

**А.Н. Молдагулова¹, Д.Н. Тұрмаханбет^{2,3*}, М.А. Мауленбеков^{1,3}, Ж.Б. Кальпеева¹,
Р.К. Ускенбаева¹**

¹Казахский национальный технический исследовательский университет им. К.И. Сатпаева,
050013, Республика Казахстан, г. Алматы, ул. Сатпаев 22

²Казахский национальный университет имени аль-Фараби,
050040, Республика Казахстан, г. Алматы, пр. аль-Фараби, 71

³ТОО «Kazakhstan R&D Solutions»,
050056, Республика Казахстан, г. Алматы, ул. Кожедуба 3

*e-mail: turmahanbetdinara@gmail.com

МОНИТОРИНГ КАЧЕСТВА ВОЗДУХА В АВТОБУСАХ АЛМАТЫ С ИСПОЛЬЗОВАНИЕМ ИОТ-УСТРОЙСТВ И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Аннотация: *Качество воздуха в общественном транспорте – важный фактор, напрямую влияющий на здоровье горожан, однако для Алматы эта проблема до настоящего времени оставалась недостаточно изученной. В данном исследовании проведён систематический мониторинг концентраций CO_2 , $PM_{2.5}$, PM_{10} , температуры и относительной влажности в автобусах на трёх ключевых транспортных коридорах города. Сбор данных осуществлялся с использованием мобильного IoT-устройства собственной разработки «Tynys», откалиброванного по коммерческому эталонному датчику, что обеспечило высокую достоверность полученных измерений. Особое внимание было уделено сравнению условий в часы пик и вне их, что позволило выявить закономерности изменения параметров воздуха при различной загрузке транспорта. Анализ показал, что в утренние часы пик концентрации CO_2 и мелкодисперсных частиц значительно превышают уровни межпиковых периодов, что подтверждает ключевую роль пассажиропотока в формировании качества воздуха в салоне. Дополнительно с применением методов машинного обучения было установлено, что именно CO_2 и $PM_{2.5}$ являются наиболее информативными индикаторами загрязнённости, позволяющими надёжно классифицировать периоды с высокой нагрузкой. Полученные результаты восполняют существующий научный пробел в исследованиях качества воздуха в общественном транспорте региона Центральной Азии и формируют базу для последующих работ, направленных на модернизацию вентиляционных систем, оптимизацию пассажиропотоков и внедрение интеллектуальных технологий мониторинга микроклимата в автобусах.*

Ключевые слова: *общественный транспорт; качество воздуха; CO_2 ; $PM_{2.5}$; PM_{10} ; IoT-устройство; машинное обучение; вентиляция; пассажиропоток; микроклимат; интеллектуальный транспорт.*

Методология

Архитектура системы

Система мониторинга качества воздуха в реальном времени основана на компактном IoT-устройстве, предназначенном для эксплуатации в закрытых помещениях, включая автобусы [1, 2]. Устройство представляет собой автономный и экономически доступный сенсорный узел, способный функционировать в условиях ограниченного энергоснабжения и нестабильного соединения [3, 4]. В его состав входит сенсорный модуль, управление им и предварительная обработка данных выполняются на микрокомпьютере, где для повышения точности измерений используется алгоритм скользящего среднего [5]. Конструктивно устройство размещено в корпусе с вентиляционными отверстиями для обеспечения оптимальной циркуляции воздуха [6, 7]. Передача данных на сервер осуществляется по беспроводному интерфейсу Wi-Fi с использованием протокола MQTT, что обеспечивает низкие накладные расходы и высокую эффективность в условиях ограниченной пропускной способности [3, 8]. Каждый набор параметров снабжается временной меткой и сохраняется в формате CSV [4]. Для обеспечения непрерывности сбора данных в условиях нестабильной связи реализован механизм буферизации: измеренные значения временно сохраняются

локально и автоматически передаются на сервер после восстановления соединения [9]. На серверной стороне данные поступают в структурированную SQL-базу, что обеспечивает их долговременное хранение, быстрый доступ и последующую обработку [3]. Для пользователей предусмотрена веб-панель, позволяющая визуализировать как текущие, так и архивные показатели качества воздуха [10]. Архитектура системы является модульной и масштабируемой, что обеспечивает возможность подключения множества устройств и интеграции в более крупные инфраструктурные решения [11].

Как показано на схеме (рис. 1), архитектура разделена на четыре уровня. Первый уровень отвечает за сбор данных и их первичную обработку с помощью встроенных сенсоров и вычислительных модулей. Второй уровень представляет собой набор регистрируемых параметров окружающей среды – концентрации углекислого газа, уровня взвешенных частиц, а также температуры и влажности. На третьем уровне реализована передача информации и её хранение: данные отправляются по протоколу MQTT и сохраняются в базе SQL, что обеспечивает их доступность для анализа и визуализации. Четвёртый уровень связан с перспективами дальнейшего развития системы – он предназначен для внедрения интеллектуальных методов обработки и машинного обучения, позволяющих повысить точность и глубину аналитики. В рамках реализации этого уровня нами были опробованы алгоритмы Logistic Regression (точность $\approx 77\%$), Random Forest (точность $\approx 95\%$) и XGBoost (точность $\approx 95,5\%$), показавшие высокую эффективность при классификации и прогнозировании состояния воздуха. На основе результатов сравнения XGBoost был выбран в качестве основного алгоритма для промышленного развёртывания благодаря его стабильности и высокой точности. Разработанный ML-модуль интегрирован в серверную часть и обрабатывает поступающие данные в реальном времени, формируя прогнозы ухудшения или улучшения качества воздуха. Благодаря модульной архитектуре его возможно дообучать на новых данных, что обеспечивает адаптацию модели к сезонным изменениям и разным маршрутам автобусов.

