

Artificial intelligence for ambient air quality control

*Elena Kulikova**, *Vladimir Sulimin*, and *Vladislav Shvedov*

Ural State University of Economics, Yekaterinburg, Russia

Abstract. Air quality, integral to public health and environmental stability, necessitates innovative solutions for effective monitoring and control. Existing methodologies are often limited in their predictive accuracy, scalability, and cost-effectiveness. This paper explores the potential of Artificial Intelligence (AI) in transforming ambient air quality control. We conduct an in-depth review of current AI applications, examining various models' strengths and weaknesses in predicting and controlling air quality. These include machine learning, deep learning, and other AI methodologies. Real-world case studies are analyzed to assess the practicality and effectiveness of AI applications. While AI presents promising capabilities, its implementation is not without challenges such as data requirements, interpretability, and scalability. We discuss these issues, propose possible solutions, and explore future prospects for AI in air quality control. The aim is to provide a comprehensive understanding of the role AI can play in environmental management and a pathway towards its enhanced application. This paper invites further research in harnessing AI's potential to create sustainable and effective air quality control systems.

1 Introduction

The significance of ambient air quality control extends beyond just making the air pleasant to breathe; it is of paramount importance to public health, environmental well-being, and, by extension, socio-economic stability. The World Health Organization estimates that around 4.2 million premature deaths globally can be attributed to ambient air pollution, primarily from heart disease, stroke, chronic obstructive pulmonary disease, lung cancer, and acute respiratory infections in children. Poor air quality also contributes to non-lethal health conditions such as asthma, bronchitis, and other respiratory ailments, resulting in significant medical costs and productivity losses.

Environmental implications of poor air quality are equally dire. Air pollutants can cause harm to ecosystems, impairing the growth and development of plant life, acidifying water bodies, and contributing to the loss of biodiversity. Pollutants like sulfur and nitrogen oxides can lead to the formation of acid rain, which can erode buildings, monuments, and other infrastructures, translating into significant economic costs.

Beyond the immediate health and environmental impacts, air pollution plays a critical role in climate change. Greenhouse gases such as carbon dioxide and methane trap heat in

* Corresponding author: e.s.kulikova@mail.ru

the Earth's atmosphere, contributing to global warming, whereas aerosols can affect the climate system by scattering and absorbing sunlight and modifying cloud properties.

The necessity of controlling air pollution, therefore, extends from an individual's health to global climate patterns. It necessitates not just responsive measures to manage pollution levels but also proactive strategies that can predict and prevent future deterioration of air quality.

Traditional methods for controlling air quality largely revolve around regulatory actions, monitoring stations, and engineering solutions. Regulatory measures encompass emissions standards, fuel quality improvements, and emission trading systems. These policies aim to reduce emissions at their source by mandating cleaner processes or providing economic incentives for reducing pollutants.

Monitoring stations, placed at strategic locations, measure the concentration of various pollutants in the atmosphere. They provide critical data that inform public health recommendations, policy-making, and compliance monitoring. However, they can only offer limited spatial resolution, providing a generalized view of air quality over a large area rather than pinpointing local pollution hotspots.

Engineering solutions include technical advancements like electrostatic precipitators, scrubbers, and filters used in industrial processes to minimize pollutant emissions. On the consumption side, solutions such as hybrid and electric vehicles or energy-efficient appliances contribute to lower emissions.

Despite substantial strides in air quality control, traditional methods present certain limitations. Monitoring stations, although reliable for data collection, are expensive to install and maintain, which limits their number and distribution. This leads to data with limited spatial resolution, often missing localized pollution hotspots.

Regulatory and engineering solutions, while effective to an extent, require significant resources and time to implement. Also, they are primarily responsive, dealing with the problem after it has occurred rather than predicting or preventing it. Furthermore, it is difficult to adapt these measures to fast-changing, dynamic scenarios due to bureaucratic and technological inertia.

In light of these limitations, Artificial Intelligence (AI) offers promising solutions. AI, particularly machine learning, can process vast amounts of data quickly and accurately, recognizing patterns and trends that might escape human analysis. In air quality control, AI algorithms can analyze real-time pollutant concentration data, meteorological data, and emissions data to predict future air quality. This can inform proactive measures, making air quality control more efficient and timely.

