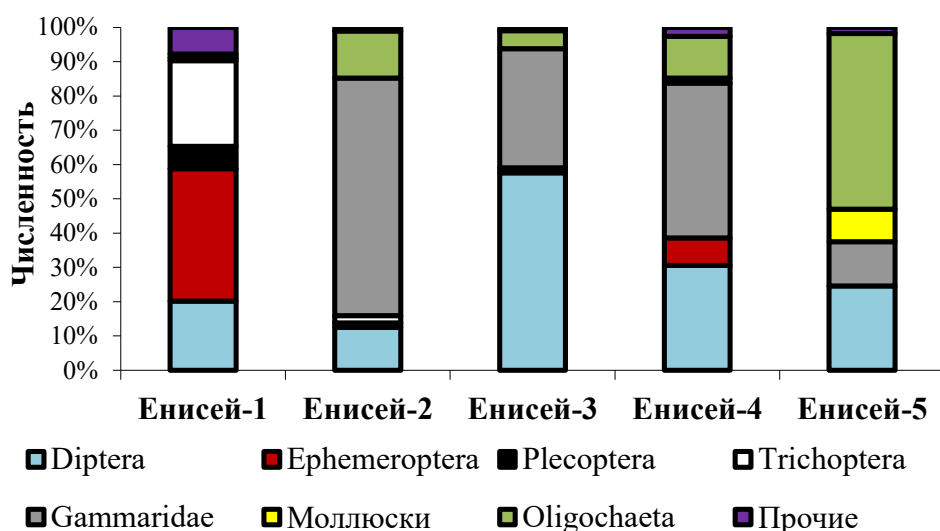


Рисунок 17 – Структурообразующие комплексы донных сообществ в р. Енисей.

Все индексы и метрики согласовано указывали на более высокое качество вод в верховье реки (участок 1) (табл. 2), что напрямую связано с доминированием в зообентосе комплекса *EPT* (68% общей численности) и высоким видовым разнообразием (2,8 бит/экз.). Дисперсионный анализ выявил, что типом грунта объяснялось от 6 до 40% вариации данных. Наиболее зависимы от типа грунта ($R^2 > 30\%$) следующие показатели: n , FBI , $BMWP$, $ASPT$, $IBGN$. Наилучшее качество вод по всем биоиндикационным показателям зафиксировано на твердых грунтах без заиливания, а наихудшее – в песчаных биотопах.

Таблица 2. Значения биоиндикационных показателей на различных участках р. Енисей.

Участок	n	H	EPT_n	EPT_N	$EPT_N\%$	EPT_n/Ch_n	EPT_N/Ch_N
1	11 ± 2	$2,8 \pm 0,2$	8 ± 2	$0,43 \pm 0,10$	68 ± 12	$5,4 \pm 1,5$	$11,2 \pm 4,1$
2	10 ± 1	$2,0 \pm 0,3$	3 ± 1	$0,19 \pm 0,05$	9 ± 2	$0,8 \pm 0,2$	$0,7 \pm 0,2$
3	11 ± 1	$2,1 \pm 0,2$	$1 \pm 0,2$	$0,04 \pm 0,02$	3 ± 1	$0,2 \pm 0,05$	$0,2 \pm 0,08$
4	11 ± 1	$2,2 \pm 0,1$	$2 \pm 0,2$	$0,14 \pm 0,03$	11 ± 2	$0,5 \pm 0,1$	$1,4 \pm 0,6$
5	8 ± 1	$2,2 \pm 0,1$	—	—	—	—	—

Таблица 2. Продолжение

Участок	FBI	$BMWP$	$ASPT$	$IBGN$	K	G
1	$3,5 \pm 0,4$	42 ± 7	$6,6 \pm 0,6$	28 ± 6	$1,5 \pm 0,8$	$1 \pm 0,8$
2	$5,1 \pm 0,3$	29 ± 4	$5,2 \pm 0,3$	15 ± 2	$1,4 \pm 0,6$	22 ± 6
3	$6,0 \pm 0,2$	16 ± 2	$3,9 \pm 0,2$	$7 \pm 0,7$	$1,6 \pm 0,4$	11 ± 3
4	$5,4 \pm 0,2$	22 ± 2	$4,3 \pm 0,2$	12 ± 1	$3,3 \pm 0,3$	16 ± 3
5	$6,7 \pm 0,1$	$10 \pm 0,6$	$2,9 \pm 0,1$	$5 \pm 0,3$	$3,5 \pm 0,4$	33 ± 4

Примечание: «—» означает, что представители группы *EPT* в зообентосе не обнаружены.

Согласованность индексов и метрик на изменение качества воды отражает корреляционный анализ. (Рис. 18). Максимальное число достоверных связей (7–8) отмечено для всех метрик *EPT* и индексов *BMWP*, *ASPT*, *IBGN*. Самая высокая степень корреляции ($r = 0.9$) зафиксирована для индекса *IBGN* в парах с *EPT_n* и *BMWP*. Наименьшую согласованность продемонстрировали показатели *n*, *H* и *G*, образующие 1–3 связи. Индекс Балашкиной *K* не образовал ни одной значимой корреляции с другими показателями качества воды, поэтому был исключен из расчета интегрального индекса.

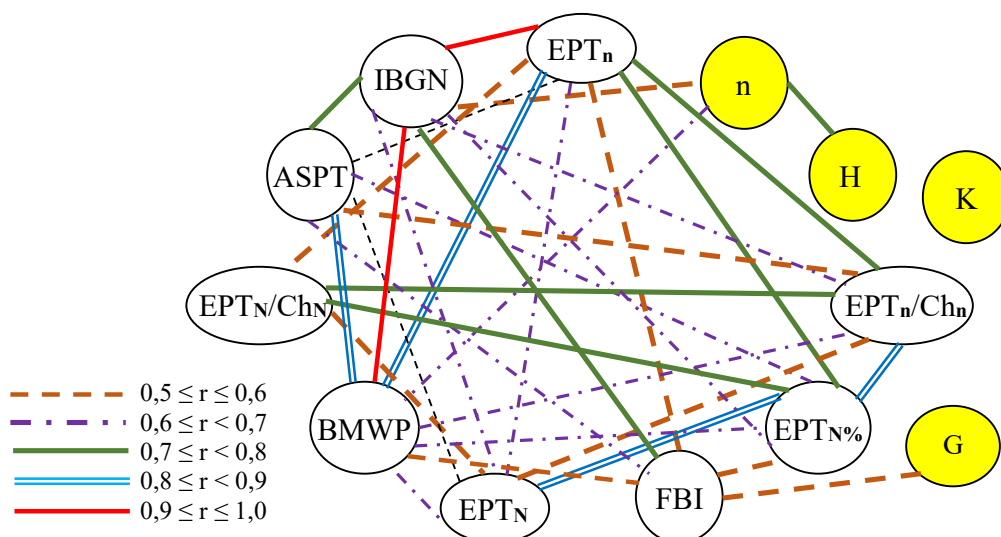


Рисунок 18 – Корреляционный граф взаимосвязей между индексами (желтым цветом выделены показатели с наименьшим количеством связей).

Метод главных компонент показал, что в первой компоненте сосредоточено 61,8% дисперсии исходных данных, на две другие компоненты приходится 12,9% и 10,2% дисперсии соответственно. Согласно правилу Кайзера, для дальнейшего анализа оставляют компоненты с собственным числом больше 1 – в нашем случае три главных компоненты PC1 – PC3. Для иерархической кластеризации, позволяющей в итоге разделить станции на категории качества воды, вместо 12 биоиндикаторов качества воды использовались полученные три главных компоненты PC1 – PC3. Это сделало кластеризацию более устойчивой и явно выраженной. Удовлетворительные результаты кластеризации получаются при выделении четырех кластеров (Рис. 19), соответствующих категориям качества вод: «условно чистая» (УЧ), «слабо загрязненная» (СЗ), «загрязненная» (З), «грязная» (Г). Основной вклад в определение класса качества дает первая главная компонента – чем правее по оси PC1 расположена станция, тем лучше на ней качество вод. Вторая и третья компоненты существенно влияют при разделении станций по категориям «слабо загрязненная» и «загрязненная».

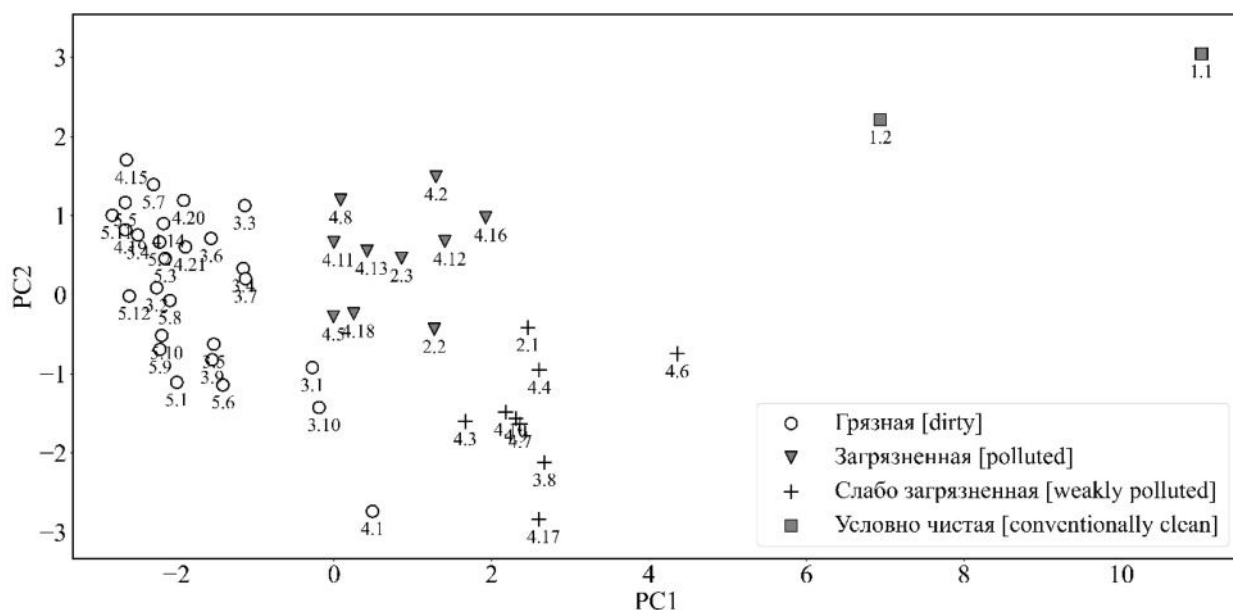


Рисунок 19 – Разделение исследованных станций в р. Енисей по качеству воды в осях первых двух главных компонент.

Качественная интерпретация главных компонент и кластеров выявила варианты несоответствия в оценке качества вод между показателями видового разнообразия (n и H) и другими индексами, а именно – EPT -метриками и показателями органического загрязнения (G , FBI). Интегральный показатель, построенный как проекция всех рассчитанных индексов и метрик на ось PC1, позволил провести оценку качества вод р. Енисей в единой системе (табл. 3).

Таблица 3. Интегральный показатель (PC1) и категории качества воды по станциям мониторинга в р. Енисей.

№ станции	1.1	1.2	2.1	2.2	2.3	3.1	3.2	3.3	3.4	3.5	3.6	3.7
PC1	11,02	6,93	2,46	1,27	0,86	-0,28	-2,25	-1,14	-1,15	-1,53	-1,56	-1,13
Качество вод	УЧ	УЧ	СЗ	З	З	Г	Г	Г	Г	Г	Г	Г
№ станции	3.8	3.9	3.10	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9
PC1	2,67	-1,55	-0,19	0,49	1,29	1,66	2,61	-0,01	4,36	2,36	0,08	2,31
Качество вод	СЗ	Г	Г	Г	З	СЗ	СЗ	З	СЗ	СЗ	З	СЗ
№ станции	4.10	4.11	4.12	4.13	4.14	4.15	4.16	4.17	4.18	4.19	4.20	4.21
PC1	2,18	0,00	1,41	0,42	-2,17	-2,64	1,93	2,60	0,25	-2,65	-1,91	-1,89
Качество вод	СЗ	З	З	З	Г	Г	З	СЗ	З	Г	Г	Г
№ станции	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	5.10	5.11	5.12
PC1	-2,00	-2,22	-2,15	-2,49	-2,65	-1,41	-2,29	-2,09	-2,21	-2,19	-2,82	-2,60
Качество вод	Г	Г	Г	Г	Г	Г	Г	Г	Г	Г	Г	Г

Примечание: УЧ – условно чистая; СЗ – слабо загрязненная; З – загрязненная; Г – грязная.

Наилучшее качество вод «условно чистые» выявлено в верховье Енисея (станции 1.1, 1.2). Далее вниз по течению происходило постепенное ухудшение качества вод, и после Красноярского водохранилища (участок 3) вода классифицирована преимущественно как «грязная». После впадения р. Ангары (участок 4) наблюдалось улучшение качества вод. И если в километре от устья Ангары (станция 4.1) еще сохранялась «грязная» категория, то далее происходило самоочищение воды до «загрязненной» и «слабо загрязненной». Начиная со станции 4.14 (пос. Бор) снова наблюдалось ухудшение качества воды, и все низовье Енисея попадало в категорию «грязная».

Список цитируемой в разделе литературы:

1. Головатюк Л.В., Зинченко Т.Д. Биотические идентификаторы в оценке качества воды эталонной реки: сравнительный анализ биоиндикационных индексов реки Байтуган (Высокое Заволжье) // Ученые записки Казанского университета. Серия естественные науки. 2020. Т. 162, № 1. С. 134–150. DOI: 10.26907/2542-064X.2020.1.134-150
2. Golubkov S.M., Balushkina E.V., Golubkov M.S. Restoration of zoobenthic communities and water quality in the river ecosystem after a decrease in the level of organic pollution // Contemporary Problems of Ecology. 2020. № 2. С. 146–155.
DOI: 10.1134/S1995425520020031
3. Baryshev I.A. Evaluating water quality in a large river system in the northern European Russia by macrozoobenthos // Water Resources. 2021. № 48(5). PP. 774–781.
DOI: 10.1134/S0097807821050055
4. Gültekin Z., Hellmann C., Aydin R., Winkelmann C. Possible indices for the assessment of ecological stream quality based on macroinvertebrates in Euphrates tributaries (Turkey) // Journal of Freshwater Ecology. 2019. № 34(1). PP. 783–806.
DOI: 10.1080/02705060.2019.1698469
5. Семенченко В.П., Разлуцкий В.И. Экологическое качество пресных вод. 2-е изд., исправл. Минск: Беларуская навука, 2011. 328 с.
6. Palatov D.M., Chertoprud M.V. Macrozoobenthic communities of compact grounds in streams of eastern Black Sea region // Russian Journal of Ecology. 2018. № 49 (1). PP. 80–86.
DOI: 10.1134/S1067413618010125
7. Tumusiime J., Uмба Tolo C., Dusabe M.-C., Albrecht Ch. Reliability of the Tanzania river scoring system (TARISS) macroinvertebrate index of water quality: a case study of the river Mpanga system, Uganda // Journal of Freshwater Ecology. 2019. № 34(1). PP. 541–557.
DOI: 10.1080/02705060.2019.1631895

Основные публикации по разделу:

1. Андрианова А.В. Оценка экологического состояния водных экосистем Восточной Сибири методами биоиндикации // Безопасность России. Правовые, социально-экономические и научно-технические аспекты. Тематический блок «Региональные проблемы безопасности». Раздел 1. Мониторинг, риски и безопасность Сибирского федерального округа. / Под. ред. Москвичева В.В. – Москва: МГОФ «Знание», 2024. – 644 стр. – ISBN 978-5-87633-206-6. – С. 510–519.
2. Шапарев Н.Я., Андрианова А.В. Река Енисей в показателях устойчивого водопользования. // Безопасность России. Правовые, социально-экономические и научно-технические аспекты. Тематический блок «Региональные проблемы безопасности». Раздел 1. Мониторинг, риски и безопасность Сибирского федерального округа. / Под. ред. Москвичева В.В. – Москва: МГОФ «Знание», 2024. – 644 стр. – ISBN 978-5-87633-206-6. – С. 485–494.
3. Андрианова А. В. Трансформация сообществ макрозообентоса в низовье р. Ангары после зарегулирования // Russian Journal of Ecosystem Ecology. 2024. Vol. 9 (1). С. 64-74. DOI: 10.21685/2500-0578-2024-1-4. [BAK]

1.6. Технология построения комплексного биоиндикационного показателя качества воды

Одной из задач отчетного периода было построение комплексного биоиндикационного показателя, с помощью которого можно оценивать качество воды.

Исходный набор данных для вычисления этого комплексного показателя формируется на основе информации, структурированной на следующие нижеперечисленные три категории (свойства объектов):

1. *Свойства локализации места взятия пробы.* Проба берется на станции, принадлежащей определенному участку реки. На одной станции берутся несколько проб в разных местах, поэтому все пробы в пределах станции перенумерованы.
2. *Характеристики места взятия пробы.* Как правило, место взятие пробы характеризуется некоторым набором свойств, которые для построения комплексного биоиндикационного показателя качества воды не используются. Однако эти характеристики могут быть полезны в предварительном или апостериорном анализе. Например, мы провели дисперсионный анализ биоиндикационных индексов и метрик, разделив пробы по типу грунта, на котором они были взяты. В таблице 4 приведены свойства, которые были в исследуемом наборе, но возможен и другой список этих свойств. Некоторые характеристики сохраняются в форме, указанной при взятии пробы, но дополняются числовым идентификатором, кодирующим ограниченный список значений, который отражает ту

или иную классификацию свойства. Например, в нашем случае грунт дополнительно описан свойством IDSoil, значение которого выбиралось из списка: «песчано-галечная смесь», «заиленный», «песок», «макрофиты».

3. *Биоиндикационные индексы и метрики качества воды в пробе.* Эти данные непосредственно участвуют в построении комплексного показателя (табл. 4). Набор метрик и индексов может быть как расширен, так и сужен. К сожалению, некоторые индексы в пробах невозможно рассчитать по причине отсутствия той или иной группы индикаторных организмов, поэтому эти свойства объекта могут быть неопределены (равны NA (Not Available), NaN (Not A Number), NULL), что не позволяет проводить анализ на основе датасета напрямую.

Для иллюстрации технологии анализ данных проведен на наборе данных, состоящем из 174 проб, взятых на 48 станциях, принадлежащих 5 участкам реки Енисей. В анализе участвовало 13 биоиндикационных индексов и метрик качества воды, рассчитанных для каждой пробы зообентоса.

Таблица 4. Перечень свойств объектов и их характеристики.

№ п/п	Название	Обозначение (сокращение)	Тип свойства	Комментарии
1. Локализация места				
1	Идентификатор участка	IDPart	Object (int)	Уникальный идентификатор участка реки. Одному участку соответствует несколько станций. Неопределенные значения недопустимы.
2	Участок реки	RiverPart	Object (str)	Название участка реки.
3	Идентификатор станции	IDStation	Object (int)	Уникальный составной идентификатор станции, на которой брали пробы. Формат идентификатора IDPart.IDStation, где IDStation – это порядковый номер, присвоенный станции, который является уникальным в пределах участка реки. Одной станции соответствуют несколько проб. Неопределенные значения недопустимы.
4	Название станции	place	Object (str)	Название станции.
5	Идентификатор пробы	IDSample	Object (str)	Уникальный составной идентификатор пробы в формате IDPart.IDStation.номер, где номер – это порядковый номер, присвоенный пробе, который является уникальным в пределах станции. Неопределенные значения недопустимы.

№ п/п	Название	Обозначение (сокращение)	Тип свойства	Комментарии
2. Характеристики места взятия пробы				
6	Берег	Bank	Object (str)	Характеристика берега (левый или правый, оконечность острова и т.д.)
7	Тип грунта	Soil	Object (str)	Текстовая характеристика грунта, описанная при взятии пробы
8	Макрофиты	Macrophytes	Object (str)	Индикатор наличия макрофитов
9	Идентификатор типа грунта	IDSoil	Object (int)	Числовой идентификатор, отражающий тип грунта из ограниченного списка значений. Список значений определяется исследователем. Неопределенные значения недопустимы.
3. Биоиндикационные индексы и метрики качества воды в пробе				
10	H	H	float64	Индекс видового разнообразия Шеннона.
11	n	n	float64	Количество видов в пробе.
12	EPTN	N-EPT	float64	Численность поденок, веснянок и ручейников
13	EPTN%	%-N-EPT	float64	Относительная численность поденок, веснянок и ручейников
14	EPTN/ChN	N-EPT/Ch	float64	Отношение численности группы ЕРТ и хирономид. Допускаются неопределенные значения.
15	EPTn	n-EPT	float64	Число видов поденок, веснянок и ручейников.
16	EPTn/Chn	n-EPT/Ch	float64	Отношение числа видов группы ЕРТ и хирономид. Допускаются неопределенные значения.
17	K	Balushkina	float64	Индекс Балушкиной
18	FBI	FBI	float64	Family Biotic Index
19	BMWP	BMWP	float64	Biological Monitoring Working Party Index
20	ASPT	ASPT	float64	Average Score Per Taxon Index
21	IBGN	IBGN	float64	Indice Biologique Global Normalize
22	G	G-Idx	float64	Олигохетный индекс Гуднайта

Предлагаемая технология построения комплексного биоиндикационного показателя состоит из следующих этапов:

1. Разведочный анализ данных исходного датасета.
2. Построение матрицы корреляций, возможное уменьшение размерности исходного датасета.
3. Осреднение показателей качества воды, переход от описания качества воды в пробе к описанию качества воды на станции.
4. Стандартизация метрик и индексов качества воды.

5. Метод главных компонент, уменьшение размерности датасета, полученного на этапе 3.
6. Кластеризация. Выделение классов качества воды. Определение класса качества воды на станции.

Опишем эти 6 этапов более подробно.

Этап 1. Разведочный анализ данных. Отбор множества метрик и индексов, на основе которых будет построен устойчивый комплексный показатель.

Поскольку все метрики и индексы описывают качество воды в пробе, то разумно ожидать, что их значения коррелированы. Однако, в силу особенностей предметной области и методики оценки качества воды по биоиндикационным показателям возможно, что не все индексы адекватно коррелированы с качеством воды в пробе, их значения могут отражать не качество воды, а некоторые географические и природные особенности участка реки. В этой связи на первом этапе для оценки общей ситуации с показателями необходим разведочный анализ датасета, который включает следующие подэтапы.

1.1. Для каждого показателя качества воды рассчитать описательную статистику, включающую 1) меры центральной тенденции: выборочное среднее, стандартную ошибку и доверительный интервал выборочного среднего, медиану; 2) меры изменчивости признака: размах, дисперсию.

1.2. Для каждого показателя качества воды построить гистограмму, по которой оценить модальность, симметрию выборки. Если необходимо, проверить гипотезу нормальности или логнормальности распределения генеральной совокупности. В настоящее время для проверки на нормальность рекомендуется использовать критерий Харке-Бера. Отметим, что в нашем случае, как правило, визуальной оценки гистограммы достаточно для отклонения гипотезы о нормальности распределения выборки. Многие показатели имеют выраженную моду в нуле и менее выраженную вторую моду, сильно распластаны. Распределение показателей, связанных с численностью, может быть близко к логнормальному, этот факт требует стандартной проверки.

1.3. Для каждого показателя качества воды определить наличие выбросов. Этот этап в основном важен для устранения возможных ошибок при расчетах показателей или считывания данных. При согласованной индикации выброса во всех показателях качества воды в пробе следует проверить исходные данные для расчета показателей и принять решение, исключить пробу из выборки, как бракованную или оставить, как соответствующую реальной ситуации. При появлении выброса в отдельном показателе следует понять причины этого с экспертом.

1.4. Построение диаграмм рассеивания для каждой пары показателей. На этом этапе можно визуально оценить попарную согласованность индикации качества воды в пробе.

Этап 2. Построение матрицы корреляций, выявление значимых корреляционных связей. Поскольку, как отмечалось выше, показатели качества воды часто не являются нормально распределенной случайной величиной, то необходимо оценить как параметрическую меру линейной связи – коэффициенты корреляции Пирсона, так и непараметрическую – коэффициенты корреляции Спирмана. Обычно в биологических оценках p -уровень значимости принимается равным 0.05.

На данном этапе следует отказаться от тех показателей, которые не коррелируют достоверно ни с одним из рассматриваемых показателей качества воды. Например, в нашем случае мы отказались от индекса Балускиной.

Этап 3. Осреднение показателей качества воды по группе проб, собранных на одной станции. Осреднение решает две задачи. Во-первых, осреднение сглаживает влияние выбора места взятия пробы на состав зообентоса в пробе. Мы будем оценивать качество воды именно станции, а не каждой отдельной пробы. Во-вторых, существуют пробы, для которых невозможно вычислить один-два индекса, но большинство метрик и индексов для этих проб может быть вычислено. Если осреднение не проводить, то такие пробы пришлось бы исключить из анализа, что с одной стороны уменьшило бы объем выборки, которая, как правило, не сильно большая, а с другой – исказило бы данные по станции.

В дальнейшем анализе объектом считается не проба, а станция. В результате далее мы анализируем датасет, каждый объект которого описывается следующими свойствами: 1) идентификатор станции (IDStation) исходного датасета, значения которого будут уникальными в рамках нового датасета; 2) биоиндикационные индексы и метрики качества воды исходного датасета, осредненные по всем пробам станции (в датасет не включаются свойства, удаленные на этапе 2).

Этап 4. Стандартизация метрик и индексов качества воды. В дальнейшем анализе вместо метрик и индексов участвуют их z -значения (у всех показателей среднее равно нулю, а значения отдельных наблюдений по этой шкале соответствуют расстоянию до среднего, выраженному в стандартных отклонениях).

Этап 5. Метод главных компонент. Цель применения метода заключается в построении очень небольшого набора показателей качества воды на основе имеющегося широкого набора метрик и индексов. Причем полученный набор показателей должен объяснять изменчивость исходного набора максимально полно.

Математически цель метода можно сформулировать следующим образом: определение в пространстве биоиндикационных индексов и метрик качества воды подпространства меньшей размерности, в ортогональной проекции на которое разброс данных (то есть среднеквадратичное отклонение от среднего значения) максимален.

Наш датасет Z состоит из M выборок $Z := \{\mathbf{Z}_i\}_{i=1}^M$. Каждая выборка $\mathbf{Z}_i \in R^N$ содержит z -значения одного осредненного по станции показателя качества воды в пробах, взятых на этой станции, т.е. выборку можно представить вектором-строкой $\mathbf{Z}_i := (z_{i1}, z_{i2}, \dots, z_{iN})$, где z_{ij} – z -значение i -го показателя на j -ой станции. Здесь M и N – количество показателей качества воды и станций в исходном датасете, соответственно. Таким образом, исходная система показателей качества воды однозначно задает начальную систему координат, в которой представлены значения показателей. Метод главных компонент находит новую ортогональную систему координат, ее оси называются главными компонентами (principal components, PC). Направление первой оси PC1 совпадает с направлением, в котором общая изменчивость (дисперсия) исходного набора данных максимальна. Направление каждой следующей оси выбирается в направлении максимальной дисперсии исходного набора при условии перпендикулярности этой оси всем предыдущим. В результате оси новой системы координат ориентированы и перенумерованы таким образом, чтобы монотонно уменьшать общую изменчивость исходного датасета. Матрица P перехода из старой системы координат в новую называется матрица нагрузок, с ее помощью можно пересчитывать старые координаты в новые:

$$PZ = Y \quad (1)$$

Матрица P ортогональная квадратная, ее строки являются нормированными собственными векторами матрицы корреляций ZZ^T z -значений осредненных показателей. Нумерация векторов совпадает с нумерацией соответствующих собственных значений, отсортированных по убыванию, т.е. первая строка матрицы является собственным вектором, соответствующим максимальному собственному значению. Собственные значения матрицы корреляций равны дисперсии датасета Z вдоль направления соответствующей главной компоненты. Согласно правилу Кайзера, для дальнейшего анализа оставляют компоненты с собственным числом больше 1. Таким образом, размерность M пространства исходных показателей качества воды сокращается до нескольких, обозначим размерность нового пространства K . Например, в анализируемом примере пространство было сокращено с $M=12$ исходно участвующих в анализе показателей до $K = 3$ главных компонент.

Матрицу P можно рассматривать как матрицу весов, с которыми исходные показатели качества воды входят в главные компоненты. На основе этих весов можно и желательно дать трактовку главным компонентам.

Отметим, что в силу того, что свойства объектов характеризуют качество воды на станции, первая главная компонента PC1 будет отражать согласованную характеристику качества воды. Веса каждого исходного показателя будут в ней примерно одинаковыми. С

малыми весами могут войти те показатели, которые слабо отражают качество воды в силу своего биологического смысла и географических и природных особенностей реки. Как правило, эти показатели выявятся еще на этапе изучения корреляции, они будут слабо коррелировать с другими показателями. На РС1 приходится максимальная дисперсия исходных данных. Смысл следующих главных компонент следует искать в отражении противоречий между некоторыми группами метрик и биоиндикаторов. Если какая-то группа индикаторов в исходных данных имеет биологически обоснованную тенденцию к завышению или занижению значений, необоснованных собственно качеством воды, то это и проявится в тех главных компонентах, которые будут соответствовать значимым по Кайзеру собственным значениям (дисперсии, объясняемой соответствующей компонентой).

Таким образом, на этом этапе будет построена небольшая система показателей качества воды. Каждый такой показатель может быть биологически обоснован и линейно зависит от исходного набора метрик и биоиндикаторов. Математически, это означает построение нового датасета $\tilde{Y}$, который строится из датасета Z следующим образом: $\tilde{Y} = \tilde{P}Z$, где матрица $\tilde{P}$, состоящая из первых K строк матрицы нагрузок P . В новом датасете каждая станция будет описываться только K свойствами, соответствующими первым K главным компонентам. Систему показателей $\tilde{Y}$ мы будем называть опорными показателями качества воды. Еще раз отметим, что первая главная компонента будет максимально возможно для исходного датасета соответствовать искомому *комплексному показателю качеству воды*.

Этап 6. Кластеризация. На этом этапе станции разделяются на кластеры. Каждому кластеру можно приписать класс качества воды. Для кластеризации используется новый датасет $\tilde{Y}$. Поскольку выделение главных компонент уменьшила зашумленность исходных данных, кластеризация датасета $\tilde{Y}$ будет более устойчивой. Предпочтение отдано иерархической кластеризации, в которой расстояние между элементами рассчитывается в евклидовой метрике, а межкластерное расстояние рассчитывается по методу Варде, что позволяет сделать кластеры наиболее плотными. Количество кластеров определяется методом локтя с использованием зависимости межкластерного расстояния, на котором происходит слияние кластеров от количества кластеров после слияния. Количество кластеров определяет возможную градацию классов качества воды. Согласно нормативным документам для оценки качества воды в водных объектах по биологическим показателям принята градация на 5 классов: «условно чистая», «слабо загрязненная», «загрязненная», «грязная» и «экстремально грязная». Однако для конкретного водного объекта эта классификация может варьировать.

Проиллюстрировать построенную классификацию можно, отобразив станции в осях главных компонент, особенно это эффективно при $K = 2$ или 3 . Таким образом, предлагаемая

технология позволяет построить комплексный показатель качества воды и на его основе распределить станции по классам качества воды.

Основные публикации по разделу:

1. Андрианова А.В., Карпова Е.Д. Оценка экологического состояния р. Енисей по структуре зообентоса при помощи комплексного биоиндикационного показателя // Вестник Томского государственного университета. Биология. (в печати) [WoS Q4]
2. ГИС «Комплексная оценка качества вод р. Енисей» / Андрианова А.В., Карпова Е.Д., Малимонов М.И., Пушкарев А.А., Якубайлик О.Э. // Федеральная служба по интеллектуальной собственности (Роспатент). Свидетельство о государственной регистрации программы для ЭВМ № 2024690444 от 16 декабря 2024 г.

2. СИСТЕМА МОНИТОРИНГА ВОЗДУХА Г. КРАСНОЯРСКА: ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНОЕ ОБЕСПЕЧЕНИЕ, ОБРАБОТКА И АНАЛИЗ ДАННЫХ

2.1. Система мониторинга воздуха Красноярского научного центра СО РАН

Исследования и разработки, посвященные задачам мониторинга и оценки состояния атмосферного воздуха г. Красноярска, ведутся коллективом авторов настоящего проекта уже более 5 лет. За эти годы создана «Система мониторинга воздуха Красноярского научного центра СО РАН» (КНЦ СО РАН), постоянно совершенствуется ее аппаратное, программно-технологическое, информационно-аналитическое обеспечение. В городе установлено и продолжает расширяться разнообразное оборудование для регистрации параметров окружающей среды: датчики концентрации взвешенных частиц в атмосфере (PM_{2.5}), метеостанции, фиксирующие атмосферное давление, температуру и относительную влажность воздуха, датчики температуры воды на различных участках реки Енисей, а также специализированные устройства, такие как метеорологические температурные профилемеры МТР-5 для измерения вертикальных профилей температуры атмосферы, термокосы для определения температуры воздуха над поверхностью Енисея, анемометры для регистрации скорости и направления ветра, видеокамеры для наблюдения за туманами над рекой и другие инструменты. Особое внимание уделяется экспериментальным исследованиям атмосферы над руслом Енисея с использованием квадрокоптеров, особенно в зимний период, над незамерзающей полыней.

«Система мониторинга воздуха КНЦ СО РАН» (далее здесь – Система) – это комплекс специального оборудования и программного обеспечения, предназначенный для сбора и анализа объективных данных о состоянии атмосферного воздуха в г. Красноярске, в составе которого, во первых, распределенные по городу сертифицированные и откалиброванные автоматические станции контроля загрязнения атмосферного воздуха (станции мониторинга), регистрирующие характеристики атмосферы и передающие эту информацию в централизованную базу данных через сотовую сеть, во вторых, специальные программные средства – встроенное программное обеспечение станций мониторинга и веб-интерфейс для их администрирования, подсистема хранения регистрируемых данных и программные интерфейсы для их загрузки через веб-сервисы, средства для оперативной обработки и визуализации данных.

Публичная часть Системы доступна по адресу: <https://air.krasn.ru/>. Уже более трех лет оперативные данные о уровне загрязнения атмосферного воздуха Красноярска, полученные с ее помощью, регулярно представляются в ежедневной рубрике «Экодозор» в вечерних выпусках новостей телеканала «Прима»; этот канал является лидером по размеру телеаудитории среди всех телеканалов Красноярска.

Программно-технологическая платформа разработки – Геопортал ИВМ СО РАН, его подсистема «Данные оперативного мониторинга». В отчетный период выполнялись работы по технической поддержке информационно-вычислительной инфраструктуры, оптимизации подсистем сбора оперативных данных и агрегации данных для быстрой веб-визуализации, разработке новых веб-интерфейсов для анализа данных мониторинга. После ввода в эксплуатацию дополнительного сервера обработки оперативных данных часть задач была перенесена на него. Это позволило снизить нагрузку на геопортал ИВМ СО РАН и более эффективно распределить нагрузку между поступающими потоками данных и алгоритмами их обработки.



Рисунок 20 – Интерфейсы пользователя «Системы мониторинга воздуха КНЦ СО РАН»: а) средние значения загрязнения воздуха, скорости и направления ветра за последние 7 дней, б) графики динамики доли ПДК (средней концентрации PM2.5 за последние 24 часа, нормированной на среднесуточную ПДК 35 мкг/м³), в) интерфейс системы на смартфоне.