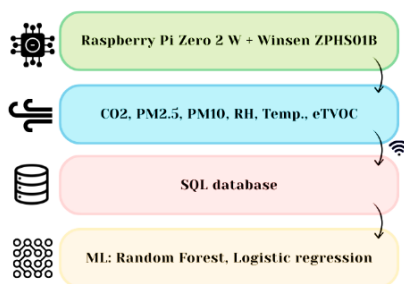


Рисунок 1 – Архитектура системы

Таким образом, ключевой особенностью системы является объединение компактного сенсорного узла, энергоэффективной обработки данных, встроенных методов машинного обучения и надёжной инфраструктуры передачи и хранения информации, что обеспечивает непрерывный и достоверный мониторинг качества воздуха в реальном времени в автобусах.

IoT устройство Breathe

В рамках исследования было использовано компактное IoT-устройство, предназначенное для мониторинга качества воздуха в салонах автобусов в режиме реального времени [10]. Оно регистрирует основные экологические показатели – концентрации CO_2 , $\text{PM}_{2.5}$, PM_{10} , а также температуру и относительную влажность, что позволило проанализировать взаимосвязь уровня загрязнения с пассажиропотоком и временем суток [7, 12, 13]. Перед началом полевых испытаний устройство прошло лабораторную калибровку: его показания были сверены с результатами эталонного коммерческого датчика Qingping, что подтвердило достоверность последующих измерений [14].

Аппаратная часть включает мини-компьютер Raspberry Pi Zero 2 W, который выступает в качестве управляющего узла; к нему подключены датчики температуры и влажности с точностью измерений $\pm 0,5^\circ\text{C}$ в диапазоне от 0 до 65°C и $\pm 3\%$ RH, лазерный сенсор для регистрации концентраций взвешенных частиц $\text{PM}_{2.5}$ и PM_{10} с минимальным порогом детекции $0,3\text{ }\mu\text{m}$, а также газовый сенсор CO_2 (NDIR-тип, диапазон 0-5000 ppm, точность $\pm(50\text{ ppm} + 5\% \text{ от значения измерения})$) [3, 9, 15]. Для обеспечения стабильной циркуляции воздуха применяются малогабаритные вентиляторы [6, 16]. Все компоненты размещены в компактном

3D-печатном корпусе с вентиляционными отверстиями и разъёмом питания, что делает устройство удобным для мобильного использования в автобусах [4]. Сбор данных осуществляется с интервалом 60 секунд, при этом для сглаживания шумов и устранения выбросов применяется алгоритм скользящего среднего [5]. Помимо сбора данных, устройство обладает функцией их визуализации в реальном времени, что позволяет оперативно отслеживать изменения качества воздуха [10]. В целом, устройство зарекомендовало себя как недорогое и надёжное решение для мониторинга воздуха в автобусах, обеспечившее эффективный сбор информации и ставшее ключевым инструментом для успешного проведения исследования [1, 3, 4].

Сбор данных

Сбор данных был организован в апреле 2025 года и охватывал три ключевых автобусных маршрута – В1, В2 и В3 (рис. 2), что позволило исследовать различные транспортные коридоры Алматы, включая как центральные участки с интенсивным пассажиропотоком, так и периферийные зоны с меньшей загрузкой. В маршрут В3 был включён участок в сторону Медео, что дало возможность оценить влияние рельефа и высоты на концентрацию загрязняющих веществ. Измерения выполнялись дважды в течение дня, чтобы зафиксировать контрастные условия: в утренний час пик (08:00-09:00), когда наблюдается максимальная плотность пассажиров и интенсивность движения, и в период относительно низкой загрузки (10:00-11:00). Такой выбор временных интервалов позволил выявить зависимость уровня загрязнения от пассажиропотока и условий вентиляции. Для снижения влияния случайных факторов измерения повторялись несколько дней подряд, а погодные условия фиксировались для последующего анализа их влияния на концентрации загрязнителей. Устройство размещалось в центральной части салона автобуса на высоте около 1,5 м – в зоне дыхания стоящего пассажира, что позволило получать репрезентативные данные, отражающие реальные условия экспозиции. Для исключения локальных аномалий устройство фиксировалось вдали от окон и дверей, чтобы минимизировать влияние кратковременных сквозняков.

