BMC Research Notes

RESEARCH NOTE

Open Access

Ecosense: a revolution in urban air quality forecasting for smart cities

Kalyan Chatterjee[1], Machakanti Navya Thara[2], Mandadi Sriya Reddy[3], N. Selvamuthukumaran[2], M. Priyadharshini[4], Tummala Abhinav Vardhan Reddy[2], Somenath Chakraborty[5], Saurav Mallik[6,7*], Mohd Asif Shah[8,10,11,12*] and Aimin Li[9]

Abstract

The Smart City (SC) framework is popular due to its advancement in enhancing lives and public safety. However, these advancements lead to many challenges due to the dependency of Internet of Things (IoT) devices in terms of electronic waste and resource consumption. To address those challenges, the integration of a weather-smart grid (WSG) with SC becomes crucial to safeguard the environment and residents' well-being. Along with these concepts, this study proposes a novel approach, *EcoSense: A Revolution in Urban Air Quality Forecasting for Smart Cities*, which incorporates Bi-directional Stacked LSTM with a Weather-Smart Grid (*BlaSt*). *BlaSt* innovatively integrates several key components: (i) the model captures intricate temporal dependencies and trends in air quality data by incorporating historical air pollutant and meteorological data. (ii) integration of the WSG component enhances the model's capability to incorporate weather data, which is critical for accurate air quality forecasting. (iii) the model computes 12-hour predictions by designing 1-hour prediction models, enabling it to provide timely forecasts with high precision. *BlaSt* demonstrates significant improvements over existing models, with enhancements of 36%, 26%, 21%, 46%, 14%, 10%, and 6% in accuracy compared to SVR, MLP, RAQP, Vlachogianni, LSTM, BLSTM, and SLSTM models, respectively. It achieves a mean absolute error (MAE) of 0.10 and a mean squared error (MSE) of 0.08. Additionally, *BlaSt* reduces computational complexity by 25%, making it more efficient in processing large-scale air quality data. The experimental results demonstrate *BlaSt*'s superior accuracy and efficiency, showcasing its potential to advance urban air quality forecasting in SCs.

Keywords Air quality, Air Pollutant Concentrations (APCs), Internet of Things (IoT), Meteorological Factors (MFs), Smart City (SC), Weather Smart Grid (WSG)

*Correspondence:
Saurav Mallik
sauravmtech2@gmail.com; smallik@hsph.harvard.edu; smallik@arizona.edu
Mohd Asif Shah
m.asif@kardan.edu.af
Full list of author information is available at the end of the article

Introduction

Smart city (SC) frameworks that integrate Internet of Things (IoT) devices, like sensors, actuators, etc., have improved public safety through data-driven decision-making and have globally revolutionized urban living [1–7]. Seen as smarter, more sustainable, and livable, SCs are pivotal in addressing urban challenges, with investments projected to reach $135 billion by 2020 [8]. Integrating modern technologies and weather smart grids (WSGs) inside SCs improves weather forecasting accuracy, environmental understanding, disaster preparedness, and urban functionality [9]. To mitigate IoT devices' ecological impact, integrating weather research and developing WSGs promise to optimize operations, reduce energy consumption, enhance urban resilience, and promote sustainability. Additionally, efforts are underway to predict and mitigate air pollution's health effects by studying interactions with meteorological factors (MFs).

This study is driven by the urgent need for reliable and accurate air quality prediction models that can keep up with the needs of quickly urbanizing areas. Cities face greater hurdles in monitoring and managing air pollution, necessitating creative solutions that maximize computational resources in addition to providing accurate forecasts.

This study addresses these challenges by integrating advanced technologies such as Bi-directional Stacked LSTM and the Weather-Smart Grid. The proposed model is designed to enhance prediction accuracy for urban areas while minimizing the computational burden, making it suitable for real-time applications in SC infrastructures. This approach is particularly motivated by the need to manage air quality in a sustainable manner, leveraging smart grids and ML to provide actionable insights for urban planners and policymakers.

Ultimately, this study aims to revolutionize how cities approach air quality management, promoting healthier living environments and contributing to the global effort to combat pollution and its adverse effects.

Promotion of sustainability and air quality openness

Promoting transparency in air quality is just one aspect of environmental stewardship that must be heavily prioritized in the creation of sustainable SCs. In this process, public involvement is essential because it gives citizens the ability to take part in monitoring activities and cultivates a sense of shared responsibility. The study by Mak & Lam [10] highlights how public involvement in environmental monitoring can drive positive outcomes in urban sustainability efforts, making it a crucial component of SC initiatives.

Climate mitigation on regional air quality

Climate change is intricately linked to air quality, and addressing this relationship is vital for SCs. Regional climate mitigation strategies can significantly influence air quality, as seen in the work by Huang et al. [11], which demonstrates the impact of climate interventions on reducing pollution levels. The study [12] demonstrates improved air quality prediction to support green SCs by developing a robust multivariate time series forecasting model. This model adeptly captures spatial and temporal dependencies in AQI and MFs, highlighting the effectiveness of advanced SLSTM architectures for precise environmental monitoring and management. Also, Neo et al. [13] provides insights into how public health and air quality can be improved through sustainable practices in SCs, offering a comprehensive approach to urban environmental management. Incorporating such strategies into the SC framework can lead to more resilient and healthier urban environments.

Despite the advancements, the environmental impact of IoT devices and the need for effective air quality management remain challenges. The integration of weather research and the development of WSGs are essential for optimizing operations and reducing energy consumption.

This study incorporates the *BlaSt* model. The *BlaSt* model features:

1. **Bi-directional Stacked LSTM:** Enhances temporal dependency capture and incorporates a weather-smart grid (WSG).
2. **Temporal Aggregation and Lagged Features:** Improves predictive accuracy by approximately 26% by capturing temporal patterns at multiple scales.
3. **Dynamic Feature Selection:** Adapts to changing conditions by selecting the most required predictive features. The model achieves good prediction performance while adjusting to dynamic settings by dynamically updating the feature set and evaluating feature relevance at various time instants up to 12 h.
4. ***1-hour Prediction Models:*** Tailored for hourly-based predictions.
5. **Spatiotemporal Correlation Coefficient:** Calculates internal spatiotemporal correlations among participating air pollutant concentrations (APCs) and meteorological factors (MFs) using the designed ρ metric.

Literature survey

Future air quality prediction is essential for SCs, where accurate predictions can guide public health and urban management. Gao et al. [14] discussed the advancements in the SC frameworks, mainly focusing

on integrating advanced sensors and IoT devices for improved urban management. This paper highlights how these technologies enhance data collection and processing for better environmental monitoring and decision-making.

Weng et al. [15] explored the application of ML techniques for SCs air quality prediction. The focus is on leveraging sophisticated algorithms to examine data from multiple sources in order to enhance the precision of predictions and facilitate efficient control of air quality. An innovative method for incorporating meteorological data into air quality forecasting models was demonstrated in this study.

Gao et al. [16] focused on how combining weather data with pollutant concentrations can enhance the accuracy of predictions and provide more reliable insights for urban planning.

Gao et al. [17] reviewed recent developments in transportation systems within SCs, mainly how data-driven approaches are utilized to optimize traffic flow and reduce pollution. This paper highlights the role of ML in improving the efficiency of transportation networks and their impact on air quality.

Bi et al. [18] addressed the challenges of integrating public engagement in SC initiatives. It discusses strategies for involving citizens in environmental monitoring and decision-making processes, emphasizing the importance of community participation in enhancing air quality management.

Qi et al. [19] examined the innovative methods for improving air quality monitoring through big data and AI. It focuses on how these technologies can be leveraged to enhance the accuracy of air quality measurements and provide actionable insights for urban sustainability.

Song & Stettler [20] introduced a Multi-AP learning network that estimates multiple air pollutants ($PM_{2.5}$, PM_{10}, O_3) at a high spatial resolution across a city using data from limited monitoring sites and urban features like land use and weather. Applied in Chengdu, the model is more accurate and computationally efficient (reducing time by two-thirds) than traditional methods. Meteorological data emerged as the most critical feature in the model's predictions.

Song et al. [21] proposed Deep-MAPS, an ML-based framework for mobile air pollution sensing, designed to monitor urban air quality with high spatial-temporal resolution. This framework, when implemented in Beijing, estimates $PM_{2.5}$ concentrations with a $1km^2$ and 1-hour resolution, attaining under 15% SMAPE. It does this by combining mobile and fixed AQ sensors. Deep-MAPS also uses urban big data to identify potential causes of pollution, supporting sustainable urban management. Additionally, the framework significantly reduces costs, saving up to 90% in hardware investment compared to traditional fixed-sensor approaches.

Song et al. [22] proposed the MCST-Tree, a multi-cascade space-time learning model based on trees that is intended for high-resolution inference of air quality in metropolitan settings. To estimate pollutant concentrations at the grid level, this model takes into account data from both stationary and mobile sensors as well as a range of urban characteristics, including traffic, population, land use, and weather. In a case study of Chengdu, the model effectively reconstructed $PM_{2.5}$ distribution maps with sparse data coverage, achieving high accuracy (SMAPE = 14.13%; $R^2 = 0.94$). This paper highlights the importance of mobile sampling, showing that increased mobile data significantly enhances model performance, demonstrating the model's potential for detailed, accurate air pollution mapping.

