

УДК 004.8:519.876.5:504.06

DOI: [10.26102/2310-6018/2026.53.2.003](https://doi.org/10.26102/2310-6018/2026.53.2.003)

Интерпретируемый прогноз загрязнения атмосферного воздуха мелкодисперсными частицами на основе данных мониторинга и методов машинного обучения

Е.В. Филюшина¹, В.А. Орлов², Л.В. Красовская², А.С. Прудкий³

¹Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнёва, Красноярск, Российская Федерация

²Сибирский федеральный университет, Красноярск, Российская Федерация

³Российский государственный аграрный университет – МСХА имени К.А. Тимирязева, Москва, Российская Федерация

Резюме. Загрязнение атмосферного воздуха мелкодисперсными частицами с аэродинамическим диаметром менее 2,5 микрометра представляет серьезную экологическую и социальную проблему в условиях урбанизированных территорий. В связи с этим актуальной является задача краткосрочного прогноза концентрации данных частиц на основе данных мониторинга качества воздуха. В работе рассматривается применение интерпретируемых методов машинного обучения для прогнозирования концентрации мелкодисперсных частиц на часовом горизонте. В качестве исходных данных использован открытый набор Beijing PM2.5 Data Set, содержащий почасовые измерения концентрации загрязняющих веществ и метеорологических параметров за период с 2010 по 2014 годы. В ходе исследования выполнена предварительная обработка данных, сформировано признаковое пространство с учетом временной структуры и автокорреляционных свойств временных рядов, а также построены модели линейной регрессии, случайного леса и градиентного бустинга. Качество прогнозирования оценивалось с использованием средней абсолютной ошибки, среднеквадратичной ошибки и коэффициента детерминации. Результаты показали, что все рассмотренные модели обеспечивают высокую точность краткосрочного прогноза, при этом различия между моделями различной сложности оказываются незначительными. Установлено, что доминирующий вклад в формирование прогноза вносит автокорреляция временного ряда концентрации загрязняющих частиц, тогда как метеорологические параметры выполняют корректирующую функцию. Полученные результаты подтверждают целесообразность использования интерпретируемых моделей машинного обучения в системах мониторинга и прогнозирования качества атмосферного воздуха.

Ключевые слова: загрязнение атмосферного воздуха, мелкодисперсные частицы, краткосрочный прогноз, машинное обучение, интерпретируемые модели, временные ряды, мониторинг качества воздуха.

Для цитирования: Филюшина Е.В., Орлов В.А., Красовская Л.В., Прудкий А.С. Интерпретируемый прогноз загрязнения атмосферного воздуха мелкодисперсными частицами на основе данных мониторинга и методов машинного обучения. *Моделирование, оптимизация и информационные технологии*. 2026;14(2). URL: <https://moitvivt.ru/ru/journal/article?id=2167> DOI: 10.26102/2310-6018/2026.53.2.003

Interpretable forecasting of fine particulate air pollution based on monitoring data and machine learning methods

Е.В. Filushina¹, В.А. Orlov², Л.В. Krasovskaya², А.С. Prudky³

¹Siberian State University of Science and Technology named after Academician M.F. Reshetnev, Krasnoyarsk, the Russian Federation

²*Siberian Federal University, Krasnoyarsk, the Russian Federation*

³*Russian State Agrarian University – Moscow Timiryazev State Agrarian Academy, Moscow,
the Russian Federation*