Модернизация оборудования обеспечила возможность беспрепятственного подключения ряда новых датчиков, устройств сбора данных (температурные профилемеры МТР-5, датчики температуры воды, сеть видеокамер для регистрации туманов, и проч.).

Подсистема регистрации данных метеонаблюдений

Сотрудниками ИВМ СО РАН уже много лет ведется работа по поддержке и развитию системы сбора оперативной информации о состоянии атмосферного воздуха в Красноярском крае. Сбор оперативных данных для научных исследований и мониторинга состояния атмосферного воздуха осуществляется в единую систему «Данные оперативного мониторинга» геопортала ИВМ СО РАН (<http://sensor.krasn.ru/sc/>). Система постоянно пополняется новыми данными, регистрируются новые посты и системы мониторинга.

Для исследования состояния атмосферного воздуха в городе Красноярске и Красноярском крае ведется сбор данных с постов наблюдения краевой сети, с собственной сети станций ФИЦ КНЦ СО РАН и с постов сети общественных проектов. Помимо основных метеорологических параметров (температура, влажность и атмосферное давление) доступны данные по содержанию взвешенных частиц в атмосферном воздухе (PM2.5, PM10). Дополнительно осуществлялся сбор оперативных данных метеонаблюдений FM-12 IX SYNOP. Данные SYNOP включают данные о ветре, осадках, влажности, температуре воздуха и др. Данные ежедневно загружались с портала ЕСИМО (Межведомственная информационная система для доступа к ресурсам морских информационных систем и комплексного информационного обеспечения морской деятельности, <http://portal.esimo.ru/>). Данные собираются по метеостанциям на территории Красноярского края и соседних с ним регионов. К недостаткам полученных данных портала ЕСИМО можно отнести пропуски данных по ряду метеостанций в некоторые периоды. По наиболее интересным для исследований метеостанциям были загружены архивы данных с сайта <https://rp5.ru/>, которые позволили заполнить пробелы в данных. Для загрузки данных был создан специальный программный модуль. Дополнительно к данным о состоянии атмосферного воздуха собираются данные о гидрологических наблюдениях и данные об осадках.

При решении ряда задач, связанных с оценкой и прогнозом изменений в состоянии окружающей среды некоторой территории, либо задач, связанных с управлением процессами сельскохозяйственной деятельности в некоторых ситуациях данных стационарных постов наблюдения недостаточно. Это связано с тем, что удаление исследуемого объекта от ближайшей метеостанции может составлять от десятков до сотен километров. Для решения этой проблемы были использованы данные глобальных моделей, имеющих сплошное

покрытие территории. В качестве примера были рассмотрены территории ряда опытных производственных хозяйств (ОПХ) ФИЦ КНЦ СО РАН.

Сегодня существует множество глобальных и региональных моделей прогноза погоды и осадков. Глобальные модели охватывают всю планету, локальные охватывают определенные области, такие как континенты, страны, горные массивы и так далее. Как глобальные, так и локальные модели различаются между собой по пространственному разрешению, которое представляет собой размер ячеек сетки и по временному интервалу.

Среди глобальных моделей погоды можно отметить бесшовную европейскую модель глобального прогнозирования ECMWF и модель Global Precipitation Measurement (GFS). GFS – самая известная глобальная модель погоды, которая обновляется каждые шесть часов американской метеорологической службой [1]. Модель ICON (Integrated COSMO model), созданная Немецкой метеорологической службой (Deutscher Wetterdienst). Совместную модель COSMO, развиваемую и поддерживаемую консорциумом ряда Европейских стран, включая Россию. В России используется база данных прогнозов COSMO-Ru. Однако данные этой модели недоступны обычным пользователям. Еще одним популярным источником метеоинформации являются сайты агрегаторы. Например, проект OpenWeatherMap (<https://openweathermap.org>) и ему подобные, которые используют данные различных моделей, данные метеорологических служб и данные домашних метеостанций. Однако у домашних станций большой разброс показаний, что может негативно влиять на общий прогноз.

Ранее уже было разработано программное обеспечение и сервисы, позволяющие ежедневно загружать актуальные данные проектов GPM и GFS. Данные автоматически загружаются в пакетном режиме, после чего выполняется их предварительная обработка с созданием производных продуктов для научных исследований. Проект GPM (Global Precipitation Measurement) – предоставляет централизованный доступ к глобальным картам осадков [2, 3]. В Институте сегодня сформирован архив данных этих проектов за прошлые годы (GPM с 2001 года, GFS с 2007 года). Результаты обработки этих данных позволяют ежедневно анализировать состояние атмосферного воздуха и поверхности почвы на выбранном участке с помощью специализированных веб-приложений.

Были разработаны веб-сервисы, позволяющие получать данные по осадкам и температуре на заданной территории в выбранный промежуток времени. Однако к данным глобальных моделей в некоторых случаях возникали вопросы в их достоверности [4, 5, 6]. С целью проверки данных моделей было создано тестовое веб-приложение, позволяющее просматривать данные по осадкам на некоторых участках исследования совместно с данными метеостанций Росгидромета. Пример интерфейса веб-приложения показан на рисунке 21.

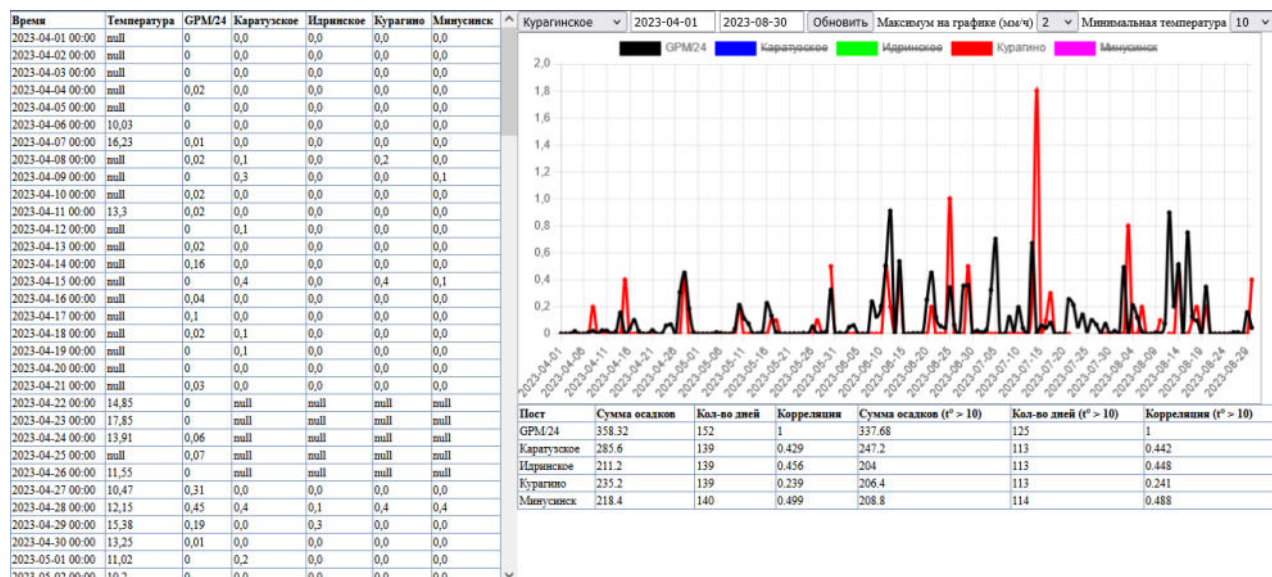


Рисунок 21 – Веб-приложение с данными моделей GPM, GFS и данными метеостанций Росгидромета.

При совместном анализе данных об осадках, полученных с метеостанций федеральной сети Росгидромета и данных модели GPM стоит учитывать ряд особенностей. Данные осадков с метеостанций Росгидромета передаются дважды в сутки (по шкале времени UTM+0) в 0 и 12 часов в виде суммы накопленных осадков за 12 часов измерений. Данные модели GPM содержат сумму накопленных осадков за полные сутки (по шкале времени UTM+0). Данные в нашей системе для метеостанций хранятся по Красноярскому времени (UTC+7).

Сравнительный анализ данных модели GPM и метеостанций Росгидромета был проведен для нескольких пилотных территорий, в качестве которых были выбраны три опытно-производственных хозяйства ФИЦ КНЦ СО РАН: Курагинское, Мининское и Михайловское. Были проанализированы данные с 2019 по 2023 годы. Результат – наблюдается слабая корреляция между модельными данными и данными наземных метеостанций. В частности, превышение суммы осадков у GPM составляет порядка 20-25 процентов над данными метеостанции.

При анализе состояния атмосферного воздуха для задач экологического мониторинга данные прогнозных моделей атмосферы позволяют выполнять оценку многолетних наблюдений за состоянием воздуха в городе или другой крупной территории. Для более детального анализа уже в рамках населенного пункта пространственного разрешения прогнозных моделей обычно недостаточно.

Все программы загрузки и подготовки файлов реализованы на языке Python 3 и языке сценариев Bash. Для хранения данных применяется СУБД PostgreSQL с модулем расши-

рения PostGIS. Веб-сервисы разработаны с использованием языка PHP 8. Клиентские веб-приложения разработаны с использованием языка JavaScript с использованием современных стандартов и технологий.

В 2024 году система сбора данных со стационарных постов наблюдений и систем, с метеостанции и данных глобальных моделей была усовершенствована. Добавлены дополнительные инструменты анализа получаемого потока данных для выявления ошибок и просмотра результатов с помощью кроссплатформенного мессенджера Telegram. Была разработана специальная программа – Бот. Это программа отправляет оперативные отчеты о возникающих ошибках оператору на почту и в мессенджер. Появилась возможность просмотра актуальных данных с ряда постов и станций через мессенджер. Просмотр актуальных данных позволяет оценить работу поста наблюдений без использования компьютера, что бывает необходимо при устранении проблем на месте установки поста. Пример сообщений программы Бота из программы Telegram показан на рисунке 22.

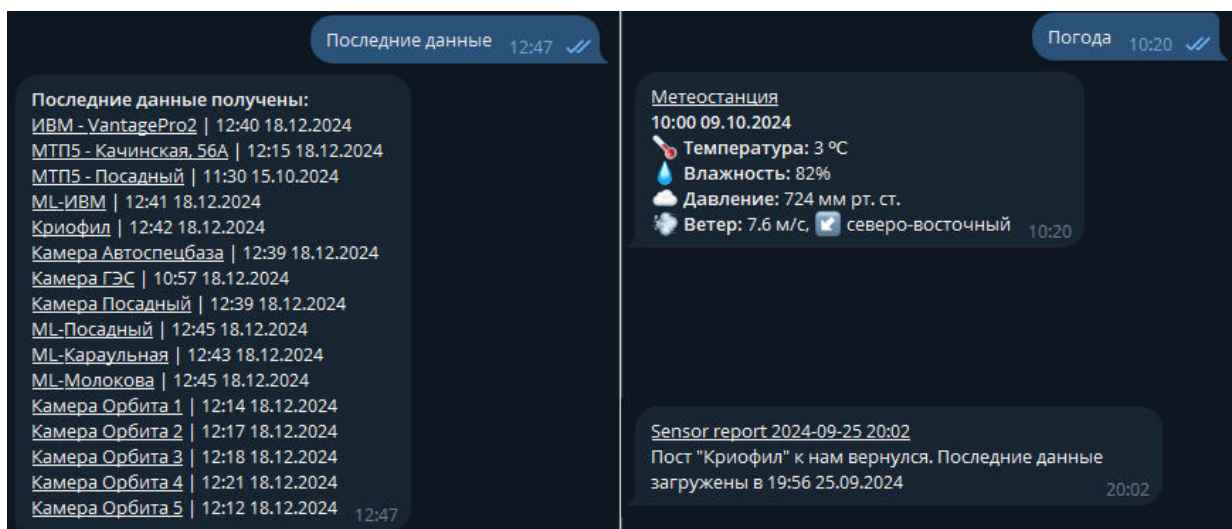


Рисунок 22 – Пример сообщений в приложении мессенджера Telegram.

Численное моделирование атмосферы Красноярск

Одной из задач отчетного периода стало исследование возможностей и специфики использования численной модели прогноза погоды WRF для моделирования динамики атмосферы г. Красноярск и окрестностей.

WRF (The Weather Research and Forecasting Model, Модель исследования и прогнозирования погоды) – это пакет прикладных программ для численного моделирования динамики атмосферы. Модель WRF применяется как для широкого круга научных исследований, так и для численного прогнозирования погоды. Это численная модель с открытым исходным кодом, разрабатываемая на основе широкой кооперации Национального центра

атмосферных исследований США (NCAR), Национального управления океанических и атмосферных исследований США (NOAA), BBC США, Военно-морской исследовательской лаборатории США, Университета Оклахомы и Федерального управления гражданской авиации США (сайт модели <https://www.mmm.ucar.edu/models/wrf>). Модель WRF является одной из самых используемых в мире региональных моделей численного прогноза погоды.

Несмотря на широкое распространение модели WRF, ее инсталляция является сложной и трудоемкой задачей. Требуется установка множества строго определенных взаимосвязанных библиотечных пакетов, компиляторов Fortran, библиотеки MPI, систем предварительной и пост обработки, файлов с географическими данными. Это приводит к нескольким известным труднопреодолимым точкам сбоя в процессе установки.

В отчетный период проведены работы по инсталляции вычислительной платформы WRF в конфигурации Advanced Research WRF Model (ARW). ARW кроме самого решателя WRF включает в себя модули географической привязки моделей прогноза (выбор области исследования, учет рельефа, распределения категорий землепользования подстилающей поверхности и т.д.); модули различных схем параметризации физических процессов (выбор моделей микрофизики влаги, описания потоков длинноволновой и коротковолновой радиации, расчета температуры и влажности подстилающей поверхности); модули задания параметров динамики атмосферы (подсеточные модели турбулентности, схемы параметризации конвективных процессов и планетарного пограничного слоя); процедуры инициализации данных, включая их интерполяцию на трехмерную сетку; пакет ассимиляции данных (WRFDA), способный обрабатывать различные типы наблюдений.

ARW была установлена и настроена на виртуальной машине под управлением операционной системы GNU/Linux с использованием системы Weather Research and Forecasting-Multi Operating System Installation Toolkit (WRF-MOSIT).

WRF-MOSIT – это набор пакетов сценариев для автоматизации, настройки и установки стандартной модели WRF и ее приложений. Набор инструментов WRF-MOSIT свободно доступен на GitHub (<https://github.com/HathewayWill/WRF-MOSIT>).

В результате было развернуто ПО в следующей конфигурации:

1. Библиотеки, инсталлированные с компилятором C GNU (zlib (1.3.1), MPICH (4.2.2), libpng (1.6.39), JasPer (1.900.1), HDF5 (1.14.4.3), PHDF5 (1.14.4.3), Parallel-NetCDF (1.13.0), NetCDF-C (4.9.2), NetCDF-Fortran (4.6.1), Miniconda);
2. Пакет WRF v4.6.0;
3. Пакеты пре- и постобработки (Development Testbed Center (DTC) Model Evaluation Tools (MET) v11.1.1, Development Testbed Center (DTC) Enhanced Model Evaluation Tools (MET-plus) v5.1.0, Development Testbed Center (DTC) Unified Post Processor (UPP) v4.1,

ARWPost v3, WRF-Python (Conda installed), OpenGrADS, GrADS, NCAR Command Language (Conda installed), Climate Data Operators (Conda installed).

Были проведены успешные запуски расчетов на тестовых данных, предоставляемых пакетом WRF.

Для успешной работы с пакетом необходимо провести настройку региональных параметров модели, учитывающих специфику исследуемой территории. Красноярск, как объект исследования, обладает рядом уникальных особенностей, которые требуют детального анализа и интеграции в модель. Во-первых, значительные перепады высот, обусловленные холмистым и ложбинным характером рельефа, оказывают существенное влияние на микроклиматические условия, распределение температур и движение воздушных масс. Во-вторых, наличие незамерзающего зимой Енисея создает дополнительные эффекты, такие как повышенная влажность и локальные температурные аномалии, которые необходимо учитывать для точности моделирования. Кроме того, особенности городской застройки, включая плотность расположения зданий, высотность, материалы строительства и распределение зеленых зон, также играют важную роль в формировании городского микроклимата. Все эти факторы требуют тщательного учета, так как их игнорирование может привести к значительным погрешностям в результатах моделирования.

Созданная информационно-вычислительная инфраструктура мониторинга атмосферы Красноярска предоставляет исходные данные, которые играют ключевую роль в процессе математического моделирования. Для обеспечения высокой точности и достоверности результатов моделирования необходимо тщательно учитывать пространственное распределение метеорологических параметров, включая их изменение по координатам (широте, долготе) и высоте над уровнем земли. Это особенно важно для учета вертикальной стратификации атмосферы, которая может существенно влиять на процессы тепло- и массообмена. Кроме того, важно корректно оценивать временную динамику этих параметров, так как их изменение во времени может оказывать значительное воздействие на итоговые результаты моделирования. Корректный учет пространственно-временной изменчивости метеорологических данных является важным этапом, обеспечивающим адекватность и надежность математических моделей. Один из примеров данных, которые должны быть усвоены моделью и в настоящее время являются предметом анализа, показан на рисунке 23.

Таким образом, настройка региональных параметров модели должна быть основана на комплексном анализе географических, климатических и антропогенных особенностей Красноярска.

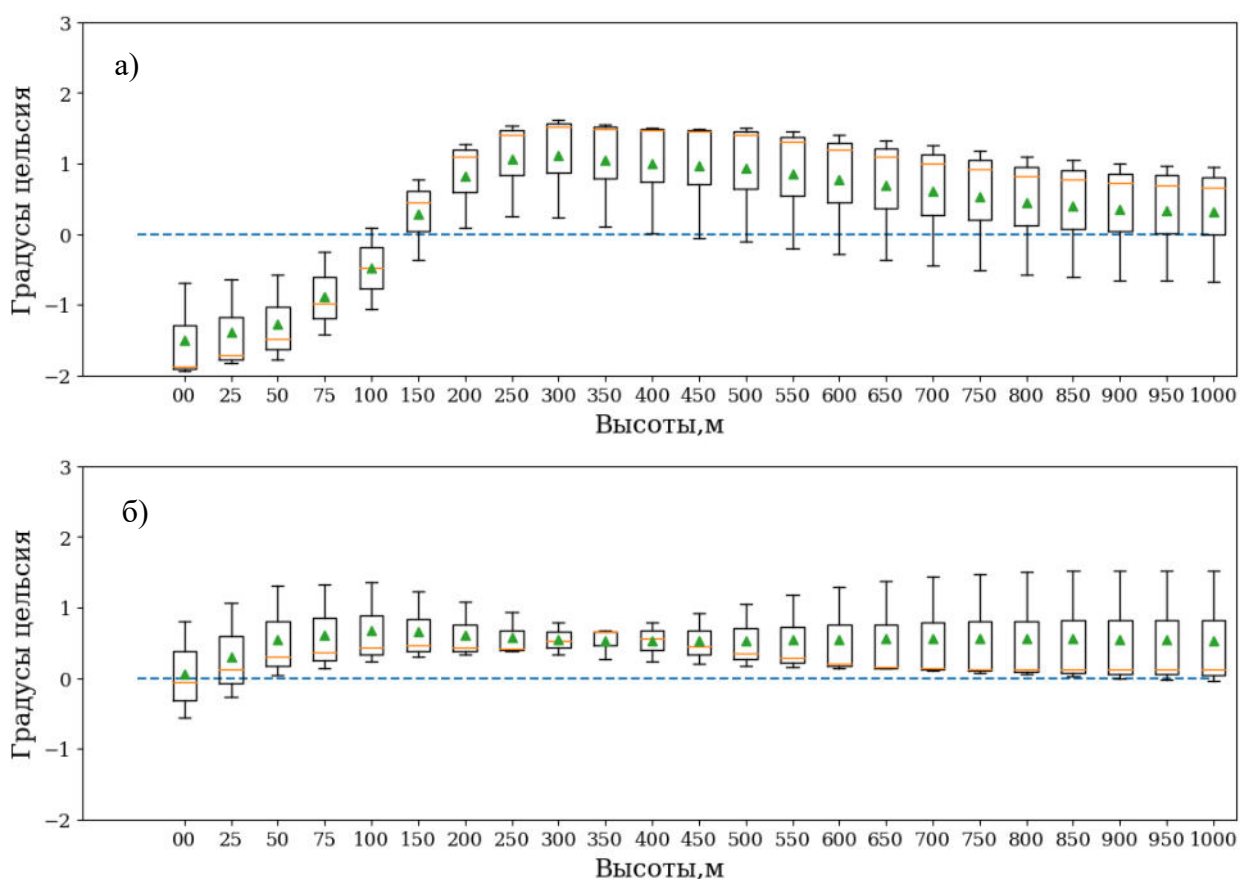


Рисунок 23 – Разница измерений двух метеорологических температурных профиломеров ($\Delta T = T_{\text{над_сушей}} - T_{\text{над_рекой}}$) на разных высотах зимой 2023-2024 гг.: а) при наличии туманов, б) в отсутствии туманов.

Программное обеспечение для реализации сводных расчётов загрязнения атмосферы

Одной из задач, которая рассматривалась в отчетный период, стало практическое знакомство и анализ возможностей и перспектив применения системы «Эколог-Город» [7]. Это программное обеспечение – «Унифицированная программа расчета загрязнения атмосферы» (УПРЗА «Эколог-Город») – разработано Санкт-Петербургской фирмой «Интеграл» для реализации сводных расчётов загрязнения атмосферы городов. В ее основе – актуальная и утвержденная приказом Минприроды России методика расчётов рассеивания выбросов вредных (загрязняющих) веществ в атмосферном воздухе. Пользователями УПРЗА «Эколог-Город» являются десятки городов России, она используется для расчетов в федеральном проекте «Чистый воздух».

В составе программного обеспечения «Эколог-Город» – базовая часть (ведение базы данных, расчет рассеивания, графический модуль ГИС) и многочисленные дополнительные модули и специализированные программы.

Система «Эколог-Город» был внедрена в 2016 году и в настоящее время с ним работают в Краевом государственном бюджетном учреждении "Центр реализации мероприятий

по природопользованию и охране окружающей среды Красноярского края". За эти годы сформирована база данных с городскими источниками выбросов, на основе которой выполняются расчеты.

Базовый модуль УПРЗА «Эколог-Город» позволяет рассчитать максимальные приземные концентрации загрязняющих веществ в атмосфере без учета влияния застройки в соответствии с приказом Минприроды РФ от 06.06.2017 № 273 «Об утверждении методов расчетов рассеивания выбросов вредных (загрязняющих) веществ в атмосферном воздухе» и «Методикой расчёта концентраций в атмосферном воздухе вредных веществ, содержащихся в выбросах промышленных предприятий (ОНД-86)». Система обеспечивает:

- создание и автоматизированное пополнение городского банка данных «Источники выбросов»;
- статистическую обработку информации;
- проведение сводных расчётов содержаний вредных веществ в атмосфере города (региона) на существующее положение и перспективу.

Расчеты выполняются на множестве точек местности, задаваемых пользователем. Это множество может включать как отдельно указанные точки, так и узлы прямоугольных сеток, что обеспечивает гибкость в выборе исследуемых участков. В составе программного комплекса УПРЗА «Эколог» используется графический модуль «ГИС Эколог», который предоставляет функционал для создания, редактирования и визуализации карт-схем предприятия и окружающей местности. На эти карты впоследствии наносятся результаты расчетов рассеивания загрязняющих веществ.

Программный комплекс поддерживает расчет приземных концентраций как для отдельных веществ, так и для групп веществ, обладающих суммирующимся вредным воздействием. Это позволяет учитывать комбинированное влияние различных загрязнителей на окружающую среду и здоровье человека, что особенно важно для комплексной оценки экологической обстановки.

Таким образом, УПРЗА «Эколог» предоставляет инструменты для детального анализа и визуализации данных, что способствует принятию обоснованных решений в области экологического мониторинга и управления.

Исследования возможностей этой программы, а также качества и полноты базы данных по источникам выбросов в Красноярске, на основе которой проводятся сводные расчеты загрязнения атмосферы Красноярска, будут продолжены.

В нижеследующих разделах этой главы представлены подробности по некоторым результатам, полученным в отчетный период.

Список цитируемой в разделе литературы:

1. Hewage, P., Trovati, M., Pereira, E. et al. Deep learning-based effective fine-grained weather forecasting model. *Pattern Anal Applic*, 2021. V. 24. P. 343–366.
2. Skofronick-Jackson G. et al. The Global Precipitation Measurement (GPM) mission for science and society. *Bulletin of the American Meteorological Society*, 2017. V. 98 (8). P. 1679-1695.
3. Hou A.Y. et al. The global precipitation measurement mission. *Bulletin of the American meteorological Society*, 2014. V. 95 (5). P. 701-722.
4. Tan, J., W. A. Petersen, and A. Tokay. A Novel Approach to Identify Sources of Errors in IMERG for GPM Ground Validation. *Journal of Hydrometeorology*, 2016, 17 (9), P. 2477-2491.
5. Tang L. et al. An improved procedure for the validation of satellite-based precipitation estimates. *Atmospheric Research*, 2015. V. 163. P. 61-73.
6. Ignatov A. V., Osipova O. P., Balybina A. S. Spatial structure of the relationships of annual precipitation amounts in Siberia and Kazakhstan. *Geography and Natural Resources*, 2018. V. 39. P. 148-152.
7. Система «Эколог-город» 4.70. Электронный ресурс. Дата обращения 27.11.2024.
URL: <https://integral.ru/shop/71/1041/>

Основные публикации по разделу:

1. Токарев А.В. Программа "Sensor Collector" для сбора, агрегации и представления данных оперативных наблюдений за состоянием окружающей среды // Свидетельство о регистрации программы для ЭВМ RU 2024667574, 26.07.2024. Заявка от 12.07.2024.
2. Кадочников А.А., Якубайлик О.Э., Токарев А.В., Заворуев В.В., Краснощеков К.В., Дергунов А.В. Информационно-аналитическая система «Характеристики атмосферы над акваторией р. Енисей» (версия 2) // Свидетельство о регистрации программы для ЭВМ RU 2024689250, 05.12.2024. Заявка от 14.10.2024.
3. Якубайлик Т.В., Токарев А.В., Якубайлик О.Э. Почасовые данные о наличии туманов над рекой Енисей в г. Красноярске по данным видеонаблюдений за 2022 – 2023 гг // Свидетельство о регистрации базы данных RU 2024623202, 19.07.2024. Заявка от 12.07.2024.
4. Пушкарев А.А., Кадочников А.А., Якубайлик О.Э. Программа для визуализации метеорологических данных реанализа на основе JPG-изображений // Свидетельство о регистрации программы для ЭВМ RU 2024666049, 09.07.2024. Заявка от 25.06.2024.

2.2. Измерение концентрации $PM_{2,5}$ в атмосфере Красноярска в 2023 году: сравнение данных четырех сетей мониторинга

Взвешенные твердые частицы с аэродинамическим размером меньше или равным $2,5 \text{ мкм}$ ($PM_{2,5}$) повсеместно обнаруживаются в приземной атмосфере. ВОЗ в сентябре 2021 года сообщила, что для здоровья людей безопасной является концентрация $PM_{2,5}$ равная или меньшая 5 мкг/м^3 . При этом, почти 91% населения мира проживает в районах, где концентрация $PM_{2,5}$ в атмосфере выше 10 мкг/м^3 [1]. При превышении этого значения наблюдается развитие различных заболеваний и возрастание смертности [2-4]. Возрастание концентрации $PM_{2,5}$ на 10 мкг/м^3 повышает риск инфаркта миокарда и увеличение числа госпитализаций [5, 6].

Существенная роль $PM_{2,5}$ в формировании благоприятной среды обитания человечества требует постоянного и объективного контроля концентрации этого загрязнителя в приземной атмосфере. Особенно актуален такой контроль в промышленных городах с населением более миллиона человек. В России самым восточным миллионным городом является Красноярск, который расположен в Восточной Сибири на берегах незамерзающей реки Енисей. Над речной акваторией концентрация $PM_{2,5}$ меньше, чем над территорией города [7]. Это вывод был сделан в результате анализа многолетних данных сети мониторинга воздуха Красноярского научного центра СО РАН. Отличительной особенностью системы является то, что только она включает пять (из двадцати двух) постов мониторинга, расположенных в русле реки Енисей. Другие три сети мониторинга (федеральная, краевая и общественная) имеют посты наблюдения за загрязнением, которые локализованы на территории города, вдали от акватории реки. Хотя все четыре сети измеряют концентрацию $PM_{2,5}$, тем не менее они различаются моделями применяемых анализаторов. По сути, это не должно влиять на корректность измерения. Величины концентраций $PM_{2,5}$, которые рассчитываются по данным датчиков каждой сети, статистически достоверно не должны отличаться. Однако, до сих пор не проведена оценка сходимости результатов измерений 4-х сетей мониторинга.

В рамках данного проекта было выполнено сравнение и анализ средних для города величин концентраций $PM_{2,5}$ (за периоды месяц и год), которые рассчитывали по данным каждой из четырех сетей мониторинга воздуха: федеральной, краевой, научной и общественной.

Характеристика сетей мониторинга

Среднесибирское УГМС (СУГМС) является структурным подразделением Федеральной службы по гидрометеорологии и мониторингу окружающей среды. На территории г. Красноярска расположено 10 стационарных постов наблюдения за загрязнением атмосферы

(ПНЗ). В 2023 году измерения концентрации $PM_{2,5}$ осуществлялись на ПНЗ №1, 3, 5, 9 с помощью анализаторов пыли EDM 180+ (Фирма «Grimm Aerosol Technik GmbH & Co.KG», Германия). Температура отбираемой пробы от -20 до +50°C.

Краевая сеть наблюдений за состоянием окружающей среды на территории Красноярского края является структурным подразделением Министерства экологии Красноярского края (далее – МЭ). В границах краевого центра расположено 7 автоматизированных постов наблюдений (АПН), которые оснащены анализаторами пыли модели E-BAM (Фирма «Met One Instruments Inc.», США). Эксплуатация E-BAM допускается при температуре окружающей среды от -30 до +50°C.

Сеть мониторинга воздуха Красноярского научного центра СО РАН (далее – КНЦ) была создана в 2018 году для исследований пространственно-временного распределения $PM_{2,5}$ в приземной атмосфере [8]. Измерение $PM_{2,5}$ осуществляется с помощью 22-х станций мониторинга воздуха CityAir (ООО «Унискан», Россия). Их эксплуатация возможна в диапазоне температур окружающей среды от -40 до + 50°C.

Общественная (краудфандинговая) сеть с названием «Небо» (далее – Небо) была создана в 2017 году и первоначально состояла из четырех станций AirVisual Node. Затем эта сеть стала расширяться за счет установки экспериментального оборудования [9]. В настоящее время общественности предлагается купить персональную станцию мониторинга воздуха (NeboAir). При её активации данные сразу же отображают сети «Небо». По всей вероятности общественная сеть состоит из двух типов станций. В 2023 на территории Красноярска в разные дни функционировало от 28 до 42 общественных датчиков $PM_{2,5}$. Следует заметить, что в инструкции, которую можно скачать на сайте <https://nebo.live/ru/krs>, нет никакой информации о температурных условиях эксплуатации станции NeboAir.

Таким образом, в трех сетях мониторинга (СУГМС, КНЦ и Небо) применяются оптические детекторы для измерения $PM_{2,5}$ и только в системе МЭ используются детекторы, работающие на хорошо известном принципе ослабления бета-излучения. Суть оптического метода состоит в том, что он позволяет определить размер и количество частиц, а для определения массы $PM_{2,5}$ требуется каким-то образом определить их плотность. В основе изотопного метода заложена закономерность того, что величина ослабления бета-излучения пропорциональна массе взвешенных частиц. По этой причине анализаторы ЕВМ 1020 и E-BAM являются первыми признанными приборами для измерения $PM_{2,5}$ и на протяжении нескольких десятков лет используются в США для мониторинга воздушной среды.

Анализ данных мониторинга за 2023 год

Среднегодовые, среднемесячные и среднесуточные концентрации $PM_{2,5}$ для приземной атмосферы Красноярска были рассчитаны для каждой сети мониторинга путем усреднения данных анализаторов, входящих в соответствующую сеть.

Анализаторы пыли EDM 180+ непрерывно измеряют концентрацию $PM_{2,5}$, однако СУГМС не представляет населению эту информацию в режиме онлайн. Более того, даже в режиме офлайн недоступна информация о среднесуточных и среднемесячных концентрациях $PM_{2,5}$ на ПНЗ. В данной работе анализируются среднемесячные концентрации на ПНЗ, которые были получены от СУГМС по специальному запросу. По этим данным с помощью метода описательной статистики были рассчитаны среднемесячные концентрации $PM_{2,5}$ и доверительный интервал для вероятности (P) равной 95%. Из-за малого количества данных (от 2 до 4) величины доверительных интервалов были сравнимы, а иногда превышали значения среднемесячных концентраций. Аналогичным методом для систем МЭ, КНЦ, Небо сначала рассчитывались среднесуточные концентрации, а потом среднемесячные концентрации $PM_{2,5}$ и доверительные интервалы. Результаты всех вычислений представлены в таблице 5.

Таблица 5. Среднемесячные концентрации $PM_{2,5}$ и доверительный интервал для $P=95\%$, N - количество измерений на постах мониторинга

Месяц	СУГМС		МЭ		КНЦ		Небо	
	$PM_{2,5}$	N	$PM_{2,5}$	N	$PM_{2,5}$	N	$PM_{2,5}$	N
I	$32,5 \pm 12,1$	4	$22,1 \pm 3,5$	198	$21,0 \pm 1,9$	682	$29,0 \pm 2,1$	1182
II	$89,3 \pm 37,6$	4	$58,9 \pm 7,3$	192	$64,0 \pm 4,1$	615	$106,0 \pm 5,4$	1077
III	$18,8 \pm 5,7$	4	$10,0 \pm 1,1$	216	$10,5 \pm 0,9$	665	$16,3 \pm 1,2$	1184
IV	$14,5 \pm 4,9$	4	$9,1 \pm 0,9$	186	$7,2 \pm 0,6$	630	$10,3 \pm 0,6$	1135
V	$17,5 \pm 4,9$	4	$11,4 \pm 0,9$	189	$9,1 \pm 0,5$	651	$14,2 \pm 0,7$	1152
VI	$17,0 \pm 8,7$	4	$13,2 \pm 1,1$	171	$8,1 \pm 0,4$	630	$13,7 \pm 0,9$	1123
VII	$9,8 \pm 10,4$	4	$12,2 \pm 0,8$	212	$7,5 \pm 0,2$	612	$11,5 \pm 0,5$	1082
VIII	$15,3 \pm 8,7$	4	$11,6 \pm 1,0$	205	$6,8 \pm 0,3$	651	$10,5 \pm 0,4$	1069
IX	$19,5 \pm 44,5$	2	$11,8 \pm 0,9$	210	$9,3 \pm 0,5$	630	$21,3 \pm 5,8$	951
X	$30,8 \pm 22,6$	4	$11,5 \pm 1,2$	215	$10,4 \pm 0,8$	651	$14,8 \pm 1,2$	939
XI	$34,7 \pm 11,2$	3	$23,0 \pm 3,9$	206	$22,6 \pm 2,3$	628	$34,6 \pm 3,1$	995
XII	$56,8 \pm 33,7$	4	$31,6 \pm 4,5$	198	$42,2 \pm 3,8$	651	$56,1 \pm 4,1$	1082

Данные, представленные в таблице 5, можно сгруппировать в два кластера по величине среднемесячных концентраций $PM_{2,5}$. В один из них вошли данные СУГМС и Небо, а другой - данные МЭ и КНЦ. Величины концентраций федеральной и общественной сетей в среднем в 1,6 раза больше, чем аналогичные концентрации краевой и научной сетей.