Chai et al. [23] focused on predicting bike-sharing demand at a grid level in SCs to improve transportation efficiency. For more precise predictions, the authors employ a deep multi-view spatial-temporal network in conjunction with a DL methodology in place of conventional time series approaches. This network incorporates both spatial and temporal data. The model performs better than current machine learning models and runs at a grid resolution of $1\ km \times 1\ km$, according to thorough testing on Beijing bike-sharing data. 6 G is projected to further improve these capabilities. Real-time control and high-frequency monitoring of bike-sharing patterns are made possible by the integration of 4 G/5 G/6 G communication technologies.

Zhang et al. [24] examined the difficulty of monitoring methane (CH_4), one of the most potent greenhouse gases, in China, the world's greatest emitter. In this work, the distribution of CH_4 across East Asia in 2017 was simulated using the Weather Research and Forecasting (WRF) model in conjunction with satellite data and a greenhouse gas module. Ground-based and satellite-based data are used to verify the correctness of the model. In order to monitor CH_4, the researchers also compared four sensor placement algorithms. They discovered that the QR pivot algorithm is the most efficient, especially in areas with high concentrations of CH_4. Gaps in the current monitoring efforts are highlighted in the research. It implies that the QR pivot algorithm, when used to strategically place 160 sensors, could increase monitoring efficiency and accuracy. These findings could be helpful for future CH_4 monitoring site planning in China.

Zaini et al. [25] review DL approaches, highlighting models like Convolutional Neural Networks (CNNs) and RNNs for their ability to capture complex patterns in environmental data. Huang et al. [26] introduced a Spatiotemporal Attention-Based RNN Network, which

improves accuracy by focusing on relevant time steps and spatial locations. Cai et al. [27] emphasized the fusion of diverse data sources to enhance predictions, aligning with our model's approach.

Advantages and disadvantages of existing methods

ANNs are effective for classifying pollutant concentrations [28, 29], providing valuable insights into current air quality levels. However, they often lack temporal modeling capabilities, which are crucial for forecasting. Time-series data forecasting is a good fit for recurrent neural networks (RNNs) [30]. However, RNNs may have trouble identifying long-term dependencies, which can reduce their predicted accuracy over long time horizons. The Multi-AP Learning Network, introduced by Song & Stettler [20], offers high accuracy and computational efficiency for predicting multiple pollutants but is limited by its dependence on existing monitoring sites, which may not provide comprehensive coverage. The Deep-MAPS Framework [21] excels in high spatial-temporal resolution and cost savings for air quality monitoring but faces constraints related to the spatial distribution of sensors, which may affect its coverage. The MCST-Tree Model [22] delivers detailed pollution mapping with high accuracy but relies heavily on the availability of mobile data, which may not always be sufficient. The strategy used by Chai et al. [23], which predicts bike-sharing demand using a deep multi-view spatial-temporal network, works well in transportation contexts but does not specifically address air quality forecasts. Finally, the methane monitoring models developed by Zhang et al. [24] focus on greenhouse gas emissions with advanced simulation and sensor placement techniques but are specific to methane rather than providing a broader air quality assessment.

Proposed BlaSt model's novel contributions

The *BlaSt* model distinguishes itself by integrating bidirectional stacked LSTM layers with a WSG. This combination overcomes the drawbacks of conventional techniques like Support support vector regression (SVR) [31], multilayer perceptron (MLP) [32], etc. by improving temporal dependency capturing and utilizing the meteorological data. The *BlaSt* model demonstrates superior performance in handling complex, multimodal datasets, offering enhanced accuracy and adaptive decision-making capabilities essential for SCs.

Research perspective

We started our research with the perspective that integrating advanced DL techniques and weather data could significantly improve urban air quality forecasting. This approach aims to bridge gaps left by existing methods, providing a more comprehensive and adaptable solution to the challenges of air quality prediction in SCs.

Model architecture & problem formulation

Modlel architecture of BlaSt

The goal of the *BlaSt* model architecture is to better manage SC resources and respond to weather-related difficulties by combining the WSG with the conventional SC framework. The architecture of *BlaSt* depicted in Fig. 1, represents a comprehensive approach to urban air quality forecasting.

The decision to use 12 *1-hour Prediction Model*s within the *BlaSt* model was reached through extensive empirical experimentation described in Table 1. This configuration was found to provide the best balance between model complexity and performance, effectively capturing temporal dependencies and patterns in the air quality data without leading to overfitting. Furthermore, the use of 12 units aligns with the 12-hour prediction window, ensuring the model effectively captures the temporal dynamics necessary for accurate air quality forecasts.

Regarding the number of *1-hour Prediction Model*s, as detailed in Table 1, it is evident that the accuracy and consistency of the *BlaSt* model improve when the number of *1-hour Prediction Model*s increases. However, the benefits plateau after reaching 12 units, beyond which additional *1-hour Prediction Model*s yield diminishing returns while increasing computational complexity. Thus, choosing 12 units represents an optimal balance, maximizing accuracy while maintaining manageable resource demands.

Therefore, the *BlaSt* model's system architecture is carefully designed to meet the challenges of urban air quality forecasting, integrating advanced methods and strategic choices to optimize both performance and resource efficiency. This detailed approach underpins the effectiveness of the *BlaSt* model within the broader SC framework.

Problem formulation

This study has the following objectives:

1. To maximize prediction accuracy by integrating weather data into the *BlaSt* model for enhanced air quality predictions in SCs.
2. Minimize prediction uncertainty by reducing the prediction interval width.
3. Minimize computational cost by decreasing both training and inference times, and minimize energy costs in WSG integration by optimizing energy consumption predictions over time t=12 h.

Therefore, the objectives are stated as follows:

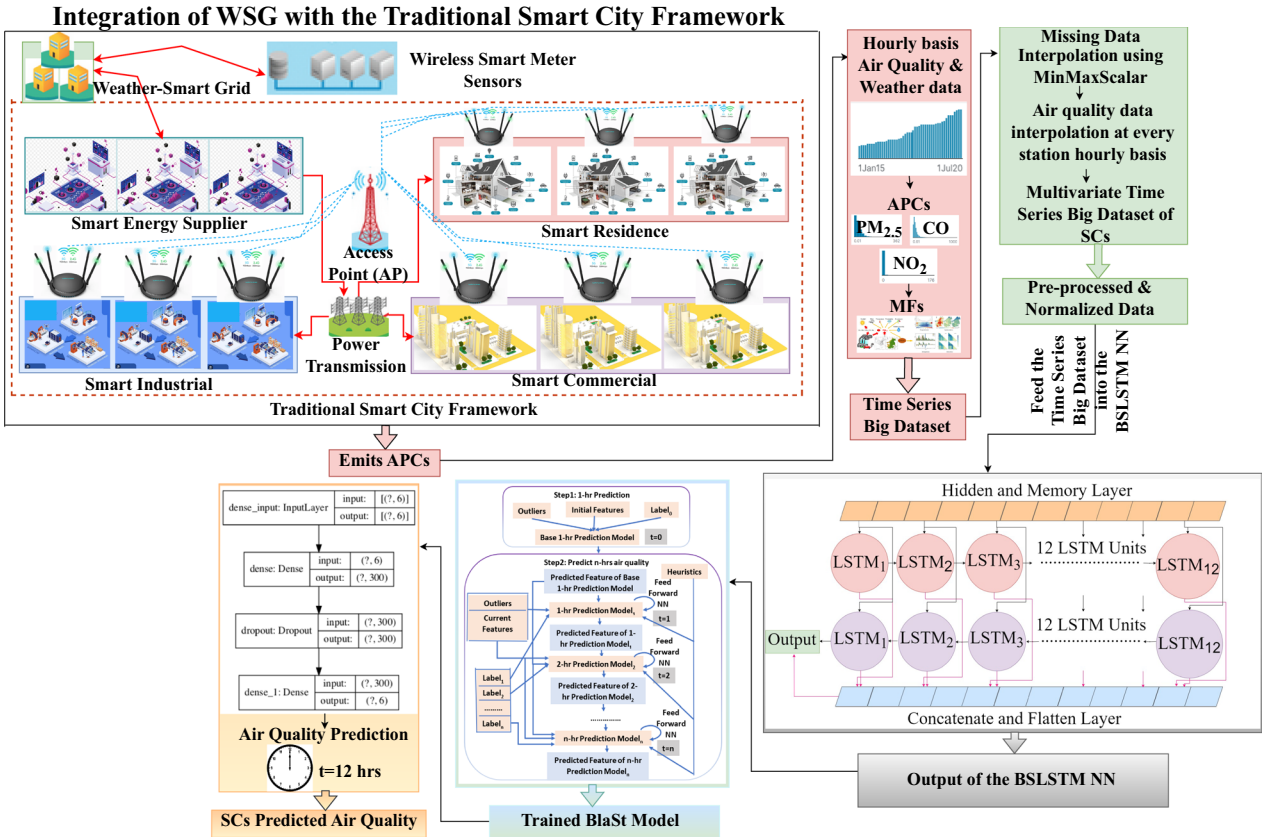


Fig. 1 Model Architecture of *BlaSt*. Here, *MinMaxScaler* scales features to a 0–1 range, aiding LSTM neural networks by ensuring consistent input feature scaling for better convergence and performance. *linear* interpolation is used to smoothly estimate missing data, preserving time series integrity by considering trends in adjacent points. Also, the choice of 12 LSTM units in the model is based on a balance between complexity and performance. This specific number was determined through empirical experimentation, where we found that 12 units provided sufficient capacity to capture the temporal dependencies and patterns in the air quality data without leading to overfitting. Additionally, 12 units aligns with the 12 time step resolution we are working with, ensuring that the model effectively captures the necessary temporal dynamics for accurate predictions

1. Maximum Air Quality Forecasting Accuracy (Max_{acc}):

$$Max_{acc} = \frac{\text{Number of Correct Forecasts}}{\text{Total Number of Forecasts}} \quad (1)$$

Here, the *BlaSt* model is enhanced with weather data for enhancing the SCs air quality forecasting accuracy.