Abstract. Atmospheric air pollution by fine particles with an aerodynamic diameter of less than 2.5 micrometers is a serious environmental and social problem in urban areas. In this context, short-term forecasting of fine particulate matter concentrations based on air quality monitoring data is of particular importance. This study investigates the applicability of interpretable machine learning methods for hourly forecasting of fine particulate air pollution. The publicly available Beijing PM2.5 data set, containing hourly measurements of particulate matter concentration and meteorological parameters for the period from 2010 to 2014, was used as the data source. Data preprocessing was performed, and a feature space was constructed with consideration of temporal structure and autocorrelation properties of the time series. Linear regression, random forest, and gradient boosting models were developed and evaluated. Forecasting performance was assessed using mean absolute error, root mean squared error, and the coefficient of determination. The results demonstrate that all considered models provide high accuracy for short-term forecasting, while differences in performance between models of varying complexity remain insignificant. It was found that the dominant contribution to the forecast is provided by the autocorrelation of the particulate matter concentration time series, whereas meteorological parameters play a corrective role. The obtained results confirm the feasibility of using interpretable machine learning models in air quality monitoring and forecasting systems.

Keywords: air pollution, fine particulate matter, short-term forecasting, machine learning, interpretable models, time series, air quality monitoring.

For citation: Filyushina E.V., Orlov V.A., Krasovskaya L.V., Prudkiy A.S. Interpretable forecasting of fine particulate air pollution based on monitoring data and machine learning methods. *Modeling, Optimization and Information Technology*. 2026;14(2). (In Russ.). URL: <https://moitvivt.ru/ru/journal/article?id=2167> DOI: 10.26102/2310-6018/2026.53.2.003

Введение

Загрязнение атмосферного воздуха мелкодисперсными взвешенными частицами PM2.5 (particulate matter с аэродинамическим диаметром менее 2,5 мкм) остается одной из наиболее серьезных экологических и социальных проблем современных урбанизированных территорий. Высокая проникающая способность частиц PM2.5 обуславливает их проникновение в альвеолярные отделы легких, что связано с повышением риска сердечно-сосудистых, респираторных и онкологических заболеваний¹. По данным Всемирной организации здравоохранения, воздействие PM2.5 ежегодно приводит к миллионам преждевременных смертей, преимущественно в крупных городах [1].

В условиях развития автоматизированных систем мониторинга и сетей IoT-датчиков формируются массивы данных с высокой временной разрешающей способностью, что создает предпосылки для применения методов машинного обучения в задачах краткосрочного прогнозирования качества воздуха [2]. В последние годы для прогнозирования концентрации PM2.5 активно применяются ансамблевые методы и модели глубокого обучения, демонстрирующие высокую точность, особенно при часовом горизонте прогнозирования [3, 4].

В то же время использование сложных моделей зачастую сопровождается снижением интерпретируемости результатов, что ограничивает их применение в системах поддержки принятия решений и экологического управления [5]. Для практических задач мониторинга качества воздуха важным является не только

¹ Air pollution. World Health Organization. URL: <https://www.who.int/health-topics/air-pollution> (дата обращения: 15.12.2025).

достижение минимальной ошибки прогноза, но и возможность объяснения влияния факторов, формирующих уровень загрязнения [6, 7]. В этой связи особый интерес представляют интерпретируемые модели машинного обучения, обеспечивающие прозрачность выводов и устойчивость к изменению входных данных.

Анализ современных исследований показывает, что краткосрочная динамика концентрации PM2.5 определяется как метеорологическими параметрами (температура, влажность, скорость и направление ветра, атмосферное давление) [8], так и выраженным автокорреляционными свойствами самого временного ряда загрязнения [9, 10]. При этом в ряде работ отмечается, что включение лаговых значений PM2.5 позволяет существенно повысить точность прогноза, а усложнение модели не всегда приводит к статистически значимому улучшению результатов [11, 12]. Однако вопрос соотношения вклада автокорреляции и метеорологических факторов, а также баланса между точностью и интерпретируемостью моделей для задачи часового прогноза PM2.5 остается недостаточно формализованным.

Научная проблема, рассматриваемая в данной работе, заключается в определении целесообразности использования интерпретируемых моделей машинного обучения для краткосрочного прогнозирования концентрации PM2.5 и оценке того, в какой степени усложнение модели повышает точность прогноза при наличии выраженной автокорреляции временного ряда.

