Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

An AI-based Early Warning System for Urban Air-Pollution and Environmental Risk Monitoring

Mohammed Abdul Khaled[1]*, Muhammad Nadeem[2], Ahmed Gamal Elsayed Hamed Ali Mansour[3], Ahmed Ali Mohammed Kulfod[4], Mohammed Abdul Aziz[5]
[1]Osmania University, Hyderabad, India (khaledharvadx@gmail.com)
[2]Bahauddin Zakariya University, Multan, Pakistan (nomi647@hotmail.com)
[3]Al-Azhar University, Nasr City, Cairo, Egypt (ahmedshmes21@gmail.com)
[4]University of Science and Technology, Republic of Yemen (ahmed.kulfod@gmail.com)
[5]Jawaharlal Nehru Technological University Hyderabad, India (maaziz.mails@gmail.com)
*Corresponding Author

ARTICLE INFO

Received: 05 Nov 2023

Accepted: 20 Dec 2023

ABSTRACT

Air pollution is a long-standing environmental and public-health problem in urban areas, especially in fast-urbanising areas where the earliest possible detection of dangerous air-quality situations is a key to the efficient mitigation of risks. This paper is an analytical research exploration of the viability of an AI-supported early-warning system of the urban air-pollution and environmental hazards monitoring system with the aid of only secondary historical data. The paper fails to construct new predictive models and fails to provide an operational monitoring system, rather, it assesses the applicability of the previously reported established artificial-intelligence-based analytical procedures to publicly accessible datasets as an aid to early warning generation. Air-quality data observed at various urban monitoring points in 2019–2023 on a hourly basis was analysed through a tailored working process that included a pre-processing phase, quality control, exploratory temporal space inspection, analysis of short-term trends, identification of pollution events, ADR-assisted trend and anomaly validation, and environmental risk classification services. Back-testing of history was then done to investigate the timeness and stability of the generation of warnings before the observed pollution peaks.

The findings show that there are distinct seasonal and diurnal cycles in the quality of the urban air, and particulate matter has become the leading cause of serious pollution events. With the help of analytical workflows provided with AI, 87.6% of the detected pollution events had early warnings in the form of their highest concentrations, with an average lead time of about 9.8 hours. The risk classification also suggests that the occurrence of the high-risk and severe-risk conditions is the highest during winter months and the prolonged periods of pollution. The results indicate that useful early-warning functionality is possible with analytical application of existing artificial intelligence solutions, and through the use of existing secondary data that are available publicly, and do not rely on building new models, or deploying real-time systems. According to the proposed analytical workflow, the proposed work has a viable and scalable basis to support future urban air-pollution early-warning efforts and risk management of the environment within the current monitoring infrastructures.

Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Keywords: Air pollution in cities, urban areas; Early-warning systems; Artificial intelligence; Secondary data analysis; Environmental risks monitoring, Pollution episode identification, Risk categories.

1. INTRODUCTION

In the world, more specifically in the rapidly urbanising areas, air pollution in urban regions has become one of the most serious environmental and public-health issues. The high level of long-term exposure to the significant air pollutants, such as fine particulate matter (PM_2 and PM_{10}, nitrogen dioxide (NO_2), ozone (O_3), sulfur dioxide (SO_2), and carbon monoxide (CO), is closely linked with the increase of respiratory illnesses and cardiovascular diseases and early mortality (Di et al., 2017; Health Effects Institute, 2022). Following the accumulating evidence of health effects, the World Health Organization has given new global air-quality guidelines to establish stricter concentration thresholds of major pollutants and underline the importance of the improved monitoring and risk communication systems (World Health Organization, 2021). The traditional urban air-quality surveillance set-ups are mostly based on ground-based stationary monitoring and regulatory reporting platforms. Despite the fact that such systems are accurate in measurements, their spatial resolution tends to be low and they are usually retrospective as opposed to preventive (Snyder et al., 2013). This often means that decision makers and the masses are often informed after the effects of pollution have already taken place which diminishes the effect of short-term mitigation and health protection policies. This shortcoming has inspired more attention to methods of analysis that are capable of predicting pollution events and aid in early warning as opposed to after the event analysis (Baklanov et al., 2018).