2. Minimal Forecasting Uncertainty (Min_{un}):

$$Min_{un} = \text{Min}\big(\text{Forecasting Interval Width}\big) \quad (2)$$

3. Minimize Computational Cost (Min_{CC}):

$$Min_{CC} = \text{Min Training Time} + \text{Min Inference Time} \quad (3)$$

4. Minimal Energy Cost with WSG Integration (Min_{EC}):

$$Min_{EC} = \text{Min}\left(\sum_{t=1}^{T} \text{Cost}_t \cdot \text{Consumption}_t\right) \quad (4)$$

Here, Cost_t denotes the energy cost at time t, while Consumption_t denotes the energy consumption at time t that the model predicts.

Also, in this study, low computational cost implies the efficient use of resources by minimizing both the training time and inference time of the proposed *BlaSt* model. This indicates that the *BlaSt* model is better suited for real-time applications in SCs where resource efficiency is crucial since it is made to produce accurate predictions while consuming less time and computing power.

Table 1 Required *1-hr Prediction Models* for the Development *BlaSt*

Units of *1-hr Prediction Models*	Accuracy (%)	Consistency (%)
1	40	30
2	55	45
3	68	58
4	80	75
5	92	88
6	93.5	90
7	94	91
8	95	92
9	95.95	92.36
10	96.5	93
11	96.68	93
12	**98.59**	**94.28**
13	99	90
14	99.23	90
15	99.23	90

Bold values indicate the efficiency as well as the superior performance of our proposed model

Dataset

Formation of the required dataset

We utilized the CityPlus SC Dataset [33] to assess pollution, weather, etc. using data from 2015 to 2020. These datasets, which focus on weather and pollution, provide hourly and daily APC and MF measures from sensors throughout various Indian cities. Each dataset is around 4 GB in size.

Analysis of spatiotemporal-correlation

The proposed *BlaSt* model considers the intrinsic spatiotemporal linkages among the included APCs and their interactions with MFs to produce an extremely accurate and precise air quality forecasting model. Identifying these relationships is a crucial first step before training since it directs the model in choosing the best set of inputs. Integrating APCs with pertinent MFs and eliminating non-essential MFs lowers the computational load. The spatiotemporal relationships between each APC and the target MFs are then calculated.

$$\rho = \frac{s\sum_{i=1}^{s}(K_i * B_i) - (\sum_{i=1}^{s} K_i + \sum_{i=1}^{s} B_i)}{\sqrt{s\sum_{i=1}^{s} K_i^2 - (\sum_{i=1}^{s} K_i)^2}\sqrt{s\sum_{i=1}^{s} B_i^2 - (\sum_{i=1}^{s} B_i)^2}} \quad (5)$$

Here, data samples are denoted by s, the internal spatiotemporal correlation coefficient is represented by ρ, and the time-series sequential historical input vectors K_i and B_i are employed in the model to find patterns and trends for multivariate AQI forecasting. These vectors are derived from the historical data, capturing relevant features that contribute to predicting future AQI values. Specifically, K_i, and B_i are obtained through the preprocessing of historical data, representing the key components in determining the internal spatiotemporal correlation coefficient ρ. The symbol $*$ represents the convolution operator.

BlaSt model

BlaSt is structured using a III-phase architecture, as illustrated in Fig. 1. Phase-I describes the integration of the WSG with the conventional SC framework. The BSLSTM NN construction is done during phase-II. Phase-III is responsible for designing the *1-hour prediction models*, and finalizing the training of *BlaSt* for producing the forecasted results.

Traditional SC framework

The SC framework is a ground-breaking method of urban management that uses data-driven insights and state-of-the-art technologies to enhance the quality of life for citizens. Fundamentally, the SC framework integrates data from multiple sources, including energy grids, public services, transportation networks, and environmental sensors, to build more sustainable, efficient, and livable urban settings. These data streams give municipal planners and administrators up-to-date information on trash management, energy consumption, public safety, traffic patterns, and air quality, allowing them to make well-informed decisions.

The framework processes a wide range of data types, including air pollution levels, weather conditions, traffic congestion, energy usage, and even social media activity. This information is collected from IoT devices, satellites, and other digital infrastructures embedded within the urban fabric. By analyzing these diverse datasets, the SC framework can predict trends, optimize resource allocation, and respond to emerging challenges in real-time.

In urban management, the SC framework plays a crucial role by facilitating smart governance, improving public services, and fostering sustainable development. It enables cities to monitor environmental conditions, manage resources efficiently, reduce carbon footprints, and enhance public safety. Additionally, the SC framework facilitates the development of more flexible and resilient urban environments that can address both short-term demands and long-term strategic objectives. By taking a comprehensive approach, SCs can better handle the challenges of contemporary urban life and guarantee a greater standard of living for all citizens.

WSG integration with SC framework

Algorithm 1 Integration Process of the WSG.

Require: Weather data, SC framework settings, Grid status
Ensure: Integrated system, optimized grid management
Initialize the SC framework with the specified settings
while Grid operations are ongoing **do**
  Retrieve weather data
  Merge weather data with the SC framework to enhance grid management
  Update control strategies for grid components
  Assess grid performance and stability
  if Anomalies are identified **then**
    Execute corrective measures
  end if
  Revise SC framework settings based on the integrated data
end while
Complete the integration of the system

Integrating WSG with the SC architecture takes lots of steps and careful considerations. The WSG was created to increase the SC's ability to use MFs more efficiently, allowing for more informed choices to be made about energy management, preparedness for emergencies, and thorough urban planning. Algorithm 1 provides specifics on integrating the WSG into the conventional SC framework. Also, this integration is done through a systematic procedure intended to improve the city's capacities for energy management, emergency preparedness, and urban planning. The way the integration algorithm functions is as follows:

1. **Initialization:** Start by setting up the SC framework with the necessary parameters and configurations tailored for integration with the WSG.
2. **Data Integration:** Integrate the weather data into the SC framework. This integration helps optimize grid operations by adjusting parameters and control actions based on the latest weather conditions.
3. **Control Actions Update:** Update the control actions for grid devices, such as generators and transformers, to reflect the integrated weather data and optimize performance.
4. **Performance Monitoring:** Regularly monitor the grid's performance and stability to ensure that the integration is functioning as expected.
5. **Parameter Adjustment:** Update the SC framework parameters as needed based on the integrated data to continuously refine and improve grid operations.
6. **Finalization:** Once the integration is complete and the system is optimized, finalize the integrated system to maintain operational efficiency.

Therefore, Algorithm 1 ensures that the WSG is effectively integrated into the SC framework, allowing for improved decision-making and operational efficiency in managing urban infrastructure and resources.

Bi-directional stacking mechanism of the traditional LSTM neural network (BSLSTM NN)

The BLSTM NN, crafted to predict future air quality in SCs, integrates bidirectional processing with multiple stacked LSTM layers to improve prediction accuracy. Processing MF and historical APC data in both backward (from the future to the past) and forward (from the past to the future) directions work in SCs to capture intricate temporal relationships and patterns of air quality. By Equations (6) and (7), Backward LSTM processes data from the future to the past, capturing context for the future, and Forward LSTM processes data from the past to the future, capturing temporal dependencies.

$$\overrightarrow{h_t} = \text{LSTM}(x_t, \overrightarrow{h_{t-1}}) \tag{6}$$

$$\overleftarrow{h_t} = \text{LSTM}(x_t, \overleftarrow{h_{t+1}}) \tag{7}$$

Here, x denotes the input at time t, $\overleftarrow{h_t}$ and $\overrightarrow{h_t}$ and the hidden states in the forward (from past to future) and backward (from future to past) directions, respectively. For layer l, the hidden states are denoted by $h_t^{(l)}$. The

next LSTM layers then merge these bidirectional hidden states.

Equation (8) describes the predictive capacity of the *BlaSt* model, which is enhanced by the BLSTM stacking process that combines many LSTM layers that have been tuned by preceding layers. This allows the BLSTM to capture nuances and higher-level features in the data.

$$h_t^{(l+1)} = \text{BLSTM}^{(l+1)}(h_t^{(l)}) \tag{8}$$

Here, $h_t^{(l+1)}$ represents the hidden states at the next layer $l+1$ after applying the BLSTM layer.

In addition to improving prediction performance, this stacking helps the *BlaSt* model catch more subtleties, higher-level features, and methodological depth in the data.

1. **Incorporate Additional Features:** Enhance the input data by integrating more relevant features such as additional MFs and external used datasets [33]. This can provide a more comprehensive view of the factors influencing air quality.
2. **Optimize Hyperparameters:** To enhance the model's performance, the required hyperparameters configuration is listed in Table 2.
3. **Advanced Regularization Technique:** Apply dropout to prevent overfitting and ensure the model generalizes well to new data.
4. **Implement Data Augmentation:** Increase the robustness of the model by noise injection to diversify the training data.

