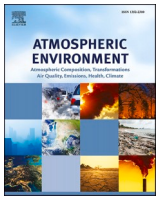
Atmospheric Environment 340 (2025) 120921

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv



An attention-based CNN model integrating observational and simulation data for high-resolution spatial estimation of urban air quality

Shibao Wang [a], Yanxu Zhang [b,*]

[a] School of Atmospheric Sciences, Nanjing University, Nanjing, 210023, China
[b] Department of Earth and Environmental Sciences, Tulane University, New Orleans, LA, 70118, USA

HIGHLIGHTS

- The model integrates observational and simulation data, outperforming traditional CNNs in air quality estimation.
- Dynamic attention weights boost efficiency by focusing on critical features.
- Model reveals fluid dynamics laws, enhancing understanding of pollution patterns.
- Model exhibits strong transferability and generalization across diverse cities.

ARTICLE INFO

Keywords:
High-resolution
Mobile monitoring
Large eddy simulation
Convolutional neural network
Attention mechanism

ABSTRACT

Machine learning, especially deep learning, can outperform traditional atmospheric models in air quality assessment, offering enhanced efficiency and accuracy without relying on detailed emission inventories and atmospheric chemical mechanisms. Despite their predictive power, deep learning models often grapple with the perception of being "black boxes" due to their intricate architectures. Here, we develop an attention-based convolutional neural network (CNN-attention) model that incorporates observational data, the parallelized large-eddy-simulation model (PALM), and urban morphology data for high-resolution spatial estimation of urban air quality. Our findings indicate that the CNN-attention model outperforms traditional CNN with higher accuracy and efficiency, achieving $R^2 = 0.987$ and root mean square error (RMSE) = 0.15 mg/m^3, while significantly reducing training time and memory usage. Compared to traditional machine learning models, the CNN exhibits higher R^2 values and lower RMSE, showcasing its adeptness at capturing complex nonlinear patterns. The inclusion of attention layer further improves the model's performance by dynamically assigning attention scores to key features, enabling the model to focus on areas of critical emissions and distinctive urban features such as highways, arterial roads, intersections, and dense building clusters. This approach also reveals fluid dynamical principles, highlighting the significant disparities in pollutant concentration across roadways caused by atmospheric turbulence, and the distinct plume formations influenced by land use and topography. When applied to various urban settings, the CNN-attention model exhibits superior generalizability and transferability. This study provides valuable scientific insights and technical support for urban planning, air quality management, and exposure risk evaluation.

1. Introduction

Currently, approximately 94 percent of world's population live in areas where air quality does not meet the updated global air quality guidelines set by the World Health Organization, presenting serious public health risks (WHO, 2016; Shiraiwa et al., 2017; Rentschler and Leonova, 2023). Monitoring and estimating urban air quality pose considerable challenges due to the high temporal and spatial heterogeneity, which stems from complex emission sources, the building environment, and non-linear atmospheric chemistry (Xu et al., 2021; Tripathi et al., 2023). The urban landscape, including variations in buildings and vegetation, also influences pollutant dispersion via atmospheric turbulent (Fu et al., 2017; Jeanjean et al., 2015). Understanding the localized dynamics of air pollution is crucial for identifying

* Corresponding author.
E-mail address: yzhang127@tulane.edu (Y. Zhang).

https://doi.org/10.1016/j.atmosenv.2024.120921
Received 23 August 2024; Received in revised form 19 October 2024; Accepted 4 November 2024
Available online 5 November 2024

high-emission zones and pollution hotspots (Deshmukh et al., 2020; Gately et al., 2017). Traffic-related emissions are often a primary source of numerous urban air pollutants, and prolonged exposure to these contaminants poses serious health risks (Chen et al., 2024). Carbon monoxide (CO) serves as a crucial indicator for urban air quality assessment and is influenced by local emission and meteorological conditions, making its accurate spatial estimation complex. The advent of mobile low-cost sensors has been transformative, providing fine-scale temporal and spatial data that complement traditional fixed-site monitoring methods (Apte et al., 2017; Van den Bossche et al., 2015). Numerous studies have deployed mobile sensor platforms to capture the complex spatial and temporal variations of urban air pollutants (Kaivonen and Ngai, 2020; Peters et al., 2014). Nevertheless, mobile monitoring still struggles to achieve comprehensive micro-scale coverage, particularly in remote streets and off-road pollution areas. To address this, we propose a novel method that integrates mobile monitoring, large eddy simulation, and machine learning to estimate urban air quality more effectively.

Large eddy simulation (LES) models represent a significant advancement in modeling atmospheric turbulence and air quality, particularly for urban environments. These models excel in capturing complex dynamics of turbulent eddies and providing high-resolution simulations (Letzel et al., 2008; Resler et al., 2021). LES outperforms traditional models such as Gaussian, CMAQ, and ADMS-Urban by explicitly resolving larger-scale turbulent motions and parameterizing smaller scales, thus offering a more comprehensive depiction of atmospheric dynamics (Sun et al., 2016; Wang and Ngan, 2021). The Parallelized Large-Eddy Simulation Model (PALM) has emerged as a powerful tool for urban air quality research, proficient in handling complex urban geometries, achieving high parallelization efficiency, and accommodating diverse physical processes (Maronga et al., 2015). Recent studies have demonstrated PALM's capacity to simulate fine-scale pollution patterns, considering the influence of urban topography and meteorological conditions (Wolf et al., 2021). For instance, Zhang et al. (2021) utilized the PALM to simulate traffic-related CO diffusion in an urban environment with an ultra-high resolution of 10 m, revealing complex pollution patterns and the influence of different factors on dispersion. Despite their strengths, LES models encounter limitations, such as high computational demands, the need for real-time weather data, and difficulties in representing multiple emission sources (Maronga et al., 2015).

Machine learning (ML) has been widely applied to urban air quality evaluation at various scales (Lim et al., 2019; Liu et al., 2020; Zhong et al., 2021). Deep learning techniques, particularly long short-term memory (LSTM) and convolutional neural networks (CNNs), have significantly advanced in estimating air pollution levels and their complex nonlinearities, outperforming traditional ML methods (Aggarwal and Toshniwal, 2021; Xu et al., 2022). CNNs, designed for hierarchical image representation, utilize layered artificial neural networks to effectively extract features and reduce estimation uncertainty via successive nonlinear transformations (Kow et al., 2020; LeCun et al., 2015). Guo et al. (2023) proposed a Convolutional LSTM (ConvLSTM) network, a deep spatio-temporal learning model that utilizes dense monitoring data to capture spatio-temporal patterns in spatial-map sequences for citywide air quality assessment. While numerous studies have focused on enhancing the accuracy of ML models for air quality estimation, they continue to struggle with several challenges, including the demand for high computational resources, model complexity, and the opacity of the "black box" nature of these algorithms (Moursi et al., 2022).

The attention mechanism enhances the trade-off between accuracy and interpretability in deep learning models. mirroring human attention processes. As a pivotal element in machine learning and natural language processing (Bahdanau et al., 2016), it allows models to dynamically prioritize input data, thereby enhancing critical information processing and model interpretability (Gu et al., 2021; Hu et al., 2022). Recent studies have incorporated the attention mechanism into various machine learning models to enhance urban air quality estimation (Wang et al., 2023). For example, Li et al. (2020) developed an attention-augmented CNN-LSTM model that improves both the accuracy and interpretability of $PM_{2.5}$ forecasts. Additionally, Hu et al. (2022) proposed an attention-based convolutional LSTM (AB-ConvLSTM) to predict urban mass transit speeds, leveraging the attention mechanism to discern and emphasize relevant historical data segments. However, these models tend to overlook diverse influencing factors, predominantly focusing on data-driven estimation while disregarding the physical dynamics of pollutant dispersion (Reichstein et al., 2019).

In this study, we develop an attention-based CNN model that integrates mobile observations, and large eddy simulations outputs to address the limitations inherent in each method while enhancing the model's accuracy, interpretability, and capacity to learn key features and physical laws. The main objective of this work is to propose an attention-based CNN model for the rapid and precise spatial estimation of high-resolution urban air quality. Initially, we explore the spatial-temporal autocorrelation of observational data and its impact on model performance. We then conduct a comparative analysis of the traditional CNN and the enhanced CNN-attention model, employing cross-validation to ensure a robust assessment of their performance. We also elucidate the principles of the attention mechanism, which enhances the model's interpretability. Finally, by applying the CNN-attention model to various urban environments, we assess the its generalizability and transferability, which are crucial for its application in diverse settings.

2. Materials and methods

2.1. Data collection and pre-processing

The dataset described here, published in Wang et al. (2021), comprises air quality data collected from the streets of Nanjing over a one-year period (October 2019 to September 2020). This data was gathered using two taxis equipped with mobile sensors (model XHAQSN-508, Hebei Sailhero Environmental Protection High-tech Co., Ltd., Hebei, China). The mobile monitoring system is equipped with a gas sensing module that detects CO, nitrogen dioxide (NO_2), and Ozone (O_3). The CO sensor (model XH-CO-50-7) operates on electrochemical principles. The system also includes a GPS module that records location data every 10 s, and a data logger that captures and transmits sensor readings to a central server. A temperature/humidity (T/H) sensor module is included for calibration purposes. To ensure the measurement accuracy of the XHAQSN-508 system, it undergoes monthly calibrations at the Station for Observing Regional Processes of the Earth System (SORPES) in Nanjing, with each calibration lasting a minimum of seven days (Ding et al., 2013). During calibration, the sensor outputs were fine-tuned using a standard measurement device, yielding CO concentration readings with a coefficient of determination (R^2) between 0.82 and 0.97. Detailed documentation of the calibration process can be found in Wang et al. (2021). Fig. 1B illustrates the spatial distribution of the processed CO concentrations at a 50 m resolution, which were obtained from mobile observations and processed with ArcMap 10.7.