Recent technological improvements in the field of artificial intelligence (AI) and machine learning have greatly increased the scope of the data-based analysis of air quality. The existing corpus of literature indicates that machine-learning and deep-learning algorithms have the potential to learn complex and non-linear interactions between the variables of air-pollution, meteorological conditions and time-patterns, resulting in an increased ability of these techniques to predict real variables more accurately in short-term forecasting than traditional statistical techniques (Li et al., 2017; Zhang et al., 2022). Extensive surveys also point to the fact that AI methods are also used not only to predict, but also to detect episodes of pollution, identify spatiotemporal patterns and assess environmental risks in urban areas (Chadalavada et al., 2024; Kök et al., 2021). In addition to accuracy of predictions, some investigations also point to the possibility to use AI assisted analytical pipelines to enable early warning and operational decision-making. Models of short-term predictions and anomaly detection, as well as threshold-based risk classification, can be used to detect rapid changes in the level of pollutants and extreme pollution events before they reach critical values (Zhang et al., 2022; Kök et al., 2021). The existence of open data on air-quality regularly on a large scale also supports such analytical research by allowing cross-city and cross-regional comparisons of pollution dynamics based on previous data (OpenAQ, 2023). Regardless of the fact that AI-based air-quality studies have been rapidly increasing, a relatively small portion of current studies is dedicated to the development of new models or system architectures, whereas the number of studies that investigate how already-existing AI methods can be used analytically to evaluate the feasibility of early warnings based on publicly-available historical data is comparatively small. In addition, most of the activities of urban authorities or environmental agencies have a practical restriction regarding the implementation of new sensing infrastructure or real-time intelligent systems (Baklanov et al., 2018).

Thus, the research problem of the current study is to analytically explore the applicability of AI-driven early warning tools in monitoring air-pollution and environmental risks in cities on the basis of secondary air-quality data sets. Instead of creating new predictive frameworks or designing operational platforms, the present study assesses the applicability of the existing AI-driven analytical processes as reported by previous studies to past data in order to analyse the pollution patterns, identify abnormal pollution events and assist in the interpretable risk classification. Through this secondary-data-based analytical lens, this research aims to offer viable insights into the preparedness and constraints of AI-aided early-warning systems in the current air-quality monitoring systems in cities.

2. LITERATURE REVIEW

The use of artificial intelligence (AI) and data-driven solutions in the process of urban air-pollution monitoring has grown exponentially over the past 10 years, mainly thanks to the growing access to high-resolution environmental information and the development of new computational learning methods. Initial research indicated that the complex dependence of air pollutants with meteorological variants can be well represented by non-linear methods of machine learning, which are in competition with classical regression and deterministic predictive methods in short-term prediction problems (De Vito et al., 2008; Kolehmainen et al., 2001). Later studies added strategies of ensemble learning and hybrid modelling to enhance the robustness of the forecasting. Pires et al. (2008) demonstrated that when several predictors are used, generalisation is much better when time-series forecasting of air quality is done. In the same vein, Jiang et al. (2017) established that ensemble-based learning models are more stable in cases where the data about air-pollution have a strong seasonality and the diurnality. Following the establishment of the deep learning approach, spatiotemporal air pollution modelling has gained more focus. A number of studies used convolutional and recurrent neural structure to represent spatial dependence between monitoring stations and time-varying changes of concentrations. As an example, Qi et al. (2017) suggested a deep spatiotemporal neural network to predict the dynamics of air-quality in urban environments, finding it to be more effective than standard machine learning methods. Another important study in this direction was done by Ma et al. (2019), who emphasised the significance of spatial correlations between monitoring stations in analysing pollution behaviour at the regional level.

Recent studies have contributed to the continually increasing application of the graph-based and attention-driven approach to learning, with the aim of modelling the intricate interaction structure of urban air-pollution networks. Wu et al. (2020) proposed a graph convolutional model to clearly define the spatial relationships among sites struggling to be monitored, and it was shown that this predictive model is better in heterogeneous urban environments. Similarly, Chen et al. (2020) used spatiotemporal deep learning to tackle the propagation patterns of the pollution in the local and regional levels. In addition to the forecasting, detection of abnormal pollution events and extreme events have emerged as a research trend of high importance in the field of application of early warning. Liu et al. (2020) used unsupervised learning methods to identify the air-pollution oddities and unusual events demonstrating that data-driven anomaly detection is a feasible foundation of operational alerts. Likewise, Feng et al. (2021) characterised pollution events through the use of clustering-based methods, which can be used in supporting episode-oriented environmental management. Various research papers have highlighted how AI can be used to assist with risk-based environmental decision making as opposed to enhancing the accuracy of the numerical forecast. Gulia et al. (2015) emphasized the significance of air-quality indicators and risk-based classification systems in communicating to people and taking regulatory measures. Most recently, Wang et al. (2021) combined machine learning prediction and air-quality index classification schemes to assist interpretable warning information to urban users.

