Deep Learning Techniques for Air Quality Prediction: A Focus on PM2.5 and Periodicity

Lakshmi Shankar[1] and Krishnamoorthy Arasu[2*]

Abstract

The rapid increase in traffic, urbanization, and industrial expansion has all contributed to a decrease in air quality, which has a vital impact on both the long-term feasibility of the environment and the health of humans, particularly in industrialized nations. Numerous studies have explored using machine learning for air quality forecasting to reduce pollution. While shallow machine learning architectures offer less accurate forecasts, deep learning, a recent advancement in computational intelligence, has immense potential in predicting air quality. Deep learning frameworks can identify intricate correlations and patterns in data on air quality, resulting in more accurate and dependable predictions. Several aspects, including climatic conditions, emission sources, and geographical characteristics, may be considered by these models, which can help one better understand and anticipate air pollution levels. This research investigates deep learning applications' periodic changes in air quality. Hybrid deep learning methods utilize optimization, data decomposition, and correlation evaluation between PM2.5 particles and other factors to overcome limitations. This study contrasts various deep learning algorithms for forecasts of air quality and demonstrates that hybrid deep learning is more accurate compared to each model alone at predicting future periods of air quality. It proposes future research directions for the future generation of models. The literature summary provides valuable insights for academics seeking future studies in this field.

Keywords: *Air Pollution, Air Quality Forecasting, Deep Learning, Hybrid Learning, PM2.5 Prediction.*

Introduction

Environmental problems have been caused by the continuous expansion of global urbanization and industrialization. The deterioration of air quality brought on by industrialization and urbanization is one of the most significant ecological problems (Kan et al., 2012; Samal et al., 2019). Examples of activities that produce and consume energy include businesses, power plants, and vehicle emissions. Other natural activities such as agriculture burning, volcanic eruptions, and wildfires have eventually contributed to the ongoing worsening of global air quality due to the requirements of transportation, manufacturing, and daily living (Biancofiore et al., 2017a). Air pollutants are quickly becoming a serious challenge for social development, and the environment. management, and economic management (Bai et al., 2019). Air pollution leads to an increase in lung cancer cases, severe respiratory illnesses, heart attacks, asthma, and a variety of skin problems. They may also result in more significant issues that impact the entire planet, such as climate change and global warming. As air pollutants Nitrogen dioxide (NO2), Carbon

---

[1] Research Scholar, School of Computer Science Engineering, Vellore Institute of Technology, Vellore, India. lakshmi.s@vit.ac.in

[2*] Assistant Professor (Sr), School of Computer Science Engineering, Vellore Institute of Technology, Vellore, India. Krishnamoorthy.arasu@vit.ac.in

monoxide (CO), Sulfur dioxide (SO2), Ozone(O3), and Particulate Matter (PM2.5 and PM10) are airborne pollutants (Chang et al., 2020). A variety of pollutants have an impact on the air, which lowers the quality of the air. The pollutant ratios examined using earlier recommended approaches are shown in Figure 1. The most prevalent pollutants are CO, O3, SO2, NO2, PM2.5, PM10, and O3. This graph examines around 40 studies that employ various pollutants in the pollution forecast technique (Kumar SV et al (2, n.d.).
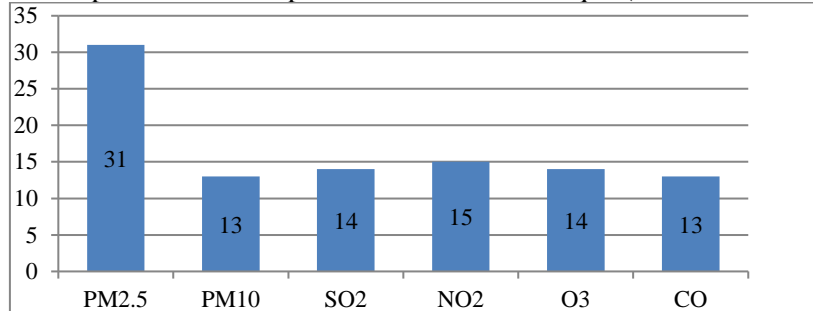

Figure 1: The Earlier technique investigated the percentage of various contaminants.

Based on the effects on human health, air pollution is categorized using the Index for Air Quality (AQI). According to the amount and duration of exposure to pollution conditions, humans may suffer a variety of fitness repercussions, which are the focus of the AQI. By each country's air quality standards, the AQI Measurements vary from one country to the next. Health problems are more likely to occur when there are greater AQI scores (Shang et al., 2019a).

Among them, the PM2.5 particle is microscopic with a diameter of under 2.5 microns. PM2.5 particles are more active than bigger particulate pollutants, which allows them to travel more rapidly. And readily through the air as well as stay in the atmosphere for a longer period. One of the most important causes of air pollution is PM2.5. Because of the small particle size, it is easily able to penetrate the human throat and nasal cavities and induce cardiovascular, bronchial, or pulmonary diseases (Vargas et al., 2018; Xing et al., 2016). Apart from these air pollutants, smoke, organic compounds, and heavy metal dust from PM2.5 People's health is seriously threatened by air pollution (X. Li et al., 2020). Estimating PM2.5 attentions is crucial for managing and lowering air pollution, which might aid in the government's capacity to effectively provide early warnings and encourage people to drive safely. The air quality prediction may also contain precise information for pollution prevention and management. The World Health Assembly of WHO estimates that PM2.5 exposure causes around seven million fatalities annually and has the most detrimental effect on human health. Numerous respiratory disorders, including problems with cardiovascular function, may arise from prolonged exposure to poor air quality. Figure 2 shows the Effects of Air Pollution (PM2.5) On the Human Body (Www.Devic-earth.com, n.d.)
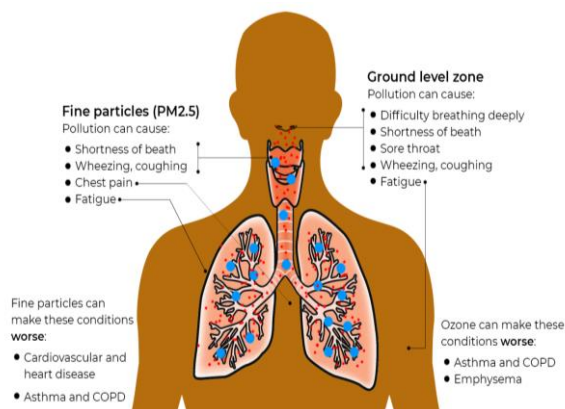

Figure 2: Effects of Air Pollution (PM2.5) On the Human Body.

Therefore, the significance of PM2.5 concentration forecasts provided by an efficient forecasting model will increase (Mehmood et al., 2022). Several natural elements often influence air quality levels, which in turn influences the analytical forecast outcomes. There is a direct relationship between weather conditions and atmospheric quality. Meteorological parameters enable rapid identification of the origins of air pollution. The following two factors may be related to the key reasons why forecasting PM2.5, an extremely complex topic in forecasting, is difficult. The primary hindrances to enhancing forecast precision are, first, the irregular and dynamic data qualities of the PM2.5 time series, which are influenced by several external factors (Ma et al., 2020; Pak et al., 2020a; X. Xu et al., 2020). Commonly occurring elements, such as climate, geography, and temporal traits, are typically shown as outside variables that contribute to the non-stationary and non-linear assets of the linked time series (Shang et al., 2019a). PM2.5 predictions, on the other hand, necessitate a high data frequency, a diverse array of PM2.5 sources, and several external variables, which may result in difficulties such as a protracted period of training and an extensive amount of data (Gugnani & Singh, 2022a).