The dataset described here, published in Zhang et al. (2021), uses the PALM-4U model (version 3689) to simulate CO dispersion, a traffic-related pollutant, across Nanjing with a high spatial resolution of 10 m. This model, developed by the team at Leibniz University of Hanover (Maronga et al., 2015), applies the Boussinesq approximation to the incompressible Navier-Stokes equations and is specifically designed for LES of turbulent flows. Optimized for high-performance computing, PALM efficiently conducts LES on parallel computing systems. The study central area is located at 118.72° E, 32.07° N. To minimize the impact of urban structures on the simulations, we use a grid stretching technique with an expansion ratio of 1.1 across 48 vertical levels. This extends the model's vertical domain to approximately 1000 m. Such a configuration requires an initial convergence phase,

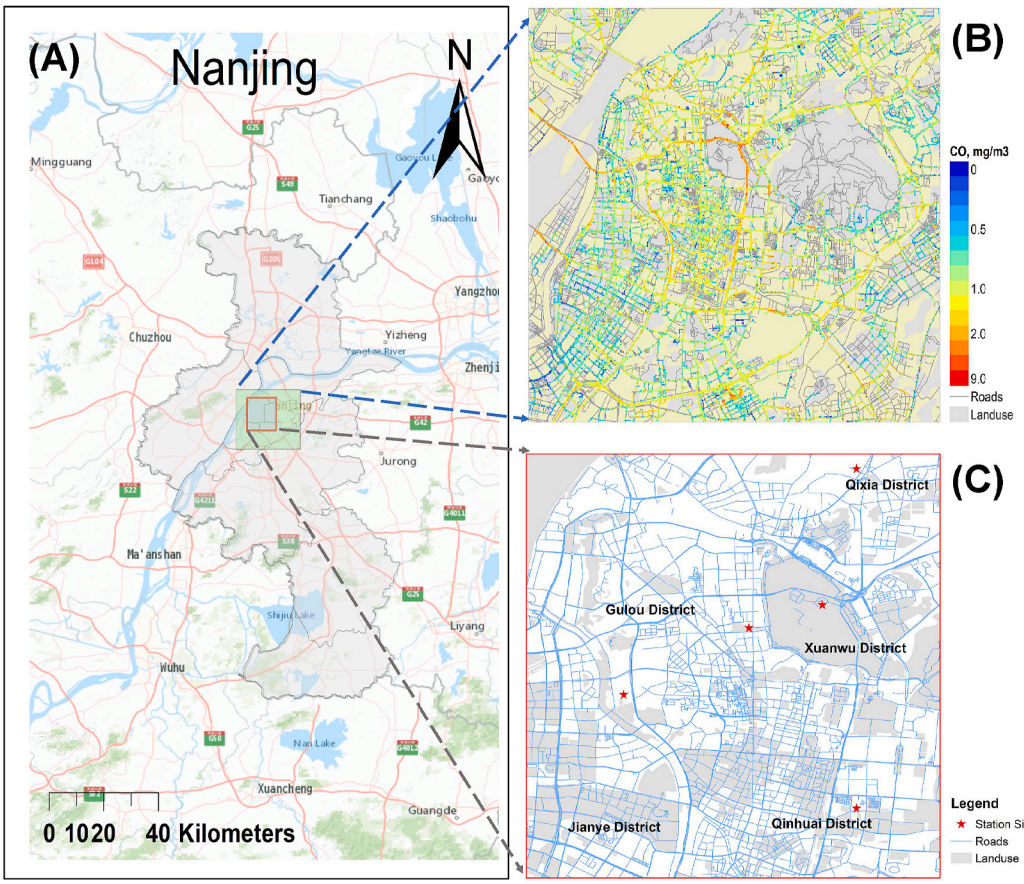


Fig. 1. Location of the monitoring areas in the city of Nanjing urban area. B means the extent of the observation area and the results of the 50 m grid point distribution of CO concentration. C represents the simulation area of the PALM model, with red stars marking the locations of background monitoring stations that serve as reference points for comparing and validating the model's outputs. The observational data include influences from both local sources and external sources transported through the atmosphere. Map credit: ESRI 2024.

during which the simulations are conducted under static boundary conditions. The model's grid resolution is set at 0.0001°, resulting a grid matrix of 960 × 960 × 48 (Fig. 1C). The simulations run for 3 h with a time step of 6 s. The initial 2 h are dedicated to allowing the turbulent field to stabilize, with the results from the final hour being averaged to indicate the model's attainment of a steady state.

In this study, the CNN-attention model incorporates four categories of input data: building heights, topography, CO emission rates, and observed CO concentrations (Figs. S1 and S3). The topographic data includes both the baseline elevation, with an original resolution of 30 m, and the building heights. The elevation data undergo linear interpolation to mathch the model's grid resolution (Computer Network Information Center). Data on building heights, obtained from Gaode Maps (https://ditu.amap.com), are estimated based on an average of 3 m per story and adjusted to fit the model's grid. CO emissions are estimated using the "standard road length" method, which distributes the total traffic emissions across individual roads according to road type and traffic volume (Zheng et al., 2009). The method for detailed calculation is elaborated in Zhang et al. (2021). We consider two emission scenarios, high and low, each estimated on a uniform wind field at the boundary layer's top. Due to computational limits, only a selection of meteorological conditions is simulated. We incorporate eight wind directions (N, NE, E, SE, S, SW, W, and NW), each 45° apart. Wind speeds of 10 m/s, 6.5 m/s, and 3 m/s to represent strong, moderate, and weak wind conditions, respectively (Yan et al., 2021; Zhang et al., 2021). This results in 48 distinct scenarios simulated using the PALM (Fig. S2). The CNN-attention model is trained on approximately 44,236,800 data points (48 scenarios × 960 × 960 grid points). For the testing phase, we use a separate set of scenarios not included in the training set to ensure an unbiased evaluation of the model's performance.

To assess the impact of observational data on the CNN-attention model's accuracy, we categorize mobile observation data into working and non-working days using the grid-based processing method (Wang et al., 2021). Employing ArcMap 10.7, we spatially integrate observation data with the corresponding grid points. The data is then sorted based on temporal characteristics, with the annual mean observed values at each grid point adjusted to match the model's resolution of 0.0001°. Fig. S3 illustrates the spatial distribution of observed CO concentrations at 10 m resolution for both weekdays and weekends. It shows that median CO levels on roads are higher during weekdays (0.86 ± 1.08 mg/m^3) than on weekends (0.82 ± 1.02 mg/m^3). This difference is primarily due to the variations in traffic emissions (Li et al., 2018). Furthermore, CO concentrations are consistently higher on major roads than on secondary roads. To enhance model convergence, we normalize all input data, effectively reducing the potential impact of differing scales and dimensions.

2.2. Attention-based CNN model

Fig. 2 shows the architecture of the CNN-attention model. In the previous section, we standardize the four input feature variables to a uniform resolution and normalize them to minimize feature discrepancies. A Lambda function is then employed to refine, rotate, and crop the input data. These preprocessed input features are then passed through a convolutional layer, where the Adam optimizer and ReLU activation function facilitate model training. L2 regularization (Lambda

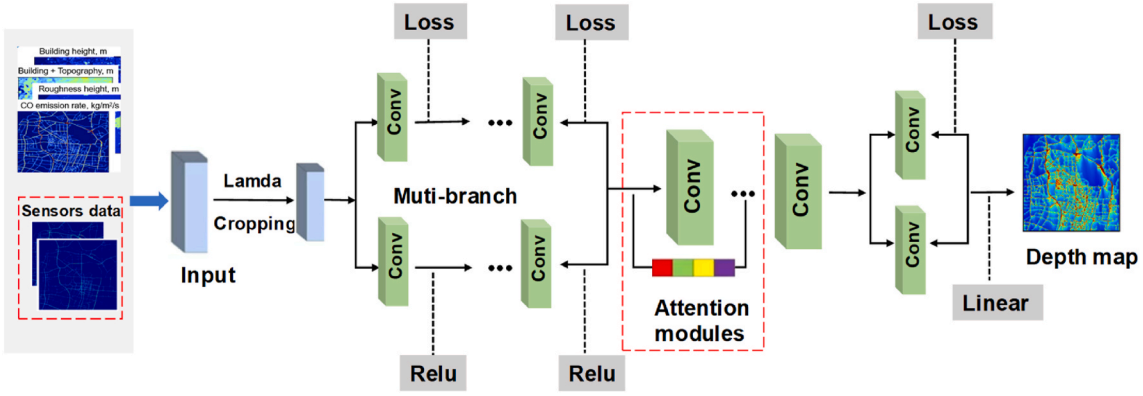


Fig. 2. CNN-attention model constructure.

= 1e-4) is applied to prevent overfitting. The model's initial layer uses 64 kernels of size 21 × 21 with ReLU activation. This is followed by a second layer with 64 kernels of size 41 × 41, and a third layer with the same number of kernels at a size of 61 × 61, both layers following the same processing protocol as the first. Subsequent to the convolutional layers, the model incorporates both a channel attention layer and a spatial attention layer to enhance feature representation. The attention layer in the model assigns weights to different channel and spatial features, allowing the model to focus on the most important aspects of the input data. After processing by the attention mechanisms, a concatenate function is used to merge all the feature maps. This ensures that all vital features are simultaneously considered in the subsequent layers. An additional convolutional layer with 64 kernels follows the attention layers to enhance feature extraction and interation. The final output of the model is produced by a Conv2D layer with a linear activation function. To thoroughly train the model on the features represented in the PALM model, we use the outputs from all 48 scenarios as our training dataset, enabling the model to estimate CO concentration distributions for each scenario accurately.

To address the high memory requirements of training the CNN-attention model, we use hardware acceleration through an NVIDIA 4090 GPU with 24 GB of memory. This powerful GPU supports a training batch size of 20, enhancing computational efficiency. Training progresses in steps, with step corresponding to one batch and an epoch encompassing all batches. We have programmed the training to span 512 epochs, with each epoch taking approximately 20 min. The model's execution and the processing of results are performed within a Python 3.10 environment. Detailed training configurations and hyper-parameters are listed in Table S1.

2.3. Attention mechanism

In this study, we use a dual attention mechanism that integrates both channel and spatial attention layers to improve model performance. Fig. 3 illustrates the workflow of this attention mechanism. Initially, the process involves global average pooling, which condenses each channel of the feature map to its mean features. These mean features serve as inputs to a dense layer that reduces their dimensionality through the ReLU activation function. Subsequently, another dense layer expands these features to the original feature space's dimensionality, with the sigmoid activation function assigning weights between 0 and 1. These weights are applied to scale the original feature map on an element-wise basis, thus adjusting the relevance of each channel. The formula for computing channel attention scores is presented as follows (Wang et al., 2020):

$$A = \text{sigmoid}(W_2 \cdot \text{ReLU}(W_1 \cdot F_c + b_1) + b_2) \quad (1)$$

Where, W_1 and b_1 represent the weights and biases of the first fully

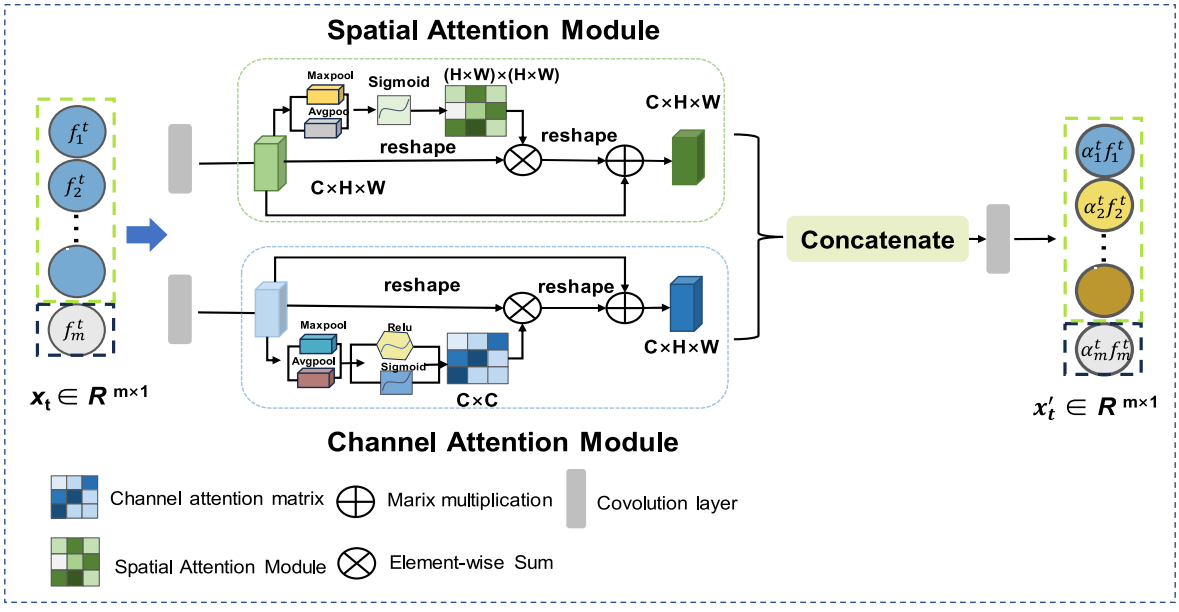


Fig. 3. Constructure of attention mechanism.

connected layer, respectively. Similarly, W_2 and b_2 correspond to the weights and biases of the second fully connected layer. *Fc* denotes the average eigenvalue of each channel obtained by global average pooling, and *c* denotes the number of channels. ReLU is a non-linear activation function, while the sigmoid activation function maps the output to a range between 0 and 1.