Гипотеза исследования состоит в том, что при часовом горизонте прогнозирования концентрации PM2.5 доминирующий вклад в формирование прогноза вносит автокорреляция временного ряда, тогда как метеорологические параметры выполняют корректирующую функцию, вследствие чего интерпретируемые модели могут обеспечивать сопоставимую точность с более сложными ансамблевыми методами.

Научная новизна работы заключается в:

- сравнительном анализе интерпретируемых и ансамблевых моделей машинного обучения в задаче часового прогноза PM2.5 в едином признаковом пространстве;
- эмпирическом подтверждении ограниченного выигрыша в точности при усложнении модели при наличии лаговых признаков концентрации PM2.5;
- обосновании применимости интерпретируемых моделей для задач мониторинга и поддержки принятия решений в области качества атмосферного воздуха.

Цель исследования – оценка эффективности интерпретируемых моделей машинного обучения при краткосрочном прогнозировании концентрации PM2.5 и их сравнение с ансамблевыми методами с точки зрения точности и интерпретируемости.

Для достижения поставленной цели в работе решаются следующие задачи:

- анализ и предварительная обработка данных мониторинга концентрации PM2.5 и метеорологических параметров;
- формирование признакового пространства с учетом временной структуры и автокорреляционных свойств временных рядов;
- построение и обучение моделей линейной регрессии, случайного леса и градиентного бустинга;
- сравнение качества прогнозирования с использованием стандартных метрик;
- анализ вклада факторов и интерпретируемости полученных моделей.

Материалы и методы

Исследование направлено на краткосрочный прогноз концентрации мелкодисперсных частиц PM2.5 на основе временных рядов данных мониторинга атмосферного воздуха. Общая схема включает следующие этапы:

- выбор и анализ исходного набора данных;

- предварительная обработка и формирование признаков;
- построение моделей машинного обучения;
- оценку качества прогнозирования;
- интерпретацию полученных результатов [5, 13].

Такой подход соответствует современным методикам экологического моделирования, основанным на применении машинного обучения и статистического анализа временных рядов загрязняющих веществ [11].

В качестве исходных данных использовался открытый набор Beijing PM2.5 Data Set, размещенный в UCI Machine Learning Repository². Он широко применяется в исследованиях по прогнозированию загрязнения воздуха [4, 10]. Датасет содержит почасовые измерения концентрации PM2.5 и сопутствующих метеорологических параметров (температура, точка росы, давление, скорость и направление ветра, осадки), зарегистрированные в Пекине с 2010 по 2014 годы.

Целевой переменной являлась концентрация PM2.5, выраженная в мкг/м³, а временные метки включали год, месяц, день и час. Такой набор обеспечивает возможность анализа сезонных, суточных и погодных закономерностей в динамике загрязнения воздуха [6].

Предобработка данных включала удаление строк с пропущенными значениями целевой переменной PM2.5, что является стандартной процедурой при анализе экологических временных рядов [2]. Категориальный признак направления ветра (cbwd) был преобразован методом one-hot кодирования.

Для учета сезонной и суточной периодичности были сформированы циклические признаки на основе синусоидальных и косинусоидальных преобразований месяца и часа наблюдения [13, 14]. Это обеспечивает корректное представление периодических зависимостей для моделей машинного обучения.

Кроме того, в признаковое пространство включены лаговые значения концентрации PM2.5 на временных интервалах t-1, t-3 и t-6 часов, что позволяет учитывать автокорреляционные свойства ряда [8, 15]. После формирования лагов строки с отсутствующими данными были удалены.

Для прогноза концентрации PM2.5 использовались три модели машинного обучения различной сложности:

Линейная регрессия (LR) – базовая интерпретируемая модель, обеспечивающая количественную оценку вклада факторов и направление влияния предикторов [16].

Случайный лес (Random Forest, RF) – ансамблевая модель на основе решающих деревьев, обладающая устойчивостью к выбросам и способностью выявлять нелинейные зависимости [17].