By tackling the current challenges, a number of driven by data strategies have been created to increase forecast accuracy in this setting (Biancofiore et al., 2017b; P. Du et al., 2020; García Nieto et al., 2018a; Shang et al., 2019b). The historical forecast models may be categorized into two categories: non-linear models for forecasting and linear forecasting techniques, depending on the construction of the data-driven framework. Due to the uncertainty of predicted inputs, autoregressive integrated models with moving averages (ARIMA) (García Nieto et al., 2018b) multiple linear regression (MLR) (Moisan et al., 2018; Vlachogianni et al., 2011), and spatially and time-related weighted regression models (GWR) (B. Guo et al., 2021) have all been widely employed to predict concentration. While the non-stationary nature and non-linearity behavior of PM2.5 cannot be fitted at that time modeled by the linearity forecasting model. Additionally, the labor and computational resources needed for these approaches, which include estimating hundreds of parameters, are expensive (Zaini et al., 2022). Furthermore, the data features of PM2.5 are routinely predicted using nonlinear forecasting models such as supporting vector regression (SVR), artificially intelligent neural networks (ANN), and extraordinary learning machines. Other pollutants, such as PM10, CO, NO2, SO2, and O3, as well as weather and temporal factors, are also linked to the spread of PM2.5 (Shang et al., 2019a). This PM2.5 prediction, therefore, requires the use of several air quality monitoring indicators, meteorological variables, and temporal variables. Specifically, this condition resulted in the training and processing of large volumes of input data for classical machine-learning models, which made it more challenging to employ those models (Qi et al., 2019). Deep learning approaches, such as variable long-short-term memories (LSTM)(X. Li et al., 2017; Ong et al., 2016; X. Xu et al., 2020), recurrent neural networks (RNN) (Pak et al., 2020a)(S. Du et al., 2020), deep belief networks(Pak et al., 2020a, 2020b, 2020c), gated recurrent unit(GRU) (Zhang et al., 2021a), and convolution neural networks(CNN) (Ravindran & Gunavathi, 2023) are extensively used to anticipate PM2.5 concentrations in a manner similar to this. Deep learning neural networks improve above more traditional machine learning techniques in producing great prediction performance. The major factor is that deep learning techniques can simultaneously collect long-term and short-term characteristics, giving them a powerful modeling capacity for more external factors.

Monitoring and Predicting Air Quality has Several Challenges

While new research is making progress in the current time surveillance of air quality, there are still problems that require further attention to be fixed (Samal et al., 2021).

- Real-time Accurate concentrations should be provided by taking into consideration varied weather circumstances; therefore, it is important to create AQ monitoring methods that are reliable and practical.(Gugnani & Singh, 2022b). It is essential to look at the nonlinearity and instability of the system.
- It is essential to look at the nonlinearity and instability of the system.

- Incorporating the machines into a constantly accessible, manageable internet connectivity.
- The requirement to use data on continuously monitored air quality to enhance short- and long-term forecasts while accounting for all factors influencing them.
- To achieve the best performance, use hybridization.

The generic Air quality forecasting model is shown in Figure 3. It forecasts future air quality indices experimentally using historical AQIs, thresholds, and meteorological data. Unidentified environmental factors may produce abrupt fluctuations in air quality that are beyond its control. Air pollution is monitored at the AQ observation station, which also gathers data and sends it to a server for study. The apparatus is equipped with humidity, temperature, atmospheric pressure, gas, and dust particle sensors. The information is sent to the cloud for analytics after being shown as an AQI in an internet application. The Air Quality Surveillance System is safeguarded by Comprehensive Operation and Maintenance Control, and IMD and EMRC do daily inspections.
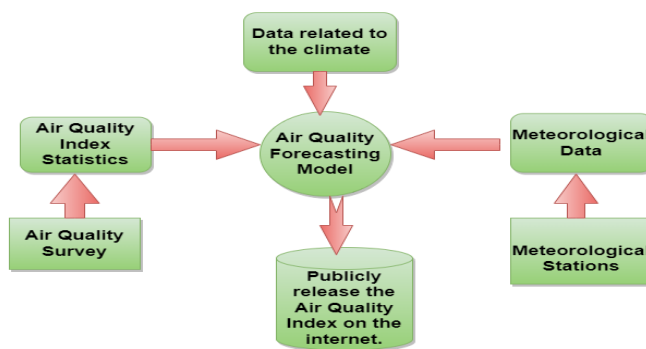

Figure 3: A generic perspective on the air quality forecasting framework

Monitoring stations transmit air quality and meteorological information to the control room. The Central Control Room converts AQMS data to AQI before transmitting it to each region's Internet site along with weather-related data and air quality predictions. The operating server of the DDS system receives the information from the file transfer protocol (FTP) server and delivers it to the Digital Display Board in each city (Hable-Khandekar & Srinath, 2017).

Related Work

There is a lengthy history of research on air quality forecasting, and most current publications use statistical methods and simple machine learning models, such as regression, ARIMA (Díaz-Robles et al., 2008), HMM (Dong et al., 2009), and Artificial Neural Networks (Q. Zhou et al., 2014), to handle air quality prediction challenges. The frequency of air quality predictions derived from large-scale information analysis has increased in several years. More researchers are aiming to apply data-driven techniques because of the dynamic and nonlinear character of time-series information linked to air quality, especially in urban analysis. Air quality has been predicted using a vast quantity of data in a variety of ways, which may help with air pollution warnings and management (Yi et al., 2018). Zhou et al. used ensemble methods to create a hybrid approach for a single-day PM2.5 forecast. Zheng et al. created a semi-supervised learning environment for air quality predicting that utilized co-training architecture with two different classifications (CRF and ANN) . Luo Zhang proposed a semi-supervised method that used EMD and BiLSTM neural networks to estimate PM2.5 levels. To effectively manage missing data in ST forecasting tasks. Junxiang Fan et al created a unique DL framework and a DRNN setup using LSTM. To forecast PM2.5 ST variability, Yanlin Qi et al [54] created a DL hybrid framework incorporating LSTM and Graph CNN. Dr. Sankari and S. Jeya estimated RMSE, MAE, and SMAPE using a bidirectional LSTM model to predict illnesses linked to PM2.5(Jeya & Sankari, 2020).

Theoretical Frameworks and Terminology of Deep Learning

Deep networks employ weighted connections between neurons in the network to create representations of knowledge in high-dimensional domains. DL has been successfully utilized in a variety of disciplines, including computer vision, natural language processing, and speech recognition, as well as the sciences of physics and chemistry. Several deep learning architectures might be created using the existing quantity of time series air quality data. These architectures may be trained to forecast air quality levels in various areas, providing useful information for environmental monitoring and pollution control operations. Deep learning algorithms may also be used to uncover patterns and connections in data that standard statistical approaches may miss, resulting in a more thorough knowledge of air pollution dynamics (Lee et al., 2018). The most often utilized deep learning approaches in air quality forecast (DBM) include long-short-term memory (LSTM) and recurrent units with gates (GRU), Gated Recurrent Unit (RNN), neural networks with convolution (CNN), deep belief networks (DBN), and deep Boltzmann methods.