Обращает внимание различие среднемесячных концентраций $PM_{2,5}$, полученных СУГМС и МЭ (Рис. 24). Подчеркнем, что эти организации работают по единым законодательным и нормативным требованиям РФ. Среднемесячные концентрации $PM_{2,5}$ различались в марте, октябре, ноябре и декабре в 1,7-2,7 раза, а июне, июле, августе в 1,1-1,3 раза. Превышение среднесуточного гигиенического норматива (ПДК_{сс}) по данным МЭ было только в феврале, а по данным СУГМС ещё и в декабре.

На рисунке 24 величина среднегодового гигиенического норматива (ПДК_{сг}) показана для того, чтобы наглядно продемонстрировать в какие месяцы формируется загрязнение, влияющее на величину среднегодовой концентрации $PM_{2,5}$. По данным МЭ «повышенное» загрязнение наблюдается в течение двух месяцев (февраль и декабрь), а по данным СУГМС – на протяжении пяти месяцев (январь, февраль, октябрь-декабрь).

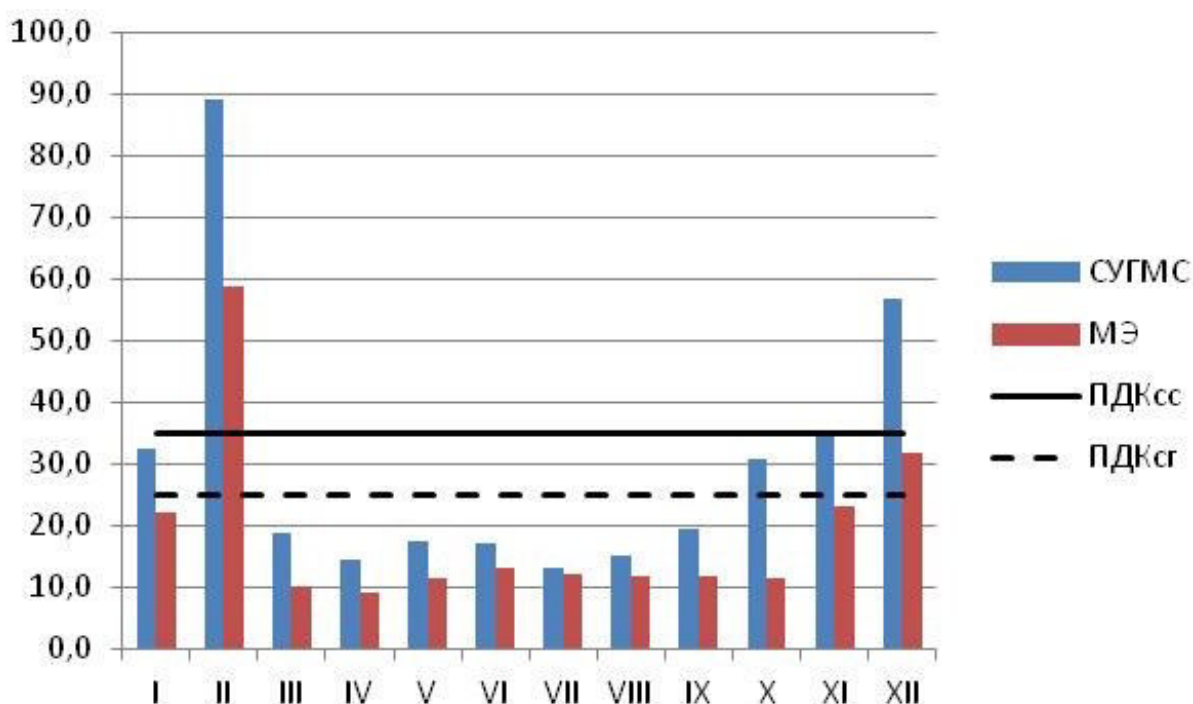


Рисунок 24 – Динамика среднемесячных концентраций $PM_{2,5}$, рассчитанных по данным СУГМС и МЭ.

Среднегодовые концентрации $PM_{2,5}$ представлены в таблице 6. Видно, что превышение безопасного уровня воздействия (5 мкг/м^3), рекомендованного ВОЗ, демонстрируют все четыре сети мониторинга. Важен другой факт. По данным СУГМС и Небо величина

среднегодового ПДК (ПДКсг) превышена и, в соответствии с законом об охране атмосферного воздуха воздушная среда Красноярска считается загрязненной. Следовательно, исполнительным органам краевой власти надо разрабатывать и осуществлять природоохранные мероприятия. Однако, среднегодовые концентрации $PM_{2,5}$, рассчитанные по данным МЭ и КНЦ, достоверно ниже величины ПДКсг, а это означает, что нет загрязнения атмосферы и не требуется расходовать бюджетные средства. Таким образом, остро встает вопрос о доверии к данным сетей мониторинга. В первую очередь объективность измерения концентраций $PM_{2,5}$ должны подтвердить СУГМС и МЭ.

Таблица 6. Среднегодовые концентрации $PM_{2,5}$

Сеть мониторинга	$PM_{2,5}$, (в $мкг/м^3$)	$PM_{2,5}$ (в долях ПДКсг = 25 $мкг/м^3$)	$PM_{2,5}$ (в долях ВОЗ = 5 $мкг/м^3$)
СУГМС	$30,7 \pm 7,5$	1,23	6,14
МЭ	$18,7 \pm 1,0$	0,75	3,74
КНЦ	$18,1 \pm 0,7$	0,72	3,62
Небо	$28,1 \pm 0,9$	1,12	5,62

Не менее важна объективность измерения концентраций $PM_{2,5}$ сетью Небо. Городские СМИ и иностранные научные издания [9] распиарили эту сеть. Доминирующая часть населения города Красноярска ориентируется на уровень загрязнения, который представляется на сайт <https://nebo.live/ru>.

Сеть КНЦ постоянно калибруется относительно анализаторов E-BAH [8], которые входят в состав приборов АПН. Таким образом, проверка объективности измерения сети МЭ автоматически отразится на сети КНЦ.

Рекомендация. Для разрешения сложившейся в г. Красноярске ситуации с измерением концентраций $PM_{2,5}$ (за периоды месяц и год), необходимо, чтобы датчики всех существующих сетей мониторинга были откалиброваны относительно эталонного гравиметрического метода (РД 52.04.893-2020).

Список цитируемой в разделе литературы:

1. Rajagopalan S., Landrigan P.J. Pollution and the heart // N. Engl. J. Med. 2021. 385. P. 1881-1892.
2. Barzeghar V., Sarbakhsh P., Hassanvand M.S., Faridi S., Gholampour A. Long-term trend of ambient air PM_{10} , $PM_{2.5}$, and O_3 and their health effects in tabriz city, Iran, during 2006–2017 // Sustain. Cities Soc. 2020. 54. 101988.

3. Brook R.D., Newby D.E., Rajagopalan S. The global threat of outdoor ambient air pollution to cardiovascular health: time for intervention // JAMA Cardiol. 2017. 2. P. 353-354.
4. Rajagopalan S., Al-Kindi S.G., et al. Air pollution and cardiovascular disease: jacc state-of-the-art review // J. Am. Coll. Cardiol. 2018. 72. P. 2054-2070.
5. Bhaskaran K, Hajat S, Armstrong B, Haines A, Herrett E, et al. The effects of hourly differences in air pollution on the risk of myocardial infarction: case crossover analysis of the MINAP database // Br Med J. 2011. 343. d5531.
6. Qiu X.Y., Wei Y.G., Wang Y., Di Q., et al. Inverse probability weighted distributed lag effects of short-term exposure to PM_{2.5} and ozone on CVD hospitalizations in New England Medicare participants-exploring the causal effects // Environ Res. 2020. 182. 7.
7. Заворуев В.В., Соколова О.В., Заворуева Е.Н., Якубайлик О.Э. Концентрация PM_{2.5} в приземной атмосфере над территорией г. Красноярска и акваторией незамерзающего участка реки Енисей // Оптика атмосферы и океана. 2023. 36. 06. С. 462–468.
8. Zavoruev V.V., Zavorueva E.N., Kadochnikov A.A., et al. Assessment of the possibility of using CityAir air monitoring station in environmental engineering // IOP Conf. Ser.: Mater. Sci. Eng. 2019. 537. 062053.
9. Labzovskii L.D., Hey J.V., Romanov A.A., Golovatina-Mora P., et al.. Who Should Measure Air Quality in Modern Cities? The Example of Decentralization of Urban Air Quality Monitoring in Krasnoyarsk (Siberia, Russia) // Environ. Sci. Policy. 2023. 140. P. 93–103.

Основные публикации по разделу:

1. Соколова О.В., Заворуев В.В., Заворуева Е.Н. Измерение концентрации PM_{2,5} в атмосфере Красноярска в 2023 году: сравнение данных четырех сетей мониторинга // Естественные и технические науки. 2024. № 6 (193). С. 192-196.
DOI: 10.25633/ETN.2024.06.11.

2.3. Концентрация бенз(а)пирена в атмосфере Красноярска в 2019-2023 годах: анализ данных краевой и федеральной наблюдательной сети

Бенз(а)пирен (BaP) относится к классу полициклических ароматических углеводородов (ПАУ). Он проявляет самую высокую канцерогенность среди широко представленных в городской атмосфере ПАУ. BaP образуется при неполном сгорании органических материалов [1]. Во многих исследованиях показаны негативные эффекты, вызываемые BaP. В зависимости от его концентрации зафиксированы повреждения ДНК человека, и эти эффекты передаются потомкам даже в третьем поколении [2]. Существует положительная

корреляция между воздействием ВаР и прогрессированием новообразований, метастазированием, частотой возникновения риска рака легких у человека [3, 4].

На территории г. Красноярска мониторинг содержания ВаР в атмосферном воздухе осуществляется государственной наблюдательной сетью (ГНС) – ФГБУ «Среднесибирское управление по гидрометеорологии и мониторингу окружающей среды» и краевой системой наблюдений (КНС) – КГБУ «Центр реализации мероприятий по использованию природных ресурсов и охране окружающей среды Красноярского края».

По данным ГНС на уровень загрязнения воздушной среды Красноярска существенное влияние оказывает концентрация ВаР [5]. Расчетным методом показано, что в административных районах города разная среднегодовая концентрация ВаР [6]. Для понимания закономерностей формирования пространственно-временного поля концентраций ВаР необходимо решить ряд задач.

В рамках данного проекта решались три задачи: изучение многолетней динамики концентраций ВаР в разных частях города по данным КНС; выявление зависимости концентрации ВаР от температуры наружного воздуха; статистическое сравнение среднегодовых концентраций ВаР, определенных ГНС и КНС в 2019-2023 годах.

В анализе и отборе проб ВаР используют различные методики измерений. В КНС суточные концентрации ВаР определяют в соответствии с методикой М 02-14-2007, а в ГНС – измеряют только среднемесячные концентрации ВаР по методике М 02-902-150-07.

Отбор проб в КНС осуществляется на четырех автоматизированных постах наблюдения за загрязнением атмосферного воздуха (АПН), наименования которых соответствуют микрорайонам расположения в Красноярске: АПН «Красноярск-Покровка» (далее – Покровка), АПН «Красноярск-Северный» (далее – Северный), «АПН «Красноярск-Солнечный» (далее – Солнечный), АПН «Красноярск-Черемушки» (далее – Черемушки). Отбор проводится 6 дней в неделю за исключением дней с осадками.

При статистической обработке данных использовали метод описательной статистики и двухвыборочный t-критерий. На рисунках и в таблицах средние значения представлены с доверительным интервалом для вероятности 95%.

По данным КНС за 2019-2023 годы рассчитаны усредненные за 5 лет среднемесячные концентрации ВаР в долях среднесуточной предельно допустимой концентрации (ПДКсс). Динамика ВаР на четырех АПН показана на рисунке 25.

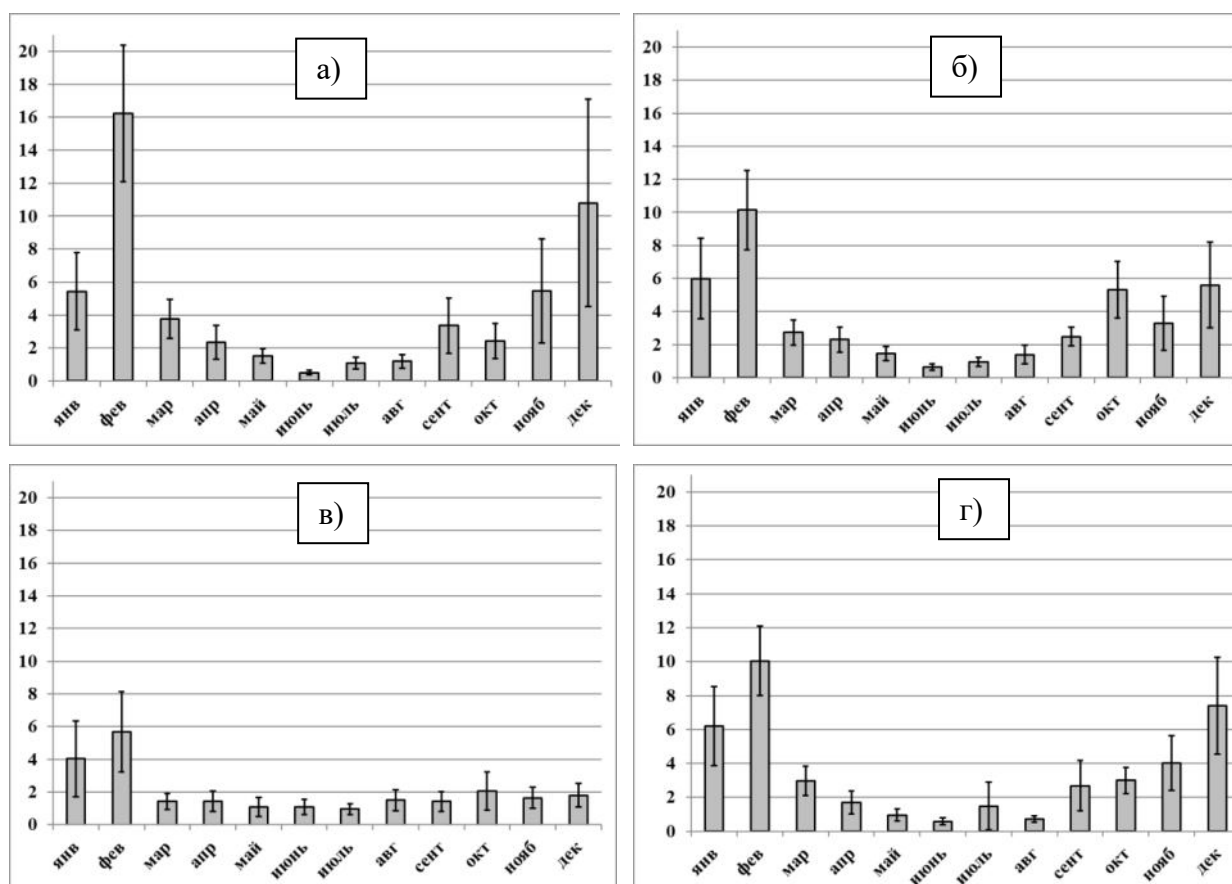


Рисунок 25 – Динамика среднесесячных концентраций ВаР (доли ПДКсс) на постах КНС: а) Покровка, б) Северный, в) Солнечный, г) Черемушки.

Представленные на рисунке 25 столбчатые диаграммы показывают, что изменения концентрации ВаР имеют схожий характер. Однако диапазон концентраций ВаР на посту Солнечный заметно меньше, чем на других трех постах. На посту Покровка самый большой диапазон среднесесячных концентраций - от 0,5 долей ПДКсс в июне до 16,0 долей ПДКсс в феврале. Обращает на себя внимание то, что на всех диаграммах самые высокие значения ВаР фиксируются в феврале. Это связано с тем, что в этот месяц постоянно возникают неблагоприятные метеорологические условия (НМУ), которые представляют собой краткосрочное сочетание метеорологических факторов (слабый ветер или штиль, вертикальная инверсия температуры в атмосфере и т.п.), способствующих накоплению загрязняющих веществ в приземном слое воздушной среды. Например, в 2022 и 2023 годах продолжительность НМУ была около 7 и 19 суток, соответственно.

В целом, в зимний период года (январь, февраль, декабрь) наблюдаются самые высокие, а летом (июль-август) – самые низкие концентрации ВаР. Из этого можно сделать заключение, что в отопительный сезон в атмосферу выбрасывается больше ВаР, чем в период прекращения отопления различных домов и зданий.

На посту Солнечный средняя за пять лет концентрация ВаР была ниже (1,9 ПДКсс), чем на постах Покровка (4,2 ПДКсс), Северный (3,3 ПДКсс), Черемушки (3,3 ПДКсс).

Для периода 2019-2023 годов были определены максимальные среднесуточные концентрации ВаР на каждом посту и для каждого месяца (Таблица 7).

Таблица 7. Максимальные среднесуточные концентрации ВаР за период 2019-2023 годов на постах КНС (доли ПДКсс)

Месяц	Покровка	Северный	Солнечный	Черемушки
Январь	46	32	40	32
Февраль	69	49,6	66	32,5
Март	28,6	15,9	10,6	22
Апрель	34	20,9	18,6	23
Май	7,4	8,1	17,9	13,2
Июнь	2,5	4,4	9,8	6,5
Июль	8,4	9,1	7,7	57
Август	7,5	18	16	5,2
Сентябрь	46	10,8	18	43
Октябрь	16	36,6	30	21,9
Ноябрь	58	29	11	27
Декабрь	93	35	21	36

Максимальные превышения ПДКсс ВаР зафиксированы на постах КНС: Покровка – 93 ПДКсс (29 декабря 2020); Северный – 49,6 ПДКсс (7 февраля 2023); Солнечный – 66 ПДКсс (17 февраля 2020); Черемушки – 57 ПДКсс (13 июля 2021). За пятилетний период количество превышений ПДКсс составило: в Северном – 473 случая, в Черемушках – 472 случая, в Покровке – 438 случаев, в Солнечном – 319 случаев.

Для выявления связи между концентрациями ВаР на постах КНС были рассчитаны коэффициенты корреляции (Таблица 8).

Таблица 8. Коэффициент корреляции (R) между концентрациями ВаР, измеренными на постах КНС.

	Покровка	Северный	Солнечный	Черемушки
Покровка	1,00			
Северный	0,36	1,00		
Солнечный	0,31	0,52	1,00	
Черемушки	0,60	0,71	0,49	1,00

Наибольшая зависимость между данными наблюдается между концентрациями ВаР, измеренными в Черемушках и Северном ($R=0,71$). Самый низкий коэффициент корреляции ($R=0,31$) наблюдается между концентрациями ВаР, измеренными в Солнечном и Покровке.

Представленный на рисунке 25 характер изменения концентрации ВаР допускает наличие зависимости содержания ВаР от величины температуры атмосферного воздуха. Для проверки этого предположения была выполнена разбивка данных по четырем температурным группам: от -20°C до -10°C , от -10°C до 0°C , от 0°C до $+10^{\circ}\text{C}$, от $+10^{\circ}\text{C}$ до $+20^{\circ}\text{C}$. Для каждой группы рассчитывали среднюю температуру и концентрацию ВаР с соответствующими доверительными интервалами. Затем были построены графики (Рис. 26) и подбирали аппроксимацию с наивысшим коэффициентом детерминации (R^2). Установлено, что экспоненциальная функция лучше всего описывает расчетные точки для данных 4-х постов.

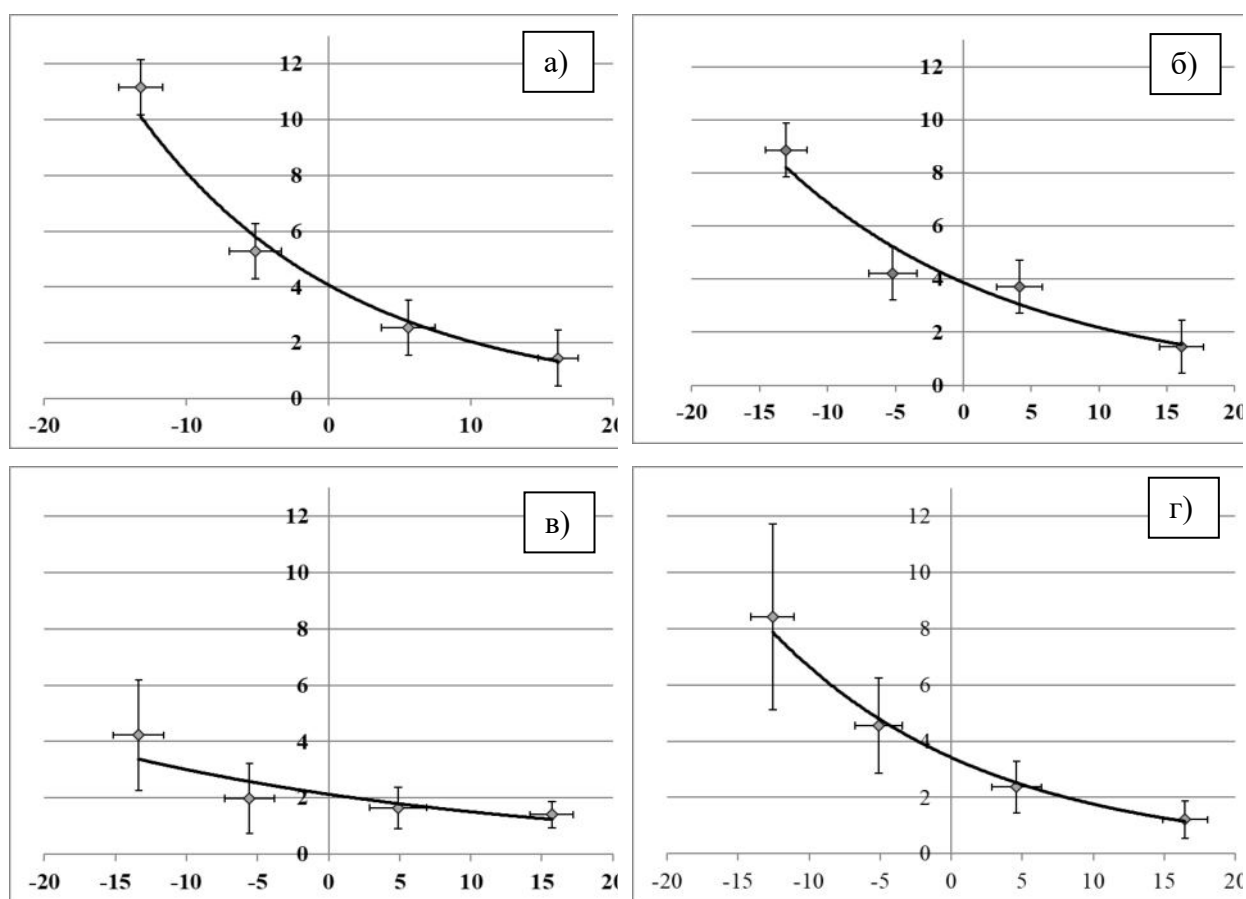


Рисунок 26 – Зависимость концентрации ВаР от температуры на постах КНС: а) Покровка, б) Северный, в) Солнечный, г) Черемушки.

Уравнения 1-4, хорошо описывают зависимость концентрации ВаР от температуры воздуха на постах Покровка, Северный, Солнечный, Черемушки, соответственно:

$$y = 3,8413e^{-0,068x}, \quad R^2 = 0,9959 \quad (1)$$

$$y = 3,8732e^{-0,058x}, \quad R^2 = 0,9375 \quad (2)$$

$$y = 2,1275e^{-0,035x}, \quad R^2 = 0,8114 \quad (3)$$

$$y = 3,4227e^{-0,066x}, \quad R^2 = 0,9924 \quad (4).$$

Для уравнений (1), (2), (4) коэффициенты детерминации от 0,9375 до 0,9959, а для уравнения (3) R^2 – чуть меньше 0,8114. Тем не менее, все коэффициенты свидетельствуют о значительной зависимости концентрации ВаР от температуры атмосферного воздуха.

Выше говорилось, что две наблюдательные сети измеряют ВаР в атмосфере Красноярска. Отбор проб осуществляется на 8 постах ГНС и на 4 постах КНС. Полученные результаты измерения, как правило, в государственных докладах и ежегодниках ГГО приводятся без доверительного интервала [7]. В связи с этим, трудно сравнивать несколько данных между собой. С учетом этого обстоятельства, была выполнена статистическая обработка данных ГНС и КНС. Для этого использовали среднемесячные концентрации на всех постах в течение года. Таким образом, в расчете использовался массив из 96 данных для ГНС и 48 данных для КНС. Результаты расчетов представлены в таблице 9.

Таблица 9. Среднегодовые концентрации ВаР по данным ГНС и КНС

Год	2019	2020	2021	2022	2023
ГНС	4,1±1,0	3,8±1,0	4,6±1,0	5,5±1,4	5,4±1,4
КНС	1,8±0,8	3,4±1,0	3,6±1,1	4,3±1,2	3,8±1,4

Таким образом, показано, что среднегодовые концентрации, за исключением 2019 года, статистически не различаются, что говорит о достоверности данных.

Решение поставленных задач позволило сделать следующие выводы.

1. Характер изменения многолетней динамики концентраций ВаР в разных частях города по данным КНС практически одинаков, но в количественном выражении (концентрации ВаР) наблюдаются различия. Наблюдается слабая корреляция между данными, полученными на четырех постах КНС.

2. Выявлена экспоненциальная зависимость концентрации ВаР от температуры атмосферного воздуха.

3. Впервые проведено сравнение среднегодовых концентраций ВаР, которые были определены ГНС и КНС. Установлено, что для периода 2020-2023 годов между этими значениями нет статистически достоверных различий.

Список цитируемой в разделе литературы:

1. Boström, C.E.; Gerde, P.; Hanberg, A.; Jernström, B.; Johansson, C.; Kyrklund, T.; Rannug, A.; Törnqvist, M.; Victorin, K.; Westerholm, R. Cancer Risk Assessment, Indicators, and Guidelines for Polycyclic Aromatic Hydrocarbons in the Ambient Air. *Environ. Health Perspect.* 2002, 110, 451–488.
2. Bukowska, B.; Sicińska, P. Influence of Benzo(a)Pyrene on Different Epigenetic Processes. *Int. J. Mol. Sci.* 2021, 22, 13453.
3. Petit, P.; Maître, A.; Persoons, R.; Bicout, D.J. Lung Cancer Risk Assessment for Workers Exposed to Polycyclic Aromatic Hydrocarbons in Various Industries. *Environ. Int.* 2019, 124, 109–120.
4. Saleh, S.A.K.; Adly, H.M.; Aljahdali, I.A.; Khafagy, A.A. Correlation of Occupational Exposure to Carcinogenic Polycyclic Aromatic Hydrocarbons (CPAHs) and Blood Levels of P53 and P21 Proteins. *Biomolecules* 2022, 12, 260.
5. Ежегодник состояние загрязнения атмосферы в городах на территории России за 2022 г.
6. Заворуев В.В., Манкевич И.В., Заворуева Е.Н. Влияние выбросов автономных источников теплоснабжения на загрязнение бенз(а)пиреном атмосферы города Красноярска // *Естественные и технические науки.* 2023. 178 (3). С 70-74.
7. Государственный доклад «О состоянии и охране окружающей среды в Красноярском крае в 2022 году».

Основные публикации по разделу:

1. Манкевич И.В., Заворуев В.В., Заворуева Е.Н. Концентрация бенз(а)пирена в атмосфере Красноярска в 2019-2023 годах: анализ данных краевой наблюдательной сети // *Естественные и технические науки.* 2024. № 6 (193). С. 178-183.
DOI: 10.25633/ETN.2024.06.11.

2.4. Сравнительный анализ данных мониторинга концентрации взвешенных частиц PM_{2.5}, регистрируемых разными датчиками

В последнее время количество и качество собираемых данных, а также их детализация имеют тенденцию к росту. Поэтому помимо непосредственной работы с данными о загрязнениях и моделями их распространения уделяется большое внимание проблемам сбора и накопления информации о загрязнениях. В связи с этим представляет особый интерес оценка эффективности использования недорогих сенсоров.

Мы анализировали данные с двух систем мониторинга: 1) с девяти автоматизированных постов наблюдений (АПН) Краевой ведомственной информационно-аналитической системы данных о состоянии окружающей среды Красноярского края (КВИАС), поддерживаемой Министерством экологии Красноярского края; 2) с 21-го поста «Системы мониторинга воздуха КНЦ СО РАН» (КНЦ СО РАН). На постах, в числе прочего, раз в 20 минут выполняется автоматическое измерение метеорологических параметров и уровня концентрации PM_{2.5} в приземном слое атмосферы. В КВИАС для мониторинга концентрации PM_{2.5} используются анализаторы пыли модели E-BAM (Met One Instruments Inc., США), принцип действия которых основан на измерении поглощения β -излучения частицами пыли, осажденными на фильтрующую ленту. Анализаторы этого класса рекомендованы для измерения содержания фракций PM₁₀ и PM_{2.5} в атмосфере, сертифицированы и аккредитованы во многих странах мира, в том числе и в России (№ 57884-14 в Госреестре средств измерений). Каждый пост системы мониторинга КНЦ СО РАН оснащен станцией мониторинга воздуха CityAir, разработанной группой компаний из новосибирского технопарка и инновационного центра Сколково. Для мониторинга концентрации PM_{2.5} используются оптические датчики (№ 75984-19 в Госреестре средств измерений), в которых проходящий через поток загрязненного воздуха фокусированный лазерный луч рассеивается на твердых частицах, что регистрируется фотодиодом и позволяет количественно оценить загрязнение. Практика применения таких датчиков показала, что они уступают по точности анализаторам E-BAM, однако пригодны для оценки уровня концентрации PM_{2.5}. Система мониторинга КНЦ СО РАН имеет около 30 постов, расположенных в разных районах г. Красноярск, что обеспечивает хорошую детализацию информации о загрязнениях.

Далее, в зависимости от принадлежности поста АПН КВИАС или системе мониторинга КНЦ СО РАН показатели будут иметь префикс «m_» или «s_», соответственно. Кроме того, для краткости датчики концентрации PM_{2.5} АПН КВИАС и станции CityAir будем упоминать как «анализаторы E-BAM» и «оптические датчики», соответственно.

По каждому посту данные представлены временными рядами измерений в приземном слое атмосферы температуры (t) в °C, давления (p) в мм рт. ст., относительной влажности воздуха (h) в % и концентрации взвешенных частиц PM_{2.5} (PM) в мкг/м³. В скобках указаны используемые далее обозначения факторов. Каждый ряд содержит до 131590 измерений (с 01.01.2019 00:00 по 31.12.2023 23:40, 3 измерения в час). Для сравнения показаний датчиков, принадлежащих разным системам мониторинга, существует 4 поста, на которых установлена измерительная аппаратура обоих типов (посты «Ветлужанка», «Покровка», «Свердловский» и «Кировский»). Далее в тексте будут приведены результаты анализа для

пары дублирующих датчиков с поста «Ветлужанка». Для краткости будем помечать s_{vet} и m_{vet} измерения станции CityAir и датчиков АПН КВИАС, соответственно.

Поскольку целью работы является построение регрессионной модели для корректировки данных о концентрациях PM_{2.5} станций CityAir, мы будем рассматривать данные не как временные ряды, а как связанные выборки случайных величин. Описательная статистика показывает, что распределение температуры имеет ярко выраженную трехмодальность (зимний, летний и демисезонный периоды), причем зимний сезон имеет большую дисперсию, чем демисезонный и летний. Распределение давления близко к нормальному, с асимметрией вправо, распределение влажности имеет большую асимметрию с «тяжелым хвостом» влево. Распределение концентрации PM_{2.5} близко к логнормальному, но все-таки таковым не является, кроме того, выборки имеют по две моды, первая из которых соответствует фоновым значениям концентрации, а вторая, менее выраженная, с большой дисперсией – периодам высоких концентраций.

Уже на этапе описательной статистики становится ясно, что имеются существенные различия в данных, полученных с помощью измерительной аппаратуры разного типа. Например, среднее значение концентрации PM_{2.5} по выборке измерений оптического датчика станции CityAir, почти в 2 раза превышает среднее значение в выборке анализатора E-BAM. Поскольку методика измерения концентрации PM_{2.5}, используемая анализаторами E-BAM, тщательно верифицирована, а оптические датчики принято использовать, как сравнительно дешевую альтернативу, то актуально построение статистически обоснованного правила корректировки первичных данных о концентрации PM_{2.5} оптических датчиков по данным, полученным от E-BAM. Вместе с тем, дополнительный анализ длительных рядов наблюдений от анализатора E-BAM, показал, что со временем меняется характер поведения измерений, что, возможно, связано с длительной эксплуатацией прибора в жестких погодных условиях и загрязненной атмосфере г. Красноярска. Мы постарались очистить ряды измерений от явных аномалий, и приняли за основу данные от E-BAM.

Диаграмма рассеяния на рисунке 27а показывает систематическое завышение концентрации оптических датчиков. Мы приняли решение для каждой пары датчиков исключить из анализа ~5% пар значений, которые соответствуют максимальным расхождениям в показаниях уровней концентраций PM_{2.5} в паре. Диаграмма рассеяния после корректировки отображена на рисунке 27б.

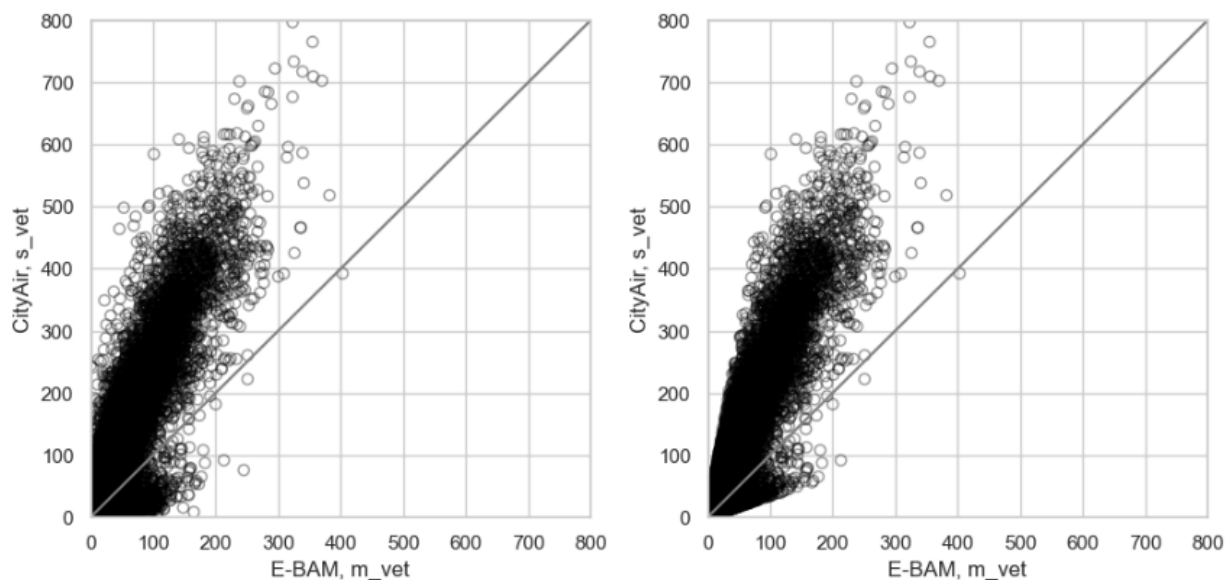


Рисунок 27 – Диаграмма рассеяния уровней концентрации взвешенных частиц PM2.5 (мкг/м³) на анализаторе E-BAM (m_vet) и оптическом датчике станции CityAir (s_vet), расположенных в Ветлужанке до (а) и после (б) очистки данных.

