

Федеральное государственное бюджетное учреждение  
«Гидрометеорологический научно-исследовательский центр  
Российской Федерации»

На правах рукописи

Борисов Данил Владимирович

**ПОСТПРОЦЕССИНГ ЧИСЛЕННЫХ ПРОГНОЗОВ КОНЦЕНТРАЦИЙ  
ВЗВЕШЕННЫХ ЧАСТИЦ (PM<sub>10</sub>) И ПРИЗЕМНОГО ОЗОНА (O<sub>3</sub>)  
С ИСПОЛЬЗОВАНИЕМ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ**

1.6.18 – Науки об атмосфере и климате

**ДИССЕРТАЦИЯ**

на соискание ученой степени  
кандидата географических наук

Научный руководитель:  
доктор географических наук  
Кузнецова Ирина Николаевна

Москва – 2025

## ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ .....	4
ГЛАВА 1. АНАЛИТИЧЕСКИЙ ОБЗОР .....	10
1.1 Современное прогнозирование качества воздуха с применением химических транспортных моделей .....	10
1.2 Методы улучшения качества численных прогнозов воздушных загрязнений .....	20
1.3 Применение методов машинного обучения для прогнозирования качества воздуха .....	22
1.4 Методы интеграции ХТМ и моделей машинного обучения .....	25
Выводы к главе 1 .....	31
ГЛАВА 2. ИСПОЛЬЗОВАННЫЕ МЕТОДЫ И ДАННЫЕ .....	33
2.1 О методе искусственные нейронные сети.....	33
2.1.1 Структура и основные составляющие искусственных нейронных сетей.....	34
2.1.2 Обучение искусственных нейронных сетей .....	37
2.1.3 Разновидности архитектур и методы улучшения оптимизации искусственных нейронных сетей.....	40
2.2 Данные измерений концентраций $O_3$ и $PM_{10}$ в московском регионе.....	43
2.2.1 Сезонная и суточная изменчивость концентраций приземного озона .....	46
2.2.2 Сезонная и суточная изменчивость концентраций $PM_{10}$ , эпизоды повышенного аэрозольного загрязнения .....	49
2.3 Численные прогнозы метеопараметров мезомасштабной системы COSMO-Ru .....	60
2.4 Численные прогнозы концентраций $O_3$ и $PM_{10}$ на территории московского региона .....	62
2.4.1 Ретроспективная оценка качества модельных прогнозов концентраций $PM_{10}$ и $O_3$ на территории московского региона .....	65
Выводы к главе 2 .....	73
ГЛАВА 3. МЕТОДИКА РАЗРАБОТКИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОСТПРОЦЕССИНГА ЧИСЛЕННЫХ ПРОГНОЗОВ ПРИЗЕМНЫХ КОНЦЕНТРАЦИЙ ЗАГРЯЗНЯЮЩИХ ВЕЩЕСТВ И ЕЕ РЕАЛИЗАЦИЯ .....	75
3.1 Общее описание и основные этапы методики .....	75
3.2 Реализация Методики в виде построения моделей машинного обучения для постпроцессинга численных прогнозов концентраций $O_3$ и $PM_{10}$ .....	79
3.2.1 Состав и объемы обучающих и тестовых выборок моделей постпроцессинга численных прогнозов концентраций $O_3$ и $PM_{10}$ .....	79

3.2.2 Экспериментальный поиск оптимальных конфигураций моделей постпроцессинга численных прогнозов $O_3$ и $PM_{10}$ .....	83
Выводы к главе 3 .....	92
ГЛАВА 4. ВЕРИФИКАЦИЯ РАЗРАБОТАННОЙ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОСТПРОЦЕССИНГА ЧИСЛЕННЫХ ПРОГНОЗОВ КОНЦЕНТРАЦИЙ ПРИЗЕМНОГО ОЗОНА .....	93
4.1 Результаты постпроцессинга численных прогнозов концентраций приземного озона с применением разработанной модели машинного обучения .....	94
4.2 Эффекты ММО-постпроцессинга численных прогнозов в периоды повышенного озонowego загрязнения .....	100
4.3 Эффекты ММО-постпроцессинга полей прогнозов концентраций приземного озона....	107
Выводы к главе 4 .....	112
ГЛАВА 5. ТЕСТИРОВАНИЕ РАЗРАБОТАННОЙ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОСТПРОЦЕССИНГА ЧИСЛЕННЫХ ПРОГНОЗОВ КОНЦЕНТРАЦИЙ $PM_{10}$ .....	113
5.1 Эффекты ММО-постпроцессинга численных прогнозов концентраций $PM_{10}$ .....	113
5.2 ММО-постпроцессинг численных прогнозов концентраций $PM_{10}$ в периоды повышенного аэрозольного загрязнения.....	119
5.3 Эффекты ММО-постпроцессинга полей прогнозов концентраций $PM_{10}$ .....	130
Выводы к главе 5 .....	134
ЗАКЛЮЧЕНИЕ .....	135
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ .....	139
СПИСОК ЛИТЕРАТУРЫ .....	141

## ВВЕДЕНИЕ

### Актуальность работы

В настоящее время активно развиваются два направления контроля загрязнения атмосферы – системы мониторинга качества воздуха, включающие наземные и спутниковые измерения концентраций загрязнителей, и системы численного моделирования и прогнозирования концентраций загрязняющих веществ (ЗВ) с применением химических транспортных моделей (ХТМ). Преимуществом использования ХТМ является получение данных о загрязнении в узлах регулярной сетки, покрывающей обширные территории, в отличие от неравномерно распределенных данных измерений [World Meteorological Organization, 2020]. Очевидным недостатком ХТМ являются отклонения прогнозов концентраций ЗВ, выявляемые при сравнении с данными измерений. Причиной отклонений прогнозов могут выступать как качество исходных данных о выбросах загрязнителей, прогноза метеопараметров, так и допущения моделирования физических и химических процессов в ХТМ [Srivastava and Nadèg, 2021; Шалыгина и др., 2020; Mailler et al., 2023].

Существует множество подходов к повышению качества прогнозов ХТМ – ансамблирование; коррекция исходных данных; задание начальных полей концентраций с использованием данных измерений; корректируются физические и химические параметризации ХТМ. Также применяется статистическая коррекция готовых прогнозов, например, по среднему отклонению от измерений концентраций.

Ввиду стремительного увеличения объема оцифрованных данных и вычислительных мощностей, в последнее десятилетие широкое применение получили модели машинного обучения (сокр. ММО), в том числе в области прогнозирования качества воздуха [Masood and Ahmad, 2021]. Преимуществами статистических моделей машинного обучения являются их универсальность, относительная простота использования, способность обучаться, т.е. накапливать информацию. Современные модели машинного обучения, в отличие от предшествующих статистических методов, способны воспроизводить сложные нелинейные связи между концентрацией ЗВ и влияющими на нее факторами [Хайкин, 2006].

Многие исследователи используют ММО обособленно для прогноза концентраций в локациях измерительных станций, либо в среднем по территории [Masood and Ahmad, 2021]. При таком подходе обучающие выборки моделей формируются из измерений концентраций и метеопараметров. Другой подход заключается в обучении ММО на численных прогнозах детерминированных моделей (ХТМ и моделей атмосферы) с использованием измерений концентраций в ячейках модельной сетки в качестве целевой переменной [Kleinert et al., 2022].

Таким образом ММО обучается предсказывать максимально близкие к измерениям концентрации, тем самым повышая качество прогноза ХТМ. Обученные таким образом ММО потенциально применимы для улучшения прогнозов концентраций ЗВ по всей территории, в том числе в ячейках модельной сетки без измерений.

В работе при разработке методики постпроцессинга численных прогнозов ЗВ в качестве маркеров воздушного загрязнения были выбраны концентрации приземного озона ( $O_3$ ) и твердых взвешенных частиц диаметром менее 10 мкм (англ. Particulate Matter, сокр.  $PM_{10}$ ). По данным Всемирной организации здравоохранения (ВОЗ)  $O_3$  и  $PM_{10}$  являются приоритетными показателями при оценке качества воздуха [World Health Organization, 2021]. Повышенное содержание в приземном слое воздуха  $O_3$  и  $PM_{10}$  вызывает острые и хронические заболевания и приводит к увеличению смертности [WHO Regional Office for Europe, 2003], приближение концентраций  $O_3$  и  $PM_{10}$  к опасным уровням происходит в основном в периоды неблагоприятных для рассеивания загрязнений метеоусловий (НМУ). Заблаговременное предупреждение о таких ситуациях имеет экологическую ценность и подчеркивает актуальность улучшения численных прогнозов концентраций  $O_3$  и  $PM_{10}$ .

**Целью** исследования является разработка на основе методов машинного обучения инструмента постпроцессинга прогнозов химической транспортной модели концентраций приземного озона ( $O_3$ ) и твердых взвешенных частиц ( $PM_{10}$ ).

Для достижения цели исследования поставлены и решены следующие **задачи**:

1. Анализ данных измерений приземных концентраций  $O_3$  и  $PM_{10}$  на территории московского региона. Изучение сезонно-суточной изменчивости концентраций  $O_3$  и  $PM_{10}$ , влияющих на нее факторов, а также условий повышения концентраций  $O_3$  и  $PM_{10}$  до опасных значений.
2. Аналитический обзор современных систем прогнозирования качества воздуха, методов повышения качества прогнозов и практик применения методов машинного обучения.
3. Изучение методов машинного обучения (МО).
4. Установление модельных отклонений прогнозов на сетке с горизонтальным шагом  $2 \times 2$  км от измеренных концентраций  $O_3$  и  $PM_{10}$  в московском регионе.
5. Формирование исследовательской базы данных для разработки моделей машинного обучения (ММО), предназначенных для постпроцессинга численных прогнозов концентраций  $O_3$  и  $PM_{10}$  на территории московского региона.
6. Разработка моделей машинного обучения для постпроцессинга численных прогнозов концентраций приземного озона (ММО- $O_3$ ) и взвешенных частиц (ММО- $PM_{10}$ ).
7. Верификация разработанных моделей постпроцессинга (ММО- $O_3$  и ММО- $PM_{10}$ ) на независимых выборках с использованием данных измерений концентраций в московском регионе.

8. Формулирование методики построения моделей постпроцессинга численных прогнозов концентраций для применения в задачах прогнозирования качества воздуха.

### **Методы исследования**

Для решения сформулированных задач исследования использованы методы обработки больших данных, статистики, оптимизации, машинного обучения, объектно-ориентированного программирования, нейронные сети, методы оценки качества численных модельных прогнозов, методы синоптического и траекторного анализа.

### **Использованные данные**

В работе использованы архивы прогнозов химической транспортной модели CHIMERE и модели атмосферы COSMO-Ru2 на регулярной сетке с горизонтальным шагом  $2 \text{ км} \times 2 \text{ км}$  для территории московского региона, архивы автоматических наземных измерений концентраций  $\text{O}_3$  и  $\text{PM}_{10}$  в московском регионе в период февраль 2019 – сентябрь 2024 г. Также использованы данные измерений метеопараметров на территории г. Москвы, архивные прогнозы глобальных моделей атмосферы GFS, ICON; архивные прогнозы XTM SILAM, данные гибридной одночастичной лагранжевой интегрированной траекторной модели NOAA.

### **Научная новизна**

1. Разработана методика построения моделей машинного обучения для постпроцессинга численных прогнозов концентраций загрязняющих веществ;
2. Разработаны и протестированы модели машинного обучения для постпроцессинга численных прогнозов концентраций приземного озона и взвешенных частиц  $\text{PM}_{10}$  на территории московского региона;
3. Разработанные модели машинного обучения применимы для постпроцессинга численных прогнозов  $\text{O}_3$  и  $\text{PM}_{10}$  на территориях, где не проводятся измерения концентраций.

### **Теоретическая и практическая значимость**

Реализованные по разработанной методике модели машинного обучения позволяют улучшить качество рассчитанных химической транспортной моделью прогнозов приземных концентраций  $\text{O}_3$  и  $\text{PM}_{10}$ , в том числе на территориях без измерений концентраций.

Разработанные модели машинного обучения могут применяться для оперативного постпроцессинга численных часовых прогнозов на первые сутки приземных концентраций  $\text{O}_3$  и  $\text{PM}_{10}$  на модельной сетке с горизонтальным шагом  $2 \text{ км} \times 2 \text{ км}$  для территории московского региона.

Созданный универсальный программный комплекс охватывает полный цикл разработки моделей машинного обучения для постпроцессинга численных прогнозов приземных концентраций загрязняющих веществ на регулярных сетках варьируемого разрешения, не имеет географической привязки и применим для регионов, обеспеченных прогнозами XTM и непрерывными измерениями концентраций.

Работа выполнялась в рамках НИТР плана Росгидромета 4.7.2 (2020–2024 гг.) и 4.9 (2025 г.).

### **Положения, выносимые на защиту**

1. Методика разработки моделей машинного обучения для постпроцессинга прогнозов концентраций загрязняющих веществ, рассчитанных химической транспортной моделью;

2. Разработанная модель машинного обучения для постпроцессинга прогнозов концентраций приземного озона на территории московского региона улучшает качество прогнозов химической транспортной модели, обеспечивая снижение среднеквадратического отклонения прогнозов от измерений в отдельные месяцы на 12–43 %, и увеличение коэффициента корреляции между прогнозами и измерениями до 0.7–0.9;

3. Разработанная модель машинного обучения для постпроцессинга численных прогнозов концентраций взвешенных частиц с диаметром менее 10 мкм ( $PM_{10}$ ) улучшает качество прогнозов химической транспортной модели, обеспечивая снижение среднеквадратического отклонения прогнозов от измерений в отдельные месяцы на 25–62 %, и увеличение коэффициента корреляции между прогнозами и измерениями до 0.4–0.6, а также успешно корректирует численные прогнозы  $PM_{10}$  при неблагоприятных для рассеивания загрязнений метеорологических условиях.

### **Достоверность научных результатов**

Достоверность полученных результатов и обоснованность выводов исследования подтверждается:

а) результатами  $k$ -блоковой валидации ( $k = 5$ ) разработанных моделей машинного обучения по обучающим выборкам, которые охватывает 5-летний период часовых прогнозов и измерений концентраций  $O_3$  и  $PM_{10}$ , включает различные сезонные условия формирования загрязнений на территории московского региона, периоды повышенных концентраций  $O_3$  и  $PM_{10}$  при НМУ;

б) тестированием разработанных моделей машинного обучения на независимых выборках, которые охватывают 7 месяцев часовых прогнозов и измерений  $O_3$  и  $PM_{10}$  и также включают различные метеоусловия формирования атмосферных загрязнений, в том числе периоды повышенных концентраций  $O_3$  и  $PM_{10}$  при НМУ. Эффективность применения ММО-постпроцессинга прогнозов концентраций  $O_3$  и  $PM_{10}$  на территориях без измерений концентраций доказана с использованием метода исключения измерительной станции из обучающей выборки ММО.

### **Апробация результатов**

Результаты диссертационной работы были представлены на следующих научных конференциях, симпозиумах и семинарах: международная молодежная школа и конференция по вычислительно-информационным технологиям для наук об окружающей среде CITES'2023, 13 –

23 июня 2023 г., Москва, Россия; Всероссийская научно-практическая конференция с международным участием «Россия в Десятилетии ООН наук об океане. Ресурсы и знания». Москва, 6-8 ноября 2024 г.; V Всероссийская конференция с международным участием, посвященная памяти академика А.М. Обухова "Турбулентность, динамика атмосферы и климата", Москва, 19–21 ноября 2024 г.; семинар Центра наук о земле МФТИ, г. Долгопрудный, Московская область, 19 сентября 2024 г.; заседание ученого совета Физического направления Института Океанологии им. П. П. Ширшова РАН, г. Москва, 8 ноября 2024 г.; семинары по краткосрочным и среднесрочным прогнозам погоды в 2023 и 2024 гг., Гидрометцентр России.

### **Публикации**

По теме диссертации опубликовано 28 работ, в том числе 10 статей в журналах, рекомендованных Высшей аттестационной комиссией (ВАК); 10 публикаций в материалах международных конференций. Получено 4 свидетельства Роспатента о государственной регистрации программы для ЭВМ.

### **Личный вклад автора**

Представленные в диссертационной работе методика, модели машинного обучения, программный комплекс полностью спроектированы, реализованы и протестированы лично автором. Все результаты исследования получены, проанализированы и описаны автором лично или в соавторстве с научным руководителем, доктором географических наук, Кузнецовой Ириной Николаевной.

### **Благодарности**

Автор выражает благодарность своему научному руководителю д.г.н. Кузнецовой Ирине Николаевне за всестороннюю поддержку на пути исследования и огромный вклад, внесенный в диссертационную работу. Автор выражает благодарность сотрудникам Гидрометцентра России к.г.н. Александру Андреевичу Кирсанову, к.ф.-м.н. Филиппу Леонидовичу Быкову, к.г.н. Ирине Юрьевне Шалыгиной, к.г.н. Юлии Васильевне Ткачевой, к.г.н. Мурату Ислемгалеевичу Нахаеву, к.ф.-м.н. Марине Владимировне Шатуновой, д.ф.-м.н. Гдалию Симоновичу Ривину и другим сотрудникам за консультационную и техническую поддержку, а также за конструктивные замечания по ходу исследования. Автор выражает благодарность сотрудникам Центра наук о земле МФТИ и Института Океанологии им. П. П. Ширшова РАН, в частности Криницкому Михаилу Алексеевичу, за ценные обсуждения и замечания, положительно повлиявшие на результаты исследования. Автор благодарит своих родителей Татьяну Ивановну и Владимира Ивановича за всестороннюю поддержку.



**Структура и объем диссертации**

Диссертация состоит из введения, 5 глав, заключения, списка используемых источников и литературы из 129 наименований. Общий объем диссертации составляет 152 страницы. Диссертация содержит 63 рисунка и 5 таблиц.

## ГЛАВА 1. АНАЛИТИЧЕСКИЙ ОБЗОР

### 1.1 Современное прогнозирование качества воздуха с применением химических транспортных моделей

История развития систем моделирования и прогнозирования загрязнения атмосферного воздуха насчитывает около 100 лет. В 1930-х годах проводились первые исследования распространения шлейфа от промышленных труб [Sutton, 1932; Bosanquet and Pearson, 1936]. В 1960-х годах практика расчетов распространения шлейфов от одиночных источников, в том числе линейных и площадных стала применяться в промышленно развитых странах с целью контроля загрязнений в промышленных районах [Turner, 1964; Holzworth, 1967].

В 70-х годах обострилась проблема переноса воздушных загрязнений, преимущественно  $SO_x$  и  $NO_x$ , на большие расстояния (трансграничный перенос), а также проблема озонового загрязнения в урбанизированных и промышленных районах. Поскольку такие типы загрязнений трудно моделировать простым шлейфом, развитие получили новые подходы – Лагранжево и Эйлерово моделирование. Лагранжево моделирование заключается в моделировании траектории загрязненной воздушной массы, при этом предполагается, что свойства воздушной массы не меняются на протяжении всего пути. Такие модели были в основном ориентированы на трансграничный перенос  $SO_2$  [Fisher, 1975; Eliassen and Saltbones, 1975]. Эйлерово моделирование заключается в расчете распространения загрязнений на плоской или трехмерной регулярной сетке. В 70-е годы подобные модели в основном использовались для воспроизведения эпизодов озонового загрязнения на урбанизированных территориях [Reynolds et al., 1967].

Ближе к 80-м годам началась разработка как двумерных моделей с регулярной сеткой глобального масштаба, в которых глобальная тропосфера усреднялась в долготном направлении [Isaksen and Rohde, 1978], так и трехмерных (3D) моделей [Peters and Jovanis, 1979].

В 1970-90-х годах активно проводились исследования по применению статистических моделей для прогнозирования воздушных загрязнений [McCollister and Wilson, 1967; Wolff and Lioy, 1978; Aron, 1980]. Статистические методы прогнозирования довольно просты в вычислительном плане, однако требуют длинного исторического ряда измеренных данных, включая метеопараметры. Основным недостатком таких моделей является то, что они не способны восстанавливать концентрации загрязнителей в периоды аномальных погодных условий, отличающихся от исторического ряда, или в периоды аномальных выбросов. Подробный обзор истории развития систем моделирования и прогнозирования загрязнений воздуха представлен в [Zhang et al., 2012].

С 1990-х годов началось стремительное развитие химических транспортных моделей (ХТМ) [Rufeger et al, 1997; Ohara et al., 1998]. Целями разработки ХТМ, в основном, были прогнозы концентраций О<sub>3</sub> [McHenry et al., 2004; McKeen et al., 2005], позже РМ [McKeen et al., 2007].

Детерминированные трехмерные химические транспортные модели на основе принципа сохранения массы явно решают математические уравнения, описывающие физические и/или химические связи между концентрациями химических веществ в атмосфере и определяющими их атмосферными процессами. Химические транспортные модели, в отличие от предшествующих методов, описывают полный жизненный цикл загрязнителей в атмосфере и учитывают максимально возможное количество факторов, влияющих на загрязнение. Одним из явных преимуществ ХТМ, в том числе перед статистическими моделями, является пространственный охват – ХТМ воспроизводят поля концентраций заданных соединений на регулярной сетке.

Основными компонентами химических транспортных моделей являются:

1. Численный прогноз метеопараметров.
2. Начальные и граничные химические условия, т.е. концентрации соединений.
3. Данные об антропогенных и биогенных выбросах соединений в атмосферу. Выбросы обычно заданы на регулярной сетке.
4. Химический блок, описывающий реакции в атмосфере.
5. Аэрозольный блок, описывающий химию аэрозолей.
6. Блок физических параметризаций, описывающий, например, процессы осаждения загрязнителей.
7. Для онлайн ХТМ – блоки, описывающие аэрозольно-радиационные и аэрозольно-облачные взаимодействия.

Существует несколько классификаций ХТМ. В ХТМ могут использоваться два метода моделирования: Лагранжево и Эйлерово, также существуют гибридные модели, в которых комбинируются оба метода.

По пространственному масштабу ХТМ подразделяются на [World Meteorological Organization, 2020]:

- локальные с шагом сетки < 1 км;
- городские с шагом сетки 1–12 км;
- региональные, или мезомасштабные, с шагом сетки 12–1000 км;
- глобальные с шагом сетки 1000–20000 км.

Одной из современных тенденций развития ХТМ является мультимасштабное моделирование, т.е. использование вложенных сеток с разными разрешениями в одной модели,

либо использование адаптивных нерегулярных сеток [Skamarock et al., 2018]. Мультимасштабные модели разрабатываются с целью учета взаимодействий крупно-, мезо- и мелкомасштабных процессов в атмосфере [Kim et al., 2018]. Современные масштабы моделирования загрязнений сужаются вплоть до внутригородских с разрешением сетки несколько метров и с учетом физики уличных каньонов [García et al., 2009; Santiago et al., 2017; Wolf et al., 2020].

По временному масштабу ХТМ подразделяются на краткосрочные, прогнозирующие атмосферные загрязнения на несколько суток вперед, и долгосрочные, прогнозирующие загрязнения на срок от недели до сезонов. Большинство современных ХТМ являются краткосрочными моделями. Долгосрочные модели также зачастую используются для ретроспективного анализа [World Meteorological Organization, 2020].

По типу соединения с метеорологическими моделями ХТМ подразделяются на «оффлайн» и «онлайн» модели. Оффлайн ХТМ [Sofiev et al., 2006, Redington et. al 2009] используют данные метеопараметров, которые заранее рассчитаны отдельной метеорологической моделью на весь срок прогноза ХТМ. Преимуществом оффлайн ХТМ является возможность проводить эксперименты с вариациями прогнозов метеопараметров разных моделей атмосферы с целью повышения качества прогноза ХТМ.

Онлайн ХТМ жестко связаны с метеорологической моделью, часто под этим термином подразумевается одна модель, которая синхронно прогнозирует атмосферные и химические процессы. Преимуществом таких моделей является то, что на каждом временном шаге прогноза учитываются взаимные обратные связи между содержанием примесей в атмосфере и атмосферными процессами. В основном это касается аэрозольно-радиационных и аэрозольно-облачных взаимодействий [Grell et al., 2000; Grell et al., 2005; Zhang, 2008; Kirsanov et al., 2020].

По типу распространения ХТМ делятся на свободно распространяемые и закрытые модели. В тестировании и доработке свободно распространяемых моделей, как правило, участвуют не только разработчики модели, но и ее пользователи. В разработке и выпуске новых версий закрытых ХТМ, как правило, участвуют только разработчики модели. Такие модели распространяются по лицензии с одноразовыми или регулярными платежами.

С середины 90-х годов началось создание и развитие национальных централизованных систем прогнозирования качества воздуха на основе ХТМ, например, CHRONOS в Канаде [Pudykiewicz et al.], NOAA-EPA в США [Davidson et al., 2008], AAQFS в Австралии [Core et al., 2004]. Также были запущены крупномасштабные интернациональные проекты, например, Европейский открытый портал прогнозирования химической погоды [Balk et al., 2011].

В обзоре [World Meteorological Organization, 2020] приводится список из 73 современных ХТМ, которые используются в мире для моделирования и прогнозирования качества воздуха. Из приведенного списка можно выявить основные современные тренды развития и применения

ХТМ: из 73 моделей 17 являются глобальными, 56 – региональными, либо городскими; 9 глобальных и 23 региональных ХТМ являются мультимасштабными; 11 глобальных и 21 региональная ХТМ являются онлайн-моделями.

Еще одним трендом является ансамблевое моделирование. При ансамблевом моделировании загрязнений в расчет берутся прогнозы нескольких разных ХТМ либо прогнозы версий одной ХТМ с разными настройками. Прогнозы ХТМ в ансамблях комбинируются различными методами – взвешенная линейная комбинация, среднее арифметическое, медиана. Ансамбли зачастую показывают лучшие результаты прогнозов О<sub>3</sub> [Delle Monache et al., 2006a; McKeen et al., 2005] и РМ [McKeen et al., 2007] по сравнению с отдельными ХТМ. Ярким примером является европейский ансамбль CAMS (Copernicus Atmosphere Monitoring Service, Служба мониторинга атмосферы Коперника), состоящий из 11 европейских ХТМ и производящий прогнозы загрязнений для всей территории Европы (<https://atmosphere.copernicus.eu>). Ансамбль проекта MarcoPolo-PANDA из 9 ХТМ используется для прогнозирования атмосферных загрязнений на территории Китая [Petersen et al., 2019]. Интернациональный ансамбль ICAP-MME состоит из 9 ХТМ и производит прогнозы глобального аэрозольного загрязнения [Sessions et al., 2015].

Важно упомянуть, что в последние десятилетия параллельно с развитием систем прогнозирования качества воздуха активно развивается сфера мониторинга качества воздуха в реальном времени. Использование данных измерений концентраций атмосферных загрязнителей для верификации прогнозов ХТМ является общепринятой практикой; применяются автоматические наземные и спутниковые измерения концентраций.

Спутниковые измерения [Kim et al., 2019] характеризуются крупными пространственными (10–100 км) и временными (фиксация 1 раз в день) масштабами, поэтому данный тип измерений может использоваться для верификации прогнозов крупномасштабных ХТМ.

Для верификации ХТМ регионального и городского масштабов, как правило, используются данные наземных автоматических станций контроля загрязнений атмосферы (АСКЗА). Во многих зарубежных странах существуют крупные сети АСКЗА:

- AirNow, более 2500 станций по всей территории США (<https://www.airnow.gov/>);
- AirBase, около 1000 станций на территории Европы (<https://www.eea.europa.eu/data-andmaps/data/aqereporting-2>);
- Китайский национальный центр экологического мониторинга CNEMC (англ. China National Environmental Monitoring Center) – густая сеть измерительных станций, в основном на территории восточного Китая [Wang et al., 2022].

В Российской Федерации сети автоматизированных наземных измерений сосредоточены в крупных городах: в г. Москве сегодня более 70 АСКЗА, в г. Санкт-Петербурге 19 станций, в

Красноярске – 27 станций, в Челябинске – 38 станций, три АСКЗА в Норильске.

Данные измерений классифицируются в соответствии с локацией измерительной станции, а именно с удаленностью измерительной станции от крупных источников антропогенных выбросов. В каждой стране существует собственная система классификации измерений, наиболее общие классы измерительных станций это:

- городские станции располагаются в жилых районах в черте города, но на удалении от крупных автомагистралей и промышленных предприятий;
- примагистральные, или транспортные, станции находятся вблизи крупных автомагистралей;
- промышленные станции контролируют уровни загрязнения в непосредственной близости от промышленных предприятий;
- фоновые станции располагаются на удалении от источников загрязнения – за чертой города или в парковых зонах.

Верификация прогнозов ХТМ проводится путем сравнения измеренной на станции концентрации с прогностическим значением в ячейке модельной сетки, в которую попадает данная станция. Не исключается, что АСКЗА различного типа, например, транспортного и городского, при высокой плотности городской инфраструктуры могут «оказаться» в одной ячейке модельной сетки ХТМ, поскольку даже ХТМ городского масштаба имеют шаг сетки в несколько километров. Учитывая, что фиксируемые на транспортных и промышленных станциях уровни загрязнений отражают воздействие локальных источников и нехарактерны для большей части территории города, как правило, транспортные и промышленные АСКЗА исключаются из верификации прогнозов ХТМ [World Meteorological Organization, 2020].

После обзора общих характеристик ХТМ, приведем конкретные примеры современных национальных и интернациональных систем прогнозирования качества воздуха на основе ХТМ с оценками качества прогнозов.

Ансамбль химических транспортных моделей CAMS создан и поддерживается Европейским центром среднесрочных прогнозов погоды (European Centre for Medium-Range Weather Forecasts, ECMWF) в рамках европейской программы Copernicus (<https://atmosphere.copernicus.eu/>).

В ансамбль CAMS входит 11 ХТМ: CHIMERE (Франция), EMEP (Норвегия), EURAD-IM (Германия), LOTOS-EUROS (Нидерланды), MATCH (Швеция), MOCAGE (Франция), SILAM (Финляндия), DENM (Дания), GEM-AQ (Польша), MONARCH (Испания) и MINNI (Италия). Прогнозы всех перечисленных ХТМ производятся на регулярной сетке с горизонтальным шагом  $0.1^\circ$  (10–20 км), охватывающей всю территорию Европы (рисунок 1.1). Ансамбль считается в каждом узле сетки как медиана прогнозов ХТМ.

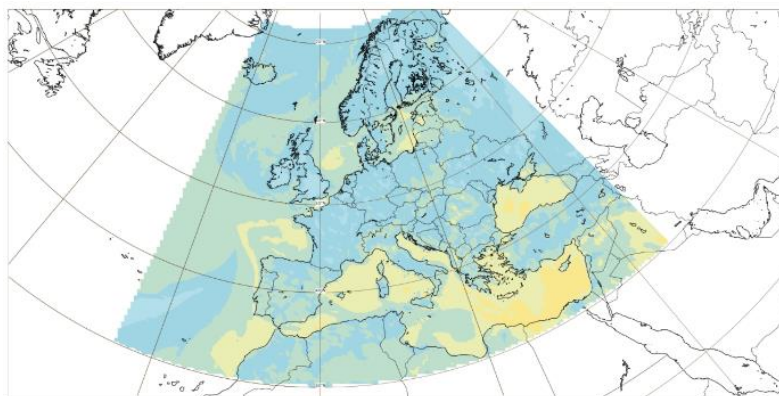


Рисунок 1.1 – Расчетный домен европейского ансамбля CAMS  
(<https://atmosphere.copernicus.eu/>)

CAMS ежедневно производит прогнозы приземных концентраций основных загрязнителей воздуха с временным шагом 1 час и заблаговременностью прогноза 4 суток (96 часов).

Регулярно производится верификация прогнозов CAMS с использованием данных автоматических измерений концентраций сети мониторинга AirBase (около 1000 станций по всей территории Европы). По представленным данным анализа за период 2014–2021 гг. ансамбль CAMS занижает прогноз  $PM_{10}$  в среднем на 20 % и завышает прогноз  $O_3$  на 10 % относительно данных измерений (<https://regional-evaluation.atmosphere.copernicus.eu/>).

В таблице 1.1 представлены среднеквадратические отклонения прогнозов от измерений ( $СКО_{П-И}$ ) и коэффициенты корреляции между измерениями и прогнозами ( $r$ ) концентраций  $O_3$  и  $PM_{10}$  ансамбля CAMS и отдельных моделей ансамбля, по анализу данных в 2023 г.

Таблица 1.1 – Метрики точности прогнозов концентраций  $O_3$  и  $PM_{10}$  ансамбля CAMS и отдельных ХТМ ансамбля, анализ за 2023 г.

Модель	$O_3$		$PM_{10}$	
	$СКО_{П-И}$ , $мкг/м^3$	$r$	$СКО_{П-И}$ , $мкг/м^3$	$r$
1	2	3	4	5
Ансамбль CAMS	12	0,78	6	0,6
CHIMERE	18	0,75	7	0,52
DEHM	18	0,63	6	0,44
EMEP	11	0,77	6	0,58
EURAD	15	0,67	–	–
GEMAQ	17	0,53	7	0,56
LOTOS	11	0,78	7	0,49

Продолжение таблицы 1.1

1	2	3	4	5
MATCH	12	0,76	7	0,45
MINNI	12	0,76	6	0,54
MOCAGE	12	0,7	9	0,34
MONARCH	14	0,62	6	0,55
SILAM	11	0,78	5	0,57

Наименьшие СКО<sub>П</sub>-и концентраций приземного озона показала финская модель SILAM, наибольшие – французская модель CHIMERE. Наименьшие СКО<sub>П</sub>-и концентраций PM<sub>10</sub> также показала модель SILAM, наибольшие – французская модель MOCAGE. Лучшую корреляцию прогнозов концентраций озона с измерениями показали модели SILAM и LOTOS, худшую – модель GEM-AQ. Лучшую корреляцию прогнозов PM<sub>10</sub> с измерениями показала модель EMEP, худшую – модель MOCAGE.

В США система национального прогнозирования качества воздуха NAQFC (англ. National Air Quality Forecasting Capability) рассчитывает почасовые прогнозы концентраций основных загрязнителей воздуха, в т.ч. O<sub>3</sub> и PM, на территории страны (в т.ч. Аляска и Гавайи) на 72 часа вперед (<https://airquality.weather.gov>). Прогнозы рассчитываются в узлах регулярной сетки с горизонтальным шагом 12 км [Campbell et al, 2022]. NAQFC создана и запущена решением Конгресса США в 2004 году, поддерживается Национальным управлением океанических и атмосферных исследований (NOAA) в сотрудничестве с Агентством по охране окружающей среды США (EPA). Основными компонентами NAQFC являются метеорологическая модель GFSv16, разработанная в NOAA, и химическая транспортная модель CMAQv5.3.1, разработанная в EPA. В качестве антропогенных эмиссий загрязнителей используются данные национального кадастра выбросов NEIC2016v1.

Качество прогнозов NAQFC оценивается с использованием данных сети измерений AirNow (2500 станций по всей территории США). К прогнозам NAQFC применяется процедура коррекции отклонений (см. раздел 1.2) с использованием данных измерений AirNow [Huang et al., 2016]. В работе [Campbell et al, 2022] представлены показатели качества почасовых прогнозов NAQFC на одни сутки вперед в сентябре 2020 года для O<sub>3</sub> и в январе 2021 г. для PM<sub>2.5</sub>. Нормализованное среднее отклонение (англ. normalized mean bias, сокр. NMB) прогнозов O<sub>3</sub> варьируется от –16 % до +7,5 % в зависимости от региона; коэффициент корреляции варьируется от 0,66 до 0,82. NMB прогнозов PM<sub>2.5</sub> – от –36 % до +27 %, коэффициент корреляции – от 0,23 до 0,83.



В работе [Benavente et al., 2023] обсуждаются эксперименты по моделированию загрязнения атмосферы в Южной Америке, а именно в агломерации Сан-Паулу (юго-восток Бразилии, население 11,3 млн., самый населенный город южного полушария) и на прилегающих к ней территориях, с применением модели WRF-Chem (англ. Weather Research and Forecasting with Chemistry, Исследование и прогноз погоды с химией). Онлайн ХТМ WRF-Chem создана и поддерживается Национальным центром атмосферных исследований США [Grell et al., 2005]. В WRF-Chem используется блок газофазных реакций MOZART-4 и блок аэрозольной химии GOCART.

Экспериментальное прогнозирование проводилось на 5 суток вперед на сетке с горизонтальным разрешением 10 км для четырех периодов – июнь 2017, июль 2018, июнь 2019 и август 2019. В использованной версии WRF-Chem включены только транспортные выбросы, полученные по данным местных организаций об использовании автотранспорта (местный автопарк, пробег автомобилей, состав топлив и др.) и по данным полевых экспериментов [Martins et al., 2006; Andrade et al., 2015; Nogueira et al., 2021]. Для верификации модельных прогнозов были использованы данные наблюдений 60 станций мониторинга, большая часть которых расположена на территории города Сан-Паулу, и несколько фоновых станций.

Результаты экспериментов показали, что WRF-Chem завышает прогнозы  $O_3$  в среднем на 30–67 % (в зависимости от периода моделирования) относительно измерений, СКО<sub>п-и</sub> составило 21–24 мкг/м<sup>3</sup>. Было установлено, что WRF-Chem достаточно хорошо восстанавливает суточную изменчивость концентраций  $O_3$  – коэффициент корреляции часовых прогнозов оказался в диапазоне 0,72–0,86. Прогнозы  $PM_{2.5}$ , наоборот, занижены моделью WRF-Chem на 39–56% от измерений при СКО<sub>п-и</sub> 14–33 мкг/м<sup>3</sup>; суточная изменчивость  $PM_{2.5}$  воспроизводится WRF-Chem хуже, чем  $O_3$  – коэффициенты корреляции 0,36–0,5.

В последние годы опубликовано много работ китайских исследователей. В работе [Wang et al., 2022] обсуждается прогнозирование загрязнения воздуха взвешенными частицами  $PM_{2.5}$  в регионах Китая, в том числе в крупных городах, зимой 2016–2017 гг., когда в нескольких регионах Китая были зафиксированы эпизоды опасных концентраций  $PM_{2.5} > 600$  мкг/м<sup>3</sup>. Авторы использовали онлайн-модель, состоящую из мезомасштабной метеорологической модели GRAPES\_Meso и химической транспортной модели CUAGE, разработанных метеорологической администрацией Китая (англ. China Meteorological Administration, CMA). ХТМ CUAGE включает в себя блок усвоения эмиссий, аэрозольный модуль CAM и модуль газофазных атмосферных реакций RADM2 (21 фотохимическая реакция, 136 газовых реакции), учтены реакции переходов газ-твердая частица на поверхностях аэрозолей. Для задания эмиссий внутри Китая использовался национальный кадастр антропогенных выбросов MEIC (сетка  $0.1 \times 0.1^\circ$ ), для задания эмиссий на внешних территориях использованы данные глобального кадастра HTAP.

Авторами [Wang et al., 2022] проводилось прогнозирование для зимнего периода 1 декабря 2016 г. – 28 февраля 2017 г. на 3 суток вперед с временным шагом 1 час на сетке с горизонтальным разрешением  $0.1 \times 0.1^\circ$ , охватывающей всю территорию Китая. При моделировании применялась процедура «теплого рестарта» – начальные модельные поля концентраций для каждого прогноза получались путем ассимиляции данных измерений  $PM_{2.5}$  (рисунок 1.2) и смоделированных полей  $PM_{2.5}$  предыдущего дня. Таким образом устранялись кумулятивные ошибки циклического моделирования.

Для верификации модельных прогнозов использовались измерения  $PM_{2.5}$  сети мониторинга CNEMC (рисунок 1.2).

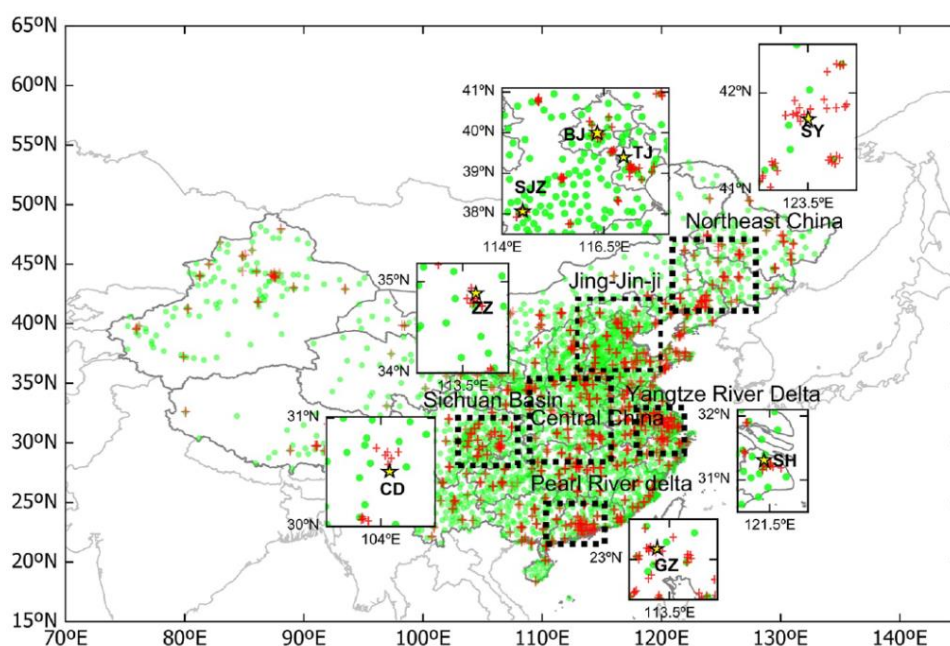


Рисунок 1.2 – Домены экспериментальных расчетов XTM CUACE (пунктирные линии, окна), локации станций метеорологических наблюдений (зеленые точки) и станций измерений  $PM_{2.5}$  (красные маркеры) [Wang et al., 2022]

Проведено несколько экспериментов, в том числе с включением аэрозольно-радиационных и аэрозольно-облачных взаимодействий в XTM. По результатам проведенных экспериментов XTM GRAPES-Meso/CUACE показала удовлетворительную временную согласованность прогнозов  $PM_{2.5}$  с измерениями на протяжении всего периода моделирования, в том числе своевременные прогнозы пиков концентраций. Удовлетворительно спрогнозировано пространственное распределение загрязнений  $PM_{2.5}$  по территории восточного Китая, в т.ч. участки с максимальным загрязнением (рисунок 1.3).

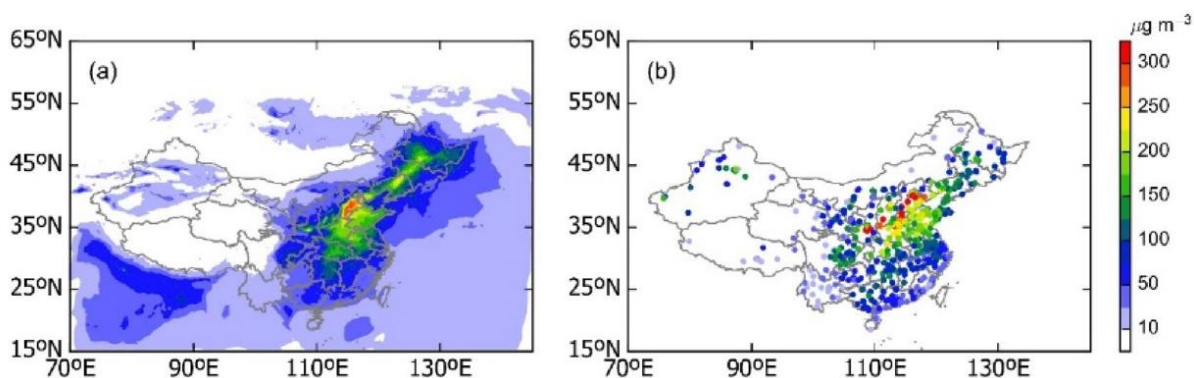


Рисунок 1.3 – Спрогнозированные (слева) и измеренные (справа) концентрации  $PM_{2.5}$  ( $\mu\text{г}/\text{м}^3$ ) в среднем по эпизодам сильного загрязнения  $PM_{2.5}$  на территории восточного Китая зимой 2016–2017 гг. [Wang et al., 2022]

Авторы [Wang et al., 2022] приводят сравнения производительности версий ХТМ с базовыми настройками и с включениями аэрозольно-радиационных и аэрозольно-облачных взаимодействий. Наилучшие результаты показывает модель с включением обоих типов взаимодействий, наибольшее влияние на улучшение прогноза оказывает включение аэрозольно-радиационных взаимодействий.

Еще один эксперимент китайских исследователей по прогнозированию загрязнения воздуха с использованием ХТМ CUAGE в январе, апреле, июле, октябре 2013, 2015, 2017 гг. обсуждается в [Zhang et al., 2021]. В качестве метеорологического блока использовалась модель WRF-ARW, связка ХТМ CUAGE и WRF-ARW производилась на основе блока WRF-Chem и была представлена в оффлайн варианте. Антропогенные выбросы заданы национальным кадастром MEIC с учетом уменьшения выбросов на территории Китая в 2013–2017 гг. Блок газовых реакций RADM2 в ХТМ CUAGE заменен на CBM-Z, содержащий 55 хим. соединений, 114 реакций, 20 фотохимических реакций. Для моделирования аэрозольной химии использовался модуль SAM с обновлением схемы сухого осаждения взвешенных частиц. Домен экспериментальных расчетов ХТМ охватывал территорию восточного Китая, горизонтальный шаг сетки составил 18 км, использованы 32 модельных высотных уровня до высоты 100 гПа.

По результатам экспериментов с использованием модели WRF-ARW/CUACE оказалось, что коэффициент корреляции прогнозов  $PM_{2.5}$  с сетевыми измерениями CNEMC изменялся от 0,59 до 0,71 в зависимости от региона прогнозов, отклонения прогнозов от измерений в основном положительные (от +12,3 % до +21,8 %), но на северо-востоке Китая в регионе Великой Китайской Равнины WRF-ARW/CUACE немного занижала расчеты концентраций  $PM_{2.5}$  – в среднем на ~5 % от измерений. Концентрации  $O_3$  модель WRF-ARW/CUACE рассчитывает удовлетворительно: коэффициент корреляции между прогнозами и измерениями составил 0,6–0,77 в зависимости от региона. При этом в двух из четырех регионов модель в среднем немного

занижает прогноз  $O_3$  по сравнению с измерениями (на 1–8 %) и сильно завышает (до 78 %) расчеты  $O_3$  на юго-западе Китая в регионе межгорной Сычуаньской впадины.

Изученные показатели успешности прогнозирования концентраций приземного озона и твердых взвешенных частиц различных зарубежных ХТМ служат ориентиром при оценках качества модельных расчетов, обсуждаемых в диссертационной работе в разделе 2.3.2 и главах 4, 5.

## 1.2 Методы улучшения качества численных прогнозов воздушных загрязнений

Последовательное и нарастающее усовершенствование численного моделирования загрязнений воздуха и сопутствующих атмосферных процессов пока еще не позволяет полностью избавиться от ошибок в прогнозах. Значительная роль в неточности прогноза загрязнения принадлежит ошибкам прогнозов метеорологических параметров, рассчитываемых моделью атмосферы [Srivastava and Nadèg, 2021]. В свою очередь, к числу главных причин погрешностей модельных прогнозов атмосферных процессов относятся неполнота и неточность знаний о текущем состоянии атмосферы и несовершенство используемых прогностических методов и моделей (<https://meteoinfo.ru/faq>). В численных моделях атмосферы учитываются процессы термогидродинамики, радиационно-облачные взаимодействия, преобразования влаги, сложные процессы в пограничном слое атмосферы и на границе с ее подстилающей поверхностью, и т.д. Некоторые физические процессы в моделях сознательно огрубляются из-за того, что их воспроизведение требует больших вычислительных ресурсов, некоторые другие процессы не учитываются, если они мало влияют на успешность прогноза.

Следует отметить и проблему предсказуемости прогнозов, поскольку даже малые ошибки задания начальных данных для расчета с течением времени могут трансформироваться в значительные. В полной мере это относится и к расчетам ХТМ, которые воспроизводят прогностические поля концентраций с использованием данных о выбросах из различных кадастров, подготовленных на основании предоставленной национальными службами информации и преобразованных в формат регулярной сетки.

Для повышения точности численных прогнозов концентраций ЗВ решаются задачи в двух направлениях; в первую очередь это усовершенствование разработчиками отдельных блоков ХТМ, воспроизводящих физические и химические процессы в атмосфере [Mailler et al., 2023; Menut et al., 2024]. Другими подходами, полезными пользователям, являются препроцессинг входных данных и постпроцессинг рассчитанных ХТМ прогнозов концентраций. На этапе препроцессинга, например, применяется региональная коррекция данных о выбросах или

коррекция типов землепользования, успешность которых показана в работах с участием автора [Шалыгина и др., 2020; Борисов и Шалыгина, 2021; Кузнецова и др., МиГ 2022; Борисов и др., 2023]. Некоторые исследователи определяют выбросы соединений по их содержанию в атмосфере, решая «обратную задачу» [Панасенко и Старченко, 2008]. Одним из распространенных способов улучшения прогнозов ХТМ является метод усвоения данных измерений для задания начальных модельных полей концентраций [Vocquet et al., 2015; Carmichael et al., 2008; World Meteorological Organization, 2020].

Принципиально иным подходом является статистический постпроцессинг – применение статистических методов для обработки рассчитанных ХТМ прогнозов. Часто используется метод «коррекции отклонения» (англ. Bias adjustment, Bias correction), когда, в частности, прогноз ХТМ корректируется средним отклонением от измерений, вычисленным по ряду прогнозов-измерений за определенный период. Также применяется коррекция путем умножения прогностической величины на отношение суммы измерений к сумме прогнозов. Существуют более сложные методы вычисления коэффициентов коррекции прогнозов с привлечением данных наблюдений [Kang et al. 2007; Kang et al., 2010a, Kang et al., 2010b; Huang et al., 2016].

В работах [Delle Monache et al., 2008; Delle Monache et al., 2006b; Kang et al., 2013; Kang et al., 2007] обсуждается применение фильтров Калмана и Колмогорова-Зурбенко для коррекции рядов прогнозов ХТМ в локациях измерительных станций. Применение фильтров значительно повышает точность прогноза, по данным исследователей метод потенциально применим к прогнозам в локациях без измерений.

Некоторые авторы [Potemski and Galmarini, 2009; Johansson et al., 2015; World Meteorological Organization, 2017; World Meteorological Organization, 2020] применяют т.н. технику «смешивания данных» (англ. data fusion). Она заключается в линейном комбинировании данных о загрязнении в определенной локации из различных источников – например, из ХТМ городского масштаба, ХТМ регионального масштаба, наземное измерение концентрации, спутниковое измерение. Прогностические и измеренные значения смешиваются исходя из предположения о том, что измерительный прибор имеет погрешность, которая может быть на том же уровне, что и отклонение прогноза. Параметры линейного уравнения подбираются для каждой локации при помощи различных статистических методов. В линейное уравнение могут добавляться различные внешние данные, имеющие отношение к загрязнению воздуха – данные о транспортной нагрузке, землепользовании, выбросах, метеоданные, демографические и социальные данные, и др.

### 1.3 Применение методов машинного обучения для прогнозирования качества воздуха

Развитие и применение инструментов машинного обучения (МО) в очень широком спектре областей – один из современных мировых трендов. Математические основы, на которых построены самые современные и успешные модели машинного обучения (ММО), были разработаны еще в середине прошлого века. Тот факт, что машинное обучение начало активно развиваться только в последние ~10 лет, объясняется двумя причинами, и обе исходят из принципов работы самих инструментов машинного обучения.

Первая причина – это значительное повышение объема данных за последние 10–20 лет. Поскольку все модели машинного обучения являются статистическими моделями, для эффективной работы они должны усваивать как можно больше данных. В последние 10 лет наблюдается взрывной рост фиксации и распространения (в т.ч. через интернет) видео-, фото-, текстовой и других видов информации. Из-за развития компьютерных технологий и гаджетов сегодня в цивилизованном мире практически каждый человек является потребителем, и, что особенно важно, источником огромного объема информации. Накопленные большие объемы данных создают возможность эффективного обучения и применения статистических моделей.

Вторая причина стремительного освоения методов МО – это существенное повышение вычислительных мощностей, крайне необходимых для обучения моделей. В прошлом веке ограниченные вычислительные мощности являлись стоп-фактором развития методов машинного обучения, поскольку даже один эксперимент по обучению модели занимал дни, а то и недели. А поскольку создание эффективной модели машинного обучения всегда требует тонкой настройки архитектуры и параметров модели под каждую определенную задачу, и, соответственно, проводится множество экспериментальных обучений, разработка таких моделей без больших вычислительных мощностей нерациональна. Современные вычислительные мощности, в особенности параллельные вычисления на графических процессорах, позволяют настраивать даже объемные ММО за адекватное время.

Тренд машинного обучения захватил и область прогнозирования качества воздуха. В работе [Masood and Ahmad, 2021] приведен обзор 90 зарубежных научных публикаций с 2003 по 2021 гг. по теме применения методов машинного обучения для прогнозирования качества воздуха. Заметим, что большая часть публикаций посвящена прогнозу РМ. В обзоре обсуждается применение 4 методов машинного обучения – машина опорных векторов; нечёткая логика; искусственные нейронные сети (ИНС); глубокие нейронные сети (ИНС с более чем 1 скрытым слоем), и приводятся сравнительная статистика успешности применения методов. Оказалось, что

наиболее точные прогнозы концентраций основных ЗВ (PM, O<sub>3</sub>, NO<sub>2</sub>, SO<sub>2</sub>, CO) в среднем по всем приведенным в обзоре [Masood and Ahmad, 2021] публикациям показали глубокие нейронные сети (рисунок 1.4).

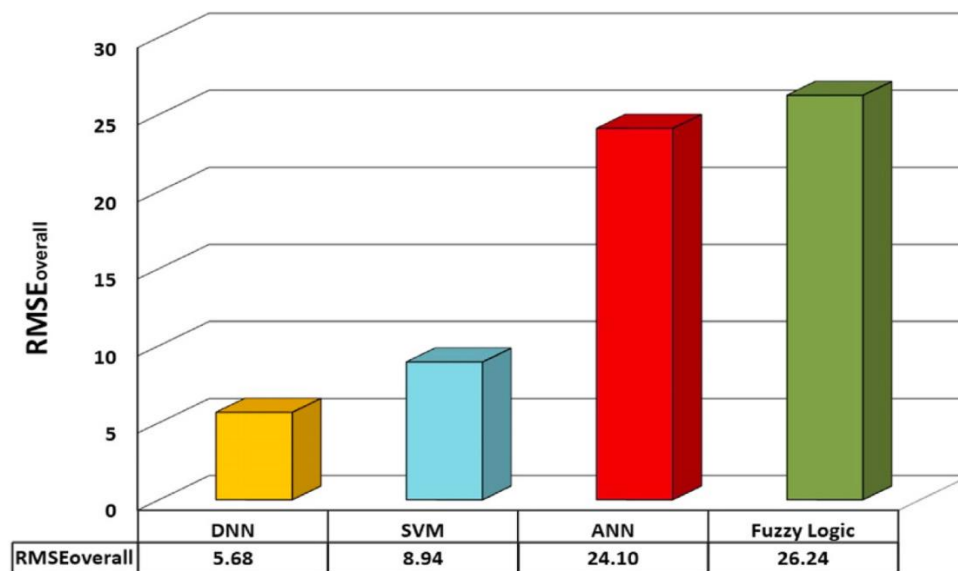


Рисунок 1.4 – Сравнительная статистика успешности применения различных методов машинного обучения для прогнозирования качества воздуха [Masood and Ahmad, 2021]

*Примечание:* RMSE<sub>overall</sub> – среднеквадратическое отклонение прогнозов концентраций (PM, O<sub>3</sub>, NO<sub>2</sub>, SO<sub>2</sub>, CO) типовых моделей от измерений в среднем по всем публикациям в обзоре [Masood and Ahmad, 2021]. DNN – глубокие нейронные сети, SVM – машина опорных векторов, ANN – искусственные нейронные сети, Fuzzy Logic – нечёткая логика.

В частности, по прогнозам концентраций PM глубокие нейронные сети показали в среднем по 6 публикациям коэффициент корреляции прогноз-измерение 0,96, СКО<sub>п-и</sub> 7,27 мкг/м<sup>3</sup>, по прогнозам концентраций O<sub>3</sub> в среднем по 4 публикациям коэффициент корреляции прогноз-измерение 0,92, СКО<sub>п-и</sub> 3,51 мкг/м<sup>3</sup>. Однако необходимо указать на то, что при таких показателях точности прогнозов с применением ММО, речь идет о локальных прогнозах в точках измерительных станций, а в качестве предикторов обучающих выборок ММО в основном использовались измерения концентраций ЗВ и метеопараметров (чаще всего скорость и направление ветра, температура, влажность [Masood and Ahmad, 2021]) на станциях за предшествующий прогнозу период. В данном случае не совсем корректно сравнивать такие прогнозы с прогнозами ХТМ в ячейках модельной сетки.

Авторы обзора [Masood and Ahmad, 2021] делают вывод, что методы машинного обучения, в особенности глубокие нейронные сети, могут производить довольно точный прогноз загрязнений за счет преимущества в виде воспроизведения сложных нелинейных связей между

предикторами и концентрацией загрязнителя. Также авторы обращают внимание на то, что ни один метод машинного обучения не может применяться обособленно для моделирования загрязнений всех масштабов – локального, регионального и глобального, и пишут о необходимости разработки гибридных методов для прогнозирования мезо- и крупномасштабных загрязнений.

В России в последнее время методы машинного обучения все чаще применяются в науках о земле [Krinitskiy, 2024], в том числе в сферах метеорологии, гидрологии, численного прогноза погоды [Чурсин и Косторная, 2024; Захваткина и др., 2024; Солдатенко, 2024; Быков, 2022; Varentsov et al., 2024]. Можно также отметить несколько публикаций с применением ММО для прогнозирования качества воздуха, а именно концентраций  $O_3$  и РМ.

Своего рода прообразом применения методов машинного обучения стала работа [Звягинцев и др., 2010] по созданию статистической модели прогноза максимальной концентрации приземного озона, в которой в качестве предикторов выбраны измеренные и спрогнозированные метеопараметры и концентрации предшественников озона. Статистическая модель была протестирована на данных для московского региона и для ряда немецких станций; модельный прогноз превысил точность «климатического» и «инерционного» прогноза, коэффициент детерминации аномалий приземного озона составил около 0,5. Авторы показали, что наиболее значимыми предикторами статистической модели явились «метеорологический потенциал загрязнения» и отклонения от норм температуры и влажности, а также отклонения от нормы концентрации приземного озона за предыдущие сутки.

Сотрудники Института оптики атмосферы им. В.Е. Зуева СО РАН [Антохин и др., 2013] разработали три статистические модели – 2 регрессионных модели и искусственную нейронную сеть, прогнозирующие среднесуточную концентрацию и дисперсию концентраций приземного озона на следующие сутки на основе ряда измерений атмосферных параметров, в т.ч. концентрации  $O_3$ , на станции в Томском академгородке. Сравнение результатов показало наиболее точный прогноз искусственной нейронной сети, с применением ИНС описано до 70 % общей дисперсии среднесуточной концентрации приземного озона и до 50 % дисперсии среднеквадратического отклонения концентраций  $O_3$ , при этом СКО<sub>П-И</sub> ИНС оказалось в пределах инструментальной погрешности измерений.

В работе [Сметанина и др., 2024] описана модель прогноза приземного озона, построенная на основе методов машинного обучения «случайный лес» и «бустинг», обученная на данных измерений приземного озона на станции Листвянка на оз. Байкал. С помощью полученной ММО и данных реанализа ERA5 было смоделировано и проанализировано содержание озона за 2017–2022 гг. Модель машинного обучения произвела более точный прогноз по сравнению с данными реанализа.



Специалисты компании «Яндекс» [Яндекс.Облако] разработали систему прогнозирования распространения вулканического пепла при извержении. ММО обучена по видеоряду с камер наблюдения, расположенных рядом с вулканом, распознавать начало извержения, высоту и направление облака пепла, класс выброса. Эти данные передаются в траекторную модель FLEXPART, которая рассчитывает рассеяние вулканического пепла на основе метеорологических прогнозов.

Авторы [Мешалкин и др., 2020] разработали однослойные нейронные сети для локаций 3 измерительных станций в г. Тула, предсказывающие концентрации твердых частиц тяжелых металлов. В качестве предикторов обучения ИНС использовались измеренные в предыдущий день метеорологические параметры и средние суточные концентрации твердых частиц. По итогам тестового прогнозирования получены средние абсолютные отклонения прогнозов ИНС от измерений концентраций на станциях от 9,89 до 12,64 мкг/м<sup>3</sup>, коэффициенты корреляции более 0,9.

#### **1.4 Методы интеграции ХТМ и моделей машинного обучения**

Большинство из приведенных в разделе 1.3 моделей обучены с использованием стационарных измерений. При таком подходе корреляционные связи, которые строит статистическая модель по ходу обучения, в итоге отражают только локальные, характерные для определенной измерительной станции, механизмы формирования загрязнений и источники выбросов. Соответственно, обученная на данных измерений статистическая модель способна производить очень точный прогноз, но только для определенной локации. К тому же, точный прогноз ММО на основе измерений предыдущего дня/дней во многом объясняется цикличностью концентраций О<sub>3</sub> и РМ<sub>10</sub> при нормальных условиях. Способность ММО точно спрогнозировать эпизоды повышенных концентраций при аномальных условиях остается под вопросом.

В данном случае за ХТМ остается преимущество перед ММО в виде прогнозов для обширных территорий с учетом пространственной неоднородности загрязнения.

Некоторые авторы используют гибридный подход, т.е. «сплав» детерминированных (ХТМ и моделей атмосферы) и статистических моделей, при котором свойства типовых моделей дополняют друг друга, что позволяет улучшить качество прогнозов загрязнения. В частности, в обучающие выборки ММО включаются прогнозы детерминированных моделей, а целевой переменной прогноза ММО задается измерение концентрации в соответствующей ячейке

модельной сетки. Формально можно сказать, что обученная таким образом модель корректирует детерминированный прогноз концентрации, приближая его к измерению.