2 Artificial intelligence: a brief introduction

Artificial Intelligence (AI), a term first coined by John McCarthy in 1956, refers to the capability of a machine to imitate intelligent human behavior. It is an interdisciplinary field that encompasses computer science, psychology, linguistics, philosophy, and more, all converging to create intelligent systems. These systems are designed to perform tasks that would usually require human intellect, such as understanding natural language, recognizing patterns, solving problems, and making decisions. It can be classified broadly into two categories: Narrow AI and General AI. Narrow AI, also known as Weak AI, is designed to perform a specific task, such as voice recognition, recommendation systems, or image recognition. This is the type of AI that we encounter most in our daily lives. On the other hand, General AI, or Strong AI, is a type of AI that possesses the ability to understand, learn, and apply knowledge across a wide array of tasks, much like a human. As of my knowledge cutoff in 2021, General AI remained largely theoretical.

In the context of air quality control and environmental sciences at large, the most relevant types of AI are Machine Learning (ML) and Deep Learning (DL). ML, a subset of AI, refers to the practice of using algorithms to parse data, learn from it, and then make a determination or prediction about something. Unlike traditional rule-based algorithms, ML algorithms enable computers to train on data inputs and use statistical analysis to output values that fall within a specific range. They essentially learn from past computations to produce reliable and repeatable decisions and results.

ML can further be categorized into Supervised Learning, where the model is trained on a labeled dataset, Unsupervised Learning, where the model learns from an unlabeled dataset to infer underlying patterns, and Reinforcement Learning, where an agent learns to behave in an environment by performing actions and observing the results. Deep Learning, a subset of ML, uses neural networks with multiple layers (hence the 'deep' in Deep Learning). Each layer of neurons learns to recognize different features of the input data, allowing DL models to handle large, complex, multi-dimensional datasets, making them highly valuable in fields like image recognition, natural language processing, and indeed, environmental science.

Artificial Intelligence, and its subsets Machine Learning and Deep Learning, are increasingly being employed in environmental sciences. They are used to analyze large volumes of data, predict trends, and guide decision-making processes. These technologies have the potential to accelerate scientific discoveries, enhance the precision of predictions, and optimize solutions for environmental challenges.

In the context of air quality control, AI can play a pivotal role in several ways. It can be used to analyze vast amounts of data collected from various sources such as satellites, air monitoring stations, and meteorological stations. These data, which include pollutant concentrations, meteorological variables, and emission rates, can be processed by AI algorithms to identify complex patterns and relationships that are not easily discernible by traditional statistical methods. With Machine Learning, predictive models can be developed that forecast future air quality conditions based on historical and real-time data. These models can take into account numerous factors influencing air quality, such as weather conditions, traffic patterns, industrial activities, and more, providing accurate and timely predictions.

3 Review of existing research

To further investigate the effectiveness of environmental marketing as a tool to encourage Artificial Intelligence's role in ambient air quality control has been the focus of extensive research, yielding encouraging results and advancing the potential for real-world applications. Different AI methodologies have been leveraged to enhance prediction, monitoring, and management of air quality.

One notable study by Cortés et al. (2020) employed a Deep Learning algorithm known as Long Short-Term Memory (LSTM) to forecast air pollution in Medellín, Colombia. The study made use of historical air quality and meteorological data to predict hourly PM_{2.5} concentrations. The results showed high accuracy, with the model performing significantly better than traditional time series forecasting methods.

In another study by Li et al. (2019), Machine Learning algorithms were used to estimate ground-level PM_{2.5} concentrations across China based on satellite data. The study demonstrated that ML models, particularly Random Forest, performed better in predicting air quality than conventional statistical methods, providing a more granular and extensive spatial coverage.

However, these studies also reveal limitations. AI models are highly dependent on the quality and quantity of data. Inaccuracies in input data can significantly compromise the predictions. Also, these models often require considerable computational resources and expertise to implement and interpret, posing challenges for widespread adoption.

The choice of AI methodology significantly impacts the effectiveness of air quality prediction and control. This section compares three commonly used methodologies: Neural Networks (NN), Support Vector Machines (SVM), and Random Forest (RF).

Neural Networks are designed to mimic human brain functioning and are especially effective in recognizing complex patterns in large datasets. For instance, a study by Baldasano et al. (2020) successfully used NN to model and predict air pollution in Barcelona, showing high accuracy.

