International Journal of Advanced Engineering: Artificial Intelligence, Robotics and Automation
ISSN 3049-396X
Vol.:2 Issue: 1
Published online on: Dec 2025
PP. 17- 24

A Comprehensive Review of Deep Learning Approaches for Air Pollution Forecasting

Ritu Gupta[1], Sandeep K Tiwari[2], Dhramandra Sharma[3], Anand kumar Singh[4]
[1] Guru Tegh Bahadur 4th Centenary Engineering College, Delhi
[2,3,4]Vikrant University, Gwalior, India

Abstract- Air pollution has become one of the most critical environmental and public health challenges worldwide, driven by rapid urbanization, industrial expansion, transportation growth, and changing climatic conditions. Accurate forecasting of air pollutant concentrations is essential for informed policy decisions, emission control strategies, and public health protection. This review examines the evolution of air pollution forecasting techniques, beginning with traditional statistical approaches and advancing toward modern data-driven deep learning methodologies. Classical models such as ARIMA, Multiple Linear Regression, and Kalman Filtering are noted for their interpretability and effectiveness under stable conditions, yet they struggle to represent nonlinear atmospheric behavior. Machine learning techniques, including Support Vector Machines, Random Forests, and Artificial Neural Networks, improved performance by capturing multivariate dependencies but still lacked the ability to model temporal dynamics effectively. Recent advancements in deep learning, particularly hybrid architectures combining convolutional networks for feature extraction with recurrent frameworks for sequential learning, have demonstrated superior predictive accuracy by capturing complex pollutant–meteorological interactions. However, challenges remain, including limited data quality, high computational cost, model interpretability concerns, and difficulties in real-time implementation. This review highlights current achievements, identifies methodological gaps, and emphasizes the need for scalable, explainable, and robust forecasting systems adaptable to diverse geographic and climatic conditions.

Keywords: Air pollution forecasting, Deep learning, 1D ConvNet, Bidirectional GRU, Time-series analysis, Hybrid neural networks, Environmental monitoring

I. INTRODUCTION

One of the most urgent environmental and societal health problems of the 21 st century is air pollution brought about by the fast industrialization, urbanization, automobile emission, and the growing energy requirements. The constant increase in the concentration of air pollutants like particulate matter (PM2.5 and PM10), nitrogen dioxide (NO 2), sulfur dioxide (SO 2), carbon monoxide (CO), and ozone (O 3) has an enormous implication on human health, ecological balance, climate change, and the general quality of life. Several epidemiological reports have evidenced a high level of correlation between the long term exposure to polluted air and the onset of serious respiratory diseases, cardiovascular diseases, neurological weakness and even early death. With the increasing population and economic activities in the cities, the ability to predict the level of air pollution has been of great importance to the policymakers, environmental agencies and the populace in general. Proper and prompt predictions enable the governments to give early warnings, develop regulatory interventions, streamline traffic movement, and eradicate dangers associated with pollution. Nevertheless, it is intrinsically difficult to forecast the air quality because of the dynamic interplay of the meteorological factors, sources of emissions, chemical processes, and geographically diverse factors [1]-[4].

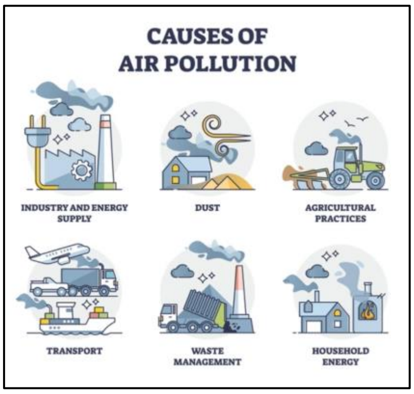

Fig. 1 Causes of Air pollution [5]