Обучение на детерминированных прогнозах может положительно сказаться на универсальности статистической модели, поскольку прогнозы в ячейках сетки отражают более общие условия формирования загрязнений, чем измерение на станции. Обученная таким образом модель потенциально применима для коррекции прогнозов во всех ячейках модельной сетки, в том числе на территориях без измерений концентраций. К тому же, детерминированные прогнозы могут содержать информацию об аномальных атмосферных условиях, при которых ожидается повышение концентраций до опасных уровней.

Ниже приводится обзор публикаций, авторы которых использовали подход комбинирования детерминированных и статистических моделей для повышения точности прогнозов качества воздуха.

Авторы [Honore et al., 2008] применили метод коррекции прогнозов XTM CHIMERE суточного максимума концентрации  $O_3$  в локациях измерительных станций на территории Франции путем обучения линейных регрессионных моделей (для каждой станции) с использованием рядов измерений. Рассчитанные статистическими моделями ошибки прогнозов затем интерполировались по территории модельного домена XTM CHIMERE. В качестве предикторов обучения регрессионных моделей использовались детерминированные прогнозы  $O_3$  и температуры. Такой подход значительно улучшил прогнозы эпизодов повышенного озонного загрязнения ( $O_3 > 180$  мкг/м<sup>3</sup>). Процедура коррекции введена в систему оперативного прогнозирования качества воздуха PREV' AIR (Франция).

Авторы [Perez and Reyes, 2006] построили и сравнили две статистические модели – линейную регрессионную модель и искусственную нейронную сеть (модель нелинейной регрессии), для прогнозирования на следующие сутки максимумов скользящих средних за 24 часа концентраций  $PM_{10}$  в локациях 5 станций мониторинга в г. Сантьяго (Чили). В качестве предикторов для обеих моделей использовались часовые измерения  $PM_{10}$  на станциях в 6 и 7 часов вечера текущего дня, разница между максимумом и минимумом измеренной температуры в текущий день и прогнозов температуры на следующий день, а также прогноз «метеорологического потенциала загрязнения атмосферы». Данные были собраны за периоды апрель–август 2001–2004 гг. Искусственная нейронная сеть показала наибольшую точность прогноза дней с высоким уровнем загрязнения  $PM_{10}$  ( $> 194$  мкг/м<sup>3</sup>) по сравнению с линейной моделью и инерционным прогнозом в 2003 году, в 2004 году инерционный прогноз оказался чуть более точным, чем идентичные по точности прогнозы статистических моделей, учитывая то, что в 2004 году было всего 11 дней с высоким уровнем  $PM_{10}$ . Авторы обращают внимание на то, что ИНС точнее прогнозирует критическое загрязнение  $PM_{10} > 239$  мкг/м<sup>3</sup> на отдельных станциях

мониторинга. По итогу исследования делается вывод, что ИНС может применяться как оперативный инструмент для прогноза качества воздуха в Сантьяго и, при должной адаптации, в других крупных мегаполисах.

В работе [Konovalov et al., 2010] с участием российских исследователей обсуждаются эффекты применения моделей линейной регрессии для прогноза на следующие сутки среднесуточной концентрации  $PM_{10}$  ( $PM_{10-CC}$ ) в локациях станций мониторинга Европейской сети AirBase (около 1000 станций). Регрессионные модели создавались для каждой станции, в предикторы моделей в различных вариациях включались:

- прогноз XTM CHIMERE  $PM_{10-CC}$  на следующие сутки для территории западной и центральной Европы с шагом сетки  $0.5^\circ \times 0.5^\circ$ ;
- прогнозы метеопараметров мезомасштабной модели MM5V3 на той же сетке;
- измерение  $PM_{10-CC}$  на станции за текущие сутки.

Прогнозы регрессионных моделей сравнивались с прогнозами XTM CHIMERE отдельно для теплых и холодных периодов 2003–2006 гг. Важной частью работы [Konovalov et al., 2010] является анализ применимости созданной для типовой станции модели для прогноза на других типовых территориях на удалении от станции. Обученная для станции модель усваивала предикторы, кроме измерений, из локации другой станции того же типа на определенном расстоянии. Либо же рассчитанное локальной моделью отклонение прогноза XTM использовалось для коррекции прогноза XTM в другой локации. Оценивалась динамика улучшения прогноза (по сравнению с прогнозом XTM) с применением локальной модели по мере удаления от станции в среднем для всех возможных пар типовых станций.

Практически все созданные локальные регрессионные модели показали улучшение прогноза по сравнению с прогнозом XTM, максимальное снижения СКО<sub>п-и</sub> составило 50 % для городских станций в холодный сезон, максимальное повышение коэффициента детерминации 85 % для станций всех типов в летний период. Наилучшие результаты показала вариация регрессионной модели с включением всех предикторов (прогнозы и измерения). Показано, что разработанные локальные модели при усвоении предикторов из другой локации улучшают прогноз  $PM_{10-CC}$  по сравнению с XTM на территориях того же типа в радиусе несколько сотен километров, кроме сельских территорий в холодный сезон, для которых улучшение прогноза наблюдается только в пределах 50 км. В теплый сезон применение локальной модели обеспечивает улучшение прогноза XTM на расстоянии вплоть до 300 км для сельских территорий и 800 км для городских территорий, в холодный сезон в радиусе более 500 км для городских территорий.

Исходя из полученных результатов, делается вывод о том, что локальные линейные регрессионные модели применимы для постпроцессинга прогнозов XTM концентраций  $PM_{10-CC}$

на большей части территории западной и центральной Европы, поскольку расстояния между станциями зачастую меньше, чем полученные критические расстояния применимости локальных моделей, особенно в теплый сезон. Комбинирование детерминированных и статистических методов прогнозирования концентраций  $PM_{10}$  обеспечивает значительное улучшение качества прогнозов. Для дальнейших исследований авторы [Konovalov et al., 2010] оставляют вопрос о методе пространственной интерполяции расчетов статистических моделей постпроцессинга.

Авторы [Zhang et al., 2024] разработали и протестировали систему прогнозирования концентраций  $PM_{10}$ ,  $NO_x$ ,  $O_3$  в Стокгольме (Швеция), включающую 3 детерминированные модели – городскую модель гауссова шлейфа, городскую модель уличных каньонов, европейский ансамбль XTM CAMS; и 4 модели машинного обучения – модель случайного леса, модель градиентного бустинга (XGBOOST), рекуррентную нейронную сеть (LSTM) и аддитивную обобщающую модель GAM. Целевой переменной прогноза моделей машинного обучения заданы часовые измерения на 4 станциях мониторинга в Стокгольме. На 3 городских станциях были собраны измерения  $PM_{10}$  и  $NO_x$  за период 10 апреля – 31 декабря 2021 г. – всего по 12 335 измерений, на 1 фоновой станции были собраны измерения  $PM_{10}$ ,  $NO_x$ ,  $O_3$  за период 5 августа – 31 декабря 2021 г. – всего 23 927 измерений.

В качестве предикторов для обучения ММО использованы:

- комбинированный прогноз трех детерминированных моделей концентраций  $PM_{10}$ ,  $NO_x$ ,  $O_3$  в ячейке сетки, соответствующей станции мониторинга;
- детерминированные прогнозы метеопараметров – скорость и направление ветра на 10 м, давление, температура, осадки, облачность, влажность, высота АПС;
- часовые измерения концентраций за предыдущие сутки;
- статистические метрики ряда измерений за предыдущие сутки – среднее, медиана, минимум, максимум, 1 и 3 квартили;
- календарные признаки – юлианская дата, год, месяц, день недели, час.

Результаты тестирования обученных ММО показали улучшение прогнозов концентраций по сравнению с прогнозами детерминированных моделей по всем загрязнителям и станциям мониторинга.

Авторы, с использованием 4 различных методов, провели анализ значимости отдельных предикторов обучения ММО. ММО улучшают прогнозы  $PM_{10}$  за счет измерений за предшествующие сутки и юлианской даты, частично отражающих сезонные вариации  $PM_{10}$ , параметризация которых некачественно задана в детерминированных моделях. Также большой вклад в прогноз ММО  $PM_{10}$  вносят детерминированные прогнозы концентраций и относительной влажности. Для прогноза ММО  $PM_{10}$  на фоновой станции важнейшими предикторами явились детерминированные прогнозы концентраций, относительной влажности, направления ветра и

атмосферного давления. ММО лучше воспроизводят суточный ход концентраций  $\text{NO}_x$  по сравнению с детерминированными моделями за счет предикторов в виде измерений за предыдущие сутки и календарных предикторов – час и день недели, которые отражают изменчивость локальных выбросов автотранспорта. Прогноз  $\text{O}_3$  на фоновой станции улучшен ММО за счет измерений прекурсоров озона  $\text{NO}_x$  за предыдущие сутки, ММО лучше воспроизводит фотохимический источник озона, чем европейский ансамбль XTM CAMS. Другими наиболее значимыми предикторами для прогноза ММО  $\text{O}_3$  явились прогнозы  $\text{O}_3$  детерминированных моделей, максимум измерений  $\text{O}_3$  за предыдущие сутки и детерминированный прогноз относительной влажности.

Авторы отмечают необходимость дополнительных разработок для улучшения качества прогноза в эпизодах повышенного загрязнения, обусловленных неблагоприятными метеословиями и/или аномальными выбросами.

Авторы [Kleinert et al., 2022] разработали и протестировали гибридную модель для прогнозирования концентраций  $\text{O}_3$  на территории Германии, состоящую из онлайн XTM WRF-Chem и модели машинного обучения. В данном исследовании был использован фреймворк MLAir [Leufen et al., 2021], предназначенный для разработки ММО, нацеленных на прогнозирование качества воздуха. Сконструирована искусственная нейронная сеть U-формы; архитектура сети включает сверточный блок и два параллельных рекуррентных LSTM блока. Подобные архитектуры ИНС ранее показывали эффективность при пространственно-временном прогнозировании [He et al., 2022].

Массив целевых переменных прогноза ИНС составили измерения  $\text{O}_3$  332 станций мониторинга качества воздуха в Германии за период январь 2009 г. – март 2010 г. Часовые измерения  $\text{O}_3$  пересчитаны в суточные максимумы скользящих средних за 8 часов концентраций (далее  $\text{O}_{3-8ч}$ ), в соответствии с рекомендациями ВОЗ [World Health Organization, 2021].

Предикторами обучающей выборки ИНС заданы псевдо-наблюдения (авторское [Kleinert et al., 2022] наименование), то есть прогнозы XTM WRF-Chem метеопараметров и концентраций, собранные в ячейках сетки, соответствующих станциям мониторинга. В конфигурации WRF-Chem задана сетка с шагом 12 км, охватывающая всю территорию Германии; эмиссии заданы кадастром TNO-MASS III. Используются прогностические концентрации  $\text{NO}$ ,  $\text{NO}_2$ ,  $\text{O}_3$ ,  $\text{CO}$ , пересчитанные в суточные максимумы средних 8-ми часовых концентраций, и прогнозы приземной температуры, испарения, давления, направления и скорости ветра на 10 м, высоты АПС, осредненные за сутки.

Объем обучающей выборки составил 92 296 пар прогнозы-измерение за период 1 января – 15 октября 2009 г., объем тестовой выборки – 32 204 пар за период 16 декабря 2009 – 31 марта 2010 г. Поскольку в архитектуру ИНС включены рекуррентные блоки (подробнее про

рекуррентные архитектуры в разделе 2.1), на вход ИНС подаются ряды прогностических предикторов за текущие сутки и за 6 предыдущих суток, ИНС рассчитывает прогнозы  $O_{3-8ч}$  на 4 суток вперед.

В работе обсуждаются эксперименты по воспроизведению в ИНС адвекции озона. Базовый эксперимент состоит в обучении ИНС на предикторах в ячейках сетки. В эксперименте 1 в обучающую выборку добавлены пространственно-осредненные предикторы (прогнозы XTM WRF-Chem), попадающие в ветровой сектор, соответствующий направлению переноса в день прогноза, на максимальном удалении 200 км. В эксперименте 2 в обучающую выборку добавлены предикторы из 3 соседних ветровых секторов, соответствующих направлению переноса в день прогноза.

По итогам проведенных экспериментов показано, что добавление в обучающую выборку предикторов из одного ветрового сектора улучшает прогноз  $O_{3-8ч}$  на 4 суток на 10 %, добавление предикторов из трех соседних ветровых секторов улучшает прогноз  $O_{3-8ч}$  на 2 суток на ~14 %. СКО<sub>п</sub>-и ИНС на тестовой выборке по базовому эксперименту составило ~19 мкг/м<sup>3</sup>, по экспериментам 1 и 2 ~13 мкг/м<sup>3</sup>.

Китайские исследователи [Xu et al., 2021] использовали несколько иной подход совмещения детерминированной модели и ММО – они обучили ИНС рассчитывать отклонения прогнозов  $PM_{2.5}$  XTM LOTOS-EUROS от измерений в г. Шанхай (Китай). XTM LOTOS-EUROS производит прогнозы качества воздуха для всей территории Китая, в качестве входных метеопараметров используются прогнозы ECMWF.

Для формирования обучающей выборки ИНС использовались часовые прогнозы концентраций XTM LOTOS-EUROS и метеопараметров модели атмосферы ECMWF на сетке с разрешением 7 км для территории г. Шанхай за период январь–август 2016 г. Период прогнозов ноябрь–декабрь 2016 г. использовался в качестве тестовой независимой выборки. Целевой переменной расчетов ИНС заданы отклонения прогнозов  $PM_{2.5}$  XTM от часовых измерений  $PM_{2.5}$  на 9 станциях мониторинга на территории г. Шанхай.

Конкретные предикторы обучающей выборки включают:

- прогнозы XTM LOTOS-EUROS концентрации  $PM_{2.5}$  в ячейке сетки, соответствующей измерительной станции, за нулевой час расчета ИНС и 9 предшествующих часов;
- прогнозы ECMWF метеопараметров – точка росы, температура, скорость ветра на 10 м – в ячейке сетки, соответствующей измерительной станции, за нулевой час расчета ИНС и 9 предшествующих часов;
- измерения  $PM_{2.5}$ ,  $O_3$ ,  $NO_2$ ,  $SO_2$  и  $CO$  на станции за 9 часов, предшествующих нулевому часу расчета ИНС.

Полученная по итогам экспериментального поиска оптимальной конфигурации ИНС имеет полносвязную архитектуру, включающую 2 скрытых слоя, каждый по 200 нейронов, обучалась на протяжении 50 эпох. Часовой прогноз  $PM_{2.5}$  созданной гибридной модели вычислялся как сумма прогноза ХТМ и расчета отклонения ИНС.

Результаты тестирования гибридной модели показали значительно улучшенный прогноз по сравнению с прогнозом ХТМ, особенно в первые несколько часов (рисунок 1.5).

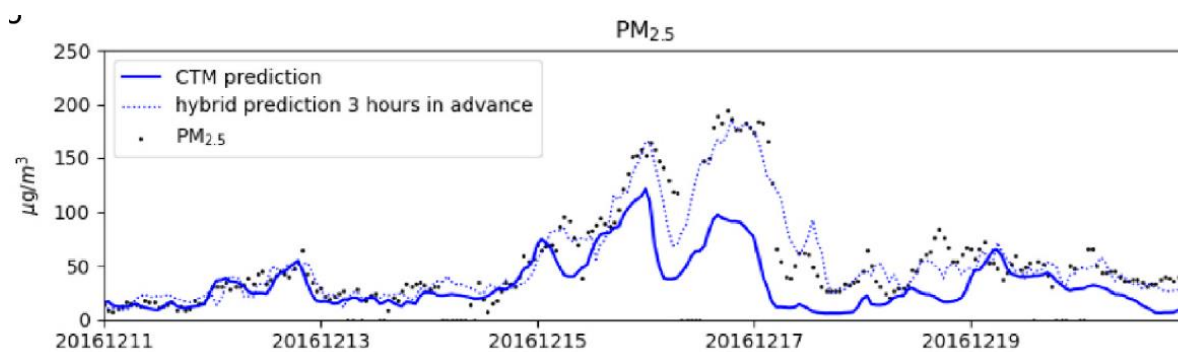


Рисунок 1.5 – Часовые концентрации  $PM_{2.5}$  ( $\text{мкг}/\text{м}^3$ ), спрогнозированные на 3-ий час ХТМ LOTOS-EUROS (синяя линия), ИНС (синяя пунктирная линия) и измеренные на станции Zhangjiang (черные точки). 11–20 декабря 2016 г. [Xu et al., 2021]

Показано, что с заблаговременностью 1 час  $SKO_{\text{П-и}}$  ХТМ LOTOS-EUROS составило  $\sim 28 \text{ мкг}/\text{м}^3$ ,  $SKO_{\text{П-и}}$  гибридной модели  $\sim 9 \text{ мкг}/\text{м}^3$ . С заблаговременностью 3 часа  $SKO_{\text{П-и}}$  гибридной модели  $\sim 16 \text{ мкг}/\text{м}^3$ , с заблаговременностью более 5 часов  $SKO_{\text{П-и}}$  гибридной модели на уровне  $\sim 20 \text{ мкг}/\text{м}^3$ .

Отдельно отмечено, что с применением гибридной модели точность прогноза т.н. нездоровых уровней загрязнения  $PM_{2.5}$  ( $> 75 \text{ мкг}/\text{м}^3$ ) с заблаговременностью 3 часа выросла с 30 % до 54 %. Сделан вывод о необходимости дополнительных разработок для улучшения прогнозов гибридной модели с заблаговременностью более 12 часов.

## Выводы к главе 1

Изучение зарубежного и отечественного опыта прогнозирования качества воздуха показало, что основой современных национальных и интернациональных систем прогнозирования являются химические транспортные модели. Специалисты по моделированию воздушных загрязнений применяют различные методы для повышения качества прогнозов, в том числе статистические методы постпроцессинга прогнозов ХТМ.

В последние годы широкое применение в области прогнозирования качества воздуха получили методы машинного обучения. ММО разрабатываются с использованием данных измерений на станциях мониторинга и применяются для локальных прогнозов, либо интегрируются в системы прогнозирования на основе ХТМ с целью повышения качества прогнозов. При интеграции в системы на основе ХТМ для обучения ММО используются детерминированные прогнозы метеопараметров и концентраций, а в качестве целевой переменной используются измерения концентраций в ячейках модельной сетки. Преимуществом такого подхода является потенциальное улучшение прогнозов по всему расчетному домену ХТМ с применением обученной ММО, в том числе в ячейках модельной сетки без измерений концентраций.

По результатам большинства приведенных в обзоре публикаций при сравнении методов МО наиболее успешный прогноз концентраций ЗВ показывают искусственные нейронные сети, за счет преимущества в виде воспроизведения сложных нелинейных связей между концентрацией ЗВ и влияющими на нее факторами.



## ГЛАВА 2. ИСПОЛЬЗОВАННЫЕ МЕТОДЫ И ДАННЫЕ

Опираясь на зарубежный и отечественный опыт прогнозирования качества воздуха и изученные способы улучшения прогнозов концентраций ЗВ (глава 1), в качестве базового метода машинного обучения выбраны искусственные нейронные сети (ИНС). В разделе 2.1 приводится описание структуры, базовых составляющих, принципов обучения, архитектур и модификаций ИНС.

Для создания ММО важен тщательный подбор, сбор и анализ данных. Основу исследовательской базы данных (БД) для обучения и тестирования ММО составили измерения приземных концентраций  $O_3$  и  $PM_{10}$  на территории московского региона, а также численные прогнозы концентраций ХТМ и метеопараметров модели атмосферы. Для подбора предикторов обучающей выборки ММО необходимо понимание основных физических и химических процессов в атмосфере, оказывающих влияние на формирование загрязнения. На начальном этапе по данным измерений концентраций  $O_3$  и  $PM_{10}$  проведен анализ сезонно-суточной изменчивости концентраций этих ЗВ, а также анализ эпизодов повышенного загрязнения в московском регионе (раздел 2.2). В разделах 2.3 и 2.4 описаны использованные в работе прогнозы метеопараметров модели COSMO-Ru и прогнозы концентраций  $O_3$  и  $PM_{10}$  по расчетам ХТМ CHIMERE. В разделе 2.4.2 представлены систематизированные оценки качества модельных прогнозов концентраций  $O_3$  и  $PM_{10}$  на территории московского региона, служащие обоснованием разработки методики постпроцессинга прогнозов с использованием ММО.

### 2.1 О методе искусственные нейронные сети

Понятие искусственных нейронных сетей возникло из попыток ученых воспроизвести процесс обработки информации головным мозгом в виде математической модели [Хайкин, 2006]. В упрощенном виде структура человеческого мозга представляет собой сеть связанных между собой клеток – нейронов. Нейроны связаны между собой синапсами, через которые передаются сигналы, т. е. мозг можно определить как нейронную сеть. Скорость реакции нейронов на 5–6 порядков меньше скорости реакции кремниевых логических элементов – базовых составляющих компьютерной техники, однако, масса нейронов и количество связей между ними компенсируют эту относительную медлительность. В коре головного мозга насчитывается около 90–96 миллиардов нейронов [Williams and Herrup, 1988] и около 60 триллионов синапсов [Shepherd and Koch, 1990].

### 2.1.1 Структура и основные составляющие искусственных нейронных сетей

На рисунке 2.1 схематично представлена искусственная нейронная сеть. Базовыми составляющими искусственной нейронной сети являются искусственные нейроны (кружки на рисунок 2.1), формирующие слои ИНС. Все нейроны соседних слоев связаны между собой, связи между нейронами характеризуются весами  $w$  – аналог синаптических связей в человеческом мозге. На вход нейрону подается массив входных сигналов – это значения предикторов (в случае нейронов первого слоя) или выходные сигналы нейронов предыдущего слоя. В нейроне вычисляется взвешенная линейная комбинация входных сигналов, к которой прибавляется смещение  $b$  (bias). Последняя операция, происходящая в нейроне – применение к полученному значению функции активации. Функция активации масштабирует значение, обычно в интервалах  $[0, 1]$  или  $[-1, 1]$ , поэтому ее еще называют функцией сжатия. В качестве функции активации обычно используются различные нелинейные функции, например, сигмоида, гиперболический тангенс.

Так, можно формализовать вычисления в искусственном нейроне следующим образом:

$$u_k = \sum_{j=1}^m w_{kj} x_j, \quad (2.1)$$

$$y_k = f(u_k + b_k), \quad (2.2)$$

где  $u_k$  – взвешенная линейная комбинация входных сигналов;

$w_{k1}, w_{k2}, \dots, w_{km}$  – веса нейрона  $k$ ;

$x_1, x_2, \dots, x_m$  – входные сигналы нейрона  $k$ ;

$y_k$  – выходной сигнал нейрона  $k$ ;

$f()$  – функция активации;

$b_k$  – смещение.

Сигналы между нейронами передаются последовательно по слоям, начиная с вектора предикторов и заканчивая последним слоем, в котором рассчитывается прогноз ИНС. В случае регрессионной задачи, когда из набора предикторов требуется спрогнозировать одно значение, выходной слой содержит один нейрон, в котором не прибавляется смещение и не применяется функция активации.

Технически все вычислительные операции в ИНС происходят не последовательно, а в виде операций с матрицами, что значительно увеличивает скорость вычислений. Переходя от понятия искусственного нейрона к понятию слоя ИНС, можно представить обработку информации в ИНС как набор операций над матрицами. На рисунке 2.2 изображен тот же процесс обработки данных в ИНС, что и на рисунке 2.1, только в виде матричных операций.

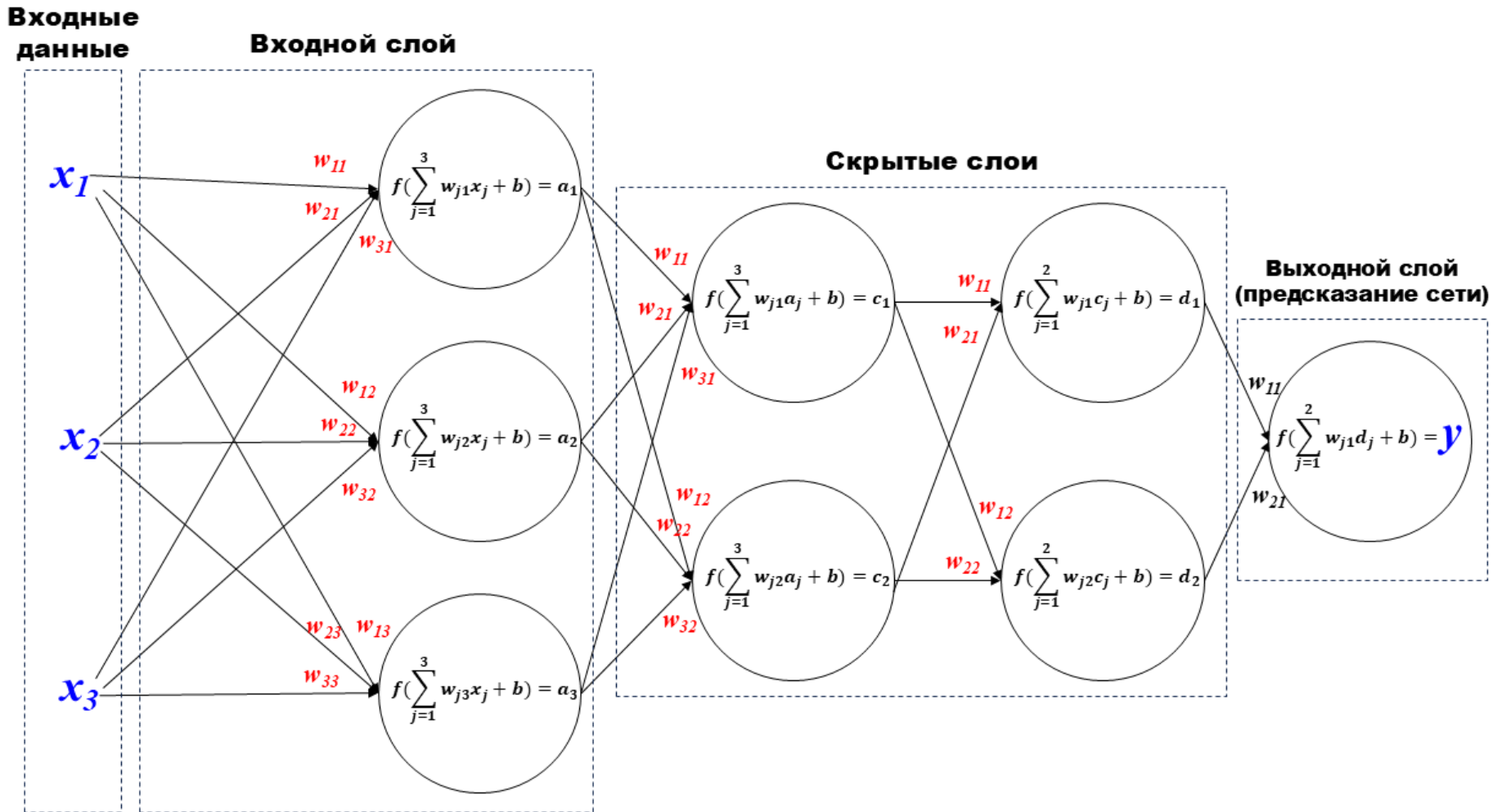


Рисунок 2.1 – Схема полносвязной искусственной нейронной сети

Примечание: Кружками обозначены искусственные нейроны, стрелками – связи между нейронами, которым соответствуют веса  $w$ .

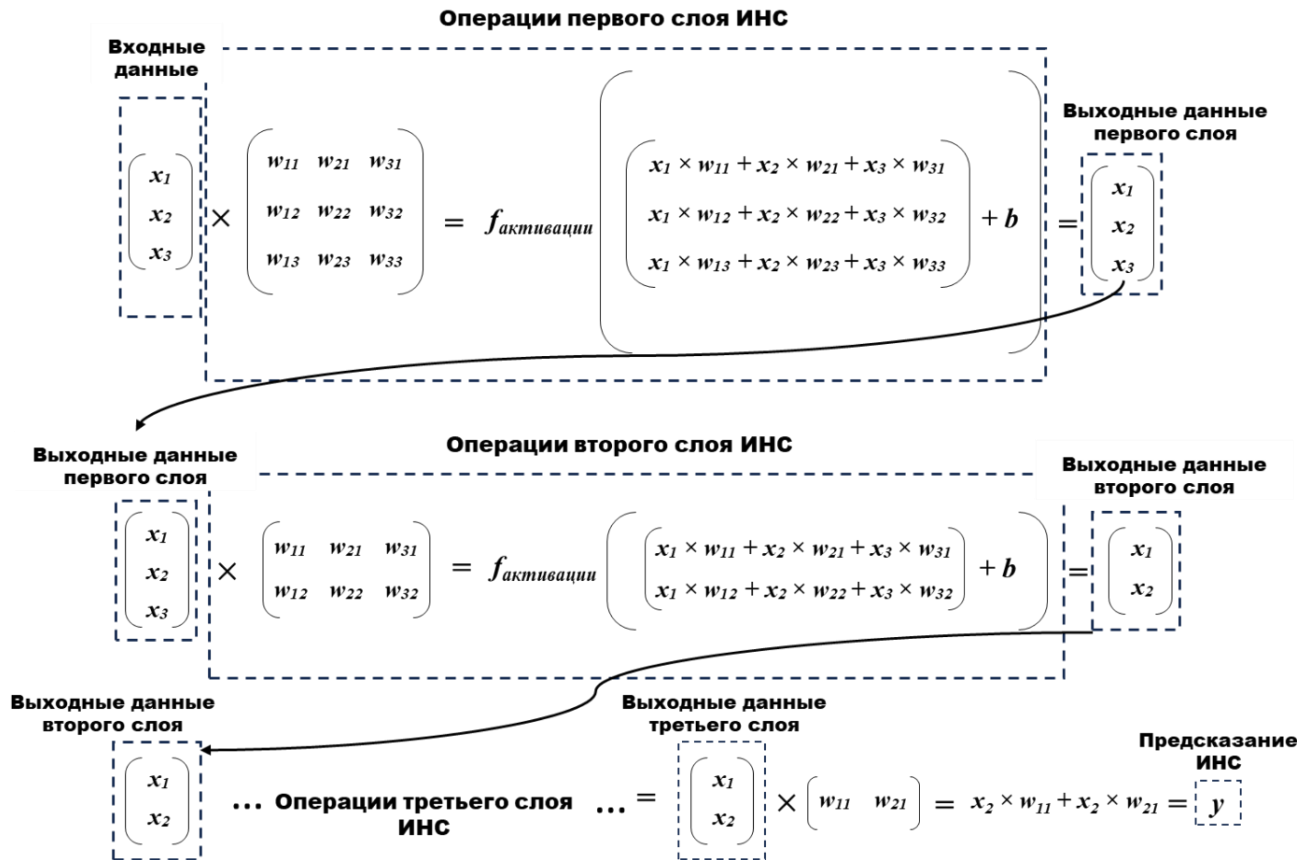


Рисунок 2.2 – Схема обработки данных в ИНС в виде операций над матрицами

*Примечание:*  $x$  – данные,  $w$  – веса в слоях ИНС.

В каждом слое (рис. 2.2) вычисляется скалярное произведение вектора входных данных и матрицы весов формы (количество нейронов в слое, количество нейронов в предшествующем слое). Затем к каждому элементу вектора скалярного произведения прибавляется смещение и применяется функция активации. Операцию в слое ИНС можно записать как:

$$X_{\text{выход}} = f_{\text{активации}}(X_{\text{вход}} \times W^T + b), \quad (2.3)$$

где  $X_{\text{вход}}$  и  $X_{\text{выход}}$  – вектора входных и выходных данных;

$f_{\text{активации}}()$  – функция активации;

$W^T$  – матрица весов;

$b$  – вектор смещения.

Весь процесс вычисления прогноза ИНС с любым количеством слоев и нейронов можно формализовать в общем виде:

$$Y = f(X, W, b), \quad (2.4)$$

где  $Y$  – прогноз ИНС;

$X$  – вектор предикторов;

$W$  – веса синаптических связей в слоях ИНС;

$b$  – смещения в слоях ИНС.

Вычисления в слоях нейросети можно представить как обработанные комбинации исходных предикторов, и можно сказать, что в слоях вычисляются новые представления предикторов, или «новые предикторы». Искусственная нейронная сеть последовательно, путем вычисления новых представлений, преобразует вектор предикторов в значение прогноза. Такая структура и принципы вычислений в ИНС позволяют воспроизводить сложные нелинейные связи между предикторами и целевой переменной прогноза, что является главным преимуществом ИНС.

### 2.1.2 Обучение искусственных нейронных сетей

Созданием искусственной нейронной сети как статистической модели прогнозирования является процесс обучения ИНС. Обучение ИНС состоит в настройке связей в слоях ИНС – весов  $w$  и  $b$  (рисунки 2.1, 2.2).

Для обучения ИНС необходим обучающая выборка. Обучающую выборку обычно формируют в виде матрицы формы ( $n$  наблюдений,  $m$  предикторов), каждому наблюдению (вектору предикторов) должна соответствовать целевая переменная, т.е. в обучающую выборку должен входить массив целевых переменных. Целевой переменной называется значение, которое пытается спрогнозировать ИНС из вектора значений предикторов. В задачах обучения с учителем значения целевой переменной обычно измерены в реальности или оценены специалистом.

Принятой практикой является обучение пакетами данных (англ. batch), так, чтобы за один шаг обучения ИНС обрабатывала не одно наблюдение, а сразу несколько. На каждом шаге обучения на вход ИНС подается пакет данных – матрица формы (размер пакета,  $n$  предикторов). Вычисление прогнозов ИНС по пакету данных называется прямым проходом. После прямого прохода на каждом шаге обучения по пакету данных вычисляется функция потерь путем сравнения прогнозов ИНС и истинных значений целевой переменной (соответствующих пакету данных). В регрессионных задачах в качестве функции потерь часто применяется средняя абсолютная ошибка (англ. mean absolute error, сокр. MAE):

$$MAE = \frac{\sum N_{пакета} |y_{i \text{ истина}} - y_{i \text{ прогноз}}|}{N_{пакета}}, \quad (2.5)$$

где  $N_{пакета}$  – размер пакета данных;

$y_{i \text{ истина}}$  – истинное значение целевой переменной;

$y_{i \text{ прогноз}}$  – спрогнозированное ИНС значение целевой переменной.

Опуская детали, весь процесс обучения ИНС сводится к оптимизации весов  $w$  и  $b$  (рисунки 2.1, 2.2), направленной на минимизацию функции потерь, т.е. ошибки прогноза. Для минимизации функции потерь применяется метод стохастического градиентного спуска. Метод заключается в вычислении на каждом шаге обучения, на т.н. обратном проходе, градиента функции потерь по весам ИНС  $\frac{\partial L}{\partial W}$ , где  $L$  – функция потерь,  $W$  – веса ИНС. Не вдаваясь в подробности, укажем только на то, что, если представить цепь операций в слоях ИНС при прямом проходе как набор вложенных функций (рисунок 2.2), то на обратном проходе градиенты  $\frac{\partial L}{\partial W}$  для матриц весов в слоях вычисляются по правилу расчета производной вложенных функций, то есть перемножением производных отдельных функций [Вейдман, 2021]. Последним этапом шага обучения является обновление весов ИНС с использованием вычисленных градиентов в сторону снижения функции потерь по правилу:

$$W_t = W_{t-1} - \eta \cdot \frac{\partial L}{\partial W}, \quad (2.6)$$

где  $W_t$  – обновленные веса;

$W_{t-1}$  – веса на предыдущем шаге обучения;

$\eta$  – гиперпараметр скорости обучения;

$\frac{\partial L}{\partial W}$  – градиент функции потерь по весам ИНС.

Таким образом, прямой и обратный проходы пакета данных через ИНС составляют один шаг обучения, на каждом шаге обучения происходит оптимизация весов ИНС. Перед обучением задается размер пакета данных и, соответственно, вся обучающая выборка проходит через ИНС за  $n = \frac{\text{Размер обучающей выборки}}{\text{Заданный размер пакета данных}}$  шагов. Проход всей обучающей выборки через ИНС называется *эпохой обучения*. Обычно обучающая выборка «прогоняется» через ИНС несколько раз с перемешиванием данных перед каждой эпохой. После каждой эпохи вычисляются показатели точности прогнозов ИНС по всей обучающей выборке (в регрессионных задачах обычно рассчитывается средняя абсолютная ошибка), что позволяет отслеживать эффективность процесса обучения ИНС. Поскольку ИНС в той или иной степени «подстраивается» под обучающие данные, оценки точности по обучающей выборке не совсем репрезентативны, поэтому в некоторых программных инструментах предоставлена возможность вычисления показателей точности ИНС после каждой эпохи на независимых данных.

Поскольку конечной целью построения любой ИНС является прогнозирование на новых входных данных, для верификации ИНС необходима тестовая независимая выборка. В задачах с временными рядами, как правило, последовательный ряд делится на периоды обучающей и тестовой выборок. Именно показатели точности прогноза на независимой выборке отражают эффективность обученной ИНС.

Обобщенность ИНС – способность прогнозирования на новых данных – является важнейшим целевым свойством ИНС, наравне с точностью прогноза. При обучении ИНС может идеально «подстроиться» под обучающую выборку, т.е. выстроить все заложенные в обучающей выборке связи между предикторами и целевой переменной, включая шумы, тогда точность прогноза ИНС на обучающих данных будет максимальной. Но, в таком случае ИНС не сможет рассчитать точный прогноз при усвоении новых данных, не участвовавших в обучении. Такой эффект называется переобучением модели. Противоположно, эффект недообучения возникает, когда при обучении не были выстроены достаточные для точного прогноза связи между предикторами и целевой переменной, и ИНС показывает неудовлетворительную точность прогноза даже на обучающей выборке [Шолле, 2018].

Так, обучение ИНС всегда нацелено на достижение баланса между точностью и обобщенностью модели. Для этого необходимы:

- максимизация объема обучающей выборки;
- разнообразие обучающей выборки, в случае данного исследования – вариативность условий загрязнения;
- тщательный выбор предикторов обучающей выборки;
- проведение экспериментов по поиску оптимальных архитектуры и гиперпараметров ИНС. Гиперпараметрами называются параметры ИНС, которые не участвуют непосредственно в обучении (как параметры  $w$  и  $b$ ) – это, например, размер пакета данных, количество эпох обучения, скорость обучения (формула 2.6).

Основными преимуществами искусственных нейронных сетей, как статистических прогностических моделей являются [Хайкин, 2006]:

- способность выстраивать сложные нелинейные связи между предикторами и целевой переменной прогноза;
- способность обучаться, т.е. накапливать информацию;
- универсальность и гибкость. Базовая структура ИНС сильно упрощает настройку моделей под разные задачи по сравнению с детерминированными моделями.

Из недостатков ИНС можно выделить:

- требовательность к вычислительным ресурсам. Большие объемы обучающих выборок и объемные архитектуры современных ИНС требуют внушительных вычислительных ресурсов;
- требовательность к формату данных. ИНС может работать только с данными формата предикторы-цель, при определенных задачах приведение данных к такому формату может быть затруднительным;
- неочевидная настройка ИНС. Под каждую определенную задачу необходим поиск архитектуры и гиперпараметров ИНС путем экспериментальных переобучений, что требует

времени и вычислительных ресурсов.

Резюмируем – искусственные нейронные сети используются для прогноза целевого значения из набора предикторов. Создание искусственной нейронной сети заключается в ее обучении – накоплении опытной информации из обучающей выборки путем выстраивания связей между предикторами и целевой переменной. Обучение ИНС происходит путем итеративного вычисления градиентов функции потерь по весам ИНС и обновления весов в сторону минимизации функции потерь (ошибки прогноза). Показателем точности прогноза ИНС служат результаты тестирования ИНС на независимых выборках.

### 2.1.3 Разновидности архитектур и методы улучшения оптимизации искусственных нейронных сетей

Основой функционирования любой ИНС является описанный выше в разделах 2.1.1, 2.1.2 математический аппарат. Существуют вариации архитектур ИНС, использующиеся для специфических задач, и методы улучшения процесса оптимизации ИНС, некоторые из которых, использовавшиеся в данном исследовании, описаны в данном подразделе.

Метод исключения нейронов (англ. Dropout) применяется для решения проблемы переобучения ИНС [Srivastava et al., 2014]. На каждом шаге обучения ИНС на заданных слоях в случайном порядке исключается заданная доля нейронов, то есть сигналы от исключенных нейронов не передаются ни в прямом, ни в обратном направлениях (схематично на рисунке 2.3).

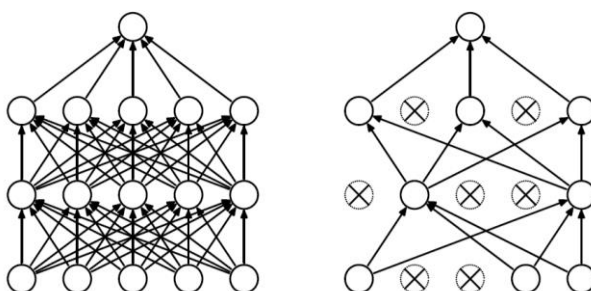


Рисунок 2.3 – Схематичное изображение метода исключения нейронов в ИНС

*Примечание:* слева – стандартная ИНС с 2 скрытыми слоями, справа – пример исключения нейронов (перекрещенные кружки) в слоях ИНС, изображенной слева [Srivastava et al., 2014].

При этом при прогнозе обученной ИНС все нейроны включены. Такой подход сглаживает процесс оптимизации ИНС и не позволяет усвоить шумы, содержащиеся в обучающих данных. Инструменты библиотеки машинного обучения Keras (<https://keras.io/>), использованной в данном



исследовании для построения ИНС, позволяют прописывать исключения не только нейронов в слоях, но и входных предикторов.

Еще одним методом борьбы с переобучением моделей машинного обучения является аугментация данных [Guilhaumon et al., 2024]. Метод аугментации заключается в добавлении шума из нормального распределения в данные обучающей выборки – зашумленные данные снижают риск переобучения и увеличивают обобщающую способность модели. В библиотеке машинного обучения Keras реализован метод аугментации – на каждом шаге обучения в стандартизированные входные данные добавляется гауссовский шум с заданным пользователем стандартным отклонением.

Рекуррентные нейронные сети (англ. Recurrent neural network, сокр. RNN), в том числе подтип LSTM (англ. Long short-term memory, Сети долгой краткосрочной памяти), широко используются в задачах прогнозирования последовательностей – временных рядов, речи, звука [Sak et al., 2014]. В то время как полносвязные ИНС вычисляют прогноз цели на временном шаге  $t$  по набору предикторов на том же временном шаге  $t$ , рекуррентные ИНС обладают свойством «памяти», то есть способны запоминать последовательности предикторов на временных шагах  $(t, \dots, t-n)$  и по ним производить прогноз целевой переменной на шаге  $t$ . Полносвязные ИНС усваивают на каждом шаге оптимизации двумерные массивы формы (размер пакета,  $n$  предикторов), для оптимизации рекуррентных ИНС требуются трехмерные массивы формы (размер пакета, длина последовательности,  $n$  предикторов). Запоминание последовательности в RNN обеспечивает дополнительный массив весов, включенный в архитектуру RNN; подробный обзор рекуррентных, в том числе LSTM, нейронных сетей представлен в [Вейдман, 2021]. Поскольку в данной работе для обучения ИНС используются последовательные часовые прогнозы-измерения концентраций, проводились эксперименты с обучением рекуррентных LSTM нейронных сетей.

Скорость обучения – один из важнейших гиперпараметров процесса оптимизации ИНС, отражающий масштаб обновления весов на каждом шаге обучения в сторону минимума функции потерь (формула 2.6). Если скорость обучения слишком мала, веса могут «застрять» в локальном минимуме функции потерь, а если слишком большая – «перепрыгивать» истинный минимум (обе ситуации схематично изображены на рисунке 2.4 а). Метод затухания скорости обучения [Вейдман, 2021] заключается в постепенном снижении скорости обучения с каждой эпохой. При этом в начале обучения при большой скорости веса стремятся к минимуму функции потерь, не застревая в локальных минимумах, а на последних эпохах при маленькой скорости обучения веса стремятся к истинному минимуму, не «перепрыгивая» его. Существуют различные схемы затухания скорости обучения, наиболее часто используемые – ступенчатая, линейная и полиномиальная (рисунок 2.4 б).

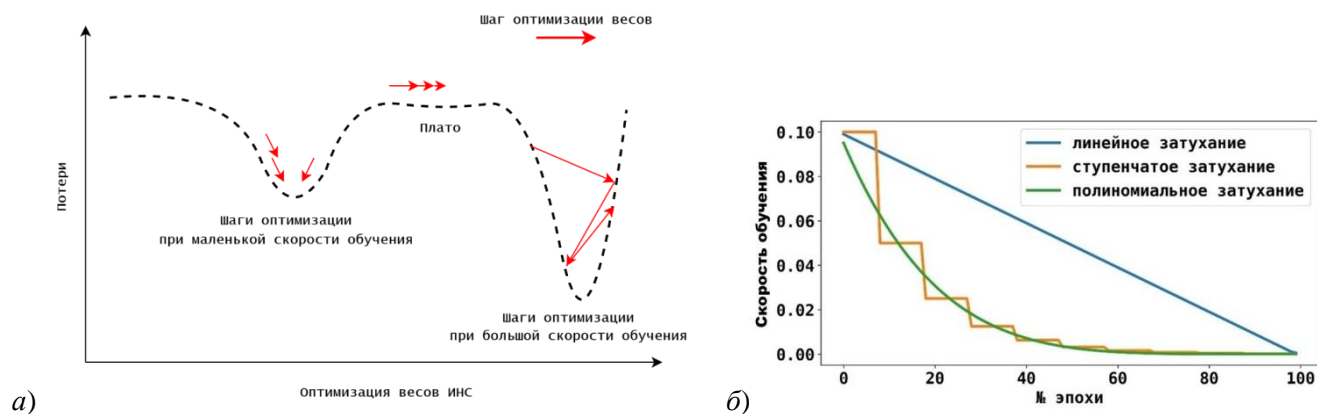


Рисунок 2.4 – Схематичное изображение (а) шагов оптимизации весов ИНС при разных скоростях обучения и градиентах функции потерь; (б) затухания скорости обучения ИНС

Еще одной проблемой, возникающей при оптимизации весов ИНС, является т.н. выход на плато функции потерь (схематично изображено на рисунке 2.4 а), когда градиенты функции потерь по весам ИНС принимают минимальные значения и веса обновляются незначительно или перестают обновляться вовсе, тогда обучение ИНС сильно замедляется или прекращается. Данную проблему может решать метод моментной оптимизации [Вейдман, 2021]. В правило обновления весов (формула 2.6) вводится т.н. импульс:

$$W_t = W_{t-1} - \left( \eta \cdot \frac{\partial L}{\partial W} + \beta * change_{t-1} \right), \quad (2.7)$$

где  $W_t$  – обновленные веса;

$W_{t-1}$  – веса на предыдущем шаге обучения;

$\eta$  – гиперпараметр скорости обучения;

$\frac{\partial L}{\partial W}$  – градиент функции потерь по весам ИНС;

$\beta$  – коэффициент импульса обновления весов;

$change_{t-1}$  – изменение весов на предыдущем шаге обучения.

При моментной оптимизации на каждом шаге обновления весов учитываются обновления весов на предыдущих шагах с коэффициентом импульса  $\beta$  (формула 2.7,  $\beta$  задается пользователем, варьируется от 0 до 1). Тогда при больших градиентах «накапливается импульс» и при выходе на плато, или попадании в локальный минимум, веса продолжают стремиться к истинному минимуму функции потерь.

Результаты применения метода затухания скорости обучения и метода моментной оптимизации при обучении ИНС для постпроцессинга численных прогнозов концентраций приземного озона приведены в авторской публикации по теме диссертации [Борисов и Кузнецова, 2023].

## 2.2 Данные измерений концентраций $O_3$ и $PM_{10}$ в московском регионе

Основу системы мониторинга качества атмосферного воздуха г. Москвы и Московской области составляет сеть из 78 автоматических станций контроля загрязнения атмосферы (АСКЗА) ГПБУ «Мосэкомониторинг» (<https://mosecom.mos.ru>). АСКЗА круглосуточно раз в 20 минут записывают в базу данных измерения концентраций 16 воздушных примесей, в т. ч. концентрации основных загрязняющих веществ – оксид углерода (CO), диоксид азота ( $NO_2$ ), оксид азота (NO), сумма углеводородных соединений ( $CH_x$ ), озон ( $O_3$ ), взвешенные частицы  $PM_{10}$  и  $PM_{2.5}$ , диоксид серы ( $SO_2$ ).

АСКЗА расположены по всей территории города и охватывают следующие функциональные зоны:

- территории, находящиеся под непосредственным влиянием транспортных магистралей («примагистральные» или «транспортные» АСКЗА);
- территории, находящиеся под влиянием выбросов крупных промышленных объектов («промышленные» АСКЗА);
- территории жилых районов на удалении от антропогенных источников загрязнения («жилые» АСКЗА);
- территории, находящиеся под влиянием смешанных источников загрязнения («смешанные» АСКЗА);
- территории, находящиеся за чертой города или в парковых зонах на удалении от источников выбросов («фоновые» АСКЗА).

Таким образом обеспечиваются максимально возможные пространственный охват и репрезентативность измерений концентраций загрязняющих веществ на территории г. Москвы. Помимо Москвы АСКЗА функционируют в 12 городах московской области.

Измерения на станциях осуществляется в соответствии с федеральными требованиями к единству измерений (Федеральный закон «Об обеспечении единства измерений» от 26.06.2008 г. № 102-ФЗ), приборы регулярно калибруются и проходят поверку и межлабораторные сравнительные испытания. По обеспеченности автоматическими станциями, контролируемым параметрам, методам и средствам контроля московская система мониторинга соответствует требованиям российского законодательства и рекомендациям Всемирной организации здравоохранения [World Health Organization, 2021]. В рамках контроля качества данных специалистами «Мосэкомониторинга» регулярно проводятся анализ и обработка поступающих с АСКЗА рядов измерений.

Для измерения концентраций приземного озона используются газоанализаторы Horiba APOA-370, Thermo 49i, ME9810B, Environnement O342M, использующие метод ультрафиолетовой фотометрии. Границы допускаемой относительной приборной погрешности измерения концентрации озона варьируются от  $\pm 15\%$  до  $\pm 20\%$ .

Для измерения концентраций взвешенных частиц PM<sub>10</sub> применяются пылемеры TEOM 1405, TEOM 1400, использующие гравиметрические принципы измерения массы пыли. Допускаемая относительная приборная погрешность измерения концентрации PM<sub>10</sub>  $\pm 20\%$ .

В работе использованы почасовые (за 00 минут каждого часа) измерения концентраций O<sub>3</sub> и PM<sub>10</sub> на АСКЗА ГПБУ «Мосэкомониторинг» за почти 6-летний период 1 января 2019 г. 00:00 – 30 сентября 2024 г. 23:00 – всего 1 058 974 измерений PM<sub>10</sub> и 983 974 измерений O<sub>3</sub>. На рисунке 2.5 представлены АСКЗА, на которых в указанный период измерялись концентрации O<sub>3</sub> (рисунок 2.5 а) и PM<sub>10</sub> (рисунок 2.5 б) с указанием типа АСКЗА и заполненности периода часовыми измерениями (в %).

Основу исследовательской базы данных составили измерения O<sub>3</sub> на 25 АСКЗА, из которых:

- 9 АСКЗА фоновое типа;
- 7 АСКЗА жилого типа;
- 5 АСКЗА примагистрального типа;
- 4 АСКЗА смешанного типа.

А также измерения PM<sub>10</sub> на 32 АСКЗА, из которых:

- 13 АСКЗА смешанного типа;
- 9 АСКЗА жилого типа;
- 5 АСКЗА фоновое типа;
- 5 АСКЗА примагистрального типа.

Большая часть АСКЗА, измерения которых включены в исследование, расположена на территории г. Москвы, в том числе Новой Москвы (Троицк, Рогово, Семеново), несколько АСКЗА расположены на территории городов московской области – Звенигород, Павловский посад, Подольск, Красногорск, Одинцово.

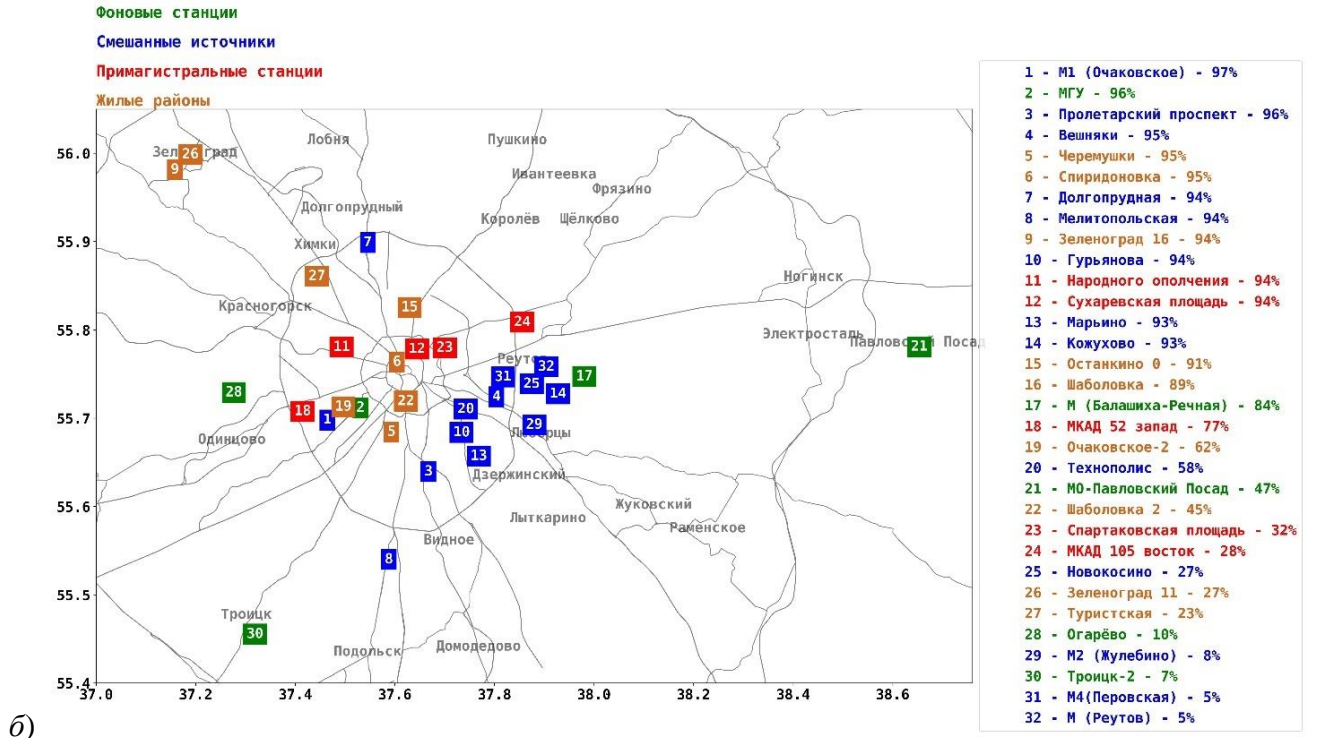
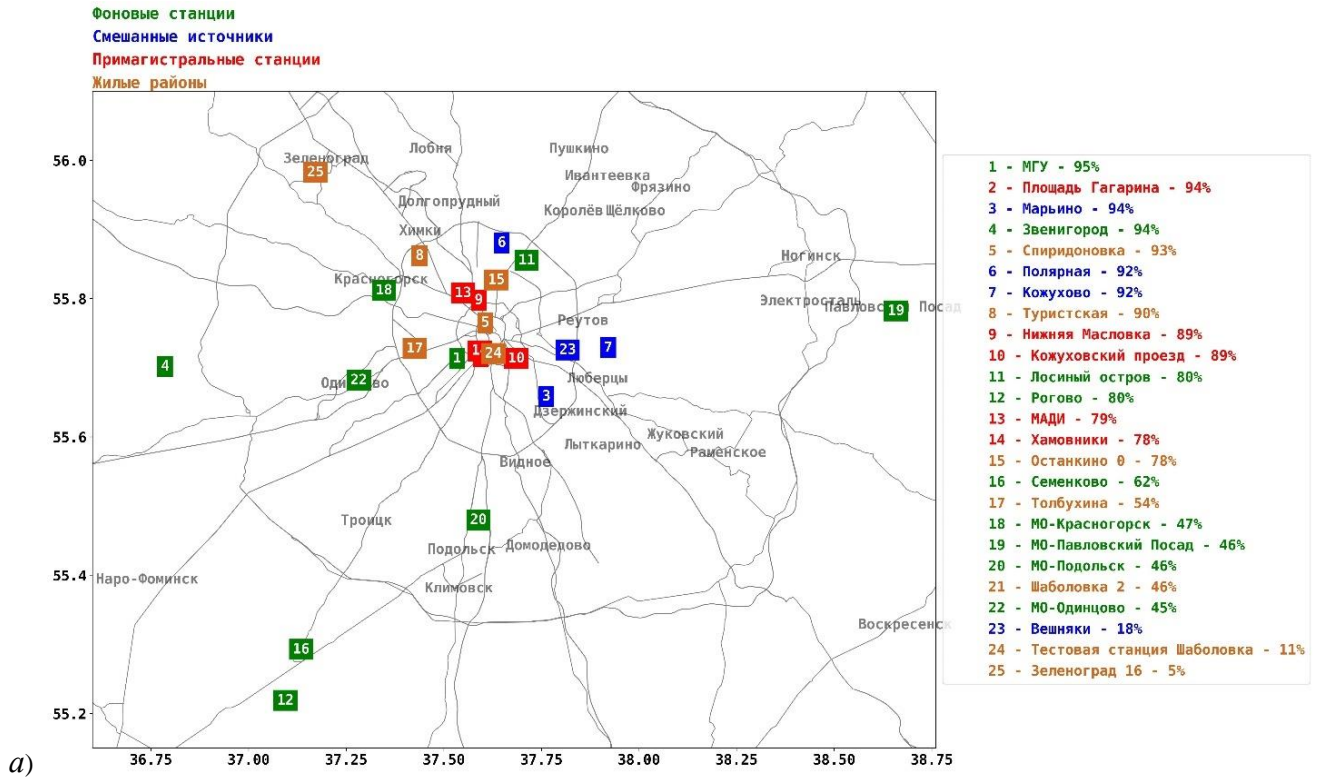


Рисунок 2.5 – АСКЗА, измеряющие концентрации  $O_3$  (а) и  $PM_{10}$  (б)

в период 1 января 2019 г. 00:00 – 30 сентября 2024 г. 23:00 с указанием заполненности периода почасовыми измерениями (в %)

*Примечание:* Желтым цветом – жилые АСКЗА, красным – примагистральные АСКЗА, синим – смешанные АСКЗА, зеленым – фоновые АСКЗА.

### 2.2.1 Сезонная и суточная изменчивость концентраций приземного озона

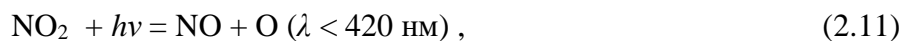
Озон ( $O_3$ ) – это, при нормальных условиях, светло-голубой газ. Содержание озона в приземном воздухе определяется двумя основными механизмами – это приток озона из стратосферы, и фотохимическая генерация озона в приземном слое в присутствии соединений-прекурсоров оксидов азота  $NO_x$  и летучих органических соединений (ЛОС) [Белан, 2010]. Основной объем атмосферного озона генерируется в стратосфере на высоте 20–25 км под воздействием ультрафиолетового солнечного излучения путем фотолиза:



где  $\lambda$  – длина волны.

В результате стратосферно-тропосферного обмена часть стратосферного озона поступает в тропосферу; наиболее интенсивный приток озона в приземный слой происходит в зонах атмосферных фронтов и струйных течений.

Фотохимическая генерация озона в приземном слое происходит в ходе следующих реакций:



где  $\lambda$  – длина волны;

$M$  – любая третья молекула, которая абсорбирует энергию реакции.

Усиливать фотохимическую генерацию приземного озона путем окисления  $NO$  до  $NO_2$  могут пероксильные радикалы ( $RO_2\cdot$ ,  $HO_2\cdot$ ), которые образуются при окислении ЛОС в атмосфере. Основным источником  $NO$  и ЛОС в мегаполисах являются транспортные выбросы [Белан, 2010].

Деструкция (сток) озона происходит на подстилающей поверхности и объектах инфраструктуры. Разрушение озона наблюдается также в сильно загрязненном городском воздухе. Надо заметить, некоторые растения способствуют озоновому насыщению: так, хвойные смолы, испаряющие ЛОС (изопрены, терпены и т.п.), способствуют генерации озона.

На рисунке 2.6 представлены рассчитанные автором средние месячные величины концентрации озона на территории московского региона по измерениям АСКЗА ГПБУ «Мосэкомониторинг» за 5-летний период 1 января 2019 г. – 31 декабря 2023 г. Как видно, в холодный период года приземные концентрации озона минимальны (16–36  $мкг/м^3$ ), как следствие ослабленного вертикального обмена приземного воздуха с вышележащими слоями

тропосферы из-за малой высоты АПС и практического отсутствия фотохимической генерации озона при недостатке УФ-радиации.

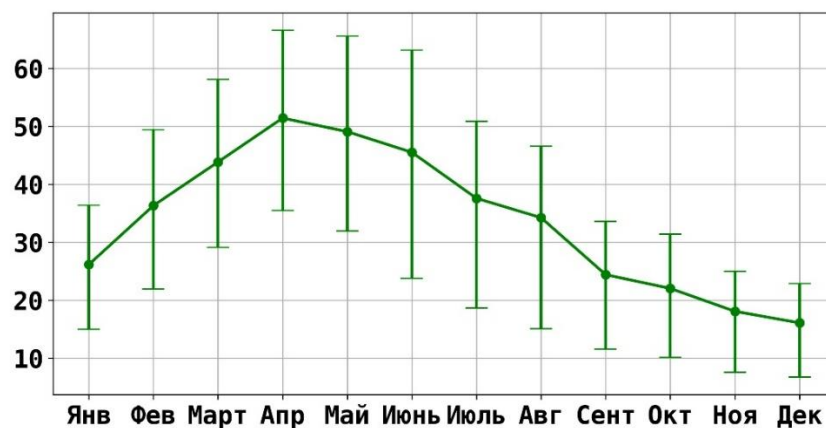


Рисунок 2.6 – Среднемесячные концентрации  $O_3$  (мкг/м<sup>3</sup>) на территории московского региона по измерениям АСКЗА ГПБУ «Мосэкомониторинг» за период 1 января 2019 г. – 31 декабря 2023 г.