RNN: Recurrent Neural Network

This kind of network feeds its current step's output into the preceding stage. In a typical neural network, each node's identification is its input/output (I /O). However, in certain scenarios-like anticipating a statement's next phrase the past nodes are also required, necessitating the recall of those nodes. To address this issue, RNN was developed using a hidden layer. The hidden state, which retains very little information about a topic, is the primary and most important feature of RNN. It has an "archive" where all measurement data is kept. This generates the output by performing a similar operation on all the inputs or invisible layers; hence, it applies the same parameters to each input. This network minimizes parameter complexity, in contrast to others. Figure 5 illustrates RNN's network architecture.
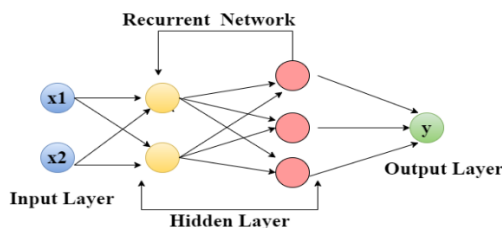

Figure 4: RNN's network architecture

RNNs are capable of processing sequential data and producing sequential output. Relationships are discovered using RNN using an activation function that is nonlinear, and the backpropagation process is used to update network values. (Sánchez-Balseca & Pérez-Foguet, 2020). Discovering connections between distant events is very difficult using the BPTT methods due to its susceptibility to vanishing gradients caused by several derivative runs, which yields very little update.

In recent years, RNN has gained global recognition for the quality of air predictions due to its capability to absorb serial data. (Arsov et al., 2020). When it comes to predicting PM2.5 concentrations, a periodically already-trained deep neural network (RNN) forecasting model outperforms existing ANN models like RNN and FFNN. If training takes too long, RNNs may have trouble detecting long-term dependent state in input data and may experience vanishing and ballooning gradients. (Liao et al., 2020).

Long- and Short-term Memory (LSTM) and Gated Recurrent Unit (GRU)

The RNN by Hochreiter and Schmidhuber has been updated into LSTM. To solve the problem and enhance network performance, notably in air quality forecasting, better frameworks have been developed, such as LSTM as well as GRU. Implementing gating

systems like LSTM fixed the vanishing gradient issue during BPTT updates. These systems let nodes forget or pass memory when not in use, maintaining enough error for updates. On input data, the LSTM employs trained gates and feedback loops (Freeman et al., 2018). The multiplicative gates determine whether incoming data must be stored in memory and control how the blocks operate. While the input gate controls the cell activation flow from input into the memory block's function. While the gate of input via the memory cell into further nodes (Freeman et al., 2018). These gates are essential for managing the information flow within a neural network. They enable the network to retain important information and filter out irrelevant input by selectively allowing or blocking the passage of data. For the network to efficiently process and store information in its memory cells, this gating mechanism is necessary. Figure 5 depicts the generic network architecture of an LSTM unit, which consists of a memory cell and three memory gate units.

Although GRU networks are a more straightforward version of LSTM networks, they are still useful for modeling sequential data. GRU networks have a single state that combines memory and hidden states, as opposed to LSTM networks, which have separate memory and hidden states. In comparison to LSTM networks, GRU networks are therefore simpler to train and have lower computational expenses (Huang et al., 2021). Regarding faster calculation times and better performance, GRU can also be superior to the LSTM algorithm. With correlation values ranging from 0.93 to 0.97 for five different input parameter combinations, LSTM also generates precise air pollution predictions (Samal et al., 2021). The GRU network, which comprises the update gate and reset gate, is seen in Figure 6.

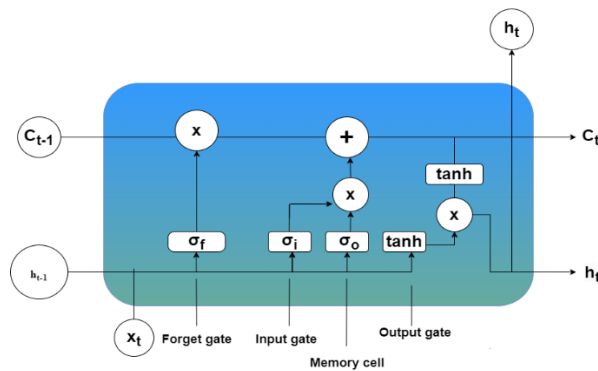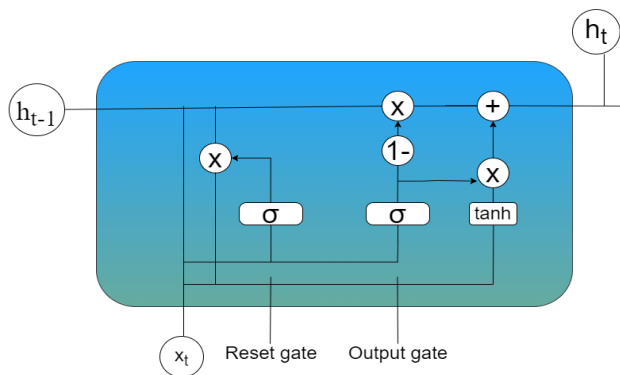

Figure 5: Generic network architecture of the LSTM

Figure 6: Network structure of the GRU

The Neural Network of Convolution

Convolutional neural network, also referred to as CNN, is widely regarded as the most effective method for spatial feature extraction, and it has been utilized in cutting-edge Computer Vision research (H. Zhou et al., 2021a). While CNN is most recognized for its

spatial or 2D data performance on 2D or spatial data, it can also function on 1D and 3D arrays. CNN networks consist of multiple layers, involving convolution, dropout, max pooling, and fully interconnected layers, with each layer having a three-dimensional structure consisting of height, breadth, and depth. CNN model aids in learning the latent features of a time series dataset and identifies spatiotemporal correlation within the dataset, CNN's two primary functions are weight distribution and time series data compression(Tariq et al., 2021a). CNN is suitable for time-series forecasting because it provides extended convolutions, which allow filters to compute extensions between cells. The amount of space between cells helps the neural network better understand the links between the time-series observations. The ability of CNN to recognize patterns and trends over time depends on this understanding of temporal dependencies. Further enhancing their forecasting abilities, CNNs are well-suited for time-series analysis due to their capacity to learn pertinent features from the data automatically. CNN-based models accurately estimate pollution levels and produce excellent forecasting results in terms of error assessment. The general network architecture of CNN is seen in Figure 7.
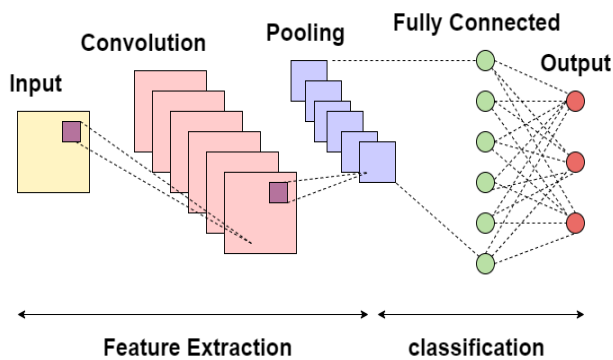

Figure 7: The general architecture of CNN

Deep Belief Networks (DBN)

