

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертационную работу **Борисова Данила Владимировича**

“ПОСТПРОЦЕССИНГ ЧИСЛЕННЫХ ПРОГНОЗОВ КОНЦЕНТРАЦИЙ
ВЗВЕШЕННЫХ ЧАСТИЦ (PM₁₀) И ПРИЗЕМНОГО ОЗОНА (O₃) С
ИСПОЛЬЗОВАНИЕМ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ”,

представленную на соискание ученой степени кандидата географических наук

по специальности 1.6.18 – Науки об атмосфере и климате

Диссертационная работа Борисова Д.В. представляет собой логично выстроенное замкнутое исследование, продуманное и выполненное автором от начала и до конца. Оно посвящено изучению пространственной и временной изменчивости загрязнения приземного воздуха озоном (O₃) и аэрозольными частицами (PM₁₀) в Московском регионе. Акцент в этом исследовании ставится на современных методах численного прогнозирования загрязнения воздуха, их анализе и практической реализации. Демонстрируются результаты предложенного оригинального подхода к конструированию и реализации моделей машинного обучения (ММО) для повышения качества прогнозов загрязнения воздуха, рассчитанных с помощью химико-транспортных моделей (ХТМ).

Актуальность темы исследования обусловлена, с одной стороны, объектом исследований – это два загрязнителя воздуха O₃ и PM₁₀, которые, по данным ВОЗ, являются приоритетными при оценке качества воздуха с точки зрения рисков для здоровья человека. Поэтому заблаговременность и достоверность информации о повышении их содержания в приземном воздухе являются необходимыми задачами прогнозирования. С другой стороны, в работе предложен очень актуальный и современный подход к повышению достоверности прогнозов качества воздуха, основанный на анализе больших объемов метеорологической и экспериментальной информации. Предложенный автором работы подход и разработанные методика и модели постпроцессинга, с применением бурно развивающегося направления машинного обучения, позволяют совершенствовать результаты прогнозирования уровней концентрации O₃ и PM₁₀, получаемых с помощью ставших уже традиционными расчетов на базе ХТМ.

Научная новизна полученных результатов состоит в следующем: 1) разработана методика построения моделей машинного обучения для постпроцессинга численных прогнозов загрязнителей воздуха; 2) на основе этой методики разработаны и протестированы модели машинного обучения для постпроцессинга численных

прогнозов концентраций приземного озона и взвешенных частиц PM_{10} на территории Московского региона; 3) разработанные модели машинного обучения применимы для постпроцессинга численных прогнозов концентраций O_3 и PM_{10} на территориях, где не проводятся измерения их концентраций.

Обоснованность и достоверность результатов диссертационной работы базируются на длительном (5-летнем) объеме исходной обучающей выборки (для разработанных моделей машинного обучения) часовых прогнозов и измерений O_3 и PM_{10} , включающей различные метеоусловия формирования атмосферных загрязнений, в том числе периоды повышенных концентраций O_3 и PM_{10} при неблагоприятных метеорологических условиях (НМУ) для рассеивания примесей в атмосфере, и подтверждаются тестированием результатов постпроцессинга на независимых выборках аналогичных данных для последующих 7 месяцев. Эффективность применения ММО-постпроцессинга прогнозов концентраций O_3 и PM_{10} на территориях, где отсутствуют измерения концентраций, доказана с использованием метода исключения станции мониторинга из обучающей выборки ММО. Разработка ММО проводилась с использованием метода k-блоковой валидации моделей по обучающим выборкам, которые охватывают 5-летний период часовых прогнозов и измерений концентраций O_3 и PM_{10} .

Теоретическая и практическая значимость выводов и рекомендаций диссертанта определяется тем, что реализованные по разработанной методике модели машинного обучения позволяют улучшать качество прогнозов, рассчитанных химической транспортной моделью, приземных концентраций O_3 и PM_{10} , в том числе на территориях, где отсутствуют измерения концентраций. Разработанные модели машинного обучения могут применяться для оперативного постпроцессинга численных часовых прогнозов приземных концентраций O_3 и PM_{10} на модельной сетке с горизонтальным шагом $2 \text{ км} \times 2 \text{ км}$ для территории Московского региона. Созданный автором 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 ХТМ и непрерывными измерениями концентраций в конкретных пунктах.