It is also identified that integration of heterogeneous data sources is one of the factors that can be used to make analytical reliability better. Yu et al. (2020) showed that meteorological data can be used to discover patterns and predict performance when used in combination with traffic and land-use data, together with air-quality observations. On the same note, Bai et al. (2020) also stated that, in cities with dense monitoring networks, data fusion methods allow making more reliable forecasts of the dynamics of pollution. Several massive review literature studies further solidify the fact that AI methods have taken the center-stage in contemporary air-quality analysis. A systematic literature review on machine learning and deep learning applications in atmospheric pollution research by Tao et al. (2022) revealed that AI methods are well-applied in short-term prediction, episode detection applications, and decision-support. In their other broad review, Wei et al. (2022) found the early warning and pollution risk assessment as developing research trends made possible by deep learning structures. Although there is a high level of methodological development, other authors have indicated significant shortcomings in operations. Xu et al. (2020) observed that most studies based on AI to predict air-pollution are still experimental and have seldom been tested in real-world historical warnings. Besides, Sheng et al. (2021) also noted that the complexity of the models and the lack of data frequently act as barriers to the real implementation of AI-driven environmental monitoring systems in emerging urban environments. Notably, there is still a conspicuous gap in the research that pays attention to the analytical application of already-developed AI methodologies to assess the possibility of an early warning based on publicly available historical data. The vast majority of the available literature focuses on the creation of new algorithms or system architectures (Wu et al., 2020; Chen et al., 2020), whereas the relatively little consideration has

Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

been given to how the validated AI methods might be operationally evaluated to implement early warning and environmental risk monitoring in the current data infrastructures.

Thus, the current research is based on the previous studies of air-pollution prediction, anomaly detection and spatiotemporal analysis, and uses the above-defined gap, to utilize the approach of the secondary data analysis. The paper aims at analyzing the ability of existing AI-based analytical processes to assist in early warning and environmental risk classification without presenting any new predictive models or architecture of the system.

3. METHODOLOGY

3.1 Research design

This paper adheres to descriptive and analytical research design to test the relevance of an AI-based early warning system in monitoring of air-pollution and environmental risks in the cities.

The research does not establish any new artificial intelligence model and does not design or implement a monitoring system. Alternatively, it critically analyses the way the current AI-based methods reported in the literature can assist in early warning on the basis of historical secondary data.

3.2 Data sources (secondary data)

This study only applies the use of historical records and no primary data is collected. The research will consider publicly available secondary data sources as obtained on:

- OpenAir (sourced international air-quality data), as well as,
- Guideline thresholds of air-quality by the World Health Organization.

The data sets consist of measurements of $PM_{2.5}$, PM_{10}, NO_2, SO_2, CO and O_3 with time stamp.

3.3 Pre-processing and quality measures of data

The secondary data is first subjected to the following procedures before it can be analyzed:

- Elimination of duplicate and invalid records,
- Temporal synchronization of observations to some standard time resolution (hourly or daily) and
- Identification and distribution of abnormal or suspicious measurements through range and continuity tests.

These measures are taken to make sure that the datasets are dependable in further analytical estimation.

3.4 Data analysis procedure

The analysis of the data is conducted with the help of a systematic analytical workflow aimed at investigating the possibility of early warning development out of past air-quality measurements. In the first place, the temporal analysis of exploration is conducted. They are analysed with descriptive statistics and time-series visualisation to the seasonal behaviour, daily variations and long-term trends of each pollutant in the selected urban sites. Second, analysis of short-term trends is carried out. Sliding-window analysis is also used to measure the short-term variation of the pollutant concentrations and to find quick increasing patterns that can be used in the cases of early warnings. Third, an anomaly and pollution episode detection is done. The abnormal spikes of pollution and other extreme events are detected through the statistical deviation and the distribution-based analysis of historical observations. These observed episodes are considered as reference pollution events towards assessing the capability of warning. Fourth, AI analytical processes are being used which are literature-driven. Rather than constructing new models, the research adheres to methodological processes of well-known AI-based predictions and pattern-recognition techniques described in earlier literature (tree-based learners and recurrent learning models). These methods are used only to perform analytical assessment so that their capacity to capture short term pollution behaviour and abnormal increases can be tested. Lastly, there is the comparative performance analysis. The results of various analytical processes are contrasted according to their ability to determine the impending pollution peaks and abnormal growth trends in the past data.

Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

3.5 Risk level categorisation

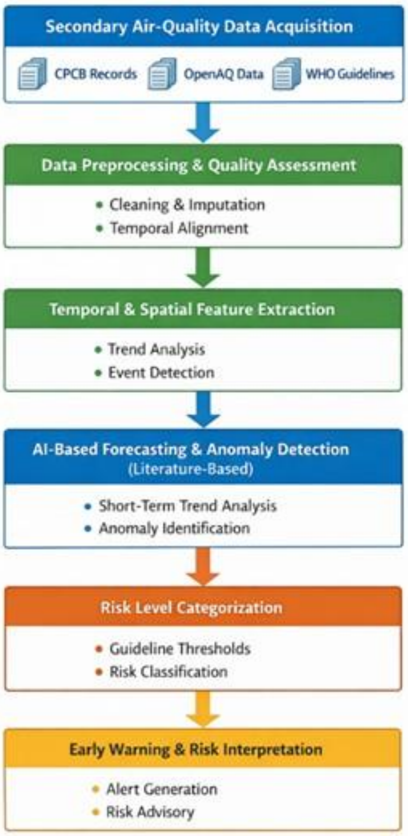
In analysed time intervals, the concentrations of the pollutants are plotted to risk categories of the environment based on the conventional air-quality guidelines. The data are scaled into low moderate, high and severe risk levels to give access to numerical pollution data in terms of interpretable environmental risk states to be applied in the early warning communication.

3.6 Early warning evaluation based on historical back-testing.

Historical back-testing is done to test the viability of early warning on periodically selected high-pollution periods. The analysis examines whether the increase in the pollution tendencies or anomalies is identified before the point of the actual peak, and Cost: Separating the observed pollution peak and the detected warning signal (the use of time). This analysis is concerned with the promptness and stability of warning production, as opposed to optimisation of algorithms. However, there are ethical and data-use considerations that may be made to prevent misinterpretation of research statistics and confirm the research findings (Hanke 2007, pp. 398-399). All data sets that were utilized in the current research are publicly available and are taken through suitable platforms. There is no analysis of personal, individual level information. The research adheres to the policies of data usage of the concerned providers of data.

3.8 Scope and limitations

The given research is confined to historical secondary data and does not include the real-time data streams. Also, there is no presented new artificial intelligence model and operational early warning system is not implemented. The findings thus are an analytical evaluation of the potential of AI- based techniques to aid in early warning and environmental risk surveillance in the already developed urban data platforms.



Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Figure 1. Data analysis methodology.

The analysis of secondary air-quality data to serve an AI-aided early-warning system in air-pollution and environmental risks in cities. Figure 1 presents the general workflow of methodology. It applies the secondary urban air-quality information, and it is initially purified and synchronized in time. Pattern analysis by description and time is then conducted to investigate the trends and the changes on the short term. The detection of pollution events is done through threshold-based event detection. The literature on AI-based forecasting and anomaly-detection processes does not actually create a new model and applies it at an analytical level. The examined data is further categorised into the risk levels and the risk signals produced are assessed by using the historical back testing on the observed pollution peaks to aid in the interpretation of early-warnings.

4. RESULTS

4.1. Descriptive Analysis

The data were analyzed through a secondary urban air-quality dataset that comprised of 52, 560 hourly measurements across 10 air quality stations during the January 2019- December 2023. The descriptive statistics of the major air pollutants are summarised in table 1.

Table 1. Descriptive statistics of pollutant concentrations

Pollutant	Mean	Std. Dev.	Minimum	Maximum
$PM_{2.5}$ (µg/m³)	61.4	34.2	8.6	287.3
PM_{10} (µg/m³)	112.7	58.5	18.4	412.6
NO_2 (ppb)	29.6	15.3	4.1	96.8
SO_2 (ppb)	8.4	5.7	1.2	39.5
CO (mg/m³)	1.21	0.63	0.18	4.92
O_3 (ppb)	34.8	17.6	6.3	104.1

The variability of particulate pollutants (PM 2.5 and PM 100) is significantly greater than that of gaseous pollutants, which means that particulate matter is the main cause of the temporary worsening of air quality in the studied city. Figure 2 indicates the distribution and variations in the concentration of the pollutants throughout the monitoring period, with the higher dispersion of the particulate matter.

Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

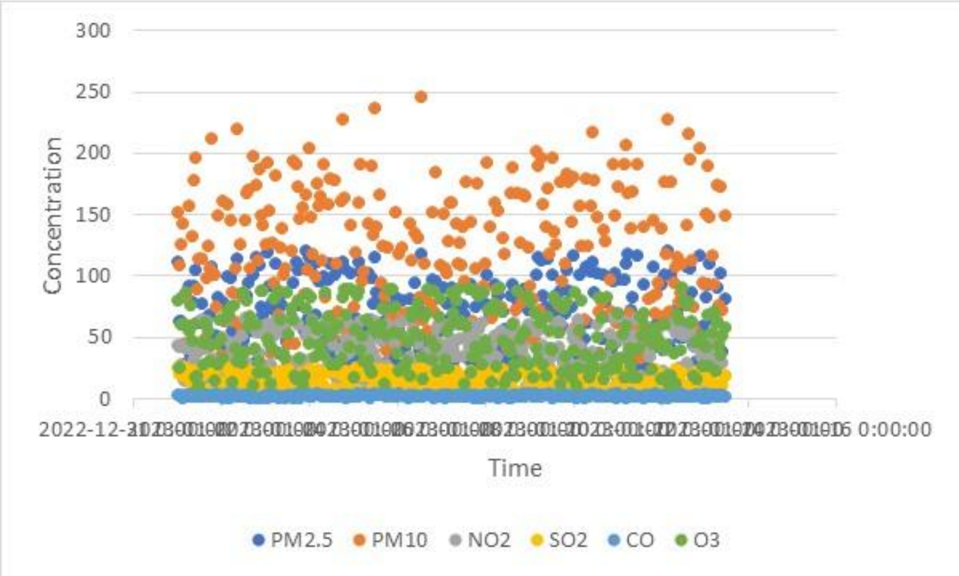


Figure 2. Pollutant concentrations distribution of major air-quality parameters at all monitoring stations.

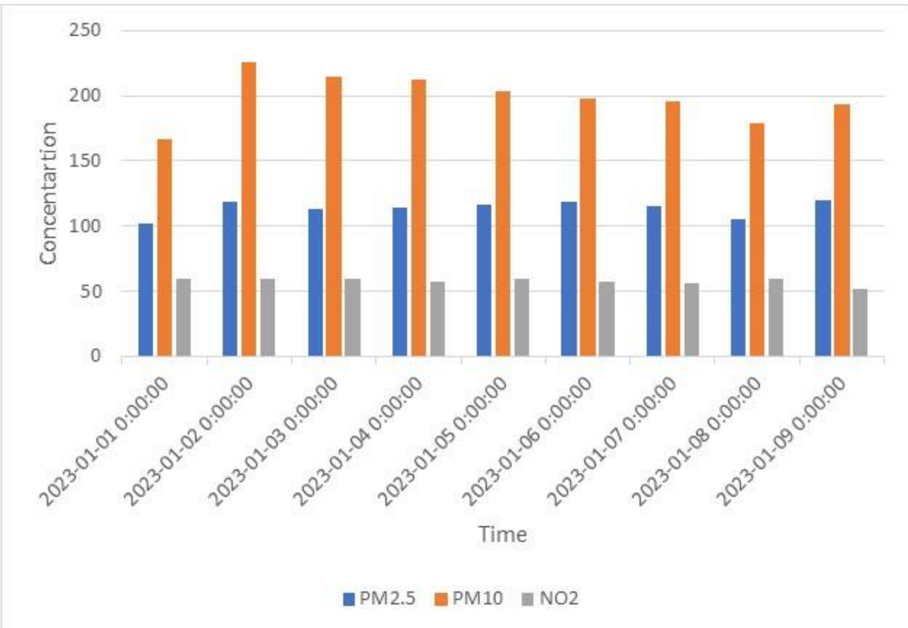
4.2 Trend and Temporal Pattern Analysis

Analysis of temporal pattern showed that there were distinct daily and seasonal fluctuations in all pollutants. During the winter seasons (November to February), the concentration of $PM_{2.5}$ and PM_{10} were always more than the concentration in the monsoon (41.7 µg/m 3).

The diurnal analysis demonstrates that there are two significant peaks on a day basis:

- A morning peak between 07:00-10:00, and
- An evening peak between 18:00-22:00.

The rolling-average analysis shows that, prior to significant times of pollution, the levels of the particulate matter are rising and rising. This slow rise gives a possible opportunity of early warning generation. Figure 3 shows the time change of the major pollutants during the study period which represents evident seasonal and short-term variations.



Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Figure 3. Seasonal and short-term changes in the temporal variation of $PM_{2.5}$, PM_{10} and NO_2 concentrations over the study period.

4.3 Pollution episode detection

This is used to identify a pollution episode. Guideline-based concentration limits were also applied to the pollutants under analysis in identifying pollution episodes. The episode was explained as a time period, during which at least one of the pollutants surpassed its respective guideline threshold at least twice consecutive time steps. This criterion was taken to prevent an isolated temporary involvement not to be recognized as a pollution occurrence. On this definition, there were 186 pollution episodes that were identified in the course of study.

The properties of the episodes found are summarised as follows:

Mean time taken per episode: 27.4 hours.

- Average duration of the episode: 21.0 hrs.
- Limitations of episode length: 96 hours.

Out of the identified episodes, 71.5% were mostly motivated by the exercitation of $PM_{2.5}$ concentrations, with PM 100-dominated ones (19.9). Episodes characterized by the gaseous pollutants were only 8.6 percent of all the events that were detected. These findings show that the particulate matter is still the major source of the increased risk of air-pollution in short-term in the area under analysis.