Corresponding Author’s E-mail ID: rituguupta414@gmail.com

Classical machine learning approaches and traditional statistical techniques often fail to describe these nonlinear temporal relationships, spatial heterogeneity and multivariate relationships found in atmospheric pollutant data. Over the last few years, the development of deep learning has introduced new possibilities to create highly efficient predictive models that are able to learn complex patterns using large amounts of data in the environment. When used in a 1D format (1D ConvNets), convolutional neural networks (CNNs) which were originally created as image processors have demonstrated good performance in extracting deep spatial and structural features out of sequential numerical data. Equivalently, recurrent neural network (RNN) versions, specifically, Gated Recurrent Units (GRU) and Bidirectional GRU (BiGRU) are very proficient at long-run temporal dependencies of time-series data, owing to their ability to process sequences forward and backward. The combination of 1D convolutional layers with Bidirectional GRU network can provide a potent hybrid deep learning model that will apply the advantages of two architectures: 1D ConvNets will most effectively extract local pattern of trends and short-term variability of pollutant concentration sequences, whereas the BiGRU component will effectively learn time-dependent features since it will be able to learn both past and present states in parallel. A hybrid architecture such as this is thus more holistic and robust in its prediction than single model approaches. The combination of the meteorological variables like temperature, humidity, speed of wind, atmospheric pressure, and season aspects also add to the input representation of the model in the air pollution forecasting context and help the model to be effective in generalizing the environmental conditions across different conditions. Besides, the capability of deep learning models to learn raw data will decrease the reliance on the domain-specific feature engineering, enhance the scaling and transferability of the model to cities and regions with varying pollution processes. Many research works have confirmed the high performance of hybrid deep learning models in time-series prediction tasks in environmental forecasting, energy demand forecasting and financial trend forecasting [6] - [8].

One recommendation that can be offered in this regard is a deep learning model consisting of 1D ConvNets and Bidirectional GRU since it does not only process large multidimensional data but also tackles problems such as overfitting, vanishing gradients, and limited long sequence memory. The proposed framework works under the principle of applying 1D convolutional filters to extract important temporal characteristics in the trends of pollutant concentrations and injecting the learnt characteristics into BiGRU units to gain a better understanding of the sequential relationships. The bidirectional structure makes sure that the model takes into consideration the connection of every time step with the past and future time point which is crucial in the representation of temporal pollutant formation processes that depend on atmospheric reactions and external emission variations. Moreover, the architecture can be also improved with the introduction of normalization, dropout regularization, and adaptive learning optimizer to enhance the model stability and generalization. The higher computational capabilities of the current deep learning models allow the implementation of these forecasting models into real-time monitoring systems, enabling the use of smart cities applications, air quality management portals, and IoT-based environmental surveillance systems. In addition to the role of pollution forecasting in the management of health and the environment, precise pollution forecasting will be used to reduce the burden of the economy and make the right decisions in transportation, preparedness to health, industry control, and community awareness initiatives. With the increased attention to the problem of air pollution in the whole world, the creation of intelligent forecasting systems using deep learning algorithms is not just a technical task but also a social need. ConvNets combined with Bidirectional GRU into a single predictive framework is a first-time move in the right direction towards having high-resolution, dependable and versatile forecasting solutions that can address the rapidly changing circumstances in the urban environment. Therefore, this paper is aimed at designing, developing, and assessing a high-quality deep learning based air pollution predictor model that makes use of the synergetic potential of 1D convolutional neural networks and bidirectional gated recurrent units in offering practical, precise, and real-time forecasts that can ultimately result in improved environmental sustainability and protection against diseases.[9].

II. LITERATURE REVIEW

He 2025 et.al Pollution of the air in industrial areas is a serious environmental and human health issue, especially in the fast-growing regions, requiring the use of good predicting systems to aid in reducing the situation. He et al. (2025) developed a hybrid model Transformer times Net, which is optimized using the Optuna algorithm to predict six key air pollutants in the Xinyang Industrial Zone in China. The model takes advantage of the fact that the Transformer is capable of learning long-range temporal dependencies and TimesNet of learning complex periodic structure in time series. Based on air quality data of 20192023, the hybrid strategy showed greater predictive power than the traditional statistical, machine learning, and deep learning models and provides a useful instrument in the policy planning and pollution control. [10].

Dairi 2025 et.al solved the increasing global health problem with air pollution by introducing a forecasting model based on deep learning that can enhance the prediction of ambient pollutants. Their approach combines Variational Autoencoders (VAE) with an Innovative Multiple Directed Attention (IMDA) mechanism and creates the IMDA-VAE architecture. The model was tested on the dataset of four states of the U.S. and measured by six statistical measures of accuracy. It was found that IMDA-VAE performed better than traditional models, including LSTM, GRU, BiLSTM, BiGRU and ConvLSTM. The paper brings out the success of integrating the attention process to improve the temporal pattern of pollutants learning and predictability across different sites. [11].

Corresponding Author’s E-mail ID: rituguupta414@gmail.com

Oldenburg 2024 et.al carried out a proof-of-concept analysis to predict concentrations of important pollutants, such as NO 2, O 3, PM 10 and PM 2.5, based on multivariate time-series data and meteorological variables at two sites. The researcher has contrasted various deep learning models with more focus on LSTM and GRU. A multi-task learning structure was introduced that was hierarchical and reflected the behavior of the atmosphere and relationship between different pollutants. The hierarchical GRU proved to be the most efficient and accurate (in terms of forecasting) of the tested models. The results indicate that hierarchical temporal modeling has the ability to improve prediction performance of smog related air quality indicators as well as the capacity to capture complex interaction of pollutants.[12].