Restricted Boltzmann Machines (RBMs), with an observable layer and a hidden layer are layered to produce DBNs, which are probabilistic generative models (Kow et al., 2020). There are no connections inside any of the visible or concealed layers, just connections between the layers. By training the weights between neurons, we can maximize the likelihood that the entire neural network will generate training data. The DBN model may be used to create data in addition to classifying and identifying characteristics in data. According to Le Roux and Bengio's research, if the buried layer's number of neurons is high enough, RBMs can suit any discrete distribution (Janarthanan et al., 2021). DBNs can thus potentially model intricate and varied datasets, making them an effective tool for a variety of machine learning tasks. It can also be created by stacking RBMs, enabling hierarchical data representations and the capture of higher-level features. Figure 8 shows the Structure of Deep Belief Networks.
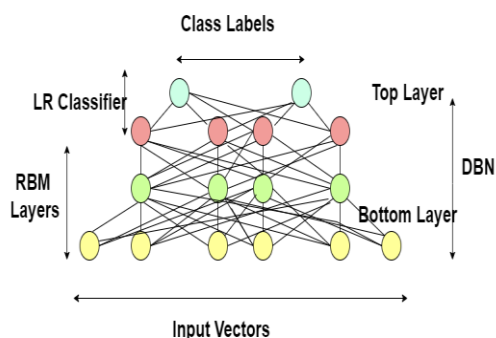

Figure 8: The Structure of Deep Belief Networks

A Summary of Deep Learning Techniques

Deep learning models, which are more accurate prediction models due to their adaptability, have been created by artificial intelligence. These models extract latent features from large datasets and conduct temporal analysis, enticing time series analysis researchers and policymakers. The text investigates the air quality forecasting capabilities of deep learning classifiers including RNN, LSTM, CNN, GRU, and DBN. The most often used assessment criteria are RMSE (Root Mean Square Error) and MAE (Mean Absolute Error). The most often used assessment criteria are RMSE and MAE. In terms of predicting accuracy, the research finds that algorithms using deep learning work better than predictive machine learning models (Jiang et al., 2021a). For example, by lowering the root-mean-square error (RMSE) by 18.35% and forecasting hourly PM2.5 levels, LSTM improves the XGBoost model's performance. When selecting hyperparameters for deep learning algorithms, caution must be taken since huge networks might impede training and cause overfitting. The learning rate, batch size, and regularization intensity are important characteristics. To prevent overfitting, it's essential to strike a balance between model complexity and generalization potential. Using cross-validation and early halting are two methods that may be used to identify the ideal hyperparameter values. The appropriate network design for dependable input characteristics must be chosen to improve prediction accuracy and decrease error. Results may be enhanced by experimenting with activation functions and optimization methods. Preprocessing processes and consideration of training data availability and quality may improve input characteristics. Table 1 demonstrates the Previous research on deep learning-based air pollution prediction.

Hybrid designs may handle complex problems and increase forecasting accuracy. CNN and RNN combined with external inputs like meteorological data or satellite images may give context and increase predicting accuracy while simultaneously capturing spatial and temporal connections.

Table 1: Previous research on deep learning-based air pollution prediction.

Ref	Examine Region	Method Applied	Target Pollutants	Granularity of Time	Predictive standards
(Ma et al., 2019a)	Beijing	LSTM and GRU, (SHAP) method	PM2.5, PM10, SO2, Metrological condition.	Hourly basis	MAE, MSE, RMSE, R2
(Yang et al., 2022)	Shanghai	iDeep Air architecture	PM2.5	Hourly basis	RMSE, MAE
(W. Du et al., 2023)	Beijing	TSVR, MTSVR	PM2.5, PM10, O3	Hourly basis	RMSE, NMGE
(Tariq et al., 2021b)	Talcher, India, Beijing	KNN, CNN–LSTM, BIGRU, Convolutional LSTM–SDAE (CLS) Sparse Denoising Autoencoder.	PM2.5	Time (Day)	MSE, RMSE, MAPE.
(Xayasouk et al., 2020)	South Korea	TL-ResNet, RNN	PM2.5	Hourly basis	RMSE, MAE, MAPE, R2
(Le et al., 2020)	South Korea	(LSTM+DAE) model	PM2.5, PM10	Hourly basis	RMSE
(C. Guo et al., 2020)	Korea	ConvLSTM	PM2.5	12 h	RMSC
(Zhang et al., 2021b)	Shanghai	(RNN+GRU+LSTM) model	PM2.5	Hourly Basis	MAE, MAPE
(S. Li et al., 2022)	Beijing	EMD-BiLSTM.	PM2.5	24 h	RMSE
(Ma et al., 2019b)	China	LSTM with Bayes	PM2.5	Hourly basis	MAE, MAPE

The Hybrid Deep Learning Model for Forecasting

The models that integrate deep learning with machine learning (CNN-BPNN), and several deep learning methods (CNN-LSTM), ensemble learning (RNN-Bootstrap), and an optimization algorithm (LSTM-BO) are examples of hybrid deep learning-based models (Jiang et al., 2021b). These models can be used as an alternative to more difficult

computational problems and can improve the forecasting capabilities of more basic models. Convolution neural networks (CNN) and Backpropagation neural networks (BPNN) a hybrid model, are utilized to estimate regional multistep ahead PM 2.5 concentrations. It was put out by Know et al. (2020) and has demonstrated improved learning performances with significant characteristics from numerous input variables and generating precise air pollution forecasting. Regarding RMSE & R2 accuracy, comparative tests reveal that CNN-BPNN performs better than the separate RF, BPNN, and LSTM models.   Additionally, CNN reduces the MAE value of the BPNN model for forecasting 10 hours in advance by 27% during training and 28% during testing. The AQI in a large metropolis is classified and predicted using a combination of the LSTM and SVR techniques. According to Janarthanan et al's publication in 2021, it outperforms other models based on deep learning, such as LSTM, RNN, and a hybrid EMD-CNN with the greatest merit of R (0.97) and least RMSE (10.9) when it comes to AQI predictions (Janarthanan et al., 2021). A new hybrid deep-based ambient air quality prediction technique (DAQFF) for PM2.5 forecast is introduced by Shengdong Du et al. DAQFF combines networks of LSTM with one-dimensional CNNs to learn patterns of spatial-temporal dependencies and correlations from multivariate data. Studies reveal that standard shallow learning and deep learning models perform better (S. Du et al., 2021). Table 2 outlines the deep learning-based hybrid forecasting model that has been developed for air quality forecasting.

Table 2: Previous research on the prediction of air pollution using Hybrid deep learning.

Ref	Examine Region	Method Applied	Target Pollutants	Granularity of Time	Predictive standards
(S. Du et al., 2021)	Beijing	DAQFF	PM2.5	Hourly basis	MAE, MSE, RMSE, SMAPE
(Bhanja & Das, 2021)	Delhi	HDNN	PM2.5	Hourly basis	RMSE, MAE, SMAPE
(R. Xu et al., 2023)	China	WT former	PM2.5, PM10	Hourly basis	RMSE, MAE, SMAPE
(Kow et al., 2020)	Taiwan	BPNN-CNN	PM2.5	Time (Day)	MAE, RMSE, R2
(Janarthanan et al., 2021)	India	SVR-LSTM	PM2.5, NO2, SO2, CO, O3.	Hourly basis	R2, RMSE, MAE, MAPE
(Ma et al., 2019c)	China	BLSTM-TL	PM2.5	Hourly basis	RMSE, R2
(Sharma et al., 2020)	Australia	CNN-LSTM	PM2.5	1 Hourly basis	RMSC, MAE, MAPE
(L. Li et al., 2020)	USA	Bootstrap method -AE	PM2.5	Hourly Basis	RMSE, R2
(Dun et al., 2022)	Fushun, China	DGRA, STCN	PM2.5	Hourly Basis	RMSE, MAE, R2
(Zeng et al., 2022)	Beijing, China	NLSTM	PM2.5	Hourly Basis	MAE, RMSE, R2