Представленные на рисунке 28 диаграммы рассеяния для температуры (а), давления (б) и влажности (в), полученные с АПН КВИАС и станции CityAir на poste «Ветлужанка», иллюстрируют согласованность показаний температуры и давления (коэффициент детерминации показаний одного типа датчика относительно другого в этом случае $R^2 > 0.98$). Однако измерения влажности разными приборами существенно отличаются (Рис. 28в). Полиномиальная регрессия третьего порядка ($R^2 = 0.78$) улучшает согласованность показаний (Рис. 28г), но все равно при повышенной влажности разброс показаний станции CityAir относительно АПН КВИАС остается более 40%. В дальнейшем анализе влажность не участвовала, хотя из физических соображений учет влияния влажности важен для построения моделей прогноза уровня концентраций.

Климат г. Красноярска континентальный с относительно морозной малоснежной зимой и жарким летом с малым количеством осадков. Более того, в холодный сезон большой вклад в уровень концентрации загрязнителя в атмосфере вносят работающие на полную мощность ТЭЦ и печное отопление. Поэтому естественна гипотеза о разном проявлении связи значений концентрации PM2.5 в атмосфере и метеорологических параметров в разные сезоны, что неизбежно будет влиять на модель корректировки датчиков.

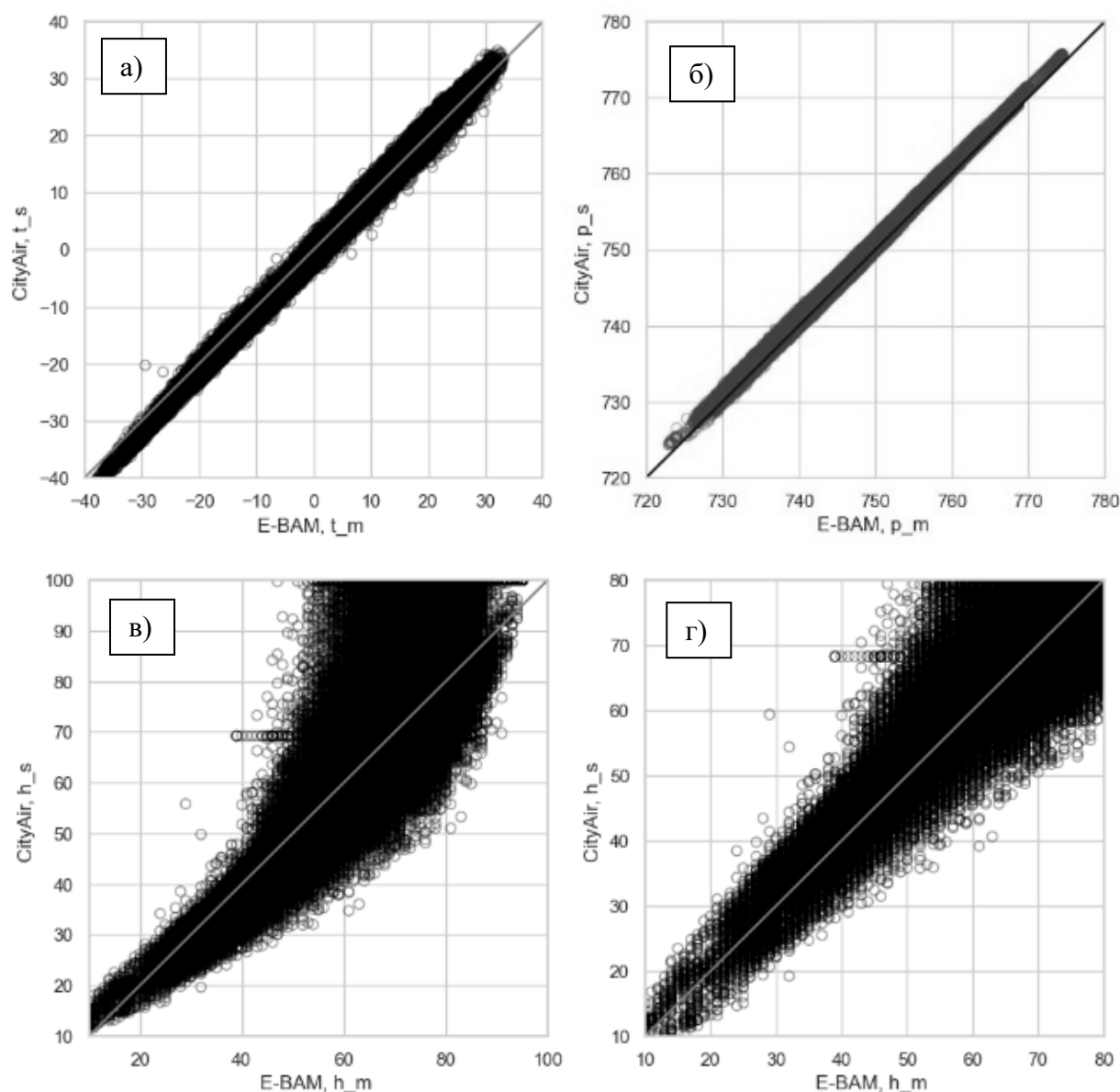


Рисунок 28 – Диаграммы рассеяния метеорологических параметров, полученных со станции CityAir и АПН КВИАС на посту Ветлужанка: температура (а), давление (б), влажность до (в) и после (г) коррекции.

Взятые по месяцам значения метеорологических параметров, распределены нормально. На основе корреляционного анализа можно сделать следующие предварительные выводы. Во-первых, в холодное время года существенны отрицательная корреляция концентрации PM_{2.5} с температурой и положительная с давлением и влажностью, а поздней весной и летом эти корреляции практически отсутствуют. Это легко объясняется следующими причинами. Низкие температуры, отсутствие ветра и высокая влажность являются причиной температурных инверсий в нижних слоях атмосферы, что затрудняет рассеяние загрязняющих веществ и приводит к периодам устойчивой повышенной концентрации их маркера PM_{2.5}. Более того, с понижением температуры повышается интенсивность печного отопле-

ния и работы ТЭЦ, что тоже увеличивает загрязнение нижних слоев атмосферы в холодный сезон. Есть и более общее замечание. Летом в целом концентрации $PM_{2.5}$ ниже, чем в холодный период, а точность измерений обоих видов датчиков при низких концентрациях взвешенных частиц ниже, чем при высоких. Таким образом, дисперсия обоих измерений растет, что плохо отражается на их согласованности. Во-вторых, в августе наблюдается слабая положительная корреляция загрязнений с температурой, тогда как в другие летние месяцы эта связь отсутствует. Такое поведение косвенно подтверждает, что измерения при низких концентрациях плохо согласованы, поскольку в августе в Красноярске часто возникают периоды повышения концентрации $PM_{2.5}$, связанные с лесными пожарами.

Таким образом, из результатов корреляционного анализа следует, что параметры регрессионной модели согласования показаний парных датчиков должны зависеть от сезона. Для последующего анализа нами были условно выделены три периода: 1) зимний (октябрь -- март) с сильной корреляцией между температурой и концентрацией $PM_{2.5}$; 2) летний (июнь – август); 3) демисезонный (апрель, май, сентябрь).

Был проведен регрессионный анализ для выявления зависимости уровней концентрации $PM_{2.5}$, полученных с помощью оптического датчика станции CityAir, от уровней, полученных от анализатора E-BAM, с учетом значений метеорологических параметров. Кроме того, мы сравнили следующие методы обучения регрессионной модели: линейная регрессия на основе метода наименьших квадратов (МНК), LASSO, эластичная сеть, метод опорных векторов, метод k ближайших соседей, дерево решений, случайный лес.

Имеющиеся данные были разбиты на обучающую (80% объема) и тестовую (20% объема) выборки. Качество модели оценивалось коэффициентом детерминации R^2 , который показывает, какую долю дисперсии тестовой выборки концентраций $PM_{2.5}$ объясняет модель. Поскольку значение коэффициента детерминации зависит от разбиения данных на обучающую и тестовую выборки, то процедура повторялась для 100 случайных разбиений. Далее в таблицах приведен средний коэффициент детерминации по всем попыткам, при этом среднеквадратичное отклонение не превышает процента.

В регрессионном анализе рассматривались различные комбинации следующих факторов: концентрации $PM_{2.5}$, полученные анализатором E-BAM (PM_m); температура (t_s) и давление (p_s), полученные с помощью датчиков станции CityAir. В качестве отклика рассматривались значения концентрации $PM_{2.5}$, полученные с оптического датчика CityAir (PM_s). Сделаем несколько замечаний относительно всех построенных нами линейных регрессионных моделей. Во-первых, все остатки имеют нулевое среднее, медиану в районе $0,8 \text{ мкг/м}^3$, слабую отрицательную симметрию ($\sim -0,2$) и умеренный эксцесс (~ 15). По критерию Дарбина-Ватсона автокорреляции первого порядка у остатков отсутствуют. Тем

не менее, статистическую проверку на нормальность остатки не проходят. Во-вторых, t -статистика с 5% уровнем значимости показывает, что во всех случаях коэффициенты значимо отличны от нуля. F -статистики с 5% уровнем значимости показывает, что во всех случаях существуют коэффициенты отличные от нуля, т.е. в этом смысле линейная модель приемлема. В-третьих, с помощью информационного критерия Акаике AIC и байесовского критерия BIC проведено сравнение качества линейных регрессий с различным набором факторов, построенных на одном и том же обучающем множестве. Анализ показал, что наименьшие AIC и BIC имеет регрессия, учитывающая PM_m , t_s и p_s .

В таблице 10 представлены R^2 , вычисленные для каждой из рассмотренных регрессионных моделей, обученных на всем объеме обучающей выборки. R^2 оценивался на основе полного объема данных тестовой выборки с учетом множественности факторов. Отметим, что регуляризация МНК не дала видимого улучшения R^2 , поэтому результаты методов LASSO и эластичной сети в таблице не приведены.

Таблица 10. Коэффициент детерминации R^2 регрессионных моделей, обученных на полном объеме данных обучающей выборки.

Факторы	Модели обучения				
	МНК	Опорных векторов	к ближайших соседей	Дерево решений	Случайный лес
PM_m	0,844	0,831	0,830	0,847	0,848
PM_m, t_s	0,856	0,850	0,870	0,789	0,864
PM_m, t_s, p_s	0,859	0,760	0,883	0,815	0,893

На основе результатов регрессионного анализа можно сделать следующие выводы. Во-первых, множественная линейная регрессия с помощью наименьших квадратов (МНК) даёт хорошее приближение, сравнимое по точности с более сложными и вычислительно-емкими методами машинного обучения. При этом линейная регрессия позволяет в явном виде получать коэффициенты, отражающие зависимость значения отклика от значений факторов. Во-вторых, для всех пар дублирующих датчиков лучшую точность предсказания отклика дают непараметрические методы «случайный лес» и « k ближайших соседей». В-третьих, добавление в анализ зависимости факторов влажности и давления не дает значительного улучшения точности моделей. Это можно объяснить двумя причинами: 1) всесезонность выборки гасит разнонаправленное влияние этих факторов в разные сезоны; 2) обсуждаемая ранее некорректность измерений влажности.

Построенные по полной обучающей выборке регрессионные модели мы оценили для отдельных групп значений тестовой выборки (табл. 11). Во-первых, коэффициент детерминации R^2 был вычислен по группам значений тестовой выборки, относящимся к одному сезону (столбцы «зима», «лето» и «деми-»). Во-вторых, R^2 был вычислен для наблюдений из тестовой выборки, соответствующих моментам, когда скользящее среднее за сутки значение концентрации PM2.5 не превышало принятого в России среднесуточного значения ПДК PM2.5 (35 мкг/м³) и наблюдениям, в которых среднесуточная концентрация превышала ПДК. В табл. 11 соответствующие столбцы помечены «ПДК, меньше» и «ПДК, больше», соответственно. В этом случае для того, чтобы избежать запаздывания периодов роста и спада концентрации PM2.5 текущее среднее значение вычислялось по наблюдениям, взятым за период 12 часов до текущего значения и 12 часов, начиная с текущего значения. Наконец, R^2 был вычислен для усредненных скользящим средним данных с окном 1 час, 6 часов, сутки.

Таблица 11. Коэффициент детерминации R^2 , рассчитанный для групп значений уровня концентрации PM2.5 на посту Ветлужанка тестовой выборки, для линейной регрессии (МНК), построенной на всём объеме данных обучающей выборки.

Факторы	Обучающая выборка							
	Скользящее среднее за			ПДК		ПО сезонам		
	сутки	6 ч.	1 ч.	меньше	больше	зима	деми-	лето
PM_m	0,939	0,936	0,895	0,324	0,739	0,862	0,220	0,582
PM_m, t_s	0,944	0,941	0,901	0,304	0,770	0,863	0,314	0,600
PM_m, t_s, p_s	0,954	0,947	0,907	0,356	0,771	0,867	0,367	0,603

Из данных табл. 11 следует, что линейная регрессия хорошо приближает скользящее среднее, и тем лучше, чем больше окно. В то же время ожидаемо предсказания отклика в период высоких концентраций точнее, чем в период низких. Это подтверждается и очень низкой точностью модели для демисезонного, и, особенно, летнего периодов. Кроме того, учет в модели температуры несколько повышает ее точность (R^2 увеличивается на проценты), добавление в модель давления еще незначительно улучшает точность прогноза отклика (R^2 увеличивается на десятые доли процента). Введение в модель влажности нецелесообразно.

Из проведенного исследования следует, что для повышения точности корректировки датчика необходимо обучать модели не на всей совокупности обучающей выборки, а предварительно выделять из всего множества данных целевую группу.

В таблице 12 представлены значения коэффициента детерминации R^2 после обучения трех моделей: линейной регрессии на основе метода наименьших квадратов (LR), случайного леса (RF) и « k ближайших соседей» (k -N). В названии строки указана группа данных, на которой проходило обучение. В частности, из табл. 2 следует, что обучение на осредненных данных с учетом всех факторов дает R^2 близкий к единице. Кроме того, мы видим значительное улучшение моделей на данных с небольшими значениями концентрации PM2.5.

Таблица 12. Коэффициент детерминации R^2 , рассчитанный для значений уровня концентрации PM2.5 оптического датчика «Ветлужанка» тестовой выборки для нескольких моделей регрессии, построенных на указанной в строке группе данных обучающей выборки.

Обучающая выборка	PM_m			PM_m, t_s			PM_m, t_s, p_s		
	LR	RF	k-N	LR	RF	k-N	LR	RF	k-N
Скользящее среднее									
за сутки	0,945	0,945	0,949	0,959	0,991	0,988	0,959	0,997	0,997
за 6 часов	0,944	0,944	0,945	0,955	0,978	0,977	0,955	0,984	0,983
за час	0,904	0,904	0,898	0,914	0,930	0,931	0,915	0,943	0,938
ПДК									
не превышает	0,616	0,621	0,577	0,631	0,654	0,683	0,631	0,725	0,712
превышает	0,772	0,769	0,752	0,795	0,796	0,799	0,807	0,840	0,820
Зима	0,889	0,893	0,883	0,890	0,879	0,889	0,890	0,898	0,893
Лето	0,795	0,782	0,788	0,797	0,787	0,799	0,799	0,874	0,823
Демисезон	0,693	0,709	0,667	0,715	0,712	0,723	0,716	0,773	0,737

Таким образом, результаты анализа, представленные в таблицах 10–11 показывают, что корректировка концентраций PM2.5 для оптического датчика CityAir относительно анализатора Е-ВАМ должна выполняться, по крайней мере, по сезонным данным. Для оперативной корректировки PM_s удобно использовать параметрическую модель множественной линейной регрессии методом наименьших квадратов, учитывающую показания температуры и давления. В этом случае фактором является концентрация PM2.5, измеренная оптическим датчиком CityAir, а для обучения модели в качестве отклика используются концентрации PM2.5, измеренные анализатором Е-ВАМ.

В результате описанного статистического анализа мы имеем подсистему из четырех оптических датчиков, откалиброванных по показаниям эталонных анализаторов Е-ВАМ. Будем использовать эти датчики для корректировки оптических датчиков CityAir всех других постов системы мониторинга КНЦ СО РАН. На основе корреляционного анализа для

каждого оптического датчика найден откалиброванный, у которого с ним максимальный коэффициент корреляции. На рисунке 29 все датчики нанесены на карту и одним цветом закрашены топографические области, объединяющие датчики, коэффициенты корреляции которых максимальны с одним из четырех откалиброванных датчиков.



Рисунок 29 – Области максимальной корреляции показаний датчиков сети КНЦ СО РАН с показаниями скорректированных датчиков постов «Ветлужанка», «Покровка», «Свердловский», «Кировский».

Полученные топографические области хорошо объясняются следующими географическими фактами. Во-первых, датчик поста «Покровка» расположен на хорошо проветриваемом высоком холме, вдали от транспортных развязок. Ожидается, что этот датчик будет наименее коррелирован с другими. Поэтому в зону его действия не попал ни один датчик. Во-вторых, самая большая зона влияния (Зона 1) оказалась у датчика поста «Ветлужанка». Это можно объяснить преобладанием северо-западных ветров и расположением датчика на северо-западе в жилом районе Красноярска. Все датчики Зоны 1 относятся к левому берегу Енисея. Эти датчики будут аналогичным образом реагировать на суточные колебания уровней концентрации $PM_{2,5}$, связанные, например, с пробками на дорогах, и длительные периоды высокой концентрации $PM_{2,5}$, связанные с неблагоприятными метеорологическими условиями. Может показаться удивительным, что в зону влияния (Зона 2) датчика «Свердловский», расположенного на правом берегу Енисея, попали датчики с левого берега. Однако все эти датчики расположены в хорошо проветриваемой относительно чистой части

города, расположенной в долине реки Енисей. Поэтому понятна и их хорошая корреляция с датчиком, находящимся в аналогичных условиях. Зона влияния (Зона 3) датчика «Кировский» объединяет небольшую группу датчиков, расположенных в наиболее загрязненной части долины реки Енисей, а также на островах, через которые проходит мост с плотным транспортным потоком.

Заключение. Предложена статистически обоснованная корректировка первичных данных об уровне концентрации взвешенных частиц PM_{2.5} в приземном слое атмосферы г. Красноярска, получаемых оптическими датчиками станций CityAir. Для построения регрессионных моделей эталонными считались измерения, получаемые от анализаторов Е-ВАМ, расположенных на тех же постах наблюдения, что и корректируемые датчики.

На основе проведенного анализа можно сделать следующие выводы.

1. При корректировке показаний датчиков необходимо учитывать метеорологические параметры. В нашем случае в линейной регрессионной модели лучше учитывать зависимость от температуры и давления. Точность измерения влажности не позволяет учитывать этот показатель при корректировке, однако не исключает его влияния на показания датчика.
2. Параметры корректировки уровня концентрации PM_{2.5} с помощью регрессионной модели существенно зависят от сезона.
3. Показания оптических датчиков станций CityAir полностью отражают тренды показателей загрязнения, поэтому эти датчики могут использоваться для описания сценариев и прогнозов периодов повышенных концентраций взвешенных частиц в атмосфере городов.
4. Для оперативной корректировки значений концентрации PM_{2.5} датчиков станций CityAir достаточно использовать линейную многофакторную регрессионную модель на основе метода наименьших квадратов. Для научных ретроспективных исследований временных рядов концентраций PM_{2.5} рекомендуется использовать многофакторную регрессию, основанную на методе случайного леса.

Основные публикации по разделу:

1. Карпова Е.Д., Петракова В.С. Статистически обоснованная корректировка показаний датчиков станций CityAir уровня концентрации взвешенных частиц PM_{2.5} в приземном слое атмосферы города // Информатика и автоматизация. 2024. Т. 23(2). С. 352–376.
DOI: 10.15622/ia.23.2.2 [Scopus Q4, Белый список-2]

2.5. Формирование и последствия температурных инверсий в г. Красноярске

В атмосфере ряда промышленных городов России отмечаются повышенные концентрации загрязняющих веществ, обусловленные индустриальными и транспортными выбросами, дымами лесных пожаров и частных домовладений, зарегулированностью рек и неблагоприятными метеоусловиями. К неблагоприятным метеоусловиям, прежде всего, следует отнести ситуации, связанные с формированием инверсных температурных слоев в атмосфере, что приводит к затруднению конвективного и турбулентного переноса примесей от поверхности Земли в верхние слои атмосферы. В работах [1–3] указывалось, что особо опасны условия накопления загрязнения в атмосфере возникают при наличии приземных и приподнятых инверсий, уменьшающих толщину слоя перемешивания примесей, которые в свою очередь определяются метеорологическими условиями и рельефом местности. Большое влияние стационарных зимних антициклонов, при которых создаются условия застоя воздуха, благоприятные для накопления вредных примесей в воздухе, отмечается в ряде исследований [4–8]. Анализ инверсий и их влияния на уровни загрязнения атмосферного воздуха рассматриваются в работах [9–11]. Первые исследования вертикальных профилей температуры атмосферы в г. Красноярске представлены в работе [12]. В работах [13,14] исследовались вертикальные распределения температуры атмосферы для городских островов тепла.

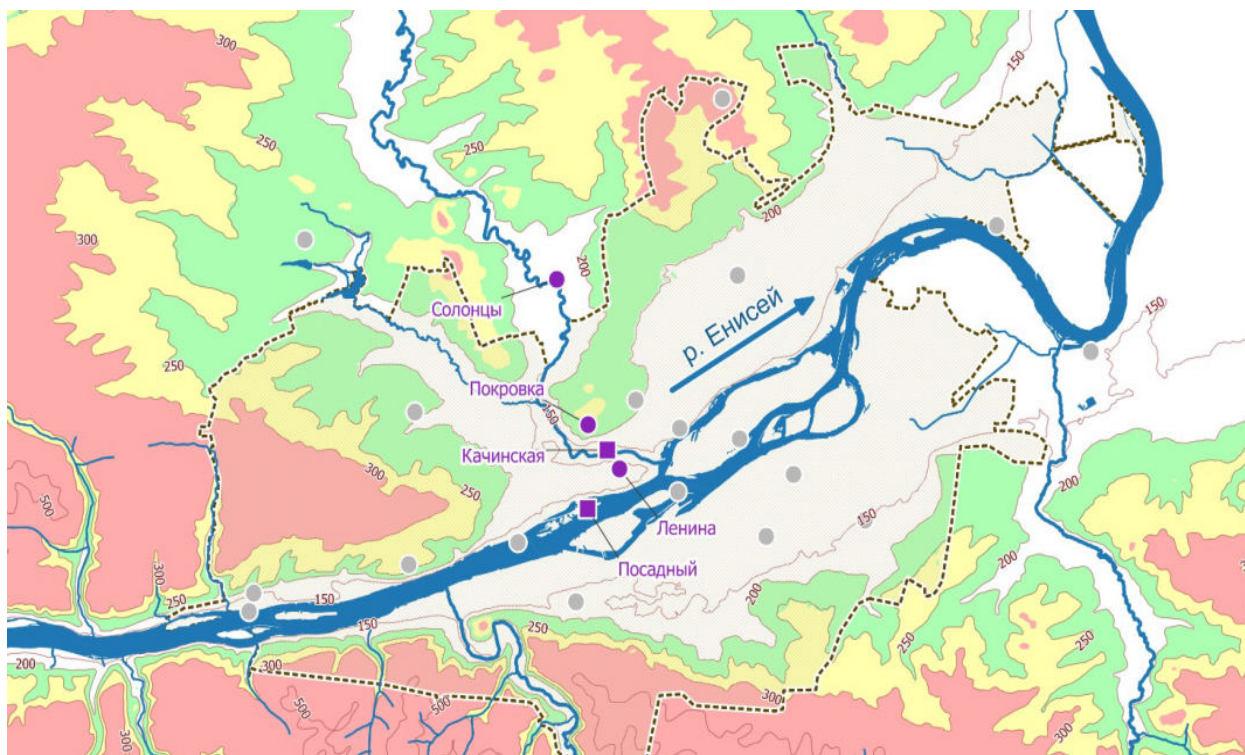


Рисунок 30 – Карта области исследования (г. Красноярск) с постами мониторинга воздуха.

Объектом исследования являлось состояние атмосферы в г. Красноярске в зимнее время года, зависящее от метеоусловий, выбросов в атмосферу, рельефа местности, зарегулированности р. Енисей плотиной Красноярской ГЭС. Мониторинг состояния загрязнения атмосферного воздуха в г. Красноярске ведётся рядом организаций федерального и региональных уровней. Использовались данные системы мониторинга КНЦ СО РАН на основе сертифицированных станций CityAir, которые измеряют температуру, атмосферное давление, относительную влажность воздуха, концентрацию частиц PM2.5 (Рис. 30, посты «Солонцы», «Ленина», «Покровка») Скорость и направление ветра измерялась на краевой станции наблюдения «Покровка». Вертикальные распределения температуры в атмосфере регистрировались двумя температурными профилемерами МТП-5, размещёнными на постах «Качинская» и «Посадный». Данные получают каждые 5 минут и передаются на сервер в ИВМ СО РАН [15, 16]. В дальнейшем на их основе вычисляются периоды температурных инверсий в атмосфере.

Формирование температурной инверсии в атмосфере.

Будем рассматривать типичную ситуацию – период с 23 декабря 11:00 по 24 декабря 14:00 2022 г. На рисунке 31 представлена динамика температуры атмосферы, скорости и направления ветра.

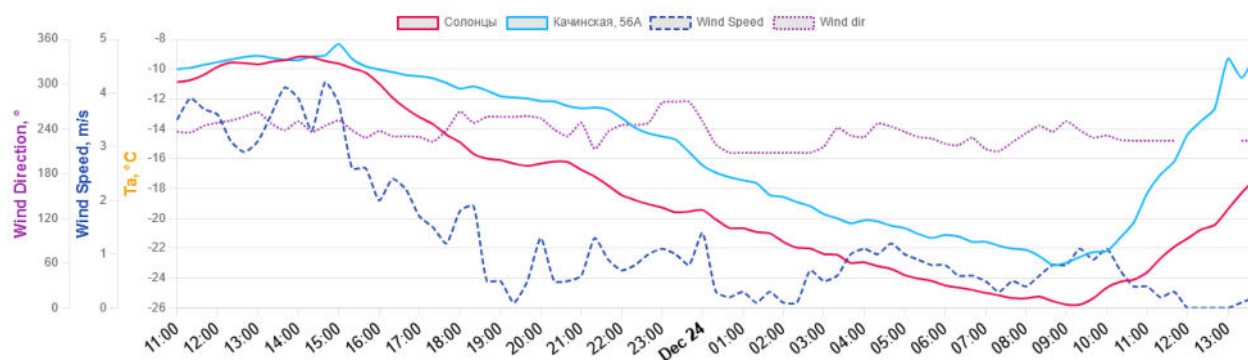


Рисунок 31 – Динамика температуры воздуха на постах Солонцы и Качинская, а также скорость (WS) и направление ветра (WD) на посту Покровка в период 23-24 декабря 2022 г.

В начале периода температура атмосферы в черте города превышает температуру в пригороде, что указывает на формирование «острова тепла», вызванное социально-экономическими процессами и приводит к нагреванию нижней части приземного слоя атмосферы. В дальнейшем холодный воздух с юго-западного направления (240°) со скоростью около 1 м/с поступает в черту города по руслам рек Енисей и Кача. Это приводит к формированию холодной волны в воздушном «острове тепла» и инициирует создание инверсного слоя температуры в атмосфере.

На рисунке 32 показаны данные с профилемера (Качинская): температура атмосферы на высоте 2м (T_a), нижняя и верхняя границы инверсного слоя, величина инверсии ΔT и концентрация $PM_{2.5}$ с ближайшего метеопоста (Ленина).

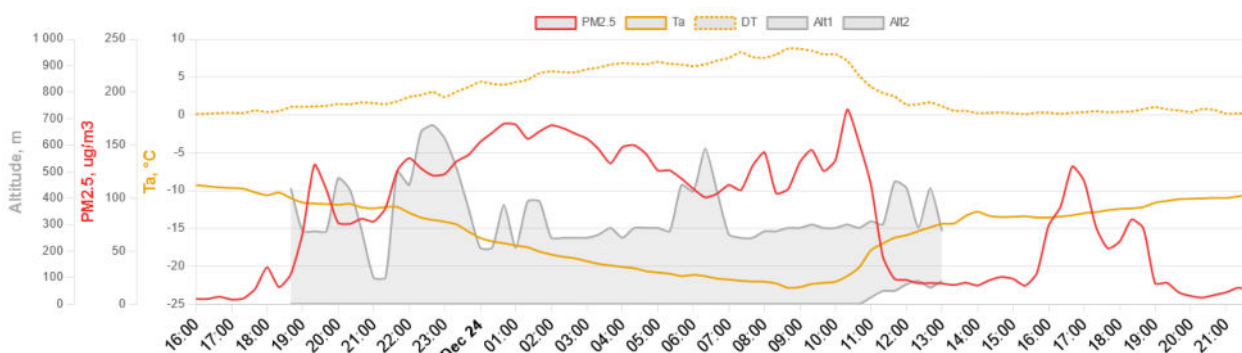


Рисунок 32 – Температура атмосферы на высоте 2м (T_a), нижняя и верхняя границы инверсного слоя, величина инверсии ΔT и концентрация $PM_{2.5}$.

Начальная инверсия образуется при приземной температуре -10°C с верхней границей инверсии до 100м (Рис. 33), в дальнейшем температура воздуха только понижается на всех высотах, но особенно сильно вблизи поверхности, где находится охлаждающий слой. После достижения минимума температуры воздуха T_a (и максимальной величины инверсии), основное влияние начинает оказывать городской остров тепла. Начинает нагреваться нижняя граница приземного слоя воздуха и происходит теплопередача в более высокие слои атмосферы. Это приводит к преобразованию инверсии в приподнятую форму с дальнейшим переходом в нормальное распределение температуры с высотой. Соответственно это приводит к рассеиванию примесей в атмосфере и очищению атмосферы города.

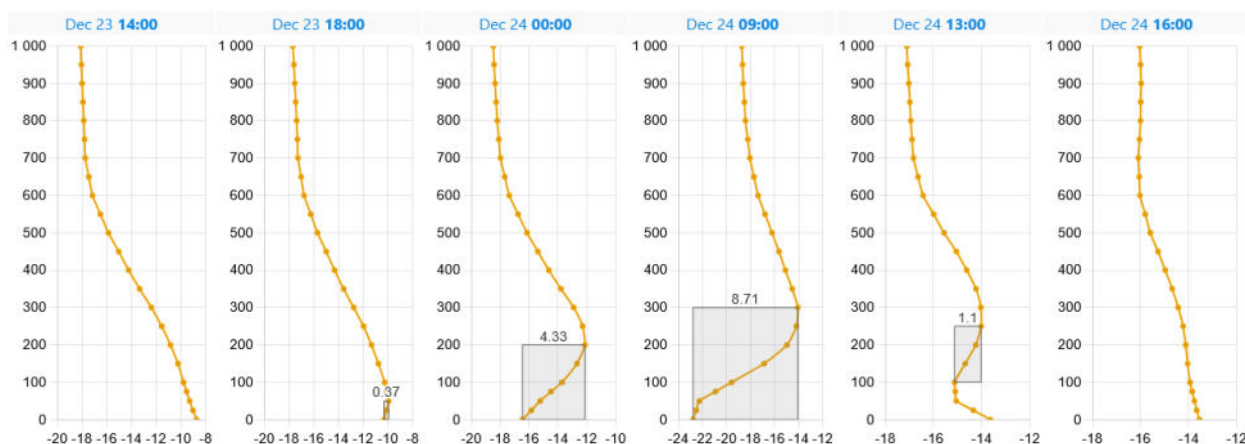


Рисунок 33 – Температурный профиль атмосферы и величина инверсии (Качинская).

На рисунке 34 представлена динамика вертикального распределения температуры воздуха на острове Посадный. Начальный график соответствует 16:00 23 декабря 2022 года и представляет распределение температуры для нормальных метеоусловий: уменьшение температуры с высотой. Формирование инверсии начинается только в 22:00 23 декабря, которая сразу переходит в приподнятую форму. Величина инверсии достигает максимального значения 4°C в 09:00 24 декабря с нижней границей на высоте 50м и верхней на высоте 300м. В последующем величина инверсии уменьшается и в 21:00 исчезает.

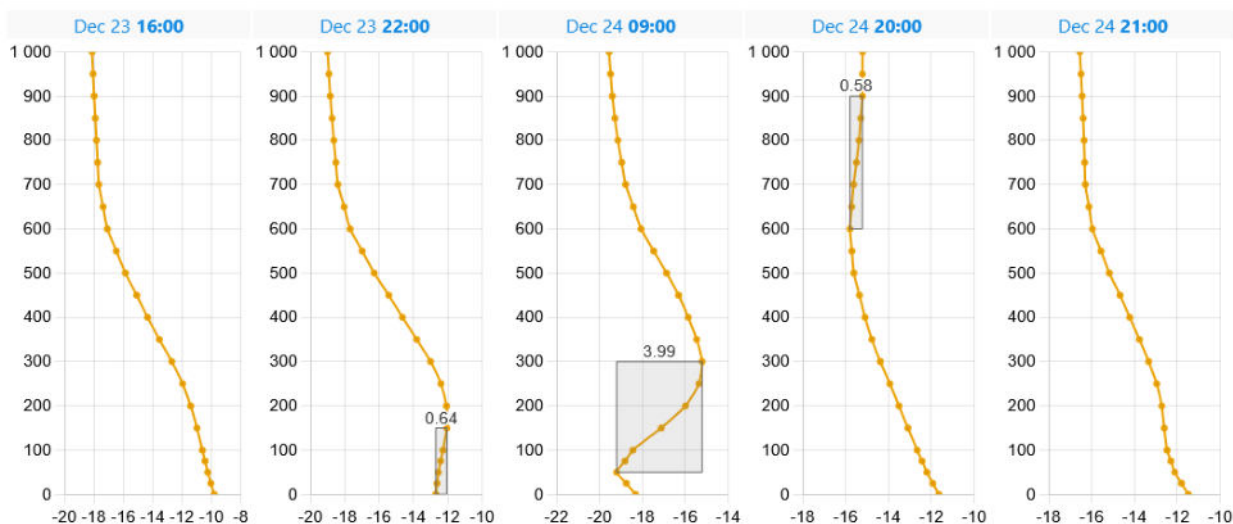


Рисунок 34 – Температурный профиль атмосферы и величина инверсии (Посадный).