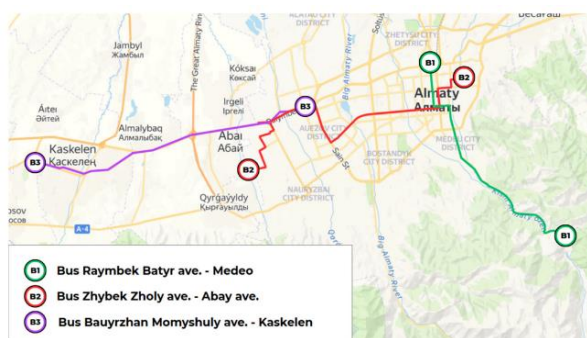


Рисунок 2 – Маршрут автобусов

Перед началом анализа данные прошли процедуру очистки и стандартизации. Были удалены записи с некорректными показаниями датчиков (например, возникающими при резком изменении питания устройства или кратковременном отключении Wi-Fi). Все числовые признаки нормализовались с использованием стандартизации, что обеспечило сопоставимость масштабов параметров при дальнейшем анализе и обучении моделей машинного обучения. Для оценки информативности собранных параметров был проведён анализ важности признаков с использованием модели Random Forest (рис. 3). Результаты показали, что наибольший вклад в вариации качества воздуха вносит тип транспортного средства, что связано с различиями в системах вентиляции и пассажиропотоке на маршрутах. Существенное влияние также оказывают температура и относительная влажность, отражающие микроклиматические условия в салоне. Концентрация CO₂ и уровень загруженности дороги показали меньший, но значимый вклад, что подтверждает необходимость их учёта при прогнозировании. Такой анализ позволил подтвердить, что выбранный набор параметров является репрезентативным и достаточным для построения моделей классификации и прогнозирования. Полученные результаты служат основой для

формирования обучающей выборки и дальнейшего применения методов машинного обучения на следующем этапе исследования.

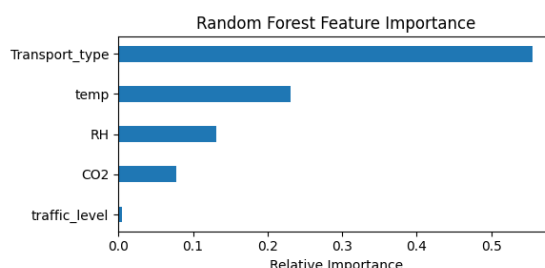


Рисунок 3 – Важность признаков в модели Random Forest

Результаты

Анализ собранных данных показал систематическую и выраженную зависимость качества воздуха в салоне автобусов от уровня пассажирской нагрузки. В периоды высокой заполняемости концентрации основных загрязнителей возрастали резко и существенно. Среднее значение CO_2 в часы пик достигало 1520 ppm, что почти в два раза превышает уровень, зафиксированный при малой загрузке (≈ 850 ppm). Концентрация мелкодисперсных частиц $\text{PM}_{2.5}$ в те же периоды составляла ≈ 48 мкг/м³, что более чем втрое выше показателей при низкой загрузке (≈ 16 мкг/м³). Для PM_{10} наблюдалась аналогичная динамика, при этом пиковые значения превышали рекомендуемые санитарные нормы. Такая зависимость однозначно указывает на прямое влияние пассажиропотока и скорости воздухообмена на качество воздуха: при росте числа пассажиров увеличивается интенсивность дыхания и количество источников загрязнителей, а естественная вентиляция не успевает обеспечить требуемый воздухообмен, что приводит к накоплению CO_2 и частиц.

Для более детального изучения вклада отдельных параметров в формирование качества воздуха был проведён анализ значимости признаков с использованием модели Random Forest (рис. 3). Горизонтальная столбчатая диаграмма показала, что основным предиктором состояния воздуха является концентрация CO_2 , вносящая наибольший вклад в способность модели различать периоды с высокой и низкой нагрузкой. Вторым по значимости фактором оказался уровень $\text{PM}_{2.5}$, также тесно связанный с количеством пассажиров и вентиляцией. Вклад PM_{10} оказался несколько меньшим, но остаётся статистически значимым, отражая присутствие более крупных частиц, которые быстрее оседают на поверхностях и менее чувствительны к кратковременным изменениям потоков воздуха. Температура и влажность также участвуют в классификации, однако их вес в модели заметно ниже, что соответствует физической природе этих параметров: они характеризуют микроклимат, но не являются прямыми индикаторами антропогенного загрязнения.

Сравнительный анализ показал, что Logistic Regression продемонстрировала наименьшую точность (accuracy ≈ 0.77), что связано с её линейной природой и ограниченной способностью моделировать сложные нелинейные зависимости между параметрами качества воздуха. Decision Tree и Random Forest показали значительно более высокие результаты: обе модели превысили порог точности 95%. При этом Decision Tree незначительно опередила Random Forest по accuracy (0.9547 против 0.9524), однако проявила склонность к переобучению, что выражалось в снижении обобщающей способности при классификации редких классов. Наиболее сбалансированные результаты показала модель Random Forest, обеспечив accuracy = 0.9524, precision = 0.6113, recall = 0.6425 и F1 = 0.6257 (табл. 1). Эти показатели указывают на устойчивую способность модели корректно классифицировать временные интервалы с разной нагрузкой и делают её оптимальным выбором для задачи прогнозирования состояния воздушной среды. На рисунке 4 представлена confusion matrix модели Random Forest. Модель демонстрирует высокую точность классификации для класса Low: верно классифицировано 647 наблюдений из 666 (специфичность более 97%). Для класса High модель правильно определила 174 из 182 наблюдений, допустив лишь семь ошибок, отнес их к Low, и одну – к Medium. Класс Medium оказался наиболее сложным для корректной идентификации: все 14 наблюдений были отнесены преимущественно к классу High, что свидетельствует о значительном пересечении признаков между этими состояниями. Несмотря на это, общее распределение предсказаний

подтверждает, что Random Forest уверенно различает периоды низкой и высокой нагрузки, что делает её надёжным инструментом для применения в системах мониторинга качества воздуха в реальном времени.