Optimization Algorithm

Several elements, such as the number of layers, neuronal networks, and the function of activation, may influence the architecture of deep learning forecasting models that are built. An essential step in creating a deep learning model is fine-tuning the parameters or hyperparameters. It entails fine-tuning the model's hyperparameters to attain the setting hyperparameters optimally, preventing overfitting. Li et al (2017) conducted experiments to find the ideal neuron count and delay times for network architecture. Error analysis is used to assess the impact of every hyperparameter combination, and the parameter values are selected at random. This process requires time since the evaluation has to be done again to get the desired outcome (Zaini et al., 2022). Optimization methods can considerably reduce the time required to execute hyperparameter searches and improve forecasting models. Ma et al. (2020a) developed a hybrid layer linked to latency LSTM (Lag-FLSTM) and Bayesian BO to forecast PM2.5 concentration. The optimal Lag-FLSTM parameters were found using BO, and an enhanced LSTM design among an added lag layer was put into practice. The study discovered that BO lowered the RMSE of Lag-LSTM in air pollution forecasting by 16.18%. Deep learning optimization methods outperform models without algorithms in terms of forecasting accuracy. In current architectures, however, metaheuristic algorithms are not especially optimized for deep learning hyperparameters.

Data Decomposition for Deep Learning Strategies

Meteorology, industrial emissions, transportation movement, and air pollution build-up all impact air quality predictions. Land use activities, particularly in cities, influence air quality. Data decomposition methods may be used to address bigger and more complicated data sequences to increase forecasting accuracy. Discrete wavelet transformation (DWT), variational mode decomposition (VMD), and Empirical mode decomposition (EMD) are some of the methods applied for data decomposition. EMD is a simple decomposition approach that extracts feature frequency without the need for predefined fundamental functions. Based on local time scales, it decomposes complicated characteristics into intrinsic mode functions (IMF).

In terms of prediction accuracy, EMD-BiLSTM outperforms BiLSTM alone, with a 38% improvement in RMSE. Compared to GRU models, hybrid EMD-GRU reduces prediction error by 44.5% for RMSE, 40.82% MAE, and 11.63 SMAP. When handling non-stationary time series, EMD significantly improves forecasting abilities and removes the temporal lag that occurs within GRU. As a means of overcoming EMD's sampling and noise sensitivity constraints, VMD splits real values into larger no sub-signals with different frequencies. The best-predicted outcomes are generated by VMD-SE-LSTM, with an R of 0.99. DWT is a technique that separates real-time series data into a number of smaller signals, also known as low-nearness signals and high-frequency detailing motions. Compared to LSTM, hybrid DWT-LSTM performs better. WPD is yet another single decomposition method that has been created to improve signal processing performance and address the disadvantage of WD in high-frequency zones. In general, the efficiency of disintegration techniques for forecasting, and data processing models may be increased by using data decomposition approaches for air quality prediction. Alternative methods for example complementary EEMD (CEEMD) and enhanced CEEMD with adaptive noise (ICEEMDAN), are also available (Zaini et al., 2022).

Correlation Evaluation between PM2.5 Particles and Other Factors

An air quality forecasting approach that is greater in precision requires a correlation evaluation between meteorological substances and contaminants. Numerous variables influence PM2.5 concentrations is influenced through multiple variables. Even so, every element has a substantial impact on the problem of air pollutant prediction. For instance, high atmosphere wind speed and pressure often reduce the quantity of pollutants in the air, whereas excessive humidity might exacerbate the quality of the air. Therefore, the features of meteorological parameters are important when it comes to predicting air quality (H. Zhou et al., 2021b). It has been ascertained that Yemen's PM2.5 levels exhibit a cyclical pattern that is impacted by meteorological elements including temperature and sun radiation. In both the fall and summer samples, there is a positive link connecting temperature and PM2.5 concentrations, according to statistical analysis. Nevertheless, the investigation has also revealed a negative correlation in summer data and a positive correlation in fall samples between relative humidity and PM2.5 concentration (Hael, 2023). In addition, several other variables, including geography, temperature, and precipitation, may have a big impact on PM2.5 concentrations. Because of the complex relationships between these factors, it is essential to consider their overall impact when creating an effective forecasting model PM2.5 and PM10/SO2/CO/NO2 exhibited a link with pollutant components, with lower-frequency spectrum correlations being weaker and higher-frequency spectrum correlations being greater. When O3, temperature, pressure, speed of the wind, and moisture were mostly exhibited in the low-frequency band, humidity was reflected primarily in the slower high-frequency signals band. Figure 9 shows the PM2.5 and other contributing variables' correlation across several frequency bands (R. Xu et al., 2023). The attention matrix analysis explained the perdition behavior of the model and revealed the time-frequency law between the variables.
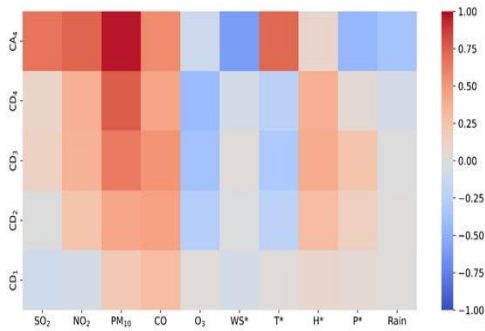
Figure 9: Correlation of PM2.5 and other relevant factors over multiple frequency bands.

Comparison Between Different Models

A complete comparison with twelve different kinds of benchmarks is performed with a focus on the predicted findings. Table 3 and Figure 10 display the forecasting comparative results from the experiences of MAPE, MAE, and RMSE, accordingly. The MAPE analysis in Table 3 shows that linear regression (LR) does a poor job of predicting PM2.5, with an average of 28% of MAPE. This information is presented because of the study. This indicates that depending entirely on LR might result in incorrect forecasts and reduce the efficiency of the forecasting system. It is vital to investigate different models that can provide forecasts that are more accurate and can improve the effectiveness of the system that is used for forecasting. DeepTCN which is the only method of predicting surpasses the other eight, and it comes in at number four in terms of expected performance. On the other hand, hybrid models such as EMD-GRU, EEMD-LSTM, and CEEMDAN-DeepTCN exceed it. These models incorporate the best aspects of a few different made by hybrid models and show a better level of resilience and stability when compared to those made by single forecasting methodologies (Jiang et al., 2021a).

