

ОТЗЫВ

на автореферат диссертации Борисова Данила Владимировича «Постпроцессинг численных прогнозов концентраций взвешенных частиц (PM10) и приземного озона (O3) с использованием моделей машинного обучения», представленной на соискание ученой степени кандидата географических наук по специальности 1.6.18 – Науки об атмосфере и климате

Работа **Борисова Данила Владимировича** посвящена крайне актуальной и востребованной задаче повышения точности прогнозов качества воздуха в мегаполисе путем синтеза «классического» гидродинамического моделирования (моделей COSMO-Ru и CHIMERE) и современных методов машинного обучения. Подобный подход находится на переднем крае современных исследований в области наук об атмосфере и экологического прогнозирования. Особую значимость работе придает тот факт, что в части применения методов машинного обучения в науках о Земле, и, в частности, в метеорологии, российская наука заметно отстает от ведущих зарубежных стран. Поэтому данное исследование не только актуально, но и стратегически важно для развития отечественных технологий.

Работа производит впечатление целостного, глубокого и масштабного исследования, сфокусированного на разработке и всесторонней апробации новой технологии постпроцессинга прогнозов концентраций ключевых загрязнителей воздуха (PM10 и O3) на основе методов машинного обучения. В числе несомненных достоинств работы можно выделить:

- Методологию построения моделей машинного обучения, предполагающая одновременный учет пространственной и временной изменчивости концентраций загрязняющих веществ. Это позволяет улучшать прогнозы не только в точках измерений, но и на всей территории расчетного домена, включая районы без станций мониторинга.
- Детальное и разностороннее исследование качества разработанной модели в различных условиях, включая эпизоды наиболее повышенных концентраций.
- Впечатляющие результаты по улучшению точности прогнозов O3 и PM10 за счет применения разработанных моделей постпроцессинга – снижение СКО на десятки процентов, значительно улучшение прогнозируемого суточного хода.
- Привязку к оперативным технологиям Гидрометцентра, позволяющая рассчитывать на внедрение результатов исследования в оперативную практику.

В то же время, к работе имеется и ряд замечаний. Их можно разделить на замечания по сути работы и в части формулировок.

Замечания по сути работы:

1. Как отмечено выше, постановка задачи предполагает одновременный учет пространственной и временной изменчивости концентраций загрязняющих веществ в моделях машинного обучения, что позволяет применять эти модели на сетке по пространству и является важным шагом вперед по сравнению с более простыми поточечными моделями. Однако, для учета пространственной неоднородности загрязнения в мегаполисе используется только один предиктор – доля площади урбанизированных территорий для по данным Globcover, определяемая для ячеек 2-х километровой сетки, вмещающих АСКЗА. Такой подход представляется крайне упрощенным. Во-первых, такой предиктор вряд ли характеризует неоднородности

локальных условий загрязнения, так как не учитывает расположение АСКЗА относительно локальных источников загрязнения - дорог, промышленных предприятий и пр. Во-вторых, он не позволяет учесть нелокальные особенности поля загрязнения – например, влияние шлейфа загрязнений от источника на соседние ячейки. Непонятно, как без учета нелокальных эффектов получилось воспроизвести шлейфы загрязнений (раздел 4.3). В связи с этим не хватает более детальных оценок, в какой степени предлагаемый подход позволяет воспроизводить основные пространственные закономерности поля загрязнений в мегаполисе.

2. В конфигурации модели машинного обучения для постпроцессинга прогнозов озона в списке предикторов фигурирует «Среднесуточное общее содержание озона в атмосферном столбе» (табл. 1). Во-первых, в тексте не приводится никакой информации об этих данных, не понятно, идет ли речь о наблюдениях или моделировании, какое их разрешение и пр. Если это наблюдения в ЦАО, то их использование в качестве предиктора существенно ограничивает применимость разработанной модели для других регионов, где нет подобных наблюдений. В связи с этим в работе не хватает оценок, насколько велико влияние данного предиктора на качество разработанной модели. Можно ли от него отказаться или чем-то заменить?
3. Соискателем проведена большая работа по подбору оптимальных конфигураций нейронных сетей. В разделе 3.2 отмечается, что в ходе исследования рассматривались различные конфигурации ИНС, включая рекуррентные, однако особенности работы различных моделей МО в тексте не рассматриваются. Такое сравнение представляется очень интересным, ему стоило бы уделить больше внимания. Также было бы целесообразным включить в список рассматриваемых моделей машинного обучения не только нейросетевые – например, алгоритмы градиентного бустинга, которые по опыту рецензента в рассматриваемой постановке задачи могут показать лучшее качество.

Замечания в части изложения материала и используемых формулировок:

4. Удивляет включение в список решаемых пункта про «Изучение и освоение методов машинного обучения». Это образовательная, но не научная задача. Более уместным было бы включить, например, задачу «выбор методов машинного обучения для задач исследования».
5. В разделе «Научная новизна» не приводится никакой информации на тему, в какой степени и по сравнению с чем приводимые там достижения являются новыми – т.е. что сделано впервые в мире, что впервые в России, что впервые в оперативной практике Гидрометцентра?
6. Несколько неудачной выглядит формулировка II и III защищаемых положений – они оба написаны по одному шаблону, меняются лишь загрязняющие вещества и конкретные метрики качества. Если бы методика применялась для 10 загрязняющих веществ, делать 10 защищаемых положений вряд ли было бы корректно.
7. Основным источником натуральных данных в работе является сеть АСКЗА, при этом автор нигде не упоминает ни конкретный источник этих данных, ни кто является оператором данной сети. Судя по всему, речь идет о сети АСКЗА ГПБУ "Мосэкомониторинг". Отсутствие ссылки на эту организацию удивляет, особенно с учетом активного сотрудничества между ней и Гидрометцентром.

Указанные замечания не умаляют значимости и практической ценности диссертационного исследования. Диссертация 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.).

Таким образом, соискатель Борисов Данил Владимирович заслуживает присуждения ученой степени кандидата географических наук по специальности 1.6.18 – Науки об атмосфере и климате.

Я, Варенцов Михаил Иванович, даю согласие на включение своих персональных данных в документы, связанные с работой диссертационного совета, и их дальнейшую обработку.

Кандидат географических наук,
старший научный сотрудник лаборатории
суперкомпьютерного моделирования природно-климатических процессов
Научно-исследовательского вычислительного центра
ФГБУО ВО «Московский государственный университет имени М. В. Ломоносова»

Варенцов Михаил Иванович

04.06.2025

Подпись Варенцова М.И. Верна
Вед. науч. по атмосфере, Вост. О.В.

УФР 103770327

Контактные данные:

Тел. +7 495 939-5424, e-mail: mikhail.varentsov@srcc.msu.ru

Ленинские Горы ул., д.1, стр.4, Москва, 119234

Специальность, по которой защищена диссертация: 25.00.30 - Метеорология, климатология, агрометеорология