Then, channel reweighting is calculated.

$$X_c^{'} = X_c \cdot A_c \tag{2}$$

Where, A_c represents the attention weight for channel c, and $X_c^{'}$ is the reweighted channel feature map.

Attention scores enable researchers to gain an intuitive understanding of a model's decision-making process by highlighting which features are prioritized through increased attentional weights. This insight reveals the types of information that the model considers most critical. Within the scope of channel attention, the softmax function can normalize weights across channels, making their total equal to one, thus facilitating comparability. The calculation of attention proportions is determined by the following formula:

$$Attention\ Propotion_i = \frac{e^{Attention\ Score_i}}{\sum_j e^{Attention\ Score_j}} \tag{3}$$

Where, $Attention\ Score_i$ is the raw attention score for the *i*th channel, and $Attention\ Score_j$ is the raw attention score for the *j*th feature channel. Attention Proportion$_i$ represents the normalized attention score proportion for the *i*th channel. The term $\sum_j e^{Attention\ Score_j}$ sums the $e^{Attention\ Score_i}$ values across all channels, ensuring that the total sum of all $Attention\ Propotion_i$ equals 1. This normalization is crucial for comparability of attention weights across channels.

The spatial attention layer aims to assess the importance of each spatial location by examining the characteristics of the feature map (Wang et al., 2023). Initially, the channel-wise average and maximum values are computed from the feature map. These statistics are concatenated along the channel axis, forming an augmented feature map. This map is then processed by a 2D convolutional layer (Conv2D), yielding a spatial weight map that assigns a relative importance to each spatial location. The original feature map, denoted as (*F*), is then scaled on an element-wise by the spatial attention weight map, denoted as (*S*), thus adjusting the prominence of each spatial position.

$$F_{avg} = \frac{1}{C}\sum_{c=1}^{C} Fc \tag{4}$$

$$F_{max} = max_{c=1}^{C}\, F_c \tag{5}$$

$$F_{new} = Concat\left(F_{avg}, F_{max}\right) \tag{6}$$

$$S = Conv2D(F_{new}) \tag{7}$$

$$S_{norm} = sigmoid(S) \tag{8}$$

$$F' = F \odot S_{norm} \tag{9}$$

Where, F_{avg} denotes the average value of the channel dimension and F_{max} denotes the maximum value; F_{new} is the new feature map obtained by splicing F_{avg} and F_{max} over the channel dimension. Concat refers to the splicing operation. S_{norm} denotes the normalized spatial attention weight map. F′ denotes the final spatially reweighted feature map. $\odot$ represents element-by-element multiplication. By calculating the spatial attention weights, the model can focus more on spatial regions that are crucial for estimation, thus improving the efficiency and accuracy of feature extraction.

2.4. Model evaluation

To comprehensively assess the CNN-attention model's performance and the transferability, we use two cross-validation methods, including 10-fold cross-validation and out-of-sample cross-validation (Wang et al., 2024). We perform a 10-fold cross-validation by dividing the data into 10 different folds. In 10-fold cross-validation, the data is partitioned into 10 subsets, with each subset serving once as the validation set and the remaining 9 as the training set. For out-of-sample cross-validation, we test the model's robustness using three unseen scenario datasets (SSE, 2 m/s; WSW, 4.75 m/s; and NNE, 8 m/s) that are separate from the training data (Wang et al., 2024). The model's accuracy is quantified using the coefficient of determination (R^2) and root mean square error (RMSE). By combining these two cross-validation methods, we validate the model's reliability on historical data, and its stability on unseen datasets. This approach allows for a more comprehensive evaluation of the model's generalization ability.

3. Results and discussion

3.1. Spatio-temporal autocorrelation analysis

To investigate the spatial and temporal autocorrelations of pollutants and improve the estimation accuracy of our model, we analyze CO concentration from mobile observation data (Crocchianti et al., 2021). This autocorrelation analysis is crucial as it elucidates the model's error structure, and reveals potential confounding factors, thereby accurately reflecting the inherent characteristics of the dataset. Fig. 4 shows the autocorrelation function (ACF) and the partial autocorrelation function (PACF) for CO concentrations over a one-year period, obtained from these mobile sources. The raw data, recorded at 10 s intervals, shows minimal variation over short periods, especially when vehicular speed is low. Thus, we use a 5-min sliding average for this analysis. As depicted in Fig. 4a, CO concentration data at the minute level exhibit a significant positive autocorrelation within short periods, such as within 1 h, which diminishes with increasing lag time (Crocchianti et al., 2021). The absence of substantial periodic fluctuations in the ACF plots suggests that minute-to-minute variations in pollutant concentrations may lack significant periodicity, possibly due to the random nature of taxi routes. Fig. 4b indicates that the partial autocorrelation at lag 1 is notable, it decreases rapidly and approaches 0 as the lag increases. This pattern shows a strong short-term influence of prior pollutant concentrations. To reduce the impact of these temporal dynamic on the CNN model's estimation accuracy, we calculate the annual average CO concentration for each 10 m grid points.

In this research, we utilize Moran's index and hotspot analysis to investigate the spatial autocorrelation of CO concentration data, which is critical for enhancing model estimation accuracy. Fig. 5 illustrates that the Moran's index for CO concentrations decreases with increasing distance, signifying a strong spatial autocorrelation at shorter ranges that weakens over longer distances. Specifically, the index value of 0.54 at near 0 distance indicates significant autocorrelation for nearby observations, which diminishes beyond 200 m, suggesting that CO concentrations are spatially autocorrelated within a 200 m range. This is consistent with Crocchianti et al. (2021), who reported a similar spatial influence range for particulate matter larger than 1.1 μm decreased to 60 m and to 155 m for NO. Leveraging these spatial autocorrelation findings, we refine our data processing to reflect the model's resolution, computing the annual average CO concentrations at each 10 m grid point. This approach preserves spatial variability and enhances the model's capacity to capture spatial features and identify local pollution sources and emission hotspots, thereby improving the its estimation accuracy.

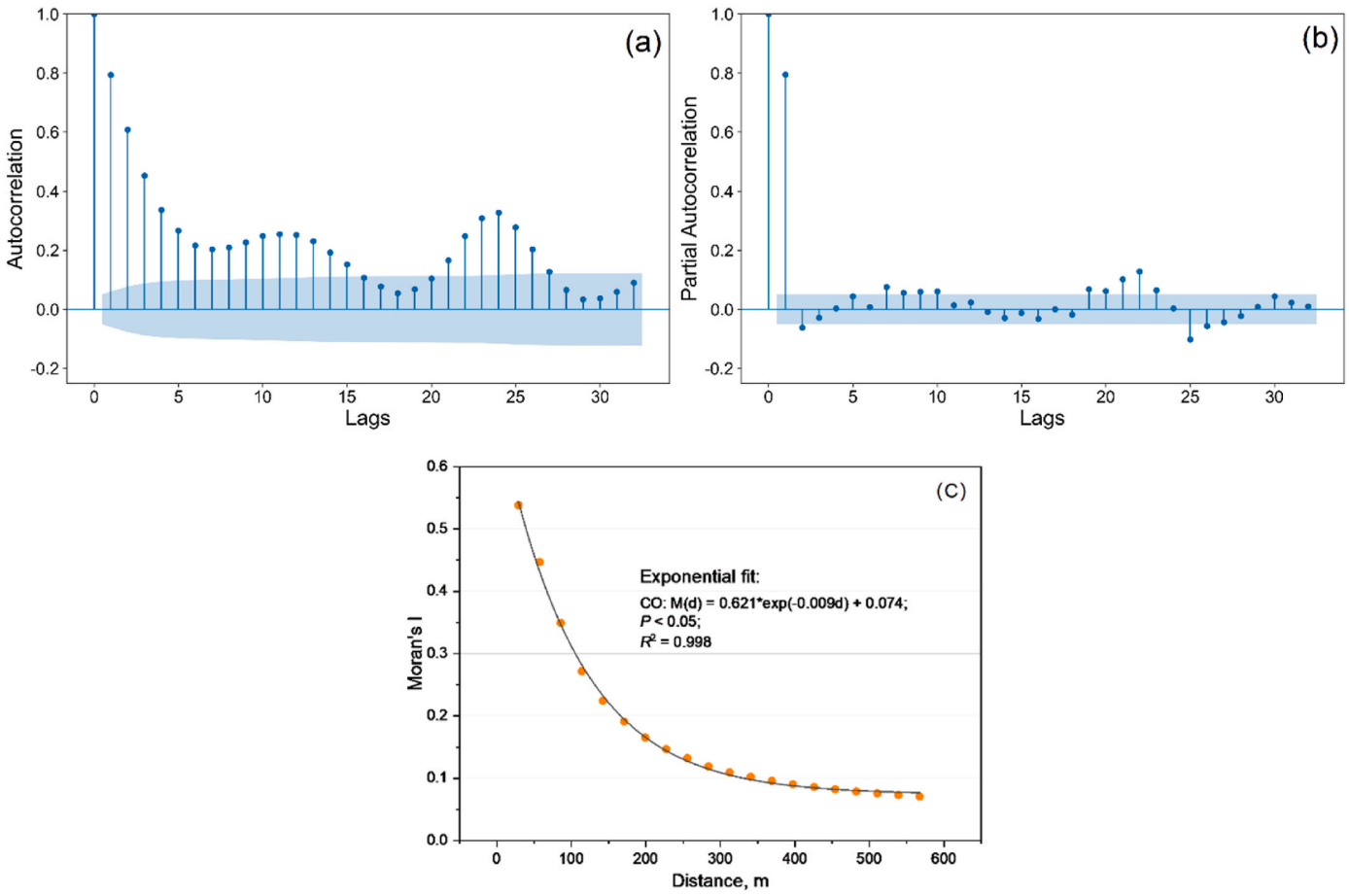


Fig. 4. Variation of the autocorrelation (a) and partial autocorrelation (b) coefficients of CO from mobile observations with lag values, with shaded areas representing the 95% confidence intervals. c depicts the change in incremental spatial autocorrelation coefficients for CO in relation to distance.