Table 3: Comparing the performance of various methodologies based on criteria

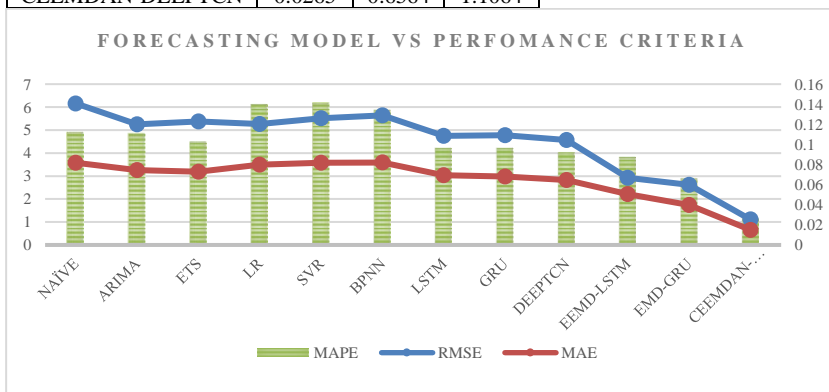| Forecasting Model	Performance Criteria		
	MAPE	MAE	RMSE
NAIVE	0.1125	3.5822	6.1606
ARIMA	0.1112	3.2614	5.2491
ETS	0.1028	3.1850	5.3791
LR	0.1402	3.4900	5.2590
SVR	0.1418	3.5814	5.5224
BPNN	0.1343	3.5870	5.6460
LSTM	0.0965	3.0390	4.7490
GRU	0.0963	2.9800	4.7750
DEEPTCN	0.9200	2.8290	4.5710
EEMD-LSTM	0.0874	2.2077	2.9203
EMD-GRU	0.0661	1.7422	2.6230
CEEMDAN-DEEPTCN	0.0265	0.6564	1.1064

Figure 10: Graph for the visual observation of various methods in terms of criteria. (a)MAPE (b)RMSE. (c)MAE.

Discussion and Recommendations

This study demonstrates how deep learning neural networks can be employed to predict quality of air using multiple modeling techniques, including data processing, correlation analysis connecting PM2.5 and other variables, and individual and hybrid models. It examines their traits, issues, and goals. Deep learning (DNN) has been shown to outperform superficial machine learning architecture. However, RNN-based models such as LSTM as well as GRU are favored due to their capability to capture time-dependent input, their ability to manage vanishing gradient problems, and their incredibly straightforward development. For air quality forecasting, stacking LSTM, bidirectional and stacked bidirectional LSTM, and stacked GRU have also been developed. CNN has gained popularity for its capacity to extract features from datasets, and autoencoders have been employed to do so. Due to their capacity to estimate air quality utilizing complex datasets and adapt to varied forecasting objectives, hybrid models that incorporate the benefits of several approaches have received attention. Decomposition techniques can enhance forecasting performance by up to 80% a hybrid models can reduce prediction error by 60-80%. Hybrid models have been employed in a variety of industries, including wind power, traffic, and temperature estimation. Hybrid sequence-to-sequence architectures outperform solo models and offer remarkable perdition abilities. The creation of advanced hybrid models, on the other hand, may enhance computational complexity while diminishing model time efficiency. Because air quality cannot be improved with a few techniques to manage complex forecasting of times series challenges, balanced model development methodologies must be considered when developing optimal hybrid models to meet diverse forecasting needs.

Future Development of a Model for Air Quality Forecasting

The application of RNN-based frameworks to the field of air quality forecasting has grown in popularity, primarily attributed to their adaptability with time series forecasting. Enhanced iterations of RNN-based models, including LSTM, GRU, and CNN, are gaining prominence in the domain of air quality forecasting. Subsequent studies should examine the effectiveness and generalizability of various methods, such as unsupervised deep learning techniques like DBM. DBN, and DRL. Exploring optimization strategies such as evolutionary algorithms, and metaheuristic algorithms, Bayesian optimization should be considered in hybrid forecasting models, but the review study reveals a scarcity of findings concerning this type of model. Additional methods for optimization, such as the Harris Hawks optimizer, Bonobo optimization algorithm, group teaching optimization algorithm, and Antlion optimizer, make it possible to investigate hybrid deep learning models.

In the realm of secondary decomposition or hybrid data decomposition methods in deep learning, the use of hybrid data decomposition techniques has become common. It is possible to boost the accuracy of forecasting models by integrating different data decomposition approaches in two-level decomposition procedures, and further studies should investigate the efficacy of hybrid models in calculating long-term periodic changes in air quality using larger input datasets. Because of their restricted space, sparse distribution, and greater operating and maintenance expenses, ground air quality monitoring can be difficult to obtain data from. For this reason, the majority of evaluated research relies on datasets from these sites. Acquiring air quality data using satellite photography, which provides wide and temporal observation at a reasonable cost, may be judged appropriate for future inquiry (Imani, 2021).

Conclusions

This study delivers a thorough analysis of many deep–learning forecasting models for the time-periodic PM 2.5 airborne particles in different metrics were analyzed to determine the effectiveness of these models. Theoretically developed modeling methodologies are systematically aligned to provide a summary of several different deep-learning techniques. This research described the combination of numerous parts that resulted in strong hybrid

models for forecasting. In addition, this work sequentially summarized the key elements of a deep learning model for forecasting, including data decomposition, correlation evaluation between PM2.5 particles and other factors, and feature extraction. This research is followed by a comparison of models for forecasting that have been constructed to highlight the features of estimating air quality using various frameworks. The analysis reveals that in comparison to standalone and machine learning approaches, hybrid deep learning technology has been effectively used to predict air containments and meteorological datasets having greater accuracy. A comprehensive review study could be undertaken to evaluate the use of hybrid applications and deep learning in various aspects of air quality prediction, including the estimation of gas emissions and their environmental impact, as well as the assessment of deep learning models from several angles, including the management of values that are missing in the dataset and model parameter optimization which may lead to the development of new theories and significant discoveries.

Acknowledgments

Declarations

Ethical approval: Not Applicable.

Competing interests: The authors declare that they have no competing interests**.**

Authors' Contributions

Lakshmi Shankar: Concepts, Investigation, and Writing – original draft**.**

Krishnamoorthy Arasu: Review and editing.

Availability of data and materials: The data used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

Arsov, M., Zdravevski, E., Lameski, P., Corizzo, R., Koteli, N., Mitreski, K., & Trajkovik, V. (2020). *Short-term air pollution forecasting based on environmental factors and deep learning models*. 15–22.

Bai, Y., Zeng, B., Li, C., & Zhang, J. (2019). An ensemble long short-term memory neural network for hourly PM2.5 concentration forecasting. *Chemosphere*, *222*, 286–294.

Bhanja, S., & Das, A. (2021). A hybrid deep learning model for air quality time series prediction. *Indonesian Journal of Electrical Engineering and Computer Science*, *22*(3), 1611.

Biancofiore, F., Busilacchio, M., Verdecchia, M., Tomassetti, B., Aruffo, E., Bianco, S., Di Tommaso, S., Colangeli, C., Rosatelli, G., & Di Carlo, P. (2017a). Recursive neural network model for analysis and forecast of PM10 and PM2.5. *Atmospheric Pollution Research*, *8*(4), 652–659.

Biancofiore, F., Busilacchio, M., Verdecchia, M., Tomassetti, B., Aruffo, E., Bianco, S., Di Tommaso, S., Colangeli, C., Rosatelli, G., & Di Carlo, P. (2017b). Recursive neural network model for analysis and forecast of PM10 and PM2.5. *Atmospheric Pollution Research*, *8*(4), 652–659.

Chang, Q., Zhang, H., & Zhao, Y. (2020). Ambient air pollution and daily hospital admissions for respiratory system–related diseases in a heavy polluted city in Northeast China. *Environmental Science and Pollution Research*, *27*(9), 10055–10064.

Díaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., & Moncada-Herrera, J. A. (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment*, *42*(35), 8331–8340.

Dong, M., Yang, D., Kuang, Y., He, D., Erdal, S., & Kenski, D. (2009). PM2.5 concentration prediction using hidden semi-Markov model-based times series data mining. *Expert Systems with Applications*, *36*(5), 9046–9055.