Структура и основные результаты работы. 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.

Во введении обоснована актуальность работы, сформулированы цель и задачи исследования, приведены положения, выносимые на защиту, научная новизна, теоретическая и практическая значимость работы.

Конечной целью исследования названа разработка, на основе методов машинного обучения, инструмента усовершенствования прогнозов приземных концентраций озона и частиц PM_{10} , полученных с помощью химико-транспортной модели.

Для достижения цели автором были поставлены и решены восемь задач – от анализа имеющейся информации по теме работы (данных измерений приземных концентраций O_3 и PM_{10} на территории Московского региона, современных систем прогнозирования качества воздуха, методов повышения качества прогнозов, практик применения методов машинного обучения) до разработки моделей машинного обучения для постпроцессинга численных прогнозов приземных концентраций озона (ММО- O_3) и взвешенных частиц (ММО- PM_{10}) на территории Московского региона, их применения и верификации. В конце работы автор логично замыкает исследование, окончательно формулируя основные принципы построения ММО для постпроцессинга численных прогнозов концентраций загрязняющих веществ в задачах прогнозирования качества воздуха, которые используют регулярные пространственные сетки данных варьируемого разрешения, не имеют географической привязки и применимы для регионов, обеспеченных прогнозами ХТМ и непрерывными измерениями концентраций в конкретных пунктах.

В первой главе диссертации сжато изложена история развития методов прогнозирования качества воздуха, рассматриваются современные системы прогнозирования на основе химических транспортных моделей. Описаны несколько национальных и интернациональных систем на основе ХТМ с приведением оценок качества прогнозов концентраций O_3 и PM_{10} . Обсуждаются различные методы повышения качества прогнозов ХТМ и практики применения моделей машинного обучения для прогнозирования концентраций загрязняющих веществ. При сравнении методов машинного обучения делается вывод, что наиболее успешный прогноз концентраций загрязняющих веществ показывают искусственные нейронные сети – за счет воспроизведения сложных нелинейных связей между концентрацией атмосферной примеси и влияющими на нее факторами.

Во второй главе обсуждаются использованные в работе методы и данные. Автором был изучен и описан применяемый в работе метод искусственных нейронных сетей, его достоинства, ограничения и недостатки. Сформирована исследовательская база данных, включающая не только часовые измерения концентраций O_3 и PM_{10} на территории Московского региона и модельные прогнозы концентраций $3B$ и метеопараметров на территории Московского региона на сетке с горизонтальным шагом 2×2 км с временным шагом 1 час за почти 6-летний период (2019-2024), но и данные о распределении городской застройки по территории Московского региона на той же сетке с горизонтальным шагом 2×2 км.

Подробно проанализированы и описаны 7 эпизодов повышенного аэрозольного загрязнения в Московском регионе в условиях дальнего переноса и/или неблагоприятных метеорологических условий НМУ рассеивания примесей.

Анализируются использованные в работе архивные численные прогнозы метеопараметров и концентраций O_3 и PM_{10} на территории Московского региона за те же годы. Подробно разбирается и анализируется качество прогнозов с использованием ХТМ концентраций O_3 и PM_{10} на территории Московского региона за 2023 г.

Глава 3 посвящена разработанной автором методике построения моделей машинного обучения для постпроцессинга численных прогнозов приземных концентраций загрязняющих веществ. Здесь особый интерес и уважение вызывает поиск оптимальных конфигураций ММО для O_3 (полносвязная 4-слойная (по 7 нейронов в слое) искусственная нейронная сеть) и для PM_{10} – (полносвязная 3-слойная (по 50 нейронов в слое с исключениями) искусственная нейронная сеть). Обучение модели осуществляется на наборе численных часовых прогнозов концентраций и метеопараметров (всего 19 и 20 параметров), в качестве целевой переменной используется часовая концентрация загрязнителя в соответствующем прогнозам узле регулярной сетки. Важно, что разработанные таким образом ММО потенциально применимы для улучшения прогнозов концентраций во всем расчетном домене ХТМ, в том числе на территориях (в узлах регулярной сетки), где отсутствуют измерения концентраций.