Soulie 2024 et.al concerned itself with the enhancement of worldwide emission inventory necessary to model the air quality and predict atmospheric composition. The paper presented CAMS-GLOB-ANT which is a high-resolution global efforts dataset of 36 pollutants across 17 anthropogenic sectors between the year 2000 and 2023. The inventory gives monthly data on a 0.1° 0.1 X grid that can be used in both regional and global atmospheric model. The consistency of the dataset was established through methodological transparency and comparisons with already existing inventories and made the dataset appropriate in the research and operational modeling. This study plays a major role in comprehending the long term emission patterns and also increases the credibility of atmospheric chemical transport models applied to policy and climate analysis. [13].

Africa 2024 et.al Investigated macroeconomic effects of air pollution in Hungary with the help of the economic modeling process based on the Cobb-Douglas production function and Solow-Swan model of growth. The analysis measured the economic performance in the country due to health impairment brought about by pollution, especially the productive labor force. It was found that over the next fifty years, air pollution may lower GDP by 4.1-9.4 per cent every year, which will be accompanied by increasing healthcare spending. This study highlights that not only is the enhancement of air quality a priority of the people, but also a critical economic investment. Less pollution leads to a direct sustainable development in the long term and the resilience of the national economy [14].

TABLE: 1.LITERATURE SUMMARY

Authors / Year	Methodology	Research Gap Identified	Key Findings
Yue et al., 2024 [15]	Scenario-based projections using PM2.5 datasets from 11 global climate models under Shared Socioeconomic Pathways (SSPs).	Existing projections lack precision in linking air quality improvements with actionable urban emission control strategies and localized forecasting models.	Even optimistic sustainability-driven pathways fail to fully meet SDG 3.9; aggressive regional mitigation and accurate forecasting are essential.
Nitinattrakul & Lalitaporn, 2023 [16]	Comparative analysis of ground-based air quality stations and satellite-derived pollution datasets during COVID-19 lockdown phases.	Forecasting frameworks often ignore period-specific activity disruptions and seasonal climate influences on pollution.	Pollution levels fell during lockdown due to reduced mobility, but seasonal climatic factors still caused variation — demonstrating the dynamic nature of pollution systems.
Vitali et al., 2023 [17]	Development of standardized evaluation and benchmarking procedures for short-term air quality forecasting models across Europe.	Lack of common validation benchmarks makes it difficult to compare performance across forecasting models and regions.	Proposed benchmark-based model evaluation improves transparency, comparability, and policy decision support in environmental forecasting systems.
Soleimanpour & Alizadeh, 2023 [18]	Long-term climate reanalysis using ERA5 and MERRA-2 datasets to assess planetary boundary layer height and ventilation effects on PM2.5.	Previous forecasting models rarely incorporate boundary layer dynamics and long-term climatic variability, which strongly affect pollution levels.	Winter pollution is highest due to shallow PBLH; gradual climate warming may marginally improve ventilation, influencing long-run pollution patterns.
Ghose & Anthopoulos, 2022 [19]	Designed a Hybrid 1D-CNN + BiGRU deep learning model with missing data handling for AQI forecasting in smart cities.	Need for models that jointly capture spatial correlations and temporal dependencies in pollution time-series while maintaining robustness to missing values.	Hybrid 1D-CNN + BiGRU model provided superior forecasting accuracy compared to standalone DL models, proving effectiveness of integrated architectures.

III. UNDERSTANDING AIR POLLUTION DYNAMICS

Air pollution is a complex environmental issue that occurs when the combination of natural and man-made objects alters the amount or distribution and transformation of impurities in the air. There are the emission sources, atmospheric processes, weather conditions, geographic features, all of which contribute to changing air pollution with time. Some of the natural causes are dust storms, volcanic activity, forest fires, and discharge of pollen. Vehicle emissions, industrial activities, burning fossil fuels, burning garbage, and building work are some of the key causes which are made by humans. Some of the most significant air pollutants include particulate matter (PM2.5 and PM10), nitrogen dioxide (NO 2), sulfur dioxide (SO 2), carbon monoxide (CO), ozone (O 3), ammonia (NH 3) and several volatile organic compounds (VOCs). On the air, these pollutants are capable of chemically and physically reacting with one another. This has the potential to create other types of pollutants known as the secondary pollutants, including ground level ozone and the secondary particulate matter, which can be even more harmful than the original emissions. [20].