Du, P., Wang, J., Hao, Y., Niu, T., & Yang, W. (2020). A novel hybrid model based on multi-objective Harris hawks optimization algorithm for daily PM2.5 and PM10 forecasting. *Applied Soft Computing*, *96*, 106620.

Du, S., Li, T., Yang, Y., & Horng, S.-J. (2020). Multivariate time series forecasting via attention-based encoder–decoder framework. *Neurocomputing*, *388*, 269–279.

Du, S., Li, T., Yang, Y., & Horng, S.-J. (2021). Deep Air Quality Forecasting Using Hybrid Deep Learning Framework. *IEEE Transactions on Knowledge and Data Engineering*, *33*(6), 2412–2424.

Du, W., Chen, L., Wang, H., Shan, Z., Zhou, Z., Li, W., & Wang, Y. (2023). Deciphering urban traffic impacts on air quality by deep learning and emission inventory. *Journal of Environmental Sciences*, *124*, 745–757.

Dun, A., Yang, Y., & Lei, F. (2022). Dynamic graph convolution neural network based on spatial-temporal correlation for air quality prediction. *Ecological Informatics*, *70*, 101736.

Freeman, B. S., Taylor, G., Gharabaghi, B., & Thé, J. (2018). Forecasting air quality time series using deep learning. *Journal of the Air & Waste Management Association*, *68*(8), 866–886.

García Nieto, P. J., Sánchez Lasheras, F., García-Gonzalo, E., & de Cos Juez, F. J. (2018a). PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study. *Science of The Total Environment*, *621*, 753–761.

García Nieto, P. J., Sánchez Lasheras, F., García-Gonzalo, E., & de Cos Juez, F. J. (2018b). PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study. *Science of The Total Environment*, *621*, 753–761.

Gugnani, V., & Singh, R. K. (2022a). Analysis of deep learning approaches for air pollution prediction. *Multimedia Tools and Applications*, *81*(4), 6031–6049.

Gugnani, V., & Singh, R. K. (2022b). Analysis of deep learning approaches for air pollution prediction. *Multimedia Tools and Applications*, *81*(4), 6031–6049.

Guo, B., Wang, X., Pei, L., Su, Y., Zhang, D., & Wang, Y. (2021). Identifying the spatiotemporal dynamic of PM2.5 concentrations at multiple scales using geographically and temporally weighted regression model across China during 2015–2018. *Science of The Total Environment*, *751*, 141765.

Guo, C., Liu, G., & Chen, C.H. (2020). Air Pollution Concentration Forecast Method Based on the Deep Ensemble Neural Network. *Wireless Communications and Mobile Computing*, *2020*, 1–13.

Hable-Khandekar, V., & Srinath, P. (2017). Machine Learning Techniques for Air Quality Forecasting and Study on Real-Time Air Quality Monitoring. *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, 1–6.

Hael, M. A. (2023). Unveiling air pollution patterns in Yemen: a spatial–temporal functional data analysis. *Environmental Science and Pollution Research*, *30*(17), 50067–50095.

Huang, G., Li, X., Zhang, B., & Ren, J. (2021). PM2.5 concentration forecasting at surface monitoring sites using GRU neural network based on empirical mode decomposition. *Science of The Total Environment*, *768*, 144516.

Imani, M. (2021). Particulate matter (PM2.5 and PM10) generation map using MODIS Level-1 satellite images and deep neural network. *Journal of Environmental Management*, *281*, 111888.

Janarthanan, R., Partheeban, P., Somasundaram, K., & Navin Elamparithi, P. (2021). A deep learning approach for prediction of air quality index in a metropolitan city. *Sustainable Cities and Society*, *67*, 102720.

Jeya, S., & Sankari, L. (2020). Air Pollution Prediction by Deep Learning Model. *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 736–741.

Jiang, F., Zhang, C., Sun, S., & Sun, J. (2021a). Forecasting hourly PM$_{2.5}$ based on deep temporal convolutional neural network and decomposition method. *Applied Soft Computing*, *113*, 107988.

Jiang, F., Zhang, C., Sun, S., & Sun, J. (2021b). Forecasting hourly PM$_{2.5}$ based on deep temporal convolutional neural network and decomposition method. *Applied Soft Computing*, *113*, 107988.

Kan, H., Chen, R., & Tong, S. (2012). Ambient air pollution, climate change, and population health in China. *Environment International*, *42*, 10–19.

Kow, P.-Y., Wang, Y.-S., Zhou, Y., Kao, I.-F., Issermann, M., Chang, L.-C., & Chang, F.-J. (2020). Seamless integration of convolutional and back-propagation neural networks for regional multi-step-ahead PM2.5 forecasting. *Journal of Cleaner Production*, *261*, 121285.

Kumar SV et al (2. (n.d.). *No Title Data analysis for predicting air pollutant concentration in Smart city Uppsala. Proc 2nd ACM SIGSPATIAL Int. Work. Geo Streaming.*

Le, V.-D., Bui, T.-C., & Cha, S.-K. (2020). Spatiotemporal Deep Learning Model for Citywide Air Pollution Interpolation and Prediction. *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 55–62.

Lee, P., Saylor, R., & McQueen, J. (2018). Air Quality Monitoring and Forecasting. *Atmosphere*, *9*(3), 89.

Li, L., Girguis, M., Lurmann, F., Pavlovic, N., McClure, C., Franklin, M., Wu, J., Oman, L. D., Breton, C., Gilliland, F., & Habre, R. (2020). Ensemble-based deep learning for estimating PM2.5 over California with multisource big data including wildfire smoke. *Environment International*, *145*, 106143.

Li, S., Wang, G., Wang, B., Cao, S., Zhang, K., Duan, X., & Wu, W. (2022). Has the Risk of Outpatient Visits for Allergic Rhinitis, Related to Short-Term Exposure to Air Pollution, Changed over the Past Years in Beijing, China? *International Journal of Environmental Research and Public Health*, *19*(19), 12529.

Li, X., Chen, H., & Li, Y. (2020). The effect of air pollution on children's migration with parents: evidence from China. *Environmental Science and Pollution Research*, *27*(11), 12499–12513.

Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., & Chi, T. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environmental Pollution*, *231*, 997–1004.

Liao, Q., Zhu, M., Wu, L., Pan, X., Tang, X., & Wang, Z. (2020). Deep Learning for Air Quality Forecasts: a Review. *Current Pollution Reports*, *6*(4), 399–409.

Ma, J., Ding, Y., Cheng, J. C. P., Jiang, F., Gan, V. J. L., & Xu, Z. (2020). A Lag-FLSTM deep learning network based on Bayesian Optimization for multi-sequential-variant PM2.5 prediction. *Sustainable Cities and Society*, *60*, 102237.

Ma, J., Ding, Y., Cheng, J. C. P., Jiang, F., & Wan, Z. (2019a). A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM2.5. *Journal of Cleaner Production*, *237*, 117729.

Ma, J., Ding, Y., Cheng, J. C. P., Jiang, F., & Wan, Z. (2019b). A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM2.5. *Journal of Cleaner Production*, *237*, 117729.

Ma, J., Ding, Y., Cheng, J. C. P., Jiang, F., & Wan, Z. (2019c). A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM2.5. *Journal of Cleaner Production*, *237*, 117729.

Mehmood, K., Bao, Y., Saifullah, Cheng, W., Khan, M. A., Siddique, N., Abrar, M. M., Soban, A., Fahad, S., & Naidu, R. (2022). Predicting the quality of air with machine learning approaches: Current research priorities and future perspectives. *Journal of Cleaner Production*, *379*, 134656.