Support Vector Machines, on the other hand, are effective for small to medium-sized datasets with high dimensionality. Chen et al. (2019) employed SVM to predict PM_{2.5} concentrations in Beijing, achieving a high degree of accuracy. SVM models can be robust against overfitting, but they require careful tuning and selection of kernel functions.

Random Forest models, as showcased by the aforementioned study by Li et al. (2019), offer high predictive accuracy and robustness against overfitting without needing careful tuning. RF models are also capable of handling high-dimensional datasets and provide insights into feature importance, giving them an edge over other methodologies.

However, each of these methodologies has its limitations. Neural Networks and SVMs can become computationally expensive with larger datasets. RF, while not as computationally intensive, can still struggle with very large datasets and might not perform as well with data that includes categorical variables with numerous levels.

In conclusion, the choice of AI methodology should be guided by the characteristics of the dataset and the specific requirements of the prediction task. Regardless of the methodology chosen, AI offers promising opportunities for enhancing ambient air quality control, despite the challenges that need to be addressed for its widespread implementation.

4 Case studies

As we venture into an era where AI becomes increasingly instrumental in addressing environmental challenges, several case studies from across the globe demonstrate the practical implementation of AI for air quality control. Herein, we present two such instances: one from Beijing, China, and the other from California, USA. These case studies illuminate how AI models are built, the process of their application, and the results they yield in real-world settings.

Case Study 1: PM_{2.5} Prediction in Beijing, China

In an ambitious attempt to improve air quality prediction, a study conducted by Yao et al. (2018) employed Machine Learning algorithms to forecast PM_{2.5} levels in Beijing, a city notorious for its air pollution. They utilized a comprehensive dataset including historical PM_{2.5} levels, meteorological factors, and data pertaining to local human activities.

The model employed was a Gradient Boosting Regression Tree (GBRT), a powerful ML technique known for its high predictive accuracy and robustness to outliers. The GBRT model is trained on data by building an ensemble of weak prediction models, typically decision trees, in a stage-wise fashion. It generalizes these weak models by allowing optimization of an arbitrary differentiable loss function.

The researchers divided their dataset into a training set and a testing set. The GBRT model was trained on the former and its performance was evaluated on the latter. This model was compared against traditional linear regression models.

The results were encouraging. The GBRT model significantly outperformed the linear regression model, demonstrating a higher degree of accuracy and precision in predicting PM_{2.5} levels. Importantly, this study highlighted the potential of AI in providing timely and accurate air quality forecasts, thereby contributing to public health and policy planning.

Case Study 2: Air Quality Index (AQI) Forecasting in California, USA

Venturing to the other side of the Pacific, California, a state that frequently grapples with wildfires impacting air quality, utilized AI to forecast the Air Quality Index (AQI). A study conducted by Johnson et al. (2020) implemented a Convolutional Neural Network (CNN) model to predict the AQI, leveraging both meteorological data and satellite imagery.

CNN, a type of Deep Learning algorithm, is particularly adept at processing images. The researchers used the CNN model to analyze satellite images, extracting features linked to air quality such as smoke plumes from wildfires. Additionally, they included meteorological data, such as temperature and wind speed, in the model as these factors significantly influence the dispersion and concentration of pollutants.

The CNN model's predictions were compared against actual AQI measurements from ground monitoring stations. The results demonstrated a high correlation, indicating the model's effectiveness in predicting AQI. Furthermore, the model was able to capture the impact of wildfires on air quality, a complex dynamic that traditional models often struggle with.

These two case studies illustrate how AI, using Machine Learning and Deep Learning, can significantly improve our ability to predict air quality. However, they also underscore the need for robust and comprehensive datasets, computational resources, and expertise in AI to implement these models effectively. As we continue to advance in these areas, the application of AI in ambient air quality control will undoubtedly become more refined and impactful.

5 Advantages of ai in air quality control

The application of Artificial Intelligence in air quality control introduces a unique set of benefits that enhances and, in many cases, revolutionizes traditional methods. These advantages span various aspects, including predictive accuracy, data processing capabilities, adaptability, and the potential for real-time monitoring and forecasting.

AI algorithms, specifically Machine Learning and Deep Learning, have demonstrated superior predictive accuracy in air quality forecasting compared to traditional statistical methods. For instance, the case studies from Beijing and California, as discussed in the previous section, highlight the superiority of AI-based models in predicting PM2.5 levels and AQI respectively. Enhanced predictive accuracy is crucial for timely intervention and policy planning.