Полученные данные показывают, что величина инверсии над рекой Енисей меньше чем над городом и она имеет приподнятую форму, поэтому ее действие более ослабленное. Более того, иногда инверсии присутствуют при низких концентрациях $PM_{2.5}$, потому что нижняя граница инверсного слоя приподнята достаточно высоко. И связано это с теплой поверхностью реки и дополнительным ветром, индуцированным течением реки.

Температурные инверсии и загрязнение атмосферы частицами $PM_{2.5}$

В этом разделе представлена взаимосвязь степени загрязнения атмосферы частицами $PM_{2.5}$ в зависимости от вида и величины температурной инверсии в атмосфере. Концентрация частиц $PM_{2.5}$ измерялась на посту мониторинга, расположенном рядом с профилемером (Рис. 30, Ленина).

Использовались данные регулярных измерений (каждые 20 мин) за 2022-2023 гг. На рисунке 35 показана повторяемость нижней границы инверсионного слоя в зависимости от концентрации $PM_{2.5}$ в атмосфере.

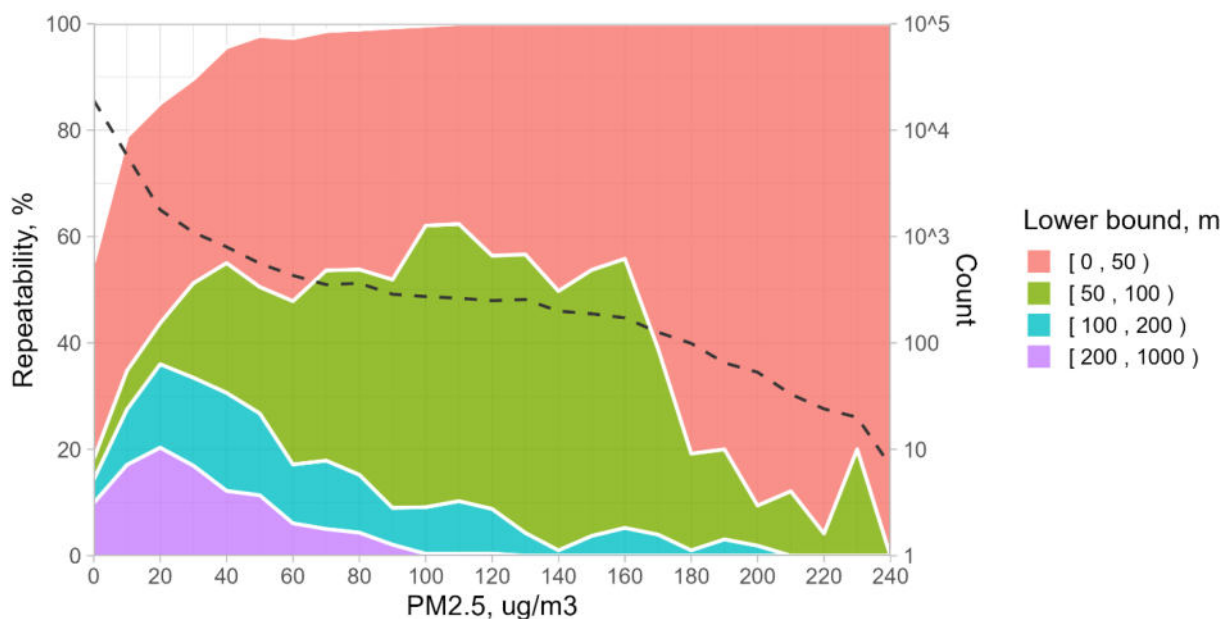


Рисунок 35 – Взаимосвязь концентрации $PM_{2.5}$ и нижней границы температурной инверсии (по правой вспомогательной шкале пунктирной линией показано количество соответствующих 20 минутных интервалов).

Из полученных результатов следует, что в области допустимых среднесуточных концентраций $PM_{2.5}$ равных 35 мкг/м^3 , вероятность инверсии составляет 93%. Повторяемость нижней границы инверсионного слоя: до 50 м – 40%, 50-100 м – 21%, 100-200 м – 18%, более 200 м – 14%. С увеличением концентрации частиц $PM_{2.5}$ уменьшается роль приподнятых температурных инверсий с нижней границей более 100 м. А при концентрациях более 180 мкг/м^3 определяющий вклад дает приземная инверсия.

В целом можно сделать вывод, что определяющую роль в загрязнении атмосферы играют инверсии 1 и 2 категории, нижняя граница которых расположена от 0 до 100 м. При таких условиях и восточном ветре, промышленные, транспортные выбросы и дымы от печей отопления частного сектора запираются в котловине города. Отметим, что в черте города направление ветра неравномерно, что связано с ландшафтом местности и городской застройкой высотными зданиями. Если рассмотреть аналогичные данные, но с добавлением условия на минимальную величину инверсии ($\Delta T > 1^\circ\text{C}$) – характер зависимости величины загрязнения от категории атмосферной инверсии сохраняется.

Более полная взаимосвязь концентраций $PM_{2.5}$ и величины инверсии показана на рисунке 36. Повышенные величины температурной инверсии обуславливают возникновение более высоких концентрации $PM_{2.5}$ в атмосфере.

Таким образом, инверсии 1 и 2 категории с величиной инверсии более 5°C оказывают наиболее значимое влияние на загрязнение атмосферы.

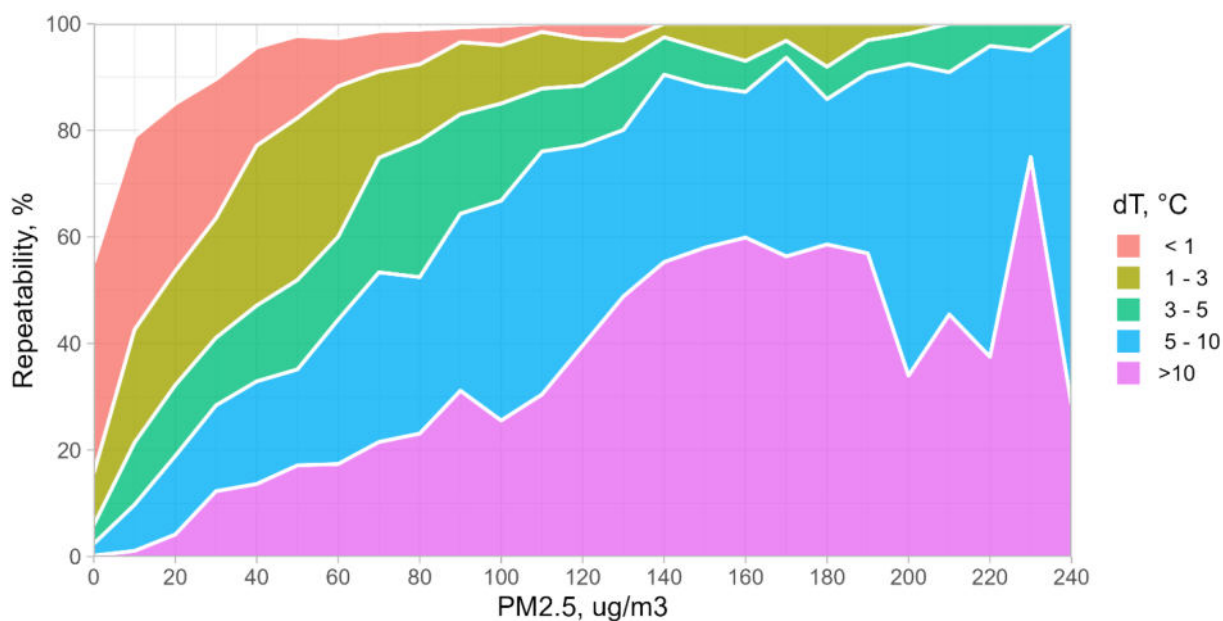


Рисунок 36 – Взаимосвязь концентрации $PM_{2.5}$ и величины температурной инверсии.

Влияние метеоусловий на величину загрязнения частицами $PM_{2.5}$

В разделе представлены статистические данные зависимости концентрации $PM_{2.5}$ в атмосфере от температуры, относительной влажности воздуха, атмосферного давления, направления и скорости ветра. Метеопараметры регистрировались на посту «Ленина», за исключением ветра, который измерялся на посту «Покровка». Использовались данные за холодный период 2022 года (январь, февраль, ноябрь, декабрь), усредненные по 20 минутным интервалам. Отметим, что концентрации $PM_{2.5}$ ниже 25 мкг/м^3 фактически являются фоновыми.

На рисунке 37 представлена зависимость $PM_{2.5}$ от различных метеопараметров. Для $PM_{2.5} > 35 \text{ мкг/м}^3$ высокие концентрации $PM_{2.5} > 100 \text{ мкг/м}^3$ наблюдаются при температурах атмосферы ниже -10°C . Значения относительной влажности воздуха находятся в диапазоне $55\% < RH < 85\%$.

Из диаграммы видно, что высокие концентрации наблюдаются при скоростях ветра менее 2 м/с . Большие скорости ветра не позволяют сформироваться инверсному слою в атмосфере. Т.е. турбулентный теплообмен слабее ветрового перемешивания воздуха. Обратим внимание, что даже при штилевых условиях конвективно-турбулентный теплообмен формирует инверсию в атмосфере несмотря на отсутствие адвективного притока холодного воздуха.

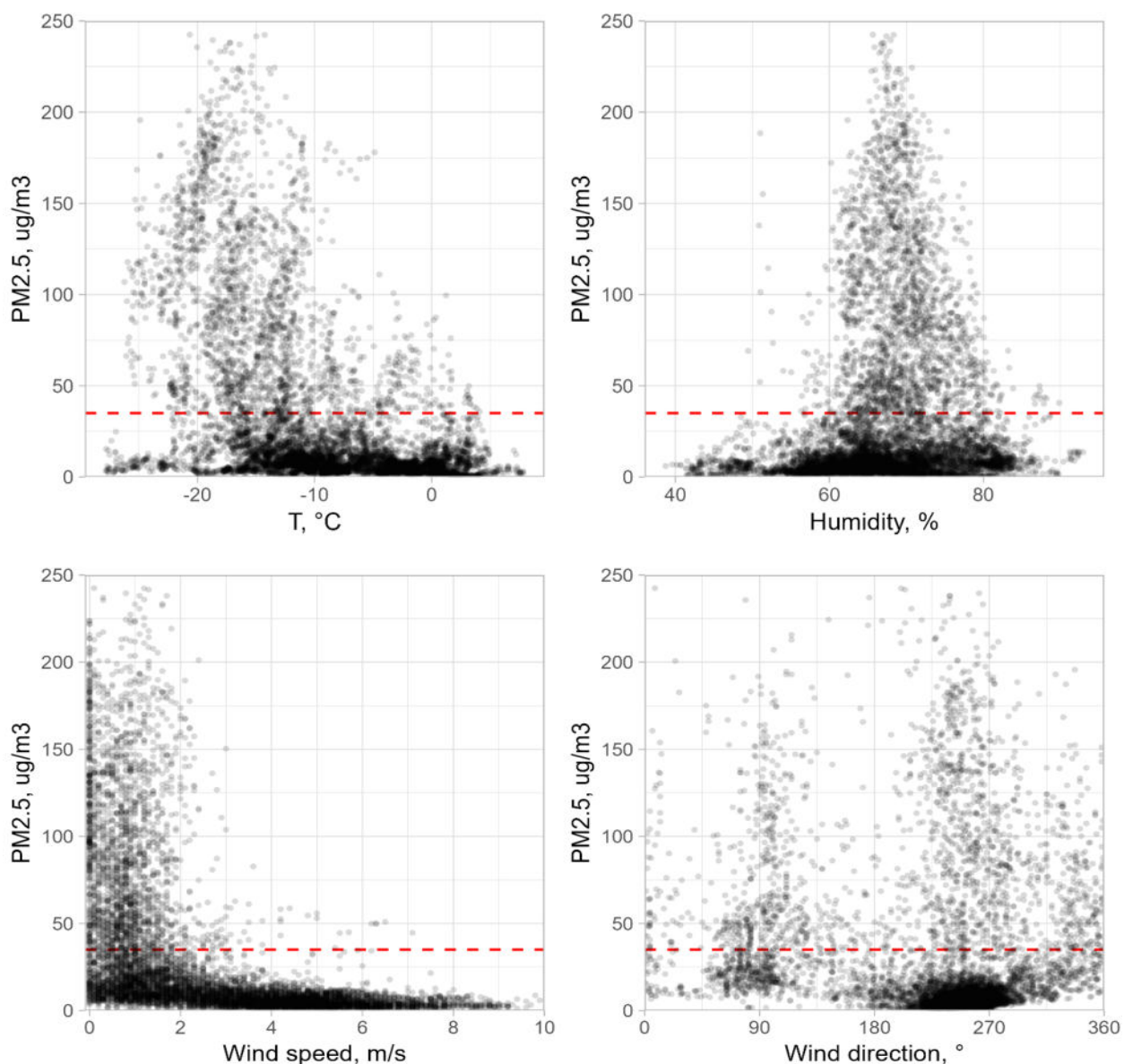


Рисунок 37 – Точечные диаграммы зависимости $PM_{2.5}$ от температуры воздуха (а), относительной влажности воздуха (б), скорости ветра (в) и направления ветра (г).

Заключение. Основная причина формирования инверсных слоев – подтекание холодного воздуха под приземный «остров тепла» в атмосфере Красноярска. На повышение концентрации $PM_{2.5}$ определяющую роль играют приземная инверсия и приподнятая инверсия с нижней границей до 100 м и величиной инверсии более $5^{\circ}C$. Инверсный слой над р. Енисей значительно слабее и его влияние меньше на формирование загрязнения. Величина инверсии растет с уменьшением температуры воздуха. При повышенных значениях концентрации $PM_{2.5} > 35 \text{ мкг/м}^3$ в зимнее время обычно наблюдаются следующие метеоусловия: $-25^{\circ}C < T_a < 0^{\circ}C$; $55\% < RH < 85\%$; $750 \text{ мм рт. ст.} < P < 770 \text{ мм рт. ст.}$; ветер до 2 м/с восточного или западного направления.

Список цитируемой в разделе литературы:

1. Holzworth G.C. Estimates of Mean Maximum Mixing Depths in the Contiguous United States // Mon. Weather Rev. 1964. Т. 92, № 5. С. 235–242.
2. Kauper E.K., Hopper C.J. The Utilization of Optimum Meteorological Conditions for the Reduction of Los Angeles Automotive Pollution // J. Air Pollut. Control Assoc. 1965. Т. 15, № 5. С. 210–214.
3. Берлянд М.Е., Генихович Е.Л., Ложкина В.П., Оникул Р.И. Численное исследование атмосферной диффузии при нормальных и аномальных условиях стратификации // Труды ГГО. 1964. № 158.
4. Garnett A. A Survey of Air Pollution in Sheffield under Characteristic Winter Anticyclonic Conditions // Int. J. Air Water Pollut. 1963. Т. 7. С. 963–968.
5. Holzworth G.C. Some meteorological aspects of community air pollution // Air Eng. 1964. Т. 6. С. 26–28 and 33–37.
6. Miller M.E. Semi-objective forecasting of atmospheric stagnation in the Western United States // Mon. Weather Rev. 1964. Т. 92, № 1. С. 23–32.
7. Niemeyer L.E. Forecasting air pollution potential // Mon. Weather Rev. 1960. Т. 88, № 3. С. 88–96.
8. Безуглая Э.Ю. Инверсии нижней тропосферы и их влияние на загрязнение воздуха Москвы // Труды ГГО. 1968. № 207. С. 202–206.
9. Ячмёнева Н.В., Гольвей А.Ю. Повторяемость инверсий и их влияние на уровень загрязнения атмосферного воздуха в г. Челябинске // Вестник ЧелГУ. 2011. № 5 (220).
10. Локощенко М.А., Богданович А.Ю., Еланский Н.Ф., Лезина Е.А. Температурные инверсии в Москве и их влияние на состав приземного воздуха // Известия Российской академии наук. Физика атмосферы и океана. 2021. Т. 57, № 6. С. 641–650.
11. Яушева Е.П., Гладких В.А., Камардин А.П., Шмаргунов В.П. Экстремальные аэрозольные загрязнения атмосферы в зимний период в Академгородке г. Томска. // Оптика атмосферы и океана. 2023. Т. 36, № 9(416). С. 711–717.
12. Битехтина М.А., Михайлюта С.В., Леженин А.А., Тасейко О.В. Эволюция пограничного слоя и особенности загрязнения атмосферы в условиях города // Вестник КемГУ. 2012. Т. 2, № 4 (52). С. 12–16.
13. Хайкин М.Н., Кадыгров Е.Н., Кузнецова И.Н. Влияние высокой концентрации аэрозоля на термическую структуру пограничного слоя атмосферы // Изв. РАН. 2006. Т. 42, № 6. С. 778–784.

14. Еланский Н.Ф., Лаврова О.В., Мохов И.И., Ракин А.А. Структура острова тепла над городами России по наблюдениям с передвижной лаборатории // ДАН. 2012. Т. 443, № 3. С. 366–371.
15. Якубайлик О.Э., Токарев А.В., Кадочников А.А., Якубайлик Т.В., Заворуев В.В. Информационно-вычислительное обеспечение мониторинга атмосферы г. Красноярска // Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов. Новосибирск: ФИЦ ИВТ, 2023. С. 26–32.
16. Yakubailik O.E., Kadochnikov A.A., Tokarev A.V. WEB Geographic Information System and the Hardware and Software Ensuring Rapid Assessment of Air Pollution // Optoelectron. Instrum. Data Process. 2018. Т. 54, № 3. С. 243–249.

Основные публикации по разделу:

1. Шапарев Н.Я., Токарев А.В., Якубайлик О.Э. Формирование температурных инверсий и их влияние на состояние атмосферы // Аэрозоли Сибири. XXXI конференция: Тезисы докладов. – Томск: Изд-во ИОА СО РАН, 2024. – ISBN 978-5-94458-203-4. – С. 43.

2.6. Прогнозирование величины концентрации загрязняющих веществ в атмосфере г. Красноярска с помощью моделей машинного обучения

Красноярск является одним из нескольких городов России с самым грязным воздухом, концентрация вредных веществ в котором часто превышает допустимые нормы, установленные Российским законодательством по охране окружающей среды [1]. Уровень загрязнения атмосферного воздуха в Красноярске обусловлен многочисленными факторами – пространственной структурой и временной динамикой загрязняющих выбросов, а также особенностями метеорологических условий, которые на территории города также обладают значительной временной и пространственной изменчивостью. Возникает необходимость не только в мониторинге, но и в прогнозировании величины концентрации вредных веществ в атмосфере города.

Одними из наиболее эффективных и применяемых моделей прогнозирования загрязняющих веществ являются: модели множественной регрессии, модели временных рядов ARIMA и ARIMAX, ансамблевые модели случайного леса и градиентного бустинга [2].

Целью работы являлось применение различных моделей машинного обучения для прогнозирования величины концентрации твердых взвешенных частиц PM_{2.5} в атмосфере города Красноярска.

Районом исследования в настоящей работе является г. Красноярск. Для прогнозирования были взяты данные наземных станций «Системы мониторинга КНЦ СО РАН» за пять

лет с января 2019 г. по март 2024 г. К рассматриваемым признакам были добавлены температурные инверсии из модели реанализа National Centers for Environmental Prediction Global Forecast System (NCEP GFS), так как они имеют сильную положительную корреляцию с величиной концентрации PM_{2.5}. Используемые данные представлены в таблице 13.

Таблица 13. Данные для анализа с помощью моделей машинного обучения.

Наземные станции мониторинга, на высоте 2 м	NCEP GFS
Концентрация частиц PM _{2.5}	Разность температур на изобарических слоях 1000 мбар – 925 мбар
Температура воздуха	Разность температур на изобарических слоях 925 мбар – 850 мбар
Влажность воздуха	Разность температур на изобарических слоях 1000 мбар – 850 мбар
Атмосферное давление	
Скорость ветра	
Направление ветра	

Данные были усреднены по суткам и по следующим датчикам: Николаевка, Овинный, Песчанка, Телевизорная и Ветлужанка. В этих районах датчики обычно показывают наиболее высокие концентрации PM_{2.5}, что является индикатором наличия загрязнений атмосферного воздуха. Также данные были разделены на временные периоды, в зависимости от величины концентрации PM_{2.5}, приблизительно соответствующие сезонам года [3].

Модель множественной линейной регрессии имеет вид:

$$y(x) = w_0 + \sum_{i=1}^n w_i x_i,$$

где $y(x)$ – целевая функция от выборки x ; w_0 – свободный коэффициент; x_i – признаки; w_i – веса признаков; n – количество используемых признаков.

Для прогнозирования использовались шесть моделей регрессии: линейная, линейная с L1-регуляризацией (Lasso), с L2-регуляризацией (Ridge), полиномиальные регрессии 2й и 3й степени, градиентный бустинг регрессий. Прогноз проводился на различных вариациях выборок, в том числе, когда в качестве признаков использовались главные компоненты [4].

Модель временных рядов ARIMA состоит из двух частей: авторегрессионной (AR) части и скользящего среднего (MA):

$$y_t = \beta + \sum_{i=1}^p a_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t,$$

где y_t – значение временного ряда в момент времени t ; ε_t – шумовая компонента в момент времени t ; p – количество предыдущих значений временного ряда, используемых в моде-

ли; q – количество запаздывающих ошибок прогноза; a_i , θ_i , β – параметры модели, которые необходимо оценить.

Расширенная модель ARIMAX позволяет также учесть дополнительные внешние факторы (экзогенные переменные), оказывающие влияние на целевую переменную. В качестве экзогенных переменных использовались значения температуры и влажности воздуха, направления и скорости ветра, атмосферного давления.

Модель случайного леса является ансамблевым подходом к построению моделей машинного обучения и использует ансамбль решающих деревьев:

$$\tilde{y}(x) = \frac{1}{k} (b_1(x) + \dots + b_n(x)),$$

где $\tilde{y}(x)$ – прогноз модели на выборке x , $b(x)$ – модель решающего дерева, k – количество решающих деревьев в лесу.

В модели градиентного бустинга построение базовых моделей происходит последовательно, т.е. каждое решающее дерево стремится уменьшить ошибку предыдущего, что позволяет улучшить качество комбинации моделей. Для подбора гиперпараметров моделей использовался поиск по сетке с применением кросс-валидации.

Для оценки точности прогноза моделей использовались четыре метрики качества: *MAE* (средняя абсолютная ошибка), *MSE* (средняя квадратичная ошибка), *MAPE* (средняя абсолютная процентная ошибка), R^2 (коэффициент детерминации).

Для каждого класса используемых моделей была выбрана модель, которая показывает наилучшее качество прогноза. Для моделей множественных регрессий лучшее качество прогноза в зимний и весенний периоды показывают полиномиальная регрессия 2-й степени, в летний период – Lasso регрессия, в осенний период – градиентный бустинг на регрессиях. Для моделей временных рядов ARIMA и ARIMAX лучшее качество прогнозирования на 6 дней вперед практически на всех временных периодах показывает класс моделей ARIMAX. При сравнении ансамблевых моделей случайного леса и градиентного бустинга лучшее качество прогноза на всех временных периодах показывает градиентный бустинг.

Сравнение моделей класса ARIMAX с моделями градиентного бустинга и регрессионными моделями показало, что они справляются с задачей прогнозирования лучше за исключением летнего периода, где схожее качество прогноза также показывают модели градиентного бустинга и регрессионные модели.

Примеры прогноза моделями градиентного бустинга и ARIMAX представлены для зимнего и летнего периода на 3 дня вперед на рисунках 38-41.

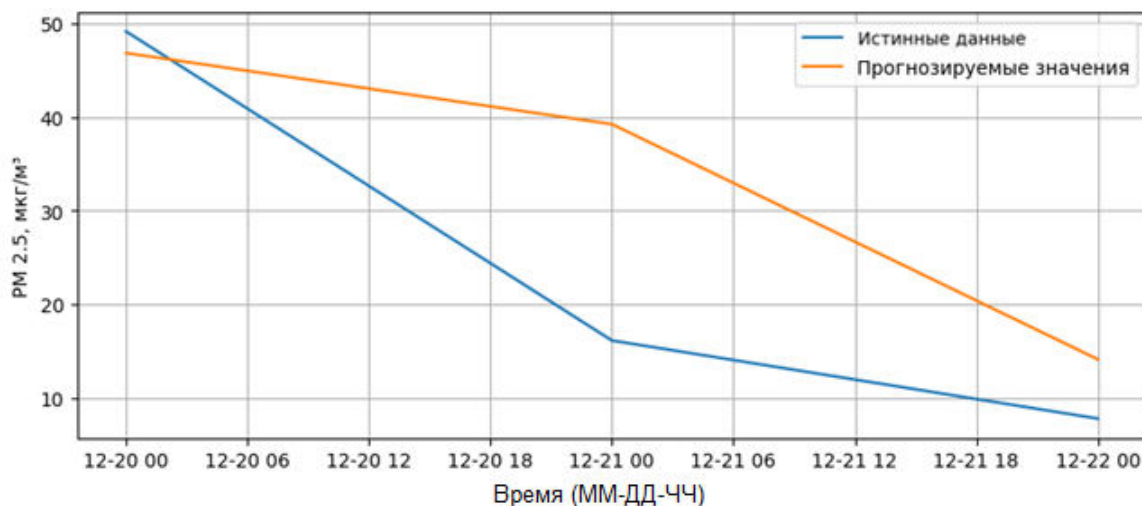


Рисунок 38 – Прогнозирование величины $PM_{2.5}$ моделью градиентного бустинга.
Срез зимнего периода: с 20 по 22 декабря 2023 г. ($MSE = 193,25$; $MAPE = 0,76$).

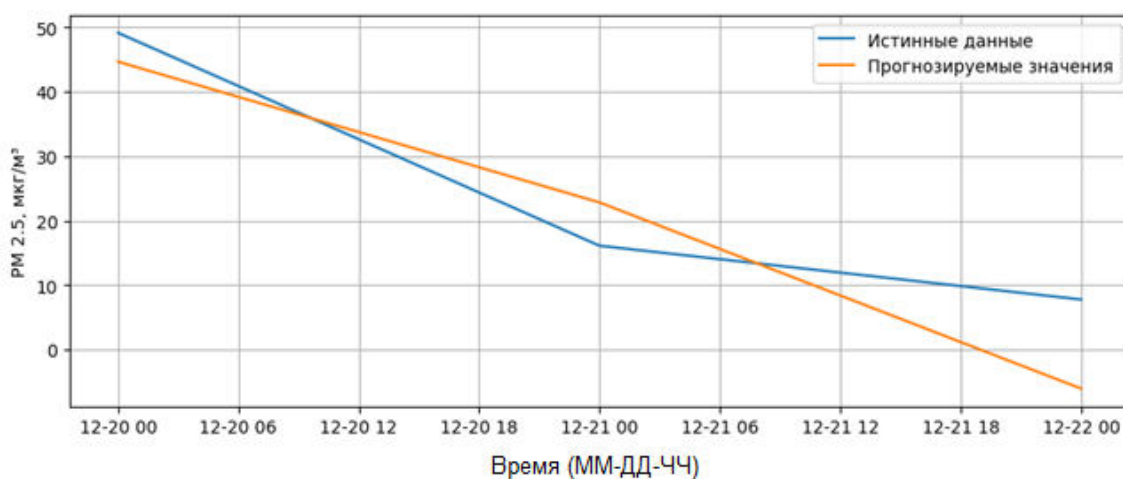


Рисунок 39 – Прогнозирование величины $PM_{2.5}$ моделью временных рядов ARIMAX.
Срез зимнего периода: с 20 по 22 декабря 2023 г. ($MSE = 85,4$; $MAPE = 0,76$).

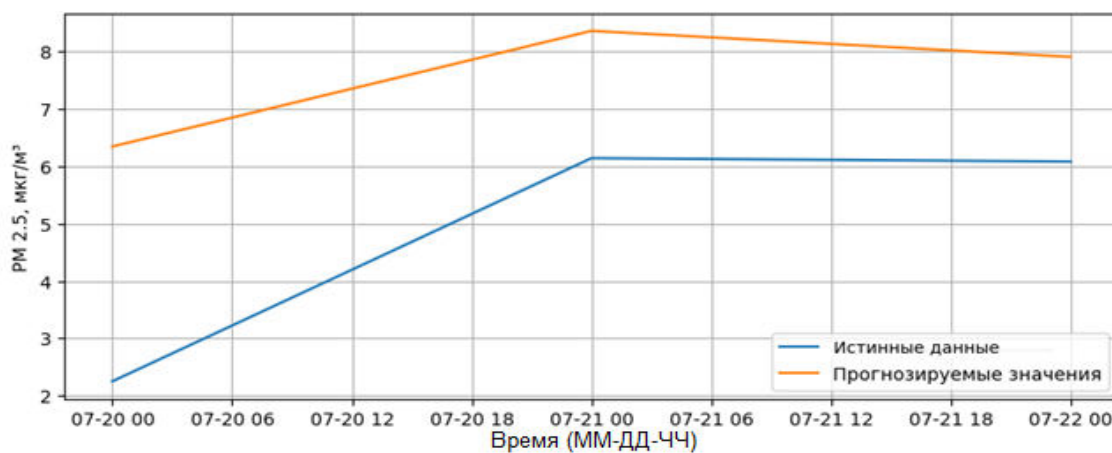


Рисунок 40 – Прогнозирование величины $PM_{2.5}$ моделью градиентного бустинга.
Срез летнего периода: с 20 по 22 июля 2023 г. ($MSE = 8,33$; $MAPE = 0,83$).

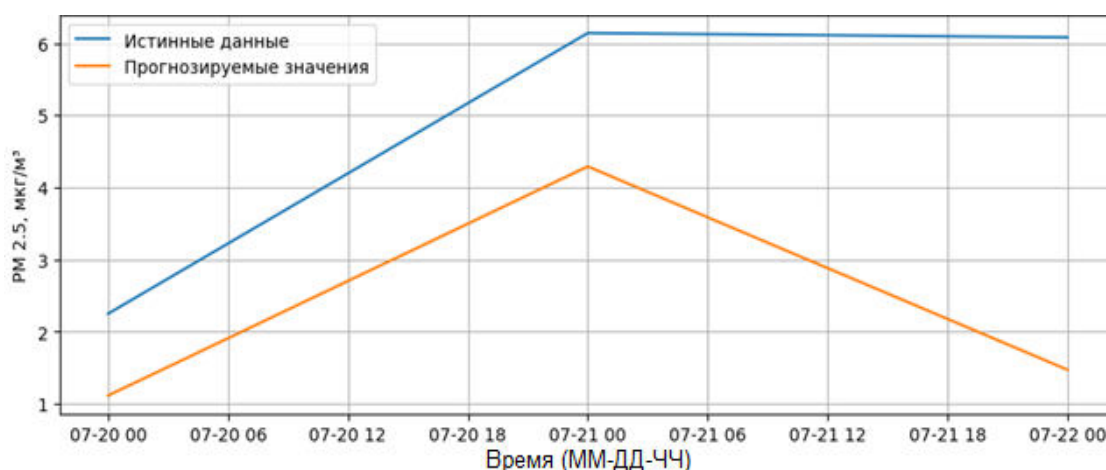


Рисунок 41 – Прогнозирование величины $PM_{2.5}$ моделью временных рядов ARIMAX. Срез летнего периода: с 20 по 22 июля 2023 г. (MSE = 8,64; MAPE = 0,52).

Таким образом, при исследовании различных моделей машинного обучения для прогнозирования величины концентрации $PM_{2.5}$ было определено, что лучшее качество показывают модели временных рядов ARIMAX, что позволяет рекомендовать их использование для прогнозирования уровня загрязнения воздуха твердыми взвешенными частицами с достаточно высоким уровнем точности.

Список цитируемой в разделе литературы:

1. Санитарные правила и нормы СанПиН 1.2.3685-21 "Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания". <https://docs.cntd.ru/document/573500115> (дата обращения 21.12.2023).
2. Liu H. Intelligent modeling strategies for forecasting air quality time series: A review // Applied Soft Computing Journal. 2021. Vol. 102. P. 106957.
3. Volodko O., Yakubailik O., Lapo T., Dergunov A. Influences of meteorological conditions in $PM_{2.5}$ levels in Krasnoyarsk city atmosphere // E3S Web of Conferences. 2023. Vol. 392. P. 02022.
4. Володько О. С., Буряк Н. А., Полянчикова Д. В., Дергунов А. В. Модели регрессии для прогноза уровня загрязнения атмосферного воздуха города Красноярска // Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов. 2023. P. 226-230.

Основные публикации по разделу:

1. Володько О. С., Лев Н. А., Полянчикова Д. В. Прогнозирование величины концентрации загрязняющих веществ в атмосфере г. Красноярска с помощью моделей машинного обучения // В сборнике: Региональные проблемы дистанционного зондирования Земли. Материалы XI Международной научной конференции. Красноярск, 2024. С. 352-356.

3. ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММ ДЛЯ ПРОБЛЕМНО-ОРИЕНТИРОВАННЫХ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИХ СИСТЕМ И СЕРВИСОВ

3.1. Разработка организационной модели управления безопасностью территорий

Предложена модель информационной поддержки принятия оперативных и стратегических решений по обеспечению безопасности территорий. Построены экземпляры модели для разных видов ситуаций, позволившие обосновать структуру информационных ресурсов комплексного мониторинга и состав сервисов интеллектуальной обработки для формирования решений.

Анализ функционирования органов государственной власти и местного самоуправления показал, что территориальное управление в большинстве ситуаций требует принятия согласованных решений несколькими акторами. Постановку задач, их реализацию и информационное обеспечение в вертикали управления целесообразно рассматривать с точки зрения фрактальности. Организация взаимодействия по горизонтали реализуется с учётом специфики функционирования ведомств, органов управления при достижении общей цели территориального управления [1].

С 1990-х годов информационно-управляющие системы эволюционировали в сложные многофункциональные аппаратно-программные комплексы (АПК), обеспечивающие информационную поддержку принятия решений в разных сферах экономики [2]. Развёртывание АПК «Безопасный город», интегрирующего технологии и информационные ресурсы разных ведомств направлено на повышение эффективности реагирования экстренных служб, а также совершенствование планирование и обоснование превентивных мероприятий [3].

Информационную поддержку управления относят к задачам повышенной сложности [4, 5]. Можно выделить три уровня информативности моделей, реализованных в автоматизированных системах. Модели физических процессов позволяют описать динамику воздействия опасных факторов, оценить возможные последствия их развития, визуализировать ситуацию на карте или в виде комбинированного отчёта.

Ситуационное моделирование в дополнение предоставляет лицу, принимающему решения (ЛПР) дополнительную информацию о рациональных способах реагирования, привлекаемых формированиях и необходимых ресурсах в соответствии с масштабом проявления опасных факторов и степенью уязвимости защищаемых объектов. Такие модели используют формализованные описания ситуаций в виде сценариев или баз знаний [4, 6]. Реализация классических подходов ситуационного моделирования и экспертных систем не получила широкого распространения в РСЧС («Единая государственная система предупрежде-

ния и ликвидации чрезвычайных ситуаций») из-за сложности пополнения и корректировки информационных ресурсов.