Таблица 1 – Сравнительные метрики качества моделей классификации

	Accuracy	Precision	Recall	F1
Logistic Regression	0.772622	0.424631	0.335995	0.297632
Random Forest	0.952436	0.611319	0.642505	0.625737
Decision Tree	0.954756	0.611575	0.644837	0.626939
XGBoost	0.954756	0.611575	0.644837	0.626939

Для количественной оценки ошибок классификации была построена матрица ошибок для модели XGBoost (рис. 4). Она показала, что алгоритм с высокой точностью распознаёт интервалы низкой загрязнённости: доля верных классификаций превышает 94%. Наибольшее число ошибок наблюдается при классификации интервалов со средней загрязнённостью: часть таких случаев отнесена к категории «низкие», часть – к «высокие». Это объясняется значительным перекрытием диапазонов концентраций CO_2 и $\text{PM}_{2.5}$ между указанными классами. Для интервалов высокой загрязнённости модель достигает более чем 70% правильных классификаций, что можно считать удовлетворительным результатом для задачи с тремя классами и сравнительно ограниченным объёмом исходных данных.

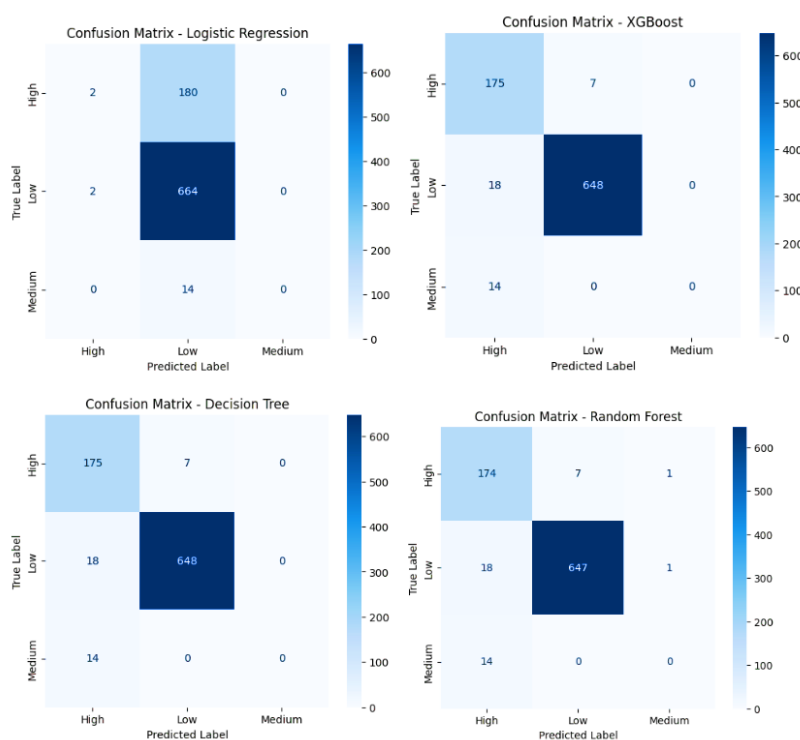


Рисунок 4 – Матрицы ошибок для различных моделей классификации

На рисунке 5 представлена корреляционная матрица, отражающая взаимосвязи между всеми измеренными параметрами. Наиболее сильная зависимость наблюдается между концентрациями PM_{10} и $\text{PM}_{2.5}$ ($r = 0.97$), что подтверждает их общую природу и согласованное поведение в условиях городской транспортной среды. Существенная положительная корреляция выявлена между содержанием CO_2 и относительной влажностью ($r = 0.44$), что связано с накоплением углекислого газа и увеличением влажности при высокой плотности пассажиропотока. Температура, напротив, демонстрирует отрицательную корреляцию с влажностью ($r = -0.39$) и eTVOC ($r = -0.34$), что отражает физико-химические процессы сорбции и испарения органических соединений. Временной фактор (Time) показывает отрицательную зависимость от CO_2 ($r = -0.48$), указывая на тенденцию снижения концентраций в ходе эксперимента благодаря естественному проветриванию и изменению пассажиропотока.

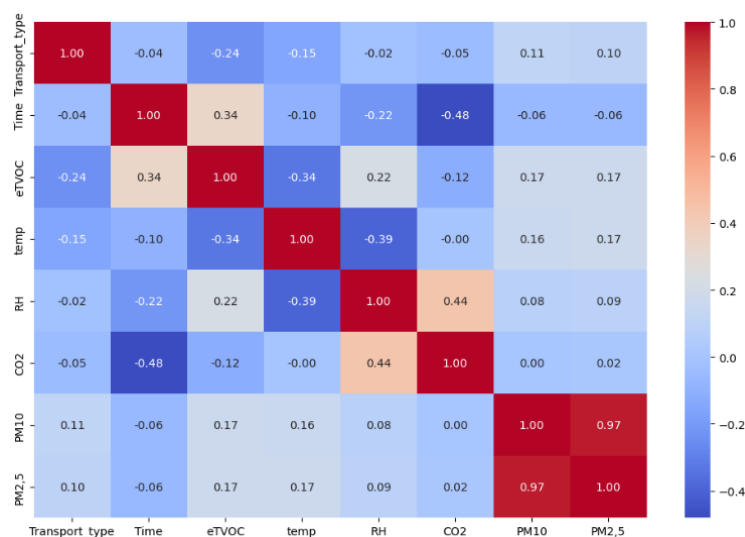


Рисунок 5 – Матрица корреляции признаков

Таким образом, корреляционный анализ позволил выделить ключевые группы взаимосвязанных параметров, наиболее значимых для построения моделей классификации и прогнозирования качества воздуха. Результаты дают основания для нескольких важных выводов. Во-первых, качество воздуха в автобусах Алматы демонстрирует устойчивую зависимость от уровня пассажиропотока, что требует разработки стратегий по оптимизации вентиляции, особенно на наиболее загруженных маршрутах. Во-вторых, параметры CO₂ и PM_{2.5} выступают надёжными индикаторами текущего состояния воздушной среды и могут быть использованы при создании автоматических систем оповещения и регулирования воздухообмена. Наконец, показано, что даже на основе сравнительно простых IoT-узлов возможно построение предсказательных моделей, способных с высокой точностью классифицировать периоды с различной нагрузкой, что открывает путь к интеграции подобных решений в интеллектуальную транспортную инфраструктуру города.