1-hr prediction models

Each *1-hour prediction model* is developed by using

1. Hourly basis recurrent heuristic rules,
2. Inserting outliers generated through random sampling of the designed dataset.

At time t=0, the *Base 1-hour prediction model* is created using

1. Outliers,
2. Initial Features, and
3. Label_0.

Following this, the remaining 12 units of *1-hour prediction models* are constructed recursively, with each preceding *1-hour prediction model* passing its output to the subsequent subordinate *1-hour prediction model*, as described in Equations (9), (10), and (11).

$$H = h_t = \text{BSLSTM}(h_{t-1}, x_t, c_t) \tag{9}$$

$$Y = y_t = \text{Dense}(h_t) \tag{10}$$

$$r_o = \sum_{i=1}^{k} (Y_k + H_k) \tag{11}$$

Here, the hidden state h_t is the output of the BSLSTM at time t. It includes the input data x_t, the previous hidden state h_{t-1} at time $t-1$, and the information obtained from the cell state c_t. This hidden state contains the temporal features and patterns that the LSTM unit extracted and that indicate the model's understanding of the sequence up to time t. It retains the context and temporal connections that were learned from earlier time steps, which is essential for forecasting. The created BSLSTM NN, which modifies the hidden state, is represented by the function *BSLSTM*(). Simultaneously, the dense layer that generates the anticipated output y_t at time t is *Dense*(). For the purpose of making predictions, such predicting future air quality values, this output shows

Table 2 Hyperparameters Configuration

Hyperparameter	Description	Value
Number of LSTM Layers	Number of stacked LSTM layers in the model	25
Number of Units per Layer	Number of LSTM units per layer	25
Batch Size	Number of samples per gradient update	64
Learning Rate	Step size for weight updates	0.01
Dropout Rate	Fraction of input units to drop during training	0.03
Recurrent Dropout Rate	Fraction of recurrent units to drop during training	0.05
Sequence Length	Number of time steps in each input sequence	12
Optimizer	Algorithm used to optimize the loss function	Adam
Activation Function	Function used to introduce non-linearity	Softmax

how the model interprets the hidden state. r_o represents the recurrent output.

Therefore, the hidden state h_t in Equations (9) and (10) implicate the internal representation of the input data at time t, encoding the temporal context and patterns learned by the *BlaSt* model, which is then used to produce the final predictions. Also, the recurrent nature described in Equation (11) summarizes the functionality of the *1-hour prediction model* and is shown in Equations (9) and (10). These formulas ensure that the input data x_t, the previous hidden state h_{t-1}, and any relevant contextual information c_t are factored in order to update the hidden state h_t at time t. After that, the output y_t is generated using the modified hidden state.

Formulation of the 1-hour prediction model

The construction of the *1-hour prediction model* follows a formulated rule expressed in Equation (12). Additionally, error computation is governed by Equation (13).

$$x_{t+1} = \alpha x_t + \epsilon_{t+1} \tag{12}$$

$$\text{Error}_t = |x_t - x_{t+1}| \tag{13}$$

Note, (i) In Equation (13), the term Error_t represents the absolute difference between the actual value x_t at time t and the predicted value x_{t+1} at time $t+1$. Specifically, Equation (12) expresses the prediction model where, x_{t+1} is the forecasted value, α is the autoregressive coefficient (a value between 0 and 1), x_t is the actual value at time t, and ϵ_{t+1} is the prediction error at time $t+1$. (ii) Equation (13) measures the absolute error between the actual value x_t and the predicted value x_{t+1}. (iii) Thus, the prediction error at time $t+1$, denoted as ϵ_{t+1} in Equation (12), is not directly represented by $Error_t$ in Equation (13). Instead, $Error_t$ represents the magnitude of the prediction error made at time t when comparing the actual value at t to the predicted value for the subsequent time $t+1$. The term ϵ_{t+1} is part of the prediction model, and the absolute difference $Error_t$ quantifies how far off the prediction was from the actual value.

The *BlaSt* model also encompasses error calculation and the computation of total APCs emissions and their atmospheric dispersion described in Equations (14), and (15), respectively.

$$\text{Emission}_{\text{total}}(\text{Emi}_{\text{total}}) = \sum_{i=1}^{n} \text{EF}_i \times \text{AL}_i \tag{14}$$

Here, EF_i represents the emission factor for pollution source i. AL_i defines the activity level of pollution source i.

$$\text{Conc} = \frac{\text{Emi}_{\text{total}}}{2\pi\sigma_y\sigma_z} \exp\left(-\frac{(y - y_{\text{Receptor}})^2}{2\sigma_y^2} - \frac{(z - z_{\text{Receptor}})^2}{2\sigma_z^2}\right) \tag{15}$$

Here, the horizontal distance from the source is shown by y, while the vertical distance is indicated by z. The horizontal and vertical dispersion parameters are defined by σ_y and σ_z. The horizontal and vertical distances of the receptor location are denoted by y_{Receptor} and z_{Receptor}, respectively.

Feature engineering and fusion

The process of feature engineering for the BSLSTM framework include selecting pertinent characteristics from several modalities, such as meteorological variables and past air quality data. Consistent scaling across modalities is ensured by normalization and scaling (Equation (16)). Temporal aggregation (Equation (17)) captures dependencies by aggregating historical observations. Lagged features (Equation (18)) represent past conditions, while weighted fusion methods (Equation (19)) combine modalities, preserving their unique traits.

$$x' = \frac{x - \mu}{\sigma} \tag{16}$$

Here, x is the original feature, μ is the mean, and σ is the standard deviation.

$$X_{\text{agg}} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{17}$$

Here, X_{agg} is the aggregated feature and x_i are the observations.

$$x_{t-l} = X_{\text{agg}_{t-l}} \tag{18}$$

Here, x_{t-l} represents the feature value at time t that is lagged by l time steps, where $X_{\text{agg}_{t-l}}$ is the actual feature value observed at time $t-l$. Therefore, Equation (18) indicates how past values of a feature are used in modeling current or future values.

$$x_{\text{fused}} = \sum_{i=1}^{l} w_i x_{t-i} \tag{19}$$

Here, x_{t-i} represents the feature at time step $t-i$, and w_i are the weights for each time step, with $\sum_{i=1}^{l} w_i = 1$. This approach combines both temporal aggregation

and modality fusion, potentially enhancing the model's performance.

Experimental design

Experimental targets

The experiments aimed to evaluate the *BlaSt* model's performance in predicting air quality under different data scenarios. The targets were:

1. Assessing the model's flexibility with respect to different data types and amounts.
2. Evaluating predictive accuracy and efficiency with different dataset configurations.
3. Comparing performance across multiple models and configurations to demonstrate *BlaSt*'s superiority.

Data preparation

The designed dataset was divided into three distinct subsets: training, testing, and validation.

1. **Training Set (80%)**: utilized to demonstrate the *BlaSt* model.
2. **Testing Set (10%):** Evaluated the prediction accuracy of the *BlaSt* model by testing its performance on hypothetical data.
3. **Validation Set (10%):** Used in order to prevent overfitting and optimize hyperparameters during the *BlaSt* model building phase.

Experiments

Three experiments were conducted to evaluate the *BlaSt* model's adaptability under different data scenarios and its predictive capability. The experiments varied in data quantity and quality:

1. *Experiment 1*: Used a moderate dataset (i.e., approximately 300,000 data samples) with missing values to test the *BlaSt* model's effectiveness under standard conditions.
2. *Experiment 2*: Employed a large dataset (i.e., approximately 500,000 data samples) without missing values to assess the *BlaSt* model's time efficiency and performance with significant data volumes.
3. *Experiment 3*: Organized the training and testing datasets (i.e., approximately 700,000 data samples) chronologically to evaluate the *BlaSt* model's practical ability to predict future data trends.

Experiments 1 and 2 used historical data with random selection of training and prediction datasets, while Experiment 3 utilized a time series approach for large datasets.

Comparison methodology

The support vector regression (SVR) [31], multilayer perceptron (MLP) [32], recurrent air quality predictor (RAQP) [34], vlachogianni [4], long short-term memory (LSTM) [35], bi-directional long short-term memory (BLSTM) [36], and stacked long short-term memory (SLSTM) [37] were the models against which the performance of the *BlaSt* model was compared. Multiple evaluation metrics served as the basis for the comparisons.

Evaluation metrics

1. **ROC (Receiver Operating Characteristic) and PR (Precision-Recall) Curves** [38, 39]: The ROC and PR curves are adopted for statistical evaluation because they offer valuable insights into the model's performance, especially in imbalanced datasets, which are common in spatial-temporal assessments.

    (a) **ROC Curve:** The True Positive Rate (TPR) against False Positive Rate (FPR) at various thresholds is plotted on the ROC curve. It provides a sensitivity versus specificity score, which is helpful in assessing how well the model separates into classes. Higher values indicate stronger classification abilities. The Area Under the ROC Curve (AUC) is a crucial measure of overall model performance.

    (b) **PR Curve:** When utilizing imbalanced data where one class is significantly more abundant than the other-the PR curve, which graphs Precision against Recall, is especially helpful. It focuses on how well the model predicts the minority class, which is frequently of more importance. Determining the optimal threshold for classification is aided by the PR curve, which illustrates the trade-off between precision and recall.