Corresponding Author’s E-mail ID: rituguupta414@gmail.com

Weather conditions have a significant influence on pollutant dispersion and accumulation since temperature, humidity, speed of wind, air pressure, and rainfall are significant factors. As an example, pollutants cannot be distributed much when the wind is less in strength, and temperature inversion makes pollutants near the surface, which may result in smog. Seasonal variations are also very significant. As an example, in the winter season, the levels of pollution are typically greater due to an increased amount of heat consumed by people, lesser mixing of air, and stagnant air. The role of geographic considerations is also involved. As an illustration, cities bordered by mountains or those in the lowland regions can experience reduced airflow thus making the pollutants remain in the air. In addition, variations in air pollution with time are different at varying times. An hourly change could be brought about by change in traffic or industry, a daily or seasonal change could be brought about by weather and human habits. Some chemical reactions of the pollutants also require time, in particular, the ones which are the result of sunshine, like the formation of ozone of other chemicals. Due to the interaction of many components that are continuously varying, air pollution is extremely difficult to predict due to the distortion of its concentrations which vary nonlinearly. You should be able to understand these dynamics so that you can develop good forecasting models with the ability to capture time patterns, variability across regions and a complex behaviour of pollutants. It has become increasingly popular with advanced models of computation, particularly deep learning models, which are capable of revealing hidden relationships in large and multivariate datasets of the environment. With a combination of historical pollution data and weather factors, these types of models can identify trends and relationships that are difficult to identify by traditional statistical techniques. This will enable more precise predictions and intelligent environmental management.[21], [22], [23].

IV. TRADITIONAL STATISTICAL AND MACHINE LEARNING APPROACHES

Long time air pollution level prediction relied on diverse statistical as well as classical machine learning models which attempt to model changes in concentrations of pollutants with time. Some of the initial techniques included Autoregressive Integrated Moving Average (ARIMA), Multiple Linear Regression (MLR) and Kalman Filtering since they were easy to use and understand. ARIMA models achieve the best results in the case of short term predictions when the levels of pollution are consistent between seasons or in cycles. Yet their greatest failure is that they cannot faithfully reflect nonlinear associations or unpredictable variations in air quality in reaction to the varying weather or unpredictable emissions. Also, the regression-based techniques presuppose that there is a straight-line relation between variables, which is not necessarily there in the changing atmospheric interactions[24], [25], [26], [27], [28]. The increasing popularity of machine learning models such as Support Vector Machines (SVM), Random Forest (RF), Artificial Neural Networks (ANN), and k-Nearest Neighbors (kNN) are increasingly being used by more and more people as computers become more intelligent at their tasks. The models are more robust when addressing nonlinear relationships and multivariate relationships and therefore are more accurate in prediction of things as compared to the traditional statistical methods. SVM is, say, a useful method when dealing with a very dimensional dataset and it is insensitive to outliers. Prediction by Random Forest, however, is more reliable since the several decision trees were combined to prevent overfitting. Introduction ANN-based models added the feature of modeling intricate dependencies between pollutants and meteorological parameters, but usually require huge training sets and parameter optimization. Despite these technologies, the temporal relationships which are induced within air pollution time-series data are frequently difficult to learn using typical machine learning models. They typically examine the input values in sequence and they lack mechanisms to detect the occurrence of trends. Due to this, their prediction capabilities might be impaired in cases where they are required to make predictions over a long time or in weather with a rapidly changing character. Such issues gave rise to the shift to deep learning-based algorithms, particularly the recurrent neural network (RNN), Long Short-Term Memory (LSTM) network, and the Gated Recurrent Unit (GRU), which are designed with the express purpose of modeling sequential dependencies [29], [30], [31].

TABLE 1. COMPARISON OF TRADITIONAL STATISTICAL AND CLASSICAL ML MODELS FOR AIR POLLUTION FORECASTING

Model Type	Example Models	Key Strengths	Major Limitations
Statistical Models	ARIMA, MLR, Kalman Filter	Simple, interpretable, good for short-term stable trends	Poor handling of nonlinearity, limited adaptability to rapid changes
Machine Learning Models	SVM, Random Forest, ANN, kNN	Better modeling of nonlinear and multivariate relationships, higher accuracy	Lack of temporal dependency modeling, sensitive to data volume and parameter tuning

The table presents the key differences between the classical and statistical machine learning models. The statistical models are effective where the pollutants act in a stable manner, but not in a dynamic atmospheric condition. Machine

Corresponding Author’s E-mail ID: rituguupta414@gmail.com

learning models are improved in nonlinearity and multiple-feature learning, yet even they are not able to capture fully sequential patterns. It is the reason why more developed deep learning architectures are required.

