Research Output Journal of Engineering and Scientific Research 3(1):101-104, 2024

ROJESR Publications **ISSN: 1115-6155**

https://rojournals.org/roj-engineering-and-scientific-research/

Urban Air Quality and Machine Learning

Ugwu Chinyere N., Val Hyginus Udoka Eze and Ogenyi Fabian C.

Department of Publication and Extension, Kampala International University Uganda

ABSTRACT

Urbanization has led to significant air quality degradation due to increased emissions from industrial activities, fossil fuel combustion, and vehicular traffic. This degradation poses severe health risks, including respiratory and cardiovascular diseases. Predicting and monitoring urban air quality is complex, influenced by both human activities and natural processes. Machine learning (ML) techniques offer promising solutions for modeling and predicting air quality metrics. This paper reviews the current state of urban air quality monitoring, explores the applications of ML in this field, and discusses the associated challenges and limitations. It highlights successful case studies and suggests future research directions to improve urban air quality monitoring using advanced ML techniques.

Keywords: Urban air quality, machine learning, air pollution, predictive modeling, health impact.

INTRODUCTION

Urbanization has pushed the world's population towards residing in urban regions, catering to their increasing needs for materials and energy. There are also consequences of this meteoric pace of urbanization. The rise in the emission of industrial effluents, burning of fossil fuels, and the exponential growth in the usage of vehicles are also leading to a burgeoning deterioration of air quality in the cities over the past few decades. This increase in pollution directly translates to damage in lung function, a rise in the rate of childhood asthma exacerbation, and a rise in the susceptibility of asthmatics to develop chronic obstructive pulmonary disease. Cognitive function has also been seen to decrease as outdoor PM2.5 increases. Extremely premature mortality is seen due to atmospheric pollutants and causes and diseases like strokes, heart diseases, chronic obstructive pulmonary diseases, pneumonia, lung cancer, mesothelial cancer, and asthma [1, 2]. Predictive analysis of air quality is difficult. Air quality is affected by actions of humans and what nature performs as an autopilot. Both entities can produce or not produce an action at any time of the day and without any prior warning. In recent years, it has been increasingly popular to harness artificial intelligence to model and predict complex processes in diverse domains. In particular, for urban air quality, there has been an increasing trend to use AI-based prediction models, but are limited to primarily regression algorithms. A survey of the different AI-based techniques used in both urban and rural areas conducted by Mohan et al., shows that urban areas primarily rely on technical tools that employ machine learning. In addition, studies have been conducted to model spatial and temporal predictability of air quality data. They have focused on studying spatial and temporal predictability of air quality metrics and have used a variety of machine learning tools like Random Forests, Neural Networks, Support Vector Machines, and Kriging [3].

IMPORTANCE OF URBAN AIR QUALITY MONITORING

Urban air quality monitoring is a key issue for a wide number of applications. In this introductory section, we provide some information about the reasons why good air quality is fundamental. First of all, we have to say that the rapid process of urbanization has been one of the most important features in recent decades. Sandy urban has led to an increase in air pollution and invalid emissions. In addition, in the urban environment, short-range weather and atmospheric turbulence play an important role in the buildup of pollution [4]. Definition of urban air quality and its importance: Aerodynamics of urban areas is characterized by a complex flow pattern which changes continually and produces a sequence of instantaneous flow fields. Every urban area thus has its own air flow pattern, resulting from its unique combination of shaped characteristics, such as building packing arrangements and building roof geometries, a place of building openings (accessible for individuals and environment), etc. Urban air quality refers to the amount of pollution that affects the air within large urban areas. Broadly, outdoor air pollution is caused by primary pollutants created by human activity. Further, the pollutants get converted into their dangerous derivatives when they are susceptible to air or sunlight. The presence of precarious components can reduce air quality and is typical of incomplete, inefficient, or uncommon combustion. The most common sources of pollution in urban areas comprise motor vehicles, assigning, power generation, industry, residential combustion, and industrial processes [5]. Clearly, there is a need to monitor the quantity of pollutants directly in the urban environment. Poor urban air quality and the resulting smog have been shown to deteriorate a large range of personal and environmental interests, from health to the economy, and have prompted interest in devising forecast systems that can predict smog to build curative strategies together with environment and quality standards. Estimating the pollutant from remotely sensed data is a complex task and has been explored by various individuals using data mining methods and basic machine learning techniques. Although data mining techniques have been widely used for pollution monitoring, there is still a need to develop an automated decision support system for managing street pollution sources. In this paper, an attempt to structure the importance of majority pattern in pollution monitoring and environment protection guidelines is based on the most majority pattern, as well as a data mining model using various techniques by the authors is put to thrash for in the end [6].