    Therefore, a more thorough assessment of the model's performance is obtained by utilizing both ROC and PR curves, which enables

Table 3 Rationale for Selecting 30 Iterations in Model Evaluation

Number of iterations	Computational cost	Statistical reliability
20	Lower	Moderate
30	Balanced	High
40	Higher	Slightly Higher
50	Significantly Higher	Marginally Higher

improved decision-making in spatial-temporal assessments where maintaining a balance between false positives and false negatives is crucial.

2. **Loss Computation**

   (a) **Mean Squared Error (MSE):** Measures the mean squared difference between the expected and actual values. MSE values below.001 suggest higher accuracy.

   (b) **Mean Absolute Error (MAE):** Calculates the mean size of the forecast errors. Higher forecast accuracy is reflected in a lower MAE.

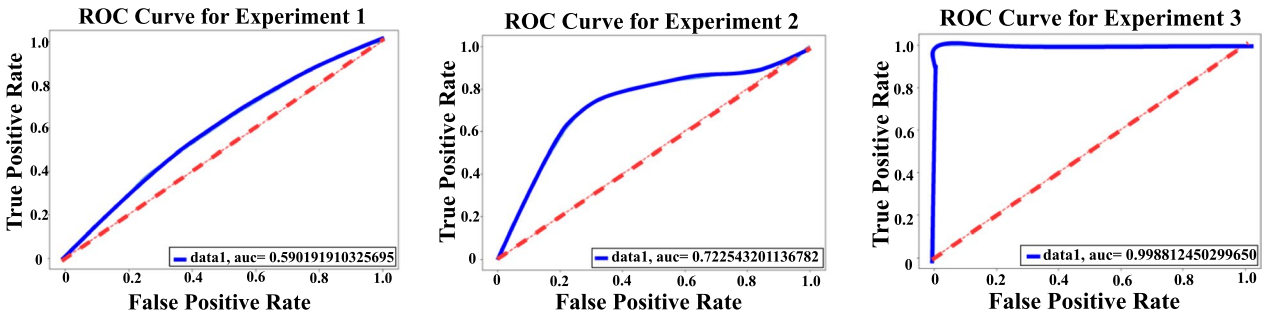
Performance evaluation

The designed dataset was randomly divided for a thorough evaluation of the *1-hour Prediction Model*s, with 30 iterations conducted to ensure the robustness and reliability of the *BlaSt* model's performance. This process helps for variability and minimizes the impact of any single random split, resulting in more consistent and generalizable outcomes as detailed in Table 3.

Experimental results analysis & discussion

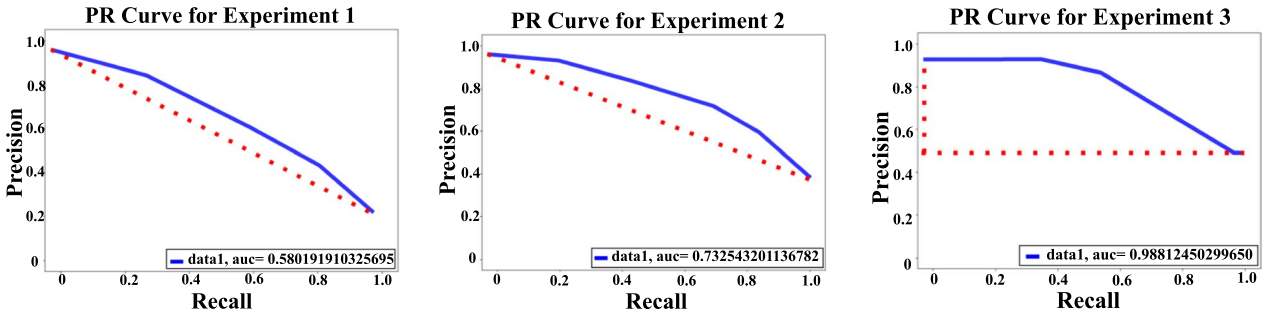
The proposed *BlaSt* model achieved superior TPR, FPR, and TNR metrics, reflecting its robust evaluation and integration with WSG with SC framework. Compared to seven existing models like SVR [31], MLP [32], RAQP [34], Vlachogianni [4], LSTM [35], BLSTM [36], and SLSTM [37] the *BlaSt* demonstrated significant accuracy improvements across experiments 1, 2, and 3, as shown in Figs. 2a–c, 3a–c.

The PR curves in Fig. 3 represent the relationship between precision and recall for each experiment, highlighting the *BlaSt* model's performance in correctly identifying positive instances. Technically, these curves



(a) (b) (c)

Fig. 2 ROC curve for (**a**) Experiment 1. (**b**) Experiment 2. (**c**) Experiment 3



(a) (b) (c)

Fig. 3 PR curve for (**a**) Experiment 1. (**b**) Experiment 2. (**c**) Experiment 3