V. ADVANCEMENTS IN DEEP LEARNING AND HYBRID ARCHITECTURES

Recent improvements in deep learning have created a profound shift in the predictive models of atmospheric and environmental time-series modeling, as well as air pollution forecasting. Conventional machine learning models tend to be inefficient in modeling nonlinearity, time-varying performance of pollutant concentrations due to meteorological conditions and variable emissions. On the other hand, the idea with deep learning models is that they are programmed to learn complex patterns of features, temporal relationships and hidden dependencies through raw or minimally processed data and achieve significant gains in predictive performance and generalization. [32].

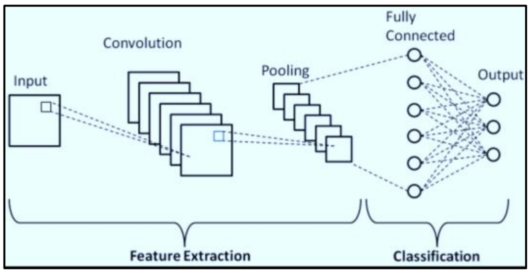

Fig. 2 CNN Architecture

Deep Learning for Capturing Nonlinear Atmospheric Behavior Air pollution forecasting needs the models, which will be able to deal with the nonlinear and highly dynamic relations between pollutants concentrations and meteorological conditions. This is especially appropriate in deep learning models that are able to learn hidden features and complicated temporal dependencies based on raw data. In contrast to classical models that are based on extensive use of handcrafted feature engineering, deep learning models are used to analysis multivariate time-series data to identify patterns associated with the pollutant trends, sources of emission, and seasonal impacts, as well as boundary-layer changes in the atmosphere. Deep learning can be a very potent tool in environmental monitoring because it is capable of modeling nonlinearity and uncertainty.[33].

Role of Convolutional Neural Networks (CNNs) in Feature Extraction Conventional Neural Networks (CNNs), particularly 1D CNNs, have been found to be very useful in the field of environmental time-series. They can be used to identify the localized temporal properties of sudden spikes in emissions, short-period of time pollution accumulation, and diurnal variation tendencies by applying filters to continuous chains of pollutant and meteorological data that simplify input complexity. This renders them a perfect front-end feature extractor to hybrid forecasting architectures which demand powerful initial data representation.

1. Recurrent Neural Network Variants for Modeling Sequential Dependencies The purpose of recurrent Neural Networks (RNNs) and their extended versions LSTM and GRU is to operate with sequential data when previous states affect the final results. LSTM networks address the problem of vanishing gradient and are able to capture long-duration temporal dependencies as opposed to GRUs, which train much faster and require fewer parameters with only minor performance degradation. Subsequent developments resulted in a Bidirectional LSTM (BiLSTM) and Bidirectional GRU (BiGRU) models that model sequences in both directions and therefore are able to encode the past and detect the future simultaneously. This two-way processing is very useful in the prediction of the pollutant formation cycles that are conditional on the change in atmospheric conditions [34].
2. **Transformer-Based Models for Global Temporal Attention** Recently, Transformer architectures, which include self-attention, have become popular in air quality prediction. Transformers are not based on sequential processing of data and are able to directly learn long-range temporal dependencies on large datasets. They are also better at forecasting with their capacity to prioritize significant time steps, especially with complicated seasonal and meteorological variations. Their complexity of computation renders them, however, too difficult to use in practice on a large scale in real-time [35].
3. **Hybrid 1D CNN–BiGRU Models for Enhanced Prediction Performance** The hybrid models combine the Congruence of CNNs and Bidirectional GRUs to determine higher forecasting precision. In these structures, 1D CNN layers extract meaningful time and trend-related features and the representations obtained are transferred to BiGRU layers to acquire contextual temporal patterns. This combination has an increased level of robustness, lower levels of noise sensitivity, and further information about the behavior of pollutants over time. Therefore, crossbreed deep learning models are the best prospects of creating precise, consistent, and scalable air pollution prediction models. [36].

VI. CHALLENGES AND LIMITATIONS

Despite the fact that forecasting models based on deep learning have shown substantial progress in terms of modeling intricate pollutant dynamics and patterns across time, a number of challenges and limitations still extend their performance, practical application, and generalizability. These limitations can only be understood on how to enhance further research and how to enhance the reliability of models within the real life system of environmental monitoring[37].