*Примечание:* Вертикальные линии – межквартильный размах распределения концентраций.

Годовой основной максимум концентраций озона приходится на апрель-май. Отметим, что его величина (около 50 мкг/м<sup>3</sup> в среднем за 2019–2023 гг.) оказалась примерно на 5 мкг/м<sup>3</sup> больше, чем по расчетам за период 2005–2020 гг. [Еланский и др., 2022]. В этот период миграция высотной фронтальной зоны над умеренными широтами, частое перемещение высотных струйных течений способствуют интенсивному перемешиванию в атмосфере и выносу стратосферного озона в нижние слои, в том числе благодаря активизации конвективного перемешивания. Можно утверждать, что годовой максимум концентраций приземного озона весной в основном обусловлен природным источником – притоком стратосферного  $O_3$ .

Формирование вторичного летнего максимума концентраций озона (34–45 мкг/м<sup>3</sup>) в Москве значительно зависит от преобладающих крупномасштабных процессов и формируется не каждый год [Андреев и др., 2023]. При аномальных погодных условиях летом (волны жары, застойные процессы в блокирующем антициклоне, адвекция озона из районов лесных пожаров и т.п.) концентрации приземного озона могут значительно превышать не только характерные для свободной тропосферы значения (на верхней границе АПС 110–140 мкг/м<sup>3</sup>), но и критерии безопасных уровней, как, например, летом 2010 г., когда приземный озон в Москве превысил более чем в 2 раза ПДК м.р., равную 160 мкг/м<sup>3</sup> [Звягинцев и др., 2011; 2016]. Летом дополнительным источником озона служит его интенсивная фотохимическая генерация в загрязненном выбросами автотранспорта воздухе мегаполиса–при высокой ультрафиолетовой

облученности днем. При этом фотохимическая генерация озона происходит во всем АПС за счет выноса  $\text{NO}_x$  наверх при интенсивном вертикальном перемешивании. Количество образования «антропогенного» озона может оказаться сравнимым с природным только в самых неблагоприятных для очищения городского воздуха метеорологических условиях; в таких случаях говорят о фотохимическом «смоге».

В теплый сезон приземный озон характеризуется выраженной суточной цикличностью (рисунок 2.7 а). Ярко-выраженный дневной максимум  $\text{O}_3$  (14–16 ч) совпадает с максимальным развитием конвективного/вертикального перемешивания и формируется примерно на 6 часов позже утреннего максимума городского загрязнения и на 2–3 часа позже максимума УФ-радиации.

В холодный период по самым грубым оценкам приемного озона почти в 2–3 раза меньше, чем летом; в суточном ходе проявляется ночной (3–5 ч) максимум концентраций озона (рисунок 2.7 б), а в зимние месяцы амплитуда осредненного суточного хода практически нивелируется.

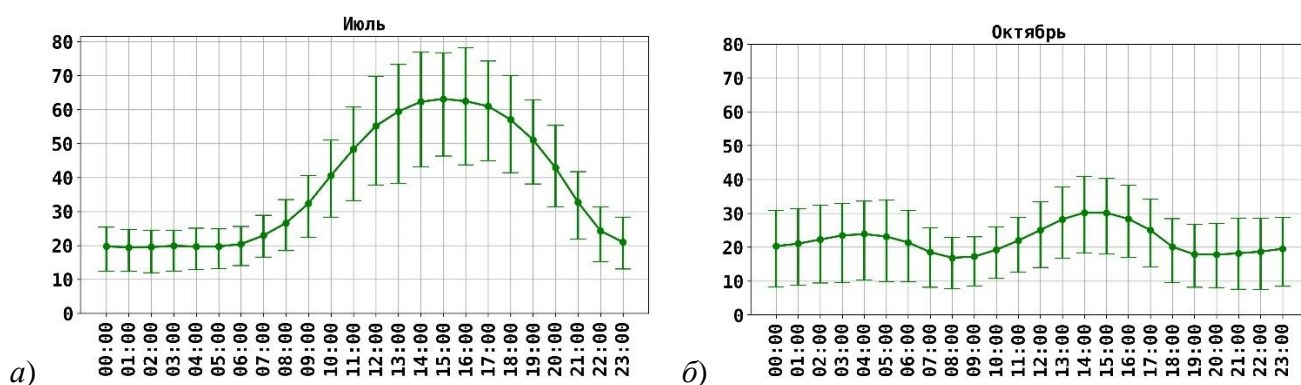


Рисунок 2.7 – Среднечасовые концентрации  $\text{O}_3$  ( $\text{мкг}/\text{м}^3$ ) в июле (а) и в октябре (б) на территории московского региона по измерениям АСКЗА ГПБУ «Мосэкомониторинг» за период 1 января 2019 г. – 31 декабря 2023 г.

*Примечание:* Вертикальные линии – межквартильный размах распределения концентраций.

Средние характеристики приземного озона по наблюдениям в период 2019–2023 гг. в целом совпадают с опубликованными ранее результатами мониторинга озона в московском регионе [Еланский и др., 2007; 2022]. Эпизоды повышенных концентраций озона, опасных для здоровья населения, с превышениями предельно допустимой концентрации (ПДК) в Москве формируются весной-летом в периоды жаркой и сухой погоды. В такие периоды днем  $\text{O}_3$  интенсивно генерируется фотохимическим путем в загрязненном воздухе мегаполиса; в главе 4 обсуждаются условия формирования нескольких эпизодов повышенных концентраций озона в 2024 г.



### 2.2.2 Сезонная и суточная изменчивость концентраций $PM_{10}$ , эпизоды повышенного аэрозольного загрязнения

Концентрация  $PM_{10}$  названа ВОЗ одним из главных показателей качества воздуха [World Health Organization, 2021]. Подчеркнем, что  $PM_{10}$  – это смесь (в отличие от  $O_3$ ) разных видов частиц органического и неорганического происхождения, включая пыль, пыльцу, сажу, дым, химикаты, сульфаты, минеральную пыль.

Источники  $PM_{10}$  отличаются от источников приземного озона, хотя они также разделяются на природные – ветровая эрозия почвы, пыление вспаханных сельхозугодий, вулканическая пыль, морская соль и т.п. – и антропогенные. К числу последних относятся выбросы мелких частиц промышленных предприятий и транспорта, пыление при дорожных, строительных работах и с открытых городских поверхностей [Еланский и др., 2022]. Вторичные  $PM_{10}$  могут образовываться в атмосфере в результате реакций с участием тяжелых металлов,  $SO_x$ ,  $NO_x$ , кислотных и мутагенных активных частиц [WHO Regional Office for Europe, 2003].

Одним из факторов загрязнения воздуха  $PM_{10}$  в московском регионе является дальний перенос почвенных частиц из пустынных районов Африки и Азии. Большое число зарубежных публикаций (например, [Cesari et al., 2018; Galindo et al., 2018]) посвящено выносу песчаного аэрозоля из Африки в Европу, который приводил к сильному ухудшению видимости, нарушениям работы транспорта и стал причиной роста респираторных заболеваний [Middleton, 2019]. Под воздействием выноса песчаного аэрозоля из центральноазиатских пустынь и Малой Азии, опустыненного Приаралья нередко оказываются южные регионы нашей страны. В некоторых ситуациях почвенные частицы, перемещаясь с воздушными массами над «черными» землями и степями Прикаспия, приводят к значительному росту  $PM_{10}$  в умеренных широтах России и даже в северных регионах [Gubanova et al., 2024].

Уровень загрязнения  $PM_{10}$  сильно зависит от погодных условий, в первую очередь, от скорости переноса и термической устойчивости. При интенсивном вертикальном перемешивании частицы выносятся в верхние слои атмосферы, при сильном ветре – рассеиваются. Основными механизмами удаления  $PM_{10}$  из атмосферы являются сухое и влажное осаждение. При повышенной относительной влажности атмосферы вода конденсируется на поверхности частиц с образованием т.н. гидрометеоров, которые коагулируются и утяжеляются, ускоряя влажное осаждение частиц. Осадки могут резко понизить концентрации  $PM_{10}$ , вымывая их из приземного слоя [Demuzere et al., 2009; Varnpadimos et al., 2012].

Данные мониторинга  $PM_{10}$  на станциях отражают влияние локальных источников и формируемое крупномасштабными атмосферными процессами аэрозольное загрязнение.

Преобладающим фактором формирования уровня  $PM_{10}$  в городской атмосфере является поступление антропогенных частиц, но в отдельные периоды/сезоны природные источники существенным образом влияют на колебания приземной концентрации  $PM_{10}$ .

Расчитанные автором по данным измерений на АСКЗА за период 2019–2023 гг. средние месячные концентрации  $PM_{10}$  (рисунок 2.8) практически совпали с приведенными в [Еланский и др., 2022] величинами, полученными по данным измерений на АСКЗА за период 2005–2020 гг.; следует отметить и совпадение формы годового хода  $PM_{10}$ .

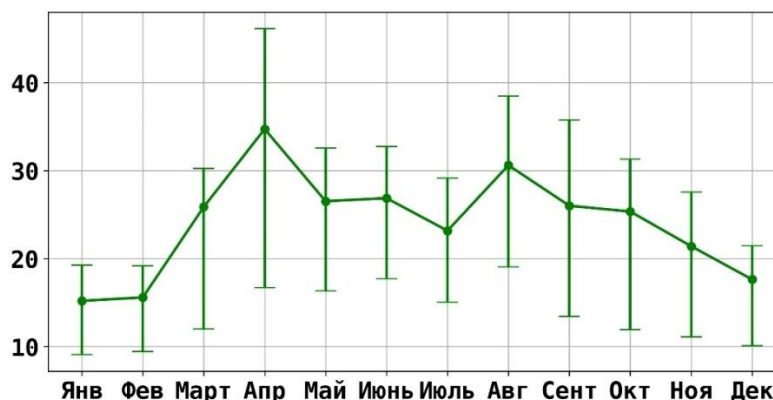


Рисунок 2.8 – Среднемесячные концентрации  $PM_{10}$  (мкг/м³) на территории московского региона по измерениям АСКЗА ГПБУ «Мосэкомониторинг» за период 1 января 2019 г. – 31 декабря 2023 г.

*Примечание:* Вертикальные линии – межквартильный размах распределения концентраций.

Как видно на рисунке 2.8, годовой максимум  $PM_{10}$  (около 35 мкг/м³) приходится на апрель – период схода снежного покрова, когда ветровая эрозия почвы в отсутствие плотного травяного покрова становится решающим фактором аэрозольного загрязнения.

Средние месячные концентрации  $PM_{10}$  в летние месяцы колеблются в пределах 24–30 мкг/м³, аэрозольное загрязнение снижается в периоды дождливой ветреной погоды и достигает максимума при устойчивой жаркой и сухой погоде. Летом чаще чем зимой наблюдаются слабые скорости ветра, приводящие к накоплению  $PM_{10}$  в приземном слое атмосферы.

В холодный период снежный покров блокирует поверхность почвы и препятствует поднятию дорожной пыли, также зимой проводится меньше строительных и дорожных работ. Низкий уровень  $PM_{10}$  с ноября по февраль (рисунок 2.8) является, в первую очередь, следствием уменьшения источников загрязнения.

Внутрисуточная изменчивость концентраций  $PM_{10}$ , определяемая влиянием источников и метеорологическими условиями рассеивания, имеет ярко выраженные сезонные различия. Зимой

на низком фоне суточные колебания  $PM_{10}$  незначительны (рисунок 2.9 а). Весной наблюдается вечерний (~21–23 ч) и утренний (~ 8–10 ч) максимумы концентраций  $PM_{10}$  (рисунок 2.9 б), причем вечерний максимум значительно превышает утренний. Вечерний максимум начинает формироваться с заходом солнца (около 19 час) в условиях усиления термической устойчивости с участием оседания частиц из пограничного атмосферного слоя, куда они переносятся благодаря дневной конвекции.

Летом, благодаря короткому периоду термической устойчивости из-за длительного светового дня, значимое повышение уровня  $PM_{10}$  наблюдается в среднем в период с 22 час до 2–3 часов ночи (рисунок 2.9 в), и этот суточный максимум значительно меньше весеннего вечернего максимума (на 15–20 мкг или 30–45 %). Максимально развитая дневная конвекция в летние месяцы обеспечивает и более низкий уровень  $PM_{10}$  в последполуденные часы (рисунок 2.9 в) по сравнению с весной.

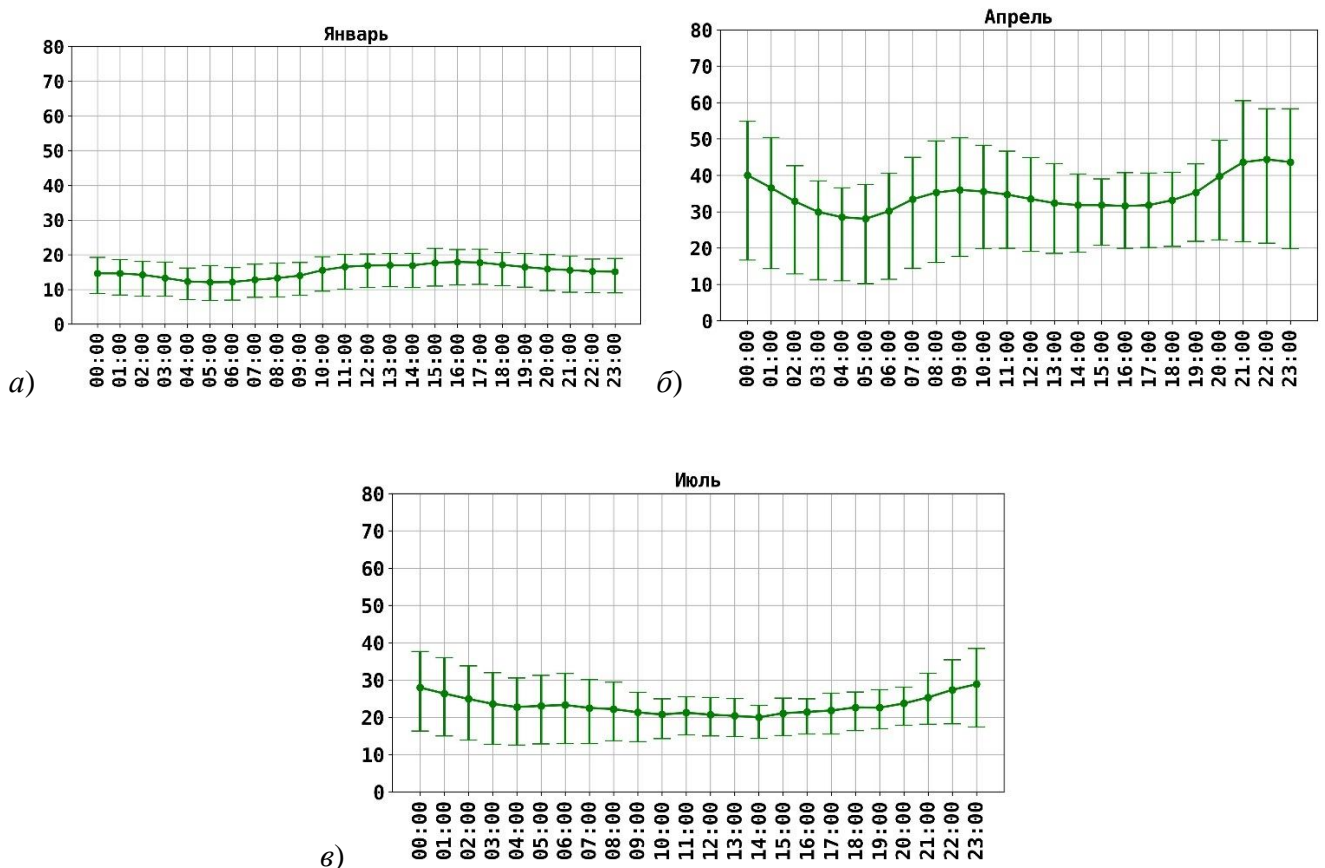


Рисунок 2.9 – Среднечасовые концентрации  $PM_{10}$  ( $\mu\text{г}/\text{м}^3$ ) в январе (а), апреле (б) и июле (в) на территории московского региона по измерениям АСКЗА ГПБУ «Мосэкомониторинг» за период 1 января 2019 г. – 31 декабря 2023 г.

*Примечание:* Вертикальные линии – межквартильный размах распределения концентраций.

Эпизоды повышенного загрязнения  $PM_{10}$  в московском регионе принято классифицировать по двум основным группам источников взвешенных частиц – локальные и

удаленные источники. Повышенное загрязнение  $PM_{10}$  локальными источниками в московском регионе зачастую происходит весной в периоды схода снежного покрова – в такие периоды, до появления растительности, подстилающая поверхность открыта и при усилении ветра происходит интенсивный подъем почвенных частиц с поверхности. Накоплению частиц в уже загрязненном приземном слое способствуют локальные неблагоприятные для рассеивания метеорологические условия (НМУ) – слабый ветер, низкая влажность, отсутствие осадков, термическая устойчивость в приземном слое, вечерние-ночные температурные инверсии, блокирующие вынос частиц в верхние слои атмосферы.

При этом нужно отметить: несмотря на то, что безветренная погода считается фактором накопления  $PM_{10}$  в приземном слое атмосферы, большие скорости порывов ветра, например, перед прохождением холодного фронта, могут также приводить к резким всплескам концентраций  $PM_{10}$  до опасных значений за счет поднятия пыли с поверхности и ветровой эрозии почвы. Будем называть такие случаи «предфронтальными» повышениями концентраций  $PM_{10}$ .

Другой тип – загрязнение от удаленных источников, связан с дальним переносом частиц из районов пыльных/песчаных бурь, в частности, в Прикаспии (рисунок 2.10). При прохождении над засушливыми районами активного холодного атмосферного фронта в нижние слои атмосферы поступает большое количество почвенных частиц, часть их оседает в ближней зоне, а самые мелкие частицы ( $PM_{10}$  и размером меньше) в антициклональных условиях при наличии мезоструйного течения нижнего уровня могут за сутки-двое переноситься на тысячи километров [Виноградова и др., 2024].

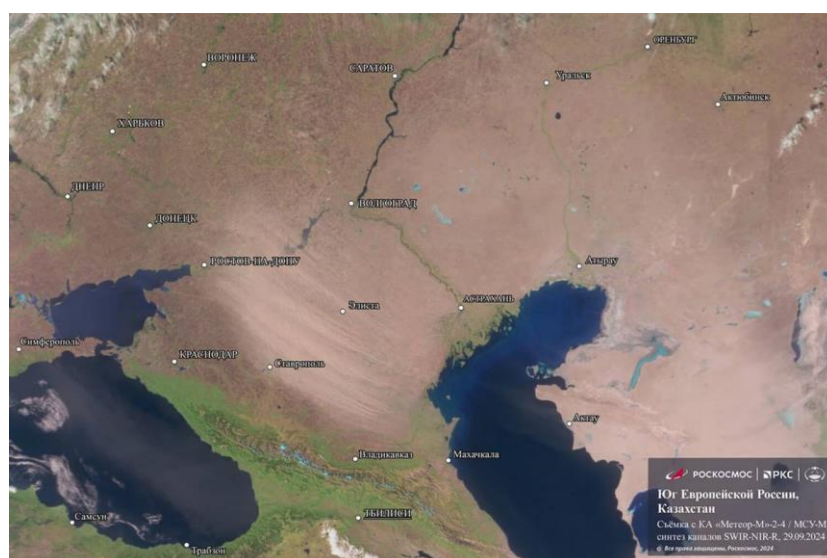


Рисунок 2.10 – Спутниковый снимок переноса песчаных аэрозолей из районов пыльных бурь в Прикаспии

Часто перенос загрязненных  $PM_{10}$  воздушных масс из удаленных районов совпадает с периодами локальных НМУ в московском регионе; слабое рассеивание частиц в принесенной

загрязненной воздушной массе приводит к максимальным концентрациям  $PM_{10}$  в приземном слое. Такие периоды можно назвать эпизодами аэрозольного загрязнения смешанного происхождения.

К эпизодам дальнего переноса  $PM_{10}$  можно также отнести загрязнение пирогенными аэрозолями в периоды лесных пожаров – летом 2010 г. [Звягинцев и др., 2011], в конце лета 2022 г. [Борисов и Кузнецова, 2024].

Отличительной чертой эпизодов дальнего переноса взвешенных частиц является одновременное повышение концентраций на преобладающем большинстве АСКЗА, поскольку принесенная загрязненная воздушная масса покрывает обширную территорию. Также в эпизодах дальнего переноса максимальный уровень  $PM_{10}$  может наблюдаться в нехарактерное время – днем; по длительности такие эпизоды могут сохраняться до 1–3 суток.

Ниже приведен анализ нескольких эпизодов повышенных концентраций  $PM_{10}$  различного генезиса в московском регионе, выявленных по данным измерений на АСКЗА за период 2019–2023 гг. Эпизоды идентифицированы по рекомендованному ВОЗ нормативу среднесуточной концентрации  $PM_{10}$  45  $\text{мкг}/\text{м}^3$  [World Health Organization, 2021] – для анализа выбирались периоды с превышением норматива на более чем 40 % АСКЗА, измеряющих  $PM_{10}$ . Приводимый анализ аэрозольных эпизодов за период 2019–2023 гг. также представлен в авторской публикации [Борисов и Кузнецова, 2024].

Так, 10–14 октября 2020 года каждый день были зафиксированы превышения норматива среднесуточной концентрации  $PM_{10}$  45  $\text{мкг}/\text{м}^3$  на более чем 90 % измерительных станций на территории московского региона. На рисунке 2.11 (а) изображены измеренные на АСКЗА концентрации  $PM_{10}$  в период 9–16 октября 2020 г.

К ночи 11 и 12 октября измеренные концентрации практически на всех АСКЗА достигали уровня 200–250  $\text{мкг}/\text{м}^3$ . Траекторный анализ (<https://www.ready.noaa.gov/hypub-bin/trajtype.pl>) указал на поступление в эти дни мелких частиц в Москву из районов Прикаспия (рисунок 2.11 б, в). Перенос происходил по западной/юго-западной периферии антициклона, при этом скорости ветра на высоте 925 гПа в районе Каспийского моря достигали 22 м/с, сильный юго-восточный ветер в АПС обеспечил перенос частиц в течение суток в московский регион. 13 и 14 октября днем зафиксированы концентрации  $PM_{10}$  около 100  $\text{мкг}/\text{м}^3$  на большинстве станций, обусловленные интенсивным перемешиванием в загрязненном воздухе при аномально теплой погоде в эти дни (+ 18 °С, + 20 °С).

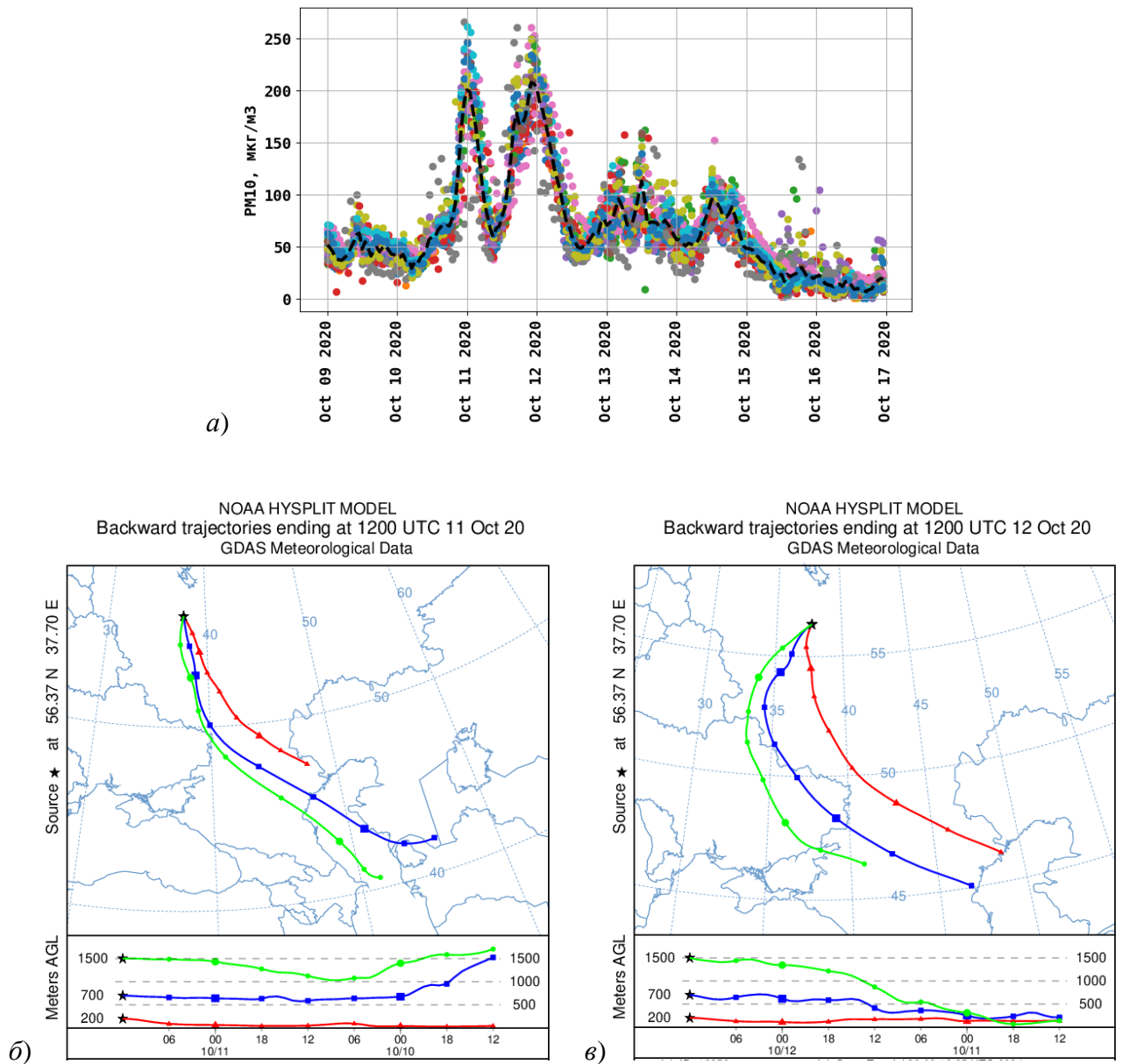


Рисунок 2.11 – Измеренные часовые концентрации  $PM_{10}$  в московском регионе 9–16 октября 2020 г. (а); траектории воздушных частиц в московский регион (б) 10–11 октября 2020 г. и (в) 11–12 октября 2020 г.

Похожие эпизоды дальнего переноса  $PM_{10}$  имели место в московском регионе в марте и апреле 2021 года (рисунок 2.12).

Начиная с второй половины дня 17 марта концентрации  $PM_{10}$  синхронно повышались практически на всех АСКЗА до 150-250  $\mu\text{g}/\text{m}^3$  и держались на таком уровне второй половины 18 марта. Наличие снежного покрова в Москве в эти дни практически исключает повышенное загрязнение  $PM_{10}$  локальными источниками. Траекторный анализ (рисунок 2.13 а) показал перенос частиц из районов Прикаспия 15–17 марта, перенос происходил по юго-западной периферии антициклона с центром над Казахстаном, скорости переноса в АПС (925 гПа) в районах Прикаспия достигали 29 м/с.

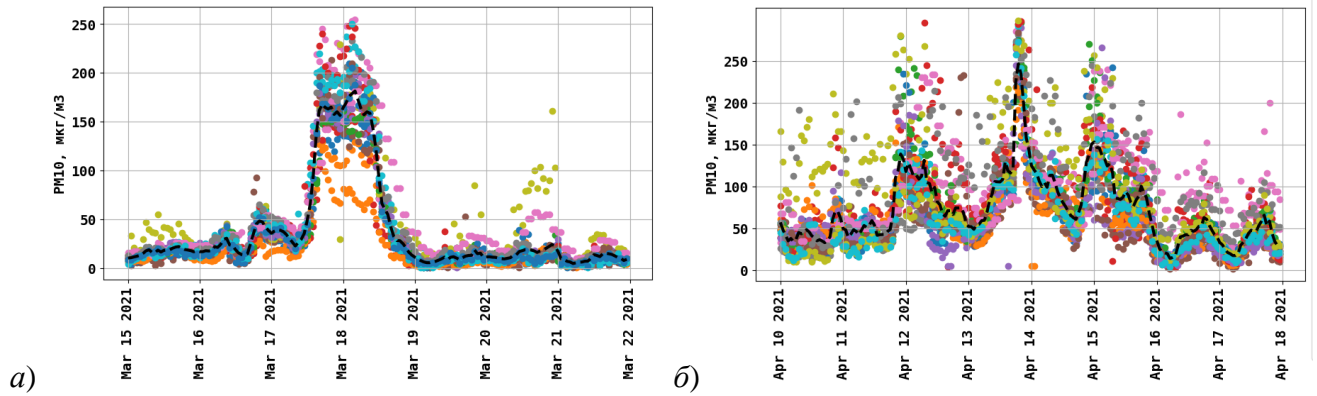


Рисунок 2.12 – Измеренные часовые концентрации  $PM_{10}$  в московском регионе  
(а) 15–21 марта 2021 г. и (б) 10–17 апреля 2021 г.

Косвенным подтверждением наличия значительного количества аэрозольных частиц в переносимой воздушной массе может служить аномальный рост температуры в нижних слоях атмосферы в дневные часы в результате прямого нагрева пылевых частиц. Так, ярко выраженный аномальный рост температуры по данным радиозондирования в 12 ВСВ наблюдался на пути переноса частиц из Прикаспия в Саратове в слое 500–700 м и в Рязани в слое 900–1300 м 16 марта 2021 г., при этом точки начала роста температуры совпадают с максимумами скорости ветра (16–18 м/с) в мезоструйном течении (рисунок 2.13 б). Максимальная величина «аэрозольного нагрева» над Рязанью составила 1,5 °С на 100 м, над Саратовом 0,6 °С на 100 м.

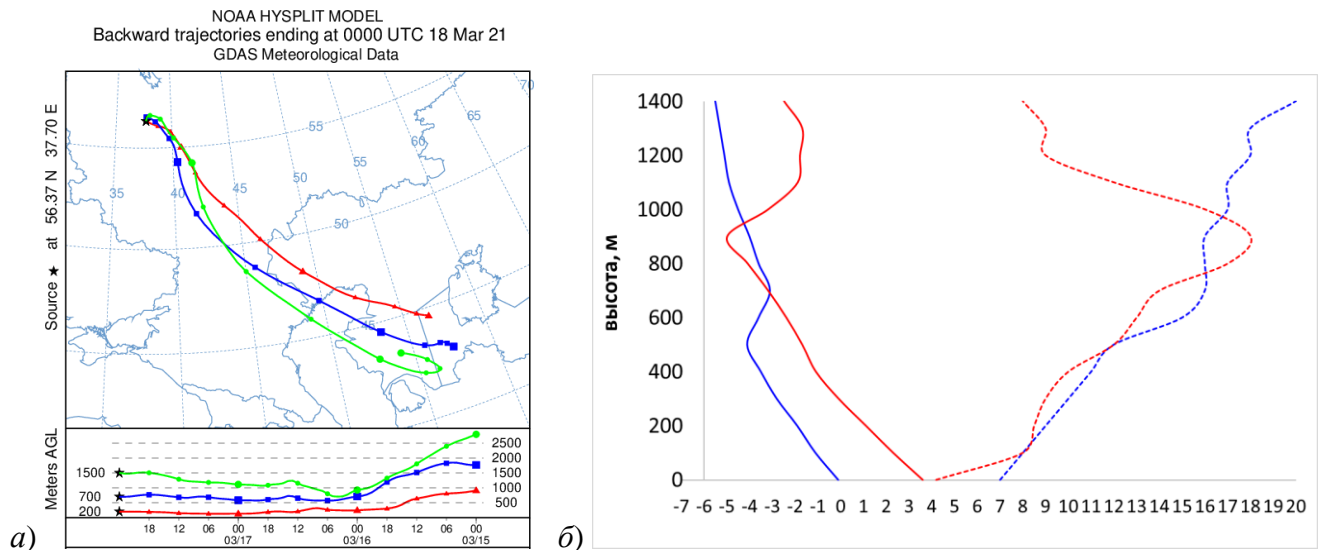


Рисунок 2.13 – (а) Траектории воздушных частиц 15–17 марта 2021 г.;  
(б) температура (°С, сплошная линия) и скорость ветра (м/с, штриховая линия) по данным радиозондирования в 12 ВСВ 16 марта 2021 г. в п. Рязань (красные линии) и в п. Саратов (синие линии)

Вечером 17 марта Москва оказалась в барической ложбине, что при слабых скоростях ветра привело к накоплению принесенных из Прикаспия частиц  $PM_{10}$ . Днем 18 марта начался приток воздуха с севера, сменилась воздушная масса, загрязнение  $PM_{10}$  резко снизилось (рисунок 2.12 а).

В другом эпизоде вечером 13 апреля 2021 г. происходил синхронный рост концентраций  $PM_{10}$  до уровня 200–300  $мкг/м^3$  практически на всех АСКЗА (рисунок 2.12 б). Как показал траекторный анализ, 12 и 13 апреля в центральные области происходил перенос воздушных масс из Прикаспия по юго-западной периферии антициклона с центром над Верхней Волгой, скорости ветра в АПС достигали 17 м/с.

Нужно отметить, что в московском регионе в середине апреля 2021 г. уже сошел снежный покров, но еще не взошла растительность, поэтому незадернованная подстилающая поверхность могла служить источником локального загрязнения. При штиле в АПС дневные температуры в Москве повышались до 16–18 °С при относительной влажности 30 %, ночью формировались приземные инверсии. Такие погодные условия, очевидно, могли способствовать накоплению  $PM_{10}$  от локальных источников. На рисунке 2.12 (б) видно, что пики концентраций наблюдались вечером 11, 13 и 14 апреля, что характерно при накоплении частиц в устойчивом АПС при НМУ, при этом фактор дальнего переноса усугублял загрязнение.

К вечеру 15 апреля направление переноса сменилось на северо-восточное, смена воздушной массы и усиление ветра в АПС в московском регионе (до 15 м/с на 925 гПа) привели к снижению концентраций  $PM_{10}$  до уровня 20–50  $мкг/м^3$  к ночи 16 апреля (рисунок 2.12 б).

Еще один эпизод аэрозольного загрязнения, в котором присутствовал фактор дальнего переноса частиц, имел место 23–25 марта 2022 года. 24 и 25 марта были зафиксированы превышения норматива среднесуточной концентрации  $PM_{10}$  45  $мкг/м^3$  на 95 % станций в московском регионе, максимальные часовые измеренные концентрации  $PM_{10}$  достигали 300  $мкг/м^3$  (рисунок 2.14 а).

20–22 марта Москва находилась на восточной периферии антициклона с центром над северо-востоком Европы. С северо-западным переносом в Москве началось аномальное для марта потепление до 10–15 °С при относительной влажности днем 21–29 %. При потеплении происходит интенсивный сход снежного покрова – 21 марта высота снега 12–14 см, 26 марта уже 3 см. С 22 по 25 марта каждую ночь наблюдаются приземные температурные инверсии 4–7 °С. 23 и 24 марта Москва находилась вблизи центра антициклона, скорости ветра на 925 гПа составляли 2–6 м/с. Пики концентраций  $PM_{10}$  наблюдаются вечером 23 и 24 марта в стабилизированном приземном слое при НМУ, при этом в эти дни во время схода снежного покрова, при все еще влажной почве, не могло происходить интенсивного выноса почвенных частиц с поверхности.



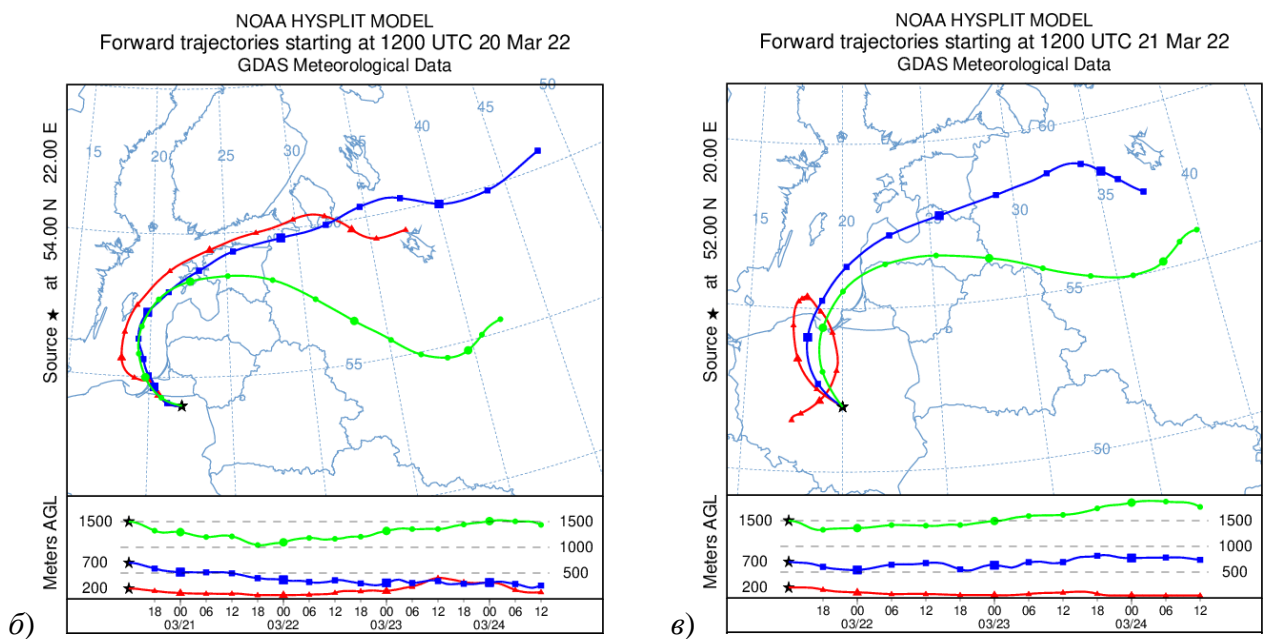
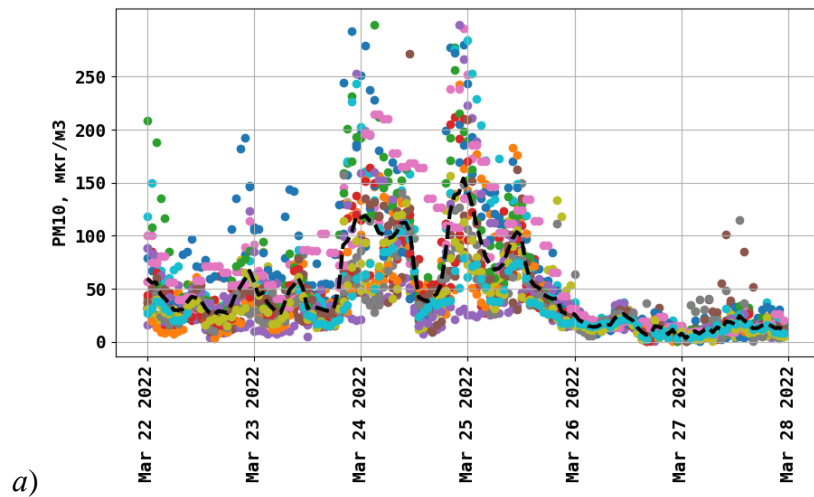


Рисунок 2.14 – (а) Измеренные часовые концентрации  $PM_{10}$  в московском регионе 22–27 марта 2022 г.; траектории воздушных частиц из Польши в центральные области России (б) 21–24 марта 2022 г. и (в) 22–24 марта 2022 г.

Траекторный анализ (рисунки 2.14 б, в) показал перенос частиц из районов центральной Европы 21–24 марта. Сделано предположение, что с началом пахотного сезона в Польше воздушная масса на западной периферии антициклона насыщалась почвенными частицами, и мелкая фракция ( $PM_{10}$ ) могла поступать в ЦФО.

Еще один аэрозольный эпизод 26–29 августа 2022 г. – на более чем 95 % АСКЗА среднесуточные концентрации превышали норматив  $45 \text{ мкг/м}^3$ , на большинстве АСКЗА суточные максимумы концентраций держались на уровне  $100\text{--}200 \text{ мкг/м}^3$ , на отдельных АСКЗА – до  $250\text{--}300 \text{ мкг/м}^3$  (рисунок 2.15 а).

Отметим, что в августе 2022 года в Рязанской, Владимирской, Нижегородской областях, а также в Приокско-Тerrasном заповеднике происходили обширные лесные пожары. Рязань находится на юго-восточном направлении от Москвы на расстоянии 182 км, Приокско-Тerrasный заповедник – в 112 км к югу от Москвы.

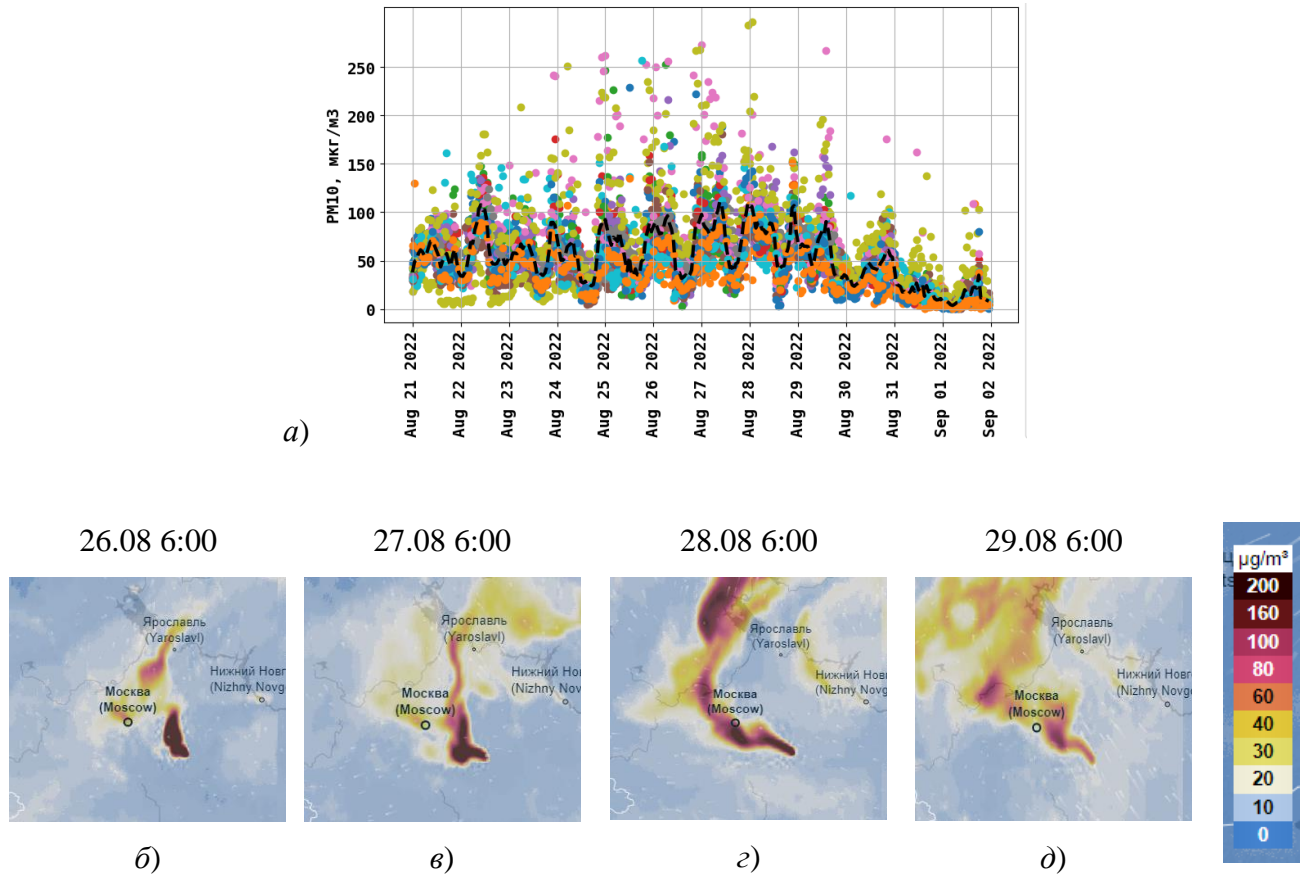


Рисунок 2.15 – (а) Измеренные часовые концентрации  $PM_{10}$  в московском регионе 21 августа – 1 сентября 2022 г.; (б-д) модельный расчет XTM SILAM переноса  $PM_{10}$  из районов лесных пожаров в Рязанской области от 6:00 26-29 августа 2022 г., справа – шкала концентрации  $PM_{10}$  в  $\mu\text{г}/\text{м}^3$

С 20 по 28 августа Москва находилась на западной/юго-западной периферии антициклона, преобладал юго-восточный перенос. Скорости ветра на 925 гПа в период 21–23 августа достигали 13–15 м/с, начиная с 23 августа и до 28 августа ветер ослабевал до 4–7 м/с. 22–28 августа температура повышалась до 31–33 °С при относительной влажности 21–31 %, каждую ночь были инверсии 3–9 °С до высоты 700 м.

Нужно отметить, что для августа нехарактерно сильное загрязнение  $PM_{10}$  локальными источниками даже в условиях НМУ, средняя многолетняя концентрация в августе в московском регионе около 30  $\mu\text{г}/\text{м}^3$  (рисунок 2.8). По расчетам XTM SILAM 26–29 августа происходил перенос в Москву загрязненного воздуха из районов лесных пожаров в рязанской области и

Приокско-Тerrasном заповеднике (рисунок 2.15 б-д). Принесенные пирогенные аэрозоли при НМУ накапливались в приземном слое на территории московского региона. Дополнительным подтверждением переноса пирогенных аэрозолей является повышение концентраций  $PM_{10}$  в дневные часы 25–29 августа, что может быть обусловлено интенсивным вертикальным перемешиванием в загрязненном шлейфом от пожаров АПС (рисунок 2.15 а).

Некоторые эпизоды повышенного загрязнения  $PM_{10}$  в Москве обусловлены исключительно локальными источниками и НМУ. Характерной чертой такого типа загрязнения является большие различия концентраций на типовых АСКЗА, тогда как при дальнем переносе концентрации повышаются синхронно практически на всех АСКЗА, как, например, в октябре 2020 г. (рисунок 2.11 а) и марте 2021 г. (рисунок 2.12 а).

Эпизоды повышенного загрязнения  $PM_{10}$  при НМУ могут формироваться осенью и весной при открытой подстилающей поверхности без снежного покрова или растительности, которые блокируют ветровой подъем почвенных частиц зимой и летом. В пример приведем 2 апрельских эпизода повышенного загрязнения  $PM_{10}$  при НМУ в 2019 и 2023 годах.

24–28 апреля 2019 года большинство АСКЗА фиксировали суточные максимумы концентрации  $PM_{10}$  100–150  $мкг/м^3$ , на отдельных, в основном примагистральных, станциях концентрации  $PM_{10}$  достигали 200 и более  $мкг/м^3$  (рисунок 2.16 а).

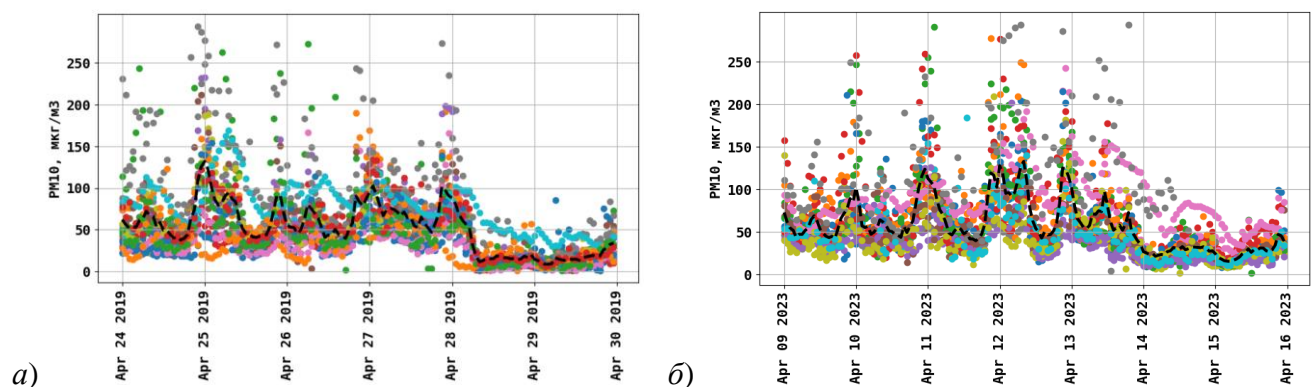


Рисунок 2.16 – Измеренные часовые концентрации  $PM_{10}$  в московском регионе

(а) 24–29 апреля 2019 г. и (б) 9–15 апреля 2023 г.

25–27 апреля 2019 г. Москва находилась в барической ложбине со скоростями ветра в АПС 6–8 м/с (на 925 гПа). При этом стояла не по сезону теплая (днем 21–25 °С) и очень сухая (относительная влажность 20–23 %) погода, каждую ночь наблюдались приземные температурные инверсии до высоты 400 м. В таких условиях в связи с усилением вертикального перемешивания в загрязненном АПС повышенные концентрации  $PM_{10}$  наблюдались и в дневное время (рисунок 2.16 а). Конец эпизода связан с прохождением холодного фронта ночью 28 апреля.

В похожих метеоусловиях сформировался аэрозольный эпизод 9–13 апреля 2023 года, когда концентрации  $PM_{10}$  повышались до 100–150  $\mu\text{кг}/\text{м}^3$ , на отдельных примагистральных АСКЗА до 200–300  $\mu\text{кг}/\text{м}^3$  (рисунок 2.16 б). Москва находилась в зоне малых барических градиентов со скоростями ветра в АПС 0–4 м/с – практически штиль. Дневные температуры составляли 15–18 °С при относительной влажности 23–43 %. Каждый вечер в условиях приземных температурных инверсий 2–7 °С до высот 200–270 м происходило повышение уровня  $PM_{10}$ . В отсутствии снежного покрова основным источником  $PM_{10}$  служил ветровой подъем почвенных частиц, 12 и 13 апреля можно наблюдать дневные пики концентраций, связанные с усилением вертикального перемешивания в загрязненном АПС.

Эпизод закончился с усилением северо-восточного ветра (до 14–15 м/с на 925 гПа) после прохождения холодного атмосферного фронта 13–14 апреля. При этом следует отметить «предфронтальное» повышение концентраций  $PM_{10}$  вечером 13 апреля (рисунок 2.16 б), связанное с усилением ветра во время прохождения фронта и подъемом частиц.

### 2.3 Численные прогнозы метеопараметров мезомасштабной системы COSMO-Ru

В исследовательскую БД включены численные прогнозы метеопараметров мезомасштабной системы COSMO-Ru. Ниже кратко описывается система прогнозирования используемых метеорологических характеристик.

С сентября 2009 года Росгидромет является полноправным членом консорциума по мезомасштабному моделированию COSMO (Consortium for Small-scale Modeling, Консорциум мезомасштабного моделирования, <http://cosmo-model.org/>). Модель COSMO адаптирована в Росгидромете в рамках системы COSMO-Ru.

Мезомасштабная система прогноза погоды COSMO-Ru с апреля 2011 г. признана Центральной методической комиссией Росгидромета по гидрометеорологическим и гелиогеофизическим прогнозам в качестве базовой отечественной технологии численного регионального краткосрочного прогноза погоды.

COSMO основана на системе термодинамических уравнений, описывающих сжимаемый поток воздуха во влажной атмосфере. Уравнения составляют замкнутую систему, позволяющую рассчитывать и прогнозировать основные метеорологические величины: температуру и влажность воздуха, скорость ветра, давление, осадки. [Ривин и др., 2015; 2019]. Для задания начальных и боковых граничных условий в модели COSMO-Ru используются данные, подготавливаемые в Гидрометцентре России на основе данных со свободным доступом (с 2022 года) и ряда метеопараметров, вычисляемых также в Гидрометцентре России.

Пользователи получают результаты численного прогноза погоды COSMO-Ru ежедневно в 00, 06, 12, 18 ч ВСВ.

Расчеты метеопараметров ведутся на сетках с повернутой сферической системой координат. В настоящее время ведутся расчеты нескольких конфигураций модели COSMO-Ru (рисунок 2.17), в том числе:

- COSMO-Ru66ENA с шагом сетки 6,6 км охватывает почти всю территорию Евразии, производит прогноз на 120 часов (5 суток);

- COSMO-Ru22ETR с шагом сетки 2,2 км охватывает всю Европейскую территорию России и территории ряда прилегающих стран, включая Белоруссию; производит прогноз на 48 часов (2 суток). В качестве начальных и боковых граничных условий в COSMO-Ru22ETR используются прогнозы COSMO-Ru66ENA;

- COSMO-Ru10Msk с шагом сетки 1 км охватывает территорию Московского мегаполиса, производит прогноз на 48 часов (2 суток).

В данном исследовании использовались прогнозы COSMO-Ru22ETR, интерполированные на сетку XTM CHIMERE (см. раздел 2.4).

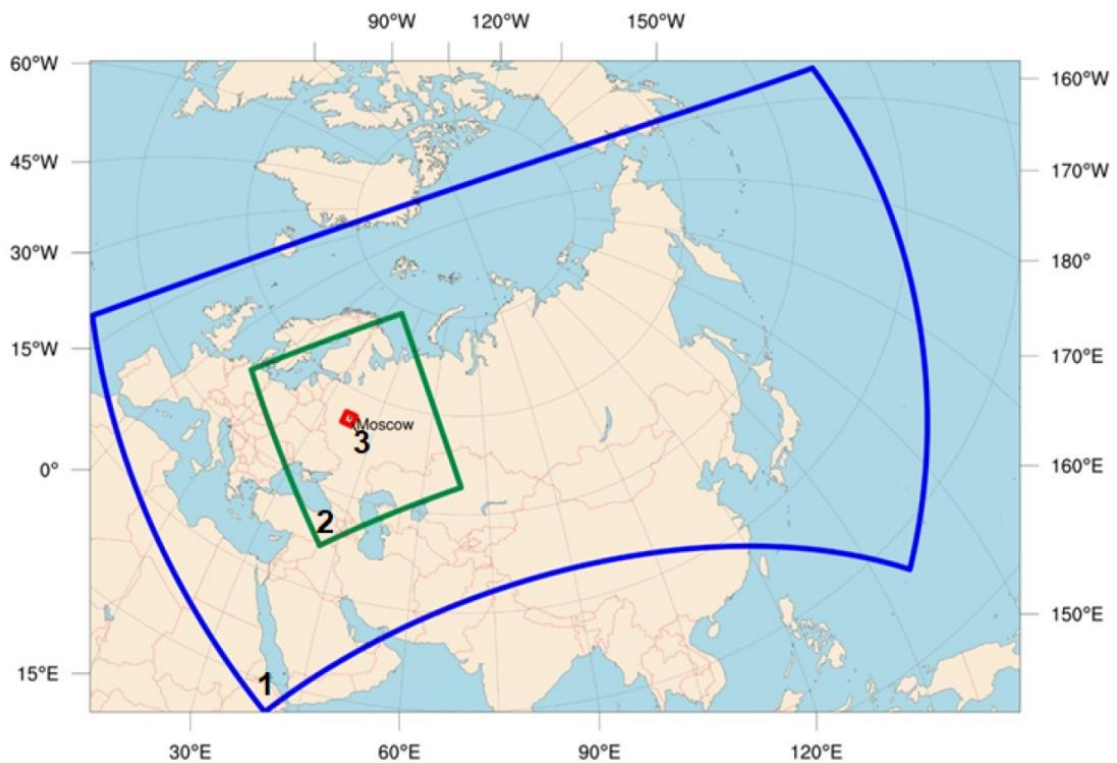


Рисунок 2.17 – Области оперативного прогноза системы COSMO-Ru в Гидрометцентре России:

1 – COSMO-Ru66ENA; 2 – COSMO-Ru22ETR; 3 – COSMO-Ru10Msk [Ривин и др., 2019]

## 2.4 Численные прогнозы концентраций $O_3$ и $PM_{10}$ на территории московского региона

В исследовательскую БД также включены прогнозы приземных концентраций  $O_3$  и  $PM_{10}$ , рассчитываемые химической транспортной моделью CHIMERE. Именно в сравнении с расчетами ХТМ CHIMERE оценивалась эффективность разрабатываемой методики постпроцессинга прогнозов концентраций на основе методов машинного обучения. В разделе дается краткое описание ХТМ CHIMERE; по результатам ее многопланового тестирования формулируется обоснование целесообразности применения методов, улучшающих качество прогнозирования ХТМ CHIMERE.

Химическая транспортная модель CHIMERE с 1999 года разрабатывается в Национальном центре научных исследований Франции. Актуальная версия CHIMERE является современным инструментом прогнозирования качества воздуха, позволяет рассчитывать концентрации сотен газовых примесей и аэрозольных частиц в атмосфере на регулярных сетках с горизонтальным шагом от 1 км до нескольких градусов по широте-долготе, охватывающих территории от локальных до континентальных масштабов [Menut et al., 2024]. Расчеты ХТМ CHIMERE входят в ансамблевый прогноз загрязнений на территории Европы CAMS (<https://atmosphere.copernicus.eu/>). Исходный код CHIMERE в основном написан на языке программирования Fortran, находится в открытом доступе (<https://www.lmd.polytechnique.fr/chimere/>).

С 2019 года в Гидрометцентре России ежедневно составляются прогнозы концентраций загрязнителей для территории московского региона на 48 часов вперед с временным шагом 1 час на основе расчетов ХТМ CHIMERE [Кузнецова и др., 2019; Кузнецова и др., Гиип 2022]. В 2019–2020 гг. регулярная сетка CHIMERE охватывала территории г. Москвы и часть московского региона к северо-востоку от Москвы; в конце 2020 г. сетка CHIMERE была смещена на юго-запад с целью охвата территории Новой Москвы (рисунок 2.18). В обоих случаях регулярная сетка состоит из 2500 узлов (50 по долготе  $\times$  50 по широте), при этом шаг сетки практически не менялся, оставаясь на уровне 2 км по долготе  $\times$  2 км по широте.

Основу входных данных, усваиваемых CHIMERE, составляют поля метеорологических параметров, выбросы первичных загрязняющих веществ и химические граничные условия. Ключевыми процессами в CHIMERE, влияющими на концентрации химических веществ, являются перенос (адвекция и перемешивание), химия и сухое и влажное осаждение (рисунок 2.19).

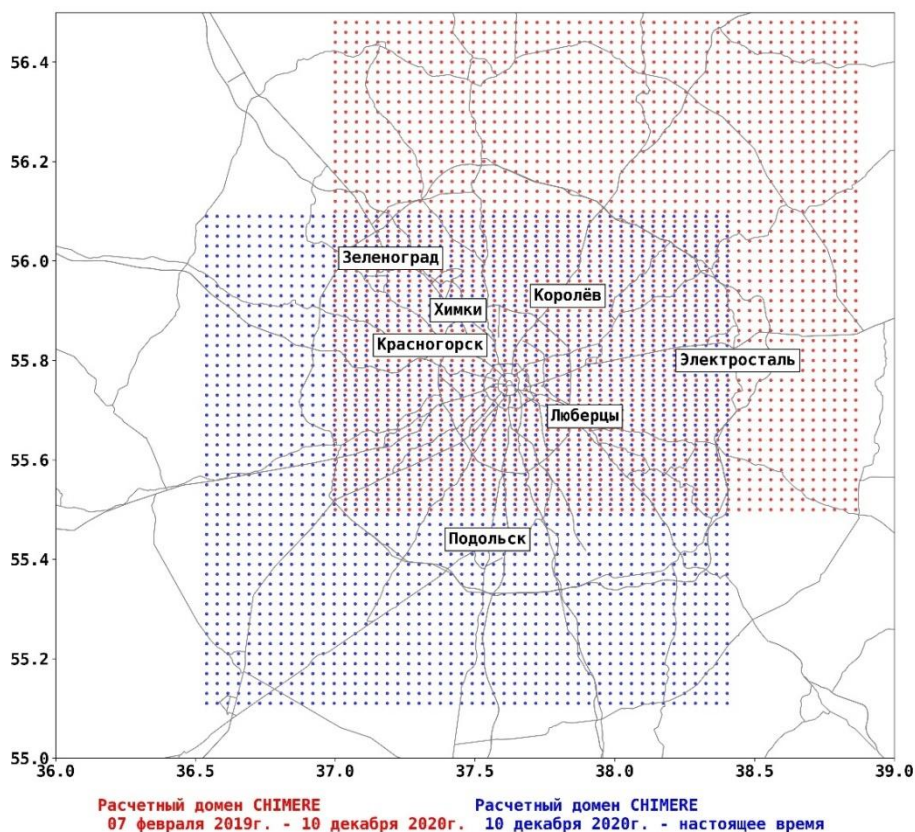


Рисунок 2.18 – Модельные сетки расчетов XTM CHIMERE с горизонтальным шагом  $2 \text{ км} \times 2 \text{ км}$  для территории московского региона за период 7 фев. 2019 г. – н.в.