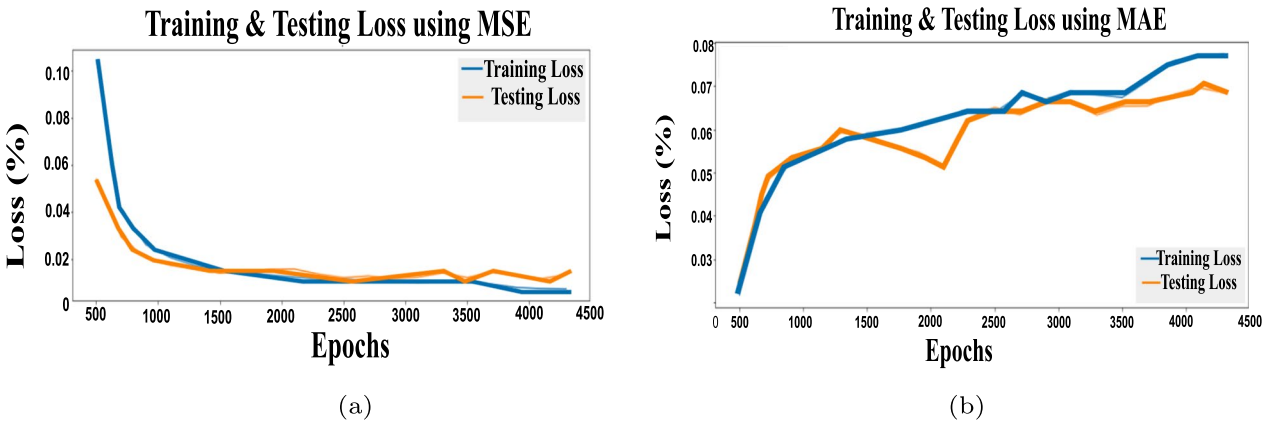


Fig. 4 (**a**) MSE and (**b**) MAE of the Proposed *BlaSt* Model

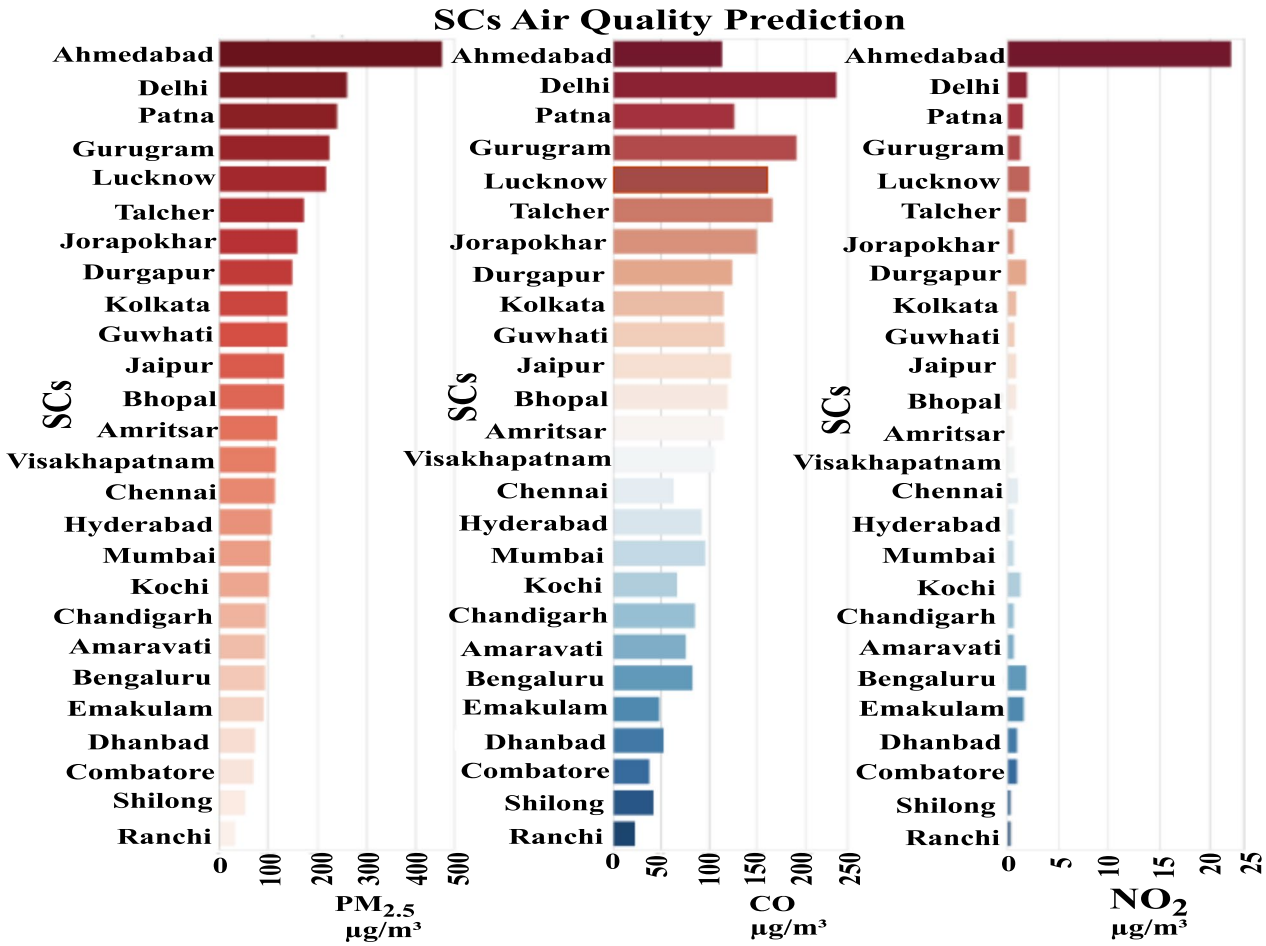


Fig. 5 SCs Future Air Quality Forecasting using the *BlaSt* Model in terms of the Future Concentration of $PM_{2.5}$, CO, and NO_2

illustrate how well the *BlaSt* model strikes a compromise between recall-the capacity to catch all pertinent positive instances-and precision-the accuracy of positive predictions-across a range of threshold values. Additionally, even when the dataset is imbalanced, a bigger area under the PR curve suggests that the *BlaSt* model achieves high recall while maintaining high precision. This means the *BlaSt* model is effective at detecting true positives with

fewer false positives, which is crucial for reliable air quality predictions in SCs. Thus, these results show the *BlaSt* model's superior ability to predict air quality in SCs, effectively mitigating challenges such as error accumulation and weakening spatiotemporal correlations over time.

The MSE and MAE losses of the *BlaSt* model are presented in Figs. 4a, b.

Our proposed *BlaSt* model is used to anticipate future air quality in SC, as shown in Fig. 5.

Figure 5, we observed the followings:

1. **Trend Analysis:** The charts reveal trends and patterns in $PM_{2.5}$, CO, and NO_2 levels over time. Analyzing these trends helps in understanding pollutant concentration fluctuations and potential sources.
2. **Peak Analysis:** Identifying peak values and their timings can highlight critical pollution events and correlate them with specific activities and environmental conditions.
3. **Seasonal Variation:** Seasonal patterns are observed, indicating variations in pollutant levels due to seasonal changes, weather conditions, or human activities.

Tables 4, 5, 6, and 7 provide a comprehensive comparison, highlighting the performance improvements achieved by *BlaSt* across various evaluation metrics.

Conclusion

The study introduces the *BlaSt* model, a pioneering air quality prediction approach in SCs, demonstrating substantial advancements through innovative methodologies and advanced techniques. It incorporates WSG, APCs, and MFs to improve accuracy and integrates a bi-directional stacking mechanism within the SC framework. *BlaSt* integrates diverse data sources through multimodal data fusion, capturing complex spatiotemporal relationships effectively while maintaining each modality's unique characteristics. *BlaSt* effectively captures temporal dependencies and trends by combining temporal aggregation, lagged characteristics, and dynamic selection. This ensures strong predictive performance even under changing conditions. Spatial relationships are

Table 4 Comparison of Accuracy between SVR [31], MLP [32], RAQP [34], Vlachogianni (Vlacho) [4], LSTM [35], BLSTM [36], SLSTM [37], and the proposed *BlaSt* Models

APCs	Method	1hr	2hr	3hr	4hr	5hr	6hr	7hr	8hr	9hr	10hr	11hr	12hr
$PM_{2.5}$	SVR	0.949	0.890	0.859	0.789	0.759	0.731	0.717	0.707	0.699	0.695	0.690	0.680
	MLP	0.958	0.910	0.890	0.886	0.805	0.778	0.748	0.739	0.723	0.715	0.699	0.690
	RAQP	0.933	0.925	0.912	0.903	0.880	0.880	0.862	0.775	0.740	0.739	0.725	0.712
	Vlacho	0.867	0.828	0.792	0.770	0.748	0.719	0.651	0.636	0.621	0.615	0.605	0.590
	LSTM	0.938	0.928	0.919	0.890	0.860	0.832	0.803	0.791	0.770	0.718	0.712	0.700
	BLSTM	0.918	0.903	0.899	0.898	0.859	0.835	0.825	0.770	0.739	0.718	0.713	0.705
	SLSTM	0.969	0.965	**0.935**	**0.912**	0.883	0.853	0.833	0.815	0.790	0.758	0.750	0.721
	BlaSt	**0.995**	**0.985**	0.920	0.900	**0.900**	**0.890**	**0.888**	**0.868**	**0.856**	**0.857**	**0.835**	**0.829**
NO_2	SVR	0.930	0.925	0.910	0.898	0.893	0.883	0.873	0.867	0.853	0.808	0.797	0.788
	MLP	0.935	0.870	0.848	0.842	0.809	0.785	0.769	0.759	0.729	0.710	0.700	0.693
	RAQP	0.966	0.952	0.938	0.910	0.894	0.877	0.862	0.838	0.799	0.788	0.758	0.708
	Vlacho	0.888	0.859	0.850	0.845	0.833	0.812	0.802	0.795	0.759	0.730	0.723	0.713
	LSTM	0.962	0.956	0.945	0.915	0.887	0.866	0.852	0.790	0.770	0.761	0.747	0.728
	BLSTM	0.987	0.964	0.940	0.915	0.894	0.865	0.847	0.810	0.799	0.770	0.756	0.728
	SLSTM	0.980	0.978	**0.960**	**0.935**	0.893	0.846	0.842	0.820	0.799	0.770	0.757	0.740
	BlaSt	**0.995**	**0.990**	0.892	0.898	**0.897**	**0.869**	**0.860**	**0.849**	**0.845**	**0.839**	**0.839**	**0.830**
CO	SVR	0.935	0.890	0.889	0.864	0.844	0.818	0.788	0.772	0.760	0.703	0.688	0.680
	MLP	0.955	0.950	0.940	0.885	0.868	0.838	0.808	0.782	0.765	0.742	0.735	0.723
	RAQP	0.962	0.943	0.910	0.895	0.883	0.863	0.847	0.829	0.812	0.784	0.760	0.748
	Vlacho	0.870	0.869	0.865	0.860	0.859	0.847	0.839	0.821	0.813	0.802	0.790	0.782
	LSTM	0.975	0.944	0.910	0.863	0.853	0.830	0.797	0.760	0.733	0.730	0.727	0.713
	BLSTM	0.991	0.964	0.940	0.908	0.894	0.866	0.827	0.810	0.790	0.770	0.747	0.718
	SLSTM	0.992	0.957	**0.938**	**0.910**	0.868	0.849	0.827	0.795	0.765	0.735	0.727	0.710
	BlaSt	**0.996**	**0.968**	0.925	0.907	**0.899**	**0.885**	**0.879**	**0.869**	**0.859**	**0.849**	**0.839**	**0.835**

Bold values indicate the efficiency as well as the superior performance of our proposed model

Table 5 Comparison of Recall between SVR [31], MLP [32], RAQP [34], Vlachogianni (Vlacho) [4], LSTM [35], BLSTM [36], SLSTM [37], and the proposed *BlaSt* Models