3.2. Model estimation performance

Fig. 5 shows a comparative analysis of the average CO concentration estimation from the CNN and CNN-attention models, with and without incorporating observational data, across 48 distinct scenarios. The results show that the inclusion of the attention mechanism enhances the model's accuracy in estimating mean CO concentrations, particularly in scenarios with high emissions and low wind speeds, resulting in a notable improvement in performance, with $R^2 = 0.972$ (Fig. 5B). This enhanced performance is primarily due to the attention mechanism's capacity to identify crucial features, particularly in areas with high emissions like motorways and arterial roads (Li et al., 2020). In addition, low wind speeds tend to hinder the dispersion of pollutants, which is especially problematic in densely built-up areas susceptible to pollutant accumulation (Fu et al., 2017). When observational data is integrated into the pre-trained CNN model, there is a slight increase in the R^2 value of 0.005 and a decrease in the residual sum of squares (RSS), indicating an enhanced estimation performance. Adding observational data to the CNN-attention model further refines its performance, leading to an improved $R^2 = 0.987$ (Fig. 5D). The augmented performance from incorporating observational data stems from two main factors: First, high-resolution observed data offers an accurate representation of air quality condition, enabling the model to adjust biases or inaccuracies inherent in theoretical or simulated inputs and align its estimates more closely with real-world observations. Second, urban air quality is influenced by a complex interplay of factors, including meteorological conditions, emission sources, topography, and building structures. Integrating observed CO concentration data enables the model to better identify critical features and their interactions, thus enhancing its overall estimation accuracy (Hu et al., 2022).

3.3. Model evaluation

To evaluate the significance of the estimation enhancements provided by models integrating observational data and attention mechanisms, we conducted a paired sample *t*-test on the estimation from three new scenarios. The analysis yield p-values less than 0.01 for each comparison, confirming that the improvements are statistically significant and robust. Further, we employ 10-fold cross-validation for training and validation to assess the adaptability and generalization of the observed data-constrained CNN and CNN-attention models. The trained models are then applied to estimate three new wind scenarios characterized by distinctly different conditions (22.5°, 2 m/s; 247.5°, 4.75 m/s; and 157.5°, 8 m/s). The results on this new dataset demonstrate that both models, when constrained by observational data, outperform their unconstrained counterparts, showing strong generalization in novel scenarios (Fig. 6). In the 4.75 m/s wind speed scenario, incorporating observed data and attention mechanisms improve estimation accuracy by 4.3 % and 1.8 %, respectively, resulting in a combined enhancement of 6.1 %. In the 8 m/s wind speed scenario, accuracy gains of 2.6 % and 1.9 % are noted for models with observed data and attention mechanisms, respectively, culminating in a total increase of 3.8 %. In contrast, the Sensor-CNN-attention model shows a slight decrease in estimation accuracy of 0.2 % compared to the CNN model alone (Table 1). These findings underscore that the assessment performance of the models varies across different wind scenarios, with a noticeable decrease in efficacy under low wind speed conditions.

The training duration for the CNN-attention model is significantly

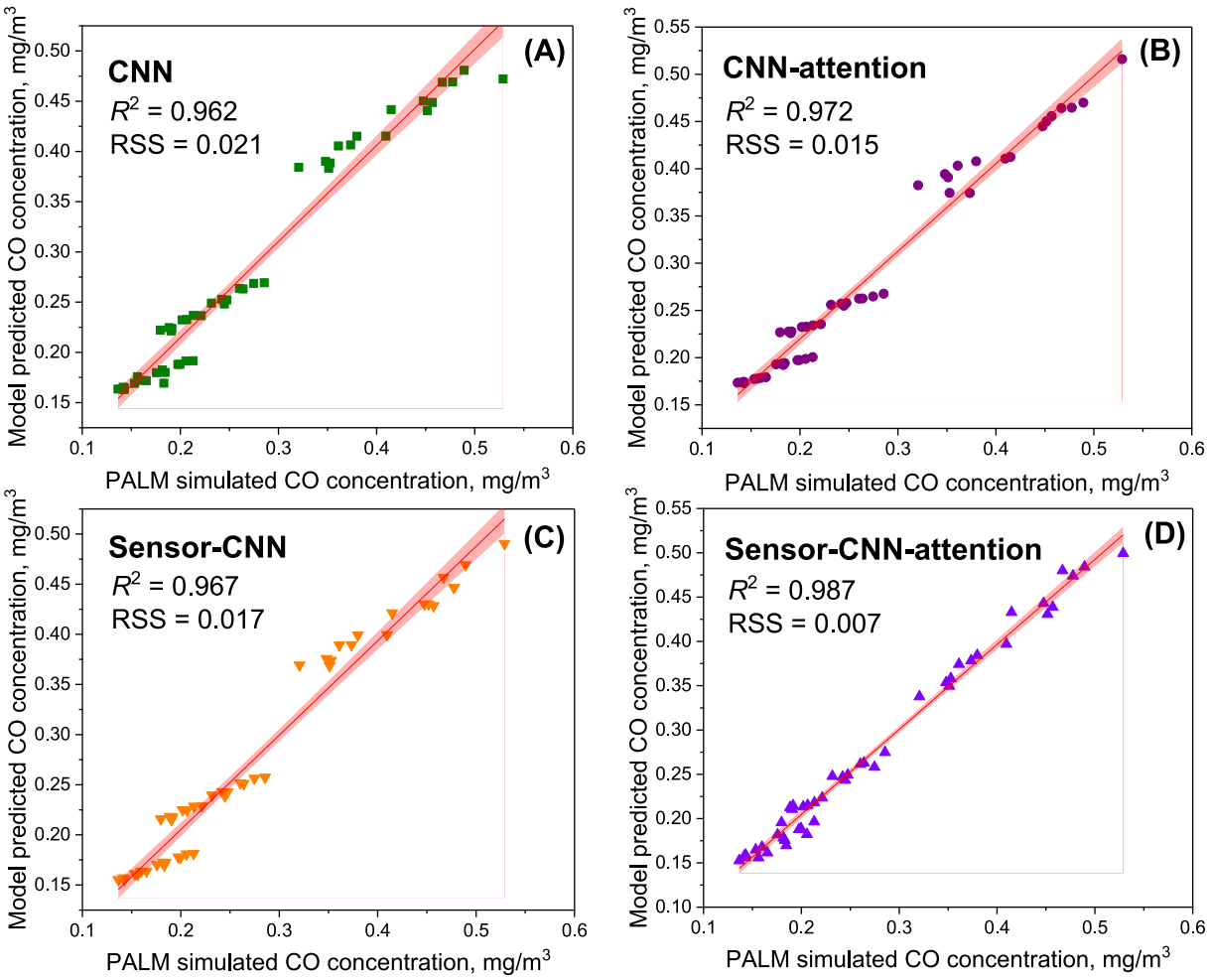


Fig. 5. Comparison of predicted average CO concentrations: CNN vs. CNN-attention models with and without observations constraints. "Sensor-CNN" indicates the CNN model enhanced with mobile observation data, while "Sensor-CNN-attention" refers to the integration of mobile observation data into the CNN-attention model. The convention applies hereinafter.

reduced to just 17.8 min, achieving a speed increase by a factor of 57 compared to the PALM simulation, and an estimated 98 % reduction in memory usage. This enhancement is attributed to the attention mechanism's parallel computation capabilities, which efficiently allocates resources to emphasize critical channel and spatial features while minimizing less significant ones. This strategic focus enhances the efficiency and precision of feature extraction within CNN models by effectively capturing interactions across different dimensions (Xie et al., 2022). To improve the accuracy of urban air pollutant assessment under varying wind conditions, several advancements are necessary. Firstly, expending the dataset with more labeled pollutant concentrations across diverse meteorological scenarios and validating with long-term observation data will bolster the model's generalizability (Liao et al., 2023). Secondly, incorporating diverse data sources, including meteorological, topographical, and traffic-related information, can enhance the model's responsiveness to environmental shifts through multimodal fusion. The current CNN model integrates meteorological data, but the absence of real-time meteorological field data curtails its ability to discern spatial meteorological features (LeCun et al., 2015). Lastly, combining traditional atmospheric dispersion and chemical transport models with deep learning techniques will enhance the model's estimation accuracy and enrich interpretative power, offering a more nuanced view of pollutant distribution both spatially and temporally.

3.4. Comparison of different ML models

This study benchmarks the estimation performance of a convolutional neural network (CNN) against traditional machine learning methods such as random forest (RF), XGBoost: Extreme Gradient Boosting (XGBoost), and linear regression (Table 2). The CNN outperforms RF, XGBoost, and linear regression with higher R^2 and lower RMSE and mean absolute error (MAE) values, underscoring its ability to capture complex nonlinear patterns. While the integration of an attention mechanism into the CNN marginally affects overall performance, it does enhance feature detection, generalization, and interpretability. Comparative analysis reveals that deep learning architectures, including CNNs, LSTMs, artificial neural networks (ANNs), and deep forests, deliver the most accurate estimation (Alimissis et al., 2018; Chen et al., 2023; Feizi et al., 2023). Ensemble methods such as extra tree (ET), RF and Light Gradient Boosting Machine (Light-GBM) also show robust performance, as indicated by high R^2 scores and low error metrics (Chen et al., 2023, 2024). Conversely, traditional machine learning models, such as multiple linear regression (MLR) and k-Nearest Neighbors (kNN), generally struggle with nonlinear complexities compared to deep and ensemble learning methods (Alimissis et al., 2018; Feizi et al., 2023). Cerezuela-Escudero et al. (2023) demonstrated that feed-forward neural networks (FFNNs) outperform traditional geostatistical methods in spatially estimating CO concentrations and also marginally surpass vector machines (SVMs) in accuracy. However, the precision of these models is limited by meteorological conditions, such as thermal