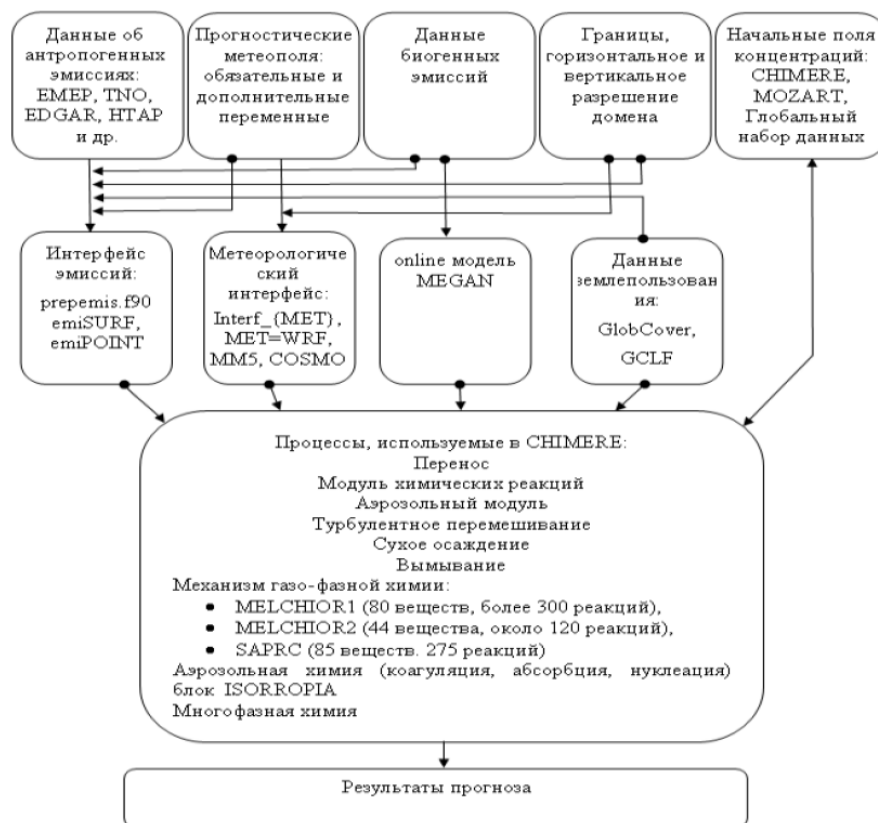


Рисунок 2.19 – Схема расчета прогнозов концентраций в XTM CHIMERE

В CHIMERE используются прогнозы метеопараметров, производимые в Гидрометцентре России с использованием модели атмосферы COSMO-Ru конфигурации COSMO-Ru22ETR с шагом сетки 2,2 км, охватывающей территорию европейской части России. Процедура расчета CHIMERE включает билинейную интерполяцию готовых на весь срок (48 часов) прогнозов метеопараметров на собственную модельную сетку (рисунок 2.18).

Для задания антропогенных выбросов первичных загрязнителей CHIMERE использует данные зарубежной базы ЕМЕП (European Monitoring and Evaluation Programme, <https://www.emep.int/>). Данные ЕМЕП представлены на регулярной сетке  $0.5^\circ \times 0.5^\circ$  для территории Европы и европейской части РФ, распределены по 11 источникам выбросов. Данные эмиссий ЕМЕП интерполируются на заданную сетку CHIMERE с учетом данных о землепользовании базы GlobCover (GlobCover Land Cover, [https://due.esrin.esa.int/page\\_globcover.php](https://due.esrin.esa.int/page_globcover.php)). Ранее в Гидрометцентре проводились эксперименты по региональной адаптации данных землепользования GlobCover для Нижнего Новгорода [Борисов и Шалыгина, 2021].

Данные землепользования базы GlobCover для территории московского региона включены в исследовательскую БД для учета пространственного распределения антропогенной нагрузки по территории. Нужно отметить, что данные GlobCover, используемые для расчетов ХТМ CHIMERE, интерполированы на актуальную модельную сетку CHIMERE (рисунок 2.18) и не требуют пространственной синхронизации с прогнозами концентраций и метеопараметров.

Конкретно использованы данные GlobCover, описывающие пространственное распределение городской застройки по территории – доли городской застройки в модельных ячейках  $2 \text{ км} \times 2 \text{ км}$  показаны на рисунке 2.20 и вполне удовлетворительно отражают современное состояние городской инфраструктуры в Москве. Так, на рисунке 2.20 на фоне плотной городской застройки в старой Москве и прилегающих окрестностях ярко выделяются парковые зоны: Лосиный остров, Измайловский, Битцевский парки, район Воробьевых гор и Крылатская пойма.

Для оценки объема биогенных эмиссий ХТМ CHIMERE использует данные модели MEGAN. На границах области расчета CHIMERE концентрации примесей задаются по результатам климатического моделирования: для газовых примесей – данные модели MOZART2, для аэрозолей – данные модели GOCARD.

Для моделирования химических преобразований ХТМ CHIMERE использует блок газофазных реакций MELCHIOR2, содержащий ~40 веществ и более 110 реакций. Аэрозоль разделяется на семь типов по составу и на несколько фракций по размерам.



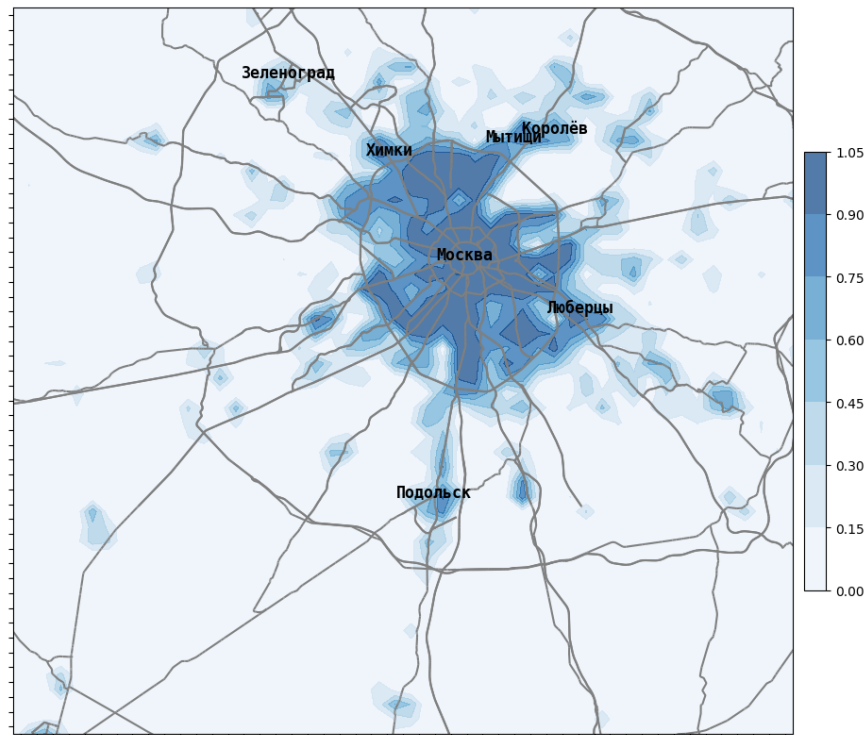


Рисунок 2.20 – Доли городской застройки в ячейках сетки с горизонтальным шагом  $2 \text{ км} \times 2 \text{ км}$  на территории московского региона по данным землепользования базы GlobCover

Архив содержит ежедневные прогностические поля концентраций XTM CHIMERE и метеопараметров модели COSMO-Ru на 1–2 суток с временным шагом 1 час на модельной сетке с горизонтальным шагом 2 км для территории московского региона (рисунок 2.18). Архив прогнозов охватывает почти 6-летний период с 7 февраля 2019 г. по 30 сентября 2024 г., по техническим причинам в архиве отсутствуют прогнозы за период 23 мая – 30 сентября 2021 г.

#### 2.4.1 Ретроспективная оценка качества модельных прогнозов концентраций $\text{PM}_{10}$ и $\text{O}_3$ на территории московского региона

В данном подразделе представлен систематизированный анализ качества численного прогнозирования концентраций  $\text{O}_3$  и  $\text{PM}_{10}$  на территории московского региона на основе XTM CHIMERE, проведенный для обоснования разработки методики постпроцессинга модельных прогнозов.

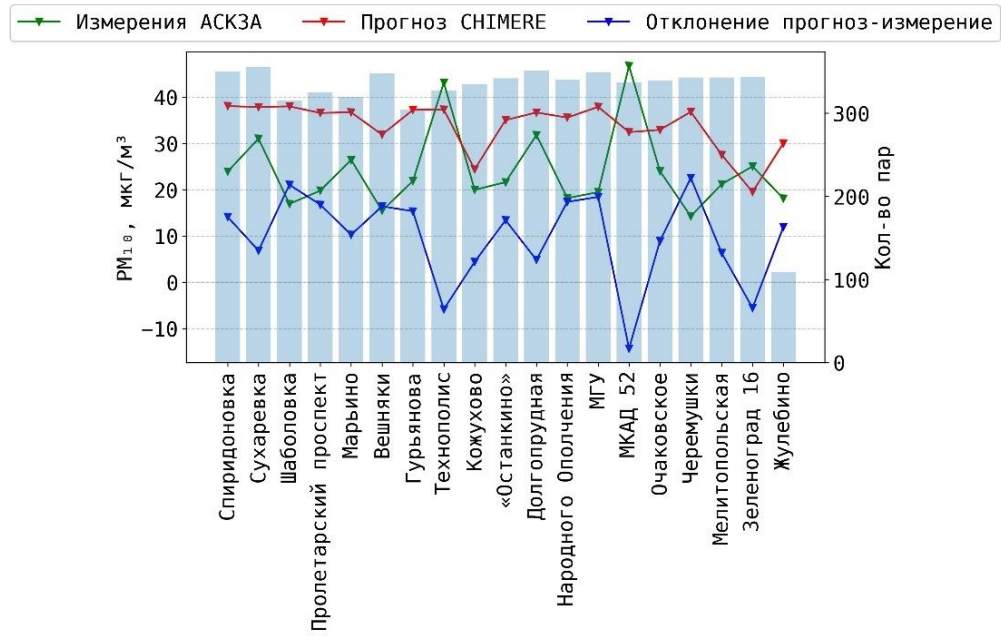
В рамках функционирующей в Гидрометцентре России технологии оценки качества прогнозов загрязнений проводятся регулярно; показатели оценок прогнозов за разные временные интервалы, начиная с 2019 года, приводятся в подготовленных с участием автора публикациях [Кузнецова и др., 2019; Кузнецова и др., Гиип 2022]. Ниже представлены некоторые метрики

качества прогнозов ХТМ CHIMERE концентраций  $O_3$  и  $PM_{10}$  на территории московского региона за 2023 год, рассчитанные путем сравнения прогностических концентраций с измеренными на АСКЗА ГПБУ «Мосэкомониторинг» по методу «ближайшего соседа» – данные измерений на АСКЗА сравнивались с прогнозом в ближайшем узле модельной сетки. Для оценки прогнозов использованы часовые измерения  $PM_{10}$  на 19 АСКЗА и  $O_3$  на 16 АСКЗА, в основном находящихся на территории г. Москвы. После синхронизации измерений АСКЗА с прогнозами ХТМ CHIMERE и удаления пропусков в оценочную выборку вошли 149 441 пара прогноз-наблюдение  $PM_{10}$  и 125 651 пара  $O_3$ .

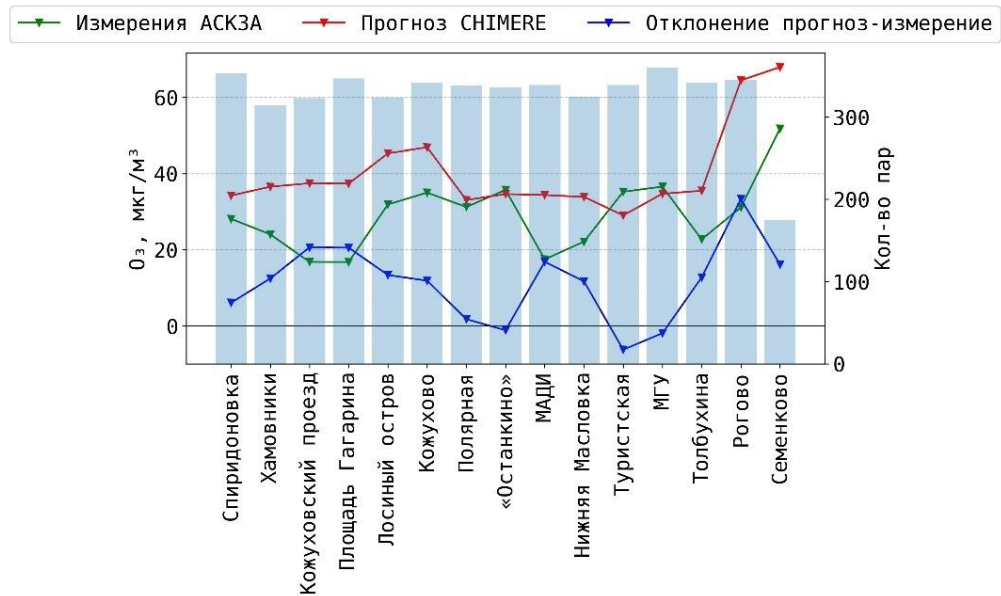
На рисунке 2.21 представлены осредненные за 2023 год измеренные (зеленой линией) и спрогнозированные (красной) среднесуточные концентрации  $PM_{10}$  (а) и  $O_3$  (б) на отдельных АСКЗА, а также средние отклонения прогнозов от измерений (далее  $\Delta_{п-и}$ ). Рисунки демонстрируют характерные для численных прогнозов меньшие межстанционные различия прогностических величин по сравнению с измеренными и меньшую пространственную изменчивость концентраций по территории города. При этом заметим (рисунок 2.21 а), что смещенный в область положительных значений основной диапазон  $\Delta_{п-и}$  ( $PM_{10-CC}$ ) составил 10–20  $мкг/м^3$ , что указывает на преимущественное модельное завышение ХТМ концентраций  $PM_{10}$ . Две АСКЗА (МКАД, Технополис), для которых ХТМ рассчитывала меньшие по сравнению с измерениями концентрации  $PM_{10}$ , имеют ярко выраженную специфику – уровень загрязнения там в 1,5–2 раза больше среднего городского. Также отметим, что АСКЗА Кожухово и Зеленоград\_16, для которых спрогнозирован заметно меньший уровень загрязнения, чем для остальных АСКЗА, находятся за пределами МКАД.

Сравнение прогностических и наблюдавшихся средних за 2023 г. среднесуточных концентраций  $O_3$  (рисунок 2.21 б) указывает на небольшое общее завышение модельных расчетов, при минимальных  $\Delta_{п-и}$  ( $O_3-CC$ ) – до 10–15  $мкг/м^3$  на станциях городского типа. Положительные величины  $\Delta_{п-и}$  ( $O_3-CC$ ) до 20  $мкг/м^3$  отмечены на АСКЗА, находящихся в непосредственной близости от автомагистралей (МАДИ, Нижняя Масловка, Площадь Гагарина и др.), что указывает на недостаточную модельную деструкцию  $O_3$  в зоне транспортных выбросов. На удалении от мегаполиса (Рогово, Семеново) смоделирован более высокий, чем в городе, уровень  $O_3$ , но с превышением по сравнению с измеренными концентрациями в среднем на ~20–30  $мкг/м^3$ . При этом на других (условно) фоновых станциях, расположенных на границах мегаполиса (Лосиный остров, Кожухово), уровень  $O_3$  смоделирован вполне успешно.

Учитывая строгость показанных выше оценок прогнозов путем сравнения станционных наблюдений с узловым расчетом – средней по модельной ячейке ( $2 \times 2$  км) концентрацией  $O_3$  и  $PM_{10}$ , выполнены оценки прогнозов в среднем по городу, а также отдельно по АСКЗА транспортного типа и остальных типов, которые условно обозначим «городские» АСКЗА.



а)



б)

Рисунок 2.21 – Средние за 2023 г. измеренные (зеленой линией), спрогнозированные XTM CHIMERE (красной линией) среднесуточные концентрации (мкг/м<sup>3</sup>) (а) PM<sub>10</sub>, (б) O<sub>3</sub> и отклонения прогнозов (синей линией) на отдельных АСКЗА

*Примечание:* Синие столбы – количество пар прогноз-измерение, по которым рассчитывались оценки.

На рисунке 2.22 приводятся распределения отклонений прогнозов от измерений среднесуточной (а) и максимальной за сутки (б) концентрации PM<sub>10</sub> на транспортных и городских АСКЗА по месяцам 2023 года в виде бокс-диаграмм. Обозначим, что здесь и далее закрашенные боксы на бокс-диаграммах обозначают межквартильный размах распределения отклонений прогнозов, равный разности между третьим (75 %) и первым квартилем (25 %)

распределения. Далее будем обозначать межквартильный размах как IQR (англ. interquartile range), первый и третий квартили распределения как Q1 и Q3.

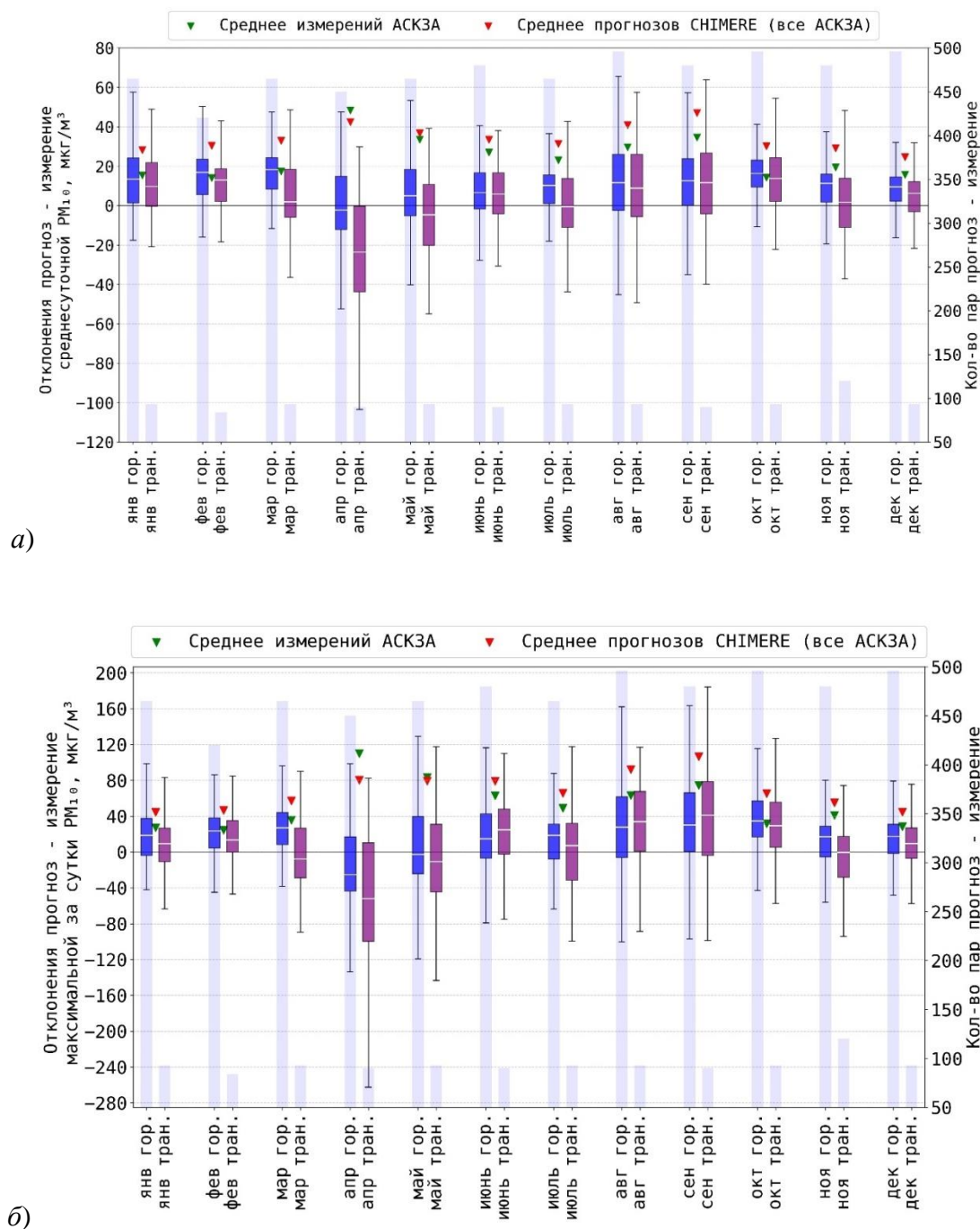


Рисунок 2.22 – Бокс-диаграммы распределений отклонений прогнозов XTM CHIMERE (а) среднесуточных и (б) максимальных концентраций  $\text{PM}_{10}$  для городских (синим цветом) и транспортных (фиолетовым цветом) АСКЗА, 2023 г.

*Примечание:* В боксе – IQR  $\Delta_{\text{п-и}}$ , линия внутри бокса – среднее  $\Delta_{\text{п-и}}$ , нижний «ус»:  $Q1 - 1.5 * \text{IQR } \Delta_{\text{п-и}}$ , верхний «ус»:  $Q3 + 1.5 * \text{IQR } \Delta_{\text{п-и}}$ . Зеленые треугольники – осредненное за месяц измерение (все АСКЗА), красные треугольники – осредненный за месяц прогноз (все АСКЗА). Синие столбы – количество пар прогноз-измерение, по которым рассчитывались оценки.

Как видно (рисунок 2.22 а), на транспортных АСКЗА средние  $\Delta_{\text{П-И}}$  ( $\text{PM}_{10\text{-CC}}$ ) в основном ниже, чем на жилых. В холодный сезон, с октября по март, ХТМ CHIMERE стабильно завышает  $\text{PM}_{10\text{-CC}}$  при достаточно низких измеренных концентрациях – 14–19  $\text{мкг}/\text{м}^3$ . В апреле и мае наблюдается максимальный разброс отклонений прогнозов от измерений как для среднесуточных, так и для максимальных концентраций (рисунок 2.22 а, б). В основном в эти месяцы CHIMERE занижает прогнозы, среднее отклонение прогнозов (все АСКЗА) среднесуточных концентраций -6  $\text{мкг}/\text{м}^3$  в апреле и 3  $\text{мкг}/\text{м}^3$  в мае, максимальных – -30  $\text{мкг}/\text{м}^3$  в апреле и -4  $\text{мкг}/\text{м}^3$  в мае.

Анализ среднеквадратического отклонения прогнозов от измерений  $\text{PM}_{10}$  по месяцам (рисунок 2.23) показал, что максимум  $\text{СКО}_{\text{П-И}}$  приходится на апрель – 30  $\text{мкг}/\text{м}^3$  для среднесуточных концентраций.

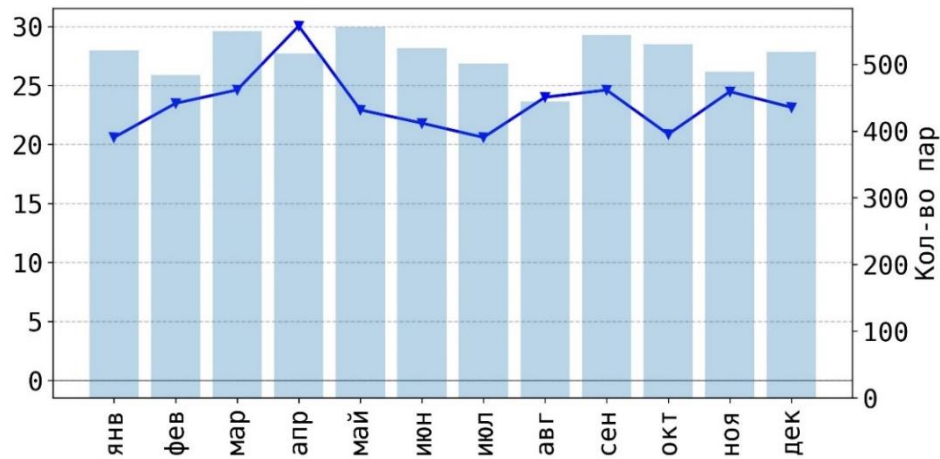


Рисунок 2.23 – Среднеквадратические отклонения прогнозов ХТМ CHIMERE ( $\text{мкг}/\text{м}^3$ ) среднесуточных концентраций  $\text{PM}_{10}$  от измерений на АСКЗА, 2023 г.

В августе и сентябре CHIMERE завышает среднесуточные концентрации  $\text{PM}_{10}$  в среднем на 11–12  $\text{мкг}/\text{м}^3$ , максимальные на 29–32  $\text{мкг}/\text{м}^3$ ; разброс отклонений, как и в апреле, достаточно большой (рисунок 2.22 а, б).  $\text{СКО}_{\text{П-И}}$  в августе-сентябре 24–25  $\text{мкг}/\text{м}^3$  для среднесуточных концентраций (рисунок 2.23).

На рисунке 2.24 представлены среднечасовые спрогнозированные CHIMERE и измеренные на АСКЗА концентрации  $\text{PM}_{10}$  в декабре и апреле 2023 года – месяцы наименьшего и наибольшего в году аэрозольного загрязнения.

С учетом низкого фона и слабых внутрисуточных флуктуаций  $\text{PM}_{10}$  зимой, модельное воспроизведение суточного хода (рисунок 2.24 а) можно считать удовлетворительным, при том, что модельный прогноз ХТМ завышает концентрации, в околополуденные часы почти в 2 раза. Зимнее модельное завышение  $\text{PM}_{10}$  следует интерпретировать, опираясь на особенность

используемых данных об эмиссиях. Здесь необходимо отметить установленную особенность БД ЕМЕР, в которой эмиссии по сезонам распределены соответственно «европейскому» стандарту: годовой максимум приходится на зимние месяцы, тогда как в Москве годовой максимум  $PM_{10}$  наблюдается весной.

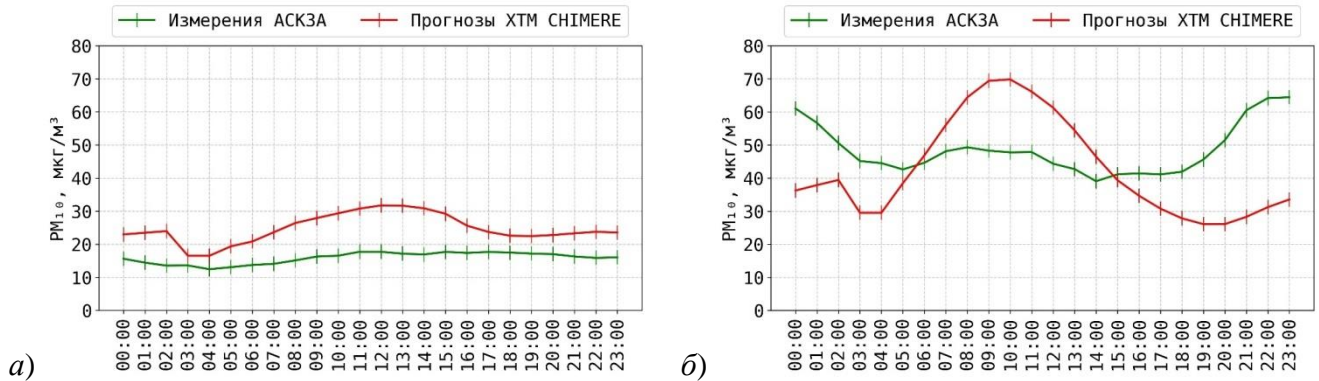


Рисунок 2.24 – Средний суточный ход концентраций  $PM_{10}$  по измерениям на АСКЗА и прогнозам ХТМ CHIMERE в декабре (а) и апреле (б) 2023 г.

В суточном ходе  $PM_{10}$  весной (рисунок 2.24 б) модельные расчеты показывают единственный и основной суточный утренний максимум, ослабевающий после полудня. Натурные измерения фиксируют в это время слабый вторичный суточный максимум. Более существенно то, что модельные расчеты плохо отражают процесс вечернего повышения уровня  $PM_{10}$  и образования ночного главного суточного максимума.

Распределения отклонений прогнозов среднесуточных (а) и максимальных за сутки (б) концентраций  $O_3$  за весь 2023 г. для АСКЗА городского и транспортного типа иллюстрируются на рисунке 2.25.

Видно (рисунок 2.25 а, б), что прогнозы  $O_3$  для транспортных АСКЗА в большей степени завышены по сравнению с городскими АСКЗА. С сентября по февраль CHIMERE немного завышает среднесуточные концентрации  $O_3$  – в среднем на 1–11  $µг/м^3$  при среднем измерении 17–34  $µг/м^3$ . Средние  $\Delta_{п-и}$  максимальных концентраций  $O_3$  в этот период небольшие – в пределах -7-5  $µг/м^3$  при среднемесячных измерениях 32–59  $µг/м^3$ .

Весной средние  $\Delta_{п-и}$  ( $O_3-сс$ ) тоже незначительные: 3  $µг/м^3$  в апреле при максимуме измеренных концентраций, 14  $µг/м^3$  в мае. Отклонения прогнозов максимальных концентраций  $O_3$  лежат в диапазоне от небольшого занижения в апреле -6  $µг/м^3$ , при среднем измерении 77  $µг/м^3$ , до завышения на 12  $µг/м^3$  в мае, при среднем измерении 70  $µг/м^3$ .

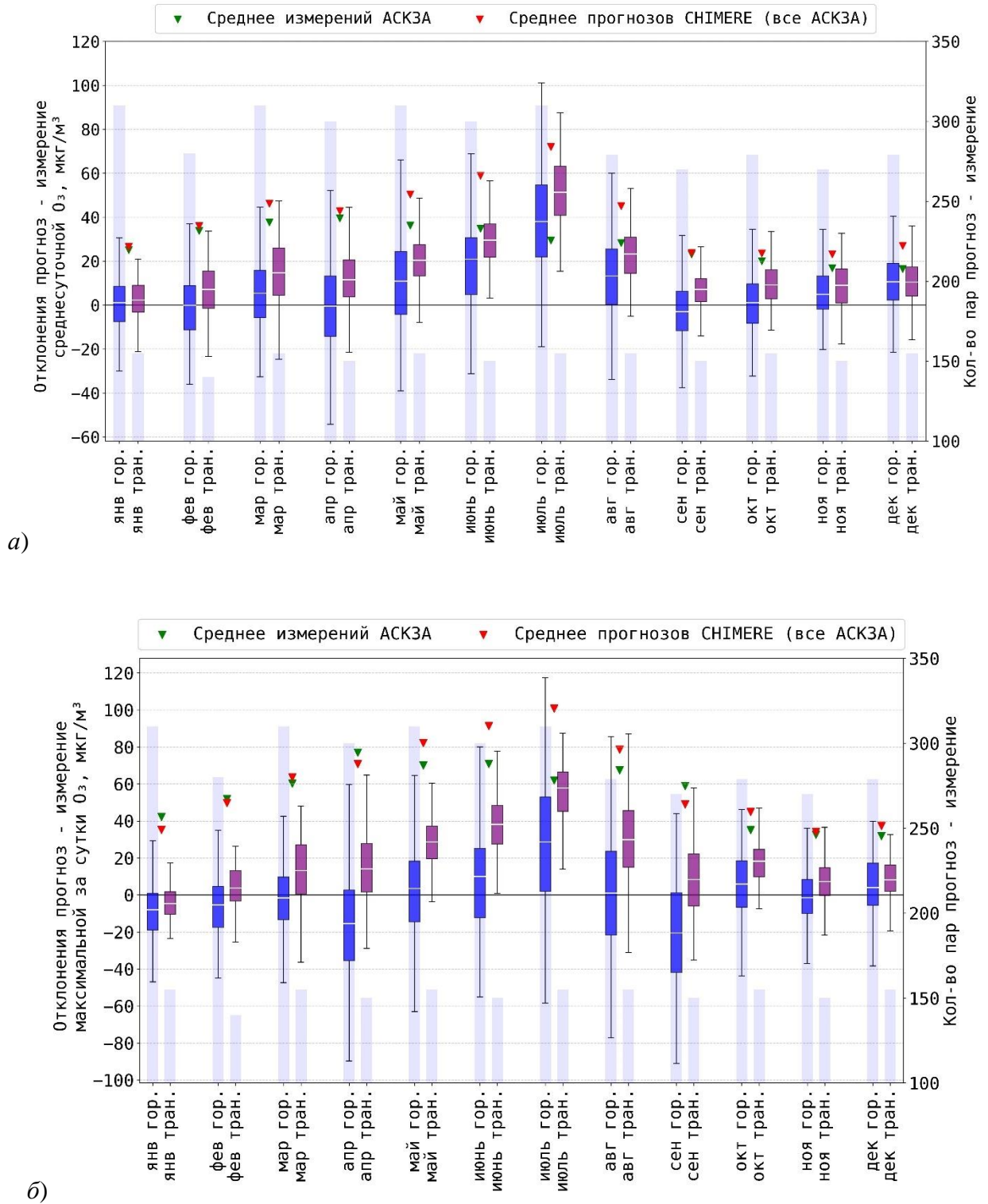


Рисунок 2.25 – Бокс-диаграммы распределений отклонений прогнозов ХТМ CHIMERE (а) среднесуточных и (б) максимальных концентраций O<sub>3</sub> для городских (синим цветом) и транспортных (фиолетовым цветом) АСКЗА, 2023 г.

*Примечание:* В боксе – IQR Δ<sub>п-и</sub>, линия внутри бокса – среднее Δ<sub>п-и</sub>, нижний «ус»: Q1 – 1.5 \* IQR Δ<sub>п-и</sub>, верхний «ус»: Q3 + 1.5 \* IQR Δ<sub>п-и</sub>. Зеленые треугольники – осредненное за месяц измерение (все АСКЗА), красные треугольники – осредненный за месяц прогноз (все АСКЗА). Синие столбы – количество пар прогноз-измерение, по которым рассчитывались оценки.

Для весеннего периода, когда могут возникать эпизоды опасных концентраций  $O_3$ , характерен довольно большой разброс отклонений прогнозов CHIMERE (рисунок 2.25). СКО<sub>П</sub>-и весной немного выше, чем в холодный период: для среднесуточных концентраций  $O_3$  19–24 мкг/м<sup>3</sup>, для максимальных 20–32 мкг/м<sup>3</sup> (рисунок 2.26 а, б).

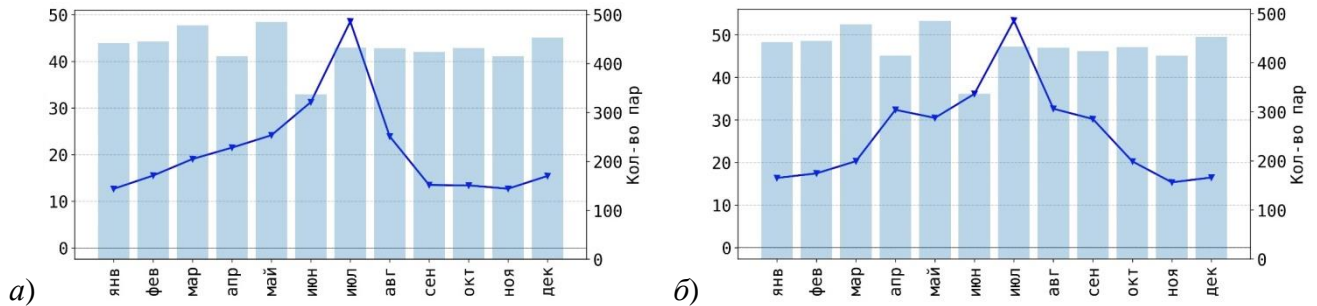


Рисунок 2.26. Среднеквадратические отклонения прогнозов XTM CHIMERE (мкг/м<sup>3</sup>) среднесуточных (а) и максимальных за сутки (б) концентраций  $O_3$  от измерений на АСКЗА, 2023 г.

XTM CHIMERE наибольшие в году концентрации  $O_3$  прогнозирует в летний период, летом прогнозы в основном завышены; среднее  $\Delta_{П}$ -и среднесуточных и максимальных концентраций  $O_3$  в июле оказалось близким к 40 мкг/м<sup>3</sup>, в августе 15–20 мкг/м<sup>3</sup> (рисунок 2.25). Также летом отмечается наибольший в году разброс отклонений прогнозов: СКО<sub>П</sub>-и среднесуточных и максимальных за сутки концентраций  $O_3$  в июле близко к 50 мкг/м<sup>3</sup> (рисунок 2.26).

При отмеченных значительных завышениях модельных прогнозов концентраций  $O_3$  необходимо отметить, что в Москве летом 2023 года содержание озона в приземном воздухе было пониженным из-за преобладающей прохладной и дождливой погоды. Недостаток летнего тепла частично компенсировало длительное «бабье лето» в сентябре, средняя за месяц концентрация  $O_3$  в этом месяце оказалась на 10–15 % больше многолетних значений [Иванова и др., 2024].

При сравнении модельного и наблюдавшегося среднего суточного хода  $O_3$  в холодный сезон года (рисунок 2.27 а) видно их удовлетворительное совпадение при незначительном временном сдвиге экстремумов.

Хорошо коррелирует модельный и наблюдавшийся суточный ход концентраций  $O_3$  в апреле (рисунок 2.27 б), при этом по расчетам XTM суточные максимумы  $O_3$  на 15–20 мкг/м<sup>3</sup> отличаются от наблюдавшихся: главный дневной максимум занижен и сдвинут на чуть более позднее время, слабо выраженный ночной максимум завышен.



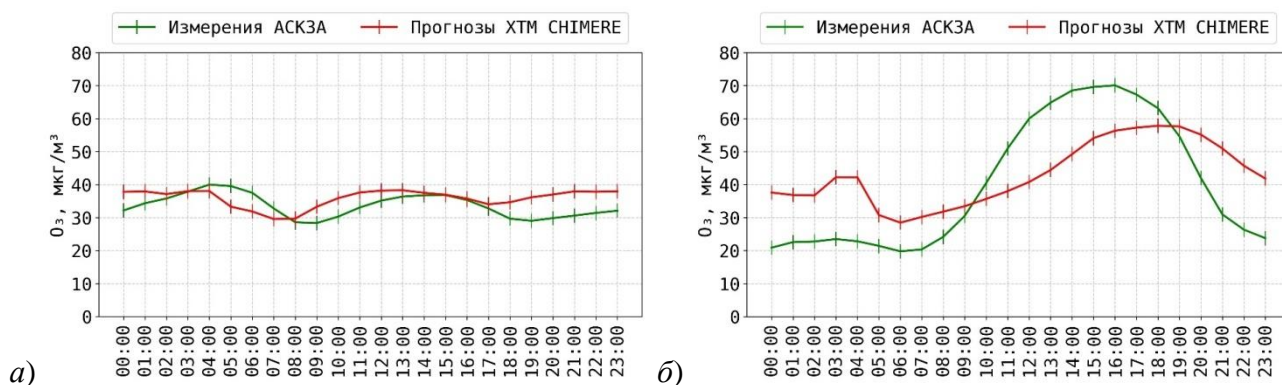


Рисунок 2.27 – Средний суточный ход концентраций  $O_3$  по измерениям на АСКЗА и прогнозам XTM CHIMERE в феврале (а) и апреле (б) 2023 г.

Проведенный анализ качества численных прогнозов концентраций  $PM_{10}$  и  $O_3$  по расчетам XTM CHIMERE на территории московского региона в 2023 году позволил выявить сезонные и суточные особенности отклонений прогнозов от измерений на АСКЗА, определить величины и знаки отклонений прогнозов каждого загрязнителя. По результатам проведенного анализа делается вывод о целесообразности применения современных методов повышения точности численных прогнозов, в частности, путем разработки и применения моделей машинного обучения.

## Выводы к главе 2

В главе 2 описан использованный в работе метод машинного обучения – искусственные нейронные сети. Основным преимуществом ИНС является выстраивание по мере обучения сложных нелинейных связей между предикторами и целевой переменной. Для повышения точности прогнозов и универсальности ИНС важно подобрать предикторы, максимизировать объем и разнообразие обучающей выборки, настроить гиперпараметры ИНС; возможно применение вариаций архитектур ИНС и методов улучшения оптимизации ИНС.

Сформирована исследовательская база данных, включающая:

- часовые измерения концентраций  $O_3$  и  $PM_{10}$  на территории московского региона за почти 6-летний период 1 января 2019 г. – 30 сентября 2024 г.;
- модельные прогнозы концентраций ЗВ и метеопараметров на территории московского региона на сетке с горизонтальным шагом  $2 \times 2$  км с временным шагом 1 час за практически тот же 6-летний период 7 февраля 2019 г. – 30 сентября 2024 г.;

– данные о распределении городской застройки по территории московского региона на той же сетке с горизонтальным шагом  $2 \times 2$  км.

По данным измерений проведен анализ сезонной и суточной изменчивости концентраций  $O_3$  и  $PM_{10}$  на территории московского региона за 5-летний период (2019–2024 гг.), описаны основные источники и атмосферные процессы, влияющие на изменчивость концентраций. Подробно проанализированы и описаны 7 эпизодов повышенного аэрозольного загрязнения в московском регионе в условиях дальнего переноса и/или НМУ. Проведенный анализ факторов формирования загрязнений  $O_3$  и  $PM_{10}$  служит ориентиром для подбора предикторов обучающих выборок для разработки ММО.

Описаны источники использованных в работе прогностических данных – мезомасштабная система COSMO-Ru и химическая транспортная модель CHIMERE. Выявленные по результатам многостороннего анализа отклонения прогнозов ХТМ CHIMERE от измерений концентраций  $O_3$  и  $PM_{10}$  служат обоснованием для разработки и реализации методики построения моделей машинного обучения для постпроцессинга численных прогнозов концентраций, которая обсуждается в главе 3.

### ГЛАВА 3. МЕТОДИКА РАЗРАБОТКИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОСТПРОЦЕССИНГА ЧИСЛЕННЫХ ПРОГНОЗОВ ПРИЗЕМНЫХ КОНЦЕНТРАЦИЙ ЗАГРЯЗНЯЮЩИХ ВЕЩЕСТВ И ЕЕ РЕАЛИЗАЦИЯ

#### 3.1 Общее описание и основные этапы методики

Разрабатываемая методика (далее – Методика) нацелена на повышение точности часовых численных прогнозов приземных концентраций загрязнителей атмосферы путем применения моделей машинного обучения. Точность численных прогнозов и ММО прогнозов приземных концентраций оценивается мерой их приближения к измеренным на АСКЗА значениям.

Основная суть методики заключается в использовании часовых численных прогнозов концентраций и метеопараметров в качестве предикторов обучения ММО и часовых измерений концентраций в соответствующих прогнозам ячейках модельной сетки в качестве целевой переменной прогноза ММО. Можно сказать, что таким образом ММО обучается приближать прогнозы концентраций к измерениям. При этом создается общая для территории ММО – обучающую выборку включаются пары прогноз-измерение для всех АСКЗА, а также в обучающую выборку включаются данные землепользования для воспроизведения ММО пространственной неоднородности загрязнения.

Преимущество такого подхода заключается в возможности применения обученной ММО ко всем ячейкам регулярной сетки в расчетном домене ХТМ, в том числе к тем, в которых не проводятся измерения. Предполагается улучшение первичных прогнозов ХТМ на всей территории расчетного домена с применением обученной ММО, назовем это ММО-постпроцессингом прогностических полей ХТМ.

Данный подход также положительно сказывается на обобщающей способности модели машинного обучения. Во-первых, для формирования обучающей выборки используются прогнозы в ячейках сетки (в данном случае площадью 4 км<sup>2</sup>); во-вторых, в обучающую выборку включаются прогнозы-измерения для всех станций. Так, ММО в меньшей степени будет учитывать особенности локации измерительной станции по сравнению с подходом обучения ММО на рядах измерений (см. разделы 1.3, 1.4).

На рисунке 3.1 изображена схема разработки и тестирования ММО для постпроцессинга численных прогнозов концентраций загрязнителей атмосферы. Исходя из целей работы и выбранных путей их достижения сформулированы основные этапы Методики:

1. Создание для каждого отдельного загрязнителя синхронизированного архива численных прогнозов концентраций и метеопараметров, данных землепользования и измерений

концентраций. Набор конкретных предикторов обучения ММО формируется для каждого отдельного загрязнителя исходя из представлений об основных факторах формирования загрязнения. Все доступные за заданный временной период часовые численные прогнозы концентраций и метеопараметров в узлах регулярной сетки (предикторы) синхронизируются по часу и ближайшему узлу сетки со всеми доступными часовыми измерениями концентраций (целевая переменная) на всех АСКЗА, попадающих в расчетный домен ХТМ. Данные землепользования, константные по времени, синхронизируются с измерениями на АСКЗА по ближайшему узлу сетки. Синхронизированный архив разделяется на обучающую и независимую тестовую выборки. Обучающая выборка должна включать максимально возможное количество вариаций формирования загрязнений в полном диапазоне изменчивости погодных условий – желательно, несколько лет прогнозов-измерений. Максимизация объема и вариативности обучающей выборки положительно сказывается на универсальности разрабатываемой ММО. Максимизация объема и разнообразия независимой тестовой выборки повышает репрезентативность верификации ММО.

2. Экспериментальный поиск оптимальной конфигурации модели машинного обучения производится для каждого загрязнителя по обучающей выборке методом k-блоковой валидации.

3. Верификация полученной оптимальной модели машинного обучения производится на независимой выборке не менее чем полугодового ряда. С использованием данных измерений концентраций загрязнителя на АСКЗА проводится расчет и сравнительный анализ метрик качества первичных прогнозов концентраций ХТМ и ММО-скорректированных прогнозов. Верификация включает обоснование применения ММО для постпроцессинга прогнозов ХТМ на всей территории расчетного домена, в том числе прогнозов в ячейках сетки без измерений концентраций. Для обоснования применяется метод исключения станции из обучающей выборки ММО (см. разделы 4.1, 5.1).

4. Применение разработанной ММО для постпроцессинга часовых прогнозов ХТМ приземных концентраций загрязнителя на всей территории расчетного домена (ММО-постпроцессинг прогностических полей ХТМ). Технология ММО-постпроцессинга прогностических полей ХТМ схематично изображена на рисунке 3.2 и включает следующие этапы:

а) сбор предикторов ММО из всех узлов сетки часовых прогностических полей концентраций и метеопараметров, рассчитанных ХТМ и моделью атмосферы, а также данных землепользования; переформатирование в ряды предикторов для усвоения ММО;

б) ММО на основе полученных рядов предикторов рассчитывает концентрации загрязнителя для каждого часа и узла сетки;

в) ММО-скорректированные прогнозы переформатируются в начальный вид часовых полей концентраций на регулярной сетке.

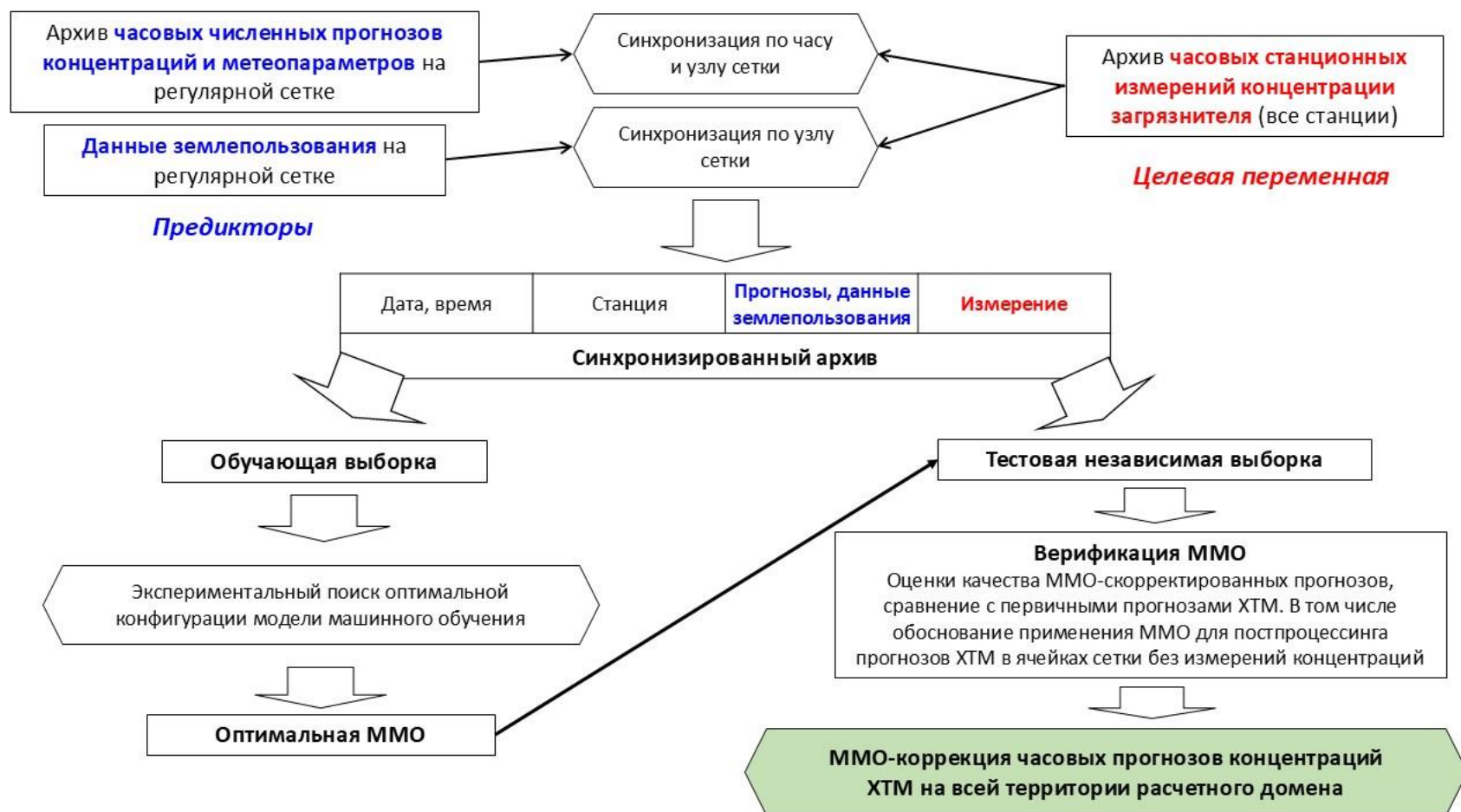
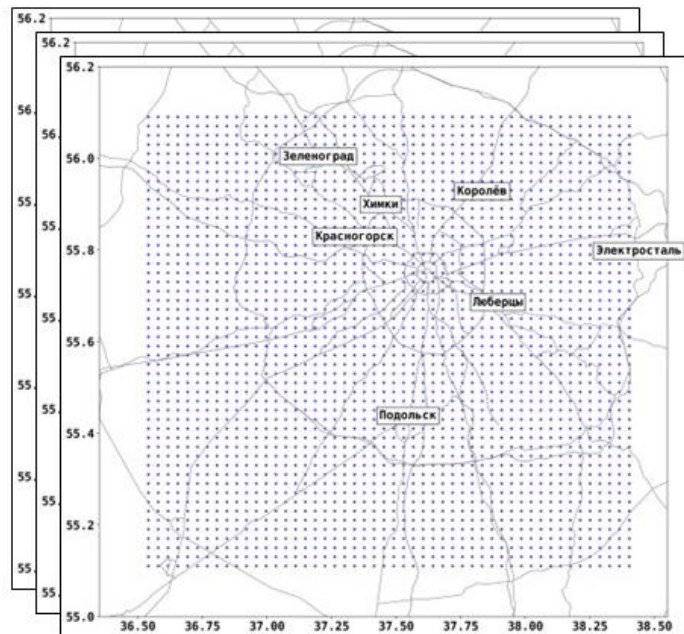


Рисунок 3.1 – Схема разработки и тестирования моделей машинного обучения для постпроцессинга численных прогнозов концентраций загрязнителей атмосферы

Часовые численные прогнозы концентраций, метеопараметров и данные землепользования в узлах регулярной сетки



Предикторы из каждого узла сетки:

- Концентрации
- Метеопараметры
- Данные землепользования

Обученная и верифицированная ММО для постпроцессинга прогнозов концентраций загрязнителя

ММО-скорректированные часовые прогнозы концентраций загрязнителя в узлах регулярной сетки

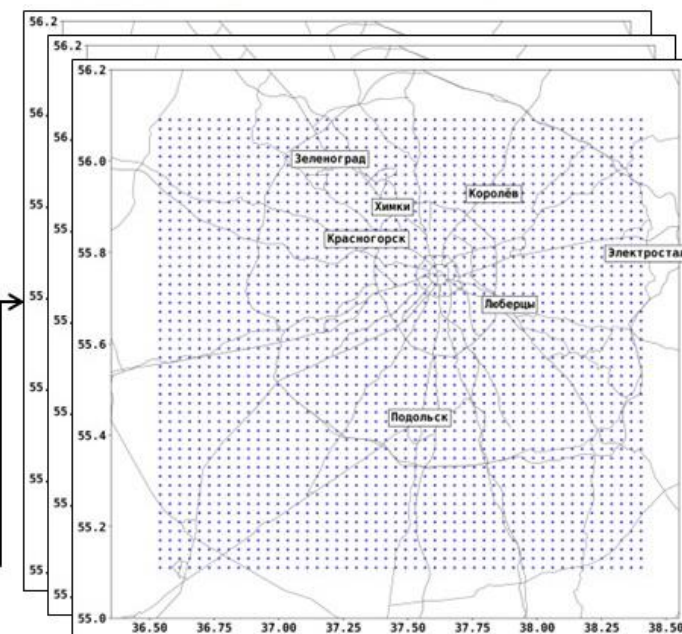


Рисунок 3.2 – Схема технологии ММО-постпроцессинга прогностических полей концентраций ХТМ

### **3.2 Реализация Методики в виде построения моделей машинного обучения для постпроцессинга численных прогнозов концентраций $O_3$ и $PM_{10}$**

Разработанная методика поэтапно реализована в ходе исследования в виде построения и верификации моделей машинного обучения для постпроцессинга численных прогнозов концентраций приземного озона и твердых взвешенных частиц  $PM_{10}$  на территории московского региона. Будем обозначать разработанные модели машинного обучения как ММО- $O_3$  и ММО- $PM_{10}$ .

В данном разделе описаны составы и объемы сформированных обучающих и независимых тестовых выборок, а также методы и результаты экспериментального поиска оптимальных конфигураций ММО- $O_3$  и ММО- $PM_{10}$ .

#### **3.2.1 Состав и объемы обучающих и тестовых выборок моделей постпроцессинга численных прогнозов концентраций $O_3$ и $PM_{10}$**

В таблице 3.1 представлены списки отобранных предикторов для разработки ММО- $O_3$  и ММО- $PM_{10}$ . Выбор предикторов ограничен описанными в главе 2 данными и осуществляется исходя из имеющихся представлений о физических свойствах  $O_3$  и  $PM_{10}$ , их источниках, сезонно-суточной изменчивости, зависимостях от крупномасштабных атмосферных процессов и локальных метеорологических условий (разделы 2.2.1, 2.2.2).

В синхронизированный архив включены часовые прогнозы приземных концентраций  $O_3$  и  $PM_{10}$ , рассчитанные ХТМ CHIMERE на первые сутки в узлах сетки с горизонтальным шагом  $2 \times 2$  км для территории московского региона (см. раздел 2.4, рисунок 2.18). Для ММО- $O_3$  также включены прогнозы прекурсоров озона – оксидов азота  $NO_x$ .

В архив также включены часовые прогнозы метеопараметров, рассчитанные моделью COSMO-Ru на первые сутки и интерполированные на сетку ХТМ CHIMERE: приземные и высотные метеорологические характеристики, включая температурные градиенты, скорости и направления ветра в АПС.

Для воспроизведения ММО- $O_3$  и ММО- $PM_{10}$  пространственной неоднородности загрязнения по территории московского региона использованы в качестве предикторов доли городской застройки в ячейках сетки ХТМ CHIMERE.

Таблица 3.1 – Предикторы, отобранные для разработки ММО-О<sub>3</sub> и ММО-PM<sub>10</sub>

Источник данных	Предикторы	
	ММО-О <sub>3</sub>	ММО-PM <sub>10</sub>
Часовые прогнозы ХТМ CHIMERE приземных концентраций на первые сутки	1. Озон (мкг/м <sup>3</sup> ) 2. Оксид азота NO (мкг/м <sup>3</sup> ) 3. Диоксида азота NO <sub>2</sub> (мкг/м <sup>3</sup> )	1. Взвешенные частицы PM <sub>10</sub> (мкг/м <sup>3</sup> )
Часовые прогнозы COSMO-Ru метеопараметров на первые сутки	4. Температура на 2 м (К) 5. Высота АПС (м) 6. Относительная влажность (0-1) 7. Осадки (кг/м <sup>2</sup> ) 8. Скорость ветра на 10 м (м/с) 9. Коротковолновая радиация (Вт/м <sup>2</sup> ) 10. Градиент температуры до 200 м (К) 11. Градиент температуры до 750 м (К) 12. Скорость ветра на 750 м (м/с) 13. Направление ветра на 750 м (sin, cos)	2. Температура на 2 м (К) 3. Высота АПС (м) 4. Относительная влажность (0-1) 5. Осадки (кг/м <sup>2</sup> ) 6. Скорость ветра на 10 м (м/с) 7. Градиент температуры до 200 м (К) 8. Градиент температуры до 400 м (К) 9. Градиент температуры до 750 м (К) 10. Скорость ветра на 200 м (м/с) 11. Скорость ветра на 400 м (м/с) 12. Скорость ветра на 750 м (м/с) 13. Скорость ветра на 1500 м (м/с) 14. Направление ветра на 750 м (sin, cos) 15. Направление ветра на 1500 м (sin, cos)
Данные землепользования базы GlobCover	14. Доля городской застройки в ячейке сетки ХТМ CHIMERE	16. Доля городской застройки в ячейке сетки ХТМ CHIMERE
Временная метка	15. Месяц (sin, cos) 16. День недели (sin, cos) 17. Час (sin, cos) 18. Заблаговременность численных прогнозов (1–24)	17. Месяц (sin, cos) 18. День недели (sin, cos) 19. Час (sin, cos) 20. Заблаговременность численных прогнозов (1-24)
Центральная Аэрологическая Обсерватория Росгидромета (Н. С. Иванова)	19. Среднесуточное содержание О <sub>3</sub> в атмосферном столбе (единицы Добсона е.Д.)	
Целевая переменная	Измерение О <sub>3</sub> на АСКЗА (мкг/м <sup>3</sup> )	Измерение PM <sub>10</sub> на АСКЗА (мкг/м <sup>3</sup> )

Для воспроизведения ММО сезонно-суточной изменчивости концентраций использовались календарные предикторы – месяц, день недели, час, а также показатели заблаговременности численных прогнозов для учета нарастания ошибок численных прогнозов со временем.

Для ММО-О<sub>3</sub> дополнительно включены ежедневные величины среднего многолетнего общего содержания озона (СО<sub>3</sub>) в Москве, предоставленные Н. С. Ивановой, Центральная Аэрологическая Обсерватория Росгидромета.



Значения направлений ветра и календарные предикторы закодированы в виде синуса и косинуса для сглаживания расстояний между крайними значениями. Для приведения всех предикторов к единому масштабу применялась стандартизация по среднему арифметическому и стандартному отклонению вектора.

Для исследований удалось сформировать синхронизированные архивы, охватывающие все часовые прогнозы-измерения  $O_3$  и  $PM_{10}$  на территории московского региона за период 7 февраля 2019 – 30 сентября 2024 гг., с небольшим интервалом отсутствия данных за июнь-сентябрь 2021 г. (отсутствуют прогнозы по техническим причинам). На рисунке 3.3 представлена количественная статистика архивов для разработки ММО- $O_3$  (а) и ММО- $PM_{10}$  (б). Синхронизированные архивы содержат ~700 и ~900 тыс. пар прогнозы-измерение  $O_3$  и  $PM_{10}$ , соответственно. Объем архива  $PM_{10}$  больше, поскольку измерения  $PM_{10}$  ведутся на большем количестве станций. В архиве  $O_3$  можно заметить (рисунок 3.3 а) увеличение количества пар от 2019 к 2024 году, что объясняется изменением расчетного домена ХТМ CHIMERE (рисунок 2.18) и открытием в течение последних 5 лет новых станций, измеряющих озон.

	2019	2020	2021	2022	2023	2024
январь		10304	12341	13427	14325	11569
февраль	5090	9035	9866	12467	13945	13578
март	6860	10265	13086	10857	15295	13192
апрель	6435	9194	12043	13411	13067	12595
май	6720	9474	6183	14432	14494	13343
июнь	6536	8243	0	10949	9989	11011
июль	6810	7568	0	10193	13459	13029
август	6910	9360	0	14909	12866	13847
сентябрь	7558	9454	0	7599	13135	13479
октябрь	7124	8962	15160	8146	13445	Тестовая выборка, объем
ноябрь	7366	9624	11749	14254	12491	90496 пар
декабрь	4382	10056	13897	12911	13640	Всего 697004 пар
	Обучающая выборка, объем 606508 пар					

а)

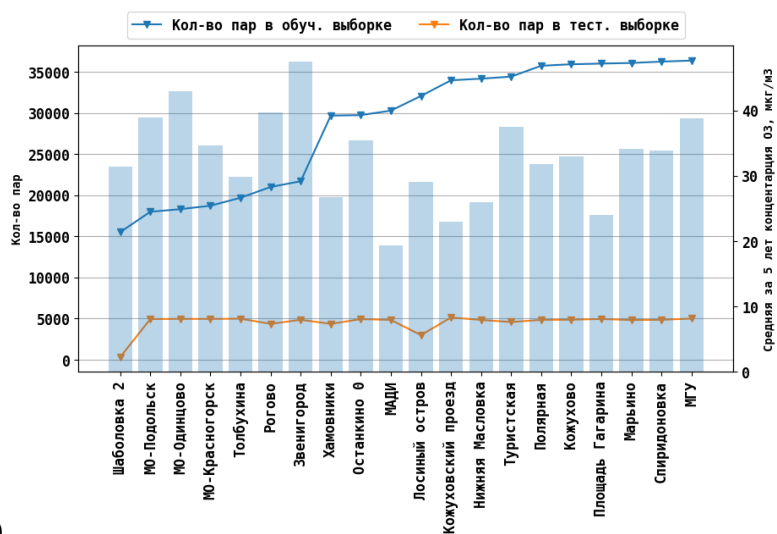
	2019	2020	2021	2022	2023	2024
январь		16908	13930	14368	14783	13801
февраль	11029	14674	12518	13334	13541	15414
март	14894	16655	14682	11496	15414	16435
апрель	14170	15904	13559	14434	14054	15202
май	14660	16889	7231	14893	14760	15917
июнь	14917	15224	0	12979	14109	15903
июль	14479	12243	0	13154	14103	15421
август	13337	10986	0	13591	13335	14481
сентябрь	14621	11643	0	7126	15348	14565
октябрь	14202	13167	13767	8386	14702	Тестовая выборка, объем 107924 пары
ноябрь	14544	11824	12133	14904	14027	
декабрь	7711	13002	15293	13009	14720	
	Обучающая выборка, объем 745366 пар					Всего 882505 пар

б)

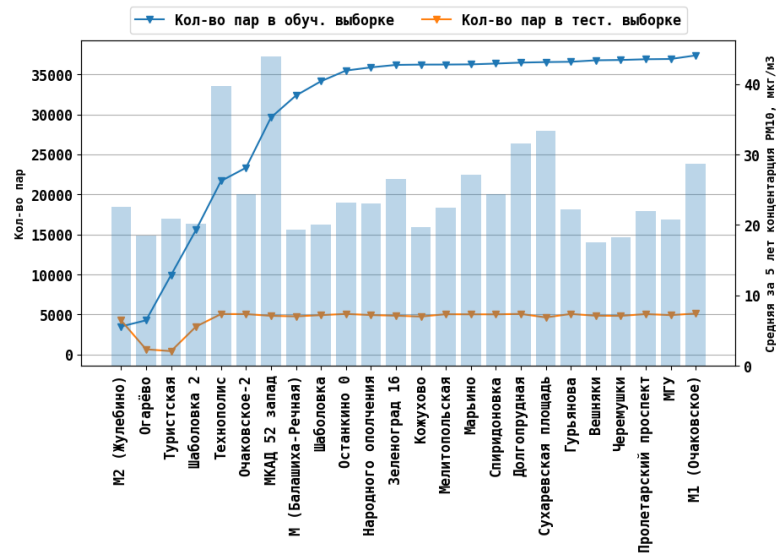
Рисунок 3.3 – Количество пар прогнозы-измерение в архивах для разработки (а) ММО-О<sub>3</sub> и (б) ММО-PM<sub>10</sub>

Для обучающих выборок ММО-О<sub>3</sub> и ММО-PM<sub>10</sub> принят период 7 февраля 2019 – 29 февраля 2024 гг., составляющий чуть более 5 лет прогнозов-измерений. Период март–сентябрь 2024 г. использован для тестовых независимых выборок. Объемы обучающих выборок составили ~600 тыс. пар прогнозы-измерение для ММО-О<sub>3</sub> и ~750 тыс. пар для ММО-PM<sub>10</sub>. Объемы независимых тестовых выборок составили ~90 тыс. пар прогнозы-измерение для ММО-О<sub>3</sub> (15 % от обучающей) и чуть более 100 тыс. пар для ММО-PM<sub>10</sub> (14 % от обучающей).