APCs	Method	1hr	2hr	3hr	4hr	5hr	6hr	7hr	8hr	9hr	10hr	11hr	12hr
$PM_{2.5}$	SVR	0.942	0.888	0.849	0.788	0.759	0.712	0.710	0.700	0.692	0.680	0.678	0.662
	MLP	0.958	0.911	0.899	0.856	0.815	0.790	0.738	0.713	0.701	0.700	0.695	0.650
	RAQP	0.923	0.915	0.899	0.895	0.875	0.858	0.820	0.799	0.765	0.725	0.710	0.700
	Vlacho	0.865	0.840	0.825	0.788	0.760	0.749	0.715	0.706	0.681	0.655	0.615	0.610
	LSTM	0.951	0.950	0.930	0.889	0.855	0.832	0.820	0.799	0.778	0.740	0.720	0.711
	BLSTM	0.969	0.943	0.911	0.880	0.851	0.833	0.815	0.789	0.760	0.732	0.723	0.712
	SLSTM	0.975	0.954	**0.943**	**0.922**	0.890	0.870	0.850	0.812	0.783	0.775	0.765	0.758
	BlaSt	**0.995**	**0.980**	0.930	0.910	**0.910**	**0.900**	**0.890**	**0.880**	**0.865**	**0.849**	**0.835**	**0.829**
NO_2	SVR	0.925	0.913	0.905	0.892	0.890	0.885	0.870	0.860	0.850	0.805	0.799	0.785
	MLP	0.955	0.875	0.853	0.832	0.819	0.795	0.779	0.769	0.739	0.720	0.710	0.701
	RAQP	0.969	0.960	0.930	0.920	0.895	0.870	0.860	0.848	0.795	0.785	0.768	0.720
	Vlacho	0.895	0.868	0.840	0.835	0.823	0.810	0.805	0.797	0.769	0.740	0.739	0.723
	LSTM	0.982	0.970	0.940	0.910	0.880	0.860	0.840	0.795	0.775	0.765	0.745	0.725
	BLSTM	0.973	0.960	0.945	0.910	0.890	0.860	0.840	0.815	0.790	0.775	0.750	0.730
	SLSTM	0.991	0.970	**0.965**	**0.930**	0.890	0.840	0.835	0.825	0.790	0.775	0.750	0.735
	BlaSt	**0.993**	**0.991**	0.895	0.890	**0.892**	**0.867**	**0.865**	**0.840**	**0.842**	**0.830**	**0.829**	**0.825**
CO	SVR	0.930	0.895	0.885	0.865	0.840	0.810	0.780	0.770	0.765	0.695	0.689	0.681
	MLP	0.960	0.945	0.942	0.880	0.860	0.830	0.801	0.780	0.750	0.740	0.730	0.722
	RAQP	0.960	0.940	0.911	0.890	0.873	0.853	0.843	0.832	0.810	0.780	0.765	0.740
	Vlacho	0.872	0.859	0.858	0.849	0.843	0.840	0.832	0.811	0.805	0.801	0.790	0.779
	LSTM	0.972	0.940	0.915	0.880	0.851	0.835	0.790	0.765	0.740	0.735	0.720	0.710
	BLSTM	0.973	0.960	0.945	0.900	0.890	0.860	0.820	0.811	0.795	0.775	0.740	0.708
	SLSTM	0.980	0.950	**0.934**	**0.900**	0.860	0.840	0.820	0.790	0.760	0.730	0.720	0.707
	BlaSt	**0.990**	**0.960**	0.920	0.895	**0.890**	**0.880**	**0.870**	**0.860**	**0.850**	**0.840**	**0.830**	**0.820**

Bold values indicate the efficiency as well as the superior performance of our proposed model

captured via spatial embeddings, considering localized variations and correlations. Through *1-hour Prediction Models*, real-time decision-making is made possible by enhanced sensitivity to environmental influences thanks to domain-specific information and customized feature engineering. Evaluation using the Spatiotemporal Correlation Coefficient metric (ρ) confirms the model's strength, with comparative experiments demonstrating its superiority over existing methods like SVR, MLP, RAQP, and the Vlachogianni model, resulting in significant improvements in predictive metrics.

Limitation

The primary limitation of the *BlaSt* model is that, because of its reliance on non-linear characteristics of APCs and MFs, it ignores chemical reactions and unavoidable natural disasters like acid rain. However, the absence of relevant data significantly impacts the *BlaSt* model's capacity to address these factors in air quality prediction. Thus, the potential ways to address these limitations are stated as follows:

1. **Integration of New Data Sources:** Future research could focus on integrating datasets that include chemical reactions (e.g., formation of secondary pollutants like ozone) and data related to natural calamities (e.g., acid rain, dust storms, wildfires). These datasets could be sourced from environmental monitoring agencies, satellite data, or predictive simulations.
2. **Development of Hybrid Models:** To improve the model's accuracy in scenarios involving chemical reactions and natural calamities, we propose the development of hybrid models that combine the strengths of the *BlaSt* model with additional modules specifically designed to handle these factors. For example, incorporating a chemical transport model (CTM) could simulate chemical reactions, while integrating meteorological models could predict the impact of natural calamities on air quality.
3. **Collaboration with Environmental Scientists:** Collaborating with experts in atmospheric chemistry and environmental science could provide deeper

Table 6 Comparison of Precision between SVR [31], MLP [32], RAQP [34], Vlachogianni (Vlacho) [4], LSTM [35], BLSTM [36], SLSTM [37], and the proposed *BlaSt* Models

APCs	Method	1hr	2hr	3hr	4hr	5hr	6hr	7hr	8hr	9hr	10hr	11hr	12hr
$PM_{2.5}$	SVR	0.938	0.904	0.892	0.862	0.802	0.798	0.753	0.735	0.710	0.697	0.655	0.633
	MLP	0.958	0.927	0.895	0.856	0.823	0.790	0.770	0.758	0.733	0.714	0.710	0.703
	RAQP	0.920	0.912	0.904	0.890	0.879	0.853	0.839	0.790	0.768	0.746	0.713	0.699
	Vlacho	0.867	0.844	0.813	0.785	0.734	0.716	0.687	0.650	0.629	0.610	0.605	0.599
	LSTM	0.948	0.923	0.919	0.893	0.863	0.833	0.813	0.793	0.753	0.740	0.725	0.711
	BLSTM	0.966	0.920	0.890	0.865	0.843	0.825	0.815	0.793	0.773	0.743	0.723	0.703
	SLSTM	0.979	0.954	**0.934**	**0.913**	0.880	0.863	0.825	0.801	0.790	0.765	0.743	0.723
	BlaSt	**0.995**	**0.975**	0.930	0.910	**0.896**	**0.892**	**0.880**	**0.860**	**0.860**	**0.850**	**0.840**	**0.828**
NO_2	SVR	0.915	0.913	0.903	0.888	0.883	0.873	0.863	0.857	0.843	0.798	0.783	0.778
	MLP	0.900	0.878	0.838	0.832	0.808	0.783	0.760	0.753	0.724	0.720	0.703	0.701
	RAQP	0.966	0.958	0.928	0.910	0.884	0.867	0.839	0.828	0.798	0.763	0.748	0.707
	Vlacho	0.868	0.848	0.840	0.835	0.823	0.802	0.797	0.790	0.720	0.709	0.705	0.703
	LSTM	0.963	0.943	0.940	0.905	0.884	0.856	0.847	0.784	0.760	0.751	0.737	0.718
	BLSTM	0.975	0.954	0.930	0.905	0.884	0.857	0.837	0.800	0.789	0.760	0.746	0.718
	SLSTM	0.980	0.978	**0.950**	**0.925**	0.883	0.856	0.837	0.810	0.790	0.772	0.747	0.728
	BlaSt	**0.992**	**0.980**	0.903	0.889	**0.884**	**0.874**	**0.865**	**0.858**	**0.848**	**0.843**	**0.838**	**0.836**
CO	SVR	0.923	0.889	0.869	0.854	0.834	0.808	0.778	0.762	0.720	0.697	0.688	0.678
	MLP	0.952	0.946	0.888	0.858	0.828	0.798	0.778	0.760	0.739	0.730	0.712	0.706
	RAQP	0.960	0.933	0.898	0.898	0.893	0.873	0.869	0.865	0.852	0.843	0.820	0.778
	Vlacho	0.872	0.867	0.863	0.860	0.859	0.857	0.849	0.841	0.813	0.804	0.790	0.784
	LSTM	0.973	0.934	0.900	0.875	0.843	0.820	0.787	0.750	0.738	0.722	0.717	0.703
	BLSTM	0.989	0.954	0.930	0.905	0.884	0.856	0.817	0.800	0.780	0.760	0.737	0.708
	SLSTM	0.980	0.954	**0.933**	**0.908**	0.864	0.839	0.817	0.790	0.760	0.730	0.717	0.701
	BlaSt	**0.995**	**0.965**	0.925	0.897	**0.896**	**0.886**	**0.878**	**0.865**	**0.857**	**0.847**	**0.836**	**0.834**

Bold values indicate the efficiency as well as the superior performance of our proposed model

Table 7 Comparison of Loss Metrics across Different Compared Models

Model	MSE	MAE
SVR [31]	0.035	0.043
MLP [32]	0.032	0.041
RAQP [34]	0.030	0.039
Vlachogianni [4]	0.033	0.042
LSTM [35]	0.029	0.038
BLSTM [36]	0.027	0.036
SLSTM [37]	0.026	0.035
BlaSt (Proposed)	**0.08**	**0.10**

Bold values indicate the efficiency as well as the superior performance of our proposed model

insights into the mechanisms of chemical reactions and natural calamities. This collaboration could lead to the incorporation of domain-specific knowledge into the *BlaSt* model, enhancing its predictive capabilities.

Hence, future research directions include integrating new data sources, developing hybrid models, working with environmental scientists, and using simulated datasets to improve the model's ability to account for chemical reactions and natural disasters in air quality forecasting. The current version of the *BlaSt* model has limitations because relevant data is not readily available. We anticipate that these methods will greatly increase the *BlaSt* model's accuracy and resilience in subsequent iterations.

Institutional review board

Not applicable.

Author contributions

K.C. and M.S.R. contributed in the conceptualization, formal analysis, investigation, methodology, and writing of the original draft. S.C., S.M., A.L. and M.A.S. reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

The authors received no funding from their institutes.