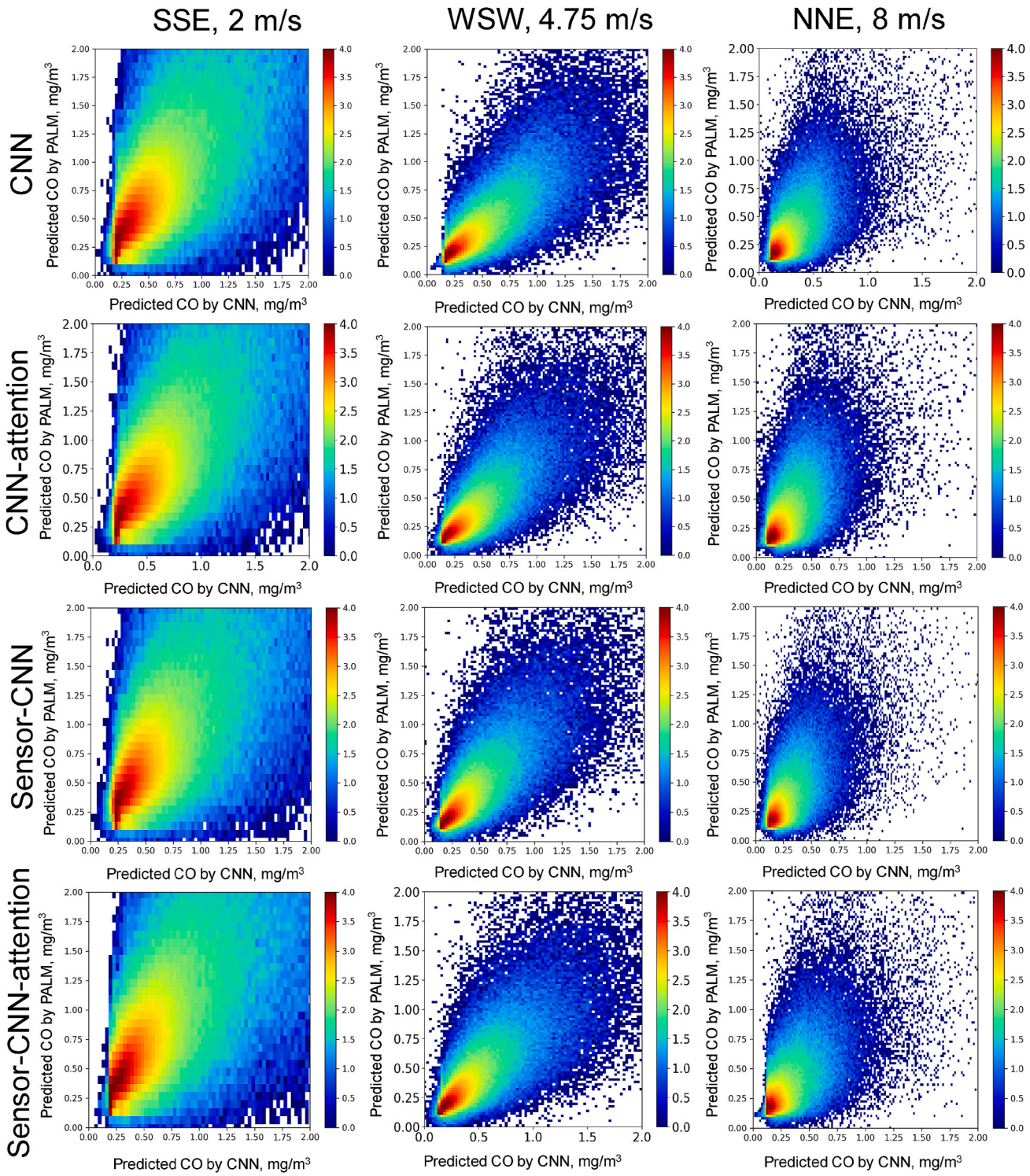


Fig. 6. Comparison of validation results for CNN vs. CNN-attention models with and without integration observation data.

inversions and low wind speeds, which induce rapid spatial variations in CO levels. To address this limitation, our study proposes the integration of CNN model with large eddy simulation model, which enables a more effectively capture of the microscale dynamics and fluctuations in pollutant concentrations, thus enhancing the model's estimation accuracy.

3.5. *Model interpretability*

To further explore the attentional layer's role in explaining the model's mechanism, we examine the contribution of attention weights across different features. Fig. 7 shows the channel attention weight contributions for 64 filters in scenarios with and without observational constraints, at a wind speed of 4.75 m/s and a WSW wind direction. Fig. 7(1) shows that the attention mechanism de-prioritizes emissions and topography in response to lower observed CO concentrations, with a

Table 1
Evaluation metrics for CNN and CNN-attention models with and without observation constraints.

Models	Indicators	SSE, 2 m/s	WSW, 4.75 m/s	NNE, 8 m/s
CNN	R^2	0.806	0.830	0.672
	RMSE, mg/m^3	0.529	0.163	0.156
CNN-attention	R^2	0.801	0.835	0.684
	RMSE, mg/m^3	0.561	0.160	0.152
Sensor-CNN	R^2	0.795	0.837	0.687
	RMSE, mg/m^3	0.588	0.156	0.153
Sensor-CNN-attention	R^2	0.802	0.842	0.693
	RMSE, mg/m^3	0.530	0.153	0.150

Table 2
Comparison of the results of different machine learning models for CO assessment.

Models	R^2	RMSE, mg/m^3	MAE, mg/m^3	References
CNN	0.67-0.84	0.20-0.56	0.18	This research
CNN-attention	0.68-0.84	0.15-0.17	0.17	
ET	0.68	0.40	0.22	Chen et al. (2023)
RF	0.63	0.42	0.23	
XGBoost	0.61	0.45	0.25	
Light-GBM	0.69	0.40	0.22	
Deep-Forest	0.75	0.35	0.17	
SW-K	0.53	0.21	–	Lu et al. (2021)
PLS-K	0.38	0.24	–	
ML-K	0.71	0.16	–	
RF	0.41	0.12	0.01	Miao et al. (2024)
GBDT	0.72	0.10	–	Liu et al. (2021)
NN	0.70	0.31	0.18	Rakholia et al. (2023)
MLPR	0.62	–	–	Fabregat et al. (2022)
FFNNs	0.51-0.83	0.26-0.81	0.14-0.62	Alimissis et al. (2018)
MLR	0.50-0.82	0.28-1.06	0.14-0.90	
ANN	0.77	0.06	0.05	Nourani et al. (2021)
SVM	–	0.03	0.02	Xu et al. (2017)
ANFIS	0.55-0.79	0.13-0.6	–	Prasad et al. (2016)
ANN	–	0.28-0.34	0.18-0.23	Feizi et al. (2023)
kNN	–	0.26-0.53	0.17-0.33	
SVR	–	0.22-0.39	0.14-0.26	
LSTM	–	0.19-0.43	0.13-0.33	
FFNNs	0.57	0.16	0.09	Cerezuela-Escudero et al. (2023)
SVM	0.49	0.18	0.09	
ET	0.68	0.18	0.11	Chen et al. (2024)

Note: ET: extra tree; RF: Random Forest; XGBoost: Extreme Gradient Boosting; Light-GBM: Light Gradient Boosting Machine; SW-K: Stepwise Regression + kriging; PLS-K: Partial Least Squares + kriging; ML-K: Machine Learning + kriging; GBDT: Gradient Boosting Decision Tree; NN: Neural Network; MLPR: Multilayer Perceptron Regressor; FFNN: Feed-forward Neural Network; MLR: Multiple Linear Regression; ANN: Artificial Neural Network; SVM: Support Vector Machine; ANFIS: Adaptive neuro-fuzzy inference system; kNN: k-Nearest Neighbors; SVR: Support Vector Regression. "-" represents no data.

significant reduction in the weights assigned to emissions. This adjustment is primarily attributed to the fact that observed CO concentrations are influenced by both emissions and topography, prompting the attention mechanism to modulate the weight contributions for each, particularly in areas with secondary roads. Conversely, the attention mechanism increases the focus on emissions in areas with higher observed concentrations, such as major roads and within tunnels (Fig. 7 (2)). Additionally, the model is capable of dynamically adjusting the weights for topography and emissions, even when special terrain features do not align with high observed concentrations (Fig. 7(D4′)). Nevertheless, the model does have its shortcomings. In regions where observational data are sparse, the attention mechanism may adjust weights similarly, potentially reducing estimation accuracy (Fig. 7(5)).

Fig. 8 demonstrates the impact of the spatial attention layer on highlighting observed features. This layer precisely highlights CO concentration signatures along major roads (Fig. 8A), and enhances attention weights for distinctive areas and road structures, such as the Nanjing railway station and Hugu Road (Fig. 8B and C) (Zhang et al., 2021). The model assigns increased weights in the high-traffic Xinjiekou district, where dense construction contributes to pollutant accumulation,. Additionally, the model allocates lower attention to secondary roads and neighborhood streets, reflecting the spatial and temporal distribution patterns of CO concentrations (Fig. 8D).

The study explores the attention mechanism's role in detecting spatial features related to topography and emissions by comparing the distribution of spatial attention weights in the CNN-attention model, both with and without the constraint of observational data (Fig. 9). The inclusion of observational data enhances the model's ability to recognize features such as ground elevation and building heights. The model notably assigns higher weights to the densely built Xinjiekou area, consistent with the higher CO concentrations observed there (Fig. 9E). It also recognizes and assigns increased weight to the Qingliangshan area due to its topography that hinders pollutant dispersion (Fig. 9I). After incorporating observational data, the topography weights' contributions align more closely with actual topography features (Fig. S1), highlighting the constraints imposed by the observational data on the model (Zhang et al., 2021). However, the model still exhibits limitations in identifying complex topographies, such as the area around Xuanwu Lake, which remains inadequately characterized.

Fig. 9C reveals a marked difference in emission attention weights, with the road's left side initially exhibiting a higher weight. Yet, following a rotation, the right side presents increased weight. This shift corresponds with a transition to westerly winds, indicating that buildings on the right of the road are taller than those on the left, thereby confirming the influence of building height on pollutant dispersion as estimated by the PALM model (Zhang et al., 2021). Moreover, the model assigns increased attention weights to distinct plume features over the northwestern of Xuanwu Lake (Fig. 9G and H) and specific urban structures such as overpasses, intersections, and Huju Road (Fig. 9J and K). It also assigns higher emission weights to the Xuanwu Lake area, positioned upwind of Central Road, where the westerly winds exacerbate pollution levels, consistent with PALM simulation results (Fig. 9O). To further enhance the model's estimation accuracy and generalizability, future research should focus on improving the recognition of diverse land-use elements, including vegetation cover, and key points of interest such as schools and hospitals. These advancements will offer a more accurate data basis for assessing the impact of pollution on population exposure risks (Wen et al., 2022).