1. **Data Quality and Availability Constraints** The accessibility and quality of quality datasets are considered as one of the most important issues in the air pollution forecasting. Most monitoring stations have an intermittent reporting, lost values, sensor noise, equipment failure, or non-consistency in calibration. The scarcity of data is further amplified in the areas where there are fewer monitor networks, which results in less accurate models. The models of deep learning assume that the large, continuous and well-distributed datasets are essential to the successful training; hence, the lack of sufficient and high-quality data explicitly constrains the performance of prediction. Also, local differences in the source of emissions imply that the model that was trained in one territory might not be generalized to a different territory.
2. **High Sensitivity to Meteorological Variability** Meteorological factors that include temperature, humidity, wind direction, and wind speed, and atmospheric pressure have a significant impact on the dynamics of air pollution. These parameters vary within short time scales and are usually nonlinear. Even when the deep learning models can be trained to capture time-dependent features, abrupt changes in the weather, the unexpected weather scenarios, or other seasonal deviations may result in serious forecasting mistakes. Such random changes necessitate an adaptive nature of models to capture them, and retraining, thus complicating computational and maintenance. [38].
3. **Complexity in Model Training and Parameter Optimization** Hybrid or multi-layered deep learning architectures have many hyperparameters including learning rate, the number of layers, and filter sizes, activation functions, and dropout ratios. It can be time-consuming to find the best combinations of them through lots of experimentation and computation. Inappropriate parameters can cause overfitting, underfitting or unstable training behavior. Moreover, deep learning models are computationally expensive and can run on specialized hardware (GPUs/TPUs), which makes them both difficult to deploy and resource-intensive to researchers and agencies [39].
4. **Interpretability and Transparency Limitations** Although the deep learning models provide better accuracy of prediction, the models are usually treated as black boxes, that is, the way the models make their decisions is hard to interpret. Such inability to be interpreted is problematic to policymakers and environmental scientists who would like to have clear models to warrant regulatory or mitigation action. Knowledge of the variables that have the most significant effect on pollution trends is essential to successful intervention planning but the complexity of deep learning makes simple interpretation unfeasible unless post-hoc explainability methods are used [40].
5. **Scalability and Real-Time Deployment Challenges** Real time air quality forecasting systems involve the implementation of models that are accurate as well as efficient in terms of the speed of inference. Multifaceted designs such as BiLSTM, BiGRU and Transformer-based models might require a lot of processing power, and thus, cannot be utilized in real-time contexts that require continuous implementation, particularly those constrained by resources. Also, the combination of live sensor data streams, anomalies, and stable operation of systems also makes large-scale deployment more challenging. [41].

VII. CONCLUSION

This review has highlighted the continuous evolution of air pollution forecasting methods, demonstrating how the field has progressed from basic statistical approaches to more sophisticated deep learning and hybrid architectures. Traditional models such as ARIMA, Multiple Linear Regression, and Kalman Filtering provided initial frameworks for analyzing pollutant concentration trends, particularly under stable and predictable atmospheric conditions. However, their inability to accurately represent nonlinear interactions and rapidly changing environmental patterns limited their effectiveness in real-world urban scenarios. Classical machine learning approaches, including Support Vector Machines, Random Forests, kNN, and Artificial Neural Networks, advanced prediction capability by capturing multivariate relationships more effectively, yet they still struggled to learn long-term temporal dependencies inherent in pollution time-series data. As highlighted in this review, recent advancements in deep learning have significantly improved forecasting accuracy by integrating automated feature extraction and sequential pattern learning. Hybrid models that combine convolutional neural networks with recurrent architectures, particularly GRU or LSTM variants, demonstrate substantial potential in capturing both localized emission spikes and long-term pollutant trends influenced by meteorological conditions. Nevertheless, challenges persist in terms of data availability, sensor inconsistencies, computational cost, interpretability, and real-time deployment needs. Therefore, future research should focus on developing scalable, adaptive, and explainable air pollution forecasting systems capable of operating reliably across diverse geographic and climatic settings. Enhanced data fusion methods, incorporation of satellite and IoT data streams, and interpretable deep learning frameworks will be essential to bridge the gap between high predictive accuracy and practical applicability. Ultimately, this review emphasizes the need for forecasting models that are not only precise but also transparent, efficient, and supportive of informed environmental policy and public health decision-making.

Corresponding Author's E-mail ID: rituguupta414@gmail.com

REFERENCES

[1] S. Saminathan and C. Malathy, "Ensemble-based classification approach for PM . concentration forecasting using meteorological data," 2022.

[2] U. Im *et al.*, "Reviewing the links and feedbacks between climate change and air pollution in Europe," no. September, pp. 1–19, 2022, doi: 10.3389/fenvs.2022.954045.

[3] G. T. Cirella, "Financing Costs and Health Effects of Air Pollution in the Tri-City Agglomeration," vol. 10, no. March, pp. 1–10, 2022, doi: 10.3389/fpubh.2022.831312.