Обсуждение

Наши результаты наглядно показывают, что качество воздуха в автобусах Алматы заметно ухудшается именно в часы пик, а характер загрязнения носит резкий, «пульсирующий» характер, совпадающий с увеличением пассажиропотока. Этот вывод особенно важен для оценки реальных рисков для здоровья, поскольку усреднённые замеры часто скрывают кратковременные, но интенсивные всплески концентраций CO₂ и твёрдых частиц [17, 18]. Выявленные закономерности согласуются с результатами исследований, проведённых в Венгрии и Нигерии, где перенаселённость салонов также признана основным фактором ухудшения качества воздуха [19, 20]. Вклад нашей работы заключается в применении данных с высоким временным разрешением, что позволило отследить изменения параметров в ходе поездки и точно зафиксировать периоды наибольшего воздействия загрязнителей. Такой подход дополняет выводы систематических обзоров и исследований, выполненных в других городах с аналогичной транспортной инфраструктурой, например, в Цзинане [21]. Использование доступного по стоимости IoT-устройства Tynus показало, что подобные технологии способны обеспечить надёжный и оперативный мониторинг качества воздуха при умеренных затратах [15, 22]. Проведённый корреляционный анализ подтвердил ключевую роль параметров CO₂ и PM_{2.5} как наиболее информативных индикаторов, устойчиво связанных с пассажиропотоком и динамикой воздухообмена [23]. Эти параметры не только демонстрируют сильную корреляцию с другими факторами (например, влажностью и временем наблюдений), но и оказываются определяющими в моделях классификации. Сравнительный анализ алгоритмов показал, что традиционная Logistic Regression ограничена в способности выявлять сложные нелинейные зависимости и продемонстрировала наименьшую точность (≈ 0.77) [24]. Более сложные модели, включая Decision Tree, Random Forest и XGBoost, превысили порог 95% accuracy [25, 26]. При этом Random Forest обеспечил наиболее сбалансированные метрики (accuracy = 0.9524, precision = 0.6113, recall = 0.6425, F1 = 0.6257), что подтверждается распределением ошибок в confusion matrix. XGBoost также показал высокую точность, особенно для интервалов низкой

загрязнённости (>94%), хотя при классификации средней загрязнённости возникали ошибки из-за перекрытия диапазонов CO₂ и PM_{2.5}. В целом полученные результаты подтверждают, что даже на основе сравнительно простых IoT-узлов можно строить надёжные предсказательные модели для мониторинга качества воздуха в реальном времени. Тем не менее, исследование имеет ряд ограничений. Данные были собраны только на девяти маршрутах и в двух временных окнах одного месяца (апрель 2025 г.), что не позволяет учесть сезонную и суточную изменчивость загрязнения. Влияние внешних факторов, включая погодные условия, режимы работы вентиляции и возраст подвижного состава, не контролировалось. Кроме того, хотя перед измерениями проводилась калибровка устройства, вопросы долгосрочного дрейфа сенсоров и их чувствительности к температуре и влажности остались за пределами данного анализа. Несмотря на эти ограничения, результаты исследования могут служить основой для практических мер. В краткосрочной перспективе приоритетными являются модернизация вентиляции и фильтрации, оптимизация пассажиропотоков и снижение переполненности салонов в часы пик за счёт корректировки расписания и увеличения числа транспортных единиц. В долгосрочной перспективе целесообразно внедрить системы постоянного мониторинга качества воздуха на наиболее загруженных автобусных маршрутах, что позволит принимать решения, основанные на данных, и снижать риски для здоровья пассажиров. В совокупности это открывает возможности интеграции подобных решений в интеллектуальную транспортную инфраструктуру города и создания систем адаптивного управления вентиляцией на основе IoT и методов машинного обучения.

Заключение

Проведённое исследование стало первым систематическим анализом качества воздуха в автобусах Алматы с высоким временным разрешением. Его результаты подтверждают, что пассажиропоток является ключевым фактором, определяющим динамику загрязнителей в замкнутых транспортных пространствах, а кратковременные пики концентраций указывают на необходимость пересмотра подходов к оценке риска.

Научная значимость работы заключается в создании доступной и масштабируемой IoT-системы мониторинга, интегрируемой в инфраструктуру «умного города» и обеспечивающей непрерывный сбор данных и анализ в реальном времени. Использование методов машинного обучения позволило выявить наиболее надёжные индикаторы загрязнённости и подтвердить эффективность моделей предсказания. Сравнительный анализ алгоритмов показал устойчивость полученных результатов и подтвердил возможность практического применения подхода. При этом исследование имеет определённые ограничения, связанные с выбором маршрутов и временными рамками наблюдений, что требует дальнейшего расширения эксперимента и учёта сезонных и эксплуатационных факторов. Несмотря на это, полученные результаты уже могут служить основой для разработки стратегий по улучшению вентиляции, оптимизации пассажиропотоков и снижению воздействия загрязнителей на здоровье населения, а в долгосрочной перспективе – для внедрения систем интеллектуального мониторинга и управления качеством воздуха в городском транспорте.