3.6. Application to other cities

Given the absence of local mobile observation data and to ensure the accuracy and transferability of our models across diverse urban environments, we employ a CNN-attention model to estimate high-resolution air quality in the urban area of Dongguan (Fig. S4). For efficiency in time and computational resources, we select PALM simulation results for six different wind directions (NW, SE, E, N, S, and W) at a wind speed of 3 m/s in Dongguan. The model inputs include both topography and emission data, with the PALM outputs forming the training dataset (Fig. S5). We conduct training and evaluation using 10-fold cross-validation. Fig. S6 offers a comparative analysis of the

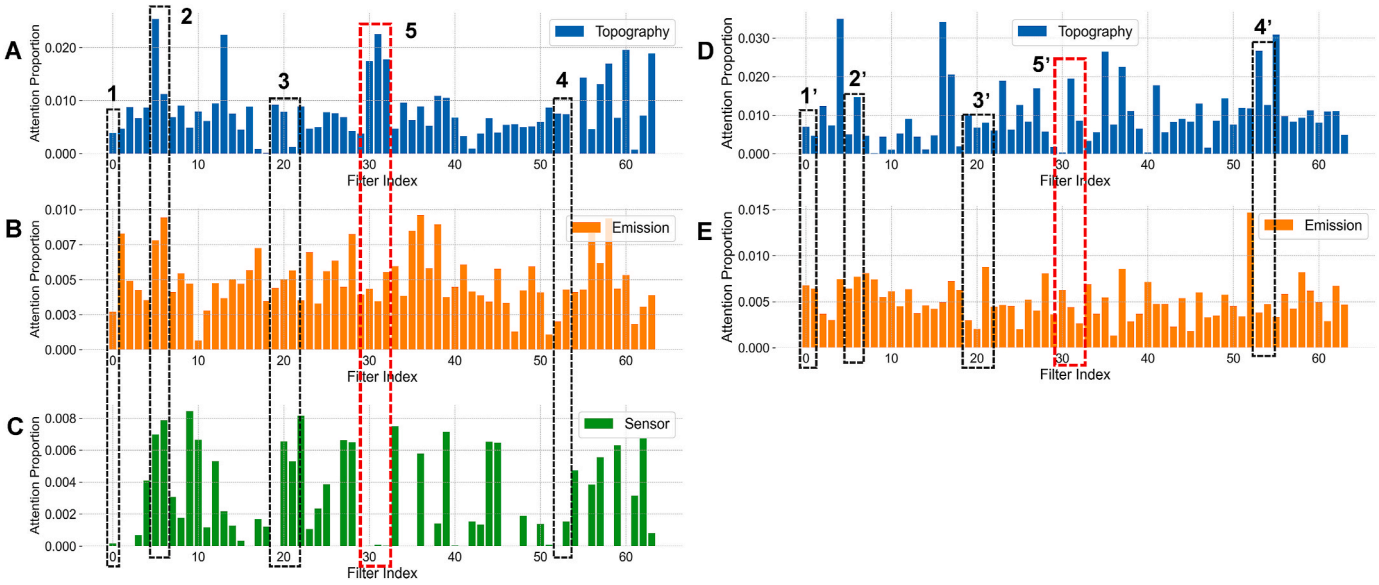


Fig. 7. Contribution proportions of attention weights for different feature channels in Nanjing with and without observation constraints: A, B, and C represent with the observation data, while D and E represent without the observation data.

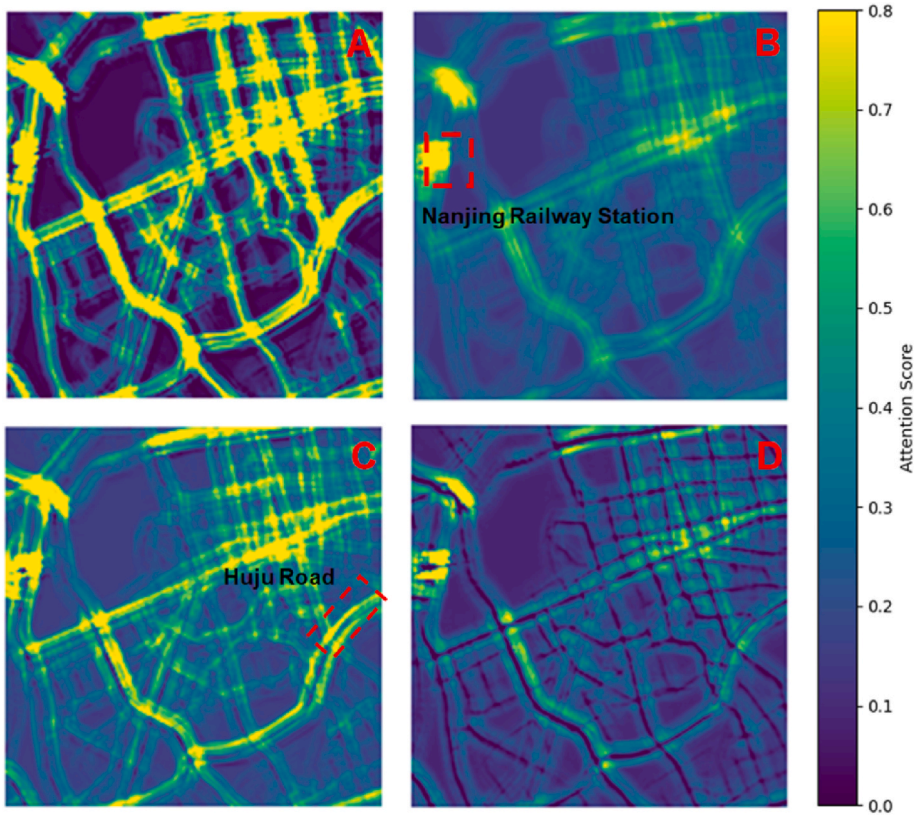


Fig. 8. Contribution of attention layer weights to observational feature space in CNN-attention model.

estimation performance and error between the CNN and CNN-attention models for the six scenarios in Dongguan. Results show that the attention mechanism significantly improves the model's CO estimation accuracy across varying conditions, with correlation coefficients (R) from 0.684 to 0.807 and RMSE ranging from 0.243 mg/m^3 to 0.288 mg/m^3. Notably, the SE and NW wind scenarios achieve higher correlation values (0.807 and 0.804, respectively) compared to other wind conditions at a wind speed of 3 m/s. The incorporation of the attention mechanism boosts the CNN-attention model's estimation precision by 10.9 %–25.7 %, indicating an enhancement in its generalization ability. Although the performance gain on the Nanjing city evaluation dataset is modest, likely because the model captures certain general features, the improved generalization becomes more apparent when applied to a new environment like Dongguan. This adaptability leads to a more significant performance boost when confronted with novel data distributions.

Fig. S7 depicts the proportion of filters in the model that accentuate

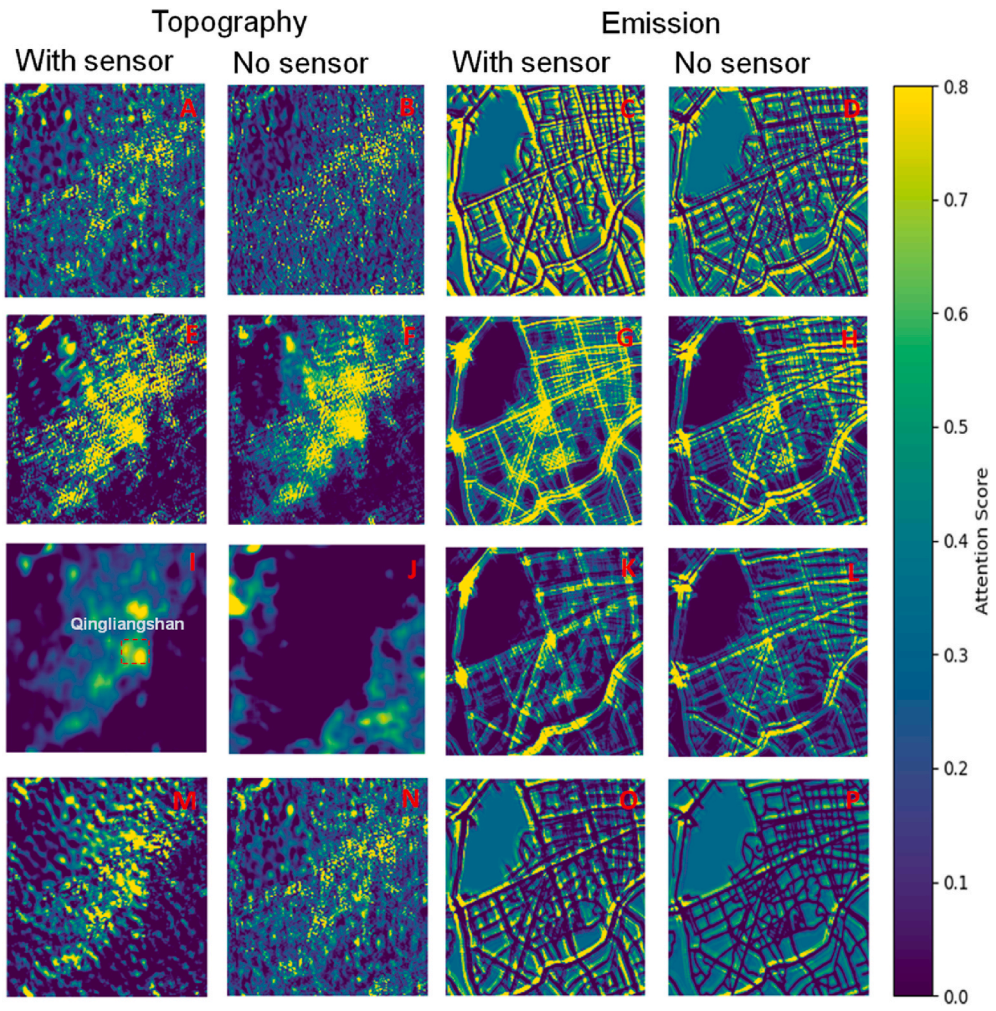


Fig. 9. Spatial attention layer weight contributions of topography and emissions in the CNN-attention model with and without observation data.

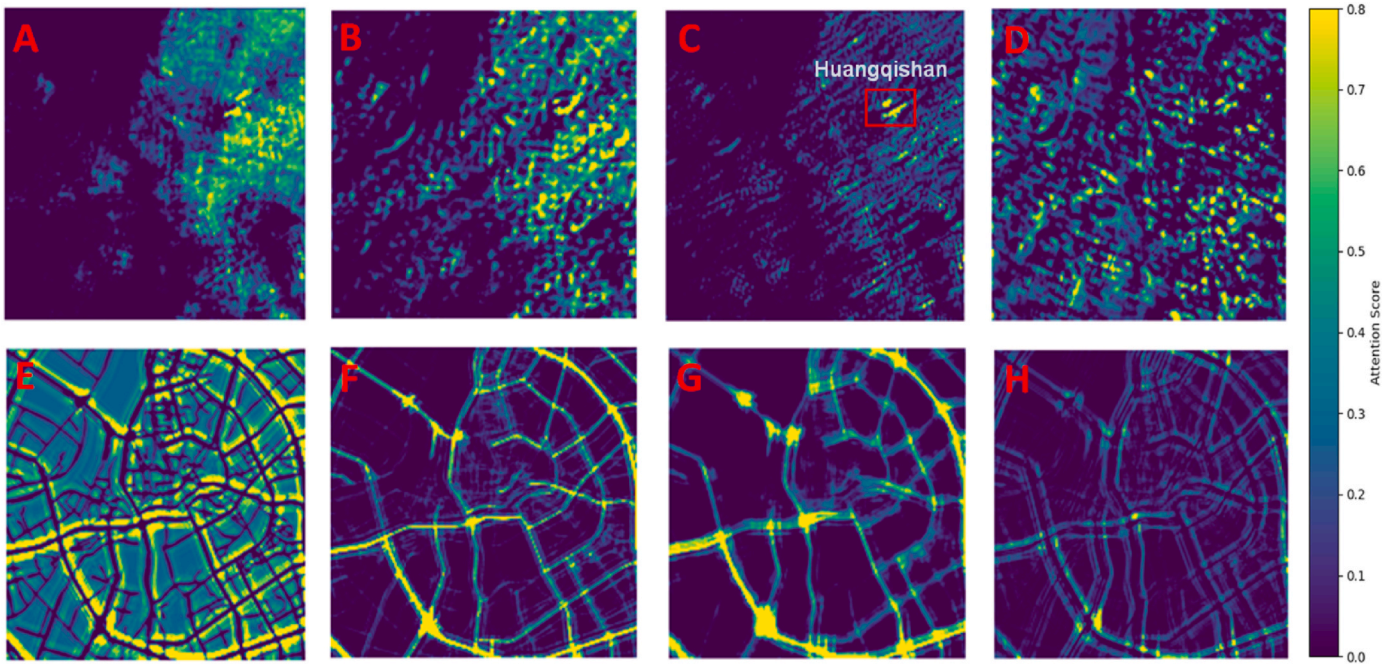


Fig. 10. Spatial attention weight contribution of different features after integrating the attention mechanism in Dongguan city. A, B, C, and D represent terrain weight contributions; while E, F, G, and H represent emission weight contributions.

topography and emission features after integrating the attention mechanism. This integration enables the model to dynamically shift its focus across different channels. The model clearly prioritizes terrain information through the channel attention mechanism. as terrain features are highlighted in more channels than emission features—17 versus 11—with an attention score above 0.01. This emphasis largely arises from the impact of topographical variations, particularly building heights, in generating atmospheric turbulence that affects pollutant concentration distribution (Fu et al., 2017).