MACHINE LEARNING APPLICATIONS IN AIR QUALITY MONITORING

State-of-the-art machine learning methods are used in various applications in urban air quality monitoring. Machine learning can be applied to air quality data on an urban scale, since the data are more readily available in urban areas. For instance, an approach of combining the CNN network with Gaussian Process (GP), Deep Learning Model including Convolutional Gaussian processes (CGPs) model, statistical models of Deep Gaussian Process (DGP), auto-encoding structures, etc., and deep convolutional neural network; and other Machine Learning methods, is proposed to detect and classify air pollution events from a set of time series sensor measurements (e.g., NO and NO2) [7]. Algorithms including support vector regression (SVR), random forest regression (RFR), Multilayer Perceptron (MLP), Neural Network autoregression (NNAR), Long short-term memory (LSTM), extreme gradient boosting machines (XGBoost), deep neural networks (DNN), etc., are used to forecast the air pollutant concentration at a given PM sensor. Zheng et al. use several Machine Learning-based methods, including the autoregressive (AR) models, the Generalized Additive Model (GAM), the Adaptive Model theory (NARX), etc. to analyze the spatiotemporal patterns of air pollutants and then monitor the stationarity of the pollutants. All the applications or techniques using machine learning methods have been demonstrated to be potentially useful tools for the monitoring and analyzing of air pollution. They are performed not only on the small-scale such as in the industrial zone or in the area where only a few sensors are deployed, but can also be widely utilized in the region that has a high density of sensor networks such as urban underground [8, 7].

CHALLENGES AND LIMITATIONS OF MACHINE LEARNING IN URBAN AIR QUALITY MONITORING

There are a number of challenges and limitations associated with using machine learning techniques for urban air quality monitoring analysis. First, the availability of high-quality reference data is crucial for model training, and machine learning models trained solely on data from low-cost air quality monitoring stations (that tend not to be officially validated) may produce biased results, especially in a deployment location objectively different from the training set. The under-dispersed real-world data also makes it difficult for a machine learning model to generalize to other experiments with diverse data [9]. Most of the current machine learning models used in air quality research mainly focus on predictive performance rather than interpretability. As new methods are developed to improve predictive accuracy, the transparent and interpretable models are often ignored because the more complex models could not easily be interpreted by humans. In this respect, ethical considerations also get raised for the transparency, accountability, and integrity of these black-box models. Black-box models cannot explain explicitly the reason for telling an air quality index value. So, when a certain ML model serves as an aid for decision-makers, their erroneous decisions due to the lack of model interpretability might result in controversy. Therefore, the development of transparent and interpretable machine learning models needs to be discussed in relation to urban air quality monitoring [10].