Список литературы

1. Systematic Review of Indoor Environmental Quality in Passenger Transport Vehicles of Tropical and Subtropical Regions / J.O. Ogundiran et al // Atmosphere. – 2025. – № 16. – P. 140. https://www.researchgate.net/publication/388417443_A_Systematic_Review_of_Indoor_Environmental_Quality_in_Passenger_Transport_Vehicles_of_Tropical_and_Subtropical_Regions.
2. Environmental Monitoring in Bus Transportation Using a Developed Measurement System / A. Todorov et al // Urban Sci. – 2023. – № 7. – P. 90. <https://doi.org/10.3390/urbansci7030090>.
3. Review of Scientific Technology-Based Solutions for Vehicular Pollution Control. Clean Technol / S. Gulia et al // Environ. Policy. – 2020. – № 22. – P. 1955-1966. <https://doi.org/10.1007/s10098-020-01952-6>.
4. Prototype of Monitoring Transportation Pollution Spikes through the Internet of Things Edge Networks / E. Nizeyimana et al // Sensors. – 2023. – № 23. – P. 8941. <https://doi.org/10.3390/s23218941>.

5. Wang Y.-C. Efficient Data Gathering and Estimation for Metropolitan Air Quality Monitoring by Using Vehicular Sensor Networks / Y.-C. Wang, G.-W. Chen // *IEEE Trans. Veh. Technol.* – 2017. – № 66. – P. 7234-7248. <https://doi.org/10.1109/TVT.2017.2655084>.
6. Ventilation Strategies for Highly Occupied Public Environments: A Review / Y. Zhang et al // *Buildings.* – 2023. – № 13. – P. 1642. <https://doi.org/10.3390/buildings13071642>.
7. CFD Investigation of Vehicle's Ventilation Systems and Analysis of ACH in Typical Airplanes, Cars, and Buses / B. Pirouz et al // *Sustainability.* – 2021. – № 13. – P. 6799. <https://doi.org/10.3390/su13126799>.
8. Innovative High-Induction Air Diffuser for Enhanced Air Mixing in Vehicles and Personalized Ventilation Applications / F.I. Bode et al // *Energies.* – 2024. – № 17. – P. 2930. <https://doi.org/10.3390/en17122930>.
9. Implementation of IoT-Based Air Quality Monitoring System for Investigating Particulate Matter (PM10) in Subway Tunnels / J.H. Jo et al // *Int. J. Environ. Res. Public Health.* – 2020. – № 17. – P. 5429. <https://doi.org/10.3390/ijerph17155429>.
10. Xu B. Air Quality Inside Subway Metro Indoor Environment Worldwide: A Review / B. Xu, J. Hao // *Environ. Int.* – 2017. – № 107. – P. 33-46. <https://doi.org/10.1016/j.envint.2017.06.016>.
11. Numerical Analysis of Air Quality Improvement and Thermal Comfort in a Classroom Using Organic Air Cleaners / E. Cruz-Octaviano et al // *Atmosphere.* – 2025. – № 16. – P. 727. <https://doi.org/10.3390/atmos16060727>.
12. The Impact of Urban Transportation Infrastructure on Air Quality / Y. Guo et al // *Sustainability.* – 2020. – № 12. – P. 5626. <https://doi.org/10.3390/su12145626>.
13. Exposure to PM_{2.5} on Public Transport: Guidance for Field Measurements with Low-Cost Sensors / K. Fameli et al // *Atmosphere.* – 2024. – № 15. – P. 330. <https://doi.org/10.3390/atmos15030330>.
14. South Coast Air Quality Management District (SCAQMD). Field Evaluation: Qingping Air Monitor; SCAQMD: Diamond Bar, CA, USA, 2023. Available online: <https://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/qingping---air-monitor---field-evaluation.pdf> (accessed on 9 June 2025).
15. Long-Term Field Evaluation of the Plantower PMS Low-Cost Particulate Matter Sensors / T. Sayahi et al // *Environ. Pollut.* – 2019. – № 245. – P. 932-940. <https://doi.org/10.1016/j.envpol.2018.11.065>.
16. Study on Dynamic Load of Air-Conditioning System in Subway Station Based on Hourly Passenger Flow / L. Wang et al // *Buildings.* – 2023. – № 13. – P. 2349. <https://doi.org/10.3390/buildings13092349>.
17. Urban Air Quality Comparison for Bus, Tram, Subway and Pedestrian Commutes in Barcelona / T. Moreno et al // *Environ. Res.* – 2015. – № 142. – P. 495-510.
18. How Can Ventilation Be Improved on Public Transportation Buses? Insights from CO₂ Measurements / X. Querol et al // *Environ. Res.* – 2022. – № 205. – P. 112451. <https://doi.org/10.1016/j.envres.2021.112451>.
19. Case Studies of Aerosol Pollution in Different Public Transport Vehicles in Hungarian Cities. / E. Papp et al // *Atmosphere.* – 2022. – № 13. – P. 692. <https://doi.org/10.3390/atmos13050692>.
20. Indoor Environmental Quality Assessment of Train Cabins and Passenger Waiting Areas: A Case Study of Nigeria / J.O. Ogundiran et al // *Sustainability.* – 2023. – № 15. – P. 16533. <https://doi.org/10.3390/su152316533>.
21. Change in Air Quality during 2014-2021 in Jinan City in China and Its Influencing Factors. *Toxics.* – 2023. – № 11. – P. 210. <https://doi.org/10.3390/toxics11030210>.
22. Kaivonen S. Real-time Air Pollution Monitoring with Sensors on City Bus / S. Kaivonen, E.C.-H. Ngai // *Digit. Commun. Netw.* – 2020. – № 6. – P. 23-30. <https://doi.org/10.1016/j.dcan.2019.03.003>.
23. Indoor Air Quality in Subway Microenvironments: Pollutant Characteristics, Adverse Health Impacts, and Population Inequity / S. Wang et al // *Environ. Int.* – 2024. – № 190. – P. 108873. <https://doi.org/10.1016/j.envint.2024.108873>.
24. Development and Validation of a Logistic Regression Model for the Diagnosis of Colorectal Cancer / C. Li et al // *Sci. Rep.* – 2025. – № 15. – P. 14759. <https://doi.org/10.1038/s41598-025-14759-3>.