Fig. 10C shows that the spatial attention layer preferentially focuses on areas with eleveted baseline topography and notable locations, such as Huangqishan park. Additionally, the model identifies building height as a key factor influencing CO concentration (Zhang et al., 2021). In Dongguan, the distribution of building height displays a scattered pattern of high-rise structures (Fig. S5). The model detects significant height variations between buildings, which can intensify atmospheric turbulence and lead to pollutant accumulation, thereby marking these areas accordingly (Fig. 10D) (Zhang et al., 2021). Furthermore, the model differentiates between emissions from major and minor roads, assigning higher attention scores to the former and noting the impact of varying building heights adjacent to these roads (Fig. 10E and F). It also closely examines road intersections, a feature identification trend observed in Nanjing as well (Fig. 10G). This selective focus highlights the model's sophisticated approach to feature prioritization, contributing to its enhanced estimation accuracy.

4. Conclusions

To enhance the estimation capabilities of our model for urban air quality, we integrate an attention mechanism into the CNN model and use mobile observation data as constraints, achieving high-resolution and rapid, accurate estimation of urban pollutants. Our findings indicate that the inclusion of observational data improved the CNN model's accuracy by margins ranging from 1.8 % to 4.3 %. Moreover, the synergy of the attention mechanism with the CNN further increases estimation accuracy by 3.8 %–6.1 %. The attention mechanism adeptly emphasizes various features, thereby enhancing the model's proficiency in pinpointing critical attributes, such as major emission sources including highways and arterial roads, along with distinctive topography features. This refinement not only boosts the model's estimation accuracy but also substantially increases its interpretability. When applied to air quality estimation in various urban environments, the enhanced model demonstrates improvements in accuracy between 10.9 % and 25.7 %, underscoring its robustness and transferability.

Despite its effectiveness, the CNN-attention model face certain limitations. One such limitation is its reliance on mobile observations, which can introduce uncertainty due to the sparse and uneven coverage provided by taxis. Additionally, integrating results from the PALM into the CNN-attention framework adds its own layer of inherent uncertainty. The complexity introduced by the attention mechanism also raises the risk of overfitting. To address these issues, future research should focus on enhancing the model's capacity to assimilate dynamic observational data more effectively. This could involve incorporating long-term observational datasets, diversifying simulation inputs from various urban environments, and applying advanced regularization techniques to reduce uncertainty and prevent overfitting. Moreover, enriching the model with a LSTM (Seng et al., 2021) and an atmospheric chemical transport model (CTM) could more effectively capture the nonlinear dynamics of atmospheric chemistry and physics. Such enhancements could improve the precision of estimation and computational efficiency (Xing et al., 2020).

CRediT authorship contribution statement

Shibao Wang: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Yanxu Zhang:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to the High Performance Computing Center (HPCC) of Nanjing University for doing the numerical calculations in this paper on its blade cluster system.

Financial support. This research has been supported by the National Key R&D Program of China (grant no. 2019YFA0606803), National Natural Science Foundation of China (grant no. 71974092), and Collaborative Innovation Center of Climate Change of Jiangsu Province.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2024.120921.

Data availability

All study data are included in the article and/or SI Appendix. Code and data will be made available on request.

References

Aggarwal, A., Toshniwal, D., 2021. A hybrid deep learning framework for urban air quality forecasting. J. Clean. Prod. 329, 129660. https://doi.org/10.1016/j.jclepro.2021.129660.

Alimissis, A., Philippopoulos, K., Tzanis, C.G., Deligiorgi, D., 2018. Spatial estimation of urban air pollution with the use of artificial neural network models. Atmos. Environ. 191, 205–213. https://doi.org/10.1016/j.atmosenv.2018.07.058.

Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., Hamburg, S.P., 2017. High resolution air pollution mapping with Google street view cars: exploiting big data. Environ. Sci. Technol. 51, 6999–7008. https://doi.org/10.1021/acs.est.7b00891.

Bahdanau, D., Cho, K., Bengio, Y., 2016. Neural machine translation by jointly learning to align and translate. arXiv:1409.0473.

Cerezuela-Escudero, E., Montes-Sanchez, J.M., Dominguez-Morales, J.P., Duran-Lopez, L., Jimenez-Moreno, G., 2023. A systematic comparison of different machine learning models for the spatial estimation of air pollution. Appl. Intell. 53, 29604–29619. https://doi.org/10.1007/s10489-023-05109-y.

Chen, B., Hu, J., Song, Z., Zhou, X., Zhao, L., Wang, Y., Chen, R., Ren, Y., 2023. Exploring high-resolution near-surface CO concentrations based on Himawari-8 top-of-atmosphere radiation data: assessing the distribution of city-level CO hotspots in China. Atmos. Environ. 312, 120021. https://doi.org/10.1016/j.atmosenv.2023.120021.

Chen, B., Hu, J., Wang, Y., 2024. Synergistic observation of FY-4A&4B to estimate CO concentration in China: combining interpretable machine learning to reveal the influencing mechanisms of CO variations. npj Clim. Atmos. Sci. 7, 9. https://doi.org/10.1038/s41612-023-00559-0.

Crocchianti, S., Sarto, S.D., Ranalli, M.G., Moroni, B., Castellini, S., Petroselli, C., Cappelletti, D., 2021. Spatiotemporal correlation of urban pollutants by long-term measurements on a mobile observation platform. Environ. Pollut. 268, 115645. https://doi.org/10.1016/j.envpol.2020.115645.

Deshmukh, P., Kimbrough, S., Krabbe, S., Logan, R., Isakov, V., Baldauf, R., 2020. Identifying air pollution source impacts in urban communities using mobile monitoring. Sci. Total Environ. 715, 136979. https://doi.org/10.1016/j.scitotenv.2020.136979.

Ding, A., Fu, C., Yang, X., Sun, J., Zheng, L., Xie, Y., 2013. Ozone and fine particle in the western Yangtze River Delta: an overview of 1 yr data at the SORPES station. Atmos. Chem. Phys. 13, 5813–5830. https://doi.org/10.5194/acp-13-5813-2013.

Fabregat, A., Vernet, A., Vernet, M., Vázquez, L., Ferré, J.A., 2022. Using Machine Learning to estimate the impact of different modes of transport and traffic restriction strategies on urban air quality. Urban Clim. 45, 101284. https://doi.org/10.1016/j.uclim.2022.101284.

Feizi, H., Sattari, M.T., Prasad, R., Apaydin, H., 2023. Comparative analysis of deep and machine learning approaches for daily carbon monoxide pollutant concentration estimation. Int. J. Environ. Sci. Technol. 20, 1753–1768. https://doi.org/10.1007/s13762-022-04702-x.

Fu, X., Liu, J., Ban-Weiss, G.A., Zhang, J., Huang, X., Ouyang, B., Popoola, O., Tao, S., 2017. Effects of canyon geometry on the distribution of traffic-related air pollution