[4] L. Montalvo, D. Fosca, D. Paredes, M. Abarca, and C. Saito, "An Air Quality Monitoring and Forecasting System for Lima City With Low-Cost Sensors and Artificial Intelligence Models," vol. 4, no. July, pp. 1–15, 2022, doi: 10.3389/frsc.2022.849762.

[5] M. J. Wolf *et al.*, "New Insights for Tracking Global and Local Trends in Exposure to Air Pollutants," 2022, doi: 10.1021/acs.est.1c08080.

[6] K. Vohra *et al.*, "Rapid rise in premature mortality due to anthropogenic air pollution in fast-growing tropical cities from 2005 to 2018," vol. 4435, no. April, 2022.

[7] S. Aguilar-gomez, J. S. G. Zivin, M. J. Neidell, and H. Dwyer, "No Title," 2022.

[8] K. Theodorakos *et al.*, "Island Transpeciation : A Co-Evolutionary Neural Architecture Search , applied to country-scale air-quality forecasting," vol. 14, no. 8, pp. 1–16, 2021.

[9] G. Huang, C. Ge, T. Xiong, and S. Song, "Large scale air pollution prediction with deep convolutional networks," vol. 64, no. September, 2021.

[10] S. He, G. Jang, and Y. Liu, "Hybrid Transformer-TimesNet Model for Accurate Prediction of Industrial Air Pollutants : A Case Study of the Xinyang Industrial Zone , China," vol. XX, no. 2, pp. 1–21, 2025, doi: 10.15244/pjoes/203348.

[11] A. Dairi, F. Harrou, S. Khadraoui, and Y. Sun, "Integrated Multiple Directed Attention-based Deep Integrated Multiple Directed Attention-based Deep Learning for Improved Air Pollution Forecasting," 2025, doi: 10.1109/tim.2021.3091511.

[12] V. Oldenburg and J. C. Matias, "Forecasting Smog Clouds With Deep Learning : A Proof-Of-Concept," 2024.

[13] A. Soulie *et al.*, "Global anthropogenic emissions (CAMS-GLOB-ANT) for the Copernicus Atmosphere Monitoring Service simulations of air quality forecasts and reanalyses," pp. 2261–2279, 2024.

[14] S. Africa, C. Republic, and N. Surgery, "Heliyon Possibilities and limits of modelling of long-range economic consequences of air pollution – A case study," vol. 10, no. July 2023, 2024, doi: 10.1016/j.heliyon.2024.e26483.

[15] H. Yue *et al.*, "Substantially reducing global PM 2 . 5 -related deaths under SDG3 . 9 requires better air pollution control and healthcare," pp. 1–13, 2024, doi: 10.1038/s41467-024-46969-3.

[16] N. Nitinattrakul and P. Lalitaporn, "Engineering and Applied Science Research Study the change in air pollution after the COVID-19 outbreak in Thailand," vol. 50, no. 2, pp. 137–148, 2023, doi: 10.14456/easr.2023.15.

[17] L. Vitali *et al.*, "A standardized methodology for the validation of air quality forecast applications (F-MQO): lessons learnt from its application across Europe," pp. 6029–6047, 2023.

[18] M. Soleimanpour and O. Alizadeh, "Analysis of diurnal to seasonal variations and trends in air pollution potential in an urban area," *Sci. Rep.*, pp. 1–9, 2023, doi: 10.1038/s41598-023-48420-x.

[19] B. Ghose and L. Anthopoulos, "A Deep Learning based Air Quality Prediction Technique Using Influencing Pollutants of Neighboring Locations in Smart City," vol. 28, no. 8, pp. 799–826, 2022, doi: 10.3897/jucs.78884.

[20] A. Humpe, "Forecasting Air Pollution in Munich : A Comparison of MLR , ANFIS , and SVM," vol. 2, no. Icaart, pp. 500–506, 2021, doi: 10.5220/0010184905000506.

[21] A. Dutta and W. Jinsart, "Air Pollution in Indian Cities and Comparison of MLR , ANN and CART Models for Predicting PM 10 Concentrations in Guwahati , India," vol. 15, no. 1, 2021.

[22] M. G. Vivanco *et al.*, "Assessment of the Effects of the Spanish National Air Pollution Control Programme on Air Quality," 2021.

[23] J. De Bont *et al.*, "Articles Ambient air pollution and daily mortality in ten cities of India : a causal modelling study," *Lancet Planet. Heal.*, vol. 8, no. 7, pp. e433–e440, 2021, doi: 10.1016/S2542-5196(24)00114-1.