Помимо вопросов консолидации больших объёмов данных, которые невозможно собрать в рамках одного ведомства, не решена проблема доступности результатов машинной обработки информации. Как показывает анализ процессов управления, использование систем поддержки принятия решений на уровне ЦУКС и ЕДДС (Центр управления в кризисных ситуациях, Единая дежурно-диспетчерская служба) не достаточно для снятия неопределённостей принятия решений у всех ЛПР и исполнителям мероприятий. Поскольку обеспечение природно-техногенной безопасности основывается на согласованных действиях множества акторов, необходимо исследовать потребность в информации всех элементов РСЧС, распределение задач управления между ними с учётом полномочий, располагаемых ресурсов, взаимодействия, а также ограничений, установленных нормативно-правовыми актами.

В качестве инструмента исследования предложена организационная модель управления. Цели её создания – систематизация информационных и управляющих процессов для организации информационной обеспеченности взаимодействия при решении задач предупреждения и ликвидации ЧС различных масштабов. Повышение качества и оперативности принятия решений обеспечивается за счёт обоснования содержания информации, определение необходимости и достаточности её объёма и способов представления для конкретного ЛПР.

Анализ структур организационных систем исследован в работе М.В. Буйневича (Санкт-Петербургский университет ГПС МЧС России) с соавторами [7]. Показаны перспективы использования методов сетевого моделирования, кластеризации и модулярности, а также графового описания в процессах проектирования информационных систем. На основе анализа функций исполнителей и степени их взаимодействия выполнен расчёт трудозатрат на создание программного обеспечения. Масштабирование описанных решений с уровня отдельной организации до иерархии территориального управления потребует доработки математического аппарата.

Организационная модель управления отражает результаты систематизации всех акторов, участвующих в решении конкретной задачи и связей, иллюстрирующих распределение информации. На основе анализа нормативных и методических документов разработана структура и собраны данные, описывающие процессы решения задачи акторами. Результат – графовое и матричное представления для задачи управления. Определены состав и структура исходных данных, процессы их обработки, формирования решений, а также шаблоны решений.

Информационное отображение иерархии управления есть:

$$A = \langle V(l, w), G(d, m) \rangle (T, H),$$

где $V(l, w)$ – множество элементов (вершин графа), имеющих определённые горизонтальные l и вертикальные w позиции; $G(d, m)$ – связи между вершинами – дуги графа со свойствами, описывающие виды информации d или способы управления m . Переменные T и H определяют *размерность* модели: T – задачи управления; H – виды ситуаций.

Элементами V являются органы территориального управления, подразделения МЧС России и других ведомств, входящих в РСЧС¹. Иерархия разделена на уровни управления l : федеральный l_0 , региональный l_1 , муниципальный l_2 ; объектовый l_3 . Для удобства противопожарные, аварийно-спасательные и другие формирования, реагирующие на ЧС и выполняющие превентивные мероприятия, обозначим l_4 . Согласно теории организационных систем, элементы группируются по стратам, слоям и эшелонам в зависимости от функций [8,9]. В древовидной структуре, когда элемент нижестоящего уровня подчинён только одному элементу вышестоящего уровня, характеризуется сильными связями [1]. По такому принципу построена федеральная противопожарная служба. Множественная подчинённость является слабой связи. В этом случае элементы обладают определённой свободой принятия решений. «Горизонтальный» информационный обмен между V одного уровня L основывается на слабых связях, а по вертикали управления связи G представлены разными видами.

Особенность A применительно к l заключается в возможности снижения размерности модели за счёт рассмотрения информационной обеспеченности и функционала одного-двух уровней с единым метаописанием свойств элементов для общих задач управления [10]. Метаописания V могут быть дополнены уникальными характеристиками территорий, объектов и формирований, важных для моделирования задач управления. Например, для Арктической зоны основные ресурсы для ликвидации ЧС сосредоточены на объектовом уровне l_3 , что предполагает их привлечение в качестве основного актора при реагировании на события за пределами промышленной зоны.

Задачи управления классифицированы как $T = \{t_1, t_2, t_3, t_4\}$, где t_1 – сбор информации, выявление и реагирование ситуаций, требующих управленческих действий (повседневное управление); t_2 – реагирование на ЧС и опасные ситуации (оперативное управление); t_3 – планирование и реализация долгосрочных мероприятий по снижению рисков (стра-

¹Постановление Правительства РФ от 30.12.2003 N 794 (ред. от 16.02.2023) «О единой государственной системе предупреждения и ликвидации чрезвычайных ситуаций»

тегическое управление); t_4 – управление в циклических ЧС. К t_1 относится формирование информационных ресурсов, актуализация планов действий для всех V . Для поддержки стратегического управления t_3 , t_4 используются аналитические сервисы. Элементы t_x являются множествами, поскольку включают, как правило, проведение нескольких мероприятий [11, 12]. В информационно-управляющих системах планирование, реализация и контроль результативности мероприятия представляется как самостоятельная задача управления.

Множество H определено на основе классификатора МЧС России и справочников выездов экстренных служб. Для моделирования взаимодействия целесообразно рассмотрение 5-7 «базовых» видов ситуаций – наиболее характерных видов ЧС для конкретного региона. К ним относятся уникальные события с большими ущербами, длительностью, необходимостью управления значительной группировкой сил и средств. Организационные модели управления, учитывающие специфику территорий Сибири, построены для циклических рисков: весенних паводков, природных пожаров, аварий систем ЖКХ.

На рисунке 42 представлен фрагмент иерархии территориальной подсистемы РСЧС Красноярского края, H = «природные пожары».

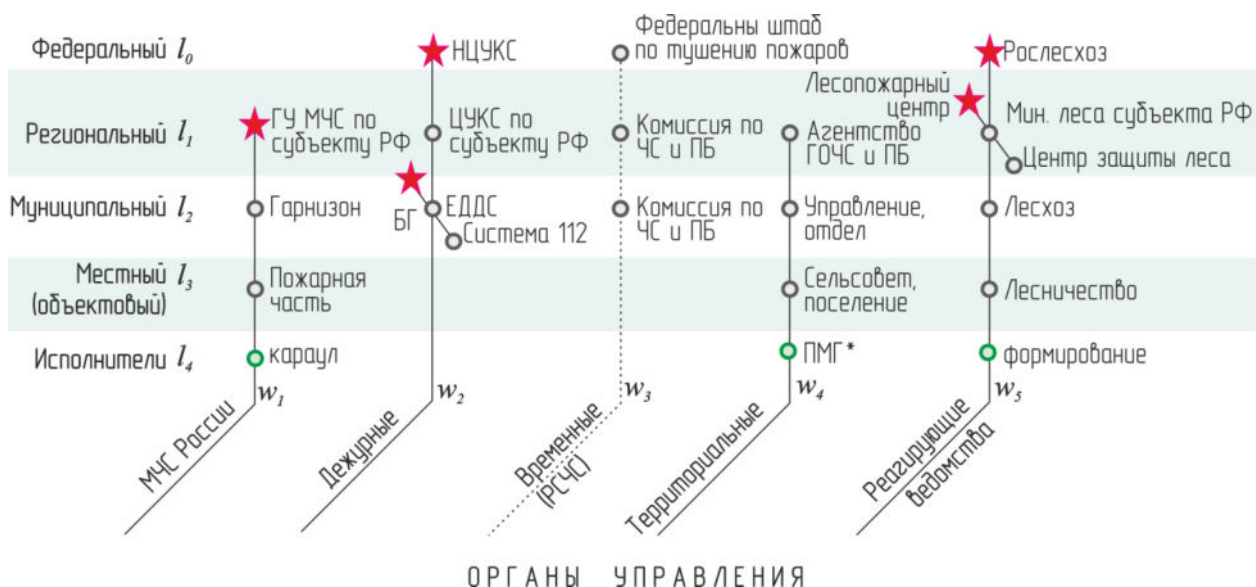


Рисунок 42 – Графическое представление модели организационного управления.

*ПМГ – патрульно-маневренные группы

Звёздочками показаны элементы РСЧС, эксплуатирующие и определяющие развитие тематических информационно-управляющих систем. Среди них: «Термические точки» и «КАСКАД» – Национальный ЦУКС; Информационная система дистанционного мониторинга (ИСДМ)» – Рослесхоз; Интегратор систем видеонаблюдения «Лесоохранитель» – АПК «Безопасный город»; система учёта работ по тушению – Лесопожарный центр Крас-

ноярского края. Структурно-функциональный анализ показал частичное дублирование функций и расхождения в порядке учёта пожаров и термических точек, что затрудняет анализ реагирования подразделений для сезонной корректировки планов действий. Процессы формирования решений при получении информации у большинства акторов не стандартизованы, что затрудняет их автоматизацию.

Ранжирование связей элементов позволило выявить узкие места в управлении и степени информационной поддержки. Например, дефицит исходных данных снижает востребованность моделей физических процессов и ситуационных моделей, предназначенных для обоснования различных тактик тушения природных пожаров, оценки вероятности перехода огня на защищаемые объекты.

Граф, представленный на рисунке 42, преобразуется в матрицу, позволяющую проанализировать информационное взаимодействие акторов, принимающих решения и исполнителей, реализующих оперативные и превентивные мероприятия. Диагональ матрицы составляют информационные процессы, выполняемые в V , а остальные элементы – состав и структуру передаваемой и принимаемой информации.

Несмотря на различия мероприятий, выполняемых разными ведомствами, информационную поддержку взаимодействия в РСЧС можно разделить на два класса: планирование действий на основе *ситуационного* моделирования; распределение ресурсов с учётом оценок рисков возникновения и масштабности потерь, получаемых с использованием *аналитического* моделирования. При этом триггером процессов решения задач t_2 является идентификация опасностей и угроз (погодных условий, грозовой активности и др.) [13]. При этом эффективность мониторинга определяется чувствительностью к слабым сигналам, позволяющим заблаговременно идентифицировать опасности и угрозы. Стратегические решения t_3 принимаются с учётом набора показателей, включающих характеристики территорий, статистику событий, реестры о проведённых мероприятиях с распределением по видам и объёмам [14].

Приведём пример экземпляра организационной модели для задачи обеспечения пожарной безопасности сельских поселений. В соответствии с Постановлением Правительства Красноярского края от 13.12.2019 № 703-п «Об утверждении Порядка предоставления и распределения субсидий бюджетам муниципальных образований Красноярского края на обеспечение первичных мер пожарной безопасности» – информационные ресурсы включают показатели паспортов безопасности территорий (наличие минерализованных полос, водоисточников, удалённость пожарных формирований различного статуса); данные о тушении пожаров (время обнаружения, прибытия формирования, локализации, ликвидации,

использование техники и огнетушащих средств); предупредительные мероприятия (информирование населения, оснащение формирований, контроль проводников горения и др.).

Структура информационных ресурсов для оценки состояния защищаемых объектов, планирования предупредительных мероприятий и мониторинга их выполнения показана на рисунке 43. Это альтернатива используемым противопожарным паспортам территорий. Описание технологий создания цифрового паспорта территорий приведено в [15].

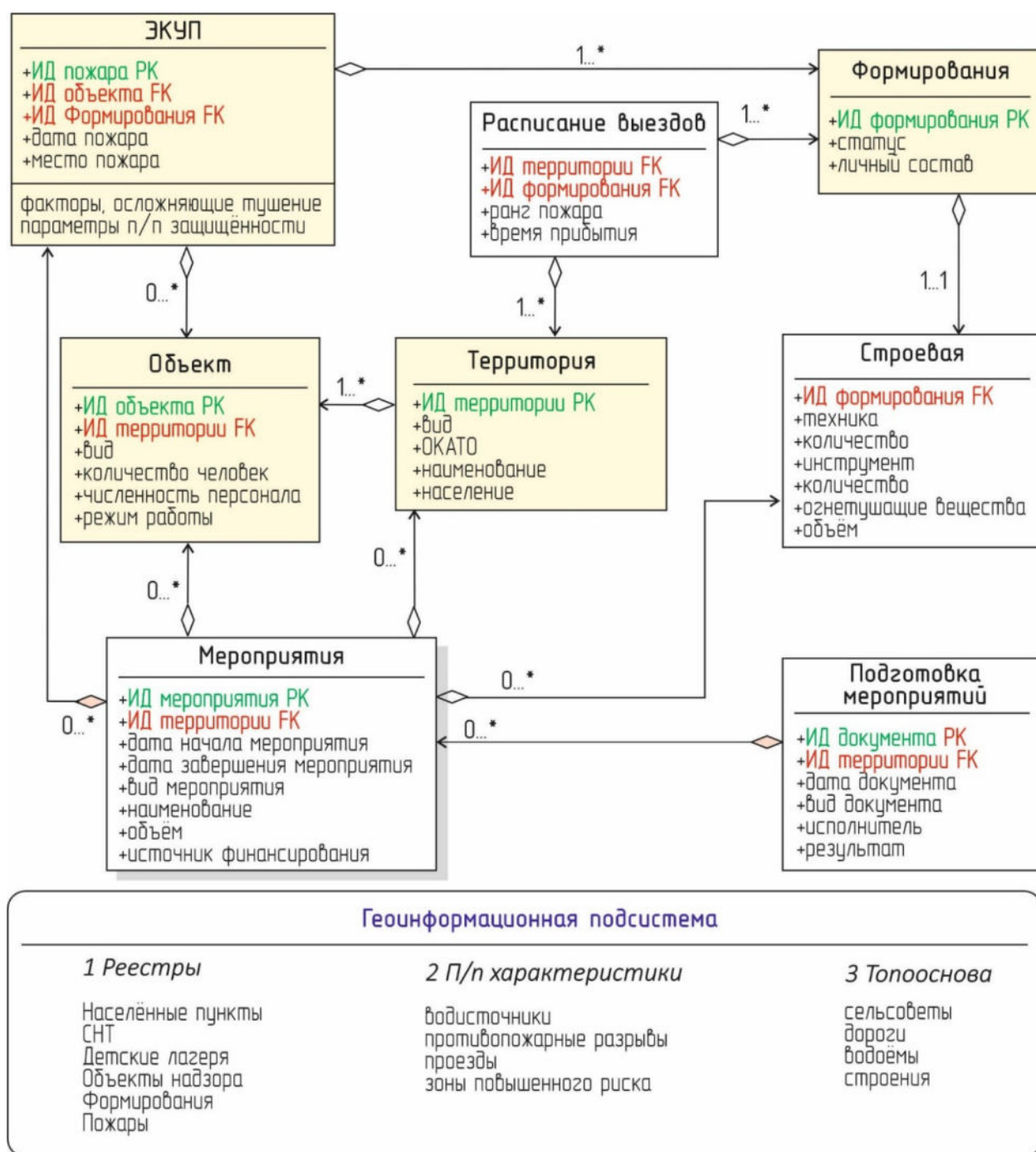


Рисунок 43 – Структура противопожарных паспортов сельских территорий

Формализованные данные, аналитическое и картографическое моделирование предназначены для автоматизации контроля подготовки к весенне-летнему пожароопасному периоду, надзора за выполнением требований Правил противопожарного режима в Российской Федерации, реализации информационной поддержки стратегических мероприятий в долгосрочной перспективе. Использование процедур агрегации применяется для масштабирования данных при решении задач управления на уровнях субъекта РФ и выше по иерархии РСЧС (свойство фрактальности).

Множество логических связей между сущностями позволяет строить разные аналитические модели для оценки защищённости территорий, эффективности управления, стратегического планирования мероприятий по обеспечению пожарной безопасности и др. Связи таблиц определены ключевыми полями РК и FK. Основная таблица содержит первичный ключ РК, зависимая вторичный – FK. Они обеспечивают логическую целостность данных.

Результаты решения задач предупреждения отражены в таблице «*Мероприятия*». Накопление и обработка достаточного объёма данных о реализованных и спланированных мероприятиях позволит сформировать базу знаний рекомендательной подсистемы, применять методы поиска по аналогиям, экстраполяции, статистической обработки и др.

Жёлтый фон (Рис. 40) имеют реестры, ведущиеся независимо. Они имеют пространственную привязку, что позволяет оперативно визуализировать все изменения.

Атрибуты событий в электронных карточках учёта пожаров (ЭКУП) разделены на три группы:

- описание объекта пожара (дата, адрес, вид строения). Данные по возгоранию автомобильного транспорта на межселенной территории рассматриваются только в случае решения задачи предупреждения пожаров, возникающих в результате ДТП.
- факторы, *осложняющие* тушение пожара. Помимо специального справочника к ним относят удалённость водоисточников, загромождение проездов, наличие взрывоопасных и легко воспламеняющихся материалов/веществ. Такие факторы идентифицируются при анализе динамики пожара [16]. Агрегированный перечень факторов используется для обоснования планов мероприятий;
- *барьеры защиты*, способствующие эффективной ликвидации пожара. Примеры: наличие сигнализации, системы пожаротушения; тушение с постановкой на водоисточник. Связь с таблицей «Мероприятия» иллюстрирует изменение процесса реагирования на пожар в результате заблаговременных мер. Например, результат участия добровольцев отражается в значении поля «Руководитель тушения пожара» и косвенно – в снижении времени реагирования.

Реестр *территорий* содержит данные о населённых пунктах, СНТ (садовое некоммерческое товарищество), детских оздоровительных лагерях. Поле «население» содержит данные о максимальном количестве людей, пребывающих на территории постоянно или временно.

Реестр *объектов* содержит данные об объектах надзора ГПН (Государственный пожарный надзор МЧС России). Это промышленные предприятия, места хранения горючих материалов, объекты с массовым пребыванием людей. Для исключения дублирования целесообразна синхронизация с программным обеспечением «Электронный инспектор» и порталом контрольно-надзорной деятельности РФ. Наличие ссылок на электронные карточки пожаротушения позволяет решать задачи оперативного характера.

Реестр *формирований* включает данные всех подразделений – пожарных частей, добровольных пожарных команд и дружин, постов пожарной охраны. При решении задач предупреждения других видов риска, в реестр необходимо включить все подразделения территориальной системы РСЧС не зависимо от их ведомственной принадлежности.

В *суточную строевую записку* необходимо включить помимо пожарных частей информацию о всех противопожарных формированиях. Данные о виде техники, инструмента и огнетушащих веществах, хранятся в отдельных справочниках, не показанных на схеме. Время прибытия формирований «привязано» не к территории, а к расписанию выездов. Совместный анализ данной таблицы и ЭКУП позволяет выявить «нестыковки» в планировании оперативных мероприятий.

Данные о *подготовке мероприятий* используются при оценивании эффективности деятельности органов местного самоуправления. Для построения аналитических моделей целесообразна организация информационного обмена с системами документооборота МЧС России и администрации субъектов РФ.

Для *картографического отображения* и анализа противопожарных паспортов в ГИС добавляются тематические слои водоисточников, зон повышенного риска, противопожарных преград и др. Состояние проездов хранится в атрибутивной информации слоёв «дороги» и «улицы». К зонам повышенного риска относятся разрывы минерализованных полос; участки застройки, удалённые от водоисточников на расстояние, превышающее нормативное; а также строения, непосредственно примыкающие к лесу и бесхозные земельные участки. Расстояние между формированиями и защищаемыми территориями вычисляется автоматически. В СНТ полномасштабная оцифровка строений не проводилась и для ориентировки целесообразно отображать данные о проездах и границах участков.

Проведённые мероприятия изменяют атрибутивную информацию тематических слоёв. Поскольку для работы с информационными ресурсами используется хранилище

данных, все корректировки фиксируются в служебных полях «дата изменения записи», «основание для изменения записи». Таким образом осуществляется двойной контроль, а мероприятия отображаются на карте.

Наличие конкретных данных о плановых и фактических показателях позволяет ранжировать защищаемые объекты уровню рисков, защищённости по степени готовности. Качество управления возможно повысить за счёт анализа обратных связей – сопоставления процессов подготовки к ЧС с каталогом событий, позволяющих выявить слабые места и тиражировать успешный опыт по защите территорий.

Алгоритм аналитического моделирования представлен на рисунке 44.

Описание шагов:

1. Определение масштаба задачи. Для уровня субъекта РФ l_1 ранжируются муниципальные образования, причём городские округа и сельские территории анализируются отдельно. Если решения принимаются на уровне муниципалитета l_2 , то виды, приоритеты и объёмы мероприятий планируются с учётом различий поселений и объектов.
2. В аналитическую модель загружаются ЭКУП, формализованные паспорта территорий, реестр формирований и проведённые мероприятия. Далее инициируется цикл с анализом и обоснованием каждого вида мероприятий. Их перечень сформирован на основе нормативных документов².
3. Формирование числового вектора признаков на основе загруженных данных. Отдельная процедура нормализующая множественность представлений для ЭКУП и характеристик формирований с единичными значениями паспортов территорий (объектов). Использование аналитических методов обработки многомерных данных позволяет более обоснованно ранжировать исследуемые территории, по сравнению с «традиционными» способами, когда учитываются 1-3 признака. Норматив подушевого финансирования мероприятий с коэффициентами отдалённости территорий, использующийся в настоящее время, слишком груб для принятия обоснованных решений. Инструментарий аналитического моделирования позволяет исследовать вклад отдельных компонентов вектора признаков на «рейтинг» исследуемых территорий. Например, включение величин ущербов от пожаров позволяет сконцентрировать внимание лиц, принимающих решения, на поселениях с наибольшими потерями.

² Постановление Правительства Красноярского края от 03.02.2023 № 97-п «Об утверждении распределения иных межбюджетных трансфертов бюджетам муниципальных образований Красноярского края на обеспечение первичных мер пожарной безопасности». URL: <http://www.krskstate.ru/docs/0/doc/93384/yes>

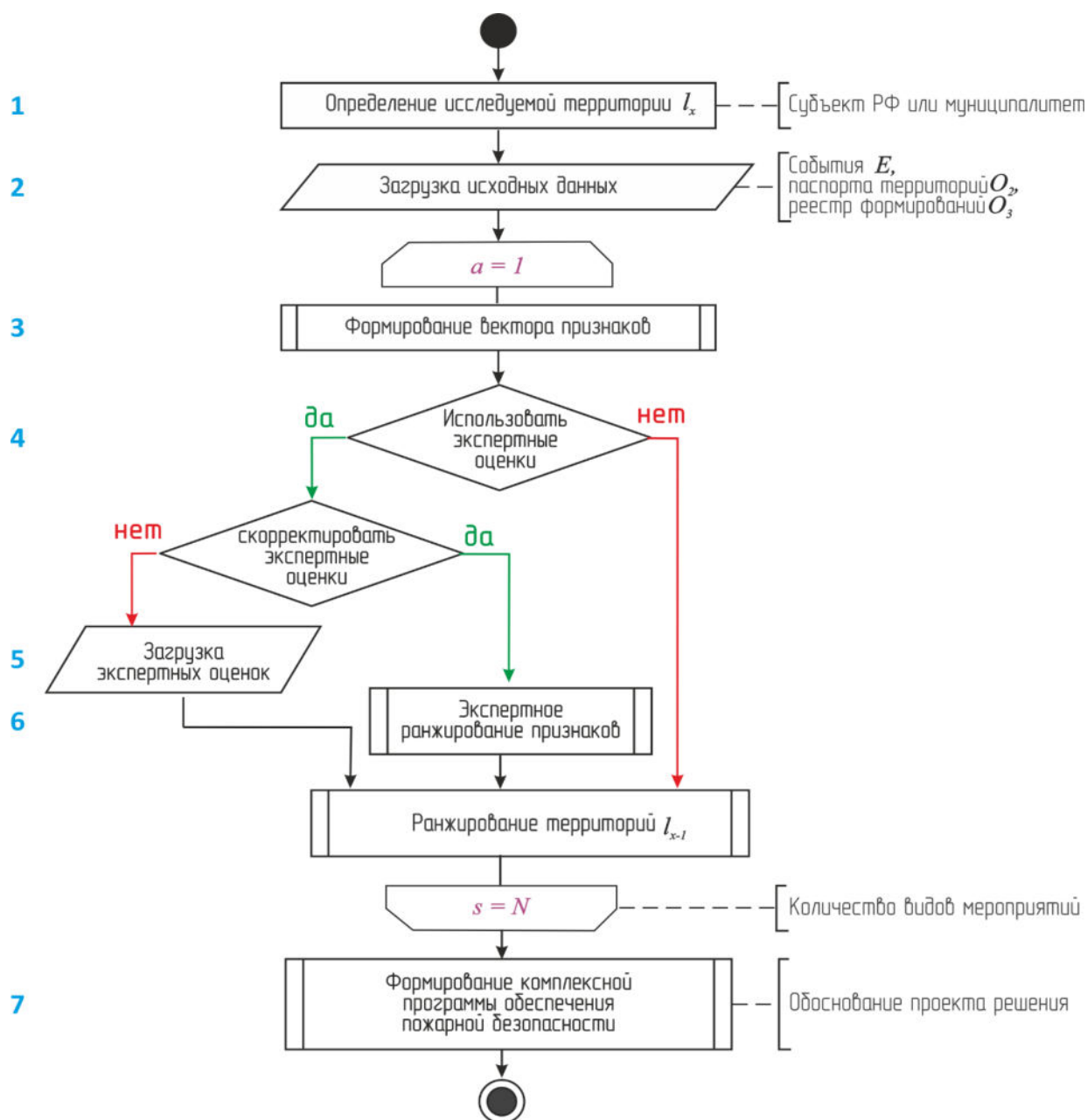


Рисунок 44 – Алгоритм формирования решений по обеспечению первичных мер пожарной безопасности.

4. Кроме работы с ненормированными данными возможно применение экспертных оценок для исходных показателей. Например, при оценке потерь в ЭКУП один погибший в пожаре приравнивается к пяти травмированным либо миллиону рублей ущерба. Также возможности пожарной части и эквивалентны 5 добровольным формированиям. Подобные допущения сильно субъективны, однако позволяют снизить размерность и упростить понимание исходного вектора признаков пожарной безопасности территории.
- 5-6. Предусмотрены функции сохранения и загрузки в аналитическую модель экспертных оценок. Их значения и оправданность использования можно обосновать практикой применения метода в работе органов управления.

7. После перебора всех видов мероприятий результаты ранжирования сопоставляются с имеющимся объёмом ресурсов. Комплексная программа должна состоять из двух разделов: перечень приоритетных мероприятий на бюджетный период, для которых разрабатываются критерии и способы контроля; перечень перспективных мероприятий для достижения приемлемых уровней безопасности на долгосрочный период. Последний может быть пересмотрен с учётом эффекта приоритетных мер.

Предложенный алгоритм позволяет работать с вектором признаков любой размерности. Состав показателей и экспертные оценки могут изменяться в допустимых пределах, с сохранением преемственности решений – важнейший этап подготовки и исследования данных. Ошибки ранжирования территорий возможно из-за разной достоверности показателей, их частичного отсутствия и других причин. Качество и обоснованность решений будет обеспечено за счёт сбора большого количества показателей (в том числе взаимосвязанных), гибкого инструментария аналитического моделирования, наличия системы тестов, транспарентности рассуждений для всех участников управления. Распределённая архитектура с возможностью поиска аналогичных решений и их повторного использования позволит итерационно повысить эффективность действий по обеспечению пожарной безопасности.

Список цитируемой в разделе литературы:

1. Новиков Д.А. Теория управления организационными системами. – М: МПСИ, 2005. – 584 с.
2. Бычков И.В., Ружников Г.М., Хмельнов А.Е. и др. Инфраструктура информационных ресурсов и технологии создания информационно-аналитических систем территориального управления. Новосибирск: Изд-во СО РАН, 2016. 240 с.
3. Нехорошев С.Н. и др. Отчёт о научно-исследовательской работе «Исследования в области построения и развития аппаратно-программного комплекса технических средств «Безопасный город» (заключительный). – М. ВНИИ ГОЧС (ФЦ), 2023. 414 с.
4. Кульба В.В., Кононов Д.А., Косяченко С.А., Шубин А.Н. Методы формирования сценариев в развитии социально-экономических систем. – М.: СИНТЕГ, 2004. 296 с.
5. Дранко О.И., Новиков Д.А., Райков А.Н., Чернов И.В. Управление развитием региона. Моделирование возможностей. – М.: ЛЕНАРД, 2023. 452 с.
6. Ноженкова Л.Ф., Ничепорчук В.В. Технологии комплексной поддержки управления природно-техногенной безопасностью // Вычислительные технологии. 2023. Т. 28. № 4. С. 109-121. DOI:10.25743/ICT. 2023.28.4.009.

7. Наумов В.Н., Наумов П.В., Буйневич М.В. Анализ структур организационных систем методами сетевого моделирования // Информатизация и связь. 2022. № 3. С. 52-58. DOI:10.34219/2078-8320-2022-13-3-52-58
8. Волкова В.Н., Денисов А.А. Теория систем и системный анализ. М.: Юрайт, 2014. 616 с.
9. Мессарович М., Тако И., Такахара Д. Теория иерархических многоуровневых систем. М.: Мир, 1973. 344 с.
10. Шокин Ю.И., Москвичев В.В., Ноженкова Л.Ф., Ничепорчук В.В. Кризисные базы данных для управления безопасностью территорий // Вычислительные технологии. 2011. Т.16. №6. С.115-126.
11. Блюмин А.М., Калянов Г.Н. Проектирование информационных систем. Уч. пособ. М.: Горячая линия-Телеком, 2023. 312 с.
12. Ничипорчук В.М., Увайсов Р.И., Олтян И.Ю. Метод процессно-ориентированной организации знаний реагирования на чрезвычайные ситуации // Информатизация образования и науки. 2020. № 1 (45). С. 128-144.
13. Nicheporchuk V.V. Method of using hazard criteria for identifying hazardous situations // CEUR Workshop Proceedings (CEUR-WS.org) Vol. 2534. ISSN 1613-0073. Proceedings of the All-Russian Conference "Spatial Data Processing for Monitoring of Natural and Anthropogenic Processes" (SDM-2019) P. 427-433.
14. Ничепорчук В.В., Пенькова Т.Г. Система аналитических показателей для стратегического контроля природно-техногенной безопасности территорий // Проблемы анализа риска, 2018. Т.15. № 1. С. 34-41.
15. Москвичёв В.В., Ничепорчук В.В., Потапов В.П., Тасейко О.В. Цифровой паспорт безопасности территории. Вычислительные технологии. 2021; 26(6): 110-132. DOI: 10.25743/ICT.2021.26.6.008
16. Батуро А.Н., Ничепорчук В.В., Бутузов С.Ю., Гилёк С.А. Управление пожарной обстановкой на основе риск-ориентированного подхода // Сибирский пожарно-спасательный вестник. 2023. № 1. С. 67-80. DOI: 10.34987/vestnik.sibpsa.2023.29.76.003

Основные публикации по разделу:

1. Ничепорчук В.В., Постникова У.С., Тасейко О.В. Проектирование интеллектуальной системы управления безопасностью территорий // Онтология проектирования. 2024. Т. 12, № 1. С. 42-54. DOI:10.18287/2223-9537-2024-14-1-42-54 [RSCI, Белый список - 4]
2. Ничепорчук В.В. Исследование информационной обеспеченности взаимодействия в РСЧС // Технологии гражданской безопасности, 2024. Т. 21. № 4(82). С. 43-47. [ВАК]

3.2. Технология применения математического моделирования движения людей при пожаре для оценки обеспечения безопасных условий эвакуации в музеях

Из логики сущности понятия «здание» («сооружение») и его функционального назначения следует, что существует иерархия средств, применяемых в проектировании и строительстве и, как следствие, образуется иерархия средств, с помощью которых достигается решение задачи по обеспечению безопасных условий эвакуации при пожаре, Таблица 14.

Таблица 14. Иерархия средств, применение которых направленно на защиту людей от воздействия опасных факторов пожара и (или) ограничение последствий их воздействия.

Уровень	Средство / организационно-техническое мероприятие		
1	Объёмно-планировочное (ОПР) и конструктивное (КР) решения (мобильные зоны безопасности) Количество людей Горючая нагрузка (количество, состав)		
2	Обнаружение пожара, оповещение	Время начала эвакуации	Управление эвакуацией (маршрут движения)
3	Ограничение единовременного количества людей в здании (сооружении)	Использование индивидуальных средств защиты	Ограничение количества (площади разлива) горючего
4	Изменение геометрических параметров ОПР и состава КР (технические решения)	Повышение огнестойкости строительных материалов (конструкций)	Снижение пожарной опасности строительных материалов
5	Зоны безопасности (стационарные)	Противодымная вентиляция	Пожаротушение, сдерживание водяными завесами

Объёмно-планировочное (ОПР) и конструктивное (КР) решения разрабатываются исходя из функционального назначения здания, в результате чего объект насыщается горючими материалами, определяется наполненность людьми (включая их характеристики мобильности). Эти же средства применяются в обеспечении пожарной безопасности. Поэтому «фундаментом» системы обеспечения пожарной безопасности в части, направленной на обеспечение безопасных условий эвакуации, являются средства первого уровня иерархии: объёмно-планировочное решение в совокупности с конструктивным решением, обеспечивающим геометрическую неизменяемость ограждающих конструкций, как минимум, в период эвакуации, с учётом функционального назначения здания (количество людей, горючая

нагрузку). Для обеспечения безопасного исхода необходимо своевременное обнаружение пожара и оповещение посетителей здания о действиях в ситуации.

Ко второму уровню отнесены организационно-технические мероприятия, неразрывно сопутствующие средствам первого уровня. Время начала эвакуации и/или маршрут могут играть ключевую роль в обеспечении безопасных условий в процессе эвакуации. Следует оговориться, что не всегда пожар сопровождается эвакуацией. Существуют такие здания, что каждое помещение, или отдельные помещения, являются по своему исполнению зоной безопасности – например, операционные, реанимационные, комнаты размещения людей с ограничением двигательной активности, групповые ячейки детских садов. В случае эвакуации необходимо управление для обеспечения движения людей безопасным маршрутом (в конкретной ситуации) из здания и/или безопасную зону.

К 3-му уровню отнесены организационно-технические мероприятия, необходимость которых выявляется в ходе проверки эффективности средств первых двух уровней. И далее, по мере необходимости и применимости, можно рассматривать средства 4-го и 5-го уровней.

Таким образом, выбор конкретных средств опирается на естественную иерархию составляющих частей здания (сооружения) и организационно-технических средств.

Сформулируем задачу определения эффективного состава средств противопожарной защиты, направленной на обеспечение безопасных условий эвакуации.

Набор средств, направленных на обеспечение безопасных условий эвакуации, выразим набором параметров $P = \{p_i\}, i = \overline{1, N_P}$.

Пусть $A = \{a_l\}, l = \overline{1, L}$, – множество условий (критериев), которые должны одновременно выполняться на участке контроля, L – количество необходимых критериев для объекта (в первую очередь – своевременность и бесперпятственность).

Определение 1. Набор средств P^0 является *эффективным*, если на каждом участке контроля выполняются заданные критерии: $P^0 = \{p_1, \dots, p_{N_P}\}: A \quad \forall i \quad \forall s, \quad i = \overline{1, N_K}, s = \overline{1, N}, N_K$ – количество участков контроля (контрольных точек) наблюдаемых (оцениваемых) характеристик, N – количество сценариев.

Утверждение 1. Для каждого здания существует эффективный набор средств противопожарной защиты: $\exists P^0 \supseteq P: P^0 = \{p_1, \dots, p_{N_P}\}: A \quad \forall i \quad \forall s, \quad i = \overline{1, N_K}, s = \overline{1, N}$.