in a large urban area: implications of a multi-canyon air pollution dispersion model. Atmos. Environ. 165, 111–121. https://doi.org/10.1016/j.atmosenv.2017.06.031.
Gately, C.K., Hutyra, L.R., Peterson, S., Wing, I.S., 2017. Urban emissions hotspots: quantifying vehicle congestion and air pollution using mobile phone GPS data. Environ. Pollut. 229, 496–504. https://doi.org/10.1016/j.envpol.2017.05.091.
Gu, J., Yang, B., Brauer, M., Zhang, K.M., 2021. Enhancing the evaluation and interpretability of data-driven air quality models. Atmos. Environ. 246, 118125. https://doi.org/10.1016/j.atmosenv.2020.118125.
Guo, R., Zhang, Q., Yu, X., Qi, Y., Zhao, B., 2023. A deep spatio-temporal learning network for continuous citywide air quality forecast based on dense monitoring data. J. Clean. Prod. 414, 137568. https://doi.org/10.1016/j.jclepro.2023.137568.
Hu, X., Liu, T., Hao, X., Lin, C., 2022. Attention-based Conv-LSTM and Bi-LSTM networks for large-scale traffic speed prediction. J. Supercomput. 78, 12686–12709. https://doi.org/10.1007/s11227-022-04386-7.
Jeanjean, A.P.R., Hinchliffe, G., McMullan, W.A., Monks, P.S., Leigh, R.J., 2015. A CFD study on the effectiveness of trees to disperse road traffic emissions at a city scale. Atmos. Environ. 120, 1–14. https://doi.org/10.1016/j.atmosenv.2015.08.003.
Kaivonen, S., Ngai, E., 2020. Real-time air pollution monitoring with sensors on city bus, Digital Commun. Networks 6, 23–30. https://doi.org/10.1016/j.dcan.2019.03.003.
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 $PM_{2.5}$ forecasting. J. Clean. Prod. 261, 121285. https://doi.org/10.1016/j.jclepro.2020.121285.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444. https://doi.org/10.1038/nature14539.
Letzel, M.O., Krane, M., Raasch, S., 2008. High resolution urban large-eddy simulation studies from street canyon to neighborhood scale. Atmos. Environ. 42 (38), 8770–8784. https://doi.org/10.1016/j.atmosenv.2008.08.001.
Li, S., Xie, G., Ren, J., Guo, L., Yang, Y., Xu, X., 2020. Urban $PM_{2.5}$ concentration prediction via attention-based CNN–LSTM. Appl. Sci. 10 (6), 1953. https://doi.org/10.3390/app10061953.
Li, Z., Fung, J.C.H., Lau, A.K.H., 2018. High spatiotemporal characterization of on-road $PM_{2.5}$ concentrations in high-density urban areas using mobile monitoring. Build. Environ. 143, 196–205. https://doi.org/10.1016/j.buildenv.2018.07.014.
Liao, H., Yuan, L., Wu, M., Chen, H., 2023. Air quality prediction by integrating mechanism model and machine learning model. Sci. Total Environ. 899, 165646. https://doi.org/10.1016/j.scitotenv.2023.165646.
Lim, C.C., Kim, H., Vilcassim, M.J.R., Thurston, G.D., Gordon, T., Chen, L.C., Lee, K., Heimbinder, M., Kim, S., 2019. Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea. Environ. Int. 131, 105022. https://doi.org/10.1016/j.envint.2019.105022.
Liu, H.Z., Liu, J.F., Liu, Y., Ouyang, B., Xiang, S.L., Yi, K., Tao, S., 2020. Analysis of wintertime O_3 variability using a random forest model and high-frequency observations in Zhangjiakou—an area with background pollution level of the North China Plain. Environ. Pollut. 262, 14191. https://doi.org/10.1016/j.envpol.2020.114191.
Liu, M., Chen, H., Wei, D., Wu, Y., Li, C., 2021. Nonlinear relationship between urban form and street-level $PM_{2.5}$ and CO based on mobile measurements and gradient boosting decision tree models. Build. Environ. 205, 108265. https://doi.org/10.1016/j.buildenv.2021.108265.
Lu, T., Marshall, J.D., Zhang, W., Hystad, P., Kim, S.Y., Bechle, M.J., Demuzere, M., Hankey, S., 2021. National empirical models of air pollution using microscale measures of the urban environment. Environ. Sci. Technol. 55 (22), 15519–15530. https://doi.org/10.1021/acs.est.1c04047.
Maronga, B., Gryschka, M., Heinze, R., Hoffmann, F., Kanani-Sühring, F., Keck, M., Ketelsen, K., Letzel, M.O., Sühring, M., Raasch, S., 2015. The Parallelized Large-Eddy Simulation Model (PALM) version 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future perspectives. Geosci. Model Dev. (GMD) 8, 2515–2551. https://doi.org/10.5194/gmd-8-2515-2015.
Miao, C., Peng, Z., Cui, A., He, X., Chen, F., Lu, K., Jia, G., Yu, S., Chen, W., 2024. Quantifying and predicting air quality on different road types in urban environments using mobile monitoring and automated machine learning. Atmos. Pollut. Res. 15 (3), 102015. https://doi.org/10.1016/j.apr.2023.102015.
Moursi, A.S.A., El-Fishawy, N., Djahel, S., Shouman, M.A., 2022. Enhancing PM2.5 prediction using NARX-Based combined CNN and LSTM hybrid model. Sensors 22 (12), 4418. https://doi.org/10.3390/s22124418.
Nourani, V., Karimzadeh, H., Baghanam, A.H., 2021. Forecasting CO pollutant concentration of Tabriz city air using artificial neural network and adaptive neuro-fuzzy inference system and its impact on sustainable development of urban. Environ. Earth Sci. 80, 136. https://doi.org/10.1007/s12665-021-09423-x.
Peters, J., den Bossche, J., Reggente, M., Van Poppel, M., Baets, B.D., Theunis, J., 2014. Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. Atmos. Environ. 92, 31–43. https://doi.org/10.1016/j.atmosenv.2014.03.039.
Prasad, K., Gorai, A.K., Goyal, P., 2016. Development of ANFIS models for air quality forecasting and input optimization for reducing the computational cost and time. Atmos. Environ. 128, 246–262. https://doi.org/10.1016/j.atmosenv.2016.01.007.
Rakholia, R., Le, Q., Ho, B.Q., Vu, K., Carbajo, R.S., 2023. Multi-output machine learning model for regional air pollution forecasting in Ho Chi Minh City, Vietnam. Environ. Int. 173, 107848. https://doi.org/10.1016/j.envint.2023.107848.
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566, 195–204. https://doi.org/10.1038/s41586-019-0912-1.
Rentschler, J., Leonova, N., 2023. Global air pollution exposure and poverty. Nat. Commun. 14, 4432. https://doi.org/10.1038/s41467-023-39797-4.
Resler, J., Eben, K., Geletič, J., Krč, P., Rosecký, M., Sühring, M., Belda, M., Fuka, V., Halenka, T., Huszár, P., Karlický, J., Benešová, N., Ďoubalová, J., Honzáková, K., Keder, J., Nápravníková, Š., Vlček, O., 2021. Validation of the PALM model system 6.0 in a real urban environment: a case study in Dejvice, Prague, the Czech Republic. Geosci. Model Dev. (GMD) 14, 4797–4842. https://doi.org/10.5194/gmd-14-4797-2021.
Seng, D., Zhang, Q., Zhang, X., Chen, G., Chen, X., 2021. Spatiotemporal prediction of air quality based on LSTM neural network. Alex. Eng. J. 60 (2), 2021–2032. https://doi.org/10.1016/j.aej.2020.12.009.
Shiraiwa, M., Ueda, K., Pozzer, A., et al., 2017. Aerosol health effects from molecular to global scales. Environ. Sci. Technol. 51 (23), 13545–13567. https://doi.org/10.1021/acs.est.7b04417.
Sun, J., Lenschow, D.H., LeMone, M.A., Mahrt, L., 2016. The role of large-coherent-eddy transport in the atmospheric surface layer based on CASES-99 observations. Boundary-Layer Meteorol. 160 (1), 83–111. https://doi.org/10.1007/s10546-016-0134-0.
Tripathi, S., Mandal, D., Chakraborty, A., 2023. Investigation of changes in atmospheric pollutants due to the cessation of anthropogenic activities: spatial heterogeneity and complex atmospheric chemistry. Aerosol Sci. Eng. 7, 237–250. https://doi.org/10.1007/s41810-023-00175-8.
Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J., De Baets, B., 2015. Mobile monitoring for mapping spatial variation in urban air quality: development and validation of a methodology based on an extensive dataset. Atmos. Environ. 105, 148–161. https://doi.org/10.1016/j.atmosenv.2015.01.017.
Wang, H., Ngan, K., 2021. Effects of inhomogeneous ground-level pollutant sources under different wind directions. Environ. Pollut. 289, 117903. https://doi.org/10.1016/j.envpol.2021.117903.
Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q., 2020. ECA-net: Efficient Channel Attention for Deep Convolutional Neural Networks. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, pp. 11531–11539. https://doi.org/10.1109/CVPR42600.2020.01155.
Wang, S., Ma, Y., Wang, Z., Wang, L., Chi, X., Ding, A., Yao, M., Li, Y., Li, Q., Wu, M., Zhang, L., Xiao, Y., Zhang, Y., 2021. Mobile monitoring of urban air quality at high spatial resolution by low-cost sensors: impacts of COVID-19 pandemic lockdown. Atmos. Chem. Phys. 21, 7199–7215. https://doi.org/10.5194/acp-21-7199-2021.
Wang, S., McGibbon, J., Zhang, Y., 2024. Predicting high-resolution air quality using machine learning: integration of large eddy simulation and urban morphology data. Environ. Pollut. 344, 123371. https://doi.org/10.1016/j.envpol.2024.123371.
Wang, Z., Wu, F., Yang, Y., 2023. Air pollution measurement based on hybrid convolutional neural network with spatial-and-channel attention mechanism. Expert Syst. Appl. 233, 120921. https://doi.org/10.1016/j.eswa.2023.120921.
Wen, Y., Wu, R., Zhou, Z., Zhang, S., Yang, S., Wallington, T.J., Shen, W., Tan, Q., Deng, Y., Wu, Y., 2022. A data-driven method of traffic emissions mapping with land use random forest models. Appl. Energy 305, 117916. https://doi.org/10.1016/j.apenergy.2021.117916.
WHO, 2016. Global Urban Ambient Air Pollution Database (Update 2022). World Health Organization. https://www.who.int/airpollution/data/cities-2016/en/. (Accessed 22 February 2022).
Wolf, T., Pettersson, L.H., Esau, I., 2021. Dispersion of particulate matter (PM2.5) from wood combustion for residential heating: optimization of mitigation actions based on large-eddy simulations. Atmos. Chem. Phys. 21, 12463–12477. https://doi.org/10.5194/acp-21-12463-2021.
Xie, Y., Chen, H., Ma, Y., Xu, Y., 2022. Automated design of CNN architecture based on efficient evolutionary search. Neurocomputing 491, 160–171. https://doi.org/10.1016/j.neucom.2022.03.046.
Xing, J., Zheng, S., Ding, D., Kelly, J.T., Wang, S., Li, S., Qin, T., Ma, M., Dong, Z., Jang, C., Zhu, Y., Zheng, H., Ren, L., Liu, T., Hao, J., 2020. Deep learning for prediction of the air quality response to emission changes. Environ. Sci. Technol. 54 (14), 8589–8600. https://doi.org/10.1021/acs.est.0c02923.
Xu, J., Zhang, M., Ganji, A., Mallinen, K., Wang, A., Lloyd, M., Venuta, A., Simon, L., Kang, J., Gong, J., Zamel, Y., Weichenthal, S., Hatzopoulou, M., 2022. Prediction of short-term ultrafine particle exposures using real-time street-level images paired with air quality measurements. Environ. Sci. Technol. 56 (18), 12886–12897. https://doi.org/10.1021/acs.est.2c03193.
Xu, Z., Huang, X., Nie, W., Chi, X., Xu, Z., Zheng, L., Sun, P., Ding, A., 2017. Influence of synoptic condition and holiday effects on VOCs and ozone production in the Yangtze River Delta region, China. Atmos. Environ. 168, 112–124. https://doi.org/10.1016/j.atmosenv.2017.08.035.
Xu, D., Zhou, D., Wang, Y., Meng, X., Gu, Z., Yang, Y., 2021. Temporal and spatial heterogeneity research of urban anthropogenic heat emissions based on multi-source spatial big data fusion for Xi'an, China. Energy Build 240, 110884. https://doi.org/10.1016/j.enbuild.2021.110884.
Yan, J., Zhao, B., Zhang, S., Lin, L., Tang, J., 2021. Observation analysis and application evaluation of wind profile radar to diagnosing the boundary layer of landing typhoon. J. Appl. Meteor. Sci. 32 (3), 332–346. https://doi.org/10.11898/1001-7313.20210306.
Zhang, Y., Ye, X., Wang, S., He, X., Dong, L., Zhang, N., Wang, H., Wang, Z., Ma, Y., Wang, L., Chi, X., Ding, A., Yao, M., Li, Y., Li, Q., Zhang, L., Xiao, Y., 2021. Large-eddy simulation of traffic-related air pollution at a very high resolution in a mega-city: evaluation against mobile sensors and insights for influencing factors. Atmos. Chem. Phys. 21, 2917–2929. https://doi.org/10.5194/acp-21-2917-2021.
Zheng, J., Che, W., Wang, X., Louie, P., Zhong, L., 2009. Road network-based spatial allocation of on-road mobile source emissions in the pearl river delta region, China, and comparisons with population-based approach. J. Air. Waste. Manage 59, 1405–1416. https://doi.org/10.3155/1047-3289.59.12.1405.
Zhong, S., Zhang, K., Bagheri, M., Burken, J.G., Gu, A., Li, B., Ma, X., Marrone, B.L., Ren, Z., Schrier, J., Shi, W., Tan, H., Wang, T., Wang, X., Wong, B.M., Xiao, X., Yu, X., Zhu, J., Zhang, H., 2021. Machine learning: new ideas and tools in

environmental science and engineering. Environ. Sci. Technol. 55 (19), 12741–12754. https://doi.org/10.1021/acs.est.1c01339.