CASE STUDIES AND SUCCESS STORIES

Machine learning and artificial intelligence have been used effectively by researchers across the world to mitigate both the key elements making up 'urban emissions' and urban air pollution as a whole. The most important machine learning techniques used for model regression analysis in air pollution studies are artificial neural networks and support vector machines. In addition, these techniques have been successfully used for pollutant forecasting, source apportionment, emission estimation, and a number of other studies [11, 12]. Increasing concentrations of heavy metals, especially in urban areas, are also a crucial aspect of air pollution research. Some researchers are using the aforementioned AI technologies to study these metals. Nevertheless, acid rain has been steadily decreasing across the region, largely due to long-term restrictions on the emission of its main components, sulfur dioxide and nitrogen oxides, as a result of public policies such as the Acid Rain Program. As a result, determining the major sources of ammonia - a common component of fertilizers that can be released into the atmosphere and cause serious health and environmental problems both directly and as a pollutant precursor - has increasingly become the focus of environmental and public health research. The relationship between human health and the amount of ammonia and PM2.5 is a subject of public health concern in densely populated urban areas. For example, used a REDD-model and several machine learning approaches to map ammonia concentration across the United Kingdom. Similarly, other machine learning regression analysis such as a matrix vector approach and deep learning were established in various studies [13].

FUTURE DIRECTIONS AND OPPORTUNITIES

There is a significant body of work that still needs to be done in the field of urban air quality and machine learning. Areas that may be explored include: - Urban Processes. Research urban processes that lead to low air quality—e.g., day of the week, season, infrastructure changes—and how machine learning can be used to help predict these events. - Observational Equipment and Off-the-Shelf Solutions. Investigate the reliability and cost of machine learning tools specifically designed to classify inexpensive air quality monitors. - High-Impact Decisions. Investigate the use of machine learning to forecast specific pollutant levels. High accuracy and precision in forecasting pollutant levels can result in immediate actionable insight. - Data Interpolations. Research the use of machine learning to interpolate air quality data. Publicly reported data is always missing data. Investigate the use of machine learning to impute data values for public presentation [14]. Technological. The hardware associated with inexpensive air quality sensors has improved significantly over the past decade. The introduction of more sensitive sensors—including Particulate Matter (PM) speciation sensors, more pinpoint temperature sensors, and a wider variety of gas sensors—could prove valuable. Policy. Political action should be considered to require better air monitoring systems on a city by city basis [15, 16]. In the future, low-cost and easy-to-use computational and visual tools could be developed that integrate weather, GPS, and satellite image data to provide citizens and policymakers with a multifaceted view of their local air quality. By knowing the content of the air that they breathe, and understanding the basic sources of pollution in their area, local citizens will be empowered to create and police policy. Until then, be aware of the limitations of low-cost sensor setups. Precautions which are recommended include: replicate on a small scale before making a significant investment, consider working with small grants rather than large contracts based on projected benefits, and check your data sources frequently. Ideally, multiple independent data sources—such as on-the-ground sensors in a variety of contexts, drone measurements, satellite measurements, and governmental monitors—would be available to verify the accuracy and relevance of each proxy [17].

CONCLUSION

Urban air quality monitoring is crucial for safeguarding public health and the environment. The integration of machine learning techniques has demonstrated significant potential in improving the accuracy and reliability of air quality predictions. However, challenges such as data quality, model interpretability, and ethical considerations must be addressed to enhance the effectiveness of these techniques. Future research should focus on developing transparent and interpretable ML models, improving sensor technology, and fostering interdisciplinary collaborations to create comprehensive and actionable air quality monitoring systems. Policymakers should consider leveraging these advancements to implement robust air quality management strategies and protect urban populations from the adverse effects of pollution.