Таким образом, задача определения эффективного состава элементов противопожарной защиты для здания сводится к преобразованию F множества средств P в множество средств P^0 , удовлетворяющих множеству заданных условий A для каждого участка здания для всех рассмотренных сценариев:

$$F: P \rightarrow P^0.$$

Утверждение 2. P^0 – необходимое и достаточное решение, если состав средств формируется согласно порядку уровней, Таблица 14.

Необходимым является набор средств, без которого здание как таковое не может существовать. Достаточный набор обеспечивает требуемые безопасные условия во время эвакуации людей в безопасную зону. При этом состав необходимых средств может являться и достаточным. Минимальность решения вытекает из последовательного добавления средств до достижения условия достаточности, которое выражено в выполнении требуемых критериев для любого участка здания.

Утверждение 3. Преобразование F не единственно.

И тогда выбор конкретного решения переходит в плоскость экономической целесообразности и применимости средств на конкретном объекте.

Алгоритм, реализующий преобразование F – применение моделирования развития пожара и эвакуации для обоснования эффективности элементов противопожарной защиты, направленных на обеспечение безопасных условий эвакуации (далее – Алгоритм), представлен на рисунке 45.

Музеи относятся к объектам повышенной опасности, поскольку являются объектами массового пребывания. В случае пожара для всех без исключения людей должны быть обеспечены безопасные условия эвакуации из здания. Критерием обеспечения безопасной эвакуации является выполнение условий бесперпятственности и своевременности на протяжении всех путей эвакуации.

Музеи имеют выраженные особенности. Как правило, часть посетителей в музеях имеют ограниченные возможности передвижения (МГН – маломобильная группа населения), им требуется помощь персонала в случае экстренной эвакуации. Для музейных комплексов характерна постоянная ротация и обновление посетителей. Чаще всего посетители не знакомы с планировкой здания, расположением эвакуационных выходов, и, именно, сотрудники должны управлять эвакуацией, направляя посетителей безопасным маршрутом. Основным ограничением при возникновении пожара может быть исключение части путей из маршрута эвакуации в силу их небезопасности. В этом случае требуется управление эвакуацией: согласованные действия персонала по распределению потоков людей по безопасным маршрутам и недопущение продолжительных скоплений.

Сохранность музейных ценностей также является важным критерием при проведении экстренных эвакуаций из здания, который достигается проведением досмотровых процедур посетителей по завершении эвакуации.

Проведение тренировочных эвакуаций в музеях очень важно: тренируется персонал, отрабатывается механизм согласованной работы всех задействованных служб музея и внешних служб безопасности.

Для успешного решения задачи обеспечения безопасных условий эвакуации из музея нужно иметь представление о характерной динамике развития пожара при различных местах расположения очага, продолжительности эвакуации. На основе этой информации решаются задачи:

- определение допустимой продолжительности эвакуации, при которой обеспечиваются безопасные условия на маршрутах эвакуации;
- установление предельной численности людей, единовременное пребывание которых в музее будет безопасным в случае пожара или другой чрезвычайной ситуации, требующей эвакуации;
- установление максимальной численности посетителей категории МГН, разрешённой к доступу на площадку единовременно, разработка стратегии их сопровождения из различных помещений музея, учитывая ограниченную скорость передвижения и возможность самостоятельного покидания здания;
- определение наиболее безопасных частей здания на этажах для использования их в качестве зон безопасности для посетителей категории МГН;
- разработка различных маршрутов эвакуации на случай действительного или предположительного блокирования эвакуационных путей и/или выходов за счёт распространения ОФП с учётом применяемых критериев безопасности;
- выявление наиболее проблемных для эвакуации участков здания, в которых образуются скопления людей с критической плотностью, проверка условия беспрепятственности эвакуации: $t_{i,ck} < 360$ с;
- разработка инструкции для персонала, подбор наиболее подходящего персонала на ключевые участки здания для управления эвакуацией, планы проведения тренировочных эвакуаций.

На примере Зимнего дворца Государственного Эрмитажа, одного из крупнейших музейных комплексов в мире, рассмотрено применение компьютерного моделирования эвакуации при пожаре с соответствием с разработанным Алгоритмом в решении задачи оценки эффективности имеющихся средств противопожарной защиты и организации управления эвакуацией из здания в случае пожара.

Применением разработанного Алгоритма показано, что для обеспечения безопасных условий эвакуации из Зимнего Дворца достаточно естественных средств противопожарной защиты при определении допустимого количества единовременно находящихся людей в

здании в количестве 4000 человек. То есть задача обеспечения безопасных условий эвакуации людей решается за счёт следующих средств: объёмно-планировочного и конструктивного решений здания, средства обнаружения пожара, оповещения, своевременного начала эвакуации и управления эвакуацией. Кроме того показано, что в качестве зоны безопасности можно использовать часть здания, удалённую от очага пожара, – динамическую зону безопасности.

Исследование показало, что задача разработки маршрута эвакуации при пожаре в зданиях музеев является многокритериальной. Важным аспектом является факт реализуемости разработанного маршрута – то есть объект может обеспечить маршрутизацию (управление потоками людей) согласно разработанной схеме в случае конкретного очага пожара.

Разработка маршрутов эвакуации при создании базы сценариев управления эвакуацией ВД_{управление эв.} должна сопровождаться соблюдением сформулированных правил, учитывающих особенности музеев:

- единообразие маршрутов эвакуации;
- из одного зала все люди направляются одним маршрутом;
- посетители анфилад (систем залов, разделённых открытыми проемами) направляются в одну сторону;
- выбираются маршруты более длинные для ускорения эвакуации в наиболее уязвимых узлах здания;
- обеспечение безопасности людей первичной задачи обеспечения безопасности музейных ценностей.

Основные публикации по разделу:

1. Кияткина Е.Н., Лобаев И.А., Кирик Е.С., Воронов С.П. Поддержка принятия управленческих решений при применении обязательных требований соразмерных риску причинения вреда // Сибирский пожарно-спасательный вестник. 2024, N2 (33). С. 34-47. DOI: 10.34987/vestnik.sibpsa.2024.84.80.021. [BAK]
2. Gravit, M.; Kirik, E.; Kozupitsa, I.; Solntseva, Y. Fire safety in malls: Curtains for fire hazards localization and safe evacuation; 2024; Construction of Unique Buildings and Structures; 112 Article No 11203. DOI: 10.4123/CUBS.112.3. [RSCI, Белый список - 4]
3. Кирик Е.С., Богданов А.В., Сушкова О.В., Витова Т.Б. Эвакуация при пожаре в музее: подготовка и проведение тренировочных эвакуаций с использованием численного моделирования // Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России. 2024. № 4. (в печати). [BAK]

3.3. Разработка и исследование непараметрических методов проверки гипотез о распределениях многомерных случайных величин большого объёма

Проверка гипотезы о независимости случайных величин является одним из основных этапов системного анализа статистических данных. На её результатах осуществляется синтез эффективных алгоритмов принятия решений. Традиционная методика проверки гипотезы о независимости случайных величин основана на использовании критерия Пирсона, которая содержит трудно формализуемый этап разбиения области значений случайных величин на многомерные интервалы. Предложена методика проверки гипотезы о независимости случайных величин, которая использует непараметрический алгоритм распознавания образов ядерного типа. Её применение позволяет обойти проблему декомпозиции области значений случайных величин на интервалы. Идея подхода состоит в формировании по исходным статистическим данным обучающей выборки для решения двухальтернативной задачи распознавания образов. Каждый класс определяется в предположении независимости либо зависимости случайных величин, что проявляется в различии их законов распределения в классах. В этих условиях появляется возможность замены исходной гипотезы на задачу проверки достоверности отличия вероятностей ошибок распознавания образов в классах. Эффективность предлагаемого подхода подтверждается результатами исследования однозначных и неоднозначных зависимостей между случайными величинами.

Ниже обосновывается методика проверки гипотезы о независимости случайных величин, основанная на непараметрическом алгоритме распознавания образов, предлагается её реализация в условиях больших объёмов статистических данных, приводятся результаты сравнения методики с общепризнанным критерием согласия Пирсона и применения при обработке информации дистанционного зондирования природных объектов.

Методика проверки гипотезы о независимости случайных величин.

Пусть имеется выборка $V = (x^i, i = \overline{1, n})$ объёма n , составленная из независимых наблюдений двухмерной случайной величины $x = (x_1, x_2)$. Выборка V извлекается из генеральных совокупностей, характеризующихся плотностями вероятностей $p(x_1)$ $p(x_2)$ или $p(x_1, x_2)$. Необходимо по статистическим данным V проверить гипотезу

$$H_0 : p(x_1, x_2) \equiv p(x_1) p(x_2)$$

о независимости случайных величин x_1, x_2 .

Для проверки гипотезы H_0 будем решать двухальтернативную задачу распознавания образов. Под классами Ω_1, Ω_2 понимаются области определения плотностей вероятностей $p(x_1) p(x_2), p(x_1, x_2)$. В этих условиях байесовское решающее правило, соответствующее критерию максимального правдоподобия, имеет вид

$$m(x) : \begin{cases} x \in \Omega_1, & \text{если } p(x_1, x_2) < p(x_1) p(x_2) \\ x \in \Omega_2, & \text{если } p(x_1, x_2) > p(x_1) p(x_2) \end{cases}.$$

В отличие от традиционной постановки задачи распознавания образов при синтезе решающего правила $m(x)$ априори отсутствует обучающая выборка, содержащая сведения о принадлежности элементов выборки V к тому или иному классу. Эти сведения должны обнаруживаться в процессе реализации методики проверки гипотезы H_0 , которая основана на выполнении следующих действий.

По выборке V восстановить плотности вероятностей $p(x_1, x_2), p(x_1) p(x_2)$, используя их непараметрические оценки типа Розенблатта – Парзена,

$$\begin{aligned} \bar{p}(x_1, x_2) &= \frac{1}{nc_1c_2} \sum_{i=1}^n \Phi\left(\frac{x_1-x_1^i}{c_1}\right) \Phi\left(\frac{x_2-x_2^i}{c_2}\right), \\ \bar{p}(x_1) \bar{p}(x_2) &= \frac{1}{n^2c_1c_2} \sum_{i=1}^n \sum_{j=1}^n \Phi\left(\frac{x_1-x_1^i}{c_1}\right) \Phi\left(\frac{x_2-x_2^j}{c_2}\right). \end{aligned}$$

В статистиках $\bar{p}(x_1, x_2), \bar{p}(x_1) \bar{p}(x_2)$ ядерные функции $\Phi(u_v)$ удовлетворяют условиям положительности, симметричности и нормированности.

Значения коэффициентов размытости $c_v, v=1, 2$ ядерных функций убывают с ростом объема n выборки статистических данных V . Тогда непараметрическое решающее правило классификации случайных величин $x = (x_1, x_2)$ запишется как

$$\bar{m}(x) : \begin{cases} x \in \Omega_1, & \text{если } \bar{p}(x_1, x_2) < \bar{p}(x_1) \bar{p}(x_2) \\ x \in \Omega_2, & \text{если } \bar{p}(x_1, x_2) > \bar{p}(x_1) \bar{p}(x_2) \end{cases}.$$

Оптимальные коэффициенты размытости ядерных функций решающего правила $\bar{m}(x)$ выбираются на основе анализа аппроксимационных свойств непараметрических оценок плотностей вероятностей $\bar{p}(x_1, x_2), \bar{p}(x_1), \bar{p}(x_2)$ из условия минимума, соответствующих им оценок среднеквадратических отклонений от $p(x_1, x_2), p(x_1), p(x_2)$. Например, для $\bar{p}(x_1)$ подобным критерием является

$$\int_{-\infty}^{\infty} \bar{p}^2(x_1) dx_1 - \frac{2}{n} \sum_{j=1}^n \bar{p}(x_1^j).$$

Определим оценки вероятностей ошибок распознавания образов $\bar{\rho}_1(\bar{c}_1, \bar{c}_2), \bar{\rho}_2(\bar{c}_1, \bar{c}_2)$ решающим правилом $\bar{m}(x)$ по исходным статистическим данным V при оптимальных

коэффициентах размытости $\bar{c}(1) = (\bar{c}_1(1), \bar{c}_2(1))$, $\bar{c}(2) = (\bar{c}_1(2), \bar{c}_2(2))$ ядерных функций статистик $\bar{p}(x_1)$, $\bar{p}(x_2)$, $\bar{p}(x_1, x_2)$ соответственно.

Значения $\bar{\rho}_t(\bar{c}(1), \bar{c}(2))$ вычисляются в режиме «скользящего экзамена» по выборке V в предположении, что её элементы принадлежат классу Ω_t ,

$$\bar{\rho}_t(\bar{c}(1), \bar{c}(2)) = \frac{1}{n} \sum_{j=1}^n 1(\delta(j), \bar{\delta}(j)), \quad t = 1, 2,$$

где $\delta(j) = t$ – указания типа $x^t = (x_1^t, x_2^t) \in \Omega_t$;

$$\bar{\delta}(j) = \begin{cases} t, & \text{если } x^j \in \Omega_t \\ 0, & \text{если } x^j \notin \Omega_t \end{cases}$$

– «решение» алгоритма $\bar{m}(x)$ о принадлежности ситуации x^j к одному из классов Ω_t , $t = 1, 2$.

При вычислении $\bar{\rho}_t(\bar{c}(1), \bar{c}(2))$ в соответствии с методикой «скользящего экзамена» ситуация $x^j = (x_1^j, x_2^j)$ из выборки V , которая подаётся на контроль в алгоритм $\bar{m}(x)$, исключается из процесса формирования статистик $\bar{p}(x_1, x_2)$, $\bar{p}(x_1)$, $\bar{p}(x_2)$.

Индикаторная функция определяется выражением

$$1(\delta(j), \bar{\delta}(j)) = \begin{cases} 0, & \text{если } \delta(j) = \bar{\delta}(j) \\ 1, & \text{если } \delta(j) \neq \bar{\delta}(j) \end{cases}.$$

Обозначим через $\bar{\bar{\rho}}_t$ минимальное значение оценки вероятности ошибки распознавания образов в предположении, что элементы выборки V принадлежат классу Ω_t , $t = 1, 2$. Сравним значения $\bar{\bar{\rho}}_1, \bar{\bar{\rho}}_2$.

Тогда гипотеза H_0 справедлива, если $\bar{\bar{\rho}}_1 < \bar{\bar{\rho}}_2$. В противном случае при $\bar{\bar{\rho}}_2 < \bar{\bar{\rho}}_1$ случайные величины x_1 и x_2 являются зависимыми.

При ограниченном объёме n выборки V возникает задача доверительного оценивания вероятностей ошибок распознавания образов. Для её решения используется традиционная методика доверительного оценивания вероятностей либо критерий Колмогорова – Смирнова.

Например, при использовании критерия Колмогорова – Смирнова отклонение $\bar{D}_{12} = |\bar{\bar{\rho}}_1 - \bar{\bar{\rho}}_2|$ сравнивается с пороговым значением

$$D_\beta = \sqrt{-\ln\left(\frac{\beta}{2}\right)/n}.$$

Здесь β – вероятность (риск) отвергнуть гипотезу $\bar{H}_0: \rho_1(c_1, c_2) = \rho_2(c_1, c_2)$. Если выполняется соотношение $\bar{D}_{12} < D_\beta$, то гипотеза $\bar{H}_0$ справедлива и риск её отвергнуть не превышает значения β . При $\bar{D}_{12} > D_\beta$ гипотеза $\bar{H}_0$ отвергается.

Формирование наборов независимых случайных величин.

Имеется выборка наблюдений $V = (x_v^i, v = \overline{1, k}, i = \overline{1, n})$ объёма n , составленная из статистически независимых наблюдений компонент многомерной случайной величины $x = (x_v, v = \overline{1, k})$. Вид плотности вероятности $p(x)$ априори неизвестен. Необходимо по статистическим данным V , используя предложенный выше критерий проверки гипотез

$$H_{vj} : p(x_v, x_j) \equiv p(x_v) p(x_j)$$

для компонент $x_v, v = \overline{1, k}, x_j, j = \overline{1, k}, v > j$, сформировать наборы независимых случайных величин $x(t) = (x_v, v \in I_t), t = \overline{1, m}$. Количество m наборов компонент случайной величины x неизвестно, а I_t – множество номеров компонент, составляющих набор $x(t)$

Предлагаемая методика основана на выполнении следующих действий.

1. В соответствии с приведёнными выше рекомендациями проверить гипотезы H_{vj} для каждой пары компонент (x_v, x_j) многомерной случайной величины $x = (x_v, v = \overline{1, k})$. Количество таких пар соответствует значению $k(k-1)/2$.
2. По результатам этапа 1 построить информационный граф $G(X, A)$, где X – множество его вершин, соответствующих компонентам случайной величины x , а A – множество ребер. Между двумя вершинами x_v, x_j имеется ребро, если выполняется гипотеза H_{vj} , т.е. компоненты x_v, x_j являются независимыми.
3. Провести анализ информационного графа $G(X, A)$ и определить его полные подграфы $G(X_t, A_t), t = \overline{1, m}$. Каждая пара вершин подграфа $G(X_t, A_t)$ имеет ребро, если компоненты случайной величины x являются независимыми. Обнаружить полные подграфы с использованием алгоритмов разрезания исходного графа, которые основаны на анализе его матрицы смежности. Компоненты $x_v, v \in I_t$, соответствующие вершинам полного подграфа $G(X_t, A_t)$, образуют набор независимых случайных величин.

Модификация методики проверки гипотезы о независимости случайных величин в условиях больших объёмов статистических данных.

При больших объёмах n статистических данных $V = (x_1^i, x_2^i, i = \overline{1, n})$ в предложенной методике используются регрессионные оценки плотностей вероятностей $\bar{p}(x_1, x_2), \bar{p}(x_1), \bar{p}(x_2)$. Эти оценки основаны на сжатии исходной информации, например, $V_1 = (x_1^i, i = \overline{1, n})$ в массив данных $\bar{V}_1 = (\bar{p}_1^j, z^j, j = \overline{1, N})$ путём декомпозиции области

значений x_1 на N интервалов. Здесь z^j – центры интервалов дискретизации значений x_1 , а $\bar{p}_1^j = \bar{P}_1^j / \Delta$ – оценка плотности вероятности в j -ом интервале; Δ – длина интервала дискретизации, $\bar{P}_1^j$ – частота встречаемости значений x_1^i из выборки V_1 в интервале под номером j . Тогда регрессионная оценка плотности вероятности $p(x_1)$ по данным $\bar{V}_1$ имеет вид

$$\bar{p}(x_1) = \frac{1}{c} \sum_{j=1}^N \bar{P}_1^j \Phi\left(\frac{x_1 - z^j}{c}\right).$$

Предлагаемый подход позволяет на порядки сократить объём n исходной статистической информации при оценивании плотностей вероятностей. Особенность статистика типа $\bar{p}(x_1)$ позволяет значительно упростить выбор коэффициентов размытости c ядерных функций в статистике $\bar{p}(x_1)$ из условия минимума критерия

$$\frac{1}{N} \sum_{i=1}^N (\bar{p}_1^i - \bar{p}(x_1^i)).$$

По аналогии осуществляется оценивание плотностей вероятностей $p(x_2)$, $p(x_1, x_2)$. Регрессионные оценки плотностей вероятностей используются при проверке гипотезы о независимости случайных величин в соответствии с предложенной методикой.

Анализ результатов вычислительного эксперимента.

Проведено сравнение эффективности предлагаемой методики проверки гипотезы о независимости двумерных случайных величин и критерия Пирсона в условиях неоднозначных зависимостей при различных объёмах статистических данных. Датчики случайных величин x_1, x_2 формировались на основе равномерного закона распределения x_1 , который использовался при вычислении значений x_2 в виде нелинейных преобразований x_1 . При этом на значения x_2 накладывались помехи с нормальным законом распределения, который имеет нулевое математическое ожидание и среднеквадратическое отклонение σ . Пример значений случайных величин x_1 и x_2 приведён на рисунке 46.

При проверке гипотезы о независимости компонент двумерной случайной величины на основе критерия Пирсона используются результаты оптимального выбора количества интервалов дискретизации

$$N^* = \left(\frac{3}{4} \Delta_1 \Delta_2 \|p(x_1, x_2)\|^2 n \right)^{1/2}.$$

Значение $\|p(x_1, x_2)\|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p^2(x_1, x_2) dx_1 dx_2$, а Δ_v – длина интервала значений случайной величины x_v , $v=1, 2$.

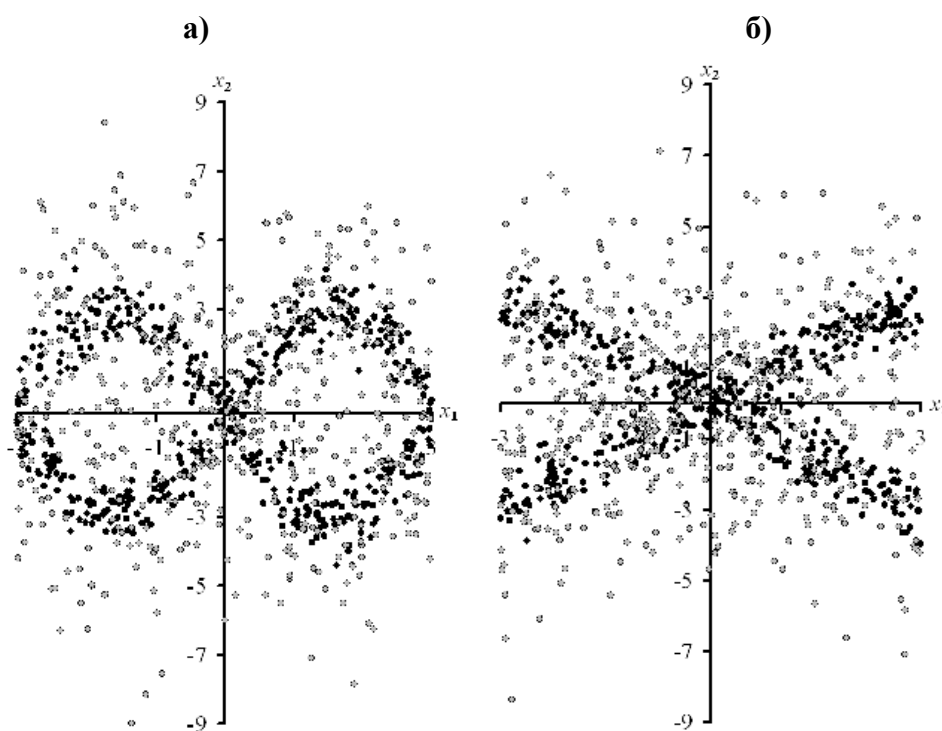


Рисунок 46 – Значения случайных величин x_1, x_2 из выборки исходных статистических данных V при $n=500$ и $\sigma = 0.5$ (темные точки), а при $\sigma = 2$ (серые точки) при использовании зависимостей различной сложности.

По результатам вычислительного эксперимента предлагаемая методика и критерий Пирсона при анализе неоднозначных зависимостей между случайными величинами в условиях малых объёмах статистических данных и средних квадратических отклонений σ помех сопоставимы и безошибочно определяют зависимость случайных величин. Данный вывод не соблюдается при зависимости между случайными величинами (рисунок 45а), когда критерий Пирсона не устанавливает зависимость в условиях $n=100$ и $\sigma \in [0.5; 2]$. С увеличением σ эффективность сравниваемых критериев снижается. Этот факт объясняется особенностями неоднозначных зависимостей и большими значениями σ , когда область определения случайных величин скрывает искомую зависимость. С увеличением объёма n исходных данных эффективность сравниваемых критериев проверки гипотезы о независимости случайных величин повышается. Этот вывод является ожидаемым, так как с ростом n повышаются асимптотические свойства непараметрических оценок плотностей вероятностей и частот встречаемости случайных величин в их двухмерных интервалах. Преимущество предлагаемой методики проверки гипотезы о независимости случайных величин наблюдается при малых значениях σ , ограниченных и больших n . При больших n и σ часто обнаруживается преимущество критерия Пирсона, если используется процедура оптимальной дискретизации области значений двухмерной случайной величины.

Применение предлагаемой методики при анализе данных ДЗЗ.

Объектом исследования являются антропогенные территории (карьер, пригородная застройка) в окрестности города Красноярска. Исходная информация формировалась по фрагментам съёмки спутника Sentinel-2 на 26.08.2021 года (Рис. 47). Использовались спектральные каналы x_j , $j = \overline{1, 9}$, которые характеризуются длинами волн (нанометры): x_1 - (458-523), x_2 - (543-578), x_3 - (650-680), x_4 - (698-713), x_5 - (733-748), x_6 - (773-793), x_7 - (785-899), x_8 - (1565-1655), x_9 - (2100-2280).

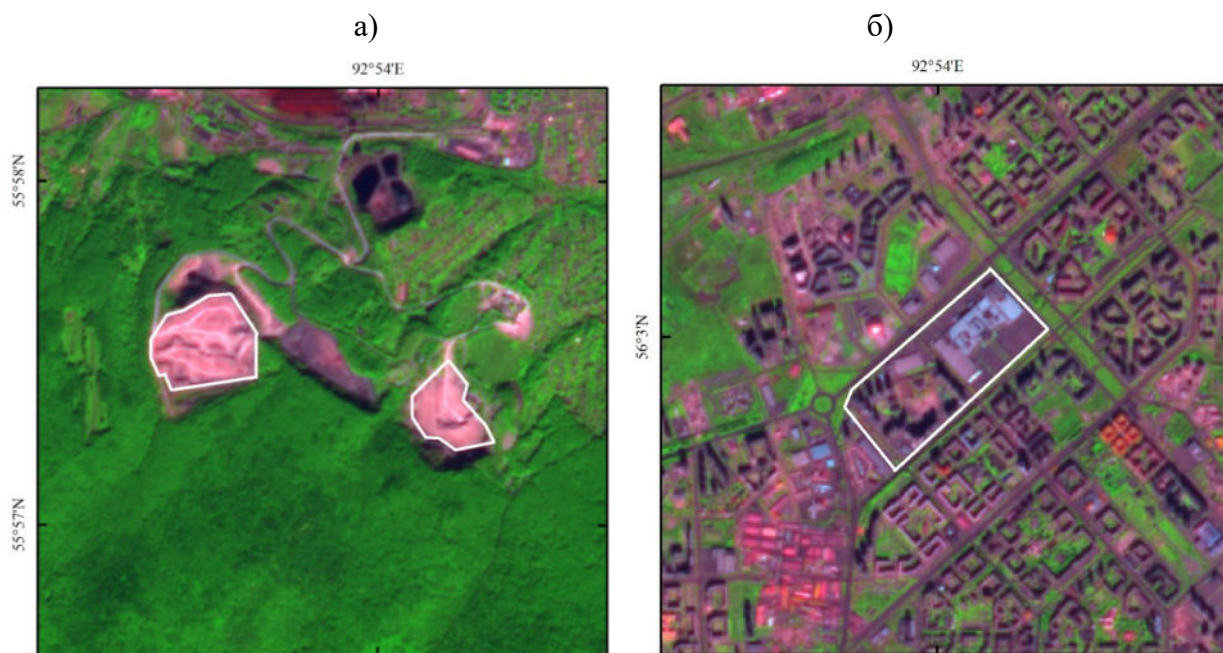


Рисунок 47 – Фрагменты спутниковой съёмки Sentinel-2. Антропогенные территории: карьер (а), пригородная застройка (б).

Предлагаемая методика позволяет формировать пары независимых и зависимых случайных величин, изменяя соотношение между её параметрами. Применение методики позволило обнаружить 31 и 29 пар спектральных признаков с сильной линейной зависимостью соответственно для объектов «карьер», «пригородная застройка». Полученные результаты представлены на рисунке 48.

Дополнительно обнаружены нелинейные зависимости между спектральными признаками для объекта «карьер»

$$(x_1, x_9), (x_1, x_8), (x_1, x_7), (x_1, x_5), (x_1, x_4)$$

и объекта «пригородная застройка» –

$$(x_7, x_9), (x_4, x_9), (x_3, x_9), (x_2, x_9), (x_1, x_9), (x_1, x_8), (x_1, x_7).$$

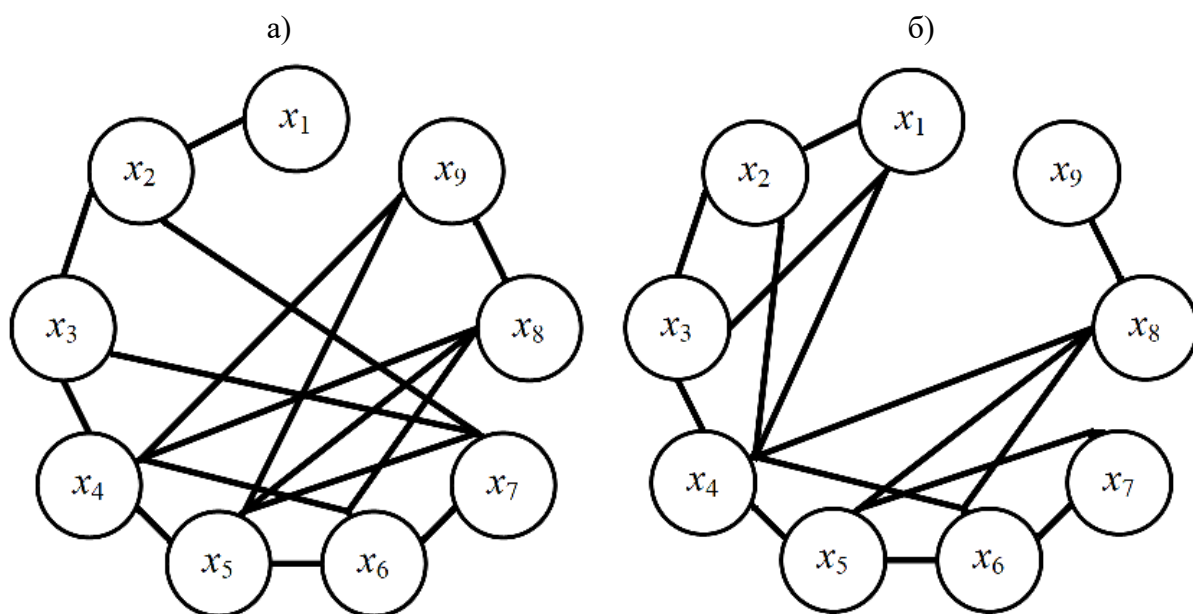


Рисунок 48 – Иллюстрация сильной линейной зависимости между парами спектральных признаков (x_i, x_j) , характеризующихся оценками коэффициентов корреляции больше 0,9; для карьера (а) и пригородная застройка (б).

Полученные результаты являются достоверными для всех пар спектральных признаков, так как соблюдается условие $|\bar{\rho}_1 - \bar{\rho}_2| > D_\beta$ при $D_\beta = 0,029$ и риске $\beta = 0,025$ отвергнуть гипотезу H_0 равенства значений ρ_1, ρ_2 .

Рассмотрена задача обнаружения антропогенных территорий по спектральным данным. Ошибка их распознавания в пространстве спектральных признаков $x = (x_j, j = \overline{1, 9})$ по обучающей выборке $V = (x^i, \sigma(i), i = \overline{1, n})$ равна 0,012, где $n = n_1 + n_2$, $n_1 = 3377$ («карьер», $\sigma(i)=1$), $n_2 = 5049$ («пригородная застройка», $\sigma(i)=2$). При исключении из обучающей выборки, например, спектральных признаков (x_4, x_5) , (x_5, x_6) , (x_4, x_5, x_6) оценки ошибок распознавания образов соответствуют значениям 0,011; 0,01; 0,008. Полученное снижение ошибок распознавания образов не являются достоверными по сравнению с оценкой ошибки в пространстве признаков $x_j, j = \overline{1, 9}$. Однако полученный результат обосновывает возможность сокращения спектральных признаков при синтезе алгоритмов принятия решений и упрощения их оптимизации.

Заключение. Методика проверки гипотезы о независимости пар случайных величин, основанная на использовании непараметрического алгоритма распознавания образов, позволяет обойти проблему дискретизации области значений случайных величин на многомерные интервалы. Эта проблема свойственна общепризнанному критерию Пирсона. Опреде-

лены условия компетентности предлагаемого метода и критерия Пирсона при анализе однозначных и неоднозначных зависимостей между случайными величинами.

С использованием аппарата теории графов предлагаемая методика разработана при формировании наборов независимых случайных величин. Полученные результаты обобщены при проверке гипотезы о независимости случайных величин для больших объемов статистических данных на основе сжатия исходной информации, что позволяет на порядки повысить вычислительную эффективность решаемых задач.

Эффективность предложенной методики подтверждена при анализе данных дистанционного зондирования антропогенных территорий и оценивании их состояний. При наличии набора спектральных признаков, характеризующихся сильной линейной зависимостью между его парами, появляется возможность сокращения количества спектральных признаков при распознавании антропогенных территорий с уменьшением оценки вероятности ошибки их распознавания.

Основные публикации по разделу:

1. Непараметрические методы проверки гипотез о распределениях случайных величин при анализе данных дистанционного зондирования / А.В. Шаруева, А.В. Лапко, В.А. Лапко; отв. ред. О.Э. Якубайлик; Рос. акад. наук, Сиб. отделение, Федерал. исслед. центр Красноярский науч. центр, Ин-т вычислит. моделир., Минобрнауки России, СибГУ им. акад. М.Ф. Решетнева. – Новосибирск: СО РАН, 2024. – 189 с.
2. Лапко А.В., Лапко В.А., Им С.Т. Методика декомпозиции спектральных данных дистанционного зондирования пожаров лесных территорий // Информатика и системы управления. 2024. Т. 81. № 3. С. 112-120. DOI: 10.22250/18142400_2024_81_3_112. [RSCI, Белый Список-3]
3. Тубольцев В.П., Лапко А. В., Лапко В. А. Модифицированный непараметрический алгоритм автоматической классификации статистических данных большого объема и его применение // Искусственный интеллект и принятие решений. 2023. № 4. С. 49-57. DOI: 10.14357/20718594230405 [RSCI, Белый Список - 3]
4. Лапко А.В., Лапко В.А. Сравнение методов оптимизации регрессионной оценки плотности вероятности одномерной случайной величины // Автометрия. 2024. Т. 60. № 3. С. 95-103. DOI: 10.15372/AUT20240309. [WoS, Scopus, RSCI, Белый Список - 2]
5. Тубольцев В.П., Лапко А.В., Лапко В.А. Непараметрический алгоритм автоматической классификации данных дистанционного зондирования // Информатика и ее применения. 2023. Т. 17. № 4. С. 23-31. DOI: 10.14357/19922264230404. [RSCI, Белый Список-3]

ЗАКЛЮЧЕНИЕ

За отчетный период выполнен полный цикл исследований, запланированных по данному проекту. В ходе его выполнения получены результаты, приоритет которых подтверждается публикациями в ведущих российских и зарубежных научных изданиях.