В этой главе приведены также результаты ранжирования параметров обучающей выборки по их значимости для разработанных ММО, что очень интересно и важно для понимания процессов формирования полей озона и аэрозолей PM_{10} в приземной атмосфере. Для ММО- O_3 – это (в порядке убывания значимости) температура воздуха, час суток (суточный ход), номер месяца (сезонный ход), относительная влажность,

среднесуточное общее содержание озона и, наконец, прогноз ХТМ концентрации прекурсора озона NO_2 (антропогенный фактор). Для ММО- PM_{10} – это номер месяца (сезонный ход), приземная температура, градиент температуры в слое 2–400 м (термическая устойчивость в нижнем слое атмосферы), далее еще пара показателей стратификации нижней тропосферы и только шестым идет час суток (суточный ход), а за ним – доля городской застройки в ячейке сетки ХТМ (антропогенный фактор). Остальные параметры дают малые поправки в расчеты. Фактор количества осадков для обоих загрязнителей расположен на 18 месте по значимости.

В главе 4 приводятся результаты верификации разработанной ММО для озона на независимой выборке за период март-сентябрь 2024 г. (7 месяцев) на территории Московского региона. Делается вывод об эффективности разработанной ММО- O_3 для приближения рассчитанных значений концентрации O_3 к наблюдаемым в первую очередь при неблагоприятных метеорологических условиях (НМУ). Также показано, что применение ММО- O_3 обеспечивает существенную коррекцию рассчитанной в ХТМ суточной динамики O_3 , и улучшение ее пространственного воспроизведения по территории Московского региона, учитывая направление распространения примеси и распределение уровня антропогенной нагрузки.

В главе 5 приведены результаты верификации разработанной ММО для PM_{10} на независимой выборке за тот же период март-сентябрь 2024 г. на территории региона. Автором проведен многосторонний анализ эффектов постпроцессинга численных прогнозов концентраций с применением ММО- PM_{10} в целом и в отдельные месяцы тестовой выборки, приведены примеры прогностических полей концентраций O_3 и PM_{10} , рассчитанных ХТМ для всей территории Московского региона. Наиболее интересные результаты этой главы следующие. ММО- PM_{10} 1) снижает завышенные ХТМ дневные концентрации PM_{10} по всей территории региона; 2) в ночное время корректирует пространственную неоднородность загрязнения PM_{10} в соответствии с распределением городской застройки; 3) успешно корректирует прогнозы концентраций PM_{10} в периоды повышенного аэрозольного загрязнения при НМУ; 4) ожидаемо недооценивает аэрозольное загрязнение, обусловленное дальним переносом частиц из удаленных районов, поскольку в ММО- PM_{10} отсутствуют соответствующие предикторы; 5) минимальные эффекты постпроцессинга наблюдаются на АСКЗА, близко расположенных к локальному источнику загрязнения.

В заключении приведены основные результаты диссертационной работы, которые кратко перечислены и оценены выше.

Список литературы (129 наименований) корректно оформлен по ГОСТ РФ.

Проводимые автором исследования и результаты были использованы при выполнении НИТР плана Росгидромета 4.7.2 (2020–2024 гг.) и 4.9 (2025 г.). По теме диссертации опубликовано 28 работ, в том числе 10 статей в журналах, рекомендованных Высшей аттестационной комиссией, 10 публикаций в материалах международных конференций. Автором получены 4 свидетельства Роспатента о государственной регистрации программы для ЭВМ.

Положения, выносимые на защиту, хорошо сформулированы и развернуто подтверждены в тексте диссертации.

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.

Автореферат достаточно полно отражает содержание диссертационной работы.

Основные замечания и комментарии к работе:

Диссертационная работа написана хорошим языком почти без опечаток.