Moisan, S., Herrera, R., & Clements, A. (2018). A dynamic multiple equation approach for forecasting PM 2.5 pollution in Santiago, Chile. *International Journal of Forecasting*, *34*(4), 566–581.

Ong, B. T., Sugiura, K., & Zettsu, K. (2016). Dynamically pre-trained deep recurrent neural networks using environmental monitoring data for predicting PM2.5. *Neural Computing and Applications*, *27*(6), 1553–1566.

Pak, U., Ma, J., Ryu, U., Ryom, K., Juhyok, U., Pak, K., & Pak, C. (2020a). Deep learning-based PM2.5 prediction considering the spatiotemporal correlations: A case study of Beijing, China. *Science of The Total Environment*, *699*, 133561.

Pak, U., Ma, J., Ryu, U., Ryom, K., Juhyok, U., Pak, K., & Pak, C. (2020b). Deep learning-based PM2.5 prediction considering the spatiotemporal correlations: A case study of Beijing, China. *Science of The Total Environment*, *699*, 133561.

Pak, U., Ma, J., Ryu, U., Ryom, K., Juhyok, U., Pak, K., & Pak, C. (2020c). Deep learning-based PM2.5 prediction considering the spatiotemporal correlations: A case study of Beijing, China. *Science of The Total Environment*, *699*, 133561.

Qi, Y., Li, Q., Karimian, H., & Liu, D. (2019). A hybrid model for spatiotemporal forecasting of PM2.5 based on graph convolutional neural network and long short-term memory. *Science of The Total Environment*, *664*, 1–10.

Ravindran, U., & Gunavathi, C. (2023). A survey on gene expression data analysis using deep learning methods for cancer diagnosis. *Progress in Biophysics and Molecular Biology*, *177*, 1–13.

Samal, K. K. R., Babu, K. S., & Das, S. K. (2021). Multi-directional temporal convolutional artificial neural network for PM2.5 forecasting with missing values: A deep learning approach. *Urban Climate*, *36*, 100800.

Samal, K. K. R., Babu, K. S., Das, S. K., & Acharaya, A. (2019). Time Series based Air Pollution Forecasting using SARIMA and Prophet Model. *Proceedings of the 2019 International Conference on Information Technology and Computer Communications*, 80–85.

Sánchez-Balseca, J., & Pérez-Foguet, A. (2020). Spatio-temporal air pollution modelling using a compositional approach. *Heliyon*, *6*(9), e04794.

Shang, Z., Deng, T., He, J., & Duan, X. (2019a). A novel model for hourly PM2.5 concentration prediction based on CART and EELM. *Science of The Total Environment*, *651*, 3043–3052.

Shang, Z., Deng, T., He, J., & Duan, X. (2019b). A novel model for hourly PM2.5 concentration prediction based on CART and EELM. *Science of The Total Environment*, *651*, 3043–3052.

Sharma, E., Deo, R. C., Prasad, R., Parisi, A. V., & Raj, N. (2020). Deep Air Quality Forecasts: Suspended Particulate Matter Modeling With Convolutional Neural and Long Short-Term Memory Networks. *IEEE Access*, *8*, 209503–209516.

Tariq, S., Loy-Benitez, J., Nam, K., Lee, G., Kim, M., Park, D., & Yoo, C. (2021a). Transfer learning driven sequential forecasting and ventilation control of PM2.5 associated health risk levels in underground public facilities. *Journal of Hazardous Materials*, *406*, 124753.

Tariq, S., Loy-Benitez, J., Nam, K., Lee, G., Kim, M., Park, D., & Yoo, C. (2021b). Transfer learning driven sequential forecasting and ventilation control of PM2.5 associated health risk levels in underground public facilities. *Journal of Hazardous Materials*, *406*, 124753.

Vargas, J. E., Kubesch, N., Hernandéz-Ferrer, C., Carrasco-Turigas, G., Bustamante, M., Nieuwenhuijsen, M., & González, J. R. (2018). A systemic approach to identify signaling pathways activated during short-term exposure to traffic-related urban air pollution from human blood. *Environmental Science and Pollution Research*, *25*(29), 29572–29583.

Vlachogianni, A., Kassomenos, P., Karppinen, A., Karakitsios, S., & Kukkonen, J. (2011). Evaluation of a multiple regression model for the forecasting of the concentrations of NOx and PM10 in Athens and Helsinki. *Science of The Total Environment*, *409*(8), 1559–1571.

Www.Devic-earth.com. (n.d.). No Title. *Https: //Www.Devic-Earth.Com/Hubfs/Effects%20of%20 Air%20pollution%20on%20human%20body%20infograph-Png.Png*.

Xayasouk, T., Lee, H., & Lee, G. (2020). Air Pollution Prediction Using Long Short-Term Memory (LSTM) and Deep Autoencoder (DAE) Models. *Sustainability*, *12*(6), 2570.

Xing, Y.-F., Xu, Y.-H., Shi, M.-H., & Lian, Y.-X. (2016). The impact of PM2.5 on the human respiratory system. *Journal of Thoracic Disease*, *8*(1), E69-74.

Xu, R., Wang, D., Li, J., Wan, H., Shen, S., & Guo, X. (2023). A Hybrid Deep Learning Model for Air Quality Prediction Based on the Time–Frequency Domain Relationship. *Atmosphere*, *14*(2), 405.

Xu, X., Tong, T., Zhang, W., & Meng, L. (2020). Fine-grained prediction of PM2.5 concentration based on multisource data and deep learning. *Atmospheric Pollution Research*, *11*(10), 1728–1737.

Yang, Y., Mei, G., & Izzo, S. (2022). Revealing Influence of Meteorological Conditions on Air Quality Prediction Using Explainable Deep Learning. *IEEE Access*, *10*, 50755–50773.

Yi, X., Zhang, J., Wang, Z., Li, T., & Zheng, Y. (2018). Deep Distributed Fusion Network for Air Quality Prediction. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 965–973.

Zaini, N., Ean, L. W., Ahmed, A. N., & Malek, M. A. (2022). A systematic literature review of deep learning neural network for time series air quality forecasting. *Environmental Science and Pollution Research*, *29*(4), 4958–4990.

Zeng, Y., Chen, J., Jin, N., Jin, X., & Du, Y. (2022). Air quality forecasting with hybrid LSTM and extended stationary wavelet transform. *Building and Environment*, *213*, 108822.

Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., & Liu, J. (2021a). Air quality predictions with a semi-supervised bidirectional LSTM neural network. *Atmospheric Pollution Research*, *12*(1), 328–339.

Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., & Liu, J. (2021b). Air quality predictions with a semi-supervised bidirectional LSTM neural network. *Atmospheric Pollution Research*, *12*(1), 328–339.

Zhou, H., Zhang, F., Du, Z., & Liu, R. (2021a). Forecasting PM2.5 using hybrid graph convolution-based model considering dynamic wind-field to offer the benefit of spatial interpretability. *Environmental Pollution*, *273*, 116473.

Zhou, H., Zhang, F., Du, Z., & Liu, R. (2021b). Forecasting PM2.5 using hybrid graph convolution-based model considering dynamic wind-field to offer the benefit of spatial interpretability. *Environmental Pollution*, *273*, 116473.

Zhou, Q., Jiang, H., Wang, J., & Zhou, J. (2014). A hybrid model for PM 2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Science of The Total Environment*, *496*, 264–274.