Availability of data and materials
The dataset generated and/or analyzed during the current study are available in the *CityPulse Dataset Collection* repository, http://iot.ee.surrey.ac.uk:8080/datasets.html

Declarations

Informed consent
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Computer Science & Engineering, Nalla Malla Reddy Engineering College, Hyderabad 500088, Telangana, India. [2]Artificial Engineering & Machine Learning, Nalla Malla Reddy Engineering College, Hyderabad 500088, Telangana, India. [3]Computer Science & Engineering, Indian Institute of Information Technology, Vadodara, Gandhinagar 362520, Gujarat, India. [4]Computer Science and Information Systems, Faculty of Science and Technology (IcfaiTech), ICFAI Foundation for Higher Education, Hyderabad 501203, India. [5]Computer Science and Information Systems, The West Virginia University Institute of Technology, Beckley, West Virginia, USA. [6]Department of Environmental Health, Harvard T H Chan School of Public Health, Boston, MA 02115, USA. [7]Department of Pharmacology & Toxicology, The University of Arizona, Tucson, MA 85721, USA. [8]Department of Economics, Kardan University, Kabul 1001, Afghanistan. [9]School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, Shaanxi, China. [10]Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, 140401, Punjab, India. [11]Chitkara Centre for Research and Development, Chitkara University, Baddi, Himachal Pradesh, 174103, India. [12]Division of Research and Development, Lovely Professional University, Phagwara, Punjab, 144001, India.

Received: 29 April 2024 Accepted: 10 January 2025
Published online: 11 February 2025

References

1. De Marco A, Mangano G. Evolutionary trends in smart city initiatives. Sustain Futures. 2021;3: 100052.
2. Kolozali Ş, Bermudez-Edo M, FarajiDavar N, Barnaghi P, Gao F, Ali MI, Mileo A, Fischer M, Iggena T, Kuemper D, et al. Observing the pulse of a city: a smart city framework for real-time discovery, federation, and aggregation of data streams. IEEE Internet of Things J. 2018;6(2):2651–68.
3. Adewuyi AA, Cheng H, Shi Q, Cao J, Wang X, Zhou B. Sc-trust: a dynamic model for trustworthy service composition in the internet of things. IEEE Internet of Things J. 2021;9(5):3298–312.
4. Vlachogianni A, Kassomenos P, Karppinen A, Karakitsios S, Kukkonen J. Evaluation of a multiple regression model for the forecasting of the concentrations of nox and pm10 in athens and helsinki. Sci Total Environ. 2011;409(8):1559–71.
5. Feng R, Gao H, Luo K, Fan J-R. Analysis and accurate prediction of ambient pm2. 5 in china using multi-layer perceptron. Atmospheric Environ. 2020;232: 117534.
6. Javidroozi V, Shah H, Feldman G. Urban computing and smart cities: towards changing city processes by applying enterprise systems integration practices. IEEE Access. 2019;7:108 023-108 034.
7. Wang Y, Kong T. Air quality predictive modeling based on an improved decision tree in a weather-smart grid. IEEE Access. 2019;7:172 892-172 901.
8. Fitzgerald S, Jimenez D, Findling S, Yorifuji Y, Kumar M, Wu L, Carosella G, Ng S, Parker R, Carter P et al., "Idc futurescape: Worldwide digital transformation 2021 predictions," *IDC, October*, 2020.
9. Fahad MGR. Towards a resilient community: application of advanced computational models and big data analytics. Birmingham: The University of Alabama at Birmingham; 2021.
10. Mak HWL, Lam YF. Comparative assessments and insights of data openness of 50 smart cities in air quality aspects. Sustain Cities Soc. 2021;69:02868.
11. Huang X, Srikrishnan V, Lamontagne J, Keller K, Peng W. Effects of global climate mitigation on regional air quality and health. Nat Sustain. 2023;6(9):1054–66.
12. Dey S. "Urban air quality index forecasting using multivariate convolutional neural network based customized stacked long short-term memory model," *Process Safety and Environmental Protection*, 2024.
13. Neo EX, Hasikin K, Lai KW, Mokhtar MI, Azizan MM, Hizaddin HF, Razak SA, et al. Artificial intelligence-assisted air quality monitoring for smart city management. PeerJ Comput Sci. 2023;9: e1306.
14. Gao H, Jiang W, Ran Q, Wang Y. Vision-language interaction via contrastive learning for surface anomaly detection in consumer electronics manufacturing. IEEE Trans Consum Electron. 2024. https://doi.org/10.1109/TCE.2024.3378771.
15. Weng Y, Dong J, He W, Liu X, Liu Z, Gao H, et al. Zero-shot cross-lingual knowledge transfer in vqa via multimodal distillation. IEEE Trans Comput Soc Syst. 2024. https://doi.org/10.1109/TCSS.2024.3402270.
16. Gao H, Wu Y, Xu Y, Li R, Jiang Z. Neural collaborative learning for user preference discovery from biased behavior sequences. IEEE Trans Comput Soc Syst. 2023. https://doi.org/10.1109/TCSS.2023.3268682.
17. Gao H, Wang X, Wei W, Al-Dulaimi A, Xu Y. Com-ddpg: task offloading based on multiagent reinforcement learning for information-communication-enhanced mobile edge computing in the internet of vehicles. IEEE Trans Veh Technol. 2023. https://doi.org/10.1109/TVT.2023.3309321.
18. Bi Z, Sun S, Zhang W, Shan M. Click-through rate prediction model based on graph networks and feature squeeze-and-excitation mechanism. Int J Web Inform Syst. 2024;20(4):341–57.
19. Qi Z, Lu T, Yue K, Duan L. Similarity search on social networks with incremental graph indexing based on probabilistic inference. Int J Web Inform Syst. 2024;20(4):395–412.
20. Song J, Stettler ME. A novel multi-pollutant space-time learning network for air pollution inference. Sci Total Environ. 2022;811: 152254.
21. Song J, Han K, Stettler ME. Deep-maps: machine-learning-based mobile air pollution sensing. IEEE Internet of Things J. 2020;8(9):7649–60.
22. Song J, Fan H, Gao M, Xu Y, Ran M, Liu X, Guo Y. Toward high-performance map-recovery of air pollution using machine learning. ACS ES &T Eng. 2022;3(1):73–85.
23. Chai J, Song J, Fan H, Xu Y, Zhang L, Guo B, Xu Y. St-bikes: predicting travel-behaviors of sharing-bikes exploiting urban big data. IEEE Trans Intell Transp Syst. 2022;24(7):7676–86.
24. Zhang X, Zhou C, Zhang Y, Lu X, Xiao X, Wang F, Song J, Guo Y, Leung KK, Cao J, et al. Where to place methane monitoring sites in China to better assist carbon management. npj Clim Atmospheric Sci. 2023;6(1):32.
25. Zaini N, Ean LW, Ahmed AN, Malek MA. A systematic literature review of deep learning neural network for time series air quality forecasting. Environ Sci Pollut Res. 2022. https://doi.org/10.1007/s11356-021-17442-1.
26. Huang Y, Ying JJ-C, Tseng VS. Spatio-attention embedded recurrent neural network for air quality prediction. Knowl-Based Syst. 2021;233: 107416.
27. Cai J, Liu T, Wang T, Feng H, Fang K, Bashir AK, Wang W. Multi-source fusion enhanced power-efficient sustainable computing for air quality monitoring. IEEE Internet of Things J. 2024. https://doi.org/10.1109/JIOT.2024.3420956.
28. Shah J, Mishra B. Analytical equations based prediction approach for pm2. 5 using artificial neural network. SN Appl Sci. 2020;2(9):1516.
29. Cordova CH, Portocarrero MNL, Salas R, Torres R, Rodrigues PC, López-Gonzales JL. Air quality assessment and pollution forecasting using artificial neural networks in metropolitan lima-peru. Sci Rep. 2021;11(1):24232.
30. Septiawan WM, Endah SN. "Suitable recurrent neural network for air quality prediction with backpropagation through time," in 2nd international conference on informatics and computational sciences (ICICoS). IEEE. 2018;2018:1–6.
31. Zhu H, Hu J. "Air quality forecasting using svr with quasi-linear kernel," in *2019 International Conference on Computer, Information and Telecommunication Systems (CITS)*. IEEE, 2019, pp. 1–5.
32. Abdullah S, Ismail M, Ahmed A. Multi-layer perceptron model for air quality prediction. Malays J Math Sci. 2019;13:85–95.
33. Puiu D, Barnaghi P, Tönjes R, Kümper D, Ali MI, Mileo A, Parreira JX, Fischer M, Kolozali S, Farajidavar N, et al. Citypulse: large scale data analytics framework for smart cities. IEEE Access. 2016;4:1086–108.

34. Gu K, Qiao J, Lin W. Recurrent air quality predictor based on meteorology- and pollution-related factors. IEEE Trans Ind Inform. 2018;14(9):3946–55.
35. Krishan M, Jha S, Das J, Singh A, Goyal MK, Sekar C. Air quality modelling using long short-term memory (lstm) over nct-Delhi, India. Air Qual, Atmosphere Health. 2019;12:899–908.
36. Zhang L, Liu P, Zhao L, Wang G, Zhang W, Liu J. Air quality predictions with a semi-supervised bidirectional lstm neural network. Atmospheric Pollut Res. 2021;12(1):328–39.
37. Manna T, Anitha A. "Forecasting air quality index based on stacked lstm," in *2022 IEEE 7th International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, vol. 7. IEEE, 2022, pp 326–330
38. Sofaer HR, Hoeting JA, Jarnevich CS. The area under the precision-recall curve as a performance metric for rare binary events. Methods Ecol Evolut. 2019;10(4):565–77.
39. Varoquaux G, Colliot O. Evaluating machine learning models and their diagnostic value. Mach Learn Brain Disord. 2023. https://doi.org/10.1007/978-1-0716-3195-9_20.

Publisher's Note