25. Predictive Modeling of Air-Quality Levels Using Decision Tree Classification: Insights from Environmental and Demographic Factors / I.G. Sudipa et al // Indones. J. Data Sci. – 2024. – № 5. – P. 251-258.
26. Prediction of Cardiovascular Disease Based on Multi-Feature Selection and XGBoost Optimisation / H. Zhang et al // Sci. Rep. – 2025. – № 15. – P. 96520. <https://doi.org/10.1038/s41598-025-96520-7>.

Финансирование

Данное исследование финансируется Комитетом науки Министерства науки и высшего образования Республики Казахстан (Грант ИРН BR24993051).

А.Н. Молдагулова¹, Д.Н. Тұрмаханбет^{2,3*}, М.А. Мауленбеков^{1,3}, Ж.Б. Кальпеева¹, Р.К. Ускенбаева¹

¹Қ.И. Сатпаев атындағы Қазақ ұлттық техникалық зерттеу университеті,
050013, Қазақстан Республикасы, Алматы қ., Сатпаев көшесі 22

²Әл-Фараби атындағы Қазақ ұлттық университеті,
050040, Қазақстан Республикасы, Алматы қ., әл-Фараби даңғылы, 71

³ТОО «Kazakhstan R&D Solutions»,
050056, Қазақстан Республикасы, Алматы қ., Кожедуба көшесі, 3

*email: turmahanbetdinara@gmail.com

АЛМАТЫ АВТОБУСТАРЫНДАҒЫ АУАНЫҢ САПАСЫН ІОТ-ҚҰРЫЛҒЫЛАР МЕН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНА ОТЫРЫП МОНИТОРИНГТЕУ

Қоғамдық көліктегі ауаның сапасы – қала тұрғындарының денсаулығына тікелей әсер ететін маңызды фактор, алайда Алматы үшін бұл мәселе бүгінгі күнге дейін жеткілікті деңгейде зерттелмеген. Осы зерттеуде қаланың үш негізгі көлік дәлізіндегі автобустарда CO_2 , $PM_{2.5}$, PM_{10} , температура және салыстырмалы ылғалдылық концентрациялары жүйелі түрде бақыланды. Мәліметтерді жинау жоғары сенімділікті қамтамасыз еткен, коммерциялық эталонды сенсорға калибрленген «Tynys» атты мобильді ІОТ-құрылғының көмегімен жүзеге асырылды. Негізгі назар жолаушылар көп болатын және аз болатын уақыттарды салыстыруға аударылды, бұл көліктің жүктелу деңгейіне байланысты ауаның сапасының өзгеру заңдылықтарын анықтауға мүмкіндік берді. Талдау нәтижелері таңғы қарбалас уақытта CO_2 мен ұсақ дисперсті бөлшектер концентрациясы айтарлықтай жоғары болатынын көрсетті, бұл жолаушылар ағынының салондағы ауаның сапасын қалыптастырудағы басты рөлін дәлелдейді. Қосымша түрде машиналық оқыту әдістерін қолдану арқылы CO_2 мен $PM_{2.5}$ ең ақпаратты индикаторлар екендігі, яғни жоғары жүктелген кезеңдерді сенімді түрде жіктеуге мүмкіндік беретіні анықталды. Алынған нәтижелер Орталық Азия аймағындағы қоғамдық көліктегі ауаның сапасы жөніндегі ғылыми олқылықты толықтырады және желдету жүйелерін жаңғыртуға, жолаушылар ағынын оңтайландыруға және автобустарда микроклиматты интеллектуалды бақылау технологияларын енгізуге бағытталған болашақ зерттеулерге негіз қалайды.

Түйін сөздер: қоғамдық көлік; ауаның сапасы; CO_2 ; $PM_{2.5}$; PM_{10} ; ІОТ-құрылғы; машиналық оқыту; желдету; жолаушылар ағыны; микроклимат; интеллектуалды көлік

A.N. Moldagulova¹, D.N. Turmakhanbet^{2,3*}, M.A. Maulenbekov^{1,3}, Zh.B. Kalpeyeva¹, R.K. Uskenbayeva¹

¹Satbayev University, 050013, Republic of Kazakhstan, Almaty, 22 Satpayev str.