Обучающая выборка ММО-О<sub>3</sub> в основном состоит из городских АСКЗА – по 30–35 тыс. пар прогноз-измерение, по 15–22 тыс. пар представлено для станций за пределами Москвы – Подольск, Одинцово, Красногорск, Рогово, Звенигород. Тестовая независимая выборка ММО-О<sub>3</sub> содержит примерно равный для всех АСКЗА объем пар – по 5 тыс., кроме АСКЗА Шаболовка и Лосиный остров (рисунок 3.4 а).



а)



б)

Рисунок 3.4 – Постанционное количество пар прогнозы-измерение в архивах для разработки

(а) ММО- $O_3$  и (б) ММО- $PM_{10}$

Обучающая выборка ММО- $PM_{10}$  также в основном состоит из городских АСКЗА – по ~35 тыс. пар прогноз-измерение (рисунок 3.4 б). Для станций с максимальным уровнем загрязнения  $PM_{10}$  МКАД 52 и Технополис (открылась в 2023 г.) представлено 30 тыс. и 22 тыс. пар, соответственно. Тестовая независимая выборка ММО- $PM_{10}$  равномерно распределена почти по всем АСКЗА – по 5 тыс. пар, минимальное количество пар в тестовой выборке представлено для станций Огарево (за городом, к западу от Москвы) и Туристская (северо-запад Москвы).

### 3.2.2 Экспериментальный поиск оптимальных конфигураций моделей постпроцессинга численных прогнозов $O_3$ и $PM_{10}$

Методика включает процедуру поиска оптимальной конфигурации модели машинного обучения для постпроцессинга численных прогнозов каждого отдельного загрязнителя. Такой подход мотивирован тем, что не существует единых правил использования определенных архитектур и параметров ММО для конкретных задач. Конфигурация модели машинного обучения в основном определяется спецификой решаемой задачи, а именно количеством и сложностью связей между предикторами и целевой переменной. В случае искусственных нейронных сетей под определенную задачу необходимо подобрать архитектуру, объем и глубину сети, а также гиперпараметры – скорость обучения, количество эпох, размер пакета данных и др.

Общепринятой практикой поиска конфигурации модели машинного обучения является метод  $k$ -блоковой валидации, схематично изображенный на рисунке 3.10:

1. Задается набор экспериментальных моделей;

2. Обучающая выборка делится на  $k$  блоков;
3. Экспериментальная модель обучается  $k$  раз на  $k-1$  блоках, на оставшемся блоке (независимая выборка) рассчитываются метрики точности модели;
4. Итоговые метрики точности экспериментальной модели получаются осреднением метрик, полученных на  $k$  блоках;
5. Метрики точности экспериментальных моделей сравниваются между собой, отбирается наиболее точная модель.

Такой подход обеспечивает надежность оценок точности экспериментальных моделей, полученных по всей обучающей выборке, исключаются случайности, которые могут возникать при тестировании модели на части выборки.

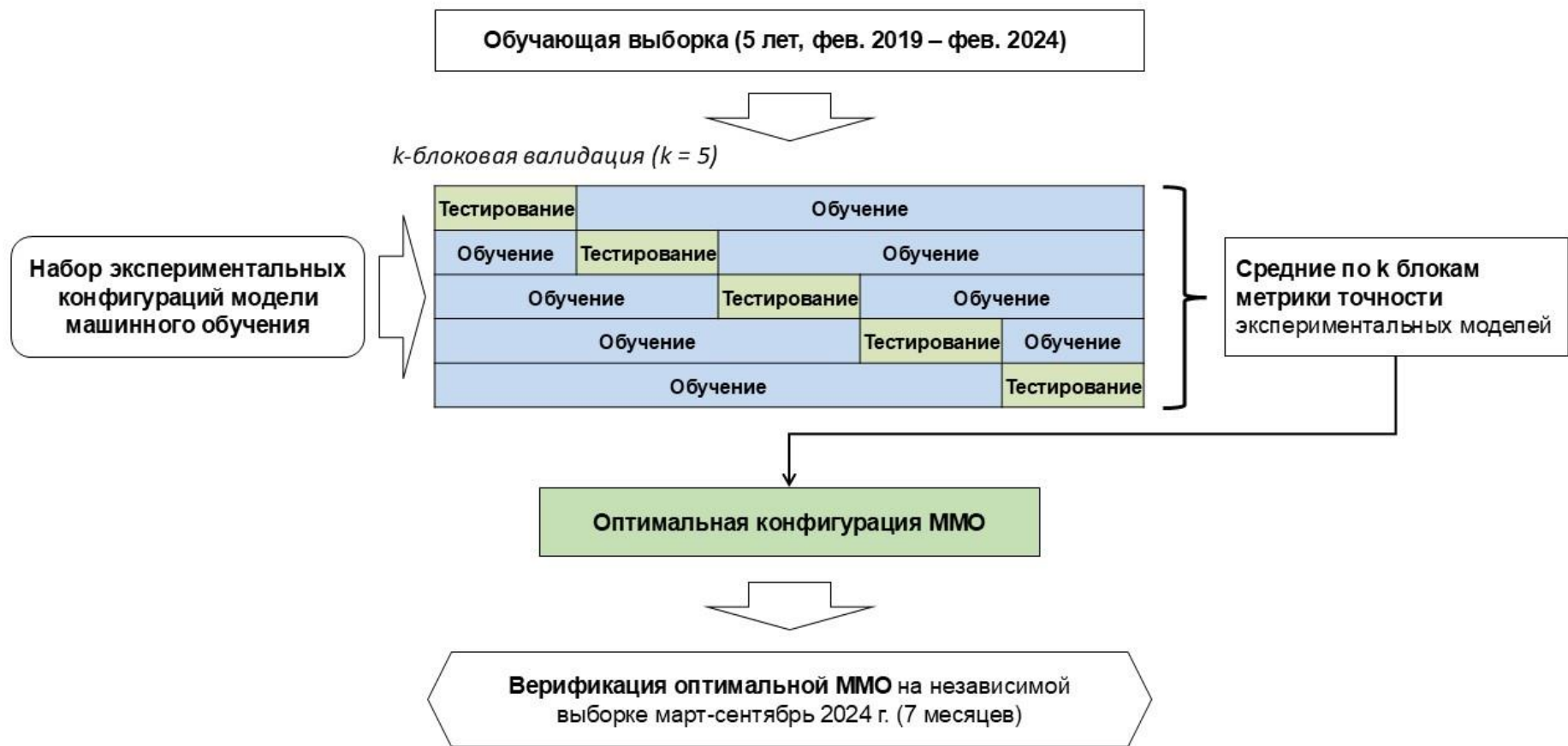


Рисунок 3.5 – Метод  $k$ -блоковой валидации для поиска оптимальной конфигурации модели машинного обучения

В данном исследовании период сформированных обучающих выборок ММО-О<sub>3</sub> и ММО-PM<sub>10</sub> составил 5 лет прогнозов-измерений, потому для  $k$ -блоковой валидации задано значение  $k = 5$ , чтобы блоки тестирования включали полную годовую изменчивость концентраций О<sub>3</sub> и PM<sub>10</sub>.

В таблицах 3.2 и 3.3 представлены списки апробированных конфигураций ММО-О<sub>3</sub> и ММО-PM<sub>10</sub>, соответственно. В качестве базовой в обоих случаях принималась модель линейной регрессии, остальные конфигурации – это искусственные нейронные сети (модели нелинейной регрессии), включая:

1. Полносвязные архитектуры ИНС с вариациями глубины (количество слоев) и ширины (количество нейронов в каждом слое);
2. Рекуррентные LSTM архитектуры ИНС (с длиной временной последовательности 12 часов);
3. Полносвязные и рекуррентные архитектуры ИНС с исключениями нейронов в слоях (dropout);
4. Вариации стандартного отклонения ( $\sigma$ ) аддитивного гауссовского шума в данных обучающей выборки (аугментация данных, см. раздел 2.1.3).

Все приведенные в таблицах ИНС обучались на протяжении 100 эпох с шагом обучения 0.001, размер пакета 128. Для всех приведенных ИНС включена моментная оптимизация с коэффициентом импульса 0.9. Всего апробировано 36 конфигураций ММО-О<sub>3</sub> (таблица 3.2) и 19 конфигураций ММО-PM<sub>10</sub> (таблица 3.3).

Для каждой экспериментальной конфигураций ММО в таблицах 3.2 и 3.3 приведены метрики точности, осредненные по  $k = 5$  тестовым блокам: среднее абсолютное отклонение прогнозов (САО); среднее квадратическое отклонение прогнозов (СКО); коэффициент корреляции прогноз-измерение ( $r$ ), в скобках приводится оценка неопределенности метрики – стандартное отклонение ( $\sigma$ ) по  $k = 5$  блокам. Экспериментальные конфигурации ММО в обеих таблицах отсортированы по возрастанию САО.

Таблица 3.2 – Экспериментальные конфигурации ММО-О<sub>3</sub>

№	Архитектура	п слоев	п нейронов в каждом слое	Dropout	$\sigma$ шума	САО ( $\sigma$ по 5 блокам)	СКО ( $\sigma$ по 5 блокам)	$r$ ( $\sigma$ по 5 блокам)
1	2	3	4	5	6	7	8	9
1	П	4	7	-	-	14,3 (1,6)	35,0 (37,7)	0,63 (0,26)
2	П	3	30	0.5	-	14,5 (1,5)	35,1 (37,6)	0,63 (0,25)
3	П	3	15	-	0.1	14,5 (1,5)	35,3 (37,5)	0,63 (0,25)

Продолжение таблицы 3.2

1	2	3	4	5	6	7	8	9
4	П	3	15	-	-	14,6 (1,6)	35,3 (37,5)	0,63 (0,25)
5	LSTM	1	10	0.5	-	14,7 (1,8)	35,6 (38,9)	0,63 (0,26)
6	П	3	30	0.5	0.1	14,7 (1,5)	35,4 (37,4)	0,63 (0,25)
7	П	3	15	-	0.5	14,8 (1,4)	35,7 (37,3)	0,62 (0,25)
8	П	4	25	-	-	14,9 (1,4)	35,8 (37,2)	0,61 (0,24)
9	LSTM	3	15	-	-	14,9 (1,7)	36,2 (38,7)	0,61 (0,25)
10	LSTM	1	30	0.5	-	14,9 (1,7)	36,5 (38,5)	0,60 (0,24)
11	П	3	20	-	-	15,0 (1,6)	35,9 (37,2)	0,61 (0,25)
12	LSTM	3	15	-	0.1	15,0 (1,5)	36,9 (38,3)	0,60 (0,24)
13	LSTM	3	30	0.5	0.1	15,0 (1,5)	37,1 (38,2)	0,60 (0,24)
14	П	4	30	0.5	-	15,1 (1,8)	36,2 (37,0)	0,60 (0,25)
15	П	3	10	-	-	15,2 (2,2)	36,6 (36,9)	0,60 (0,25)
16	П	4	14	0.5		15,3 (1,4)	36,0 (37,2)	0,60 (0,24)
17	LSTM	3	20	0.5		15,3 (1,8)	37,3 (38,2)	0,58 (0,25)
18	П	3	30	0.5	0.5	15,4 (1,4)	36,2 (37,1)	0,61 (0,25)
19	П	3	25	-	-	15,4 (2,0)	36,4 (37,1)	0,60 (0,25)
20	П	4	35	-	-	15,5 (1,5)	36,5 (37,0)	0,59 (0,24)
21	П	4	55	-	-	15,5 (1,4)	37,1 (36,7)	0,58 (0,23)
22	П	2	15	-	-	15,7 (3,0)	37,5 (36,6)	0,58 (0,26)
23	П	4	15	-	-	16,0 (3,3)	38,3 (36,4)	0,56 (0,27)
24	П	2	25	-	-	16,3 (3,7)	38,8 (36,3)	0,57 (0,26)
25	П	2	35	-	-	16,3 (3,3)	38,4 (36,3)	0,58 (0,25)
26	П	3	30	0.5	1	16,6 (1,6)	37,4 (36,6)	0,57 (0,24)
27	П	2	55	-	-	16,7 (3,4)	38,8 (36,2)	0,55 (0,26)
28	П	3	55	-	-	17,5 (4,9)	40,1 (36,0)	0,53 (0,26)
29	П	3	35	-	-	17,5 (5,5)	41,8 (36,2)	0,53 (0,28)
30	П	3	100	-	-	17,9 (4,8)	41,0 (35,6)	0,52 (0,25)
31	П	2	100, 50	-	-	18,1 (5,9)	41,6 (36,1)	0,53 (0,27)
32	П	2	100	-	-	18,2 (5,7)	41,7 (36,1)	0,53 (0,28)

## Окончание таблицы 3.2

1	2	3	4	5	6	7	8	9
33	П	3	75	-	-	19,0 (7,8)	43,5 (36,5)	0,51 (0,29)
34	П	2	75	-	-	19,8 (9,7)	46,2 (38,4)	0,50 (0,32)
35	ЛР	-	-	-	-	27 (3,5)	46,8 (30,7)	0,44 (0,18)

*Примечание: П – полносвязная ИНС; LSTM – рекуррентная LSTM ИНС; ЛР – модель линейной регрессии; Dropout – доля исключенных нейронов в каждом слое;  $\sigma$  шума – стандартное отклонение аддитивного гауссовского шума в данных обучающей выборки; CAO – осредненное по  $k = 5$  блокам среднее абсолютное отклонение прогнозов; СКО – осредненное среднее квадратическое отклонение прогнозов;  $r$  – осредненный коэффициент корреляции прогноз-измерение.*

Из таблицы 3.2 видно, что наиболее точный прогноз концентраций  $O_3$  показала полносвязная ИНС глубиной в 4 слоя и шириной в 7 нейронов без дополнительных настроек. Чуть менее точный прогноз показали 3-слойные полносвязные ИНС бóльшей ширины (№ 2–4, 6, 7 в таблице 3.2) с исключениями нейронов в слоях и с добавлением шума в данные, а также рекуррентная сеть с исключением нейронов в единственном слое (№ 5 в таблице 3.2).

Из ИНС наименее точный прогноз показали 2–3-слойные модели увеличенной ширины без исключений нейронов в слоях и без добавления гауссовского шума в данные (№ 30–34 в таблице 3.2). Линейная модель (№ 35 в таблице 3.2) показала заметно менее точный прогноз по сравнению со всеми ИНС как по среднему абсолютному отклонению, так и по коэффициенту корреляции.

Таблица 3.3 – Экспериментальные конфигурации ММО-PM<sub>10</sub>

№	Архитектура	п слоев	п нейронов в каждом слое	Dropout	$\sigma$ шума	CAO ( $\sigma$ по 5 блокам)	СКО ( $\sigma$ по 5 блокам)	$r$ ( $\sigma$ по 5 блокам)
1	LSTM	1	30	0.5	-	11,6 (0,8)	22,4 (2)	0,51 (0,08)
1	П	3	50	0.5	-	11,6 (0,7)	51,9 (66,9)	0,44 (0,2)
2	П	3	30	0.5	-	11,8 (0,8)	52,1 (66,9)	0,43 (0,2)
3	LSTM	3	15	0.5	-	11,8 (0,8)	23,4 (2,4)	0,49 (0,04)
4	П	3	15	-	0.1	11,8 (0,9)	51,9 (67)	0,44 (0,2)
5	П	3	25	-	-	11,9 (0,7)	51,9 (66,9)	0,44 (0,2)
6	П	4	14	0.5	-	11,9 (0,7)	52,3 (66,8)	0,42 (0,19)
7	П	3	25	-	0.1	12 (1)	52 (66,9)	0,43 (0,2)
8	П	4	7	-	-	12,2 (0,5)	52,5 (66,6)	0,41 (0,19)
9	П	4	7	0.5	-	12,2 (0,8)	52,6 (66,6)	0,41 (0,19)



Продолжение таблицы 3.3

1	2	3	4	5	6	7	8	9
10	LSTM	1	10	0.5	-	12,3 (1,1)	23,6 (2,3)	0,48 (0,07)
11	LSTM	3	15	-	-	12,3 (0,4)	24,1 (1,5)	0,48 (0,05)
12	П	3	15	-	-	12,4 (1)	52,4 (66,7)	0,42 (0,19)
13	П	4	25	-	-	12,4 (0,8)	52,6 (66,4)	0,42 (0,18)
14	П	4	14	-	-	12,8 (1,4)	53,2 (66,3)	0,38 (0,19)
15	П	2	35	-	-	12,8 (1,6)	52,8 (66,4)	0,41 (0,2)
16	П	2	75	-	-	13,4 (1,9)	53,1 (66,2)	0,41 (0,2)
17	П	2	100, 50	-	-	13,8 (3,1)	54,4 (65,7)	0,4 (0,21)
18	ЛР	-	-	-	-	22,1 (2,3)	61,1 (62,6)	0,25 (0,11)
19	LSTM	1	30	0.5	-	11,6 (0,8)	22,4 (2)	0,51 (0,08)

*Примечание: П – полносвязная ИНС; LSTM – рекуррентная LSTM ИНС; ЛР – модель линейной регрессии; Dropout – доля исключенных нейронов в каждом слое;  $\sigma$  шума – стандартное отклонение аддитивного гауссовского шума в данных обучающей выборки; CAO – осредненное по  $k = 5$  блокам среднее абсолютное отклонение прогнозов; СКО – осредненное среднее квадратическое отклонение прогнозов;  $r$  – осредненный коэффициент корреляции прогноз-измерение.*

Для PM<sub>10</sub> (таблица 3.3) наибольшую точность прогнозов концентраций показала рекуррентная LSTM сеть (№ 1 в таблице 3.3) с небольшой разницей CAO по сравнению с полносвязными архитектурами (№ 2, 3 в таблице 3.3). Положительно сказывается исключение нейронов в слоях ИНС (№ 1–4 в таблице 3.3). Как и в случае с MMO-O<sub>3</sub>, все ИНС заметно превосходят по CAO и  $r$  линейную модель (№ 19 в таблице 3.3).

Из таблицы 3.3 видно преимущество рекуррентных LSTM архитектур перед полносвязными при сравнении средних квадратических отклонений прогнозов и коэффициентов корреляции – СКО меньше в среднем в 2 раза,  $r$  больше на 0.05–0.1, хотя при сравнении средних абсолютных отклонений большой разницы не наблюдается. При этом нужно учитывать, что среднеквадратическое отклонение более чувствительно к выбросам в данных, чем среднее абсолютное отклонение.

При сравнительном анализе эффективности наилучших рекуррентной LSTM (№ 1 в таблице 3.3) и полносвязной (№ 2 в таблице 3.3) архитектур MMO-PM<sub>10</sub> на тестовой выборке март-сентябрь 2024 г. было выявлено переобучение рекуррентной LSTM модели – при исключении станции из обучающей выборки (см. раздел 5.1) LSTM модель показывала менее точный прогноз для этой станции, чем полносвязная модель и XTM CHIMERE, т. е. эффекты MMO-постпроцессинга отсутствовали. На основании этого можно предположить, что LSTM модель при обучении «запомнила» входящие в обучающую выборку станции, прогнозируя,

например, аномально высокие концентрации  $PM_{10}$  для примагистральной станции МКАД 52 запад. Такое переобучение негативно сказывается на обобщающем свойстве ММО- $PM_{10}$ , учитывая сформулированную задачу ММО-постпроцессинга прогнозов по всей территории, в том числе в ячейках сетки без измерений концентраций.

Лучшая из полносвязных архитектур ИНС – 3-слойная (по 50 нейронов в слое) сеть с исключениями нейронов в слоях (№ 2 в таблице 3.3) – была выбрана в качестве оптимальной ММО для постпроцессинга прогнозов  $PM_{10}$ .

Наименее точный прогноз концентраций  $PM_{10}$  показали 2-слойные полносвязные ИНС увеличенной ширины без дополнительных настроек (№ 17, 18 в таблице 3.3).

Также можно обратить внимание на то, лучшая ИНС для прогноза  $O_3$  (№1 в таблице 3.2) показала среднюю точность прогноза  $PM_{10}$  (№ 9 в таблице 3.3), что свидетельствует о разнице в связях, необходимых для точного воспроизведения концентраций  $O_3$  и  $PM_{10}$ , учитывая, что количество предикторов примерно одинаковое – 19 для ММО- $O_3$  и 20 для ММО- $PM_{10}$ .

По итогам экспериментального поиска получены оптимальные конфигурации ММО- $O_3$  – полносвязная 4-слойная (по 7 нейронов в слое) искусственная нейронная сеть; и ММО- $PM_{10}$  – полносвязная 3-слойная (по 50 нейронов в слое с исключениями) искусственная нейронная сеть.

Проведен анализ вкладов отдельных предикторов, включенных в полученные оптимальные ММО- $O_3$  и ММО- $PM_{10}$ . Оценки значимости предикторов проводились методом случайного перемешивания значений предиктора по временному ряду тестовой независимой выборки март-сентябрь 2024 г., такой подход исключает предиктор из тестовой выборки. Метрика значимости предиктора рассчитывалась как среднеквадратическое отклонение прогнозов концентраций ММО от измерений по тестовой выборке – чем выше  $SKO_{п-и}$ , т.е. чем менее точный прогноз ММО на тестовой выборке с исключенным предиктором, тем значимее данный предиктор.

В таблице 3.4 представлены ранжированные по убыванию значимости предикторы ММО- $O_3$  и ММО- $PM_{10}$ .

Больше других предикторов ММО- $O_3$  приоритизирует численный прогноз приземной температуры и текущий час (таблица 3.4,  $SKO_{п-и}$  около  $28 \text{ мкг/м}^3$ ), т.е. модель восстанавливает суточную изменчивость концентраций  $O_3$ , что будет далее подтверждено результатами верификации (глава 4). На 3 и 4 местах по значимости для ММО- $O_3$  с  $SKO_{п-и}$  около  $26 \text{ мкг/м}^3$  стоят текущий месяц (сезонная изменчивость концентраций  $O_3$ ) и прогноз относительной влажности. На 5 и 6 месте ( $SKO_{п-и}$   $24\text{--}25 \text{ мкг/м}^3$ ) – среднесуточное общее содержание озона и прогноз ХТМ концентрации прекурсора озона  $NO_2$ . Метрики значимости остальных предикторов, в том числе прогноза ХТМ концентрации  $O_3$ , варьируются в узком диапазоне  $SKO_{п-и} \approx 22\text{--}23 \text{ мкг/м}^3$ .

Таблица 3.4 – Метрики значимости отдельных предикторов в ММО-О<sub>3</sub> и ММО-PM<sub>10</sub>.

№	Предиктор	СКО <sub>П-И</sub> (мкг/м <sup>3</sup> ) ММО-О <sub>3</sub> на тестовой выборке с перемешанным предиктором	Предиктор	СКО <sub>П-И</sub> (мкг/м <sup>3</sup> ) ММО-PM <sub>10</sub> на тестовой выборке с перемешанным предиктором
1	T <sub>2м</sub>	28,35	Месяц	34,54
2	Час	28,05	T <sub>2м</sub>	33,11
3	Месяц	26,06	Tgrad <sub>400м</sub>	31,35
4	ρ	25,73	Tgrad <sub>750м</sub>	27,25
5	Среднесуточное ОСО	24,72	h <sub>АПС</sub>	27,16
6	NO <sub>2</sub>	23,75	Час	27,08
7	SRad	23,33	Доля города в ячейке	26,91
8	O <sub>3</sub>	23,23	V <sub>400м</sub>	26,81
9	Заблаговременность прогноза	23,03	ρ	26,80
10	Tgrad <sub>200м</sub>	23,02	V <sub>1500м</sub>	26,73
11	Доля города в ячейке	22,47	V <sub>200м</sub>	26,67
12	V <sub>750м</sub>	22,21	dd <sub>750м</sub>	26,67
13	h <sub>АПС</sub>	22,17	V <sub>750м</sub>	26,59
14	NO	21,91	dd <sub>1500м</sub>	26,55
15	Tgrad <sub>750м</sub>	21,89	День недели	26,53
16	V <sub>10м</sub>	21,86	V <sub>10м</sub>	26,49
17	dd <sub>750м</sub>	21,84	Заблаговременность прогноза	26,48
18	Precip	21,77	Precip	26,44
19	День недели	21,75	PM <sub>10</sub>	26,43
20			Tgrad <sub>200м</sub>	26,35

*Примечание: T – численный прогноз температуры (K); ρ – относительной влажности (0-1); Tgrad – градиента температуры (K); V – скорости ветра (м/с); dd – направления ветра (sin, cos); Precip – осадков (кг/м<sup>2</sup>); h<sub>АПС</sub> – высоты АПС (м); SRad – коротковолновой радиации (Вт/м<sup>2</sup>). O<sub>3</sub>, NO, NO<sub>2</sub>, PM<sub>10</sub> – прогнозы приземных концентраций ХТМ (мкг/м<sup>3</sup>). ОСО – общее содержание озона (e.Д.).*

ММО-PM<sub>10</sub> при прогнозе концентрации приоритизирует в первую очередь текущий месяц, учитывая сезонную изменчивость концентраций PM<sub>10</sub>. На втором месте с СКО<sub>П-И</sub> ~33 мкг/м<sup>3</sup> стоит численный прогноз приземной температуры, на третьем месте – прогноз градиента температуры в слое 2–400 м (СКО<sub>П-И</sub> 31 мкг/м<sup>3</sup>), отражающий термическую устойчивость в АПС. Метрики значимости остальных предикторов имеют небольшой разброс СКО<sub>П-И</sub> ≈ 26–27 мкг/м<sup>3</sup>, на 4–6 местах по значимости стоят прогноз градиента температуры в слое

750 м, прогноз высоты пограничного слоя и текущий час (суточная изменчивость концентраций  $PM_{10}$ ).

В главах 4 и 5 обсуждаются результаты верификации полученных оптимальных ММО- $O_3$  и ММО- $PM_{10}$  на независимой выборке март-сентябрь 2024 г., в том числе эффекты ММО-постпроцессинга прогнозов в периоды повышенного озонового и аэрозольного загрязнения, приводятся иллюстрации ММО-постпроцессинга часовых прогностических полей концентраций  $O_3$  и  $PM_{10}$ , смоделированных ХТМ.

### Выводы к главе 3

Разработана методика построения моделей постпроцессинга численных прогнозов приземных концентраций загрязняющих веществ на основе методов машинного обучения. Основу методики составляют (а) формирование обучающей и тестовой независимой выборок из синхронизированного архива часовых пар численных прогнозов (концентраций и метеопараметров) и измерений концентраций целевого загрязнителя; (б) экспериментальный поиск по обучающей выборке для каждого отдельного загрязнителя оптимальной конфигурации ММО для постпроцессинга прогнозов концентраций; (в) верификация полученной оптимальной ММО на тестовой независимой выборке; (г) применение ММО для постпроцессинга часовых прогностических полей концентраций, смоделированных ХТМ.

Разработанная методика поэтапно реализована в ходе исследования в виде построения и верификации моделей машинного обучения для постпроцессинга численных прогнозов концентраций приземного озона и твердых взвешенных частиц  $PM_{10}$  на территории московского региона (ММО- $O_3$  и ММО- $PM_{10}$ ).

С учетом факторов формирования озонового и аэрозольного загрязнения в московском регионе отобраны предикторы для разработки ММО- $O_3$  и ММО- $PM_{10}$ . Сформированы обучающие выборки ММО- $O_3$  и ММО- $PM_{10}$ , включающие 5-летний период часовых прогнозов-измерений февраль 2019 г. – февраль 2024 г., а также тестовые независимые выборки, включающие 7 месяцев прогнозов-измерений март-сентябрь 2024 г.

Путем экспериментального поиска методом  $k$ -блоковой валидации получены оптимальные конфигурации ММО- $O_3$  и ММО- $PM_{10}$ . Все экспериментальные ИНС заметно превосходят по точности прогнозов концентраций модель линейной регрессии, положительно на точности прогнозов сказываются углубление ИНС, исключение нейронов в слоях ИНС и добавление гауссовского шума в данные обучающей выборки. Проведен анализ вкладов отдельных предикторов, включенных в полученные оптимальные ММО- $O_3$  и ММО- $PM_{10}$ .

#### ГЛАВА 4. ВЕРИФИКАЦИЯ РАЗРАБОТАННОЙ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОСТПРОЦЕССИНГА ЧИСЛЕННЫХ ПРОГНОЗОВ КОНЦЕНТРАЦИЙ ПРИЗЕМНОГО ОЗОНА

Проверка разработанных моделей машинного обучения (ММО-PM<sub>10</sub> и ММО-O<sub>3</sub>) на независимых данных и ее результаты служат основанием для получения выводов об эффективности разработанной Методики. Для независимой проверки создана база данных за период 1 марта – 30 сентября 2024 г., включающая ежедневные часовые прогнозы концентраций O<sub>3</sub> и PM<sub>10</sub> на первые сутки, рассчитанные ХТМ SHIMERE, ММО-PM<sub>10</sub> и ММО-O<sub>3</sub>, и ряды часовых измерений O<sub>3</sub> и PM<sub>10</sub> на АСКЗА в московском регионе. Рассчитаны статистические метрики качества первичных ХТМ и ММО-скорректированных прогнозов O<sub>3</sub> и PM<sub>10</sub>.

С целью приведения оценок качества прогнозов в формат установленных ВОЗ гигиенических нормативов рассчитаны суточные максимумы средних за 8 часов концентраций O<sub>3</sub> (O<sub>3-8ч</sub>) и средние суточные концентрации PM<sub>10</sub> (PM<sub>10-CC</sub>), для которых нормативы ВОЗ составляют 100 и 45 мкг/м<sup>3</sup>, соответственно [World Health Organization, 2021]. В отдельных случаях проводится сравнение уровня загрязнения с российскими гигиеническими нормативами.

Проведен анализ сезонных эффектов ММО-постпроцессинга и прогнозируемой суточной изменчивости концентраций O<sub>3</sub> и PM<sub>10</sub>. С целью оценки эффектов ММО-постпроцессинга на территориях с различными уровнями антропогенной нагрузки проведен анализ для отдельных АСКЗА, при этом АСКЗА последовательно исключались из обучающей выборки ММО для обоснования ММО-постпроцессинга прогнозов ХТМ в ячейках сетки без измерений концентраций.

Особую актуальность и интерес представляют озоновые и аэрозольные эпизоды, сопровождаемые существенным повышением уровня O<sub>3</sub> и PM<sub>10</sub>; обсуждения эпизодов повышенных концентраций приземного озона и взвешенных частиц представлены автором в ряде публикаций [Борисов и др., 2020; Борисов и Кузнецова, 2023; 2024]. В диссертационной работе уделено значительное внимание причинам и условиям возникновения эпизодов повышенного загрязнения в Москве в период независимой проверки (март-сентябрь 2024 г.) и, соответственно, анализу успешности их предсказания ХТМ и разработанными моделями машинного обучения. При анализе используется термин «неблагоприятные метеорологические условия» (НМУ), закрепленный в п.3 статьи 19 Федерального закона от 04.05.1999 № 96-ФЗ «Об охране атмосферного воздуха». НМУ – это неблагоприятные для рассеивания примеси и очищения приземного воздуха от загрязнений метеорологические условия [Кузнецова и др., 2024], при наступлении которых проводится регулирование выбросов в атмосферу с целью

предотвращения опасного для здоровья уровня загрязнения. Этим определяется особая важность прогнозирования загрязнения приземного воздуха при формировании НМУ.

#### **4.1 Результаты постпроцессинга численных прогнозов концентраций приземного озона с применением разработанной модели машинного обучения**

Верификации ММО-О<sub>3</sub> предшествовал анализ данных мониторинга загрязнений и атмосферных условий в период независимой проверки. Необходимо подчеркнуть, что в 2024 году в московском регионе наблюдался нетипичный сезонный ход концентраций приземного озона. Уже в конце марта, при установлении аномально теплой погоды с повышением температуры до + 10 °С, на отдельных АСКЗА были зафиксированы случаи нехарактерного повышения приземного озона до 80 мкг/м<sup>3</sup>. Весенний максимум озона, обычно приходящийся на апрель, наблюдался в конце мая во время волн жары с повышением температуры до + 25, 30 °С [Иванова и др., 2024]. А годовой максимум озона (вместо типичного апрельского) пришелся на июль, в котором озона в приземном воздухе оказалось почти на 40 % больше, чем в среднем за предыдущие 5 лет. Фактором нетипичного повышения уровня О<sub>3</sub> в июле стала череда блокирующих антициклонов с волнами жары и сухой погоды. В августе и сентябре 2024 г. повышение озонового загрязнения также наблюдалось при аномально жаркой и сухой погоде [Иванова и др., 2025]. Не исключено, что имеется связь нехарактерных изменений приземного озона в 2024 г. с имевшими место крупномасштабными аномалиями общего содержания озона (ОСО) и наблюдаемой максимальной солнечной активностью [Иванова и др., 2025], но эти вопросы выходят за рамки исследований автора. В общем можно констатировать, что в теплый сезон 2024 г. в московском регионе содержание приземного озона было больше обычного. Отмеченные особенности приземного озона в период тестирования будут использованы при интерпретации результатов.

Результаты сравнений концентраций О<sub>3-8ч</sub>, рассчитанных по данным всех АСКЗА, с первичными прогнозами ХТМ и прогнозами ММО-О<sub>3</sub> представлены на рисунке 4.1, здесь показаны боксовые распределения отклонений «прогноз-измерение» (соответственно, Δ<sub>п-и</sub> ХТМ и Δ<sub>п-и</sub> ММО-О<sub>3</sub>) в каждом месяце тестовой выборки.

Несмотря на нетипичную сезонную изменчивость приземного озона в московском регионе в теплый период 2024 г., эффект ММО-постпроцессинга прогнозов О<sub>3</sub> проявился во все месяцы и выражен в сужении разброса отклонений прогнозов ММО-О<sub>3</sub> и уменьшении СКО<sub>п-и</sub>(ММО-О<sub>3</sub>) по сравнению с прогнозами ХТМ (рисунок 4.1). Наблюдается общая

закономерность «регулирования» модели машинного обучения: ММО-О<sub>3</sub> чаще всего «снижает» завышенные ХТМ концентрации О<sub>3</sub>.

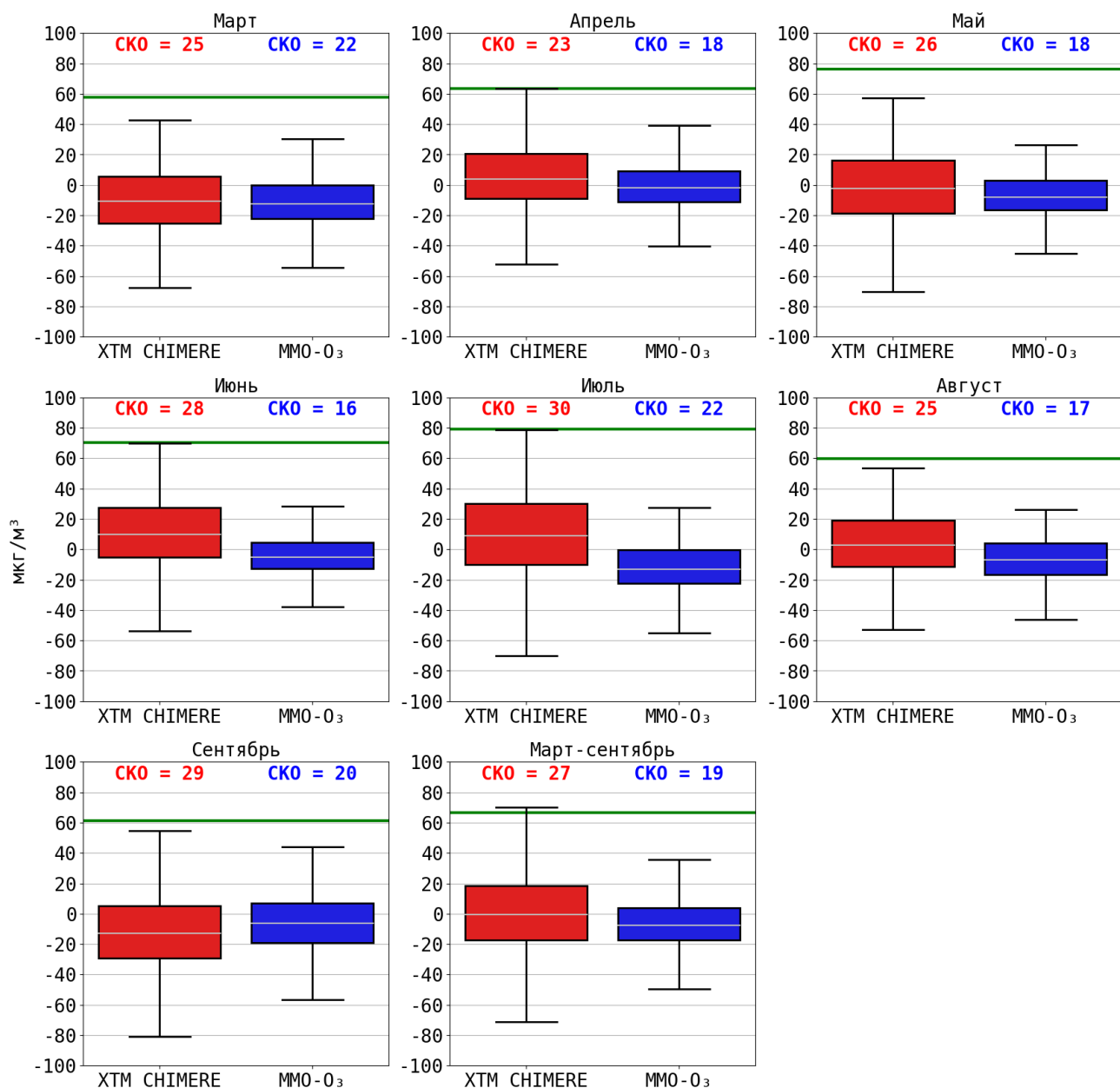


Рисунок 4.1 – Отклонения прогнозов О<sub>3-8ч</sub> ХТМ CHIMERE (красные боксы) и ММО-О<sub>3</sub> (синие боксы) от измерений (все АСКЗА). Независимая выборка, март-сентябрь 2024 г.

*Примечание:* В боксе – IQR  $\Delta_{П-и}$ , линия внутри бокса – среднее  $\Delta_{П-и}$ , нижний «ус»:  $Q1 - 1.5 * IQR \Delta_{П-и}$ , верхний «ус»:  $Q3 + 1.5 * IQR \Delta_{П-и}$ . Зеленые сплошные линии – средняя за месяц измеренная концентрация О<sub>3-8ч</sub> (все АСКЗА). Над боксами SKO<sub>П-и</sub> ХТМ и ММО-О<sub>3</sub>.

Следует отметить небольшие отличия результатов в отдельные месяцы. В марте, при средней измеренной концентрации О<sub>3-8ч</sub> около 60 мг/м<sup>3</sup>, ХТМ CHIMERE и ММО-О<sub>3</sub> почти

одинаково немного занижают концентрации  $O_{3-8ч}$  (в среднем на 11–13  $мкг/м^3$ ), но  $СКО_{П-и}(ММО-O_3)$  меньше, чем  $СКО_{П-и}(ХТМ)$ , IQR сужен с -25(Q1)-6(Q3)  $мкг/м^3$  (ХТМ) до -22-0  $мкг/м^3$  (ММО- $O_3$ ).

В апреле, мае и июне, при удовлетворительной точности прогнозов ХТМ,  $\Delta_{П-и}(ММО-O_3)$  еще меньше,  $СКО_{П-и}$  с применением ММО-постпроцессинга уменьшилось на 5–12  $мкг/м^3$ .

В июле ХТМ CHIMERE в среднем завысила прогнозы  $O_{3-8ч}$  на 9  $мкг$ , в условиях повышенного уровня приземного озона ММО- $O_3$  занижала концентрацию – в среднем за месяц на 13  $мкг/м^3$ . IQR  $\Delta_{П-и}$  сужен с применением ММО-постпроцессинга с -10-30  $мкг$  (ХТМ) до -23-0  $мкг$  (ММО- $O_3$ ),  $СКО_{П-и}$  уменьшилось на 8  $мкг/м^3$ . Примерно такое же уменьшение  $\Delta_{П-и} O_{3-8ч}$  с применением ММО- $O_3$  наблюдается и в августе.

При нетипичном повышенном приземном озоне в сентябре ХТМ CHIMERE в среднем занижала концентрации  $O_{3-8ч}$  на 13  $мкг/м^3$ , ММО- $O_3$  повышала прогнозы CHIMERE – среднее  $\Delta_{П-и}$  (ММО- $O_3$ ) смещено на уровень -6  $мкг/м^3$ . IQR  $\Delta_{П-и}$  сужен с применением ММО-постпроцессинга с -29-5  $мкг/м^3$  (ХТМ) до -19-7  $мкг/м^3$  (ММО- $O_3$ ),  $СКО_{П-и}$  уменьшено с 29 до 20  $мкг/м^3$ .

В среднем по всей тестовой выборке  $СКО_{П-и} O_{3-8ч}$  уменьшилось с применением ММО-постпроцессинга с 27  $мкг/м^3$  (ХТМ CHIMERE) до 19  $мкг/м^3$  (ММО- $O_3$ ) – практически на 30%.

Важно оценивать применимость ММО-постпроцессинга для воспроизведения концентраций озона в диапазоне нарушения гигиенических нормативов, именно поэтому в работе уделено внимание модельному суточному ходу. На рисунке 4.2 представлены осредненные за месяц почасовые концентрации  $O_3$  по расчетам ХТМ, ММО- $O_3$  и по измерениям на всех АСКЗА в тестовой выборке.

Можно отметить, что ХТМ CHIMERE в основном недооценивает максимальный/дневной уровень  $O_3$ , а в июне и июле немного завышает величину суточного максимума.

На рисунке 4.2 видны ярко выраженные эффекты ММО-постпроцессинга суточного хода концентраций  $O_3$ : а) в каждом из семи месяцев расчеты ММО- $O_3$  ближе, чем расчеты ХТМ, к измеренным концентрациям; б) ММО- $O_3$  точнее воспроизводит время наступления и величину суточного максимума  $O_3$  и значительно уменьшает завышенный ХТМ ночной озон. Коэффициенты корреляции между часовыми прогнозами и измерениями  $O_3$  увеличены в мае-сентябре до 0.8–0.9 с применением ММО-постпроцессинга, в марте и апреле – до 0.7. По всей тестовой выборке март-сентябрь коэффициент корреляции увеличен с 0.5 до 0.8.



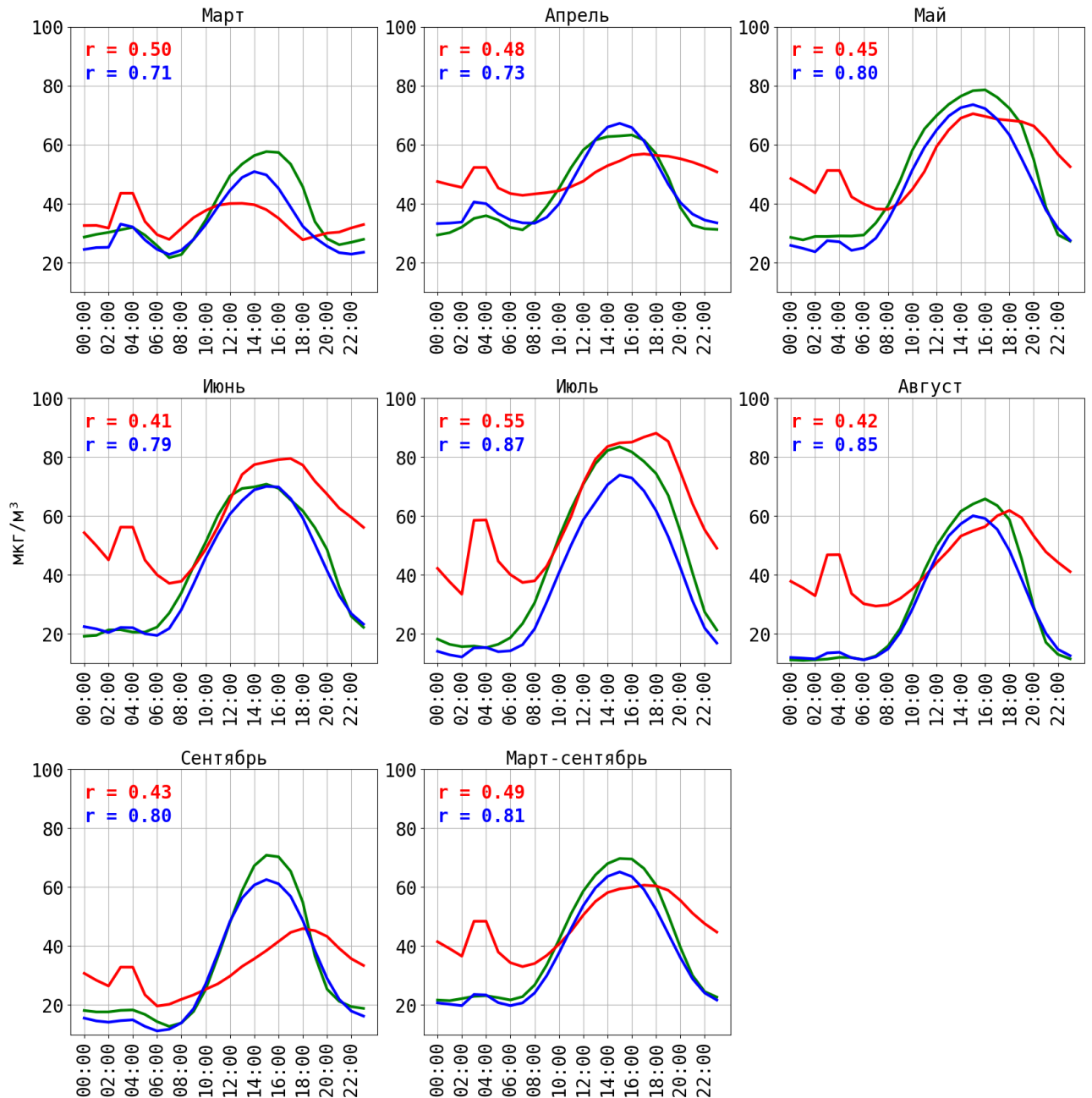


Рисунок 4.2 – Средние почасовые концентрации  $O_3$  по данным измерений на АСКЗА (зеленые линии), по расчетам ХТМ CHIMERE (красные линии) и ММО- $O_3$  (синие линии). Независимая выборка, март-сентябрь 2024 г.

*Примечание:* В левом верхнем углу – коэффициенты корреляции между часовыми прогнозами и измерениями  $O_3$  (все АСКЗА) по расчетам ХТМ CHIMERE (красным) и ММО- $O_3$  (синим).

Проведен анализ эффектов ММО-постпроцессинга численных прогнозов концентраций  $O_3$  для локаций отдельных АСКЗА (по ближайшему узлу сетки) с различными уровнями антропогенной нагрузки. На рисунке 4.3 (а) представлены  $СКО_{П-и}$  и средние отклонения прогнозов ( $СО_{П-и}$ )  $O_{3-8ч}$  по расчетам ХТМ CHIMERE и ММО- $O_3$ , рассчитанные для отдельных

АСКЗА за период март-сентябрь 2024 г. На рисунке 4.3 (б) визуализированы эффекты ММО-постпроцессинга прогнозов ХТМ концентраций  $O_3$  на отдельных АСКЗА в виде увеличения/уменьшения  $СКО_{П-и}$  (в %) по всей тестовой выборке с применением ММО- $O_3$ .

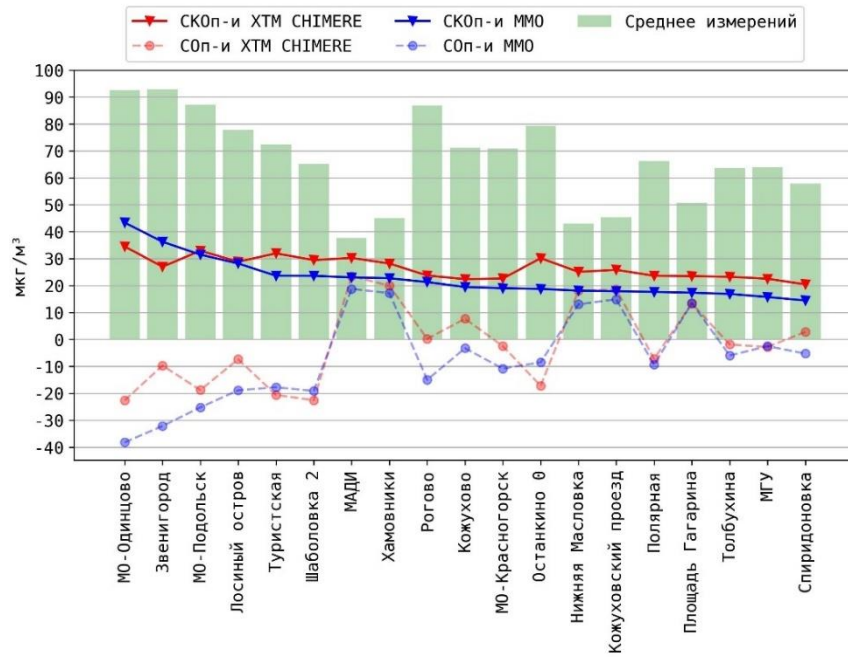
Необходимо отметить, что метрики эффективности ММО-постпроцессинга на отдельных АСКЗА (рисунок 4.3 а, б) получены с учетом исключения станции из обучающей выборки – ММО- $O_3$  последовательно переобучалась  $n = \text{количество станций}$  раз с исключением определенной АСКЗА из обучающей выборки и расчетом прогнозов концентраций для исключенной АСКЗА на тестовой выборке. Такой подход применялся, чтобы симитировать и оценить эффекты ММО-постпроцессинга численных прогнозов  $O_3$  на территориях без измерений концентраций.

При анализе динамики  $СКО_{П-и}$  (рисунок 4.3 б) эффект ММО-постпроцессинга прогнозов ХТМ наблюдается на всех АСКЗА, кроме двух – Звенигород и Одинцово. Заметим, что АСКЗА Звенигород и Одинцово находятся на удалении от Москвы к западу, в отсутствие крупных источников транспортных выбросов, разрушающих озон. На этих АСКЗА зафиксированы наибольшие из всех АСКЗА средние за период концентрации  $O_{3-8ч}$  – более  $90 \text{ мкг/м}^3$ ,  $СКО_{П-и}$  ХТМ CHIMERE на этих станциях  $28\text{--}25 \text{ мкг/м}^3$ , ММО- $O_3$  –  $36\text{--}43 \text{ мкг/м}^3$  (рисунок 4.3 а).

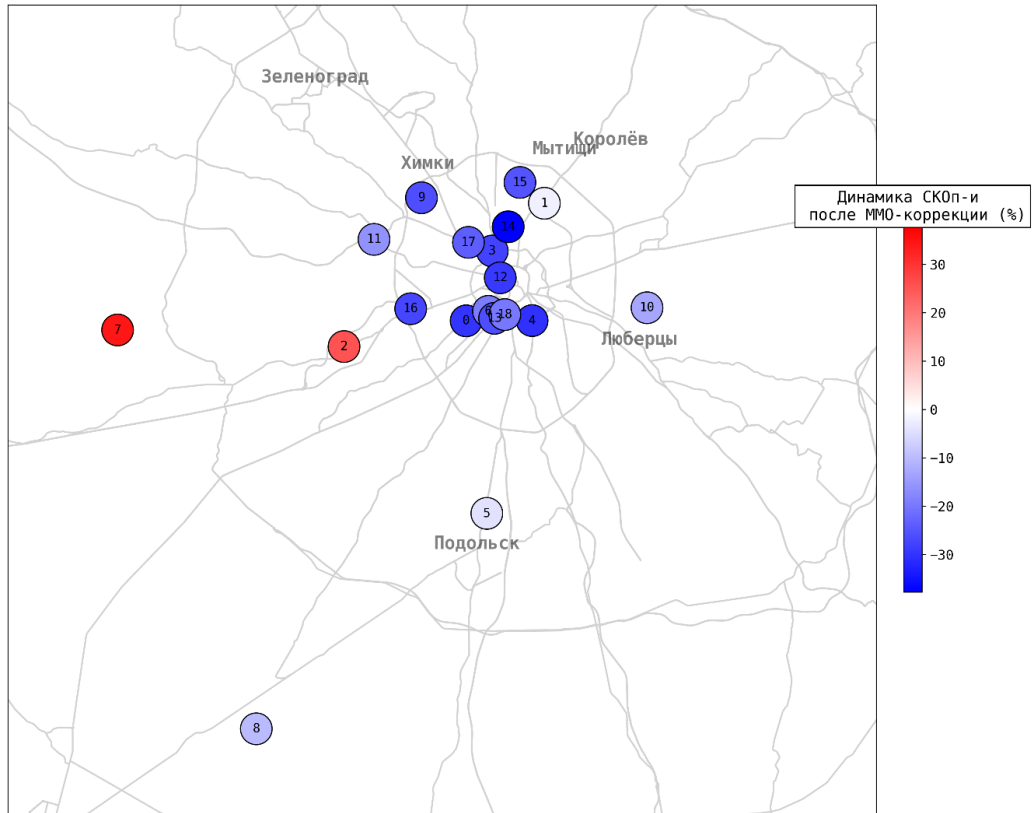
Минимальные эффекты ММО-постпроцессинга наблюдаются на фоновых станциях с высоким уровнем озонового загрязнения  $\sim 75\text{--}85 \text{ мкг/м}^3$  Подольск и Лосинный остров –  $СКО_{П-и}$  и CHIMERE, и ММО- $O_3$  около  $30 \text{ мкг/м}^3$ . При этом на фоновой станции Рогово с похожим высоким уровнем озонового загрязнения  $\sim 85 \text{ мкг/м}^3$   $СКО_{П-и}$  ММО-скорректированных прогнозов около  $20 \text{ мкг/м}^3$ . Максимальный из всех АСКЗА эффект ММО-постпроцессинга получился на городской станции Останкино с высоким средним уровнем  $O_{3-8ч}$   $80 \text{ мкг/м}^3$ .

Для всех АСКЗА на территории Москвы, за исключением АСКЗА Лосинный остров, ММО- $O_3$  успешно корректирует прогноз CHIMERE, снижая  $СКО_{П-и}$  приблизительно на 30 % – с  $25\text{--}30 \text{ мкг/м}^3$  до  $15\text{--}25 \text{ мкг/м}^3$ , вне зависимости от территориальной принадлежности и уровня загрязнения на АСКЗА (рисунок 4.3 а).

При анализе средних отклонений прогнозов видно (рисунок 4.3 а), что на фоновых станциях с отсутствующими или минимальными эффектами ММО-постпроцессинга – Одинцово, Звенигород, Подольск, Лосинный остров – ММО- $O_3$  в среднем занижает прогнозы  $O_{3-8ч}$ , ХТМ CHIMERE также занижает прогнозы для этих станций, но в меньшей степени. Завышения прогнозов  $O_{3-8ч}$  ХТМ CHIMERE видны на примагистральных станциях с низким уровнем озона – МАДИ, Хамовники, Нижняя Масловка, Кожуховский проезд; ММО- $O_3$  корректирует прогнозы, рассчитывая с средним меньшие концентрации  $O_{3-8ч}$  для этих станций. Для всех остальных станций, в том числе за пределами города – Рогово, Кожухово, Красногорск, средние отклонения прогнозов ММО- $O_3$  в пределах  $\pm 15 \text{ мкг}$ .



a)



0 - МГУ	5 - МО-Подольск	10 - Кожухово	15 - Полярная
1 - Лосинный остров	6 - Хамовники	11 - МО-Красногорск	16 - Толбухина
2 - МО-Одинцово	7 - Звенигород	12 - Спиридоновка	17 - МАДИ
3 - Нижняя Масловка	8 - Рогово	13 - Площадь Гагарина	18 - Шаболовка 2
4 - Кожуховский проезд	9 - Туристская	14 - Останкино 0	

б)

Рисунок 4.3 – (а) среднеквадратическое (SKOp-и) и среднее (COп-и) отклонение прогнозов  $O_3$ -8ч, рассчитанных XTM CHIMERE (красные линии) и MMO- $O_3$  (синие линии) от измерений на отдельных АСКЗА; (б) изменения (в %) SKOp-и прогнозов  $O_3$ -8ч с применением MMO-постпроцессинга на отдельных АСКЗА. Март-сентябрь 2024 г.

Из приведенных результатов можно сделать вывод о том, что ММО-О<sub>3</sub> значительно корректирует прогнозы ХТМ концентраций приземного озона на всех городских территориях, эффекты ММО-постпроцессинга минимальны для некоторых фоновых АСКЗА и отсутствуют только на двух АСКЗА на западе московской области с максимальными уровнями озонowego загрязнения.

#### **4.2 Эффекты ММО-постпроцессинга численных прогнозов в периоды повышенного озонowego загрязнения**

Значительное повышение концентраций озона в приземном воздухе, превышение гигиенических нормативов принято называть озоновыми эпизодами. В озоновых эпизодах возможно образование фотохимического смога; высокие концентрации озона представляют опасность для здоровья, негативно воздействуя на органы дыхания и сердечно-сосудистую систему [WHO Regional Office for Europe, 2003; Еланский и др., 2022; Андреев и др., 2023, Zhang et al., 2019]. Имеются указания на рост летальности при длительном сохранении озоновых эпизодов, в частности такие оценки сделаны на примере лета 2010 г. в работе [Звягинцев и др., 2011]. В исследовании [Achebak et al., 2024] отмечено, что в летние сезоны 2015–2017 годов в Европе 114 447 человек погибло от респираторных заболеваний, спровоцированных озоновым загрязнением воздуха, основными источниками которого авторами названы трансграничный перенос, морской транспорт и крупные промышленные объекты.

В московском регионе озоновые эпизоды формируются в теплый сезон при установлении аномально жаркой и сухой безветренной погоды и наблюдаются в послеполуденные часы. Необходимым условием фотохимического образования опасных концентраций О<sub>3</sub> является высокий уровень УФ-радиации и повышенные концентрации в городском воздухе прекурсоров озона (NO<sub>x</sub> и ЛОС) в утренние часы.

Важно подчеркнуть, что ММО-О<sub>3</sub> разрабатывалась на полной выборке метеорологических условий в московском регионе, с доминирующими условиями умеренного и интенсивного рассеивания при повторяемости НМУ в Москве в среднем 5–8 %. Проверка ММО-О<sub>3</sub> на эпизодах повышенного озонowego загрязнения априори завышает ожидания успешности ММО-постпроцессинга, поскольку в обучающей выборке таких событий статистически недостаточно.

Для идентификации повышенного озонowego загрязнения в период тестирования март-сентябрь 2024 г. использовался рекомендованный ВОЗ и используемый в России норматив 100 мкг/м<sup>3</sup> (О<sub>3-8ч</sub>) [World Health Organization, 2021], условно обозначим его ПДК<sub>8ч</sub>. Озоновые

эпизоды отбирались по критерию превышения ПДК<sub>8ч</sub> на более чем 40 % АСКЗА, измеряющих О<sub>3</sub> в московском регионе.

В эпизоде, начавшемся в конце мая, на 11 АСКЗА измеренные О<sub>3-8ч</sub> превышали ПДК<sub>8ч</sub>. С 28 мая по 2 июня приземная температура повышалась до 26–31 °С, при относительной влажности воздуха 25–31 %, и каждую ночь в нижнем 300-метровом слое формировались приземные температурные инверсии 2–7 °С. Имеет значение, что регион более недели находился под влиянием переноса очень теплой обогащенной озоном воздушной массы из южных регионов ЕТР, чему способствовал блокирующий антициклон над южным Уралом. В темное время суток по всем признакам преобладали НМУ.

На АСКЗА Звенигород, Одинцово, Лосиный остров, находящихся вдали от крупных источников транспортных выбросов, максимальные за сутки О<sub>3-8ч</sub> повышались до 150 мкг/м<sup>3</sup>, на городских станциях МГУ, Полярная, Толбухина, Туристская – до 125 мкг/м<sup>3</sup> (рисунок 4.4).

ХТМ СИМЕРЕ 28 мая – 02 июня рассчитывала суточные максимумы О<sub>3-8ч</sub> на городских АСКЗА в интервале 50–90 мкг/м<sup>3</sup> и 75–125 мкг/м<sup>3</sup> для фоновых станций Рогово и Лосиный остров (рисунок 4.4). ММО-О<sub>3</sub> рассчитывала концентрации О<sub>3-8ч</sub> в пределах 75–100 мкг/м<sup>3</sup> для всех станций. Отметим, что на большинстве городских АСКЗА ММО-О<sub>3</sub> корректирует рассчитанные ХТМ прогнозы О<sub>3-8ч</sub>, повышая их на 10–25 мкг/м<sup>3</sup> и приближая к измеренным концентрациям. Минимальные эффекты ММО-постпроцессинга в данном эпизоде наблюдались на загородных АСКЗА Рогово, Одинцово, Лосиный остров.