Градиентный бустинг (XGBoost) – ансамблевый метод, ориентированный на минимизацию ошибки за счет последовательного обучения слабых моделей и широко применяемый для прогнозирования качества воздуха [18].

Разделение выборки на обучающую и тестовую выполнялось с сохранением хронологического порядка наблюдений (без случайного перемешивания), что является корректным подходом для временных рядов [19].

Качество прогнозов оценивалось с помощью трех метрик: MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) и R² (коэффициент детерминации) [5]. Эти показатели наиболее часто используются для оценки моделей прогнозирования загрязнения атмосферы, включая задачи краткосрочного прогноза PM2.5.

² UCI Machine Learning Repository. Beijing PM2.5 Data Set. University of California, Irvine; 2017. Available from: <https://archive.ics.uci.edu/dataset/381/beijing+pm2+5+data>

Интерпретация результатов осуществлялась через анализ коэффициентов линейной регрессии и оценку значимости признаков в ансамблевых моделях, что позволило определить вклад автокорреляционных и метеорологических факторов в формирование концентрации PM2.5 [6].

Результаты

В ходе исследования были построены и протестированы три модели машинного обучения для краткосрочного прогноза концентрации PM2.5: линейная регрессия, метод случайного леса и метод градиентного бустинга. Для всех моделей использовалась одинаковая обучающая и тестовая выборка, сформированная с сохранением временного порядка наблюдений.

Качество прогнозирования оценивалось с использованием показателей средней абсолютной ошибки (MAE), среднеквадратичной ошибки (RMSE) и коэффициента детерминации R^2 . Полученные значения метрик приведены в Таблице 1.

Таблица 1 – Показатели качества прогнозирования концентрации PM2.5
Table 1 – PM2.5 prediction performance metrics

Модель	MAE, $\mu\text{g}/\text{m}^3$	RMSE, $\mu\text{g}/\text{m}^3$	R^2
Линейная регрессия	11,87	21,29	0,949
Случайный лес	11,72	21,19	0,950
Градиентный бустинг	11,53	20,85	0,951

Для анализа структуры моделей и вклада входных параметров была рассчитана важность признаков для ансамблевых методов и коэффициенты линейной регрессии. В Таблице 2 представлены десять наиболее значимых признаков для модели случайного леса.

Таблица 2 – Наиболее значимые признаки модели случайного леса
Table 2 – Top feature importances for the Random Forest model

Признак	Важность
pm25_lag1	0,961607
pm25_lag3	0,008069
pm25_lag6	0,004358
Iws	0,004225
DEWP	0,003778
TEMP	0,002566
day	0,002421
PRES	0,002255
hour	0,001877
hour_sin	0,001838

Результаты оценки коэффициентов линейной регрессии, отсортированных по абсолютному значению, приведены в Таблице 3.

Таблица 3 – Коэффициенты линейной регрессии
Table 3 – Linear regression coefficients

Признак	Коэффициент
cbwd_cv	4,443419
month_cos	3,901312
cbwd_NW	-3,396395
month_sin	3,289014
cbwd_SE	3,220523
hour_sin	-3,207123
pm25_lag1	0,994470
Ir	-0,769080
TEMP	-0,571686
DEWP	0,554638

В Таблице 4 представлены наиболее значимые признаки для модели градиентного бустинга.

Таблица 4 – Наиболее значимые признаки модели градиентного бустинга
Table 4 – Top feature importances for the gradient boosting model

Признак	Важность
pm25_lag1	0,963474
pm25_lag3	0,008086
DEWP	0,006866
Iws	0,004651
pm25_lag6	0,003385
TEMP	0,001936
hour_sin	0,001670
PRES	0,001622
hour	0,001494
cbwd_NW	0,001189

Таким образом, в разделе представлены количественные результаты прогнозирования концентрации PM2.5 и параметры обученных моделей, полученные в ходе вычислительного эксперимента.