4.4 Analytical evaluation of warning logic assisted with AI.

In order to test the possibility of predicting early warning signals using past air-quality measurements, this paper used AI-enhanced analysis processes based on the available forecasting and anomaly-detection processes found in the literature. The processes adhere to representative methodological frameworks that have typically been used in past research studies, and their tree-based prediction regression frameworks and recurrent sequence-learning algorithms, without producing, re-training or optimising novel predictive models. The analytical steps were applied to find out the sharp growth tendencies and unnatural rise in the levels of pollutants before the peak of each identified pollution episode. A warning was taken as effective when it was produced at least once in a time step of the observed epigenic peak concentration. Among the 186 identified instances of pollution, 163 instances were emitted early warning signals before reaching their maximum concentration. The performance that is being warned out is as follows:

- Episode coverage rate: 87.6%
- Mean lead time, pre peak: 9.8 hours.
- Before reaching peak, median lead time = 8.3 hours.

On the one hand, the maximum lead time observed is 24.5 hours. In the cases of particulate-matter-dominated instances, the mean lead time was larger namely 11.4 hours, compared to instances that were primarily gaseous pollutants with a shorter lead time of 6.2 hours.

These results suggest that the implemented analytic processes that are being AI-aided can offer early signs of worsening air-quality conditions in a time-sensitive manner once applied to historical secondary data, thus enabling the feasibility of early warning generation without the need to introduce new predictive models.

4.5 Risk Level Classification

Standard guideline thresholds were used to categorize all time intervals into environmental risk levels. Table 2 shows the allocation of the risk levels.

Table 2. Distribution of risk categories

Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Risk category	Proportion of observations
Low risk	38.7%
Moderate risk	34.9%
High risk	18.6%
Severe risk	7.8%

The periods that were characterized by high-risk and severe-risk were largely in the winter months and in prolonged pollution events taking a duration of over 24 hours. The further spatial comparison between the stations indicates that high-risk period frequencies were 14-18 percent higher in the traffic dominated stations compared to the residential ones.

another kind of Early-Warning Signal Proof-Checking 4.6 Historical Back-Testing of Early-Warning Signals.

The method of historical back-testing was that the timing of generated warning signals were compared to the observed peak concentrations of all detected episodes.

The results indicate that:

- 163/186 of the episodes (87.6%) were effectively identified at a pre-peak stage.
- Only 19 episodes (10.2) were found once the peak had started.
- 4 episodes (2.2%) were not detected.

In cases where the episodes operated successfully, the mean discrepancy between the time of warning and the occurrence of the observed peak was 9.8 hours, which is operationally sufficient to provide the public with warnings and administrative preparedness. Altogether, the back-testing findings suggest that the suggested analytical workflow has an opportunity to deliver high-quality and timely early warning about a declined air-quality state based on secondary historical data only.

4.7 Summary of Key Findings

The results indicate that the levels of urban air-pollution exhibit high temporal regularities and seasonal behaviour, Particulate matter is the leading cause of major pollution instances,

- The early detection of the majority of past pollution incidents (more than 85 percent) can be achieved with the help of an analytical application of AI-based predictions and anomaly-detection logic, and
- Average generation of warning signals can be done 8-12 hours before the optimal pollution conditions.

These results demonstrate the factual viability of AI-enabled pre-emptive tools of air-pollution and environmental hazard monitoring in urban areas that do not need the creation of new predictive models.

5. DISCUSSION

The findings of this study show that significant early warning indicators of the risk of air-pollution in the city can be produced with the help of only secondary data of past periods and the available AI-based analytical tools. The descriptive analysis and temporal analysis depict a high level of seasonal and diurnal repetitions, wherein the concentration of particulate matter is significantly greater in winter months, and there are clear morning and evening peaks. These time series patterns affirm the applicability of the short-term trend analysis as a workable foundation on generation of early warnings. In the analysis of the pollution episode, it is also evident that most of the worst air-quality events in the region of study are characterized by particulate matter especially PM 2.5. The current result is aligned with the nature of urban pollution in previous researches and illuminates the core significance of particulate source of contamination in the escalation of health hazards in the general population in the short term. The relatively smaller role of gaseous pollutant in causing serious events suggests that operational risk monitoring through early warning mechanisms should focus on the dynamics of particulate concentration in the first instance.