Озоновые эпизоды в июле с кратковременным повышением концентраций выше гигиенических нормативов формировались в неординарных атмосферных условиях. К ним относятся в первую очередь волны жары: с начала июля и до 18 числа днем воздух прогревался до + 28, 32 °С и до + 33, 35 °С на волне самой сильной жары 2–5 июля. Только 5–6 июля волну 30-ти градусной жары прервал атмосферный фронт с дождями, слабые осадки 13 и 15 июля прошли местами. На фоне малоградиентного барического поля почти каждую ночь наблюдались температурные инверсии при слабом ветре преимущественно из юго-западного сектора.

При том, что трудно признать линейную связь приземного озона с общим содержанием озона в атмосфере (ОСО), следует принять во внимание отмеченные в публикациях ЦАО [Иванова и др. 2024; 2025] факты, свидетельствующие о неординарности процессов в атмосфере, приведших к самому высокому за период наблюдений ОСО. Во 2 и 3 квартале 2024 практически над всей РФ средние за месяц значения ОСО на 5–10 % были выше средних многолетних значений; в мае-августе 2024 г. ОСО в Москве на 6 % превышало среднее многолетнее значение. В [Иванова и др., 2025] отмечено, что связанный с ОСО ультрафиолетовый индекс (УФИ), превышающий значение 5.5 (порог, выше которого необходимо использовать средства защиты от УФ-излучения), в Московском регионе

наблюдался 1 день в июне (28 июня УФИ = 5.8), 7 дней в июле (1, 7–10, 12 и 17). То есть Москва временами находилась в условиях повышенной УФ-облученности, что важно для понимания повышенного содержания озона в приземном воздухе.

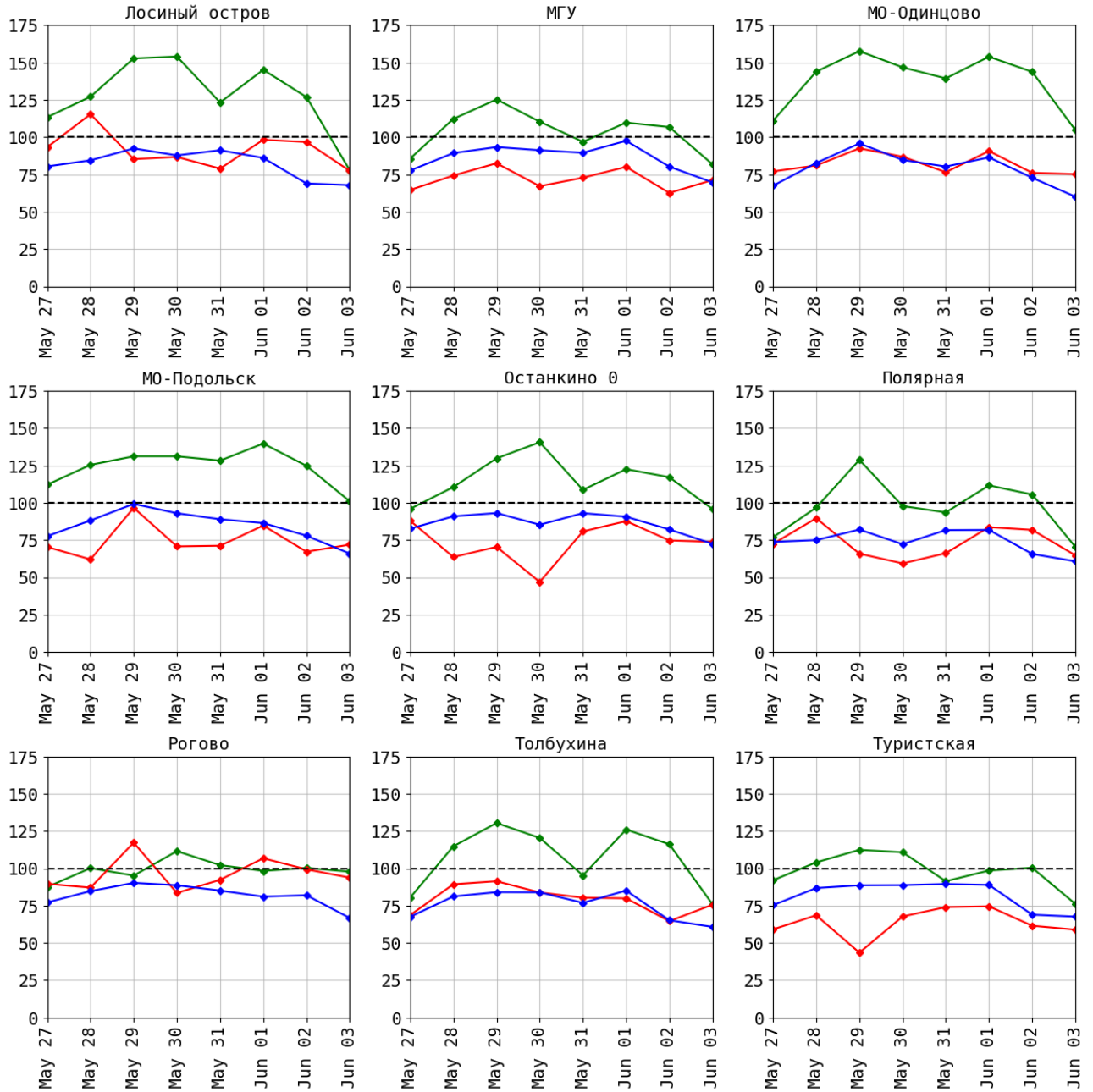


Рисунок 4.4 –  $O_{3-8ч}$  ( $\mu\text{кг}/\text{м}^3$ ) по измерениям на АСКЗА (зеленые линии), рассчитанные ХТМ SHIMERE (красные линии) и ММО- $O_3$  (синие линии) в эпизоде повышенного озонowego загрязнения 28 мая – 2 июня 2024 г. (независимая выборка)

*Примечание:* Пунктирная линия – ПДК<sub>8ч</sub>  $O_3$   $100 \mu\text{кг}/\text{м}^3$ .

Возможно, также имеет значение то, что в теплый сезон 2024 г. наблюдались необычно высокие значения полярного арктического озона (ПАО – среднее значение ОСО в зональном

поясе 63–90 ° с. ш.). Среднемесячное значение ПАО в июне 2024 г. оказалось максимальным за все годы наблюдений с 1979 г. Все дни июля, кроме 25, и почти все дни августа (кроме 3, 6, 16–20 и 30) ПАО обновлял ранее зарегистрированные с 1979 г. максимальные значения для этих дней [Иванова и др., 2025].

Авторы [Иванова и др. 2024; 2025] высказали предположение, что необычное содержание ПАО отчасти является откликом на резко возросшую солнечную активность. По данным спутника GOES в мае зарегистрирована 21 солнечная вспышка класса X – самых мощных солнечных вспышек. В июне таких вспышек было две, они произошли за сутки до начала аномалий полярного арктического озона (<https://www.spaceweatherlive.com/ru/solnechnaya-aktivnost.html>).

Подчеркнем, что на основе установленных особенностей воспроизведения суточного хода  $O_3$  моделью CHIMERE и ММО- $O_3$  в июле (рисунок 4.2), еще до анализа результатов ММО-постпроцессинга сделан обоснованный вывод о целесообразности применить к озоновым эпизодам второй нормативный критерий – ПДК<sub>М.Р.</sub> 160 мкг/м<sup>3</sup> для разовой концентрации  $O_3$ . Очевидно, что критерий ПДК<sub>М.Р.</sub> намного мягче ПДК<sub>8ч.</sub>, что вносит неоднозначность при оценке степени опасности ситуации. Так, с 1 по 18 июля на АСКЗА МГУ по критерию ПДК<sub>8ч.</sub> наблюдалось 5 дней с его нарушением, по критерию ПДК<sub>М.Р.</sub> – всего 1 день; на АСКЗА Полярная 8 и 2 дня, соответственно; на АСКЗА Толбухина 10 и 1, на АСКЗА Останкино 14 и 5.

На рисунке 4.5 представлены данные измерений и модельные расчеты концентраций  $O_3$  в длительном июльском эпизоде повышенного озонового загрязнения, в двух форматах – суточные максимумы часовых концентраций  $O_3$  в соответствии с ПДК<sub>М.Р.</sub> и суточные максимумы средних за 8 часов концентраций  $O_3$  в соответствии с ПДК<sub>8ч.</sub> При любом из критериев (ПДК<sub>М.Р.</sub>, ПДК<sub>8ч.</sub>) ХТМ CHIMERE, даже при систематических максимальных завышениях июльских концентраций  $O_3$  (!) (разд. 2.4.1, рисунок 2.25), в июле 2024 г. преимущественно занижает концентрации приземного озона. ММО- $O_3$  в основном еще более снижает прогнозы концентраций  $O_3$  ХТМ, поскольку обучена на данных за период февраль 2019 – февраль 2024 гг. корректировать завышенные прогнозы ХТМ в июле. Успешным ММО-постпроцессингом прогнозов  $O_3$  в данном эпизоде можно считать только снижение концентраций на примагистральных станциях с активной деструкцией озона (например, Кожуховский проезд на рисунке 4.5 д).

Суточные максимумы часовых концентраций O<sub>3</sub> (ПДК<sub>М.Р.</sub>)

Суточные максимумы средних за 8 часов концентраций O<sub>3</sub> (ПДК<sub>8ч</sub>)

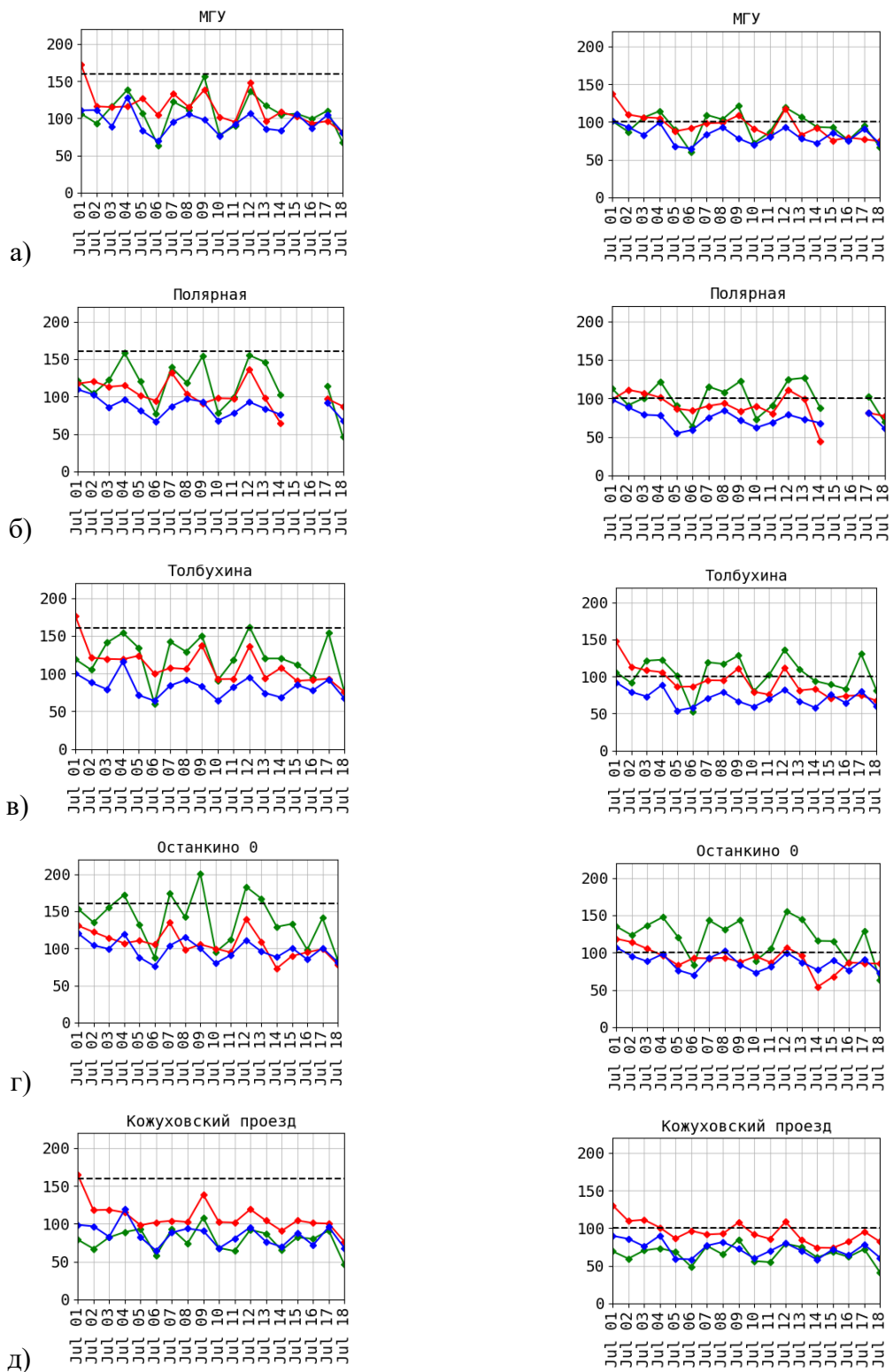


Рисунок 4.5 – Суточные максимумы часовых (слева) и средних за 8 ч (справа) концентраций O<sub>3</sub> (мкг/м<sup>3</sup>) по измерениям на АСКЗА (зеленые линии), рассчитанные ХТМ SHIMERE (красные линии) и ММО-O<sub>3</sub> (синие линии) 1–18 июля 2024 г. (независимая выборка)

Примечание: Пунктирная линия – ПДК<sub>8ч</sub> O<sub>3</sub> 100 мкг/м<sup>3</sup>.



Не исключено, что внешний фактор – повышенное ОСО летом 2024 г. – мог стать одной из причин менее успешного прогноза  $O_3$  ХТМ и ММО- $O_3$  в июле по сравнению с другими месяцами (рисунок 4.1) и прошлыми годами (раздел 2.4.1), когда ХТМ в основном превышала дневные уровни приземного озона.

На этом эпизоде подтверждается вывод о том, что модели машинного обучения, являясь по сути статистическими моделями, ограничены в возможности воспроизводить редко повторяющиеся в обучающей выборке (практически единичные) ситуации.

Озоновый эпизод в начале сентября также был связан с аномальными погодными условиями (сентябрь оказался теплее обычного на рекордных 4 °С). С 25 августа до 3 сентября в малоградиентном поле или вблизи центра антициклона в затишье почти каждую ночь наблюдались приземные температурные инверсии 2–7 °С до высоты 650 м, днем температура повышалась до 26–28 °С (26,28 августа и 1 сентября до 30 °С) при относительной влажности 42–45 %, осадков не наблюдалось вплоть до 4 сентября. То есть имел место период преобладающих НМУ. 28 августа днем проходил атмосферный фронт, температура немного понизилась, но оставалась не по сезону высокой; на рисунке 4.6 видно резкое снижение концентраций  $O_3$  28–29 августа практически на всех АСКЗА. За девять дней аномально теплой погоды (25 августа – 3 сентября)  $O_{3-8ч}$  35 раз превысил ПДК<sub>8ч</sub> (рисунок 4.6).

ХТМ CHIMERE в период 25 августа – 3 сентября прогнозировала  $O_{3-8ч}$  с постоянным занижением относительно измерений на всех АСКЗА, за исключением фоновой станции Рогово, на которой можно признать самым успешным прогноз ХТМ по близости к измерениям (рисунок 4.6).

В данном эпизоде преобладающих НМУ ММО- $O_3$  показала выраженную эффективность коррекции прогнозов ХТМ, практически на всех станциях увеличивая прогностические концентрации  $O_3$ , тем самым приближая их к измерениям (рисунок 4.6). В дни с наибольшими концентрациями  $O_{8ч}$  ММО- $O_3$  почти в 2 раза уменьшает ошибку прогноза ХТМ на городских АСКЗА (Останкино, Полярная, Туристская на рисунке 4.6).

Тем не менее, в данном эпизоде нельзя игнорировать внешний фактор – возможную адвекцию озона с юга в очень теплой и сухой воздушной массе, чем можно частично объяснить повышенные концентрации  $O_3$  в Москве.

По итогам проведенного анализа следует отметить, что в эпизодах повышенного озонового загрязнения, в которых значительный вклад вносит фотохимическая генерация озона при НМУ, ММО- $O_3$  в основном корректирует (повышает) расчеты ХТМ, приближая прогностические концентрации к измеренным.

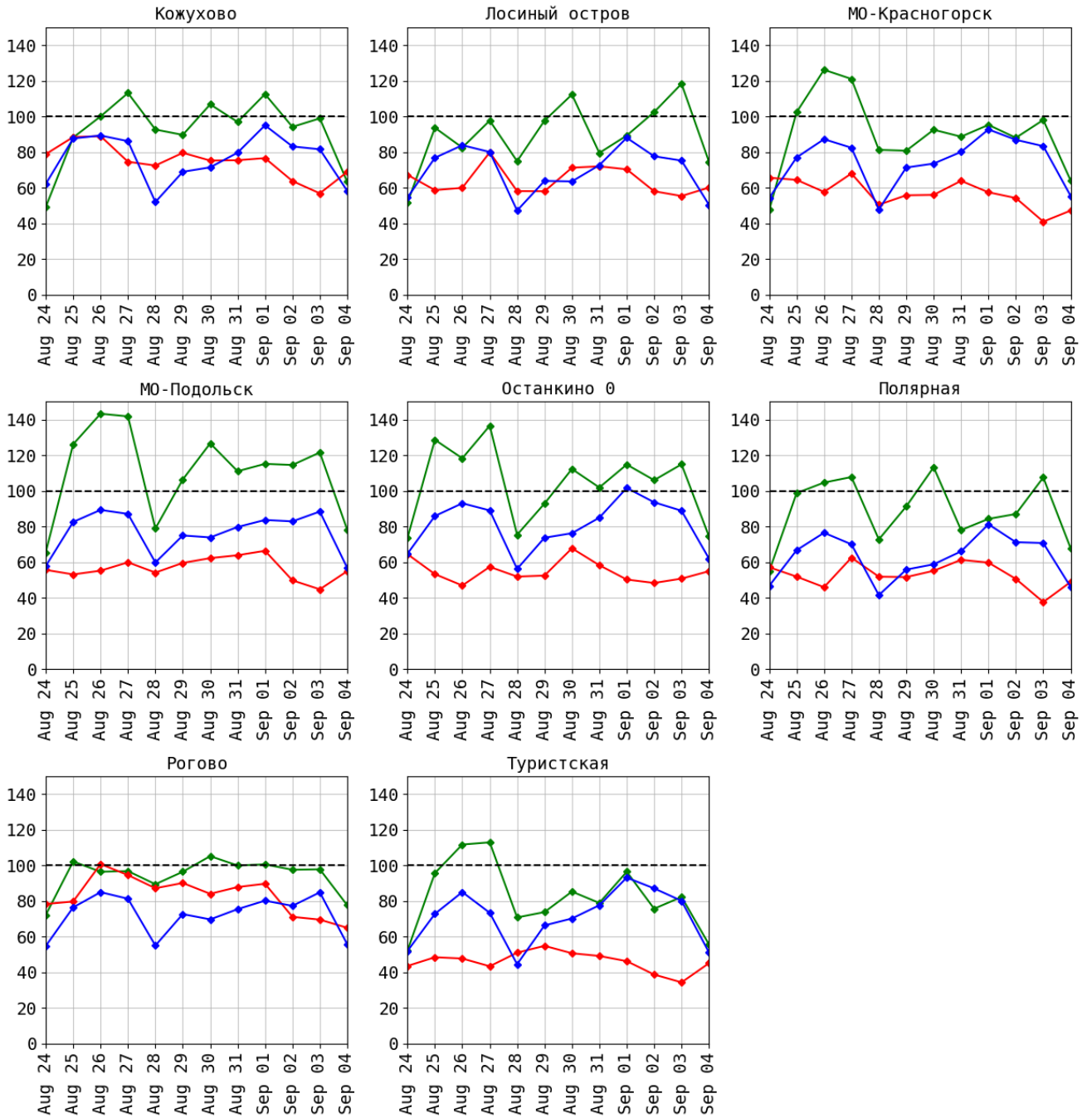


Рисунок 4.6 – O<sub>3-8ч</sub> (µг/м<sup>3</sup>) по измерениям на АСКЗА (зеленые линии), рассчитанные ХТМ SHIMERE (красные линии) и ММО-O<sub>3</sub> (синие линии) в эпизоде повышенного озонового загрязнения 25 августа – 3 сентября 2024 г. (независимая выборка)

*Примечание:* Пунктирная линия – ПДК<sub>8ч</sub> O<sub>3</sub> 100 µг/м<sup>3</sup>.

### 4.3 Эффекты ММО-постпроцессинга полей прогнозов концентраций приземного озона

Финальным этапом разработанной Методики является ММО-постпроцессинг поля модельных прогнозов концентраций в расчетной области (см. раздел 3.1). Доказанная эффективность ММО-постпроцессинга прогнозов на территориях без измерений концентраций (при исключении станций из обучающей выборки, см. раздел 4.1, рисунок 4.3) служит обоснованием такого подхода.

Проведены эксперименты с усвоением ММО-О<sub>3</sub> прогностических предикторов и данных землепользования (см. таблицу 3.1) из всех ячеек модельной сетки CHIMERE, получены расчеты ММО-О<sub>3</sub> концентраций приземного озона на первые сутки для всей территории московского региона с шагом 2×2 км и с временным шагом 1 час.

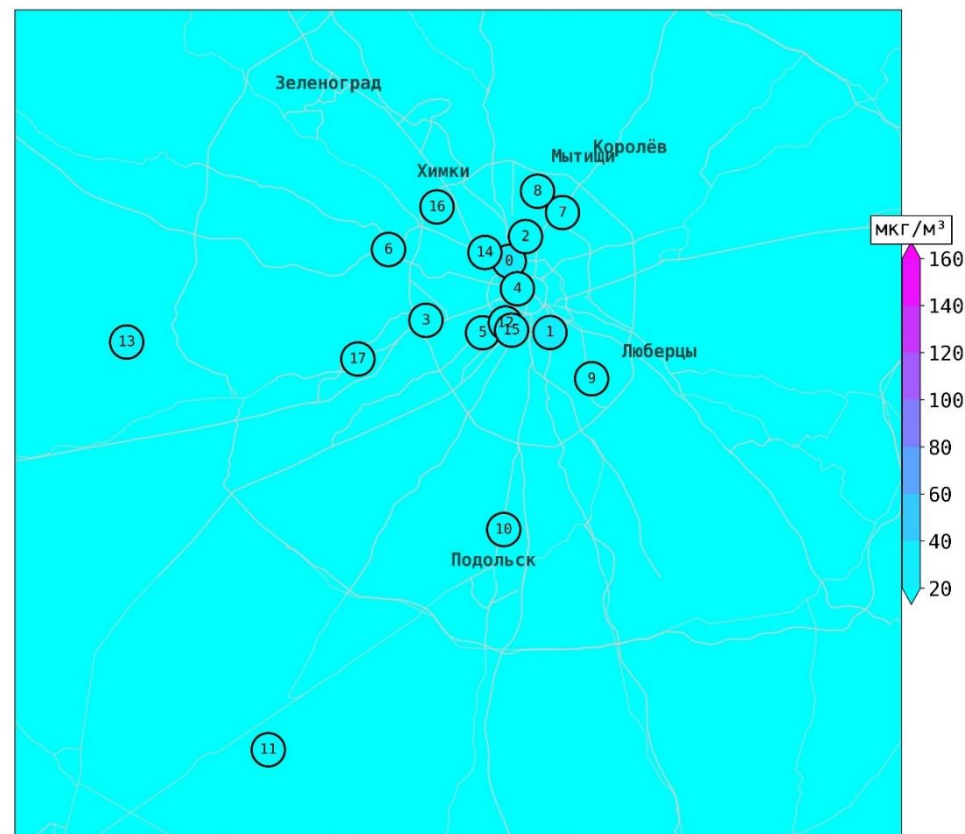
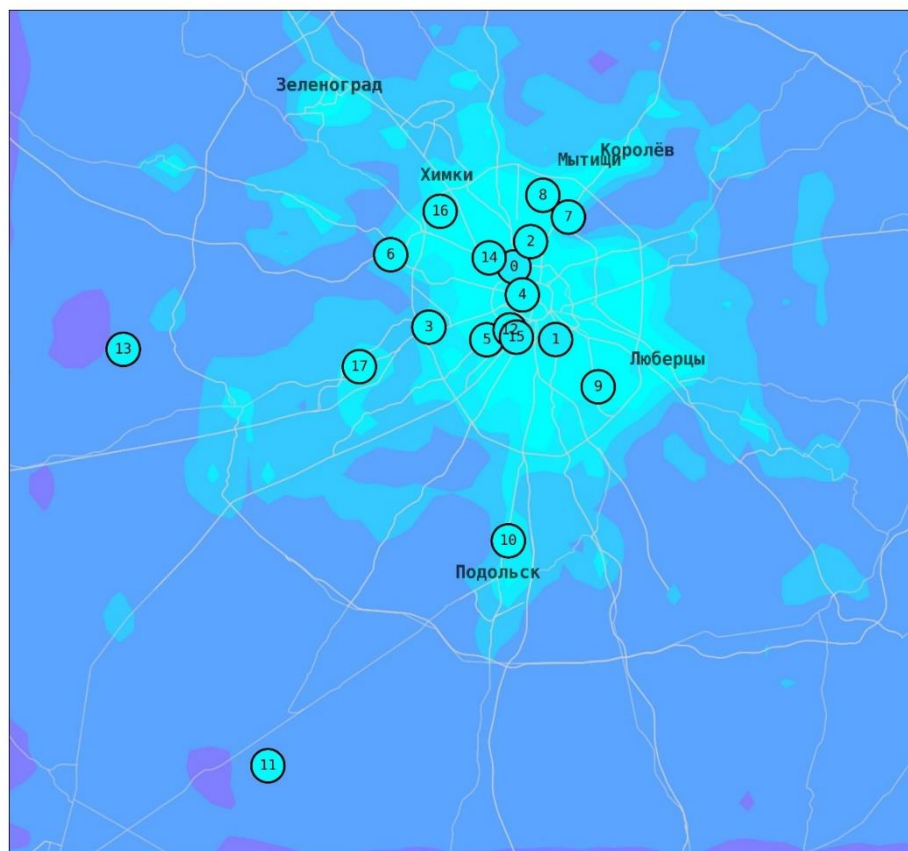
На рисунках 4.7 и 4.8 представлены рассчитанные ХТМ CHIMERE (слева) и ММО-О<sub>3</sub> (справа) часовые поля концентраций приземного озона 29 мая в эпизоде повышенных концентраций О<sub>3</sub>. На рисунки нанесены часовые измерения концентраций О<sub>3</sub> на АСКЗА в соответствующее прогнозам время.

На рисунке 4.7 представлены поля О<sub>3</sub> в 4:00 МСК – время суточного минимума О<sub>3</sub> в отсутствии солнечного излучения и большей части источников прекурсоров озона. Видно, что измеренные на всех АСКЗА концентрации О<sub>3</sub> менее 20 мкг/м<sup>3</sup>, ХТМ CHIMERE (слева) вполне успешно рассчитывает концентрации на городских территориях, они близки к измерениям, но на удалении от Москвы (фоновые станции Рогово и Звенигород) ХТМ завышает концентрации. ММО-О<sub>3</sub> в полном соответствии с измерениями прогнозирует ночной минимальный уровень озона для всей территории московского региона, в том числе на фоновых территориях.

На рисунке 4.8 представлены прогностические поля на 15:00 МСК – время, близкое к суточному максимуму О<sub>3</sub>. Как видно, ХТМ CHIMERE успешно воспроизводит повышенные концентрации О<sub>3</sub> в окрестностях Москвы и в юго-западном секторе региона, шлейф над которым обусловлен юго-западными ветрами и выносом воздушной массы с фоновой территории. Но на северо-западе Москвы видны нехарактерные низкие концентрации О<sub>3</sub>, что могло стать следствием влияния слабого юго-восточного переноса городского загрязнения в нижних слоях АПС.

XTM CHIMERE

ММО-О<sub>3</sub>

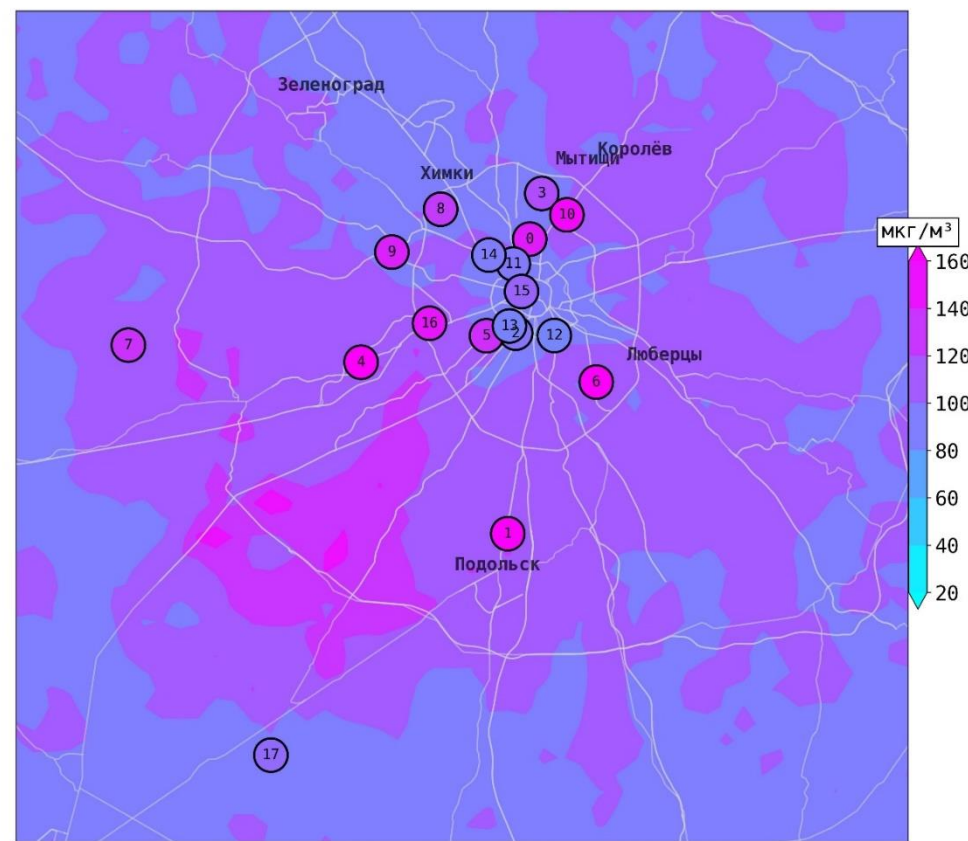
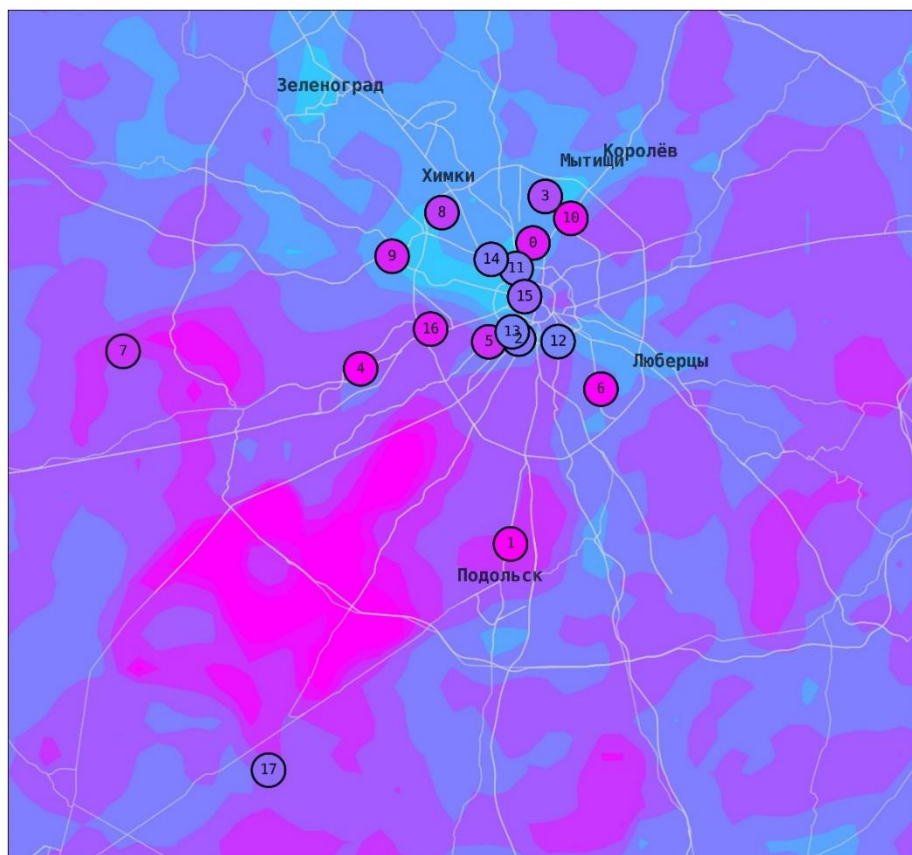


0 - Нижняя Масловка	4 - Спиридоновка	8 - Полярная	12 - Хамовники	15 - Площадь Гагарина
1 - Кожуховский проезд	5 - МГУ	9 - Марьино	13 - Звенигород	16 - Туристская
2 - Останкино 0	6 - МО-Красногорск	10 - МО-Подольск	14 - МАДИ	17 - МО-Одинцово
3 - Толбухина	7 - Лосинный остров	11 - Рогово		

Рисунок 4.7 – Прогнозы на первые сутки приземных концентраций О<sub>3</sub> на территории московского региона 29 мая 2024 г. 04:00 МСК XTM CHIMERE (слева), ММО-О<sub>3</sub> (справа). Кружками нанесены соответствующие сроку прогнозов часовые измерения концентраций О<sub>3</sub> на АСКЗА в той же цветовой гамме, что и прогностические поля

XTM CHIMERE

ММО-О<sub>3</sub>



109

0 - Останкино 0	4 - МО-Одинцово	8 - Туристская	12 - Кожуховский проезд	15 - Спиридоновка
1 - МО-Подольск	5 - МГУ	9 - МО-Красногорск	13 - Хамовники	16 - Толбухина
2 - Площадь Гагарина	6 - Марьино	10 - Лосиный остров	14 - МАДИ	17 - Рогово
3 - Полярная	7 - Звенигород	11 - Нижняя Масловка		

Рисунок 4.8 – Прогнозы на первые сутки приземных концентраций О<sub>3</sub> на территории московского региона 29 мая 2024 г. 15:00 МСК XTM CHIMERE (слева), ММО-О<sub>3</sub> (справа). Кружками нанесены соответствующие сроку прогнозов часовые измерения концентраций О<sub>3</sub> на АСКЗА в той же цветовой гамме, что и прогностические поля

Поле прогнозов ММО-О<sub>3</sub> (рисунок 4.8, справа) в общих очертаниях совпадает с прогностическим полем ХТМ CHIMERE, отражающим максимум озона на юго-западе региона и минимум на северо-западе Москвы. При этом ММО-О<sub>3</sub> увеличила рассчитанные CHIMERE концентрации, приблизив их к измерениям: прогнозы ММО-О<sub>3</sub> практически совпали с измерениями в центре города (АСКЗА Нижняя Масловка, Хамовники, Кожуховский проезд) и на окраинах Москвы (АСКЗА Полярная, Туристская, Марьино, Толбухина, пригородная АСКЗА Красногорск).

При этом в рассчитанных ММО-О<sub>3</sub> полях отображается более низкий уровень приземного озона в центре города – на территории с наибольшей транспортной нагрузкой и плотностью городской застройки. Как и ХТМ CHIMERE, ММО-О<sub>3</sub> воспроизводит повышенные концентрации О<sub>3</sub> в лесопарковой зоне Лосиный остров (северо-восток Москвы). На фоновых станциях (Звенигород, Подольск, Рогово) ММО-О<sub>3</sub> рассчитывает повышенный уровень О<sub>3</sub>, на 20–40 мкг/м<sup>3</sup> недооценивая его по сравнению с измерениями.

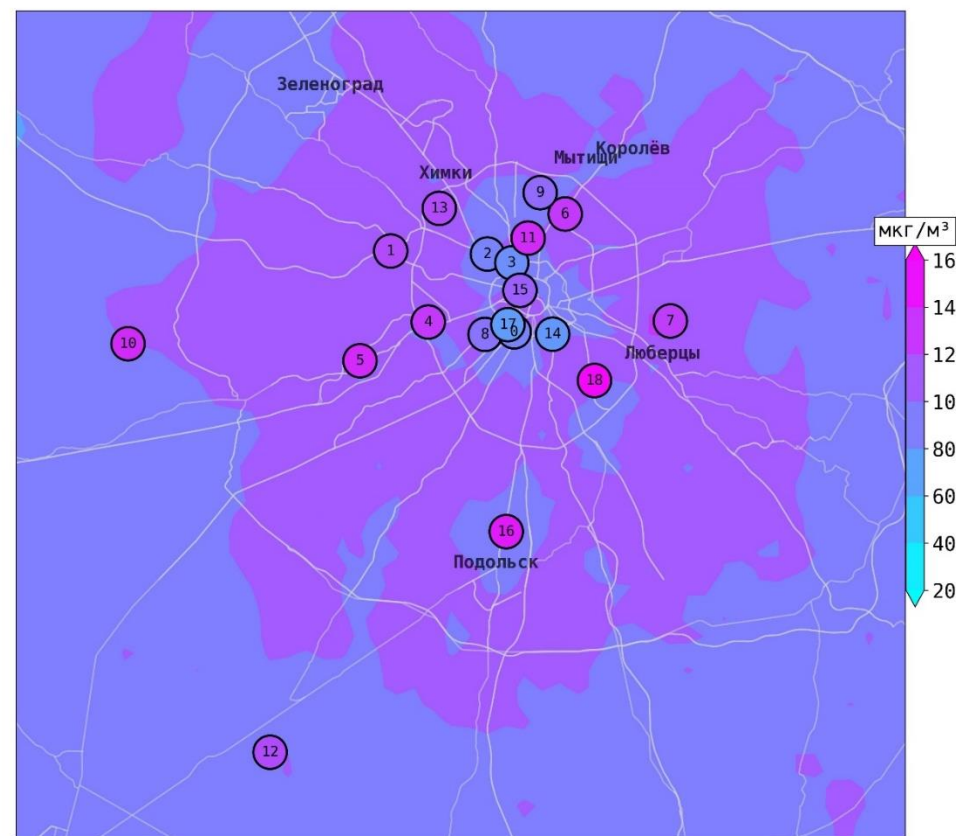
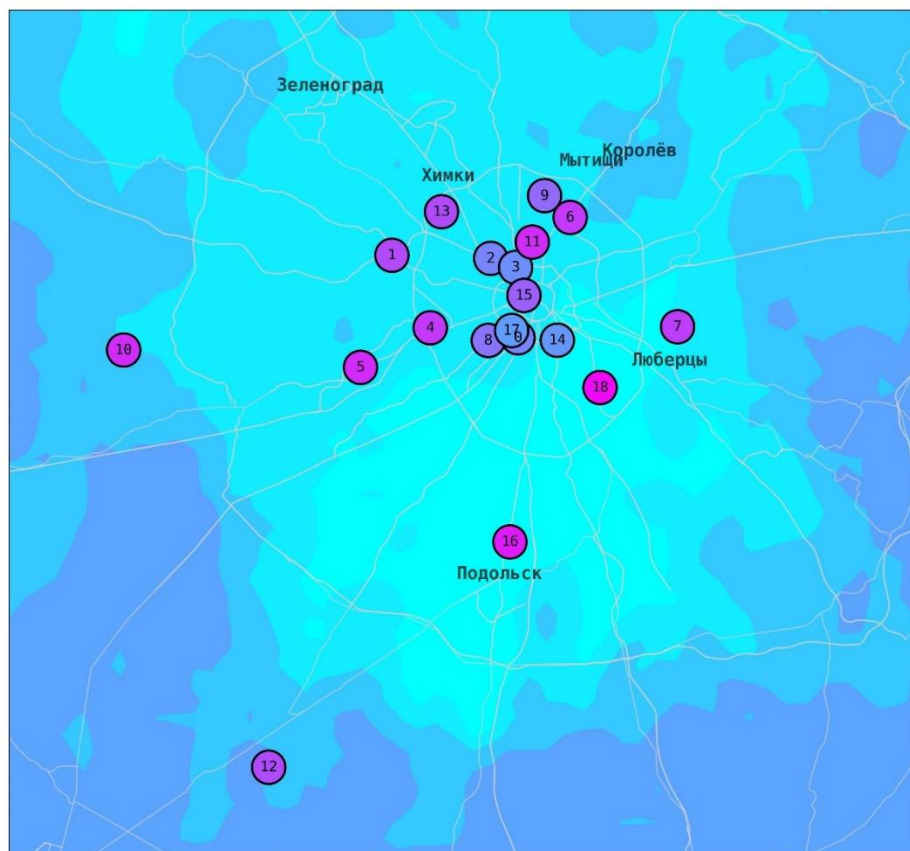
Еще один пример эффективности ММО-постпроцессинга прогностических полей концентраций озона приведен на рисунке 4.9, здесь изображены поля прогнозов О<sub>3</sub> ХТМ CHIMERE и ММО-О<sub>3</sub> на 15:00 1 сентября 2024 г. – в эпизоде повышенного содержания озона в приземном воздухе.

1 сентября в Москве стояла сухая и жаркая погода (днем до +30 °С). ХТМ CHIMERE воспроизвела явно нехарактерный для такого типа погоды низкий уровень О<sub>3</sub> в регионе (рисунок 4.9, слева): рассчитанные концентрации на городской территории до 40 мкг/м<sup>3</sup>, на фоновой территории – не больше 80 мкг/м<sup>3</sup>. В расчетах ХТМ нет также и отражения влияния юго-восточного переноса, обеспечивавшего поступление в Москву очень теплой насыщенной озоном воздушной массы из южных регионов ЕТР, где долгое время стояла 30–35-градусная жара при еще высокой УФ-радиации.

ММО-О<sub>3</sub> успешно скорректировала поле прогнозов ХТМ (рисунок 4.9, справа), повысив прогностические концентрации О<sub>3</sub> по всей территории и значительно приблизив их к измерениям на АСКЗА. При этом еще раз отметим, что ММО-О<sub>3</sub> учитывает распределение городской застройки по территории: в центре и в северной части города прогнозы ММО-О<sub>3</sub> в диапазоне 80–100 мкг/м<sup>3</sup>, на окраинах города – 100–120 мкг/м<sup>3</sup>, что почти полностью соответствует измерениям. На удаленных от границ мегаполиса территориях ММО-О<sub>3</sub> также повышает уровень концентраций по сравнению с расчетами ХТМ, хотя прогнозы ММО-О<sub>3</sub> на 20 мкг/м<sup>3</sup> (что в пределах инструментальной погрешности) меньше измеренных концентраций.

XTM CHIMERE

ММО-О<sub>3</sub>



0 - Площадь Гагарина	4 - Толбухина	8 - МГУ	12 - Рогово	16 - МО-Подольск
1 - МО-Красногорск	5 - МО-Одинцово	9 - Полярная	13 - Туристская	17 - Хамовники
2 - МАДИ	6 - Лосинный остров	10 - Звенигород	14 - Кожуховский проезд	18 - Марьино
3 - Нижняя Масловка	7 - Кожухово	11 - Останкино 0	15 - Спиридоновка	

Рисунок 4.9 – Прогнозы на первые сутки приземных концентраций О<sub>3</sub> на территории московского региона 1 сентября 2024 г. 15:00 МСК XTM CHIMERE (слева), ММО-О<sub>3</sub> (справа). Кружками нанесены соответствующие сроку прогнозов часовые измерения концентраций О<sub>3</sub> на АСКЗА в той же цветовой гамме, что и прогностические поля

#### Выводы к главе 4

Результаты верификации на независимой выборке март-сентябрь 2024 г. разработанной ММО для постпроцессинга численных прогнозов концентраций приземного озона на территории московского региона демонстрируют эффективность применения ММО-О<sub>3</sub> в виде снижения отклонений прогнозов от измерений: СКО<sub>П-И</sub>(О<sub>3-8ч</sub>) уменьшилось по всей тестовой выборке с 27 мкг/м<sup>3</sup> (ХТМ СHIMERE) до 19 мкг/м<sup>3</sup> (ММО-О<sub>3</sub>) – на 30 %, в отдельные месяцы с 23–30 мкг/м<sup>3</sup> до 16–22 мкг/м<sup>3</sup> – на 12–43 %. ММО-О<sub>3</sub> существенно корректирует суточный ход концентраций О<sub>3</sub> в соответствии с измерениями, коэффициенты корреляции между часовыми прогнозами и измерениями увеличены с 0.4–0.5 (ХТМ) до 0.7–0.9 (ММО-О<sub>3</sub>).

Анализ эффектов применения ММО-О<sub>3</sub> для отдельных АСКЗА проведен с использованием метода исключения АСКЗА из обучающей выборки ММО-О<sub>3</sub>, чтобы оценить эффекты ММО-постпроцессинга прогнозов О<sub>3</sub> на территориях без измерений концентраций. Анализ показал снижение СКО<sub>П-И</sub>(О<sub>3-8ч</sub>) на ~30 % на городских АСКЗА. Минимальные эффекты ММО-постпроцессинга прогнозов О<sub>3</sub> выявлены на некоторых фоновых АСКЗА, эффекты отсутствуют только на 2 фоновых АСКЗА на западе московской области.

Показана эффективность ММО-постпроцессинга прогнозов О<sub>3</sub> в эпизодах повышенного озонового загрязнения в московском регионе при НМУ – ММО-О<sub>3</sub> в ~2 раза снижает отклонение преимущественно заниженных прогнозов ХТМ на городских АСКЗА. По мнению автора, отсутствие эффектов ММО-постпроцессинга в июльском эпизоде в первую очередь объясняется аномальным уровнем О<sub>3</sub>, на 40 % превысившим среднюю многолетнюю концентрацию О<sub>3</sub> в июле [Иванова и др., 2025].

Приведены примеры ММО-постпроцессинга часовых полей прогнозов приземного озона на территории московского региона, смоделированных ХТМ в периоды НМУ. ММО-О<sub>3</sub> успешно корректирует поля ХТМ по уровням озонового загрязнения, при этом учитывает время суток, смоделированный ХТМ перенос и уровни антропогенной нагрузки на территориях.



## ГЛАВА 5. ТЕСТИРОВАНИЕ РАЗРАБОТАННОЙ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОСТПРОЦЕССИНГА ЧИСЛЕННЫХ ПРОГНОЗОВ КОНЦЕНТРАЦИЙ $PM_{10}$

Независимая выборка для тестирования разработанной ММО- $PM_{10}$  охватывает период 1 марта – 30 сентября 2024 г. и включает ежедневные часовые прогнозы концентраций  $PM_{10}$  на первые сутки, рассчитанные ХТМ CHIMERE и ММО- $PM_{10}$  для территории московского региона, и ряды часовых измерений  $PM_{10}$  на АСКЗА. Для составления ряда оценок производился пересчет часовых концентраций в среднесуточные (норматив ВОЗ  $PM_{10-CC} = 45 \text{ мкг/м}^3$ ).

### 5.1 Эффекты ММО-постпроцессинга численных прогнозов концентраций $PM_{10}$

При интерпретации результатов ММО-постпроцессинга необходимо руководствоваться представлениями об основных механизмах загрязнения приземного воздуха твердыми взвешенными частицами  $PM_{10}$  в московском регионе, к которым относятся поступление частиц от наземных локальных источников и адвекция аэрозоля из удаленных регионов. Отметим, что из общей статистики отклонений модельных прогнозов, приводимой в данном разделе, удалены обусловленные переносом  $PM_{10}$  из удаленных районов два периода повышенных концентраций в марте-апреле 2024 г., поскольку ММО- $PM_{10}$ , обученная на локальных прогностических предикторах, априори неспособна предсказывать подобные события. Эпизоды повышенных концентраций  $PM_{10}$  в тестовом периоде, в том числе обусловленные дальним переносом взвешенных частиц, и оценки модельных прогнозов в эпизодах рассматриваются в разделе 5.2.

На рисунке 5.1 показаны боксовые распределения отклонений прогнозов ХТМ CHIMERE и ММО- $PM_{10}$  среднесуточной концентрации  $PM_{10}$  в каждом из 7 месяцев независимой проверки; как видно, в каждом месяце наблюдается сужение разброса  $\Delta_{П-И}$  ММО- $PM_{10}$  по сравнению с  $\Delta_{П-И}$  ХТМ.

Заметим, что измеренные  $PM_{10}$  в первые месяцы независимой проверки совпадают со средним многолетним годовым максимумом  $PM_{10}$  по наблюдениям. Важно отметить, что в заданном в ХТМ CHIMERE сезонном распределении эмиссий  $PM_{10}$  годовой максимум приходится на начало весны в соответствии с европейскими стандартами, при том, что в московском регионе в это время чаще всего лежит снег или только начинается сход снежного покрова. И это в значительной степени повлияло на то, что ХТМ CHIMERE в марте больше, чем в другие месяцы, завышает прогнозы  $PM_{10-CC}$  – в среднем на  $22 \text{ мкг/м}^3$ , IQR  $\Delta_{П-И}$  (ХТМ)  $5\text{--}38 \text{ мкг/м}^3$ . ММО- $PM_{10}$  в марте успешно корректирует расчеты ХТМ: при средней ошибке

близкой к 0 межквартильный размах  $\Delta_{П-и}$  (ММО-PM<sub>10</sub>) составил -7-8 мкг/м<sup>3</sup>, а СКО<sub>П-и</sub> в марте уменьшается с применением ММО-PM<sub>10</sub> почти в 2 раза (рисунок 5.1).

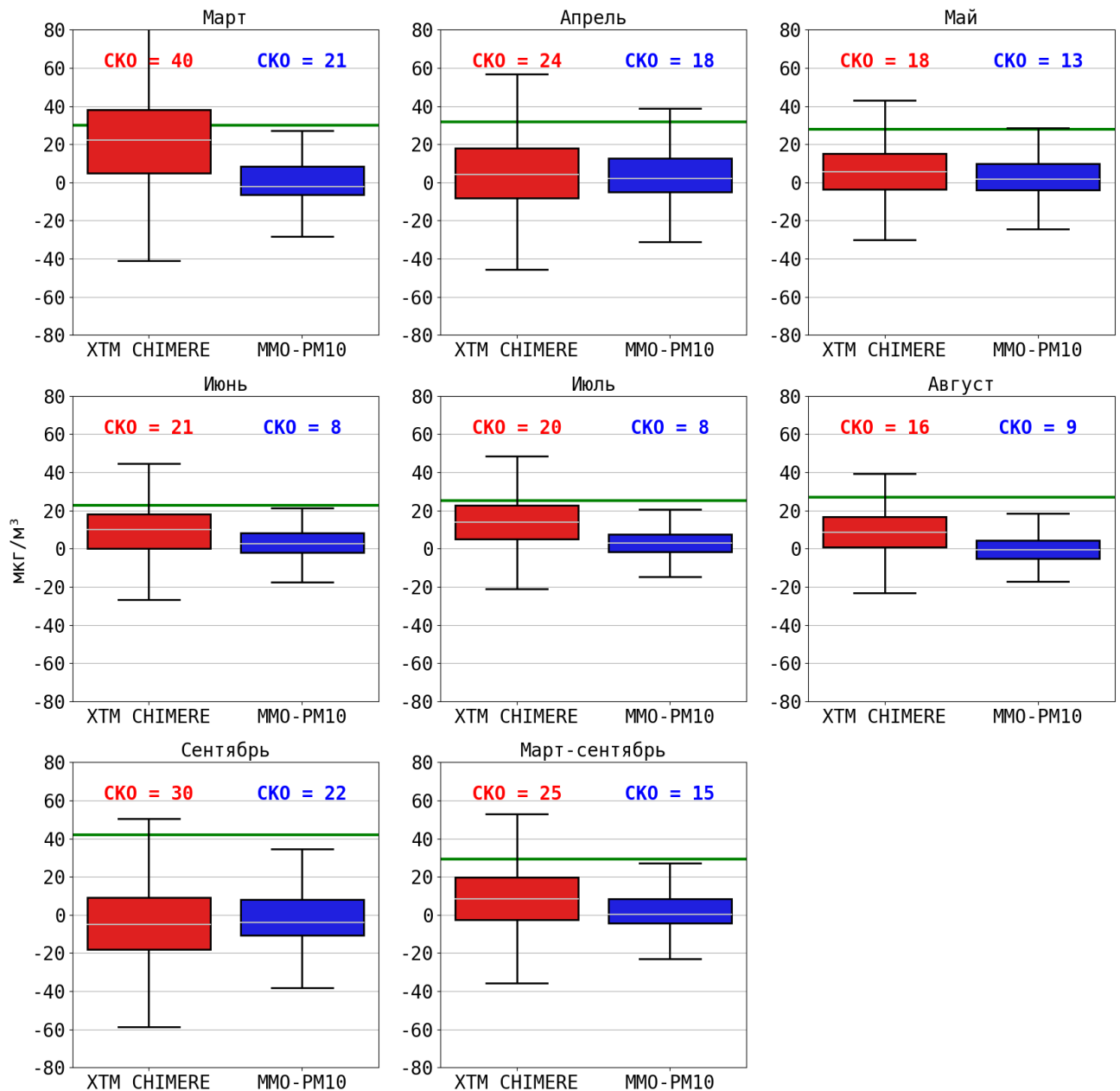


Рисунок 5.1 – Отклонения прогнозов PM<sub>10-СС</sub> XTM CHIMERE (красные боксы) и ММО-PM<sub>10</sub> (синие боксы) от измерений (все АСКЗА)

Независимая выборка, март-сентябрь 2024 г.

*Примечание:* В боксе – IQR  $\Delta_{П-и}$ , линия внутри бокса – среднее  $\Delta_{П-и}$ , нижний «ус»:  $Q1 - 1.5 * IQR \Delta_{П-и}$ , верхний «ус»:  $Q3 + 1.5 * IQR \Delta_{П-и}$ . Зеленые сплошные линии – средняя за месяц измеренная концентрация PM<sub>10-СС</sub> (все АСКЗА). Над боксами СКО<sub>П-и</sub> XTM и ММО-PM<sub>10</sub>.

В апреле (годовой максимум  $PM_{10}$ ) и мае обе модели удовлетворительно прогнозируют  $PM_{10-CC}$ : среднее отклонение прогнозов ХТМ CHIMERE  $+4-6$   $мкг/м^3$ , ММО- $PM_{10}$   $+2$   $мкг/м^3$ ; при этом с применением ММО- $PM_{10}$  заметно сужается IQR  $\Delta_{П-и}$ ,  $СКО_{П-и}$  уменьшается на 25–30 % (рисунок 5.1).

В летние месяцы измеренный уровень  $PM_{10-CC}$  понижается до 23–27  $мкг/м^3$ , в основном благодаря преобладающему конвективному перемешиванию в продолжительный световой день. Летом видны наибольшие эффекты ММО-постпроцессинга: при средних завышениях прогнозов ХТМ CHIMERE  $PM_{10-CC}$  на 9–14  $мкг/м^3$ , средние отклонения ММО- $PM_{10}$  близки к 0; с применением ММО- $PM_{10}$  значительно сужены IQR  $\Delta_{П-и}$ ;  $СКО_{П-и}$  уменьшены в июне и июле на 60 %, в августе на 44 % – до 8–9  $мкг/м^3$  (при  $СКО_{П-и}$  ХТМ 16–21  $мкг/м^3$ ).

В аномально теплом и сухом сентябре 2024 г. в Москве наблюдались максимальные за весь период независимой выборки концентрации  $PM_{10}$  – средняя за весь месяц с учетом данных всех АСКЗА концентрация  $PM_{10-CC}$  оказалась немного больше 40  $мкг/м^3$ . В условиях часто наблюдавшихся кратковременных НМУ в сентябре зафиксировано 12 дней с превышениями на более чем 40 % АСКЗА норматива ВОЗ 45  $мкг/м^3$  ( $PM_{10-CC}$ ). При данных обстоятельствах естественно ожидать модельное занижение уровня  $PM_{10}$ , но оказалось, что и ХТМ CHIMERE, и ММО- $PM_{10}$  в среднем занизили  $PM_{10-CC}$  всего на 4–5  $мкг/м^3$ . IQR  $\Delta_{П-и}$  (ХТМ) составил  $-18-9$   $мкг/м^3$ , IQR  $\Delta_{П-и}$  (ММО- $PM_{10}$ ) немного уже:  $-11-8$   $мкг/м^3$ ;  $СКО_{П-и}$  снижено с применением ММО- $PM_{10}$  с 30 до 22  $мкг/м^3$ .

По всей тестовой выборке март-сентябрь  $СКО_{П-и}$  ( $PM_{10-CC}$ ) снизилось с применением ММО- $PM_{10}$  с 25  $мкг/м^3$  (ХТМ CHIMERE) до 15  $мкг/м^3$  – на 40 %.

В проведенных ранее исследованиях [Борисов и Кузнецова, 2024] и в подразделе 2.4.1 были установлены специфические особенности расчетов ХТМ CHIMERE: преимущественное завышение уровня  $PM_{10}$  и резко отличающийся от наблюдаемого суточный ход концентрации  $PM_{10}$  в московском регионе. Проверка на независимых данных в целом подтвердила предварительные выводы: как видно на рисунке 5.2, во все месяцы ХТМ воспроизводит ярко выраженный дневной максимум концентрации  $PM_{10}$  в отсутствие такового по измерениям и преимущественно занижает главный вечерний максимум  $PM_{10}$ .

Иную, приближенную к реальным наблюдениям, динамику суточного хода  $PM_{10}$  воспроизводит разработанная ММО- $PM_{10}$ .

С апреля по август ММО- $PM_{10}$  успешно воспроизводит суточный ход  $PM_{10}$  в условиях интенсивного конвективного перемешивания днем и максимумом концентраций вечером-ночью.

Обратим внимание, что в сентябре ММО- $PM_{10}$  рассчитывает, в соответствии с измерениями, бимодальный суточный ход  $PM_{10}$ , хотя на 10–15  $мкг/м^3$  недооценивает вечернее повышение концентраций.

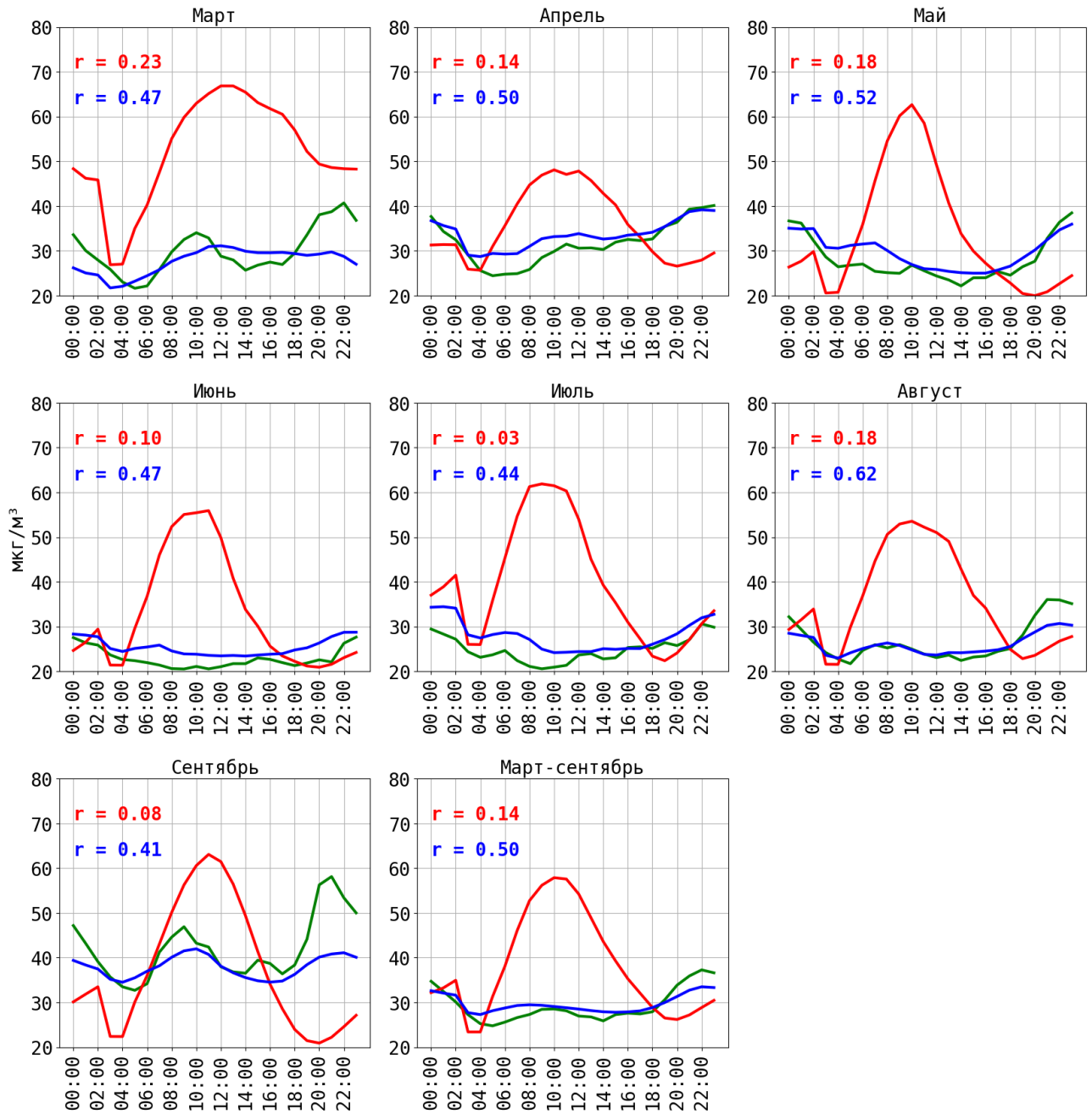


Рисунок 5.2 – Средние почасовые концентрации  $PM_{10}$  по данным измерений на АСКЗА (зеленые линии), по расчетам ХТМ CHIMERE (красные линии) и ММО- $PM_{10}$  (синие линии). Независимая выборка, март-сентябрь 2024 г.