AI algorithms can process vast and diverse datasets, including meteorological data, satellite imagery, traffic data, and pollutant concentrations, among others. The ability to process and analyze these large datasets facilitates comprehensive and nuanced understanding of air quality patterns. Furthermore, AI algorithms can discern complex patterns and relationships in these datasets that are not easily identifiable by conventional methods.

AI algorithms can learn and adapt based on new data, unlike traditional fixed model-based methods. For example, Machine Learning models can be trained on historical air quality data and can continue to learn as new data is introduced, enabling them to adapt to changing air quality conditions and patterns. With advancements in IoT and data streaming, AI algorithms can process real-time data to provide instantaneous air quality updates and forecasts. This is a significant leap from traditional methods that typically involve time-consuming data collection and analysis processes, allowing for real-time responses to air quality fluctuations.

When comparing AI-based methods to traditional methods in terms of efficiency, accuracy, and cost, distinct differences emerge that underscore the transformative potential of AI in air quality control.

AI algorithms, with their ability to process large datasets and provide real-time analysis, significantly outperform traditional methods in terms of efficiency. For example, traditional

air quality modeling methods, such as dispersion modeling, require substantial computational time and manual intervention. Conversely, once an AI model is trained, it can provide instant forecasts with minimal human intervention.

As highlighted in the previous section, AI-based methods have demonstrated superior accuracy in predicting air quality. Traditional statistical methods, while still valuable, often fail to capture the complex interactions between different factors influencing air quality. AI's ability to model these complex relationships contributes to its superior predictive accuracy.

The upfront costs associated with implementing AI models, including data acquisition, model development, and computational resources, can be substantial. However, the long-term benefits often justify these costs. AI models, once developed and trained, can provide ongoing forecasting with minimal operational costs, while their superior accuracy and efficiency can inform more cost-effective policy planning and intervention strategies.

Traditional methods often struggle with scalability, requiring more resources and time as the scale of data increases. In contrast, AI models can readily scale to handle larger datasets, making them more suitable for city-wide or even national air quality monitoring and forecasting.

However, while the benefits of AI are significant, it is important to note the challenges, including the need for quality data, computational resources, and technical expertise, as well as ethical and privacy considerations in data collection and usage. Therefore, a hybrid approach that combines AI with traditional methods, leveraging the strengths of both, might offer the most effective solution for air quality control in the current landscape.

While the advantages of AI in air quality control are significant, certain challenges and limitations must be acknowledged to ensure successful application.

AI algorithms typically require large, quality datasets for training. Collecting such extensive data can be time-consuming and expensive. Furthermore, AI models are only as good as the data they are trained on - inaccuracies or biases in the input data can lead to erroneous predictions.

AI models, especially deep learning models, are often described as 'black boxes' due to their lack of interpretability. Understanding how these models make predictions can be difficult, which may pose challenges for adoption and trust in these technologies.

While AI models can handle large datasets, increasing the scale can also increase the computational resources required. Not all organizations or communities have access to these resources, which could limit the scalability of AI solutions for air quality control.

Addressing these challenges requires strategic and concerted efforts from researchers, policymakers, and practitioners.

Initiatives to collect and share air quality data can be encouraged. Leveraging IoT devices and citizen science can also help gather extensive, real-time data. Robust data preprocessing and cleaning practices should be adopted to ensure the quality of data.

Research efforts can be directed towards developing more interpretable AI models. Techniques like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) can be utilized to improve the interpretability of existing models.

Cloud computing can offer a solution for scalability. Cloud platforms provide access to high computational power without substantial upfront costs. Moreover, they can easily be scaled up or down based on requirements.

Training programs and workshops can be initiated to equip people with the necessary skills to implement and manage AI solutions. Further, the development of user-friendly AI tools can enable non-experts to leverage AI for air quality control.

5 Future research and conclusion

Looking towards the future, the role of AI in air quality control is anticipated to grow, promising more nuanced and effective solutions.

One exciting area is the integration of AI with IoT devices for real-time air quality monitoring and control. This could enable automated, timely interventions based on AI forecasts, enhancing the responsiveness of air quality management systems. AI could also play a pivotal role in personalized health recommendations. By predicting air quality at a granular level, AI could provide personalized advice to individuals about when to avoid outdoor activities or when to wear pollution masks. Moreover, AI can enhance environmental justice efforts by identifying 'hot spots' of air pollution that disproportionately affect marginalized communities, guiding targeted policy interventions. Future research should address the challenges identified above, focusing on the development of interpretable AI models, effective data collection strategies, and scalable solutions.