REFERENCES

1. Zhang X, Han L, Wei H, Tan X, Zhou W, Li W, Qian Y. Linking urbanization and air quality together: A review and a perspective on the future sustainable urban development. Journal of Cleaner Production. 2022 Apr 20;346:130988. [HTML]

2. Liu H, Cui W, Zhang M. Exploring the causal relationship between urbanization and air pollution: Evidence from China. Sustainable Cities and Society. 2022. researchgate.net
3. Zhao B, Yu L, Wang C, Shuai C, Zhu J, Qu S, Taiebat M, Xu M. Urban air pollution mapping using fleet vehicles as mobile monitors and machine learning. Environmental Science & Technology. 2021 Mar 24;55(8):5579-88. [HTML]
4. Wu H, Gai Z, Guo Y, Li Y et al. Does environmental pollution inhibit urbanization in China? A new perspective through residents' medical and health costs. Environmental Research. 2020. [HTML]
5. Liang L, Gong P. Urban and air pollution: a multi-city study of long-term effects of urban landscape patterns on air quality trends. Scientific reports. 2020. nature.com
6. Liu X, Jayaratne R, Thai P, Kuhn T, Zing I, Christensen B, Lamont R, Dunbabin M, Zhu S, Gao J, Wainwright D. Low-cost sensors as an alternative for long-term air quality monitoring. Environmental research. 2020 Jun 1;185:109438. qut.edu.au
7. Kumar K, Pande BP. Air pollution prediction with machine learning: a case study of Indian cities. International Journal of Environmental Science and Technology. 2023 May;20(5):5333-48. springer.com
8. Harishkumar KS, Yogesh KM, Gad I. Forecasting air pollution particulate matter (PM2. 5) using machine learning regression models. Procedia Computer Science. 2020. sciencedirect.com
9. Bush T, Papaioannou N, Leach F, Pope FD, Singh A, Thomas GN, Stacey B, Bartington S. Machine learning techniques to improve the field performance of low-cost air quality sensors. Atmospheric Measurement Techniques. 2022 Jun 1;15(10):3261-78. copernicus.org
10. Zhang W, Wu Y, Calautit JK. A review on occupancy prediction through machine learning for enhancing energy efficiency, air quality and thermal comfort in the built environment. Renewable and Sustainable Energy Reviews. 2022. sciencedirect.com
11. Castelli M, Clemente FM, Popovič A, Silva S, Vanneschi L. A machine learning approach to predict air quality in California. Complexity. 2020;2020(1):8049504. wiley.com
12. Liu X, Lu D, Zhang A, Liu Q, Jiang G. Data-driven machine learning in environmental pollution: gains and problems. Environmental science & technology. 2022 Jan 27;56(4):2124-33. [HTML]
13. Bi X, Zhang M, Wu Y, Fu Z, Sun G, Shang L, Li Z, Wang P. Distribution patterns and sources of heavy metals in soils from an industry undeveloped city in Southern China. Ecotoxicology and Environmental Safety. 2020 Dec 1;205:111115. sciencedirect.com
14. Hussain A, Draz U, Ali T, Tariq S, Irfan M, Glowacz A, Antonino Daviu JA, Yasin S, Rahman S. Waste management and prediction of air pollutants using IoT and machine learning approach. Energies. 2020 Aug 1;13(15):3930. mdpi.com
15. Giordano MR, Malings C, Pandis SN, Presto AA, McNeill VF, Westervelt DM, Beekmann M, Subramanian R. From low-cost sensors to high-quality data: A summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors. Journal of Aerosol Science. 2021 Nov 1;158:105833. sciencedirect.com
16. Chojer H, Branco PT, Martins FG, Alvim-Ferraz MC, Sousa SI. Development of low-cost indoor air quality monitoring devices: Recent advancements. Science of The Total Environment. 2020 Jul 20;727:138385. up.pt
17. Concas F, Mineraud J, Lagerspetz E, Varjonen S, Liu X, Puolamäki K, Nurmi P, Tarkoma S. Low-cost outdoor air quality monitoring and sensor calibration: A survey and critical analysis. ACM Transactions on Sensor Networks (TOSN). 2021 May 28;17(2):1-44. acm.org

CITATION: Ugwu Chinyere N., Val Hyginus Udoka Eze and Ogenyi Fabian C. Urban Air Quality and Machine Learning. Research Output Journal of Engineering and Scientific Research. 2024 3(1):101-104.