*Примечание:* В левом верхнем углу – коэффициенты корреляции между часовыми прогнозами и измерениями  $PM_{10}$  (все АСКЗА) по расчетам ХТМ CHIMERE (красным) и ММО- $PM_{10}$  (синим).

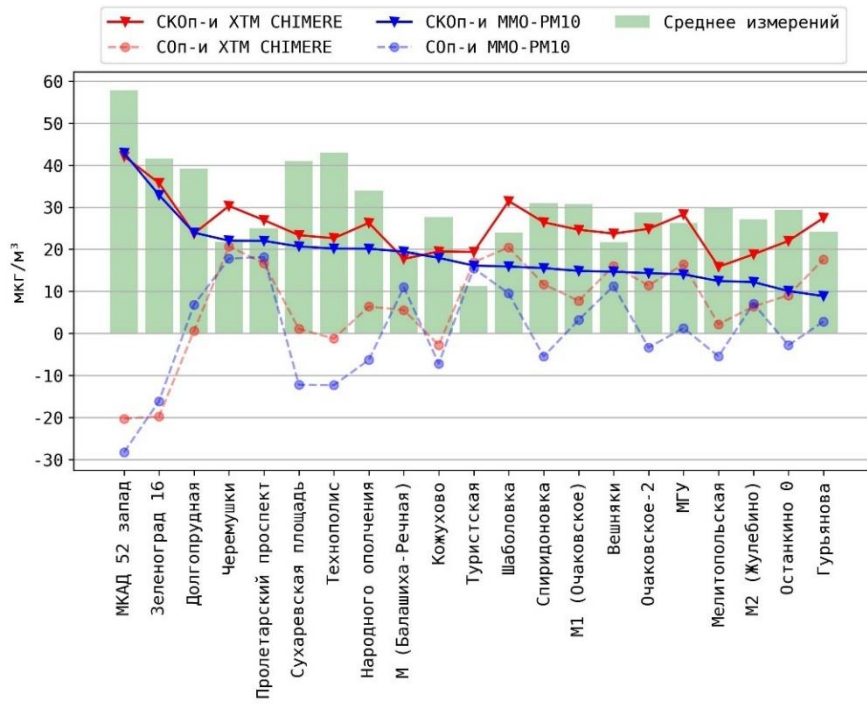
В марте ММО- $PM_{10}$  занижает вечерний максимум концентраций на ~10 мкг, а ХТМ, наоборот, в той же степени завышает вечерний максимум, поскольку, повторимся, в ХТМ заложен европейский максимум эмиссий  $PM_{10}$  в начале весны.

Нельзя не отметить существенное возрастание коэффициентов корреляции между часовыми прогнозами и измерениями  $PM_{10}$  с применением ММО- $PM_{10}$ , в большинстве месяцев они увеличились в 2–4 раза до 0.4–0.5, в августе до 0.62 (рисунок 5.2). По всей тестовой выборке март-сентябрь коэффициент корреляции увеличился с 0.14 до 0.5.

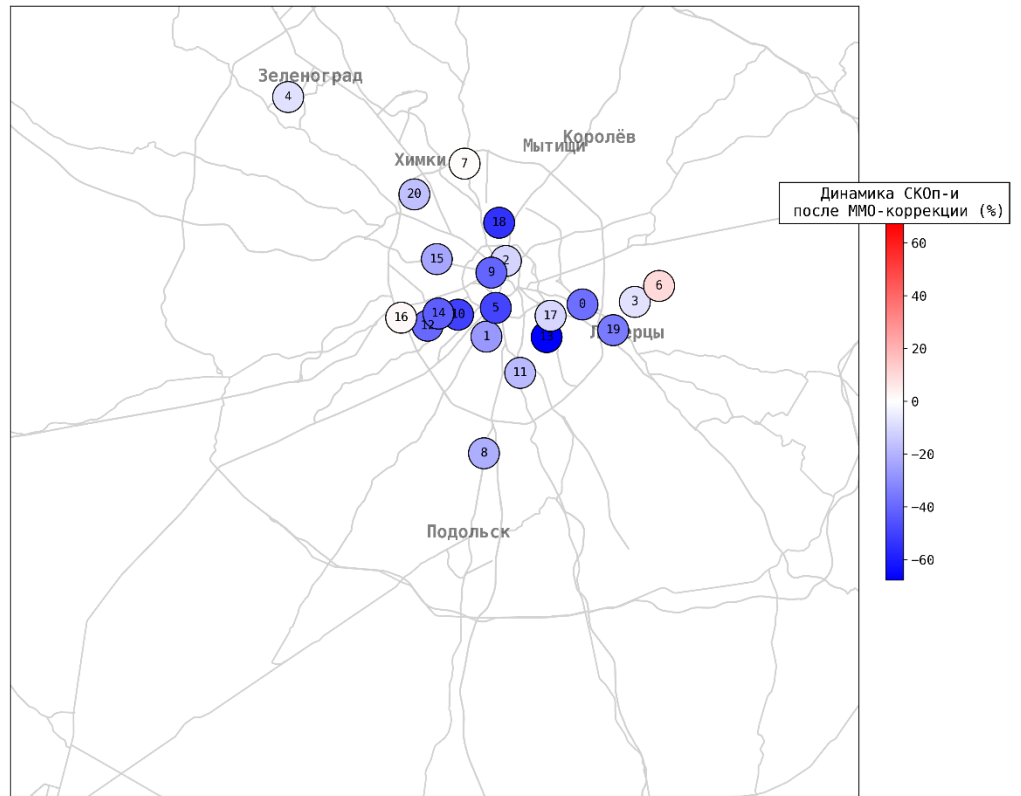
Учитывая пространственную неоднородность загрязнения  $PM_{10}$  в московском регионе, проведен анализ эффективности ММО-постпроцессинга прогнозов концентраций на отдельных АСКЗА различной типовой принадлежности. Эффекты ММО-постпроцессинга для отдельных станций оценивались с учетом исключения станции из обучающей выборки ММО- $PM_{10}$ , как и для ММО- $O_3$  (см. раздел 4.1). Полученные по всему тестовому периоду март-сентябрь 2024 г. средние квадратические ( $СКО_{п-и}$ ) и средние ( $СО_{п-и}$ ) отклонения прогнозов  $PM_{10-CC}$  по расчетам ХТМ и ММО- $PM_{10}$  для отдельных АСКЗА представлены на рисунке 5.3 (а); на рисунке 5.3 (б) визуализированы эффекты ММО-постпроцессинга прогнозов ХТМ на отдельных АСКЗА в виде увеличения/уменьшения  $СКО_{п-и}$  по всей тестовой выборке (в %).

Представленные расчеты иллюстрируют выраженные эффекты ММО-постпроцессинга прогнозов ХТМ практически на всех АСКЗА, за исключением пяти – МКАД 52 запад, Зеленоград 16, Долгопрудная, Балашиха-Речная, Кожухово. Следует отметить, что АСКЗА транспортного типа «МКАД 52 запад», находящаяся в непосредственной близости от крупной автомагистрали, отражает высокую техногенную нагрузку с нерегулярными выбросами, соответственно, и наибольший уровень  $PM_{10-CC}$  (почти  $60 \text{ мкг/м}^3$ ); ММО- $PM_{10}$ , как и ХТМ CHIMERE, ожидаемо, в среднем занижают прогнозы для этой станции (рисунок 5.3 а). АСКЗА «Зеленоград 16» с высоким измеренным уровнем  $PM_{10-CC}$  (на удалении от Москвы) в тестовый период находилась под влиянием локальных местных выбросов. Вблизи АСКЗА Долгопрудная в период тестирования проводились строительные работы, хотя средний уровень  $PM_{10}$  для данной станции и ХТМ, и ММО- $PM_{10}$  отразили успешно –  $СО_{п-и} +1-7 \text{ мкг/м}^3$ . На расположенных за МКАД на востоке от Москвы АСКЗА Балашиха-Речная и Кожухово уровень  $PM_{10}$  низкий, значимых эффектов ММО-постпроцессинга здесь не наблюдается (рисунок 5.3 а, б).

Для большинства городских АСКЗА  $СКО_{п-и}$  ММО- $PM_{10}$  находится в диапазоне 9–23  $\text{мкг/м}^3$  при  $СКО_{п-и}$  ХТМ CHIMERE 20–33  $\text{мкг/м}^3$ , средние отклонения прогнозов ММО- $PM_{10}$  преимущественно в диапазоне  $\pm 15 \text{ мкг/м}^3$ . Максимальные эффекты ММО-постпроцессинга прогнозов ХТМ со снижением  $СКО_{п-и}$  в 1,5–2 раза видны на большинстве станций внутри города со средним и низким уровнем загрязнения  $PM_{10}$  – Шаболовка, Вешняки, МГУ, Останкино, Гурьянова и др. (рисунок 5.3 а, б).



a)



0 - Вешняки	6 - М (Балашиха-Речная)	11 - Пролетарский проспект	16 - МКАД 52 запад
1 - Черемушки	7 - Долгопрудная	12 - М1 (Очаковское)	17 - Технополис
2 - Сухареvская площадь	8 - Мелитопольская	13 - Гурьянова	18 - Останкино 0
3 - Кожухово	9 - Спиридоновка	14 - Очаковское-2	19 - М2 (Жулебино)
4 - Зеленоград 16	10 - МГУ	15 - Народного ополчения	20 - Туристская
5 - Шаболоvка			

b)

Рисунок 5.3 – (a) среднеквадратическое (SKOp-и) и среднее (COп-и) отклонение прогнозов PM<sub>10-CC</sub>, рассчитанных XTM CHIMERE (красные линии) и MMO-PM<sub>10</sub> (синие линии) от измерений на отдельных АСКЗА; (b) изменения (в %) SKOp-и прогнозов PM<sub>10-CC</sub> с применением MMO-постпроцессинга на отдельных АСКЗА. Март-сентябрь 2024 г.

## 5.2 ММО-постпроцессинг численных прогнозов концентраций $PM_{10}$ в периоды повышенного аэрозольного загрязнения

Как было показано в главе 2 (подраздел 2.2.2), причинами повышения общегородского уровня  $PM_{10}$  в Московском регионе являются НМУ и дальний перенос взвешенных частиц. Фиксируемые в иных условиях короткопериодные всплески концентраций  $PM_{10}$  на отдельных АСКЗА, как правило, обусловлены влиянием локальных близко расположенных источников.

Для идентификации повышенного загрязнения  $PM_{10}$  использован рекомендованный ВОЗ норматив среднесуточной концентрации  $PM_{10}$ , равный  $45 \text{ мкг/м}^3$  [World Health Organization, 2021]. В период независимого тестирования ММО- $PM_{10}$  март-сентябрь 2024 г. почти в каждый из 7 месяцев были дни с превышениями норматива ВОЗ  $PM_{10-CC}$  в московском регионе, будем называть такие периоды аэрозольными эпизодами.

В первом в тестовый период аэрозольном эпизоде 17–18 марта 2024 г. на большинстве АСКЗА зафиксированы  $PM_{10-CC}$  около  $100 \text{ мкг/м}^3$  и более, в целом за период 15–20 марта в московском регионе отмечено 76 превышений норматива ВОЗ. При наличии снежного покрова аэрозольное загрязнение имеет адвективное происхождение; траекторный анализ подтвердил это предположение, указывая на поступление в Москву воздушной массы, ранее перемещавшейся над западным Казахстаном и Прикаспием.

На рисунке 5.4 представлены концентрации  $PM_{10-CC}$ , измеренные на АСКЗА и рассчитанные ХТМ CHIMERE и ММО- $PM_{10}$  в период 15–20 марта 2024 г. 16–19 марта, в дни с превышениями норматива, ХТМ CHIMERE рассчитывала концентрации  $PM_{10}$  в диапазоне  $30–75 \text{ мкг/м}^3$ , недооценивая реальный уровень загрязнения, а ММО- $PM_{10}$  еще более понижала рассчитанный ХТМ уровень  $PM_{10}$  для всех АСКЗА (рисунок 5.4). Обе модели не учитывают фактор дальнего переноса загрязненных  $PM_{10}$  воздушных масс в московский регион.

Неординарность эпизода резкого повышения уровня  $PM_{10}$  27–28 марта 2024 г. связана с отсутствием переноса из южных регионов во время быстрого схода снежного покрова. 27.03 уровень  $PM_{10}$  с 19 до 22 ч на большинстве АСКЗА повысился в 2–4 и более раз, также резко (за 2–3 часа) произошел спад уровня  $PM_{10}$  до фоновых значений утром 28.03. Пиковые концентрации в длившемся чуть больше 12 часов эпизоде фиксировались с 22 ч (27.03) до 2 ч ночи (28.03). Синоптические условия (слабый ветер, приземная инверсия), безусловно, способствовали накоплению примеси в приземном воздухе, но различия в концентрации  $PM_{10}$  в часы наибольшего загрязнения между городскими и транспортными АСКЗА оказались чрезвычайно большими – более  $100 \text{ мкг/м}^3$ , например, максимальная концентрация  $PM_{10}$  на АСКЗА МГУ была около  $70 \text{ мкг/м}^3$ , на Сухаревской площади –  $270 \text{ мкг/м}^3$ . При НМУ, вследствие увеличения роли локальных источников, неоднородность поля  $PM_{10}$  значительно возрастает, что не всегда корректно воспроизводится ХТМ с разрешением 2 км.

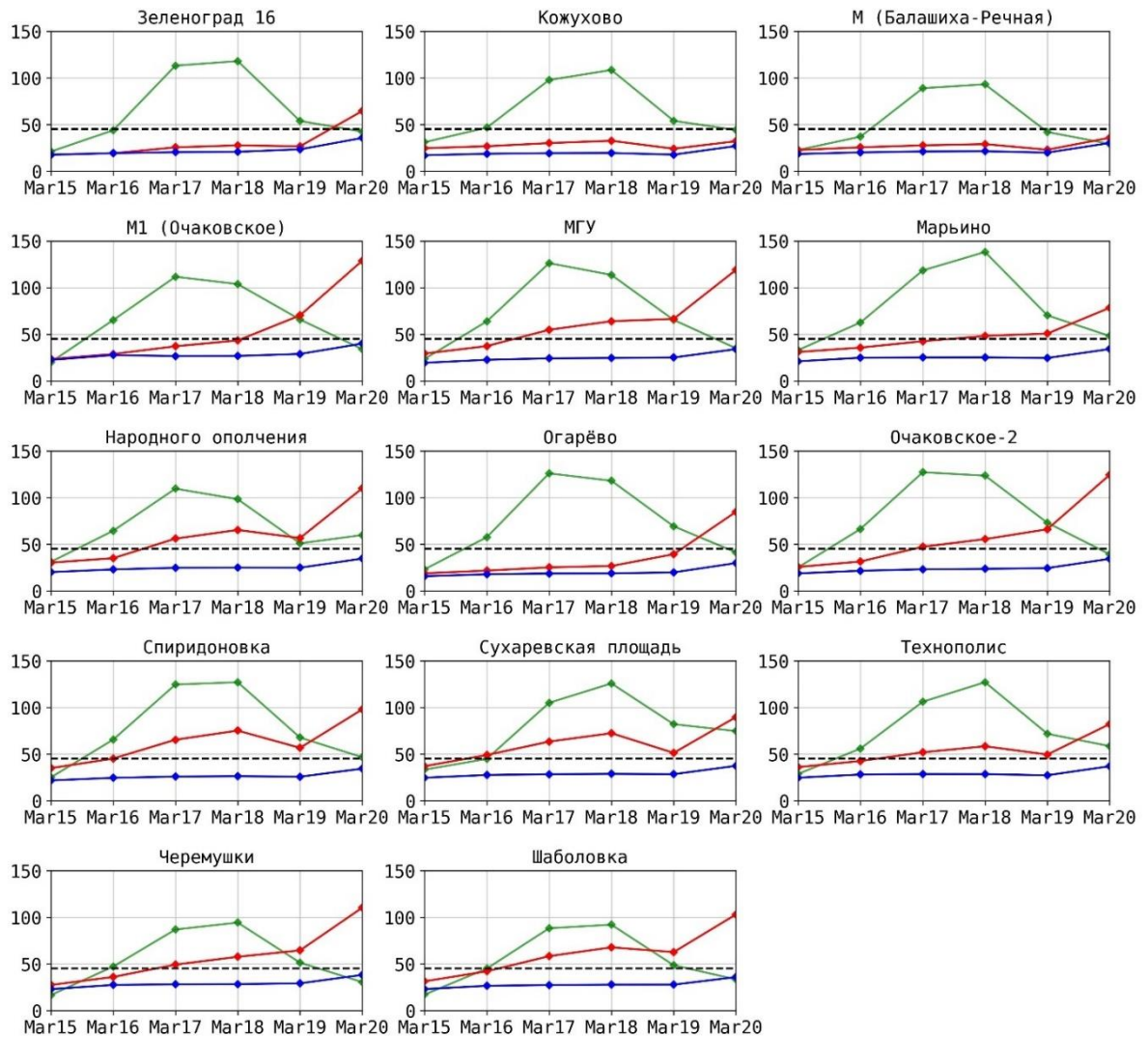


Рисунок 5.4 – Среднесуточные концентрации  $PM_{10}$ , измеренные на АСКЗА (зеленые линии), рассчитанные XTM SHIMERE (красные линии) и ММО- $PM_{10}$  (синие линии). 15-20 марта 2024 г.

*Примечание:* Пунктирная линия – норматив ВОЗ  $PM_{10-CC}$  45  $\mu\text{г}/\text{м}^3$ .

Важно отметить, что кратковременный всплеск уровня  $PM_{10}$  наблюдался при медленном перемещении над регионом оси барического гребня: слабые северные ветры 27 марта сменились на южные 28 марта (рисунок 5.5). Петлеобразное движение частиц на рисунке 5.5 указывает на то, что наблюдалось редкое явление – т. н. возврат загрязненной городскими выбросами воздушной массы, что обеспечило значительный рост концентраций  $PM_{10}$  вблизи локальных источников. Следует принять к сведению еще один факт: 28 марта по наблюдениям на метеостанциях в московском регионе еще отмечался снежный покров (высота 9–14 см), но в предшествующие 3–4 дня стояла сухая и аномально теплая погода (днем +7, 10 °С), благоприятная для освобождения городских поверхностей от снега (в первую очередь, автодорог и тротуаров, крыш и т. п.). Это, при развитой конвекции днем, приводило к поднятию пыли и насыщению воздушной массы аэрозолями.



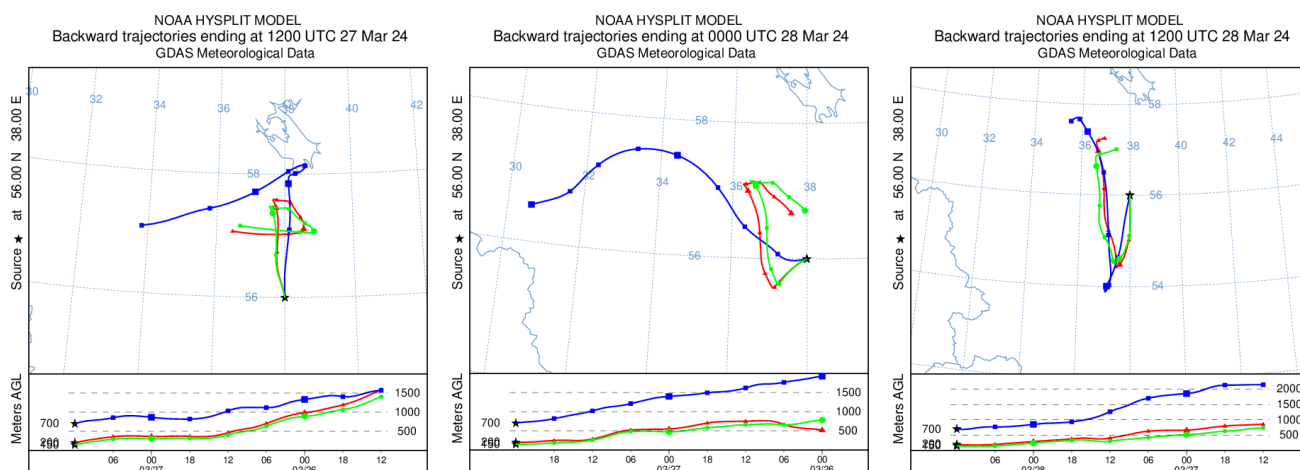


Рисунок 5.5 – Перемещение в предшествующие 3 суток воздушных частиц, поступивших в Москву 27 и 28 марта 2024 г.

На рисунке 5.8 видно, что ММО-PM<sub>10</sub> 28 марта для большинства станций корректирует прогнозы ХТМ CHIMERE (также вполне успешные), при этом учитывает сформировавшиеся НМУ и прогнозирует повышенное загрязнение PM<sub>10</sub> с превышениями норматива, недооценивая загрязнение только для примагистральных АСКЗА. Также можно обратить внимание на то, что ХТМ прогнозирует повышенное загрязнение и 27 марта, а ММО-PM<sub>10</sub> «снижает» расчеты ХТМ в соответствии с измерениями.

Зафиксированный в начале апреля 2024 г. эпизод сформировался с участием дальнего переноса пылевого аэрозоля из Северной Африки (рисунок 5.6). К 1 апреля снег сошел, в Москве юго-западные ветры (в АПС скорость до 11–15 м/с) привели к резкому потеплению. В пришедшем с юго-западными потоками субтропическом воздухе 1–2 апреля дневная температура повышалась до 21–24 °С при относительной влажности 30–40 %, но днем 3 апреля произошла смена воздушной массы с понижением температуры более чем на 10 °С.

Резкий всплеск концентраций PM<sub>10</sub> наблюдался к вечеру 2 апреля – концентрации повышались до 125 мкг/м<sup>3</sup> в среднем по городу, до 200–250 мкг/м<sup>3</sup> и выше на отдельных АСКЗА (рисунок 5.7). 3 апреля прошел холодный атмосферный фронт, ветер усиливался до 14–17 м/с, наблюдался «предфронтальный» всплеск концентраций PM<sub>10</sub>, вызванный поднятием частиц с поверхности при усилении ветра. После прохождения фронта с осадками концентрации PM<sub>10</sub> понизились до 10–40 мкг/м<sup>3</sup> к вечеру 3 апреля. Превышения норматива ВОЗ PM<sub>10-CC</sub> зафиксированы 2 и 3 апреля на 83 % и 75 % АСКЗА, соответственно.

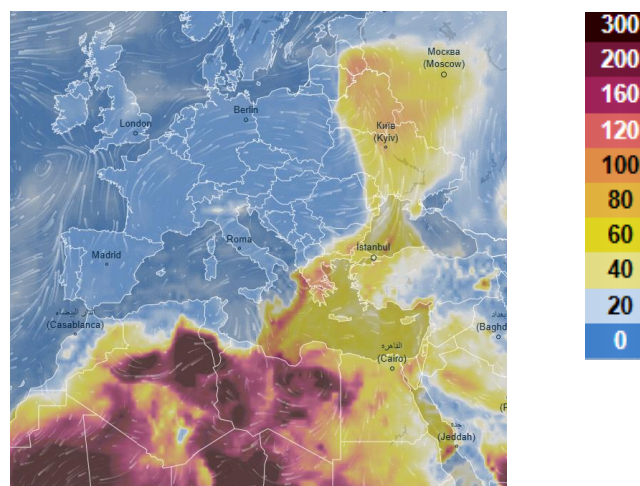
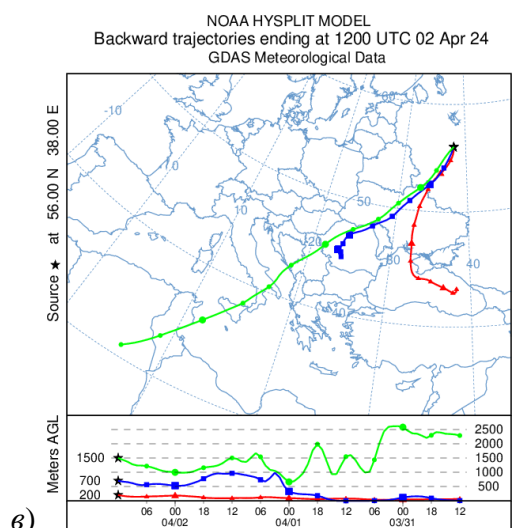
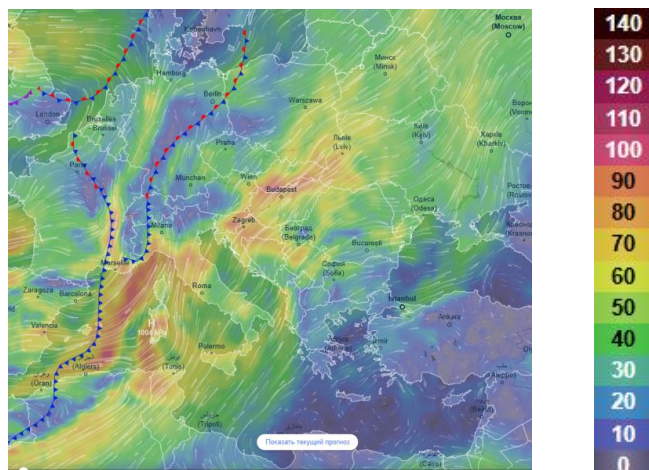
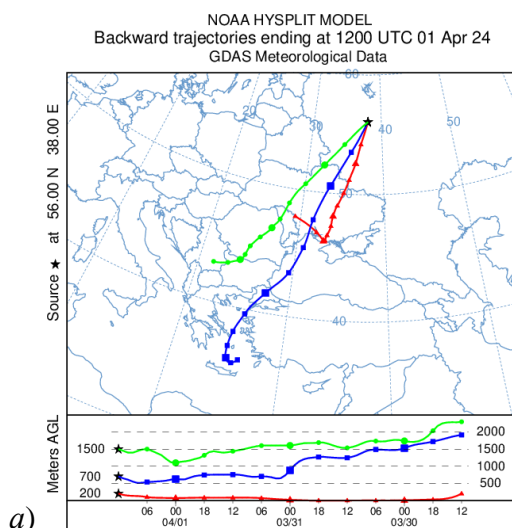


Рисунок 5.6 – Траектории воздушных частиц, поступивших в Москву 1 (a) и 2 (c) апреля; (b) поле скорости ветра на высоте 850 гПа (около 1500 м) при прохождении над Северной Африкой холодного атмосферного фронта; (d) расчет концентраций PM<sub>10</sub> XTM SILAM с ярко выраженным шлейфом африканского аэрозоля, распространяющимся в восточную Европу

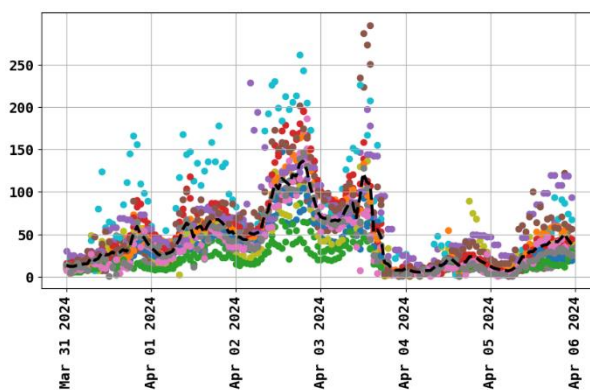


Рисунок 5.7 – Измеренные на АСКЗА в московском регионе часовые концентрации PM<sub>10</sub> (мкг/м<sup>3</sup>) 31 марта – 5 апреля 2024 г.

Хотя главным фактором повышения концентраций  $PM_{10}$  в данном эпизоде был дальний перенос частиц, ММО- $PM_{10}$  показала удовлетворительные результаты коррекции: на рисунке 5.8 видно, что 2 апреля ММО- $PM_{10}$  повышает расчеты ХТМ до превышений норматива практически на всех АСКЗА, но в основном недооценивает измеренные концентрации.

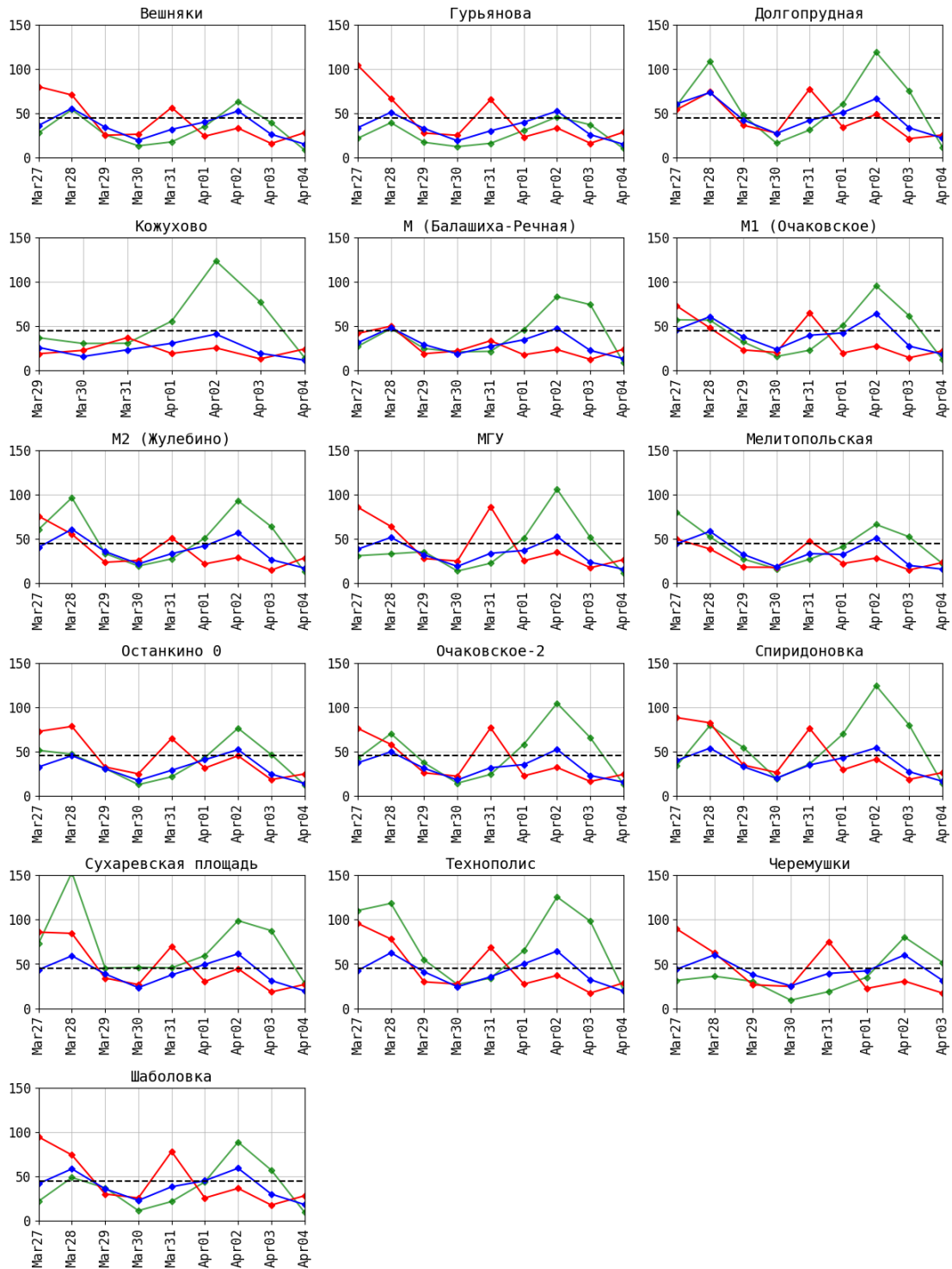


Рисунок 5.8 –  $PM_{10-CC}$ , измеренные на АСКЗА (зеленые линии), рассчитанные ХТМ СИМЕРЕ (красные линии) и ММО- $PM_{10}$  (синие линии). 27 марта – 4 апреля 2024 г.

*Примечание:* Пунктирная линия – норматив ВОЗ  $PM_{10-CC}$  45  $\mu g/m^3$ .

Из этого можно сделать вывод о том, что ММО-PM<sub>10</sub> учитывает фактор НМУ в дни максимальных концентраций (28 марта и 2 апреля), однако, ожидаемо, не учитывает фактор дальнего переноса аэрозолей.

Примечательно, что ни ХТМ, ни ММО-PM<sub>10</sub> не учли повышение PM<sub>10-CC</sub> 3 апреля, возникшее из-за «предфронтального» всплеска концентраций.

Похожий по происхождению аэрозольный эпизод зафиксирован в Москве 25–26 апреля, он был связан с приходом в московский регион субтропической воздушной массы в теплом секторе средиземноморского циклона. В эти дни в Москве температура повышалась днем до + 21, 23 °С при влажности 53–57 %. Повышение уровня PM<sub>10</sub> началось днем 25 апреля при скорости ветра в АПС (на 925 гПа) до 16 м/с. В ночь на 26 апреля ветер стих и начался рост концентраций, средняя по городу (часовая) концентрация PM<sub>10</sub> возросла до 125 мкг/м<sup>3</sup>, на отдельных АСКЗА до 250 мкг/м<sup>3</sup> (рисунок 5.9 а). 26 апреля на всех АСКЗА в московском регионе зарегистрированы превышения норматива ВОЗ PM<sub>10-CC</sub>.

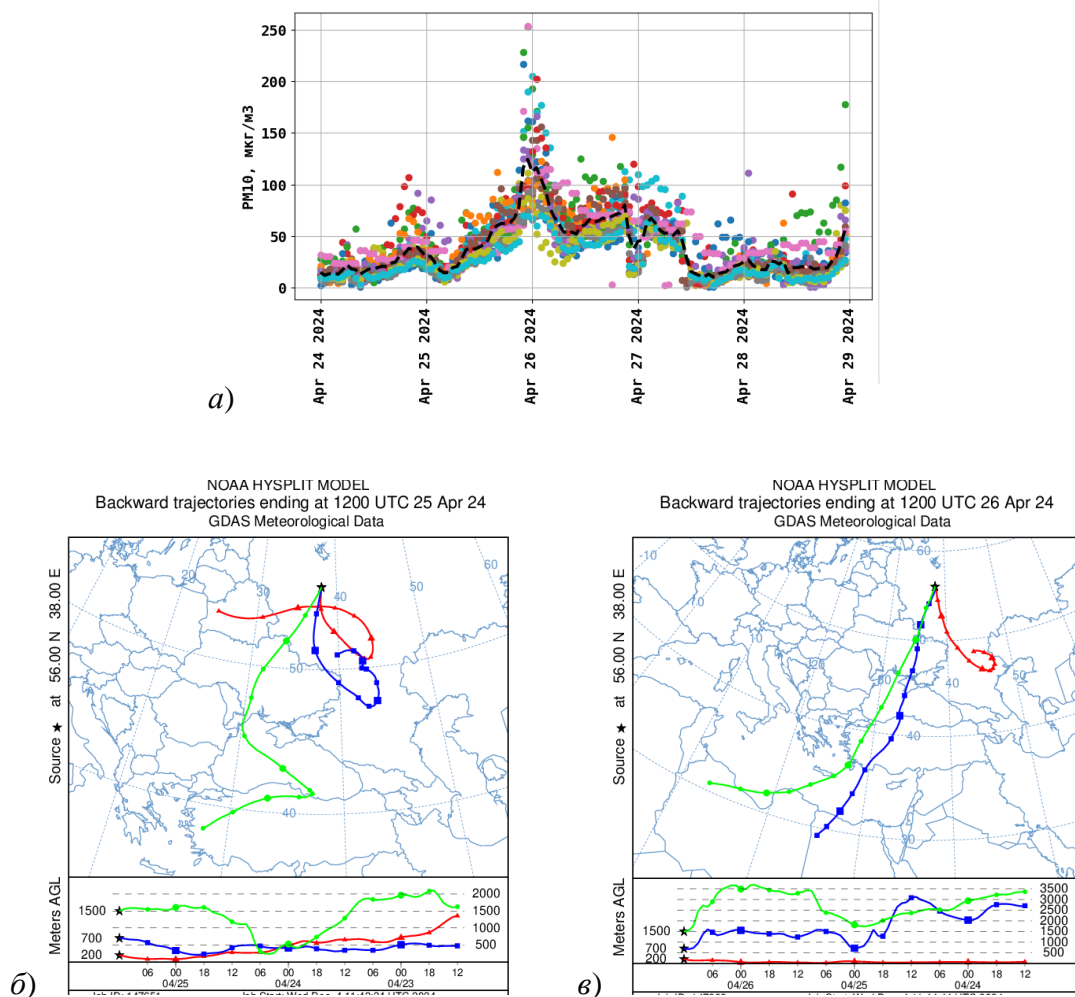


Рисунок 5.9 – (а) Измеренные на АСКЗА часовые концентрации PM<sub>10</sub> (мкг/м<sup>3</sup>) 24–28 апреля 2024 г.; траектории воздушных частиц, поступавших в московский регион (б) 25 апреля 2024 г., (в) 26 апреля 2024 г.

Основным фактором повышенного загрязнения в этом эпизоде стал, как показал траекторный анализ (рисунок 5.9 б, в), дальний перенос мелкого аэрозоля из северной Африки, Турции, где насыщение воздуха песчаными частицами происходило при прохождении холодного фронта через эти регионы 23 апреля – за двое суток до повышения уровня  $PM_{10}$  в Москве.

В данном эпизоде при НМУ (26.04) ММО- $PM_{10}$  на большинстве АСКЗА повышала рассчитанные ХТМ концентрации, приближая их к измеренным и прогнозируя превышения норматива ВОЗ (рисунок 5.10). Но все же ММО- $PM_{10}$  на многих АСКЗА недооценила измеренные концентрации из-за неучета фактора адвекции частиц.

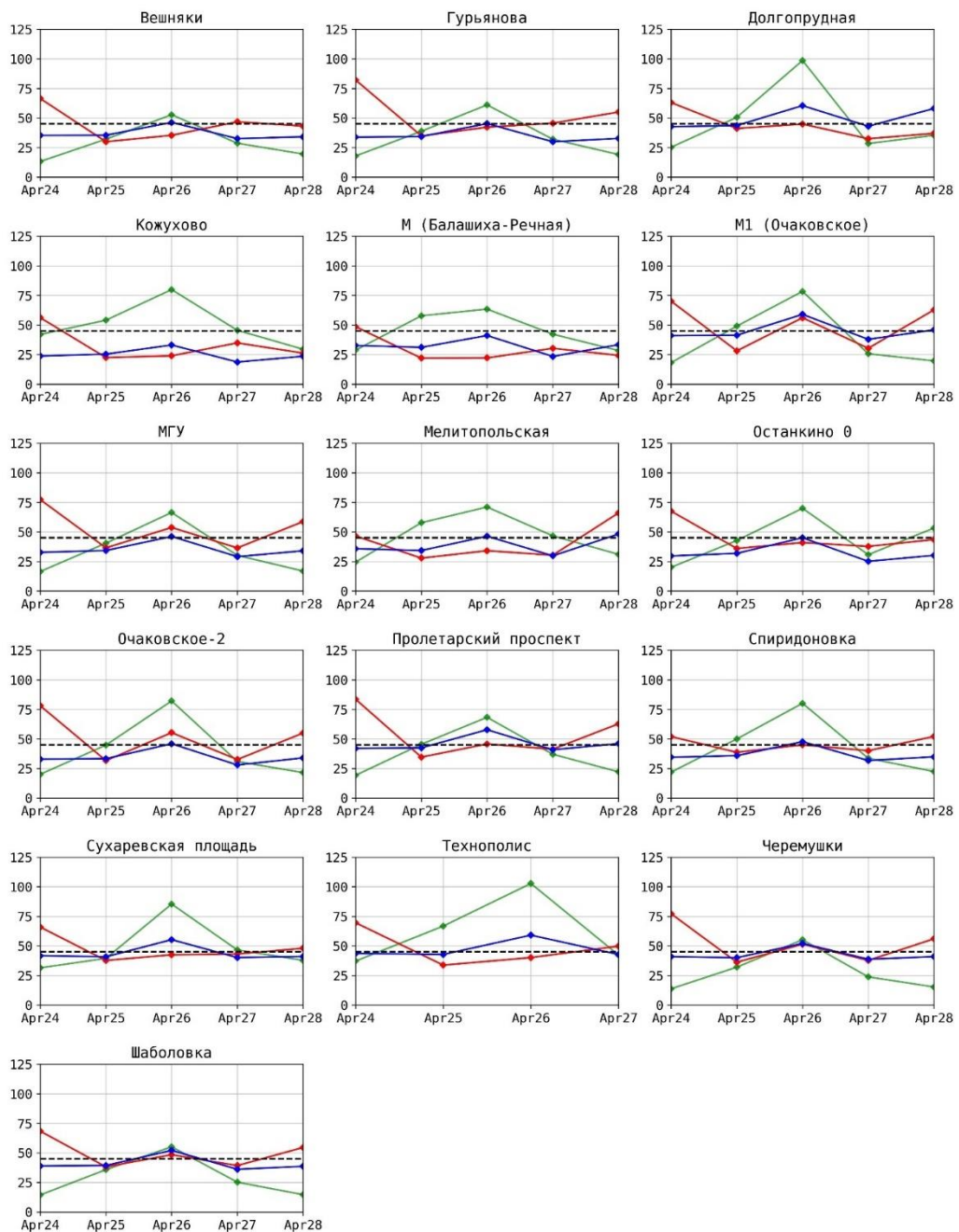


Рисунок 5.10 –  $PM_{10-CC}$ , измеренные на АСКЗА (зеленые линии), рассчитанные ХТМ SHIMERE (красные линии) и ММО- $PM_{10}$  (синие линии). 24 –28 апреля 2024 г.

*Примечание:* Пунктирная линия – норматив ВОЗ  $PM_{10-CC}$   $45 \text{ мкг/м}^3$ .

Зафиксированные в мае и сентябре эпизоды повышенного загрязнения  $PM_{10}$  (летом их не было) наблюдались в периоды НМУ при слабых скоростях ветра в малоградиентных барических полях, температурных инверсиях, жаркой и засушливой погоде и, что важно, при закрытости подстилающей поверхности травяным покровом.

В конце мая в Москве установилась аномально жаркая и сухая погода (относительная влажность днем 20–35 %) на западной периферии блокирующего антициклона (с меридиональной осью над Уралом). В дневные часы, благодаря активной конвекции, уровень  $PM_{10}$  понижался на АСКЗА всех типов до фоновых значений; в ночные часы, в штилевых условиях и при образовании температурных инверсий, концентрации  $PM_{10}$  резко возрастали (рисунок 5.11). Повышение уровня  $PM_{10}$  начиналось с заходом солнца и сохранялось до утра вследствие слабого рассеивания при НМУ. 29 мая ночью средняя по городу концентрация  $PM_{10}$  была около  $75 \text{ мкг/м}^3$ , на отдельных АСКЗА концентрация увеличивалась до  $100\text{--}150 \text{ мкг/м}^3$  (рисунок 5.11). Траекторный анализ исключил адвективный фактор повышения концентраций  $PM_{10}$ .

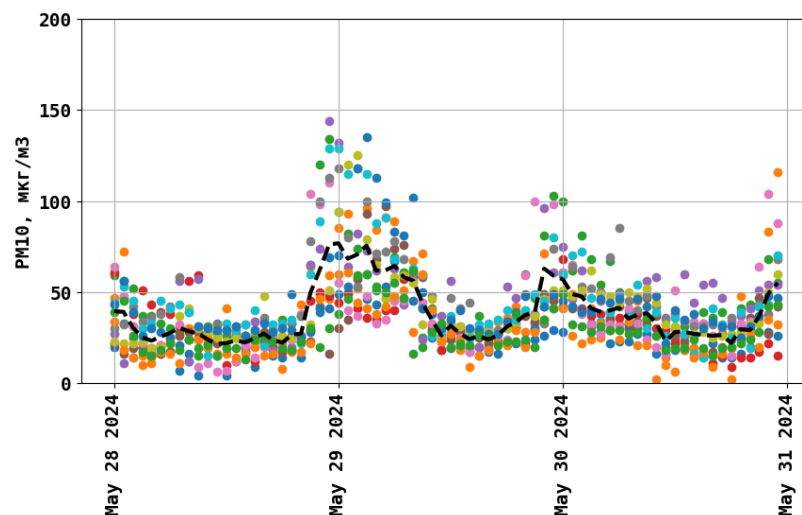


Рисунок 5.11 – Измеренные на АСКЗА часовые концентрации  $PM_{10}$  28–30 мая 2024 г.

В данном эпизоде ХТМ SHIMERE рассчитывала среднесуточную концентрацию  $PM_{10}$  в диапазоне  $60\text{--}85 \text{ мкг/м}^3$  в Москве и  $30\text{--}40 \text{ мкг/м}^3$  на подмосковных станциях (Зеленоград на рисунке 5.12). ММО- $PM_{10}$  в данном эпизоде при НМУ корректирует прогнозы, уменьшая завышенные расчеты ХТМ  $PM_{10-CC}$  на городских станциях, при этом прогнозируя превышения норматива ВОЗ или близкие к превышению концентрации  $PM_{10-CC}$  (рисунок 5.12). На подмосковной АСКЗА Зеленоград ХТМ занижает прогнозы  $PM_{10-CC}$ , ММО- $PM_{10}$  «вытягивает» концентрации и прогнозирует превышения норматива в соответствии с измерениями.

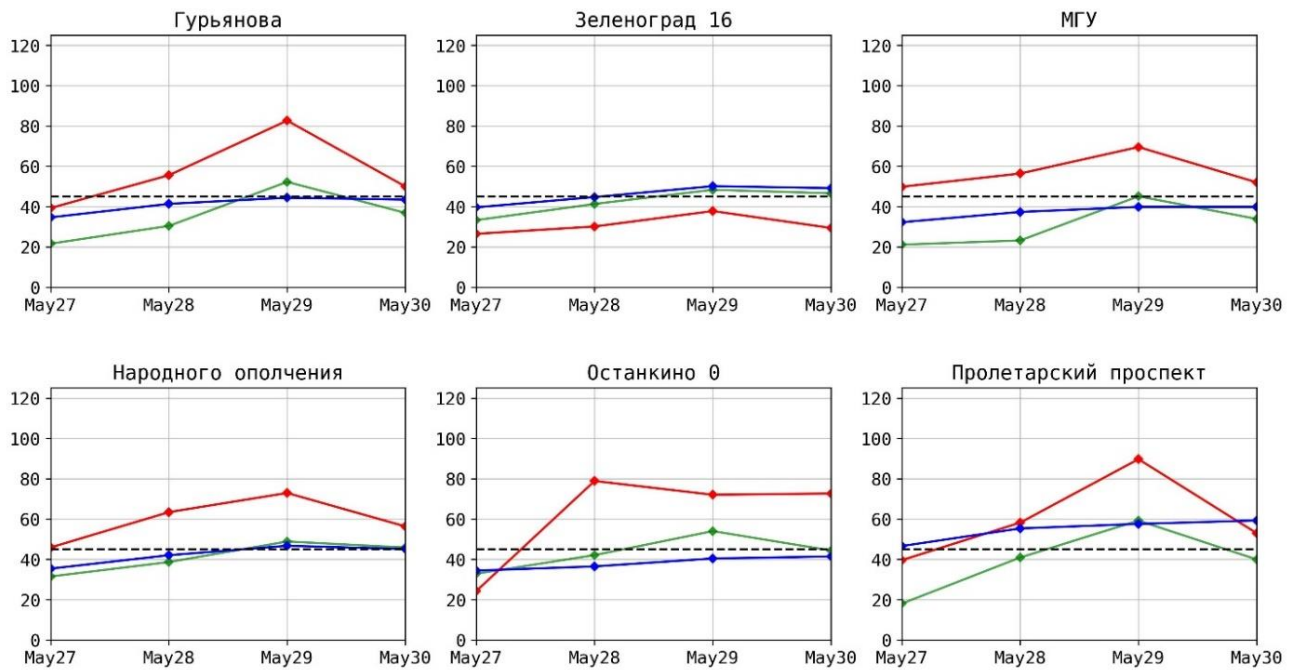


Рисунок 5.12 –  $PM_{10-CC}$ , измеренные на АСКЗА (зеленые линии), рассчитанные ХТМ CHIMERE (красные линии) и ММО- $PM_{10}$  (синие линии). 27–30 мая 2024 г.

*Примечание:* Пунктирная линия – норматив ВОЗ  $PM_{10-CC}$  45  $mcg/m^3$ .

Хотя лето в Москве оказалось более чем на 2 °С теплее обычного, но благодаря длинному световому дню и преобладавшим условиям умеренного рассеивания, значимых НМУ не формировалось и уровень  $PM_{10}$  находился в пределах многолетней изменчивости, был невысоким.

Иная ситуация сложилась в сентябре. Средний за месяц уровень  $PM_{10}$  в Москве оказался в 1,5 раза больше, чем весной, и в 1,7 раз больше летнего! Основной причиной повышенного аэрозольного загрязнения в сентябре стали аномальные погодные условия: за месяц наблюдались три продолжительные волны тепла и в целом месяц оказался на 4 °С теплее обычного (<https://meteoinfo.ru/climat-tabl3>). Следует отметить рекордно теплую и сухую погоду в начале осени и в умеренных, и в южных широтах ЕТР. К тому же, на юге с августа наблюдалась почвенная засуха. Соответственно, жара и засуха способствовали ветровой эрозии почв и стали предпосылками для насыщения почвенными частицами воздушных масс, перемещавшихся над южными регионами. Именно с волнами тепла при поступлении воздушных масс с юга в Москве наблюдалось повышение уровня  $PM_{10}$ , при наступлении затишья и наличии инверсии (т.е. НМУ) формировались наиболее высокие уровни аэрозольного загрязнения.

Такие неблагоприятные условия сложились 6–8 сентября вблизи центра антициклона, в штилевых условиях ночные инверсии усиливались до 5–7 °С. При НМУ средний по городу уровень  $PM_{10}$  повышался вечером-ночью до 60–70  $mcg/m^3$ , на отдельных АСКЗА до 100–150  $mcg/m^3$  (рисунок 5.13).

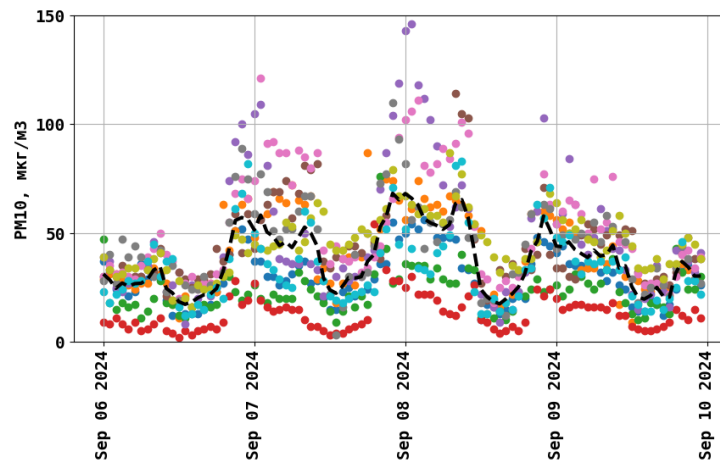


Рисунок 5.13 – Измеренные на АСКЗА часовые концентрации  $PM_{10}$  6–9 сентября 2024 г.

Как видно на рисунке 5.14, в этом эпизоде, при измерениях на большинстве АСКЗА концентрациях  $PM_{10-CC}$  50-75  $мкг/м^3$ , расчеты ХТМ СИМЕРЕ оказались существенно завышенными практически на всех АСКЗА. ММО- $PM_{10}$  снижает прогнозы ХТМ до 50–60  $мкг/м^3$ , приближая их к измерениям. При этом ММО- $PM_{10}$  прогнозирует превышения норматива ВОЗ практически на всех АСКЗА.

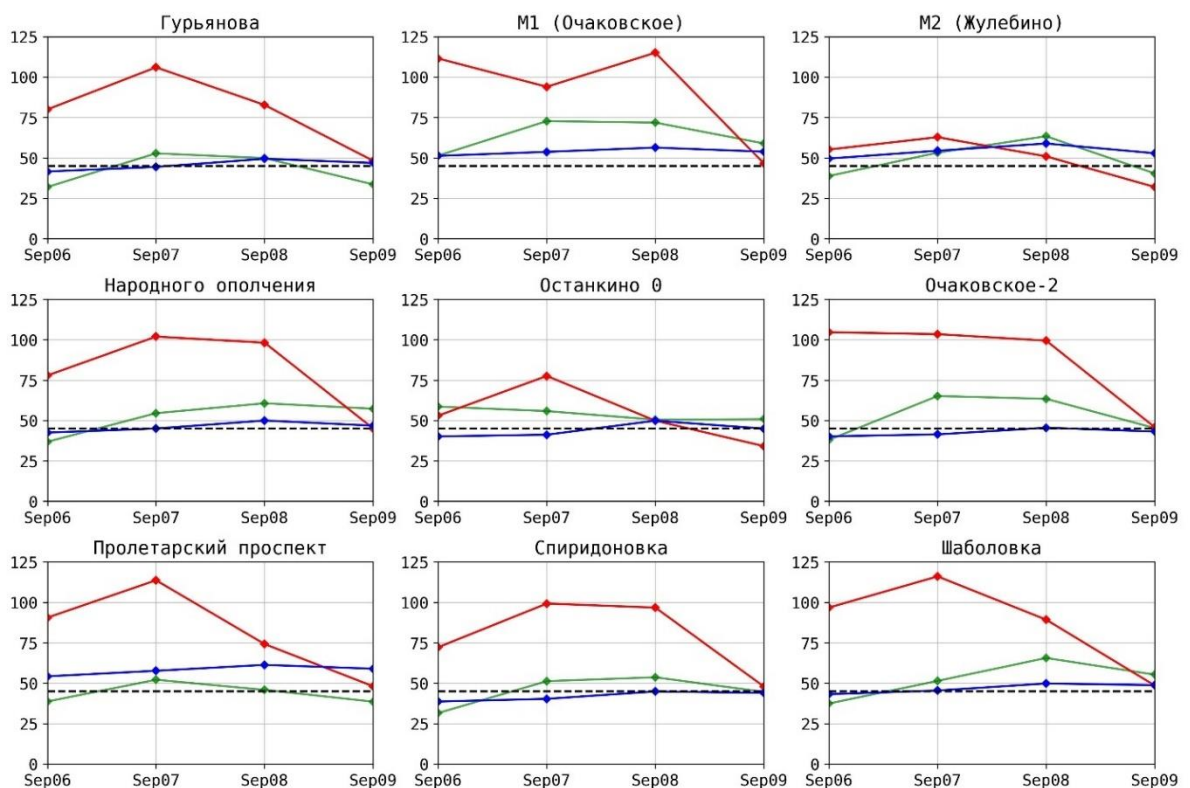


Рисунок 5.14 –  $PM_{10-CC}$ , измеренные на АСКЗА (зеленые линии), рассчитанные ХТМ СИМЕРЕ (красные линии) и ММО- $PM_{10}$  (синие линии) 6–9 сентября 2024 г.

*Примечание:* Пунктирная линия – норматив ВОЗ  $PM_{10-CC}$  45  $мкг/м^3$ .



Еще два сентябрьских аэрозольных эпизода – в середине (16–18 сентября) и в конце (26–28 сентября), формировались в условиях адвекции загрязненных  $PM_{10}$  воздушных масс из охваченных засухой юго-восточных районов ЕТР и Прикаспия (рисунок 5.15).

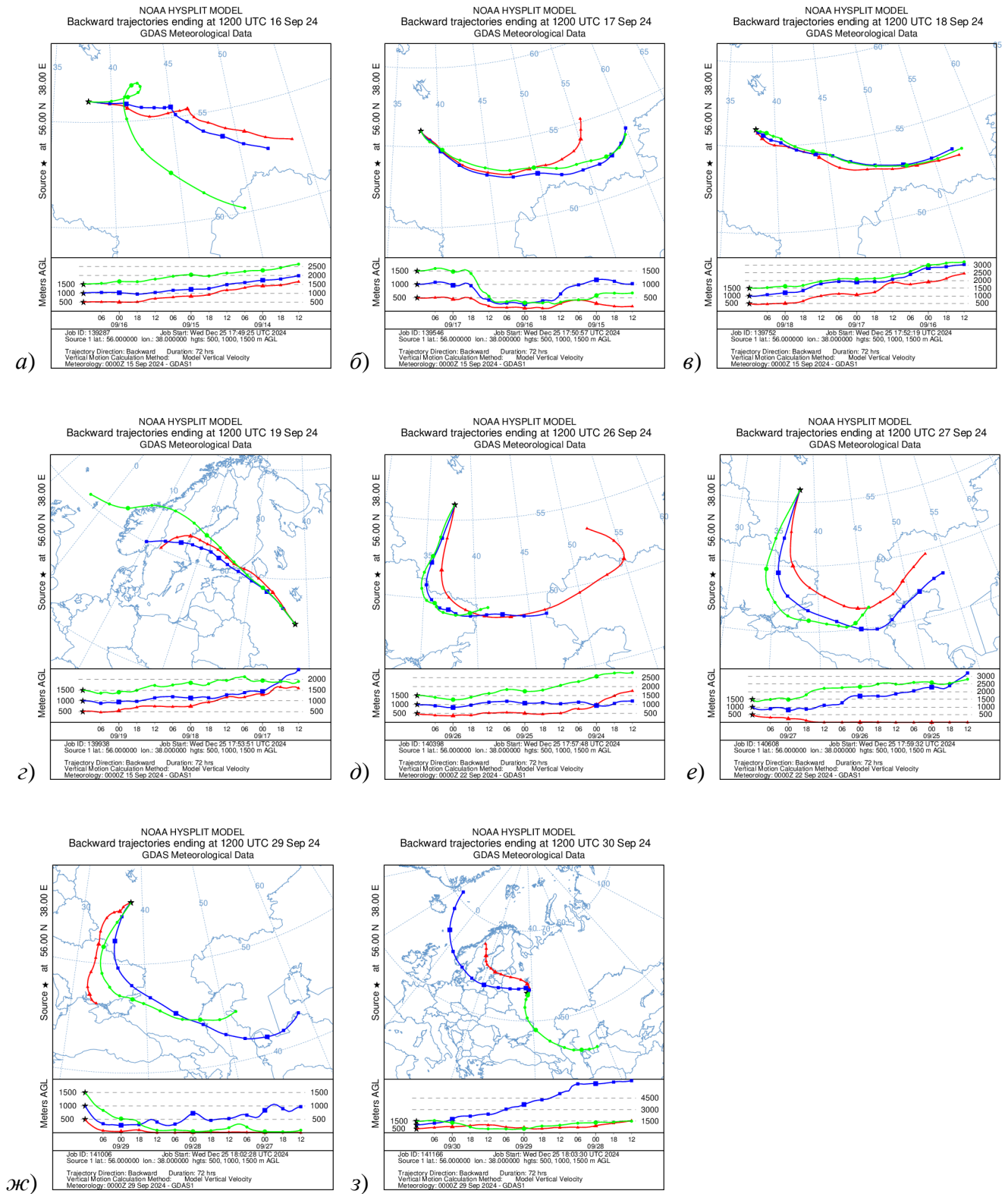


Рисунок 5.15 – Траектории воздушных частиц, поступавших в московский регион в аэрозольном эпизоде 16–18 сентября (а-в) и при его окончании 19 сентября (г); в аэрозольном эпизоде 26–29 сентября (д-ж) и при его окончании 30 сентября (з)

Эффекты ММО-постпроцессинга в подобных аэрозольных эпизодах обсуждались выше. Во избежание повторений укажем только на то, что в разработанной ММО-PM<sub>10</sub> отсутствуют предикторы, отражающие сильное загрязнение PM<sub>10</sub> на удаленных территориях с засухой и адвекцию воздушных масс оттуда в московский регион, поэтому ММО-PM<sub>10</sub> не может качественно предсказывать подобные события.

Явные эффекты ММО-постпроцессинга видны в эпизодах повышенного загрязнения PM<sub>10</sub>, обусловленных преимущественно локальными НМУ, т.е. ММО-PM<sub>10</sub> «различает» формирование неблагоприятных метеоусловий, способствующих накоплению PM<sub>10</sub> в приземном слое.

### 5.3 Эффекты ММО-постпроцессинга полей прогнозов концентраций PM<sub>10</sub>

Технология ММО-постпроцессинга поля прогнозов концентраций ХТМ подробно описана в разделе 3.1, в общем виде она представляет собой расчет ММО приземных концентраций загрязнителя для всех ячеек модельной сетки в расчетном домене ХТМ с усвоением прогностических предикторов (см. таблицу 3.1) из этих же ячеек модельной сетки. Как и для ММО-О<sub>3</sub>, обоснованием применения ММО-PM<sub>10</sub> для постпроцессинга прогностических полей является доказанная эффективность ММО-постпроцессинга прогнозов концентраций PM<sub>10</sub> на территориях без измерений концентраций (при исключении станции из обучающей выборки, см. раздел 5.1).

Отдельные иллюстрации ММО-постпроцессинга часовых прогностических полей концентраций PM<sub>10</sub> на территории московского региона представлены на рисунках 5.16, 5.17, где визуализированы прогнозы PM<sub>10</sub> ХТМ CHIMERE (слева) и ММО-скорректированные прогнозы (справа) на первые сутки 29 мая 2024 г. В этот день в малоградиентном барическом поле при безветренной погоде и ночной инверсии наблюдалось повышение концентраций PM<sub>10</sub> до превышений норматива ВОЗ (см. раздел 5.2, рисунки 5.11, 5.12). Прогнозы представлены за 2 срока – на рисунке 5.16 за срок 02:00 МСК – ночной максимум концентраций PM<sub>10</sub>, на рисунке 5.17 за срок 12:00 МСК – дневной минимум концентраций. На рисунки также нанесены часовые измерения концентраций PM<sub>10</sub> в соответствующее срокам прогнозов время.

Ночью 29 мая при НМУ ХТМ CHIMERE прогнозировала повышение концентраций PM<sub>10</sub> в северной части Москвы, достаточно точно предсказывая концентрации для станций Сухаревская площадь, Останкино, Долгопрудная (рисунок 5.16, слева). В южной части города CHIMERE в основном занизила прогнозы PM<sub>10</sub>.