²Al-Farabi Kazakh National University, Republic of Kazakhstan, Almaty
050040, Republic of Kazakhstan, Almaty, 71 Al-Farabi Avenue

³LLP «Kazakhstan R&D Solutions»,
050056, Republic of Kazakhstan, Almaty, 3 Kozheduba str.

*email: turmahanbetdinara@gmail.com

AIR QUALITY MONITORING IN ALMATY BUSES USING IOT DEVICES AND MACHINE LEARNING METHODS

Air quality in public transport is a critical factor directly affecting urban health, yet for Almaty this issue has remained understudied until now. In this study, a systematic monitoring of CO_2 , $PM_{2.5}$, PM_{10} , temperature, and relative humidity concentrations was conducted in buses across three key transport corridors of the city. Data collection was carried out using a mobile IoT device of our own design, «Tynys», calibrated against a commercial reference sensor to ensure high measurement reliability. Special attention was paid to comparing peak and off-peak hours, which allowed us to identify patterns of air quality changes under varying passenger

loads. The analysis showed that during morning peak hours, CO₂ and fine particulate matter concentrations significantly exceeded off-peak levels, confirming the key role of passenger flow in shaping in-vehicle air quality. Additionally, by applying machine learning methods, it was established that CO₂ and PM_{2.5} are the most informative indicators of pollution, enabling reliable classification of high-load periods. The obtained results fill an existing research gap in studies of air quality in public transport in Central Asia and provide a basis for further work aimed at modernizing ventilation systems, optimizing passenger flows, and implementing intelligent microclimate monitoring technologies in buses.

Key words: public transport; air quality; CO₂; PM_{2.5}; PM₁₀; IoT device; machine learning; ventilation; passenger flow; microclimate; intelligent transport

Авторлар туралы ақпарат

Айман Николаевна Молдагулова – доцент, профессор-зерттеуші, Қ.И. Сатпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы; e-mail: a.moldagulova@satbayev.university. ORCID: <https://orcid.org/0000-0002-1596-561X>.

Динара Нұрланқызы Тұрмаханбет* – бакалавр 4 курс студенті, Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы; e-mail: turmahanbetdinara@gmail.com.

Марлен Амантайұлы Мауленбеков – бакалавр 4 курс студенті, Қ.И. Сатпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы; e-mail: m.maulenbekov@stud.satbayev.university.

Жулдыз Бейшеналиевна Кальпеева – Ph.D, доцент, Қ.И. Сатпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы; e-mail: z.kalpeyeva@satbayev.university. ORCID: <https://orcid.org/0000-0002-4970-3095>.

Раиса Кабиевна Ускенбаева – доктор технических наук, профессор, Қ.И. Сатпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы; e-mail: r.k.uskenbayeva@satbayev.university. ORCID: <https://orcid.org/0000-0002-8499-2101>.

Сведения об авторах

Айман Николаевна Молдагулова – доцент, профессор-исследователь, Казахский национальный технический исследовательский университет имени К.И. Сатпаева, Алматы; e-mail: a.moldagulova@satbayev.university. ORCID: <https://orcid.org/0000-0002-1596-561X>.

Динара Нұрланқызы Тұрмаханбет* – студентка 4 курса бакалавриата, Казахский национальный университет имени аль-Фараби, Алматы; e-mail: turmahanbetdinara@gmail.com.

Марлен Амантайұлы Мауленбеков – студент 4 курса бакалавриата, Казахский национальный технический исследовательский университет имени К.И. Сатпаева, Алматы; e-mail: m.maulenbekov@stud.satbayev.university.

Жулдыз Бейшеналиевна Кальпеева – Ph.D, доцент, Казахский национальный технический исследовательский университет имени К.И. Сатпаева, Алматы; e-mail: z.kalpeyeva@satbayev.university. ORCID: <https://orcid.org/0000-0002-4970-3095>.

Раиса Кабиевна Ускенбаева – доктор технических наук, профессор, Казахский национальный технический исследовательский университет имени К.И. Сатпаева, Алматы; e-mail: r.k.uskenbayeva@satbayev.university. ORCID: <https://orcid.org/0000-0002-8499-2101>.

Information about the authors

Aiman Nikolaevna Moldagulova – Associate Professor, Research Professor, Satbayev University, Almaty, Kazakhstan; e-mail: a.moldagulova@satbayev.university. ORCID: <https://orcid.org/0000-0002-1596-561X>.

Dinara Nurlankyzy Turmakhanbet* – 4th year Bachelor's student, Al Farabi Kazakh National University, Almaty, Kazakhstan; e-mail: turmahanbetdinara@gmail.com.

Marlen Amantayuly Maulenbekov – 4th year Bachelor's student, Satbayev University, Almaty, Kazakhstan; e-mail: m.maulenbekov@stud.satbayev.university.

Zhuldyz Beishenalievna Kalpeyeva – Ph.D, Associate Professor, Satbayev University, Almaty, Kazakhstan; e-mail: z.kalpeyeva@satbayev.university. ORCID: <https://orcid.org/0000-0002-4970-3095>.

Raisa Kabievna Uskenbayeva – Doctor of Technical Sciences, Professor, Satbayev University, Almaty, Kazakhstan; e-mail: r.k.uskenbayeva@satbayev.university. ORCID: <https://orcid.org/0000-0002-8499-2101>.

Поступила в редакцию 20.09.2025
Поступила после доработки 28.11.2025
Принята к публикации 05.01.2026