[24] J. Weuve *et al.*, "Exposure to Air Pollution in Relation to Risk of Dementia and Related Outcomes : An Updated Systematic Review of the Epidemiological Literature," vol. 129, no. September, 2021.

[25] C. Menares, P. Perez, and S. Parraguez, "Urban Climate Forecasting PM 2 . 5 levels in Santiago de Chile using deep learning neural networks," vol. 38, no. June, 2021, doi: 10.1016/j.uclim.2021.100906.

[26] D. Silva, Y. De Souza, T. Antonini, and H. Valadares, "Unorganized machines and linear multivariate regression model applied to atmospheric pollutant forecasting," vol. 42, 2020.

[27] A. Kallel, "Jo ur l P re of," 2020.

[28] Q. Guo *et al.*, "Air Pollution Forecasting Using Artificial and Wavelet Neural Networks with Meteorological Conditions," pp. 1429–1439, 2020.

[29] A. Ukhov, S. Mostamandi, A. Silva, J. Flemming, and Y. Alshehri, "Assessment of natural and anthropogenic aerosol air pollution

Corresponding Author's E-mail ID: rituguupta414@gmail.com

in the Middle East using MERRA-2 , CAMS data assimilation products , and high-resolution WRF-Chem model simulations," vol. 5, pp. 9281–9310, 2020.

[30] M. Arsov, E. Zdravevski, R. Corizzo, N. Koteli, and K. Mitreski, "Short-term air pollution forecasting based on environmental factors and deep learning models," vol. 21, pp. 15–22, 2020, doi: 10.15439/2020F211.

[31] A. C. Science, "PM10 FORECASTING THROUGH APPLYING," vol. 3, no. 1, pp. 31–43, 2020, doi: 10.2495/EI-V3-N1-31-43.

[32] Y. Yarragunta, S. Srivastava, D. Mitra, and H. Chandra, "Influence of forest fire episodes on the distribution of gaseous air pollutants over," *GIScience Remote Sens.*, vol. 57, no. 2, pp. 190–206, 2020, doi: 10.1080/15481603.2020.1712100.

[33] F. Jammoul, T. Kuenzer, E. Stadlober, and H. Siegfried, "Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID- 19 . The COVID-19 resource centre is hosted on Elsevier Connect , the company ' s public news and information website . [34] Y. A. Y. Okkaoğlu and E. G. M. E. Yücel, "Estimation and forecasting of - PM 10 air pollution in Ankara via time series and harmonic regressions," *Int. J. Environ. Sci. Technol.*, vol. 17, no. 8, pp. 3677–3690, 2020, doi: 10.1007/s13762-020-02705-0.

[35] S. Lolli, "Impact of meteorological conditions and air pollution on COVID - 19 pandemic transmission in Italy," *Sci. Rep.*, pp. 1–15, 2020, doi: 10.1038/s41598-020-73197-8.

[36] M. So *et al.*, "A demonstration project of Global Alliance against Chronic Respiratory Diseases : Prediction of interactions between air pollution and allergen exposure — the Mobile Airways Sentinel NetworK-Impact of air POLLution on Asthma and Rhinitis approach," vol. 0, no. 13, pp. 1561–1567, 2020, doi: 10.1097/CM9.0000000000000916.

[37] M. T. Hasan, V. S. Chourasia, and S. M. Asutkar, "A Forecasting tool for Air Quality Monitoring Built up on Cloud and IoT," no. 10, pp. 3821–3832, 2019, doi: 10.35940/ijitee.J9992.0881019.

[38] M. Jia *et al.*, "Regional Air Quality Forecast Using a Machine Learning Method and the WRF Model over the Yangtze River Delta , East China," pp. 1602–1613, 2019, doi: 10.4209/aaqr.2019.05.0275.

[39] T. Borsdorff, J. De Brugh, S. Pandey, O. Hasekamp, I. Aben, and S. Houweling, "Carbon monoxide air pollution on sub-city scales and along arterial roads detected by the Tropospheric Monitoring Instrument," pp. 3579–3588, 2019.

[40] P. Tarín-carrasco, M. Morales-suárez-varela, U. Im, J. Brandt, and L. Palacios-peña, "Isolating the climate change impacts on air-pollution-related- pathologies over central and southern Europe – a modelling approach on cases and costs," pp. 9385–9398, 2019.

[41] G. P. Brasseur *et al.*, "Ensemble forecasts of air quality in eastern China – Part 1 : Model description and implementation of the MarcoPolo – Panda prediction system , version 1," pp. 33–67, 2019.

Corresponding Author's E-mail ID: rituguupta414@gmail.com