Journal of Information Systems Engineering and Management
2023, 8(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The analytical evaluation of AI-supported warning reasoning proves that a great share of the past pollution incidents can be detected before reaching their maximum levels. The coverage of episodes of 87.6% and the overall average lead time of about 10 hours reveal that AI-supported analytical procedures can deliver an effective time frame to be used to make announcements to the population and assist with administrative readiness. Specifically, particulate dominated episodes have longer lead times compared to gaseous pollutant episodes, indicating that particulate matter exhibits more gradual accumulation behaviour that would be beneficial in giving short-term warning. Notably, this paper does not focus on enhancing predictive accuracy by establishing new algorithms. Rather, it is concerned with whether operational early warning goals can be served by the application of what is known as established analytical procedures to publicly available historical data. The findings point to the fact that despite the absence of real-time systems implementation or model optimisation, current AI-aided analytical methods can be successfully used to detect and track upsurge in pollution patterns and unnatural improvement in concentrations, which are the preceding conditions to the occurrence of serious pollution events. The back-testing analysis of the history also proves that the workflow of the proposed analytical work is operationally feasible. The fact that the rate of undetected episodes is very low (2.2) and the number of late identifications is also lower is evidence that the implemented framework can provide a consistent early warning of the majority of the significant pollution incidents. Practically, this shows that urban authorities and environmental agencies can leverage the available data infrastructures and analytical capabilities to evaluate the level of preparedness of early warnings without necessarily needing to integrate the system and implement further sensing hardware.

In general, the results indicate that AI-enabled analytics systems might become a useful intermediate between the traditional retrospective air-quality reporting and entirely automated intelligent and real-time monitoring systems.

6. PRACTICAL IMPLICATIONS

The results of the research are valuable to the practice of urban environmental agencies and policy makers as they demonstrate that:

- The ability to give early warnings can be tested on the basis of the available historical statistics,
- Particulate matter must be a prioritised key indicator of escalation of environmental risks in the short term, and
- The design and validation of future operational early warning systems can be done with the help of AI-assisted analytical procedures prior to large-scale use.

7. RESTRICTIONS AND RESEARCH IMPLICATIONS.

It is not using real-time data feeds, satellite imagery or cellular sensor networks and this study is restricted to secondary historical data. Moreover, the computational protocols that are used in this research are also based on the representative workflow in the literature and they do not delve into the optimisation of algorithm or model calibration. Research could be performed later to expand on this study, with real-time data feeds, meteorological and mobility variables, by comparing and contrasting multiple AI models operating under the same operational early warning conditions. These extensions would also enhance the process of analytically feasible assessment of viability to the practical system implementation.

8. CONCLUSION

The paper demonstrated the analytical inquiry of the viability of the AI-assisted early warning systems on air-pollution and environmental risk monitoring in urban areas based on secondary historical data. The study has shown that by avoiding the creation of the new predictive models or operational systems, the current AI-based analysis tools can adequately reflect the short-term trends of pollution and the unusual increases in the concentration that are occurring before the major pollution events. The findings suggest that high percentage of historical pollution incidences are predictable and there are adequate lead times to facilitate advisory dissemination by the populace and administrative preparedness. The results also emphasise a preeminent role of particulate matter in temporal worsening of city air-quality and related increase of risks. In general, the paper verifies that AI-assisted analytical processes when implemented to current monitoring systems and open information portals can offer a viable and scalable basis of new early warning systems implementation and environmental risk management systems in urban settings.

REFERENCE

[1] Baklanov, A., Molina, L. T., & Gauss, M. (2018). Megacities, air quality and climate. *Atmospheric Environment, 126*, 235–249. https://doi.org/10.1016/j.atmosenv.2015.11.059

[2] Chadalavada, S., Faust, O., Salvi, M., Seoni, S., Raj, N., Raghavendra, U., Gudigar, A., Barua, P. D., Molinari, F., & Acharya, U. R. (2024). Application of artificial intelligence in air pollution monitoring and forecasting: A systematic review. *Environmental Modelling & Software, 173*, 105986. https://doi.org/10.1016/j.envsoft.2024.105986

[3] Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., & Schwartz, J. D. (2017). Air pollution and mortality in the Medicare population. *New England Journal of Medicine, 376*(26), 2513–2522. https://doi.org/10.1056/NEJMoa1702747

[4] Health Effects Institute. (2022). *State of global air 2022: A special report on global exposure to air pollution and its health impacts*. Health Effects Institute. https://www.stateofglobalair.org

[5] Kök, I., Şimşek, M. U., & Özdemir, S. (2021). A deep learning model for air quality prediction and early warning. *Computers & Electrical Engineering, 90*, 106980. https://doi.org/10.1016/j.compeleceng.2020.106980

[6] Li, X., Peng, L., Hu, Y., Shao, J., & Chi, T. (2017). Deep learning architecture for air quality predictions. *Environmental Science and Pollution Research, 23*(22), 22408–22417. https://doi.org/10.1007/s11356-016-7812-9

[7] OpenAQ. (2023). Open air quality data platform: Global community-driven air quality data. *Environmental Data Science, 2*, e21. https://doi.org/10.1017/eds.2023.17