Обсуждение

Полученные результаты показывают, что все рассмотренные модели машинного обучения демонстрируют высокое качество краткосрочного прогноза концентрации PM2.5. Значения коэффициента детерминации R^2 для линейной регрессии, случайного леса и градиентного бустинга превышают 0,94, что свидетельствует о способности моделей адекватно описывать динамику временного ряда PM2.5 на часовом интервале прогнозирования. При этом различия в точности между моделями различной сложности оказываются незначительными.

Анализ структуры моделей и важности признаков показывает, что во всех случаях доминирующий вклад в формирование прогноза вносит лаговое значение концентрации PM2.5 за предыдущий час. Данный факт указывает на выраженные автокорреляционные свойства временного ряда PM2.5 и подтверждает инерционный характер процессов загрязнения атмосферного воздуха. Подобные выводы о ключевой роли автокорреляции

при краткосрочном прогнозировании PM2.5 отмечаются в ряде современных исследований [10, 13], где также подчеркивается высокая прогностическая ценность прошлых значений концентрации для кратковременных прогнозов.

Второстепенную, но устойчивую роль в моделях играют метеорологические параметры, прежде всего скорость ветра, температура воздуха, точка росы и атмосферное давление. Их вклад отражает физические процессы рассеивания, переноса и накопления аэрозольных частиц в приземном слое атмосферы. Состояние метеорологической обстановки, особенно температурные инверсии и слабые ветры, может усиливать накопление PM2.5 и замедлять его рассеивание, что подтверждается в работах, посвященных взаимодействию метеоусловий и концентрации загрязнителей [8, 9]. Направление ветра и временные признаки, характеризующие суточную и сезонную цикличность, также оказывают влияние на прогноз, однако их значимость существенно ниже по сравнению с лаговыми характеристиками PM2.5. Подобное распределение важности факторов согласуется с результатами других авторов, указывающих, что метеорологические параметры выполняют корректирующую функцию относительно основной автокорреляционной динамики загрязнения [8, 9].

Сопоставление результатов линейной регрессии и ансамблевых моделей показывает, что усложнение алгоритма приводит лишь к незначительному росту точности прогнозирования. Градиентный бустинг демонстрирует наилучшие значения метрик, однако превосходство над линейной регрессией и случайным лесом находится в пределах нескольких процентов. Это позволяет сделать вывод о том, что в условиях краткосрочного прогноза с использованием лаговых признаков сложные нелинейные зависимости играют ограниченную роль. Аналогичные наблюдения приводятся в работах, посвященных интерпретируемому прогнозированию качества воздуха, где отмечается, что линейные и ансамблевые модели могут обеспечивать сопоставимую точность при наличии выраженной автокорреляции временного ряда [6, 16]. В частности, в [6] подчеркивается, что простые линейные модели при определенных конфигурациях признаков показывают эффективность, сопоставимую с более сложными алгоритмами.

Таким образом, результаты исследования подтверждают целесообразность применения интерпретируемых моделей машинного обучения в задачах мониторинга и прогноза загрязнения воздуха. Линейная регрессия и случайный лес обеспечивают оптимальный баланс между точностью и объяснимостью, что делает их особенно эффективными при интеграции в системы поддержки принятия решений и экологического управления. В то же время использование более сложных ансамблевых моделей, таких как градиентный бустинг, может быть оправдано в случаях, когда требуется максимальная точность, однако выигрыш по сравнению с интерпретируемыми методами остается умеренным.

В более широком контексте полученные результаты указывают, что при анализе временных рядов концентрации PM2.5 ключевым фактором успешного прогнозирования является корректный учет временной структуры данных и автокорреляционных свойств процесса загрязнения. Это подтверждает современные тенденции в области экологического моделирования, направленные на повышение интерпретируемости и устойчивости прогнозных систем [5, 18].