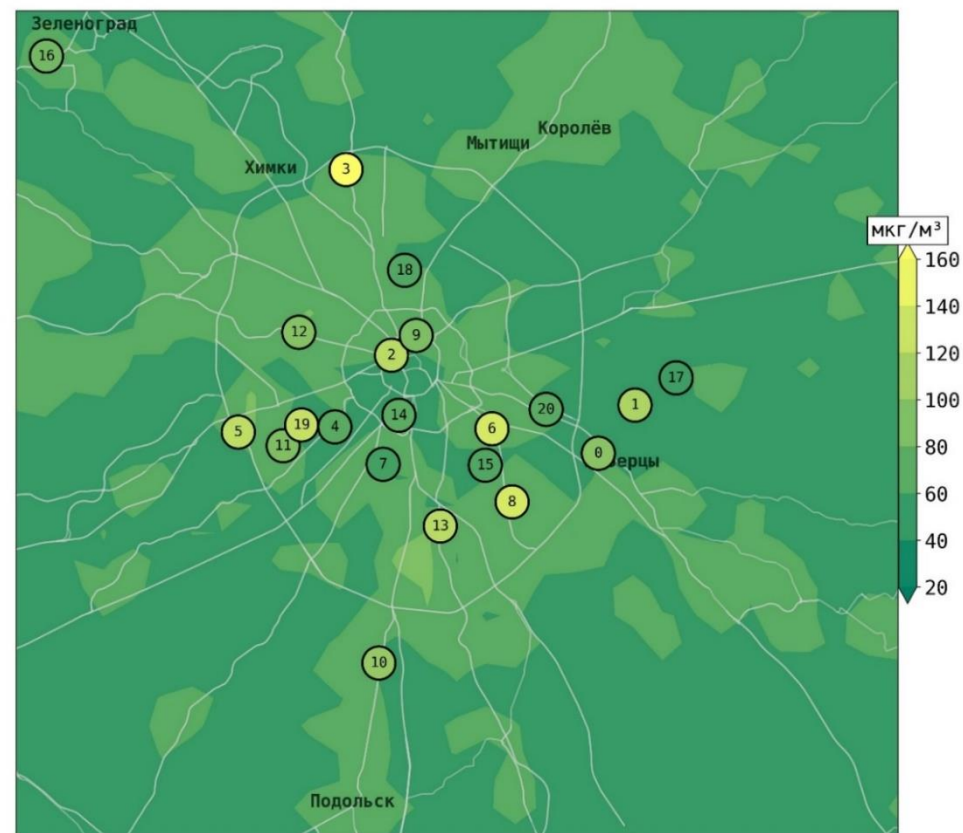
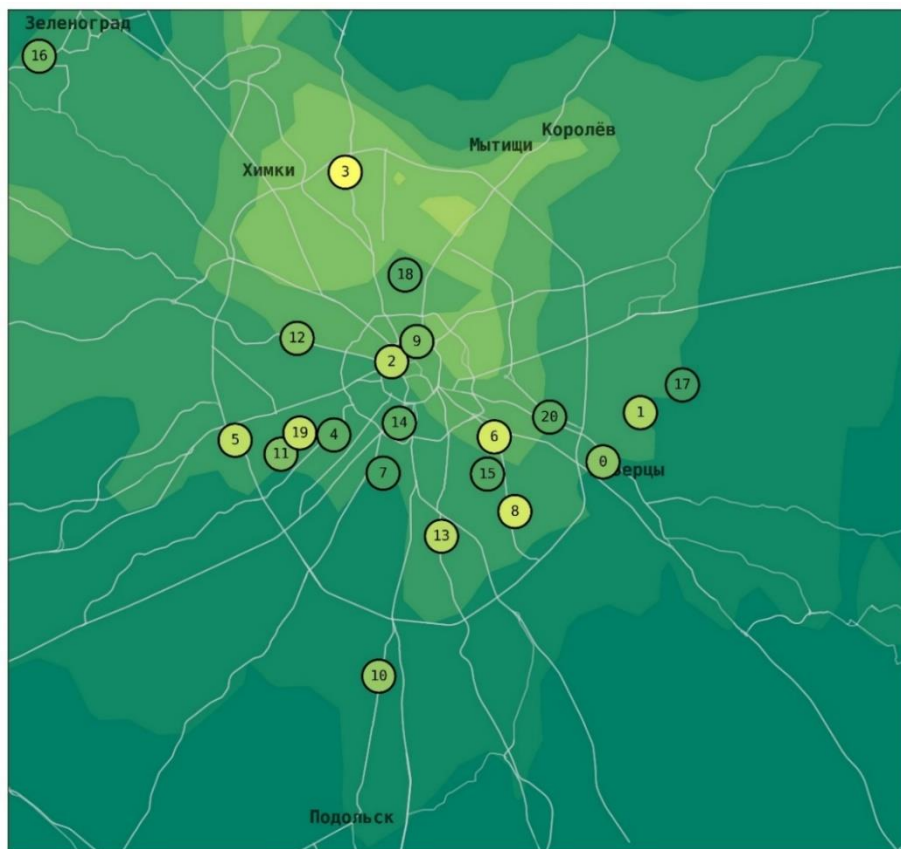
ММО-PM<sub>10</sub>, очевидно, учитывая распределение плотности городской застройки (как предиктор в обучающей выборке), корректирует прогнозы CHIMERE, повысив концентрации PM<sub>10</sub> практически на всех территориях, кроме северной части Москвы (рисунок 5.16, справа). Для территорий с городской застройкой ММО-PM<sub>10</sub> рассчитала уровень концентраций 60–80 мкг/м<sup>3</sup>, для парковых зон (Лосинный остров на северо-востоке Москвы, Измайловский парк на востоке) и фоновых территорий примерно на 20 мкг/м<sup>3</sup> меньше – 40–60 мкг/м<sup>3</sup>. В итоге ММО-PM<sub>10</sub> довольно точно рассчитала загрязнение для станций на удалении от транспортных выбросов (например, АСКЗА Останкино, Вешняки, Балашиха-Речная на рисунок 5.16, справа), но занизила прогнозы PM<sub>10</sub> для примагистральных станций (например, МКАД 52 запад, Спиридоновка, Технополис).

Днем 29 мая (рисунок 5.17) загрязнение PM<sub>10</sub> существенно уменьшилось за счет выноса частиц в верхние слои АПС при интенсивном вертикальном перемешивании. При этом ХТМ CHIMERE прогнозирует максимальное загрязнение PM<sub>10</sub> – до 160 мкг/м<sup>3</sup> и выше на северо-западе и юго-востоке Москвы, а также к югу и к юго-западу от Москвы (рисунок 5.17, слева). Минимумы концентраций – 20–40 мкг/м<sup>3</sup> и менее – CHIMERE прогнозирует для парковых зон и фоновых территорий без плотной застройки. В итоге днем почти для всех АСКЗА ХТМ значительно завышает прогнозы концентраций, за счет заданного некорректного суточного распределения объемов выбросов.

ММО-PM<sub>10</sub> успешно корректирует прогностическое поле дневных концентраций PM<sub>10</sub>, снижая концентрации до 40–60 мкг/м<sup>3</sup> на территории города, для парковых зон и фоновых территорий ММО-PM<sub>10</sub> рассчитывает концентрации на уровне 20–40 мкг/м<sup>3</sup> (рисунок 5.17, справа). В итоге ММО-PM<sub>10</sub> занижает прогноз на 20–40 мкг/м<sup>3</sup> только для некоторых примагистральных станций (Очаковское, Сухаревская площадь, Технополис на рисунке 5.17, справа), для большинства АСКЗА, в том числе фоновых, точно рассчитывая уровень загрязнения PM<sub>10</sub>.

XTM CHIMERE

ММО-PM<sub>10</sub>



0 - М2 (Жулебино)	5 - МКАД 52 запад	9 - Сухаревская площадь	13 - Пролетарский проспект	17 - М (Балашиха-Речная)
1 - Кожухово	6 - Технополис	10 - Мелитопольская	14 - Шаболовка	18 - Останкино 0
2 - Спиридоновка	7 - Черемушки	11 - М1 (Очаковское)	15 - Гурьянова	19 - Очаковское-2
3 - Долгопрудная	8 - Марьино	12 - Народного ополчения	16 - Зеленоград 16	20 - Вешняки
4 - МГУ				

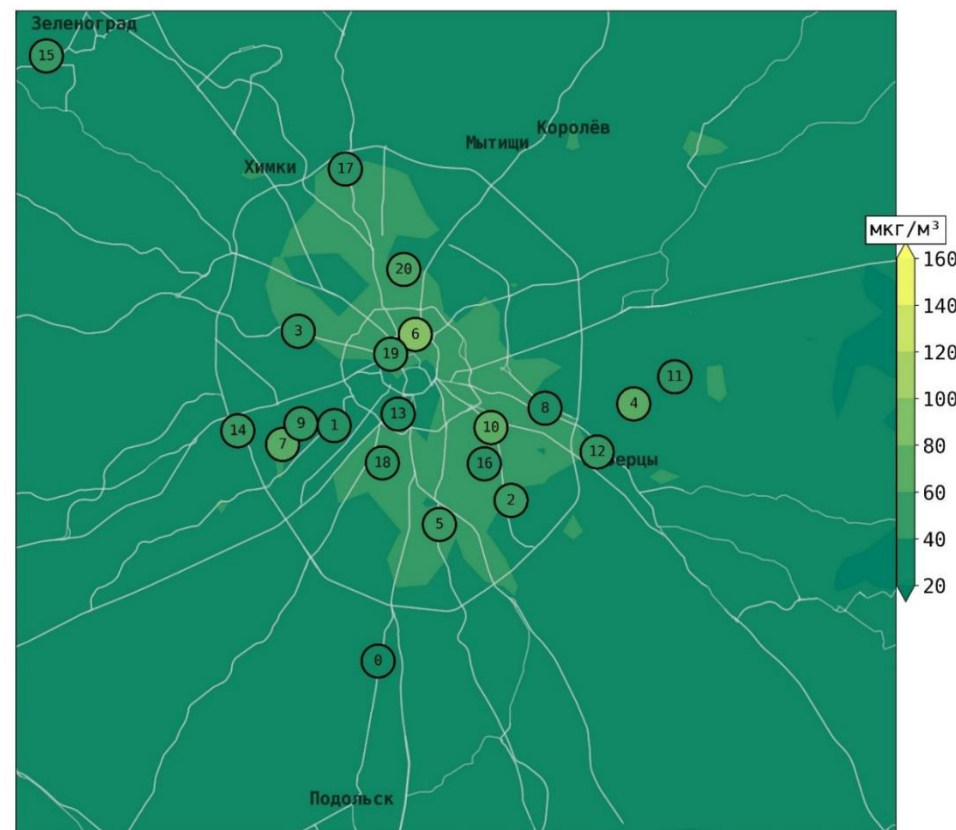
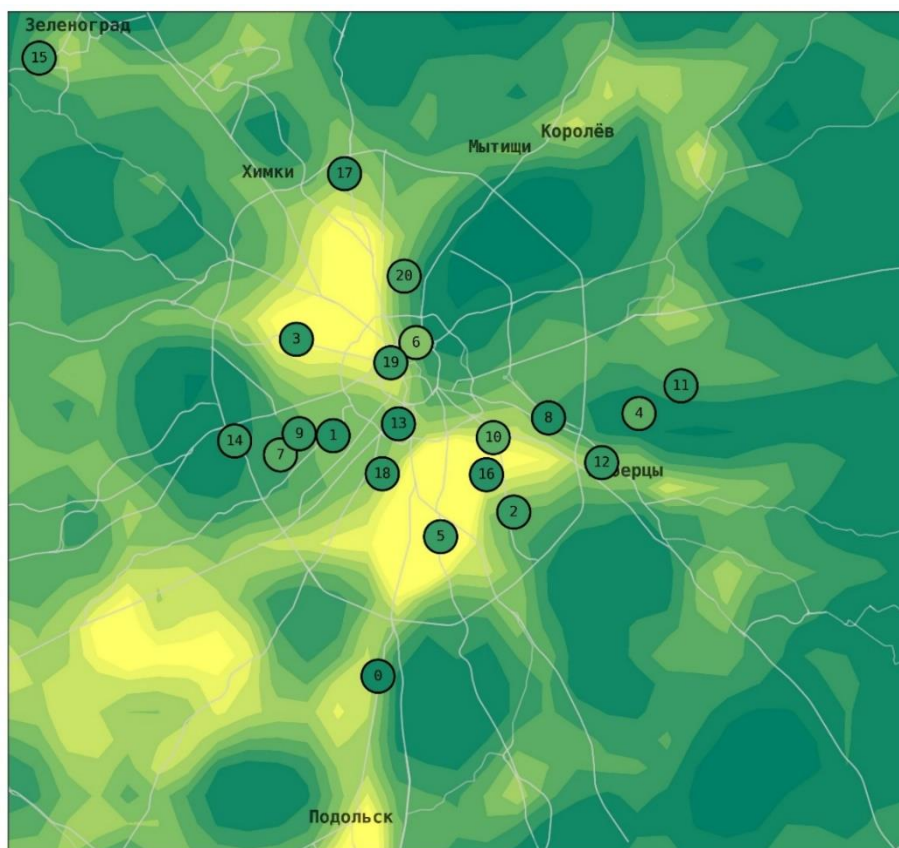
Рисунок 5.16 – Прогнозы на первые сутки приземных концентраций PM<sub>10</sub> на территории московского региона 29 мая 2024 г. 02:00 МСК

XTM CHIMERE (слева), ММО-PM<sub>10</sub> (справа). Кружками нанесены соответствующие сроку прогнозов часовые измерения

концентраций PM<sub>10</sub> на АСКЗА в той же цветовой гамме, что и прогностические поля

XTM CHIMERE

ММО-PM<sub>10</sub>



0 - Мелитопольская	5 - Пролетарский проспект	9 - Очаковское-2	13 - Шаболовка	17 - Долгопрудная
1 - МГУ	6 - Сухареvская площадь	10 - Технополис	14 - МКАД 52 запад	18 - Черемушки
2 - Марьино	7 - М1 (Очаковское)	11 - М (Балашиха-Речная)	15 - Зеленоград 16	19 - Спиридоновка
3 - Народного ополчения	8 - Вешняки	12 - М2 (Жулебино)	16 - Гурьянова	20 - Останкино 0
4 - Кожухово				

Рисунок 5.17 – Прогнозы на первые сутки приземных концентраций PM<sub>10</sub> на территории московского региона 29 мая 2024 г. 12:00 МСК  
 XTM CHIMERE (слева), ММО-PM<sub>10</sub> (справа). Кружками нанесены соответствующие сроку прогнозов часовые измерения  
 концентраций PM<sub>10</sub> на АСКЗА в той же цветовой гамме, что и прогностические поля

**Выводы к главе 5**

Сравнительный анализ первичных прогнозов ХТМ и ММО-скорректированных прогнозов концентраций  $PM_{10}$  на территории московского региона в период тестовой независимой выборки март-сентябрь 2024 г. показал успешность применения ММО- $PM_{10}$  – СКО<sub>П-И</sub> ( $PM_{10-CC}$ ) уменьшилось по всей тестовой выборке с 25 мкг/м<sup>3</sup> (ХТМ CHIMERE) до 15 мкг/м<sup>3</sup> (ММО- $PM_{10}$ ) – на 40 %. Максимальная эффективность ММО-постпроцессинга выявлена в летние месяцы – СКО<sub>П-И</sub> ( $PM_{10-CC}$ ) ММО-скорректированных прогнозов 8–9 мкг/м<sup>3</sup>, на 44–62 % меньше, чем по расчетам ХТМ.

ММО- $PM_{10}$  успешно корректирует существенные отклонения воспроизводимого ХТМ суточного хода концентраций  $PM_{10}$ , коэффициенты корреляции прогноз-измерение увеличены с 0.1–0.2 до 0.4–0.6.

С применением ММО-постпроцессинга в 1.5–2 раза снижены СКО<sub>П-И</sub> ( $PM_{10-CC}$ ) на городских АСКЗА, с учетом исключения АСКЗА из обучающей выборки ММО- $PM_{10}$ . Минимальные эффекты ММО-постпроцессинга наблюдаются на близко расположенных к локальному источнику загрязнения АСКЗА.

ММО- $PM_{10}$  успешно корректирует прогнозы концентраций  $PM_{10}$  в периоды повышенного аэрозольного загрязнения при НМУ. ММО- $PM_{10}$  ожидаемо недооценивает аэрозольное загрязнение, обусловленное переносом частиц из удаленных районов, поскольку в ММО- $PM_{10}$  отсутствуют соответствующие предикторы.

Продемонстрирована успешность применения ММО-постпроцессинга часовых полей концентраций  $PM_{10}$ , смоделированных ХТМ CHIMERE для территории московского региона. ММО- $PM_{10}$  снижает завышенные ХТМ дневные концентрации  $PM_{10}$  по всей территории, в ночное время корректирует пространственную неоднородность загрязнения  $PM_{10}$  в соответствии с распределением городской застройки.

## ЗАКЛЮЧЕНИЕ

По итогам проведенного исследования сделаны следующие основные выводы:

1. Разработана и применена в исследованиях методика построения моделей машинного обучения для постпроцессинга численных прогнозов приземных концентраций загрязняющих веществ. ММО обучаются для каждого отдельного загрязнителя, в качестве предикторов обучающей выборки используются часовые численные прогнозы концентраций, метеопараметров и данные землепользования, в качестве целевой переменной задаются часовые измерения концентраций загрязнителя в соответствующих прогнозам ячейках модельной сетки.

2. Разработаны модели машинного обучения для постпроцессинга прогнозов химической транспортной модели CHIMERE приземных концентраций озона (ММО-О<sub>3</sub>) и твердых взвешенных частиц диаметром менее 10 мкм (ММО-PM<sub>10</sub>) на территории московского региона. Для обучения ММО-О<sub>3</sub> и ММО-PM<sub>10</sub> использованы архивные часовые прогнозы на первые сутки концентраций 3В ХТМ CHIMERE и метеопараметров мезомасштабной системы ЧКПП COSMO-Ru в узлах сетки с горизонтальным шагом 2 км × 2 км для территории московского региона за 5-летний период февраль 2019 – февраль 2024. Также использованы данные доли городской застройки в ячейках той же модельной сетки. В качестве целевой переменной прогноза ММО-О<sub>3</sub> и ММО-PM<sub>10</sub> заданы архивные часовые измерения концентраций О<sub>3</sub> и PM<sub>10</sub> на всех станциях на территории московского региона.

3. Верификация разработанных ММО-О<sub>3</sub> и ММО-PM<sub>10</sub> с использованием рядов измерений концентраций на тестовой независимой выборке, охватывающей период март-сентябрь 2024 г., показала успешность применения обеих ММО в виде снижения отклонений ММО-скорректированных прогнозов О<sub>3</sub> и PM<sub>10</sub> от измерений по сравнению с прогнозами ХТМ CHIMERE:

– с применением ММО-О<sub>3</sub> СКОп-и прогнозов ХТМ суточного максимума скользящей средней 8-ми часовой концентрации О<sub>3</sub> (О<sub>3-8ч</sub>) снижено по всей тестовой выборке, включающей измерения 19 АСКЗА в московском регионе, с 27 мкг/м<sup>3</sup> (ХТМ) до 19 мкг/м<sup>3</sup> (ММО-О<sub>3</sub>) – на 30 %, на 12–43 % в отдельные месяцы тестовой выборки. Улучшено прогнозирование суточного хода концентраций О<sub>3</sub>, коэффициент корреляции между часовыми прогнозами и измерениями О<sub>3</sub> увеличен с 0.49 (ХТМ) до 0.81 (ММО-О<sub>3</sub>) по всей тестовой выборке, с ~0.4–0.5 (ХТМ) до ~0.7–0.9 (ММО-О<sub>3</sub>) в отдельные месяцы тестовой выборки;

– с применением ММО-PM<sub>10</sub> СКО<sub>п-и</sub> прогнозов ХТМ среднесуточной концентрации PM<sub>10-сс</sub> снижено по всей тестовой выборке, включающей измерения 23 АСКЗА, с 25 мкг/м<sup>3</sup> (ХТМ) до 15 мкг/м<sup>3</sup> (ММО-PM<sub>10</sub>) – на 40 %, на 25–62 % в отдельные месяцы тестовой выборки. ММО-PM<sub>10</sub> существенно корректирует суточный ход концентраций PM<sub>10</sub>, прогнозируемый ХТМ CHIMERE, увеличивая коэффициент корреляции между часовыми прогнозами и измерениями PM<sub>10</sub> с 0.14 (ХТМ) до 0.5 (ММО-PM<sub>10</sub>) по всей тестовой выборке, с ~0.1–0.2 (ХТМ) до ~0.4–0.6 (ММО-PM<sub>10</sub>) в отдельные месяцы тестовой выборки.

4. Анализ эффектов ММО-постпроцессинга прогнозов O<sub>3</sub> и PM<sub>10</sub> на отдельных АСКЗА, проведенный с использованием метода исключения АСКЗА из обучающей выборки ММО для имитации эффектов ММО-постпроцессинга на территориях без измерений концентраций, позволил выявить эффективность применения и ММО-O<sub>3</sub>, и ММО-PM<sub>10</sub> на большинстве АСКЗА:

– минимальная или отсутствующая эффективность применения ММО-O<sub>3</sub> выявлена на некоторых фоновых АСКЗА. На всех городских АСКЗА, вне зависимости от территориальной принадлежности и уровня загрязнения, применение ММО-O<sub>3</sub> обеспечивает снижение СКО<sub>п-и</sub> O<sub>3-8ч</sub> по всей тестовой выборке с 25–30 мкг/м<sup>3</sup> (ХТМ) до 15–25 мкг/м<sup>3</sup> – на ~30 %;

– минимальная или отсутствующая эффективность применения ММО-PM<sub>10</sub> наблюдается на нескольких АСКЗА с аномальными измеренными уровнями PM<sub>10</sub>, расположенных в непосредственной близости или на удалении от крупных источников загрязнения. На большинстве городских АСКЗА СКО<sub>п-и</sub> PM<sub>10-сс</sub> по всей тестовой выборке снижено с применением ММО-PM<sub>10</sub> в 1.5–2 раза по сравнению с СКО<sub>п-и</sub> PM<sub>10-сс</sub> (ХТМ) – до 10–15 мкг/м<sup>3</sup>.

5. Проведено целевое изучение эпизодов повышенного аэрозольного и озонowego загрязнения в московском регионе в период февраль 2019 – сентябрь 2024 гг. (почти 6 лет). На основании проведенного анализа 12 эпизодов повышенного аэрозольного загрязнения в московском регионе установлено, что значительное повышение уровня PM<sub>10</sub> в московском регионе наблюдается при дальнем переносе частиц и при неблагоприятных для очищения воздуха метеорологических условиях (НМУ). Отличительными особенностями эпизодов дальнего переноса, наблюдаемых в основном весной и осенью, являются:

– быстрое перемещение загрязненного PM<sub>10</sub> воздуха из районов пыльных бурь в Прикаспии и Северной Африке; из районов восточной Европы в период пахотных работ; из районов лесных пожаров (пирогенные аэрозоли);



– синхронное повышение  $PM_{10}$  на большинстве АСКЗА (до 200–250  $\mu\text{г}/\text{м}^3$  и выше в отдельных эпизодах).

Аэрозольное загрязнение при НМУ характеризуется повышением концентраций  $PM_{10}$  при теплой, сухой, безветренной погоде вечером-ночью в условиях термической устойчивости в приземном слое, днем концентрации снижаются практически до фоновых значений за счет интенсивного конвективного перемешивания. При дальнейшем переносе частиц аэрозольное загрязнение может наблюдаться и в дневное время.

Повышенное озоновое загрязнение формируется при НМУ в малоградиентных барических полях, когда ночью происходит накопление прекурсоров озона в инверсионном слое, а днем при ясной, теплой, сухой и безветренной погоде  $O_3$  образуется фотохимическим путем. Также нельзя не учитывать основной природный источник – приток  $O_3$  из верхних слоев атмосферы при усилении конвективного перемешивания –, и возможную адвекцию насыщенных  $O_3$  воздушных масс.

6. Проведен анализ эффективности применения ММО- $O_3$  и ММО- $PM_{10}$  в эпизодах повышенного озонового и аэрозольного загрязнения в период тестовой независимой выборки март-сентябрь 2024 г.:

– ММО- $O_3$  в озоновых эпизодах при НМУ увеличивает заниженные прогнозы ХТМ, снижая отклонение прогноза в 2 раза в дни с наибольшими концентрациями  $O_3$ . Отсутствие положительных эффектов ММО-постпроцессинга в июльских эпизодах, предположительно, связано с аномально высоким, нехарактерным для сезона, озоновым загрязнением – среднемесячная измеренная концентрация  $O_3$  в июле 2024 г. на 40 % превысила среднюю многолетнюю величину;

– ММО- $PM_{10}$  успешно корректирует прогнозы ХТМ в аэрозольных эпизодах при НМУ, приближая концентрации к измеренным и прогнозируя превышения гигиенического норматива. ММО- $PM_{10}$  недооценивает повышение концентраций  $PM_{10}$ , обусловленное переносом частиц из удаленных районов, поскольку не содержит предикторов, которые бы отражали такие процессы.

7. Обоснованное успешными результатами верификации применение ММО- $O_3$  и ММО- $PM_{10}$  для постпроцессинга часовых полей прогнозов концентраций ХТМ CHIMERE на первые сутки на сетке с горизонтальным шагом 2 км  $\times$  2 км для территории московского региона, с усвоением ММО прогностических предикторов из всех узлов сетки, показало:

– ММО- $O_3$  корректирует прогностические поля  $O_3$  ХТМ CHIMERE с учетом погодных условий, приближая концентрации к измеренным. ММО- $O_3$  понижает прогнозы ХТМ по всему полю в ночное время и повышает концентрации днем, при этом распределяет уровни озонового загрязнения в соответствии с уровнями антропогенной нагрузки;

– ММО-PM<sub>10</sub> также успешно корректирует прогностические поля PM<sub>10</sub> ХТМ CHIMERE с учетом погодных условий, времени суток и распределения городской застройки по территории.

Разработанные модели машинного обучения могут применяться для повышения качества производимых в Гидрометцентре России ежедневных прогнозов ХТМ CHIMERE концентраций O<sub>3</sub> и PM<sub>10</sub> на территории московского региона.

Все поставленные задачи выполнены, цель исследования достигнута.

## СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

АПС	Атмосферный пограничный слой
АСКЗА	автоматическая станция контроля загрязнения атмосферного воздуха (экологический пост)
ЕТР	Европейская территория России
ЗВ	загрязняющие вещества
ИНС	искусственные нейронные сети
ЛОС	Летучие органические соединения
ММО	Модель машинного обучения
ММО- постпроцессинг	Постпроцессинг численных прогнозов приземных концентраций загрязняющего вещества с применением модели машинного обучения
ММО-О <sub>3</sub>	Модель машинного обучения для постпроцессинга численных прогнозов концентраций приземного озона
ММО-PM <sub>10</sub>	Модель машинного обучения для постпроцессинга численных прогнозов приземных концентраций взвешенных частиц PM <sub>10</sub>
НМУ	Неблагоприятные для рассеивания загрязнений метеорологические условия
ОСО	Общее содержание озона в атмосферном столбе
ПДК	Предельно допустимая концентрация загрязняющего вещества в приземном воздухе (ПДК <sub>м.р.</sub> – максимальная разовая предельно допустимая концентрация)
СКО <sub>п-и</sub>	Среднеквадратическое отклонение прогнозов от измерений
СО <sub>п-и</sub>	Среднее отклонение прогнозов от измерений
ХТМ	Химическая транспортная модель
ЦФО	Центральный федеральный округ
IQR	Межквартильный размах распределения (англ. interquartile range)
О <sub>3-8ч</sub>	Суточный максимум средних за 8 часов концентраций приземного озона
PM <sub>10-сс</sub>	Средняя за сутки концентрация PM <sub>10</sub>
Q1	Первый квартиль распределения (25%)
Q3	Третий квартиль распределения (75%)

$r$	Коэффициент корреляции Пирсона
$\Delta_{п-и}$	Отклонение прогноза от измерения

## СПИСОК ЛИТЕРАТУРЫ

1. Белан, Б.Д. Озон в тропосфере / Б.Д. Белан; под ред. В.А. Погодаева. – Томск: Изд-во Института оптики атмосферы СО РАН, 2010. – 488 с.
2. Борисов, Д.В. Изменения кадастровых данных о выбросах в атмосферу загрязняющих веществ в московском регионе / Д.В. Борисов, И.Н. Кузнецова, М.И. Нахаев // Гидрометеорологические исследования и прогнозы. – 2023. – № 2 (388). – С. 156–173.
3. Борисов, Д.В. Интегрирование химической транспортной модели и искусственной нейронной сети для прогноза концентрации PM10 / Д.В. Борисов, И.Н. Кузнецова // Гидрометеорологические исследования и прогнозы. – 2024. – № 3 (393). – С. 42–63.
4. Борисов, Д.В. Исследование сезонной и суточной изменчивости концентраций приземного озона / Д.В. Борисов, И.Ю. Шалыгина, Е.А. Лезина // Гидрометеорологические исследования и прогнозы. – 2020. – № 3 (377). – С. 122–135.
5. Борисов, Д.В. Постпроцессинг численных прогнозов концентраций приземного озона с использованием машинного обучения / Д.В. Борисов, И.Н. Кузнецова // Гидрометеорологические исследования и прогнозы. – 2023. – № 4 (390). – С. 86–104.
6. Борисов, Д.В. Уточнение данных о землепользовании для расчетов эмиссий в химической транспортной модели CHIMERE на примере нижегородского региона / Д.В. Борисов, И.Ю. Шалыгина // Гидрометеорологические исследования и прогнозы. – 2021. – № 3 (381). – С. 150–161.
7. Быков, Ф.Л. Постпроцессинг численных прогнозов приземных метеорологических параметров на основе нейросетевых методов: дис. ... канд. физ.-мат. наук: 25.00.29 / Быков Филипп Леонидович. – М., 2022. – 244 с.
8. Вейдман С. Глубокое обучение. Легкая разработка проектов на Python / С. Вейдман. – СПб.: Питер. – 2021. – 272 с.
9. Загрязнение воздуха на Европейской части России и Украине в условиях жаркого лета 2010 г. / А.М. Звягинцев, О.Б. Блюм, А.А. Глазкова [и др.] // Известия РАН. Физика атмосферы и океана. – 2011. – Т. 47. – № 6. – С. 757–766.
10. Закономерности изменчивости концентраций малых газовых составляющих в приземном воздухе г. Москвы / Н.Ф. Еланский, М.А. Локощенко, И.Б. Беликов [и др.] // Известия РАН. Физика атмосферы и океана. – 2007. – Т. 43. – № 2. – С. 219–231.
11. Захваткина, Н.Ю. Использование метода нейронной сети для выделения разрывов в ледяном покрове арктических морей по радиолокационным данным /

Н.Ю. Захваткина, И.А. Бычкова, В.Г. Смирнов // Метеорология и гидрология. – 2024. – № 4. – С. 91–103.

12. Иванова, Н.С. Содержание озона над территорией Российской Федерации во втором квартале 2024 г. / Н.С. Иванова, И.Н. Кузнецова, Е.А. Лезина // Метеорология и гидрология. – 2024. – № 9. – С. 136–141.

13. Иванова, Н.С. Содержание озона над территорией Российской Федерации в третьем квартале 2024 г. / Н.С. Иванова, И.Н. Кузнецова, Е.А. Лезина // Метеорология и гидрология. – 2025. – № 1. – С. 136–141.

14. Иванова, Н.С. Содержание озона над территорией Российской Федерации в 2023 г. / Н.С. Иванова, И.Н. Кузнецова, Е.А. Лезина // Метеорология и гидрология. – 2024. – № 3. – С. 135–142.

15. Иванова, Н.С. Содержание озона над территорией Российской Федерации в 2024 г. / Н.С. Иванова, И.Н. Кузнецова, Е.А. Лезина // Метеорология и гидрология. – 2025. – № 3. – С. 136–141.

16. Концентрация тропосферного озона на территории России в 2022 г. / В.В. Андреев, М.Ю. Аршинов, Б.Д. Белан [и др.] // Оптика атмосферы и океана. – 2023. – Т. 36. – № 8. – С. 642–655.

17. Кузнецова, И.Н. Методы прогнозирования метеорологических условий, влияющих на загрязнение приземного воздуха / И.Н. Кузнецова, Ю.В. Ткачева, Д.В. Борисов // Метеорология и гидрология. – 2024. – № 8. – С. 87–103.

18. Мешалкин, В.П. Нейронные сети в автоматизированной системе мониторинга состояния окружающей среды / В.П. Мешалкин, В.М. Панарин, А.А. Маслова // Sciences of Europe. – 2020. – № 50-2 (50). – С. 14–20.

19. Моделирование загрязнения приземного воздуха с характерными в период COVID-19 сокращениями эмиссий в атмосферу с использованием моделей CHIMERE и COSMO-ART / И.Н. Кузнецова, Г.С. Ривин, Д.В. Борисов [и др.] // Метеорология и Гидрология. – 2022 – № 3. – С. 25–35.

20. Оценка содержания озона в атмосфере на станции Листвянка Байкальской природной территории по наземным данным и по данным реанализа / А.М. Сметанина, С.А. Громов, В.А. Оболкин [и др.] // Метеорология и гидрология. – 2024. – № 4. – С. 126–132.

21. Панасенко, Е.А. Численное решение некоторых обратных задач с различными типами источников атмосферного загрязнения / Е.А. Панасенко, А.В. Старченко // Вестник Томского государственного университета. Математика и механика. – 2008. – № 2 (3). – С. 47–55.

22. Причины и факторы положительных аномалий приземной концентрации озона в Московском регионе и на юго-восточном побережье Крыма / А.М. Звягинцев, И.Н. Кузнецова, И.Ю. Шалыгина [и др.] // Оптика атмосферы и океана. – 2016. – Т. 29. – № 6. – С. 493–502.
23. Прогноз перемещения и трансформации загрязняющих веществ в атмосфере с помощью модели COSMO-ART / Р.М. Вильфанд, А.А. Кирсанов, А.П. Ревокатова [и др.] // Метеорология и гидрология. – 2017. – № 5. – С. 31–40.
24. Пространственно-временные вариации содержания загрязняющих примесей в воздушном бассейне Москвы и их эмиссии / Н.Ф. Еланский, А.В. Шилкин, Н.А. Пономарев [и др.] // Известия РАН. Физика атмосферы и океана. – 2022. – Т. 58. – № 1. – С. 92–108.
25. Пылевой аэрозоль из районов Северного Прикаспия в приземном воздухе центра Европейской России / А.А. Виноградова, Д.П. Губанова, Е.А. Лезина, Ю.А. Иванова // Оптика атмосферы и океана. – 2024. – Т. 37. – № 6 (425). – С. 453–460.
26. Система COSMO-Ru негидростатического мезомасштабного краткосрочного прогноза погоды Гидрометцентра России: второй этап реализации и развития / Г.С. Ривин, И.А. Розинкина, Р.М. Вильфанд [и др.] // Метеорология и гидрология. – 2015. – № 6. – С. 58–70.
27. Система краткосрочного численного прогноза высокой детализации COSMO-Ru, ее развитие и приложения / Г.С. Ривин, И.А. Розинкина, Е.Д. Астахова [и др.] // Гидрометеорологические исследования и прогнозы. – 2019. – № 4 (374). – С. 37–53.
28. Система прогнозирования качества воздуха на основе химических транспортных моделей / И.Н. Кузнецова, И.Ю. Шалыгина, М.И. Нахаев [и др.] // Гидрометеорологические исследования и прогнозы. – 2019. – № 4 (374). – С. 203–218.
29. Солдатенко, С.А. Искусственный интеллект и его применение в задачах численного прогноза погоды / С.А. Солдатенко // Метеорология и гидрология. – 2024. – № 4. – С. 5–27.
30. Сравнение различных методов статистического прогнозирования суточной динамики приземной концентрации озона / П.Н. Антохин, Б.Д. Белан, Д.Е. Савкин, Г.Н. Толмачев // Оптика атмосферы и океана. – 2013. – Т. 26. – № 12. – С. 1082–1089.
31. Статистическое моделирование максимальных суточных концентраций приземного озона / А.М. Звягинцев, И.Б. Беликов, Н.Ф. Еланский [и др.] // Оптика атмосферы и океана. – 2010. – Т. 23. – № 2. – С. 127–135.
32. Тестирование и перспективы технологии прогнозирования загрязнения воздуха с применением химических транспортных моделей CHIMERE и COSMO-Ru2ART /

И.Н. Кузнецова, М.И. Нахаев, А.А. Кирсанов [и др.] // Гидрометеорологические исследования и прогнозы. – 2022. – № 4 (386). – С. 147–170.

33. Уловимые частицы. Как нейросети помогают предотвращать последствия пеплопадов [Электронный ресурс] // Яндекс.Облако: [сайт]. – Режим доступа: <https://yandex.cloud/ru/special/kamchatka> (дата обращения: 26.01.2025).

34. Хайкин, С. Нейронные сети: полный курс / С. Хайкин; пер. с англ. Н.Н. Куссуль, А.Ю. Шелестова. – 2-е изд. испр. – М.: Вильямс, 2006. – 1104 с.

35. Чурсин, В.В. Применение глубинных нейронных сетей для обнаружения вероятных зон атмосферных осадков и гроз / В.В. Чурсин, А.А. Косторная // Метеорология и гидрология. – 2024. – № 4. – С. 55–66.

36. Шолле, Ф. Глубокое обучение на Python / Ф. Шолле. – СПб.: Питер, 2018. – 400 с.

37. Эффективность коррекции эмиссий для расчетов химической транспортной модели CHIMERE в московском регионе / И.Ю. Шалыгина, И.Н. Кузнецова, М.И. Нахаев [и др.] // Оптика атмосферы и океана. – 2020. – Т. 33. – № 6. – С. 441–447.

38. A dispersion modelling system SILAM and its evaluation against ETEX data / M. Sofiev, P. Siljamo, I. Valkama [et al.] // Atmospheric Environment. – 2006. – Vol. 40. – № 4. – P. 674–685.

39. A European open access chemical weather forecasting portal / T. Balk, J. Kukkonen, K. Karatzas [et al.] // Atmospheric Environment. – 2011. – Vol. 45. – № 38. – P. 6917–6922.

40. A Kalman-filter bias correction method applied to deterministic, ensemble averaged, and probabilistic forecasts of surface ozone / L. Delle Monache, J.M. Wilczak, S.A. Mckeen [et al.] // Tellus, Series B: Chemical and Physical Meteorology. – 2008. – Vol. 60. – № 2. – P. 238–249.

41. A real-time Eulerian photochemical model forecast system: Overview and initial ozone forecast performance in the Northeast U.S. corridor / J.N. McHenry, W.F. Ryan, N.L. Seaman [et al.] // Bulletin of the American Meteorological Society. – 2004. – Vol. 85. – № 4. – P. 525–548.

42. Advanced air quality forecasting system for Chiba Prefecture / T. Ohara, A. Fujita, T. Kizu, S. Okamoto // Proceedings of the 22nd NATO/CCMS International Technical Meeting on Air Pollution Modeling and its Application, Clermont- Ferrand, France, 2-6 June 1997. – New York: Springer, 1998. – P. 625–628.

43. Air quality forecasting system for Southeastern Brazil / M.F. Andrade, R.Y. Ynoue, E.D. Freitas [et al.] // Frontiers in Environmental Science. – 2015. – Vol. 3. – P. 1–14.



44. Air quality simulation with WRF-Chem over southeastern Brazil, part I: Model description and evaluation using ground-based and satellite data / N.R. Benavente, A.L. Varavela, J.P. Nascimento [et al.] // *Urban Climate*. – 2023. – Vol. 52.

45. An Air Pollutants Prediction Method Integrating Numerical Models and Artificial Intelligence Models Targeting the Area around Busan Port in Korea / H. Hong, I. Choi, H. Jeon [et al.] // *Atmosphere*. – 2022. – Vol. 13. – № 9. – Art. 1462.

46. An improved version of the Piecewise Parabolic Method advection scheme: description and performance assessment in a bidimensional testcase with stiff chemistry in toyCTM v1. 0 / S. Mailler, R. Pennel, L. Menut, A. Cholakian // *Geoscientific Model Development*. – 2023. – Vol. 16. – № 24. – P. 7509–7526.

47. Application of a Multiscale, Coupled MM5/Chemistry Model to the Complex Terrain of the VOTALP Valley Campaign / G. Grell, S. Emeis, W.R. Stockwell [et al.] // *Atmospheric Environment*. – 2000. – Vol. 34. – № 9. – P. 1435–1453.

48. Application of the Kolmogorov-Zurbenko filter and the decoupled direct 3D method for the dynamic evaluation of a regional air quality model / D. Kang, C. Hogrefe, K. Foley [et al.] // *Atmospheric Environment*. – 2013. – Vol. 80. – P. 58–69.

49. Aron, R.H. Forecasting high level oxidant concentrations in the Los Angeles basin / R.H. Aron // *Journal of the Air Pollution Control Association*. – 1980. – Vol. 20. – P. 1227–1228.

50. Assessment of an ensemble of seven real-time ozone forecasts over eastern North America during the summer of 2004 / S. McKeen, J. Wilczak, G. Grell [et al.] // *Journal of Geophysical Research Atmospheres*. – 2005. – Vol. 112. – Art. D21307.

51. Bias adjustment techniques for improving ozone air quality forecasts / D. Kang, R. Mathur, S.T. Rao, S. Yu // *Journal of Geophysical Research*. – 2007. – Vol. 113. – № 23. – Art. D23308.

52. Bosanquet, C.H. The spread of smoke and gases from chimneys / C.H. Bosanquet, J.L. Pearson // *Transactions of the Faraday Society*. – 1936. – Vol. 32. – P. 1249–1263.

53. Chemistry-Weather Interacted Model System GRAPES\_Meso5.1/CUACE CW V1.0: Development, evaluation and application in better haze/fog prediction in China / H. Wang, X.Y. Zhang, P. Wang [et al.] // *Journal of Advances in Modeling Earth Systems*. – 2022. – Vol. 14. – № 12. – P. 1–25.

54. CMAQ-CNN: A new-generation of post-processing techniques for chemical transport models using deep neural networks / A. Sayeed, E. Eslami, Y. Lops, Y. Choi // *Atmospheric Environment*. – 2022. – Vol. 273. – Art. 118961.

55. Columnar and surface urban aerosol in the Moscow megacity according to measurements and simulations with the COSMO-ART model / N. Chubarova, H. Vogel, E. Androsova [et al.] // *Atmospheric Chemistry and Physics*. – 2022. – Vol. 22. – P. 10443–10466.
56. Combining deterministic and statistical approaches for PM10 forecasting in Europe / I.B. Konovalov, M. Beekmann, F. Meleux [et al.] // *Atmospheric Environment*. – 2010. – Vol. 43. – № 40. – P. 6425–6434.
57. Data Augmentation for Regression Machine Learning Problems in High Dimensions / C. Guilhaumon, N. Hascoët, F. Chinesta [et al.] // *Computation*. – 2024. – Vol. 12. – № 2. – P. 1–17.
58. Day-night variability of PM10 components at a Mediterranean urban site during winter / N. Galindo, E. Yubero, J.F. Nicolás [et al.] // *Air Quality, Atmosphere and Health*. – 2018. – Vol. 11. – № 10. – P. 1251–1258.
59. Deep Learning to Evaluate US NO<sub>x</sub> Emissions Using Surface Ozone Predictions / T. He, D.B.A. Jones, K. Miyazaki [et al.] // *Journal of Geophysical Research: Atmospheres*. – 2022. – Vol. 127. – № 4. – Art. e2021JD035597.
60. Development and evaluation of an advanced National Air Quality Forecasting Capability using the NOAA Global Forecast System version 16 / P.C. Campbell, Y. Tang, P. Lee [et al.] // *Geoscientific Model Development*. – 2022. – Vol. 15. – № 8. – P. 3281–3313.
61. Development of WRF/CUACE v1.0 model and its preliminary application in simulating air quality in China / L. Zhang, S. Gong, T. Zhao [et al.] // *Geoscientific Model Development*. – 2021. – Vol. 14. – № 2. – P. 703–718.
62. Development towards a global operational aerosol consensus: basic climatological characteristics of the International Cooperative for Aerosol Prediction Multi-Model Ensemble (ICAP-MME) / W.R. Sessions, J.S. Reid, A. Benedetti [et al.] // *Atmospheric Chemistry and Physics*. – 2015. – Vol. 15. – P. 335–362.
63. Dropout: A Simple Way to Prevent Neural Networks from Overfitting / N. Srivastava, G. Hinton, A. Krizhevsky [et al.] // *Journal of Machine Learning Research*. – 2014. – Vol. 15. – № 1. – P. 1929–1958.
64. Effect of Natural Forest Fires on Regional Weather Conditions in Siberia / A. Kirsanov, I. Rozinkina, G. Rivin [et al.] // *Atmosphere*. – 2020. – Vol. 11. – № 10. – Art. 1133.
65. Eliassen, A. Decay and transformation rates of SO<sub>2</sub>, as estimated from emission data, trajectories and measured air concentrations / A. Eliassen, J. Saltbones // *Atmospheric Environment*. – 1975. – Vol. 9. – № 4. – P. 425–429.

66. Emission factors for gas-powered vehicles traveling through road tunnels in São Paulo, Brazil / L.D. Martins, M.F. Andrade, E.D. Freitas [et al.] // *Environmental Science and Technology*. – 2006. – Vol. 40. – № 21. – P. 6722–6729.

67. Ensemble forecasts of air quality in eastern China e Part 2. Evaluation of the MarcoPolo-panda prediction system, version 1 / A.K. Petersen, G.P. Brasseur, I. Bouarar [et al.] // *Geoscientific Model Development*. – 2019. – Vol. 12. – № 3. – P. 1241–1266.

68. Estimating Daily PM<sub>2.5</sub> and PM<sub>10</sub> over Italy Using an Ensemble Model / A. Shtein, I. Kloog, J. Schwartz [et al.] // *Environmental Science & Technology*. – 2020. – Vol. 54. – № 1. – P. 120–128.

69. Evaluation of a CFD-based approach to estimate pollutant distribution within a real urban canopy by means of passive samplers / J.L. Santiago, R. Borge, F. Martin [et al.] // *Science of the Total Environment*. – 2017. – Vol. 576. – № 3. – P. 46–58.

70. Evaluation of several PM<sub>2.5</sub> forecast models using data collected during the ICARTT/NEAQS 2004 field study / S. McKeen, S.H. Chung, J. Wilczak [et al.] // *Journal of Geophysical Research Atmospheres*. – 2007. – Vol. 112. – Art. D10S20.

71. Evolution of Vehicle Emission Factors in a Megacity Affected by Extensive Biofuel Use: Results of Tunnel Measurements in São Paulo, Brazil / T. Nogueira, L. Kamigauti, G.M. Pereira [et al.] // *Environmental Science and Technology*. – 2021. – Vol. 55. – № 10. – P. 6677–6687.

72. Fisher, B.E.A. The long-range transport of sulphur dioxide / B.E.A. Fisher // *Atmospheric Environment*. – 1975. – Vol. 9. – № 12. – P. 1063–1070.

73. Fully coupled “online” chemistry within the WRF model / G.A. Grell, S.E. Peckham, R. Schmitz [et al.] // *Atmospheric Environment*. – 2005. – Vol. 39. – № 37. – P. 6957–6975.

74. Geographic sources of ozone air pollution and mortality burden in Europe / H. Achebak, R. Garatachea, M.T. Pay [et al.] // *Nature Medicine*. – 2024. – Vol. 30. – № 6. – P. 1732–1738.

75. Global Atmosphere Watch Expert Meeting on Measurement-Model Fusion for Global Total Atmospheric Deposition (MMF-GTAD) [Электронный ресурс] // World Meteorological Organization: [сайт]. – 2020. – Режим доступа: <https://library.wmo.int/records/item/57018-global-atmosphere-watch-expert-meeting-on-measurement-model-fusion-for-global-total-atmospheric-deposition-mmf-gtad> (дата обращения: 22.02.2025).

76. Global Atmosphere Watch Workshop on Measurement-Model Fusion for Global Total Atmospheric Deposition (MMF-GTAD) [Электронный ресурс] // World Meteorological Organization: [сайт]. – 2017. – Режим доступа: <https://library.wmo.int/records/item/55559->

global-atmosphere-watch-workshop-on-measurement-model-fusion-for-global-total-atmospheric-deposition-mmfd-gtad (дата обращения: 22.02.2025).

77. Gubanova, D. Long-Range Atmospheric Transport of Dust from the Caspian Sea Region to the Arctic Zone of the European Part of Russia in December 2023 / D. Gubanova, A. Vinogradova, E. Kotova // *Doklady Earth Sciences*. – 2024. – Vol. 519. – № 2. – P. 1–8.

78. Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide. Report on a WHO Working Group (Bonn, Germany 13–15 January 2003) [Электронный ресурс] // WHO Regional Office for Europe: [сайт]. – 2003. – Режим доступа: <https://iris.who.int/bitstream/10665/107478/1/E79097.pdf> (дата обращения: 05.02.2025).

79. Holzworth, G.C. Mixing Depths, Wind Speeds and Air Pollution Potential for Selected Locations in the United States1 / G.C. Holzworth // *Journal of Applied Meteorology* (1962-1982). – 1967. – Vol. 6. – № 6. – P. 1039–1044.

80. Improving 3-day deterministic air pollution forecasts using machine learning algorithms / Z. Zhang, C. Johansson, M. Engardt [et al.] // *Atmospheric Chemistry and Physics*. – 2024. – Vol. 24. – № 2. – P. 807–851.

81. Improving NOAA NAQFC PM2.5 predictions with a bias correction approach / J. Huang, J. McQueen, J. Wilczak [et al.] // *Weather and Forecasting*. – 2016. – Vol. 32. – № 2. – P. 407–421.

82. Isaksen, I.S.A. A two-dimensional model for the global distribution of gases and aerosol particles in the troposphere. Rep. AC-47 / I.S.A. Isaksen, H. Rohde. – Stockholm: International Meteorological Institute in Stockholm, 1978. – 65 p.

83. Kang, D. Assessment of bias-adjusted PM2.5 air quality forecasts over the continental United States during 2007 / D. Kang, R. Mathur, S.T. Rao // *Geoscientific Model Development*. – 2010b. – Vol. 2. – № 2. – P. 309–320.

84. Kang, D. Real-time Bias-adjusted O3 and PM2.5 air quality index forecasts and their evaluations over the continental United States / D. Kang, R. Mathur, S.T. Rao // *Atmospheric Environment*. – 2010a. – Vol. 44. – № 18. – P. 2203–2212.

85. Krinitskiy, M.A. An overview of machine learning and deep learning applications in Earth sciences in 2024: achievements and perspectives / M.A. Krinitskiy // *Moscow University Physics Bulletin*. – 2024. – Vol. 79. – № S1. – P. 746 - 756.

86. Leufen, L.H. MLAir (v1.0) – a tool to enable fast and flexible machine learning on air data time series / L.H. Leufen, F. Kleinert, M.G. Schultz // *Geoscientific Model Development*. – 2021. – Vol. 14. – № 3. – P. 1553–1574.

87. Limited-Area Atmospheric Modeling Using an Unstructured Mesh / W.C. Skamarock, M.G. Duda, S. Ha, S. Park // *Monthly Weather Review*. – 2018. – Vol. 146. – № 10. – P. 3445–3460.
88. Machine learning based bias correction for numerical chemical transport models / M. Xu, J. Jin, G. Wang [et al.] // *Atmospheric Environment*. – 2021. – Vol. 248. – № 23. – Art. 118022.
89. Masood, A. A review on emerging artificial intelligence (AI) techniques for air pollution forecasting: Fundamentals, application and performance / A. Masood, K. Ahmad // *Journal of Cleaner Production*. – 2021. – Vol. 322. – № 11. – Art. 129072.
90. McCollister, G. Linear stochastic models for forecasting daily maxima and hourly concentrations of air pollutants / G. McCollister, K. Wilson // *Atmospheric Environment*. – 1967. – Vol. 9. – № 4. – P. 417–423.
91. Middleton, N. Variability and Trends in Dust Storm Frequency on Decadal Timescales: Climatic Drivers and Human Impacts / N. Middleton // *Geosciences*. – 2019. – Vol. 9. – № 6. – Art. 261.
92. Multi-scale modelling of urban air pollution: development and application of a Street-in-Grid model (v1.0) by coupling MUNICH (v1.0) and Polair3D (v1.8.1) / Y. Kim, Y. Wu, C. Seigneur, Y. Roustan // *Geoscientific Model Development*. – 2018. – Vol. 11. – № 2. – P. 611–629.
93. New era of air quality monitoring from space: geostationary environment monitoring spectrometer (GEMS) / J. Kim, U. Jeong, M.H. Ahn [et al.] // *Bulletin of the American Meteorological Society*. – 2019. – Vol. 101. – № 1. – P. 1–22.
94. One decade of parallel fine (PM<sub>2.5</sub>) and coarse (PM<sub>10</sub>-PM<sub>2.5</sub>) particulate matter measurements in Europe: trends and variability / I. Barmpadimos, J. Keller, D. Oderbolz [et al.] // *Atmospheric Chemistry and Physics*. – 2012. – Vol. 12. – № 7. – P. 3189–3203.
95. Operational air quality forecasting in Canada: Numerical model guidance for ground-level ozone and particulate matter [Электронный ресурс] / J. Pudykiewicz, A. Kallaur, R. Moffet [et al.] // Preprints, Fifth Conf. on Atmospheric Chemistry: Gases, Aerosols, and Clouds, Long Beach, CA, Amer. Meteor. Soc., the 83rd Annual Meeting. – Режим доступа: <https://ams.confex.com/ams/pdfpapers/54490.pdf> (дата обращения: 18.02.2025).
96. Ozone ensemble forecasts: 1. A new ensemble design / L. Delle Monache, X. Deng, Y. Zhou, R. Stull // *Journal of Geophysical Research*. – 2006a. – Vol. 111. – № 5. – Art. D05307.
97. Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction / L. Delle Monache, T. Nipen, X. Deng [et al.] // *Journal of Geophysical Research*. – 2006b. – Vol. 111. – № 5. – Art. D05308.

98. Perez, P. An integrated neural network model for PM10 forecasting / P. Perez, J. Reyes // *Atmospheric Environment*. – 2006. – Vol. 40. – № 16. – P. 2845–2851.
99. Peters, L.K. Numerical simulation of the transport and chemistry of CH and CO in the troposphere / L.K. Peters, A.A. Jouvanis // *Atmospheric Environment*. – 1979. – Vol. 13. – № 10. – P. 1443–1462.
100. Potempski, S. Est modus in rebus: analytical properties of multi-model ensembles / S. Potempski, S. Galmarini // *Atmospheric Chemistry and Physics*. – 2009. – Vol. 9. – № 24. – P. 9471–9489.
101. Predictability of European air quality: Assessment of 3 years of operational forecasts and analyses by the PREV'AIR system / C. Honoré, L. Rouil, R. Vautard [et al.] // *Journal of Geophysical Research*. – 2008. – Vol. 113. – № 4. – Art. D04301.
102. Predicting air quality: improvements through advanced methods to integrate models and measurements / G.R. Carmichael, A. Sandu, T. Chai [et al.] // *Journal of Computational Physics*. – 2008. – Vol. 227. – № 7. – P. 3540–3571.
103. Real-Time Air Quality Forecasting, Part I: History, Techniques, and Current Status / Y. Zhang, C. Seigneur, M. Bocquet [et al.] // *Atmospheric Environment*. – 2012. – Vol. 60. – P. 632–655.
104. Representing chemical history in ozone time-series predictions - a model experiment study building on the MLAir (v1.5) deep learning framework / F. Kleinert, L. Leufen, A. Lupascu [et al.] // *Geoscientific Model Development*. – 2022. – Vol. 15. – № 23. – P. 8913–8930.
105. Reynolds, S. Mathematical Modeling of Photochemical Air Pollution / S. Reynolds, P.M. Roth, J.H. Seinfeld // *Atmospheric Environment*. – 1967. – Vol. 7. – № 11. – P. 1033–1061.
106. Rufeger, W. Applying the DYMOS system in conurbations / W. Rufeger, P. Mieth, T. Lux // *Proc. Modsim 97, 4, Hobart, TAS, Australia, International Congress on Modeling and Simulation*. – 1997. – Vol. 4. – 1797–1801.
107. Sak, H. Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition / H. Sak, A. Senior, F. Beaufays // *Neural and Evolutionary Computing*. – 2014.
108. Seasonal variability of PM2.5 and PM10 composition and sources in an urban background site in Southern Italy / D. Cesari, G.E. De Benedetto, P. Bonasoni [et al.] // *Science of the Total Environment*. – 2018. – Vol. 612. – P. 202–213.
109. Seigneur Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models / M. Bocquet, H. Elbern, H. Eskes [et al.] // *Atmospheric Chemistry and Physics*. – 2015. – Vol. 15. – P. 5325–5358.

110. Sensitivity of modeled sulphate and nitrate aerosol to cloud, pH and ammonia emissions / A.L. Redington, R.G. Derwent, C.S. Witham, A.J. Manning // *Atmospheric Environment*. – 2009. – Vol. 43. – № 20. – P. 3227–3234.

111. Shepherd, G.M. Introduction of synaptic circuits / G.M. Shepherd, C. Koch // *The Synoptic Organization of the Brain* / ed. G.M. Shepherd. – New York: Oxford University Press, 1990. – P. 3–31.

112. Srivastava, N. Impact of meteorological parameterization schemes on CTM model simulations / N. Srivastava, B. Nadèg // *Atmospheric Environment*. – 2021. – Vol. 268. – № 23. – Art. 118832.

113. Sutton, O.G. A theory of eddy diffusion in the atmosphere / O.G. Sutton // *Proceedings of the Royal Society of London*. – 1932. – Vol. 135. – № 826. – P. 143–165.

114. The Australian Air Quality Forecasting System. Part I: Project description and early outcomes / M.E. Cope, G.D. Hess, S. Lee [et al.] // *Journal of Applied Meteorology*. – 2004. – Vol. 43. – № 5. – P. 649–662.

115. The CHIMERE chemistry-transport model v2023r1 / L. Menut, A. Cholakian, R. Pennel [et al.] // *Geoscientific Model Development*. – 2024. – Vol. 17. – P. 5431–5457.

116. The comprehensive model system COSMO-ART– Radiative impact of aerosol on the state of the atmosphere on the regional scale / B. Vogel, H. Vogel, D. Bäumer [et al.] // *Atmospheric Chemistry and Physics*. – 2009. – Vol. 9. – № 22. – P. 8661–8680.

117. The extraction and fusion of meteorological and air quality information for orchestrated services / L.O. Johansson, V. Epitropou, K.D. Karatzas [et al.] // *Environmental Modelling and Software*. – 2015. – Vol. 64. – P. 143–155.

118. The impact of weather and atmospheric circulation on O<sub>3</sub> and PM<sub>10</sub> levels at a rural mid-latitude site / M. Demuzere, R. Trigo, J. Arellano, N. Lipzig // *Atmospheric Chemistry and Physics*. – 2009. – Vol. 9. – P. 2695–2714.

119. The Use of Modern Third-Generation Air Quality Models (MM5-EMIMO-CMAQ) for Real-Time Operational Air Quality Impact Assessment of Industrial Plants / R.S.J. García, J.L.P. Camaño, J.L. Morant, R.M.G. Barras // *Water, Air, & Soil Pollution: Focus*. – 2009. – Vol. 9. – № 1-2. – P. 27–37.

120. Toward a US National Air Quality Forecast Capability: Current and Planned Capabilities / P. Davidson, K. Schere, R. Draxler [et al.] // *Air Pollution Modeling and Its Application XIX. NATO Science for Peace and Security Series Series C: Environmental Security* / C. Borrego, A.I. Miranda (eds). – Dordrecht: Springer, 2008. – P. 226–234.

121. Training Materials and Best Practices for Chemical Weather/Air Quality Forecasting (ETR-26) [Электронный ресурс] // World Meteorological Organization: [сайт]. – 2020. – Режим

доступа: <https://library.wmo.int/records/item/54300-training-materials-and-best-practices-for-chemical-weather-air-quality-forecasting-etr-26> (дата обращения: 22.02.2025).

122. Turner, D.B. A Diffusion Model for an Urban Area [Электронный ресурс] / D.B. Turner // *Journal of Applied Meteorology* (1962-1982). – 1964. – Vol. 3. – № 1. – P. 83–91. – Режим доступа: <http://www.jstor.org/stable/26169533> (дата обращения: 15.02.2025).

123. Varentsov, M.I. Approximation of spatial and temporal variability of the urban heat island in Moscow using machine learning / M.I. Varentsov, M.A. Krinitskiy, V.M. Stepanenko // *Moscow University Physics Bulletin*. – 2024. – Vol. 79. – № S1. – P. 791-804.

124. WHO global air quality guidelines. Particulate matter (PM<sub>2.5</sub> and PM<sub>10</sub>), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide [Электронный ресурс] // World Health Organization: [сайт]. – 2021. – Режим доступа: <https://www.who.int/publications/i/item/9789240034228> (дата обращения: 22.02.2025).

125. Williams, R.W. The control of neuron number / R.W. Williams, K. Herrup // *Annual review of neuroscience*. – 1988. – Vol. 11. – № 1. – P. 423–453.

126. Wolf, T. A very high-resolution assessment and modelling of urban air quality / T. Wolf, L.H. Pettersson, I. Esau // *Atmospheric Chemistry and Physics*. – 2020. – Vol. 20. – № 2. – P. 625–647.

127. Wolff, G.T. An empirical model for forecasting maximum daily ozone levels in the north-eastern United States / G.T. Wolff, P.J. Lioy // *Journal of the Air Pollution Control Association*. – 1978. – Vol. 28. – P. 1034–1038.

128. Zhang, J. Ozone Pollution: A Major Health Hazard Worldwide / J. Zhang, Y. Wei, Z. Fang // *Frontiers in Immunology*. – 2019. – Vol. 10. – Art. 2518.

129. Zhang, Y. Online Coupled Meteorology and Chemistry models: History, Current Status, and Outlook / Y. Zhang // *Atmospheric Chemistry and Physics*. – 2008. – Vol. 8. – № 1. – P. 2895–2932.