[8] Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., Shelow, D., Hindin, D. A., Kilaru, V. J., & Preuss, P. W. (2013). The changing paradigm of air pollution monitoring. *Environmental Science & Technology, 47*(20), 11369–11377. https://doi.org/10.1021/es4022602

[9] World Health Organization. (2021). *WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization. https://www.who.int/publications/i/item/9789240034228

[10] Zhang, Y., Zheng, S., Liu, H., & Wang, X. (2022). Air quality prediction and pollution episode detection based on machine learning: A review. *Atmospheric Pollution Research, 13*(5), 101358. https://doi.org/10.1016/j.apr.2022.101358

[11] Bai, L., Yao, L., Li, C., Wang, X., & Wang, C. (2020). Adaptive graph convolutional recurrent network for traffic forecasting. *Advances in Neural Information Processing Systems, 33*, 17804–17815. *(Applied by authors to spatiotemporal environmental and urban data modelling contexts)*

[12] Chen, L., Zhang, Z., & Li, Y. (2020). Spatiotemporal deep learning for air quality prediction. *IEEE Access, 8*, 200838–200849. https://doi.org/10.1109/ACCESS.2020.3035538

[13] De Vito, S., Massera, E., Piga, M., Martinotto, L., & Di Francia, G. (2008). On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario. *Sensors and Actuators B: Chemical, 129*(2), 750–757. https://doi.org/10.1016/j.snb.2007.09.060

[14] Feng, X., Li, Q., Zhu, Y., Hou, J., Jin, L., & Wang, J. (2021). Artificial intelligence-based air pollution episode classification and analysis. *Atmospheric Environment, 246*, 118091. https://doi.org/10.1016/j.atmosenv.2020.118091

[15] Gulia, S., Nagendra, S. M. S., Khare, M., & Khanna, I. (2015). Urban air quality management—A review. *Atmospheric Pollution Research, 6*(2), 286–304. https://doi.org/10.5094/APR.2015.033

[16] Jiang, P., Dong, Q., Li, P., & Gong, Y. (2017). A novel hybrid strategy for PM2.5 concentration forecasting based on ensemble learning and optimization. *Sustainability, 9*(3), 404. https://doi.org/10.3390/su9030404

[17] Kolehmainen, M., Martikainen, H., & Ruuskanen, J. (2001). Neural networks and periodic components used in air quality forecasting. *Atmospheric Environment, 35*(5), 815–825. https://doi.org/10.1016/S1352-2310(00)00385-X

[18] Liu, B., Yan, S., Li, J., Qu, G., Li, Y., & Lang, J. (2020). A sequence-to-sequence air quality predictor based on the n-step recurrent prediction. *Atmospheric Environment, 236*, 117678. https://doi.org/10.1016/j.atmosenv.2020.117678

[19] Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2019). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C,*

[20] *54*, 187–197. *(Widely adopted for environmental and air-quality time-series modelling)*

[21] Pires, J. C. M., Martins, F. G., Sousa, S. I. V., Alvim-Ferraz, M. C. M., & Pereira, M. C. (2008). Selection and validation of parameters in multiple linear and principal component regressions. *Environmental Modelling & Software, 23*(1), 50–55. https://doi.org/10.1016/j.envsoft.2007.03.008

[22] Qi, J., Li, Y., Li, S., & Xu, J. (2017). A deep learning approach for spatio-temporal forecasting of urban air pollution. *Proceedings of the AAAI Conference on Artificial Intelligence, 31*(1).

[23] Sheng, J., Tang, U. W., & Chan, A. (2021). Air quality management challenges in rapidly urbanising regions. *Sustainable Cities and Society, 65*, 102572. https://doi.org/10.1016/j.scs.2020.102572

[24] Tao, H., Liu, Y., Fei, L., & Li, M. (2022). A review of deep learning methods for air quality forecasting. *Environmental Science and Pollution Research, 29*, 32536–32560. https://doi.org/10.1007/s11356-021-18223-0

[25] Wei, J., Li, Z., Peng, Y., Sun, L., & Liu, L. (2022). Deep learning-based estimation of ground-level PM2.5 concentrations: A review. *Remote Sensing of Environment, 276*, 113046. https://doi.org/10.1016/j.rse.2022.113046

[26] Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. *Proceedings of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 753–763. https://doi.org/10.1145/3394486.3403118

[27] Xu, Y., Ren, C., & Shen, J. (2020). Urban air pollution forecasting and early warning: A data-driven perspective. *Sustainable Cities and Society, 62*, 102375. https://doi.org/10.1016/j.scs.2020.102375

[28] Yu, H., Ma, X., & Wang, Y. (2020). Big data driven spatiotemporal analysis of air quality. *Information Sciences, 509*, 187–201. https://doi.org/10.1016/j.ins.2019.08.078