Тем не менее исследование имеет ряд ограничений, которые следует учитывать при интерпретации результатов. Первое ограничение связано с тем, что модели были обучены на данных, характерных для конкретной географической области (например, мегаполиса с определенной структурой источников эмиссий и метеорологическими условиями). Прямой перенос моделей на другие регионы с иной топографией и климатом может потребовать адаптации и переобучения, поскольку вклад метеорологических

факторов и характер автокорреляции может варьироваться. Во-вторых, анализ был сконцентрирован на часовом горизонте прогнозирования; для более длинных временных горизонтов (суточный и более) вклад нелинейных эффектов и метеоусловий может возрастать, что потенциально усиливает преимущества сложных моделей. Наконец, в данной работе не учитывались пространственные взаимосвязи между различными пунктами мониторинга, которые в региональном прогнозировании PM2.5 могут быть значимыми.

Следует отметить, что помимо анализа коэффициентов линейной регрессии и стандартной оценки важности признаков в ансамблевых моделях, в современной практике интерпретации моделей машинного обучения широко применяются методы объяснимого искусственного интеллекта (Explainable AI, XAI). Одним из наиболее распространенных подходов является метод SHAP (SHapley Additive Explanations), основанный на значениях Шепли и позволяющий оценивать вклад каждого признака в формирование как глобальных, так и локальных прогнозов ансамблевых моделей. Применение SHAP-анализа позволяет получать более устойчивые и детализированные оценки влияния факторов по сравнению с классической важностью признаков, а также анализировать направление влияния отдельных переменных на прогнозируемую концентрацию PM2.5. Использование подобных методов интерпретации может быть рассмотрено в качестве перспективного направления дальнейших исследований и дополнительно повысить прозрачность и практическую применимость прогнозных моделей качества воздуха.

Заключение

В работе рассмотрена задача краткосрочного прогноза концентрации мелкодисперсных частиц PM2.5 на основе данных мониторинга качества атмосферного воздуха с использованием методов машинного обучения. В рамках исследования проведен анализ временных рядов концентрации PM2.5 и метеорологических параметров, а также выполнено построение и сравнение моделей различной сложности, включая линейную регрессию, метод случайного леса и градиентный бустинг.

Полученные результаты показывают, что все рассмотренные модели обеспечивают высокое качество прогнозирования при часовом горизонте, а ключевым фактором, определяющим точность прогноза, является учет автокорреляционных свойств временного ряда PM2.5. Установлено, что использование лаговых признаков позволяет достигать сопоставимой точности как для интерпретируемых моделей, так и для более сложных ансамблевых методов, при этом выигрыш в точности при усложнении модели остается ограниченным.

Практическая значимость работы заключается в обосновании целесообразности применения интерпретируемых моделей машинного обучения для задач мониторинга и прогноза загрязнения атмосферного воздуха. Линейная регрессия и метод случайного леса обеспечивают оптимальный баланс между точностью и объяснимостью, что делает их перспективными для внедрения в интеллектуальные системы поддержки принятия решений и экологического управления.