Предложены методы морфометрического анализа агроландшафтов и оценки агроэкологических условий сельскохозяйственных территорий на примере опытно-производственного хозяйства (ОПХ) «Михайловское» ФИЦ КНЦ СО РАН. Основное внимание уделено применению цифровых моделей рельефа, в частности FABDEM, для анализа ключевых морфометрических параметров, таких как крутизна и экспозиция склонов, кривизна, топографический индекс влажности и индекс потенциала плоскостной эрозии. Использование современных технологий ГИС и данных ДЗЗ позволило создать детальную базу данных и серию тематических карт, отражающих особенности рельефа и его влияние на агроэкологические условия. Полученные данные могут быть использованы для прогнозирования эрозионных процессов, оптимизации землепользования и разработки противоэрозионных мероприятий. Предложенная методика применима для дальнейших исследований агроландшафтов в условиях цифровизации сельского хозяйства.

Проведены исследования влияния задымления от лесных пожаров на состояние природной среды, в частности, в районе г. Красноярска. Анализировались данные за август 2021 года, когда дым от масштабных пожаров в Якутии достиг города. Использовались данные метеостанций, спутниковые измерения аэрозольного индекса UVAI, показания датчиков качества воздуха и видеонаблюдение за туманами над рекой Енисей. Установлено, что дымовые аэрозоли увеличивают оптическую толщину атмосферы, снижая количество солнечной радиации, достигающей поверхности. Измерения показали, что это привело к уменьшению температуры верхнего слоя почвы на $5,6^{\circ}\text{C}$ и температуры приземного слоя атмосферы на $3,3^{\circ}\text{C}$. Также зафиксировано формирование адвективных туманов охлаждения над Енисеем при выполнении определенных условий влажности и разности температур. Результаты подтверждают значительное влияние лесных пожаров на локальный климат и экосистемы.

В рамках проекта разрабатывались инновационные методы анализа многомерных данных для моделирования сложных процессов в биосистемах, включая медицинские исследования. Основное внимание было уделено изучению структуры нуклеотидных последовательностей, в частности, анализу распределения GC-состава (доли гуанина и цитозина) в геномах хлоропластов, митохондрий и бактерий. Результаты исследования свидетельствуют о значительном потенциале разработанных методов для анализа взаимосвязей между структурой, функцией и таксономией биологических макромолекул. Эти подходы представляют

ценность для биоинформатики, а также могут быть успешно адаптированы для решения задач в других областях, включая анализ антропогенных экосистем.

Были проведены исследования, посвященные изучению влияния городской среды на формирование локальных климатических зон (ЛКЗ) и тепловых аномалий на примере Красноярска. Городская среда, включая промышленные объекты, асфальтированные дороги и плотную застройку, создает специфический микроклимат, отличающийся от окружающей местности. Результаты показали, что тепловые аномалии преимущественно сосредоточены в зонах крупногабаритной застройки и промышленных объектов. Анализ спутниковых снимков за 2013–2022 годы позволил выделить постоянные контуры тепловых аномалий, которые коррелируют с ЛКЗ. Установлено, что промышленные объекты, такие как алюминиевый завод (КраЗ), оказывают значительное влияние на локальное повышение температуры. Исследование подчеркивает важность комплексного подхода к изучению городской среды, включая анализ спутниковых данных, наземных измерений и метеорологических параметров. Полученные результаты могут быть использованы для планирования городского развития и смягчения негативного воздействия тепловых аномалий на экологическую обстановку и комфорт городской среды.

Выполнена оценка экологического состояния реки Енисей с использованием методов биоиндикации и многомерной статистики. Актуальность работы обусловлена необходимостью разработки комплексных подходов к мониторингу качества вод, особенно в условиях возрастающей антропогенной нагрузки. В качестве индикаторов использовались организмы зообентоса, а также 13 биоиндикационных индексов и метрик, включая индекс Шеннона, ЕРТ-индексы, индексы BMWP, ASPT и другие. Для анализа данных применялись методы корреляционного анализа, дисперсионного анализа и метод главных компонент. Результаты показали, что качество воды ухудшается по мере продвижения от верховья к низовью реки, с наилучшими показателями в верховье (категория «условно чистая») и наихудшими в низовье (категория «грязная»). Предложен интегральный показатель качества воды, основанный на многомерной статистике, который позволяет классифицировать станции мониторинга по четырем категориям: «условно чистая», «слабо загрязненная», «загрязненная» и «грязная». Исследование подчеркивает важность комплексного подхода к оценке экологического состояния водных объектов и необходимость дальнейшего развития методов биоиндикации и анализа данных.

Разработана технология построения комплексного биоиндикационного показателя качества воды, основанная на анализе данных, полученных из проб зообентоса. Исходные данные структурированы по трем категориям: свойства локализации места взятия пробы, характеристики места взятия пробы и биоиндикационные индексы и метрики качества

воды. Для анализа использован набор данных, включающий 174 пробы, взятые на 48 станциях реки Енисей, с расчетом 13 биоиндикационных индексов и метрик. Результатом является устойчивый комплексный показатель качества воды, позволяющий классифицировать станции по степени загрязнения. Методология может быть адаптирована для различных водных объектов и наборов данных.

В рамках настоящего проекта были продолжены работы развитию и совершенствованию «Системы мониторинга воздуха Красноярского научного центра СО РАН», которая функционирует уже более пяти лет. Система включает аппаратное, программное и информационно-аналитическое обеспечение для мониторинга состояния атмосферного воздуха в Красноярске. В городе установлены датчики для измерения концентрации взвешенных частиц (PM_{2.5}), метеостанции, температурные профилемеры, анемометры и другие устройства. В отчетный период проведены работы по оптимизации подсистем сбора и обработки данных, модернизации оборудования и разработке новых веб-интерфейсов. Также внедрены инструменты для анализа данных через мессенджер Telegram, что повысило оперативность мониторинга. Для улучшения точности прогнозов и анализа состояния атмосферы использовались данные глобальных моделей (GPM, GFS) и региональных метеостанций. Проект продолжает развиваться, уделяя особое внимание повышению точности и оперативности мониторинга, интеграции новых технологий для анализа экологической обстановки.

Были исследованы концентрации взвешенных частиц PM_{2,5} в атмосфере Красноярска по данным четырех сетей мониторинга: федеральной (Среднесибирское УГМС), краевой (Минэкологии края), научной (Система КНЦ СО РАН) и общественной («Небо»). Установлено, что среднемесячные и среднегодовые концентрации PM_{2,5}, измеренные разными сетями, существенно различаются. Данные УГМС и «Небо» показывают превышение среднегодового гигиенического норматива (ПДК_{сг}), что указывает на загрязнение атмосферы, тогда как данные Мэкологии и КНЦ СО РАН свидетельствуют о его отсутствии. Это ставит вопрос о достоверности измерений и необходимости калибровки датчиков всех сетей относительно эталонного гравиметрического метода. Результаты подчеркивают важность унификации подходов к мониторингу качества воздуха для принятия обоснованных природоохранных решений.

Был выполнен анализ концентрации бенз(а)пирена (BaP) в атмосфере Красноярска за период 2019-2023 годов на основе данных краевой и федеральной наблюдательных сетей. BaP, являясь высококанцерогенным веществом, образуется при неполном сгорании органических материалов и оказывает негативное воздействие на здоровье человека. В работе изучена многолетняя динамика концентраций BaP в разных районах города, выявлена зависимость концентрации от температуры воздуха и проведено статистическое сравнение данных

КНС и ГНС. Установлено, что наибольшие концентрации ВаР наблюдаются в зимний период, что связано с отопительным сезоном и неблагоприятными метеорологическими условиями. Показано, что среднегодовые концентрации ВаР, измеренные обеими сетями, статистически не различаются, что подтверждает достоверность данных. Результаты исследования подчеркивают необходимость дальнейшего мониторинга и контроля уровня ВаР в атмосфере Красноярска.

Проведен сравнительный анализ данных мониторинга концентрации взвешенных частиц PM_{2.5}, полученных с помощью двух систем: автоматизированных постов наблюдений системы мониторинга Минэкологии края, использующих анализаторы E-BAМ, и станций CityAir системы мониторинга КНЦ СО РАН, оснащенных оптическими датчиками. Анализ показал, что оптические датчики систематически завышают концентрацию PM_{2.5} по сравнению с E-BAМ, однако их показания могут быть скорректированы с учетом метеорологических параметров (температуры и давления). Регрессионные модели, включая линейную регрессию, случайный лес и метод k-ближайших соседей, продемонстрировали высокую точность при прогнозировании концентраций PM_{2.5}. Установлено, что параметры модели зависят от сезона, а учет влажности нецелесообразен из-за низкой точности измерений. Оптические датчики CityAir пригодны для мониторинга трендов загрязнения, для оперативной корректировки рекомендуется использовать линейную регрессию, а для ретроспективного анализа – метод случайного леса.

Проведен анализ формирования температурных инверсий в атмосфере Красноярска и их влияния на уровень загрязнения воздуха частицами PM_{2.5}. Основное внимание уделено зимнему периоду, когда стационарные антициклоны и рельеф местности способствуют застою воздуха и накоплению вредных примесей. Использованы данные мониторинга температуры, влажности, скорости ветра и концентрации PM_{2.5}, полученные с помощью сертифицированных станций и температурных профиломеров. Установлено, что приземные и приподнятые инверсии с нижней границей до 100 м и величиной инверсии более 5°C играют ключевую роль в повышении концентрации PM_{2.5}. Наиболее высокие уровни загрязнения наблюдаются при температуре ниже -10°C, относительной влажности 55–85% и слабом ветре (до 2 м/с). Инверсии над рекой Енисей оказывают меньшее влияние на загрязнение из-за их ослабленной приподнятой формы. Результаты подчеркивают важность учета метеоусловий и рельефа при прогнозировании и снижении уровня загрязнения воздуха в промышленных городах.

В проекте рассматривалось применение моделей машинного обучения для прогнозирования концентрации взвешенных частиц PM_{2.5} в атмосфере города Красноярска, где уровень загрязнения воздуха часто превышает допустимые нормы. Использованы данные

наземных станций мониторинга за период с января 2019 по март 2024 года, дополненные метеорологическими параметрами. Рассмотрены модели множественной регрессии, временных рядов (ARIMA, ARIMAX), а также ансамблевые методы (случайный лес, градиентный бустинг). Наилучшие результаты показали модели ARIMAX, которые рекомендованы для прогнозирования уровня загрязнения воздуха с высокой точностью. Оценка качества прогнозов проводилась с использованием метрик MAE, MSE, MAPE и R^2 .

Одна из задач исследований по настоящему проекту была связана с разработкой организационной модели управления безопасностью территорий, которая направлена на повышение эффективности принятия оперативных и стратегических решений в условиях чрезвычайных ситуаций (ЧС). Предложенная модель включает структурирование информационных ресурсов, разработку сервисов интеллектуальной обработки данных и анализ взаимодействия органов власти и ведомств. Особое внимание уделено фрактальности управления, где задачи распределяются между различными уровнями (федеральный, региональный, муниципальный) и акторами. Модель учитывает специфику территорий, таких как Арктическая зона или регионы с циклическими рисками (паводки, пожары). Внедрение аппаратно-программных комплексов (АПК), таких как «Безопасный город», позволяет интегрировать данные различных ведомств для улучшения реагирования на ЧС. Предложены методы ситуационного и аналитического моделирования, включая графовые и матричные представления, для оптимизации процессов управления и распределения ресурсов. На примере обеспечения пожарной безопасности сельских поселений Красноярского края показано, как формализованные данные и аналитические модели могут использоваться для автоматизации контроля, планирования и мониторинга мероприятий. Модель также предусматривает создание цифровых паспортов территорий, что способствует повышению качества управления и снижению рисков.

В проекте также рассматривалось применение математического моделирования для оценки безопасных условий эвакуации людей при пожаре в музеях. Представлена иерархия средств противопожарной защиты, включая объёмно-планировочные и конструктивные решения, системы обнаружения пожара, управление эвакуацией и ограничение количества людей в здании. Особое внимание уделено особенностям музеев, таких как наличие маломобильных групп населения, необходимость управления эвакуацией и сохранность музейных ценностей. Разработан алгоритм компьютерного моделирования для оценки эффективности противопожарных мер, который был применён на примере Зимнего дворца Государственного Эрмитажа. Показано, что безопасная эвакуация может быть обеспечена за счёт естественных средств противопожарной защиты при ограничении количества посетителей

до 4000 человек. Также подчёркивается важность разработки маршрутов эвакуации и проведения тренировочных эвакуаций для персонала.

В проекте предложена и исследована методика проверки гипотезы о независимости многомерных случайных величин, основанная на непараметрическом алгоритме распознавания образов. Данный подход позволяет избежать трудоемкого этапа дискретизации области значений случайных величин, характерного для традиционного критерия Пирсона. Методика предполагает формирование обучающей выборки для решения двухальтернативной задачи распознавания образов, где классы определяются в зависимости от независимости или зависимости случайных величин. Эффективность метода подтверждена вычислительными экспериментами, включая анализ неоднозначных зависимостей и сравнение с критерием Пирсона.

Разработанная методика адаптирована для работы с большими объемами данных за счет использования регрессионных оценок плотностей вероятностей, что значительно повышает вычислительную эффективность. Применение метода продемонстрировано на примере анализа данных дистанционного зондирования Земли (ДЗЗ) для антропогенных территорий. Результаты показали возможность сокращения числа спектральных признаков при распознавании объектов без потери точности. Методика также позволяет формировать наборы независимых случайных величин с использованием аппарата теории графов, что расширяет её применимость в задачах анализа многомерных данных.

Практическая значимость результатов исследований направлена, прежде всего, на эффективную организацию природопользования. В частности, исследование направлено на информационно-аналитическое обеспечение решения актуальных задач управления, включая количественную оценку характеристик природных и антропогенных экосистем, а также экологический мониторинг на территории Красноярского края. Информационное обеспечение играет ключевую роль в реализации широкого спектра мероприятий и организации систем поддержки принятия управленческих решений. Разрабатываемые технологии и программное обеспечение способствуют оперативному информированию населения Красноярска и прилегающих территорий о качестве атмосферного воздуха и динамике его изменений во времени.

ПРИЛОЖЕНИЕ А. Научные публикации в журналах, индексируемых в российских и международных информационно-аналитических системах научного цитирования

1. Shaparev N., Tokarev A., Yakubailik O. Cold traces of smoke from wildfires in the environment // *Environmental Processes*. 2024. V. 11. Article 5. DOI: 10.1007/s40710-024-00687-6, EDN: LXPMOC [Web of Science Q3, Scopus Q1, ядро РИНЦ, БелыйСписок 2]
2. Senashova M.Yu., Sadovsky M.G. The Distribution of the GC Content in the Spatial Structure of the Fragments of Mitochondrial, Chloroplast, and Bacterial Genomes // *Biophysics*. 2024. V. 69. № 2. P. 209-218. DOI: 10.1134/S0006350924700234, EDN: NRFUNG [Scopus Q4, ядро РИНЦ, БелыйСписок 3, ВАК]
3. Гостева А.А., Матушко А.К. Влияние городской среды на формирование локальных климатических зон и тепловых аномалий // *ИнтерКарто. ИнтерГИС*. 2024. Т. 30. № 2. С. 543-555. DOI: 10.35595/2414-9179-2024-2-30-543-555, EDN: JKSCZO [Scopus, RSCI, ядро РИНЦ, ВАК]
4. Андрианова А.В. Трансформация сообществ макрозообентоса в низовье р. Ангары после зарегулирования // *Russian Journal of Ecosystem Ecology*. 2024. Т. 9. № 1. С. 64-74. DOI: 10.21685/2500-0578-2024-1-4, EDN: RSFPZI [ВАК K2]
5. Карепова Е.Д., Петракова В.С. Статистически обоснованная корректировка показаний датчиков станций CityAir уровня концентрации взвешенных частиц PM_{2.5} в приземном слое атмосферы города // *Информатика и автоматизация*. 2024. Т. 23. № 2. С. 352-376. DOI: 10.15622/ia.23.2.2, EDN: EPWEZC [Scopus Q4, RSCI, ядро РИНЦ, БелыйСписок 2, ВАК K1]
6. Соколова О.В., Заворуев В.В., Заворуева Е.Н. Измерение концентрации PM₂₅ в атмосфере Красноярска в 2023 году: сравнение данных четырех сетей мониторинга // *Естественные и технические науки*. 2024. Т. 193. № 6. С. 192-196. DOI: 10.25633/ETN.2024.06.11, EDN: ULMPOU [ВАК K2]
7. Манкевич И.В., Заворуев В.В., Заворуева Е.Н. Концентрация бенз(а)пирена в атмосфере Красноярска в 2019-2023 годах: анализ данных краевой наблюдательной сети // *Естественные и технические науки*. 2024. Т. 193. № 6. С. 178-183. DOI: 10.25633/ETN.2024.06.10, EDN: BUWLCF [ВАК K2]
8. Сиротинин А.А., Володько О.С. Построение оптимизационной модели беспроводной внутренней сети для использования технологии интернета вещей // *Информационные и математические технологии в науке и управлении*. 2024. Т. 34. № 2. С. 135-143. DOI: 10.25729/ESI.2024.34.2.013, EDN: CQHIV [ВАК K2]

9. Ничепорчук В.В., Постникова У.С., Тасейко О.В. Проектирование интеллектуальной системы управления безопасностью территорий // *Онтология проектирования*. 2024. Т. 12. № 1. С. 42-54. DOI: 10.18287/2223-9537-2024-14-1-42-54, EDN: KRJTMQ [RSCI, ядро РИНЦ, БелыйСписок 4, ВАК K1]
10. Gravit M., Kirik E., Kozupitsa I., Solntseva Y. Fire safety in malls: Curtains for fire hazards localization and safe evacuation // *Construction of Unique Buildings and Structures*. 2024. V. 112. № 3. Article 11203. DOI: 10.4123/CUBS.112.3, EDN: PLEHHO [RSCI, ядро РИНЦ, БелыйСписок 4, ВАК K1]
11. Lapko A.V., Lapko V.A., Bakhtina A.V. Comparison of Methods for Testing the Hypothesis of Independence of Random Variables Based on a Nonparametric Classifier and Pearson's Chi-Squared Test // *Optoelectronics, Instrumentation and Data Processing*. 2024. V. 59. № 5. P. 551-560. DOI: 10.3103/S8756699023050047, EDN: BOCSZC [Web of Science Q4, Scopus Q4, ядро РИНЦ, БелыйСписок 2, ВАК]
12. Lapko A.V., Lapko V.A., Bakhtina A.V. Comparison of the Methodology for Hypothesis Testing of the Independence of Two-Dimensional Random Variables Based on a Nonparametric Classifier // *Scientific and Technical Information Processing*. 2024. V. 50. № 6. P. 572581. DOI: 10.3103/S0147688223060084, EDN: NIJSMZ [Web of Science Q4, Scopus Q4, ядро РИНЦ, БелыйСписок 3, ВАК]
13. Lapko A.V., Lapko V.A. Analysis of optimization methods for nonparametric estimation of probability density in large samples // *Measurement Techniques*. 2024. V. 66. № 11. P. 839-847. DOI: 10.1007/s11018-024-02298-x, EDN: GMYFER [Web of Science Q4, Scopus Q4, ядро РИНЦ, БелыйСписок 4, ВАК]
14. Lapko A.V., Lapko V.A. Comparison of optimization methods for the regression estimation of the probability density of a one-dimensional random variable // *Optoelectronics, Instrumentation and Data Processing*. 2024. V. 60. № 3. P. 390-397. DOI: 10.3103/S8756699024700456, EDN: MJTMIQ [Web of Science Q4, Scopus Q4, ядро РИНЦ, БелыйСписок 2, ВАК]

ПРИЛОЖЕНИЕ Б. Зарегистрированные результаты интеллектуальной деятельности

1. Якубайлик Т.В., Токарев А.В., Якубайлик О.Э. Почасовые данные о наличии туманов над рекой Енисей в г. Красноярске по данным видеонаблюдений за 2022 – 2023 гг. // Федеральная служба по интеллектуальной собственности (Роспатент). Свидетельство о государственной регистрации базы данных № 2024623202 от 19 июля 2024 г. Регистрационный номер ЕГИСУ НИОКТР 624120400040-9.
2. Токарев А.В. Программа «Sensor Collector» для сбора, агрегации и представления данных оперативных наблюдений за состоянием окружающей среды // Федеральная служба по интеллектуальной собственности (Роспатент). Свидетельство о государственной регистрации программы для ЭВМ № 2024667574 от 26 июля 2024 г. Регистрационный номер ЕГИСУ НИОКТР 624120400028-7.
3. Кадочников А.А. Программа анализа и мониторинга процесса обработки поступающей информации для системы сбора данных // Федеральная служба по интеллектуальной собственности (Роспатент). Свидетельство о государственной регистрации программы для ЭВМ № 2024667060 от 19 июля 2024 г. Регистрационный номер ЕГИСУ НИОКТР 624120400039-3.
4. Андрианова А.В., Каропова Е.Д., Малимонов М.И., Пушкарев А.А., Якубайлик О.Э. ГИС «Комплексная оценка качества вод р. Енисей» // Федеральная служба по интеллектуальной собственности (Роспатент). Свидетельство о государственной регистрации программы для ЭВМ № 2024690444 от 16 декабря 2024 г. Регистрационный номер ЕГИСУ НИОКТР 625011000535-4.

ПРИЛОЖЕНИЕ В. Прочие публикации и доклады на конференциях

1. Безопасность России. Правовые, социально-экономические и научно-технические аспекты: Тематический блок «Региональные проблемы безопасности». Раздел I. Мониторинг, риски и безопасность Сибирского федерального округа / Н.В. Абасов, Э.Ш. Акбулатов, А.В. Андрианова [и др.]. – Москва: МГОФ "Знание", 2024. – 644 с. – ISBN 978-5-87633-206-6. EDN: JERRSN [РИНЦ]
2. Безопасность России. Правовые, социально-экономические и научно-технические аспекты: Тематический блок «Региональные проблемы безопасности». Раздел II. Территориальные риски регионов Сибири. Кузбасс. Енисейская Сибирь. Байкал / Н.В. Абасов, Э.Ш. Акбулатов, А.В. Андрианова [и др.]. – Москва: МГОФ "Знание", 2024. – 624 с. – ISBN 978-5-87633-207-3. EDN: HZWUWK [РИНЦ]
3. Сенашова М.Ю. Упорядоченность значений GC-состава фрагментов в пространственной структуре геномов органелл // Информационные и математические технологии в науке и управлении. 2024. Т. 34. № 2. С. 33-40. DOI: 10.25729/ESI.2024.34.2.003, EDN: IPIWYV [БАК K2]
4. Ovchinnikova Iu., Sadovsky M., Sokolov V. Clustering of bacteria by the triplet composition of 5S RNA // ITM Web of Conferences. 2024. V. 59, 04014. DOI: 10.1051/itmconf/20245904014, EDN: AKUJRP [РИНЦ]
5. Сенашова М., Садовский М. Связь информационной емкости и таксономии геномов оспы // Биоинформатика регуляции и структуры геномов / системная биология: Четырнадцатая международная мультиконференция: Тез. докл. (Россия, Новосибирск, 5–10 авг. 2024 г.) / Ин-т цитологии и генетики СО РАН; Новосиб. нац. исслед. гос. ун-т. – Новосибирск: ИциГ СО РАН, 2024. С. 39-43. DOI: 10.18699/bgrs2024-1.1-10. URL: <https://bgrssb.icgbio.ru/2024/ru/>
6. Овчинникова Ю.И., Садовский М.Г. Сравнительное изучение структуры и таксономии на примере 5S и 16S РНК бактерий // Биоинформатика регуляции и структуры геномов / системная биология: Четырнадцатая международная мультиконференция: Тез. докл. (Россия, Новосибирск, 5–10 авг. 2024 г.) / Ин-т цитологии и генетики СО РАН; Новосиб. нац. исслед. гос. ун-т. – Новосибирск: ИциГ СО РАН, 2024. С. 974-975. DOI: 10.18699/bgrs2024-5.2-09. URL: <https://bgrssb.icgbio.ru/2024/ru/>
7. Якубайлик О.Э., Заворуев В.В., Кадочников А.А., Токарев А.В. Информационно-вычислительные системы и технологии оперативного мониторинга окружающей природной среды // Региональные проблемы дистанционного зондирования Земли: XI Междунар. науч. конф. Красноярск, 10–13 сентября 2024 г. – Красноярск: Сиб. федер. ун-т, 2024.

URL: <https://rprs.sfu-kras.ru/node/113>

8. Володько О.С., Лев Н.А., Полянчикова Д.В. Прогнозирование величины концентрации загрязняющих веществ в атмосфере г. Красноярска с помощью моделей машинного обучения // Региональные проблемы дистанционного зондирования Земли: материалы XI Междунар. науч. конф. Красноярск, 10–13 сентября 2024 г. – Красноярск: Сиб.федер. ун-т, 2024. С. 352-356. EDN: NXCYXP [РИНЦ]
9. Володько О.С., Лемешкова Е.Н. Вертикальные моды внутренних сейш в стратифицированном озере // Марчуковские научные чтения 2024: Тезисы Междунар. конф., 7–11 октября 2024 г. / Ин-т вычислит. математики и матем. геофизики СО РАН. С. 74. EDN: BLOHET [РИНЦ]
10. Голубева В.Д., Матушко А.К. Применение спутниковых снимков для получения температуры воды в реке Енисей // Региональные проблемы дистанционного зондирования Земли: материалы XI Междунар. науч. конф. Красноярск, 10–13 сентября 2024 г. – Красноярск : Сиб.федер. ун-т, 2024 С. 164-168. EDN: TSFNHW [РИНЦ]
11. Петракова В.С., Карпова Е.Д. Алгоритмы сегментации мультимодально распределенных данных // Прикладная математика и фундаментальная информатика. Материалы XIV Международной молодежной научно-практической конференции с элементами научной школы. Омск, 2024. С. 31-32. EDN: GCTCUE [РИНЦ]
12. Ничепорчук В.В. Исследование информационной обеспеченности взаимодействия в РСЧС // Технологии гражданской безопасности. 2024. Т. 21. № 4(82). С. 47-51. EDN: QTYCQE [БАК K1]
13. Ничепорчук В.В., Кобыжакова С.В. Информационная поддержка управления пожарными рисками сельских территорий // Сибирский пожарно-спасательный вестник. 2024. № 3 (34). С. 66-74. DOI: 10.34987/vestnik.sibpsa.2024.56.72.004, EDN: APDQZC [БАК K3]
14. Ничепорчук В.В., Сулова С.С. Иерархия моделей информационного обеспечения управления безопасностью территорий // Моделирование сложных процессов и систем. Сборник трудов XXXIV Международной научно-практической конференции. Химки, 2024. С. 39-45. EDN: RQFBZW [РИНЦ]
15. Мариупольский Т.И., Ничепорчук В.В., Кустов Д.В., Панфилов И.А., Храпунова Е.В. Анализ данных антропогенных пожаров // Научно-технический вестник Поволжья. 2024. № 6. С. 80-82. EDN: DETKCI [РИНЦ]
16. Ничепорчук В. В., Сулова С.С. Метод комплексной поддержки управления лесопожарной обстановкой // Материалы VI Всероссийской научно-практической конференции

- «Актуальные проблемы обеспечения пожарной безопасности и защиты от чрезвычайных ситуаций». – Железногорск: СибПСА, 2024.– С. 248-254. EDN: JBCRXТ [РИНЦ]
17. Ничепорчук В.В., Кобыжакова С.С. Специфика применения нормативно-правовых актов для предупреждения и тушения природных пожаров / Труды VI Всероссийской научно-практической конференции «Актуальные проблемы обеспечения пожарной безопасности и защиты от чрезвычайных ситуаций». – Железногорск: СибПСА, 2024.– С. 376-383. EDN: GNBKRD [РИНЦ]
18. Ничепорчук В.В., Сулова С.С. Применение анализа данных о бытовых пожарах в территориальном управлении / Решетнёвские чтения: Материалы XXVIII Междунар. науч. конф., посвящ. 100-летию генерального конструктора ракетно-космических систем академика М.Ф. Решетнёва (18–22 нояб. 2024, г. Красноярск): в 2 ч.– Красноярск, 2024. Ч. 2. – С. 188-190. EDN: CEQNDB [РИНЦ]
19. Ничепорчук В.В., Коробко А.В. Информационные ресурсы управления безопасностью территорий // Информационные технологии и нанотехнологии (ИТНТ-2024). Сборник трудов по материалам X Международной конференции и молодежной школы. В 6-ти томах. Самара, 2024. С. 41512. EDN: AIOKPS [РИНЦ]
20. Ничепорчук В.В., Кобыжакова С.В. Цифровые технологии в задачах управления безопасностью территорий // Мониторинг, моделирование и прогнозирование опасных природных явлений и чрезвычайных ситуаций: Международная научно-практическая конференция. г. Красноярск, 25 октября 2024 г., СПСА г. Железногорск, 2024. URL: <https://monitoring.sibpsa.ru/>
21. Ничепорчук В.В., Сулова С.С. Применение анализа данных о бытовых пожарах в территориальном управлении // Решетневские чтения. Материалы XXVIII Международной научно-практической конференции, посвященной 100-летию со Дня рождения генерального конструктора ракетно-космических систем академика Михаила Федоровича Решетнева. В 2-х частях. Красноярск, 2024. Ч.1. С. 188-190. EDN: CEQNDB [РИНЦ]
22. Nicheporchuk V., Penkova T. Architecture of Intelligent Platform for Territorial Safety Management // 2024 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), Novosibirsk, Russian Federation, 2024, P. 372-377, DOI: 10.1109/SIBIRCON63777.2024.10758467 [Scopus]
23. Кияткина Е.Н., Лобаев И.А., Кирик Е.С., Воронов С.П. Поддержка принятия управленческих решений при применении обязательных требований соразмерных риску причинения вреда // Сибирский пожарно-спасательный вестник. 2024. Т. 33. № 2. С. 34-47. DOI: 10.34987/vestnik.sibpsa.2024.84.80.021, EDN: YLIDEP [БАК K3]

24. Vitova T., Kirik E. Personal Space of People in Movement under Different Conditions // 15th International Conference on Parallel Processing & Applied Mathematics – Ostrava, Czech Republic, September 8-11, 2024. URL: <https://ppam.edu.pl/program>
25. Lapko A.V., Lapko V.A., Bakhtina A.V. Application of a nonparametric procedure for testing the hypothesis about the independence of random variables given a large amount of statistical data // Measurement Techniques. 2024. V. 66. № 10. P. 744-754. DOI: 10.1007/s11018-024-02288-z, EDN: RWJRK [Web of Science Q4, Scopus Q4, ядро РИНЦ, БелыйСписок 4, ВАК]
26. Lapko A.V., Lapko V.A. Procedure for decomposing the values of two-dimensional spectral features of remote sensing based on the analysis of correlation coefficient components // Measurement Techniques. 2024. V. 67. № 6. P. 427-432. DOI: 10.1007/s11018-024-02362-6, EDN: KYBFAN [Web of Science Q4, Scopus Q4, ядро РИНЦ, БелыйСписок 4, ВАК]
27. Лапко А.В., Лапко В.А. Оценивание традиционных числовых характеристик логнормальных законов распределения одномерной случайной величины в условиях большого объёма статистических данных // Измерительная техника. 2024. Т. 73. № 2. С. 23-29. DOI: 10.32446/0368-1025it.2024-2-23-29, EDN: JXKNGI [RSCI, ядро РИНЦ, БелыйСписок 3, ВАК]
28. Лапко А.В., Лапко В.А. Методика декомпозиции значений двухмерных спектральных признаков дистанционного зондирования на основе анализа составляющих коэффициента корреляции // Измерительная техника. 2024. Т. 73. № 6. С. 12-17. DOI: 10.32446/0368-1025it.2024-6-12-17, EDN: PPHUJK [RSCI, ядро РИНЦ, БелыйСписок 3, ВАК]
29. Лапко А.В., Лапко В.А. Сравнение методов оптимизации регрессионной оценки плотности вероятности одномерной случайной величины // Автометрия. 2024. Т. 60. № 3. С. 95-103. DOI: 10.15372/AUT20240309, EDN: PRRFGH [RSCI, ядро РИНЦ, БелыйСписок 2, ВАК]
30. Лапко А.В., Лапко В.А. Быстрая методика оптимизации регрессионной оценки плотности вероятности одномерной случайной величины // Искусственный интеллект и принятие решений. 2024. № 2. С. 123-131. DOI: 10.14357/20718594240210, EDN: SEACUQ [RSCI, ядро РИНЦ, БелыйСписок 3, ВАК K1]
31. Тубольцев В.П., Лапко А.В., Лапко В.А. Непараметрический алгоритм автоматической классификации данных дистанционного зондирования // Информатика и ее применения. 2023. Т. 17. № 4. С. 23-31. DOI: 10.14357/19922264230404, EDN: MPEWAW [RSCI, ядро РИНЦ, БелыйСписок 3, ВАК K1]

32. Тубольцев В.П., Лапко А.В., Лапко В.А. Модифицированный непараметрический алгоритм автоматической классификации статистических данных большого объема и его применение // Искусственный интеллект и принятие решений. 2023. № 4. С. 49-57. DOI: 10.14357/20718594230405, EDN: QHNFRU [RSCI, ядро РИНЦ, БелыйСписок 3, ВАК K1]
33. Лапко А.В., Лапко В.А., Шаруева А.В. Непараметрический алгоритм распознавания образов в задаче формирования наборов независимых случайных величин // Информатика и системы управления. 2024. Т. 79. № 1. С. 81-90. DOI: 10.22250/18142400_2024_79_1_81, EDN: BHGEQU [ВАК K2]
34. Лапко А.В., Лапко В.А. Быстрый алгоритм оптимизации многомерной регрессионной оценки плотности вероятности независимых случайных величин // Информатика и системы управления. 2024. Т. 80. № 2. С. 70-84. DOI: 10.22250/18142400_2024_80_2_70, EDN: KSZIST [ВАК K2]
35. Лапко А.В., Лапко В.А., Им С.Т. Методика декомпозиции спектральных данных дистанционного зондирования пожаров лесных территорий // Информатика и системы управления. 2024. Т. 81. № 3. С. 112-120. DOI: 10.22250/18142400_2024_81_3_112, EDN: ISVPEU [ВАК K2]
36. Данилин И.М., Лапко В.А., Кузнецов А.А., Бабий И.А., Вайсман А.О. Инновационный учебно-научный центр мониторинга лесных ресурсов Сибири на основе лазерной и микроволновой аэрокосмической съемки // Сибирский аэрокосмический журнал. 2024. Т. 25. № 1. С. 8-17. DOI: 10.31772/2712-8970-2024-25-1-8-17, EDN: DFUBIJ [RSCI, ядро РИНЦ, ВАК K2]
37. Лапко А.В., Лапко В.А. Декомпозиция значений спектральных признаков дистанционного зондирования на основе составляющих коэффициента корреляции // Решетневские чтения. Материалы XXVIII Международной научно-практической конференции, посвященной 100-летию со Дня рождения генерального конструктора ракетно-космических систем академика Михаила Федоровича Решетнева. В 2-х частях. Красноярск, 2024. Ч.1. С. 458-460. EDN: JHEGXL [РИНЦ]
38. Лапко А.В., Лапко В.А. Двухуровневая система декомпозиции области значений случайной величины // Решетневские чтения. Материалы XXVIII Международной научно-практической конференции, посвященной 100-летию со Дня рождения генерального конструктора ракетно-космических систем академика Михаила Федоровича Решетнева. В 2-х частях. Красноярск, 2024. Ч.1. С. 461-463. EDN: ZHBJOE [РИНЦ]