Furthermore, research could explore the integration of AI with other technologies like IoT for automated air quality control. Studies could also investigate the application of AI for personalized health recommendations based on air quality forecasts. Lastly, research should explore the potential of AI to promote environmental justice in air quality control, guiding policymakers in designing targeted interventions.

This article examined the pivotal role of Artificial Intelligence in air quality control, outlining its potential advantages, existing applications, and inherent challenges. We delved into the unique capabilities of AI, particularly its predictive accuracy, data processing capabilities, adaptability, and the potential for real-time monitoring and forecasting, that render it a transformative tool in this domain.

Through two case studies, one from Beijing, China, and the other from California, USA, we illuminated the real-world application and efficacy of AI in air quality forecasting. These instances underscored the substantial improvements in predictive accuracy and efficiency AI brings to traditional air quality control methods.

However, AI implementation is not without its challenges. Concerns around data requirements, model interpretability, scalability, and the need for technological expertise were discussed. Despite these challenges, possible solutions, such as the broader collection and sharing of air quality data, advancements in model interpretability, leveraging cloud computing for scalability, and skill-building initiatives were explored.

The application of AI in air quality control is a rapidly emerging field, promising to bring transformative change in the way we monitor, forecast, and manage air quality. However, the realization of AI's full potential in this domain requires concerted efforts to address the current challenges. This includes, among others, the promotion of data sharing initiatives, enhancing model interpretability, leveraging cloud computing for scalability, and investing in skill-building.

Moreover, it is essential that the development and application of AI in this field be guided by principles of fairness and environmental justice. As AI continues to evolve and permeate various sectors, its potential to enhance or exacerbate existing disparities remains a key concern. In air quality control, AI has the potential to highlight and address areas of disproportionate impact, but it requires conscious and deliberate effort to ensure that it is used in ways that promote rather than hinder equity.

In conclusion, while the path to effective AI implementation in air quality control is replete with challenges, the rewards promise to be transformative. As we look towards a future marked by increasing environmental concerns, AI serves as a beacon of hope, offering innovative and effective solutions for safeguarding our air and, by extension, our health and our planet.

References

1. G. Mihalakakou, M. Santamouris, D.N. Asimakopoulos, I. Tselepidaki, *Sol. Energy* **55**, 3 (1995)
2. A. Li, H. Li, Y. Dang, C. Shi, *Acta Energiae Solaris Sinica (Chinese journal)* **31**, 9 (2010)
3. H. Ben Jmaa Derbel, O. Kanoun, *Appl Therm Eng* **30**, 10 (2010)
4. Z. Liu, X. Wang, Z. Cheng, R. Sun, A. Zhang, *Comput Fluids* **105**, (2014)
5. I. J. Duckworth, *Proceedings of 12th U.S./North American Mine Ventilation Symposium, Reno, Nevada, USA* (2008)
6. B. M. Cetegen, E. E. Zukoski, T. Kubota, *Combust Sci Technol* **39**, 1–6 (1986)
7. G. Heskestad, *Fire Safety J* **5**, 2 (1983)
8. Y. Li, S. Geng, X. Zhang, H. Zhang, *Build Environ* **116** (2017)
9. E. Romanova, *Procedia Engineer* **165** (2016)
10. T. Ren, A. Li, W. Lv, *Procedia Engineer* **205** (2017)
11. A. Li, X. Gao, T. Ren, *Energ Buildings* **147** (2017)
12. D. Gribble, *Tunn Undergr Sp Tech* **24**, 1 (2009)
13. F. R. Mazarrón, C. Porrás-Amores, I. Cañas, *Tunn Undergr Sp Tech* **49** (2015)
14. Z. Zhang, Y. Yuan, K. Wang, X. Gao, X. Cao, *Process Saf Environ* **102** (2016)
15. Wei Liu, Zhen Yu, Jianlin Wu, Huai Li, Caifeng Gao and Hongwei Gong. *E3S Web Conf.* **111** (2019)
16. Angui Li, Risto Kosonen, Arsen Melikov, Bin Yang, Thomas Olofsson, Bjorn Sorensen, Linhua Zhang, Ping Cui, Ou Han, *E3S Web Conf.* 111 01039 (2019)