В качестве направлений дальнейших исследований целесообразно рассматривать расширение горизонта прогнозирования, учет пространственной неоднородности загрязнения воздуха, а также интеграцию данных от распределенных сетей IoT-датчиков и метеорологических моделей. Дополнительный интерес представляет исследование методов интерпретируемого машинного обучения в условиях неполноты и зашумленности данных, что позволит повысить устойчивость и практическую применимость прогнозных систем качества воздуха.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Cohen A.J., Brauer M., Burnett R., et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet*. 2017;389(10082):1907–1918. [https://doi.org/10.1016/S0140-6736\(17\)30505-6](https://doi.org/10.1016/S0140-6736(17)30505-6)
2. Deters J.K., Zalakeviciute R., González M., Rybarczyk Y. Modeling PM_{2.5} Urban Pollution Using Machine Learning and Selected Meteorological Parameters. *Journal of Electrical and Computer Engineering*. 2017;2017. <https://doi.org/10.1155/2017/5106045>
3. Fang Z., Yang H., Li C., Cheng L., Zhao M., Xie C. Prediction of PM2.5 hourly concentrations in Beijing based on machine learning algorithm and ground-based LiDAR. *Archives of Environmental Protection*. 2021;47(3):98–107. <https://doi.org/10.24425/aep.2021.138468>
4. Ma J., Yu Zh., Qu Y., Xu J., Cao Y. Application of the XGBoost Machine Learning Method in PM2.5 Prediction: A Case Study of Shanghai. *Aerosol and Air Quality Research*. 2020;20(1):128–138. <https://doi.org/10.4209/aaqr.2019.08.0408>
5. Ejohwomu O.A., Oshodi O.Sh., Oladokun M., et al. Modelling and Forecasting Temporal PM_{2.5} Concentration Using Ensemble Machine Learning Methods. *Buildings*. 2022;12(1). <https://doi.org/10.3390/buildings12010046>
6. Zhang Y., Sun Q., Liu J., Petrosian O. Long-Term Forecasting of Air Pollution Particulate Matter (PM2.5) and Analysis of Influencing Factors. *Sustainability*. 2024;16(1). <https://doi.org/10.3390/su16010019>
7. Antamoshkin O., Kukarcev V., Pupkov A., Tsarev R. Intellectual Support System of Administrative Decisions in the Big Distributed Geoinformation Systems. In: *14th International Multidisciplinary Scientific GeoConference SGEM 2014, 17–26 June 2014, Albena, Bulgaria*. Sofia: STEF92 Technology Ltd; 2014. P. 227–232.
8. Yin P.-Y., Chang R.-I., Day R.-F., Lin Y.-Ch., Hu Ch.-Y. Improving PM2.5 Concentration Forecast with the Identification of Temperature Inversion. *Applied Sciences*. 2022;12(1). <https://doi.org/10.3390/app12010071>
9. Xiao Q., Zheng Y., Geng G., et al. Separating emission and meteorological contributions to long-term PM_{2.5} trends over eastern China during 2000–2018. *Atmospheric Chemistry and Physics*. 2021;21(12):9475–9496. <https://doi.org/10.5194/acp-2021-28>
10. Karimian H., Li Q., Wu Ch., et al. Evaluation of Different Machine Learning Approaches to Forecasting PM_{2.5} Mass Concentrations. *Aerosol and Air Quality Research*. 2019;19(6):1400–1410. <https://doi.org/10.4209/aaqr.2018.12.0450>
11. Masood A., Hameed M.M., Srivastava A., et al. Improving PM_{2.5} prediction in New Delhi using a hybrid extreme learning machine coupled with snake optimization algorithm. *Scientific Reports*. 2023;13. <https://doi.org/10.1038/s41598-023-47492-z>
12. Kukartsev V.V., Boyko A.A., Mikhalev A.S., Tynchenko V.S., Rukosueva A.A., Korpacheva L.N. Simulation-Dynamic Model of Working Time Costs Calculation for Performance of Operations on CNC Machines. In: *Journal of Physics: Conference Series: Volume 1582: High-Tech and Innovations in Research and Manufacturing (HIRM-2020), 28 February 2020, Siberia, Russia*. Bristol: IOP Publishing Ltd; 2020. <https://doi.org/10.1088/1742-6596/1582/1/012052>
13. Shen J., Valagolam D., McCalla S. Prophet forecasting model: a machine learning approach to predict the concentration of air pollutants (PM_{2.5}, PM₁₀, O₃, NO₂, SO₂, CO) in Seoul, South Korea. *PeerJ*. 2020;8. <https://doi.org/10.7717/peerj.9961>
14. Fedorova N.V., Dzhioeva N.N., Kukartsev V.V., Dalisova N.A., Ogol A.R., Tynchenko V.S. Methods of Assessing the Efficiency of the Foundry Industrial Marketing. In: *IOP Conference Series: Materials Science and Engineering: Volume 734*:

II International Scientific Conference "Advanced Technologies in Aerospace, Mechanical and Automation Engineering" – MIST: Aerospace – 2019, 18–21 November 2019, Krasnoyarsk, Russia. Bristol: IOP Publishing Ltd; 2020. <https://doi.org/10.1088/1757-899X/734/1/012083>

15. Kukartsev V.V., Tynchenko V.S., Chzhan E.A., et al. Solving the Problem of Trucking Optimization by Automating the Management Process. In: *Journal of Physics: Conference Series: Volume 1333: The International Conference "Information Technologies in Business and Industry", 18–20 February 2019, Novosibirsk, Russia*. Bristol: IOP Publishing Ltd; 2019. <https://doi.org/10.1088/1742-6596/1333/7/072027>
16. Agibayeva A., Khalikhan R., Guney M., Karaca F., Torezhan A., Avcu E. An Air Quality Modeling and Disability-Adjusted Life Years (DALY) Risk Assessment Case Study: Comparing Statistical and Machine Learning Approaches for PM_{2.5} Forecasting. *Sustainability*. 2022;14(24). <https://doi.org/10.3390/su142416641>
17. Lee M., Lin L., Chen Ch.Y., et al. Forecasting Air Quality in Taiwan by Using Machine Learning. *Scientific Reports*. 2020;10. <https://doi.org/10.1038/s41598-020-61151-7>
18. Morapedi T.D., Obagbuwa I.Ch. Air pollution particulate matter (PM2.5) prediction in South African cities using machine learning techniques. *Frontiers in Artificial Intelligence*. 2023;6. <https://doi.org/10.3389/frai.2023.1230087>
19. Palanichamy N., Haw S.-Ch., S S., Govindasamy K., Murugan R. Prediction of PM_{2.5} concentrations in Malaysia using machine learning techniques: a review. *F1000Research*. 2021;10. <https://doi.org/10.12688/f1000research.73163.1>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Елена Владимировна Филюшина, кандидат физико-математических наук, доцент кафедры информационно-экономических систем Сибирского государственного университета науки и технологий имени академика М.Ф. Решетнёва, Красноярск, Российская Федерация.

e-mail: marbury@yandex.ru

Орлов Василий Алексеевич, студент кафедры цифровых технологий управления, институт управления бизнес-процессами, Сибирский федеральный университет, Красноярск, Российская Федерация.

e-mail: vasi4244@gmail.com

ORCID: [0009-0002-7542-4548](https://orcid.org/0009-0002-7542-4548)

Красовская Людмила Владимировна, аспирант кафедры цифровых технологий управления, институт управления бизнес-процессами, Сибирский федеральный университет, Красноярск, Российская Федерация.

e-mail: kraslud@yandex.ru

ORCID: [0000-0002-9674-8384](https://orcid.org/0000-0002-9674-8384)

Elena V. Filushina, Candidate of Physico-mathematical Sciences, Associate Professor at the Department of Information and Economic Systems, Siberian Federal University of Science and Technology named after Academician M.F. Reshetnev, Krasnoyarsk, the Russian Federation.

Vasily A. Orlov, Student at the Department of Digital Technology in Management, Institute of Business Process Management, Siberian Federal University, Krasnoyarsk, the Russian Federation.

Lyudmila V. Krasovskaya, Postgraduate at the Department of Digital Technology in Management, Institute of Business Process Management, Siberian Federal University, Krasnoyarsk, the Russian Federation.

Прудкий Александр Сергеевич, кандидат педагогических наук, доцент, и. о. зав. кафедрой высшей математики, Российский государственный аграрный университет – МСХА имени К.А. Тимирязева, Москва, Российская Федерация.

e-mail: prudkiy@rgau-msha.ru

*Статья поступила в редакцию 26.12.2025; одобрена после рецензирования 04.02.2026;
принята к публикации 09.02.2026.*

*The article was submitted 26.12.2025; approved after reviewing 04.02.2026;
accepted for publication 09.02.2026.*