Есть несколько замечаний и комментариев к тексту и оформлению рисунков:

1. Стр. 13. Первый абзац целесообразно было бы дополнить уточняющими комментариями автора о том, какие именно тренды развития и применения ХТМ следуют из приведённого списка.
2. Стр. 17. Если бы автор приводил при сравнении качества моделей не только СКО отклонений (в абсолютных единицах), но и коэффициенты вариации (ОСО), можно было бы сравнить качество моделей в отношении обоих загрязнителей, озона и PM_{10} , что, в принципе, любопытно.
3. На стр. 51 в третьем абзаце утверждение «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» выглядит излишне категорично. Ведь возможно классическое перевевание транспортом и/или ветром частиц PM_{10} , ранее опустившихся на поверхность, и переход их обратно в состав приземного аэрозоля (resuspention), как в пригороде, так и особенно в городе – с тротуарной плитки или дорожного покрытия.

4. Несколько замечаний по рисункам:

Рис. 2.10 не полностью выполняет свою задачу. Хотелось бы видеть, как пыль распространяется в воздухе до Москвы.

Рис. 2.14 почему-то представлен в виде прямых траекторий переноса, хотя на всех аналогичных рисунках траектории обратные. Почему? В результате этот рисунок достоверно не подтверждает перенос воздуха и примесей именно в Московский регион

5. К сожалению, верификация моделей не проводилась на данных зимних месяцев.

Почему? Возможно, там у ММО были бы свои интересные возможности и другие наиболее значимые предикторы, связанные с совершенно другими процессами в слое воздуха над заснеженной поверхностью.

Сформулирую ряд вопросов, которые возникли по мере прочтения рукописи:

Вопрос 1. Наиболее важными предикторами не только для озона, но и для аэрозолей оказались природные факторы. Если строить такие модели только для территории города (и обучать их на территории города), ранжирование используемых предикторов будет иное и, возможно, значимость антропогенных факторов возрастет?

Вопрос 2. При дальнем атмосферном переносе (например, из Прикаспия) шлейф переносимого воздуха и примесей постепенно расширяется, и в Московском регионе покрывает не только сам город, но и большую часть пригородных территорий. Может ли быть предиктором дальнего переноса пыли метрика «одновременности» возрастания и убывания концентрации PM_{10} в городе и на «фоновых» станциях?

Вопрос 3. Можно ли оценить влияние объемов обучающей и тестовой выборок, а также их соотношения на эффективность ММО постпроцессинга? Есть ли у этих показателей оптимальные и предельные значения? Понятно, что это влияние должно зависеть от изменчивости предикторов и целевых компонент. Но для полученных конкретных ММО для озона и аэрозолей PM_{10} в Московском регионе можно посчитать такие метрики?

Отмеченные замечания и комментарии по оформлению текста диссертационной работы несколько не умаляют значимости, цельности и логической замкнутости исследования Д.В. Борисова. Практическое применение результатов этой работы уже находит свое применение в совершенствовании прогнозов загрязнения приземного воздуха

Московского региона. Поставленные в отзыве вопросы не требуют их немедленного решения, но позволяют взглянуть на эту задачу под другими углами.

***Заключение.** Диссертационная работа Борисова Данила Владимировича представляет собой законченную научно-квалификационную работу и соответствует паспорту специальности 1.6.18 – Науки об атмосфере и климате. Диссертация Борисова Д.В. удовлетворяет требованиям Положения о порядке присуждения ученых степеней, утвержденного Постановлением Правительства Российской Федерации от 24.09.2013 г. № 842 (с изменениями на 25.01.2024 г.), а Борисов Данил Владимирович заслуживает присуждения искомой степени кандидата географических наук по специальности 1.6.18 – Науки об атмосфере и климате.*

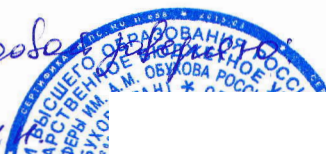
Официальный оппонент

ведущий научный сотрудник
ИФА им. А.М. Обухова РАН,
доктор географических наук



А.А. Виноградова

Подпись А.А. Виноградова
Ученый секретарь, к.ф.-м.н.



13. 05. 2025
Мисева Ю. В.

