Predicting high-resolution air quality using machine learning: Integration of large eddy simulation and urban morphology data☆

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

[a] *School of Atmospheric Sciences, Nanjing University, Nanjing, Jiangsu, China*
[b] *Allen Institute for AI, Seattle, WA, USA*

ABSTRACT

Accurately predicting air pollutants, especially in urban areas with well-defined spatial structures, is crucial. Over the past decade, machine learning techniques have been widely used to forecast urban air quality. However, traditional machine learning approaches have limitations in accuracy and interpretability for predicting pollutants. In this study, we propose a convolutional neural network (CNN) model to predict the spatial distribution of CO concentration in Nanjing urban area at 10 m resolution. Our model incorporates various factors as input, such as building height, topography, emissions, and is trained against the outputs simulated by the parallelized large-eddy simulation model (PALM). The PALM model has 48 different scenarios that varied in emissions, wind speeds, and wind directions. The results display a strong consistency between the two models. Furthermore, we evaluate the performance of our model using a 10-fold cross-validation and out-of-sample cross-validation approach. This yields a robust correlation (with both $R^2 > 0.8$) and a low RMSE between the CO predicted by the PALM and CNN models, which demonstrates the generalization capability of our CNN model. The CNN can extract crucial features from the resulted weight contribution map. This map indicates that the CO concentration at a location is more influenced by nearby buildings and emissions than distant ones. The interpretable patterns uncovered by our model are related to neighborhood effects, wind speeds, directions, and the impact of orientation on urban CO distribution. The model also shows high prediction accuracy ($R > 0.8$) when applied to another city. Overall, the integration of our CNN framework with the PALM model enhances the accuracy of air quality predictions, while enabling a fluid dynamic laws interpretation, providing effective tools for air quality management.

1. Introduction

Air pollution has been identified as the most significant environmental risk factor to health and well-being at the global scale (GBD, 2019; WHO, 2023). Long-term exposure to poor air quality can result in pulmonary disease, heart disease, lung tumors, and stroke (Zhang et al., 2014; Ghorani-Azam et al., 2016). Urban air pollutants (i.e., CO and NOx) exhibit significant spatial heterogeneity due to unevenly distributed emission sources, complex flow patterns and topography, and physicochemical transformations. It is noteworthy that these variations can be substantial, even over relatively short distances, ranging from 10 m to 1 km (Li et al., 2023; Jiang et al., 2021; Apte et al., 2017). Therefore, incorporating high-precision spatiotemporal characteristics of air pollutants into air quality prediction models is crucial for effective air pollution prevention (Yang et al., 2021; Wei et al., 2022).

Large-eddy simulation (LES) models are extensively applied in various cities to examine turbulent flows and air pollution in recent years (Letzel et al., 2008; Zhang et al., 2021). Compared to conventional air quality models, such as Gaussian model, AERMOD, CMAQ, and ADMS-Urban (Biggart et al., 2020; Rood, 2014), the LES models demonstrates enhanced proficiency in capturing turbulent structures and eddies due to its superior ability to account for turbulent eddies and higher resolution, thereby providing a significant advantage in addressing turbulence-related issues (Sun et al., 2016; Wolf et al., 2020). For instance, Zhang et al. (2021) developed an air quality model with a 10-m resolution for traffic-related carbon monoxide (CO) distribution based on parallelized large-eddy simulation model (PALM), uncovering a detailed geographic dispersion pattern of air pollution both within and

around the road network. However, the PALM model still faces challenges such as extended computation times, elevated costs, and the inability to integrate real-time weather data.

Compared to traditional atmospheric models, machine learning techniques (ML) offer a more efficient, precise, and cost-effective approach to simulate all pollutants without relying on pollution inventories (Feng et al., 2019; Bozdag et al., 2020; Zhong et al., 2021). In comparison to traditional ML models, such as support vector machines (Liu et al., 2022a, b) and decision trees (Hu et al., 2017), convolutional neural networks (CNNs) exhibit considerable strengths in feature extraction, managing high-dimensional data, and addressing complex nonlinear relationships (LeCun et al., 2015). Air pollutant forecasting faces a high uncertainty due to its complex generation mechanism and influencing factors such as meteorology, emissions, and urban morphological characteristics (Tian et al., 2014; Li et al., 2019; Wu et al., 2022). CNNs can effectively extract these features through extensive data training (Kow et al., 2020) and perform nonlinear transformations of the input features through multiple layers to reduce this uncertainty. For example, Zhang et al. (2020) proposed a deep spatio-temporal orthogonal regularization residual CNN model for air prediction, which is designed to capture the complex relationship present in dynamically biased meteorological data. Nevertheless, despite the proficiency of CNNs in feature extraction, their complex internal mechanisms often render them as "black box" models, making their decision-making process difficult to interpret. So, the current challenges in ML-based air pollution prediction include reliance on manual feature engineering, the lack of high-quality data, limited spatial resolution, and the difficulty in interpreting results (Liu et al., 2022a, b).

Our goal is to combine the CNN and PALM models to overcome limitations such as insufficient training data and the computational time and cost associated with PALM simulations. Our innovation approach uses the high-resolution, spatially complete outputs of the PALM model to train the CNN model. This configuration permits pollutants prediction across the entire modeling domain, rather than being restricted to monitoring sites alone (Liu et al., 2021a). The aim of this study is to: i) develop a high-resolution (10 m) CNN model that combines PALM results and urban morphology data for predicting traffic-related air quality in Nanjing urban area; ii) evaluate the performance and generalization of our model with cross-validation methods; iii) analyze the contribution patterns of the weights of the topography and emissions to gain a deeper understanding of how the network interpret this information. We also discuss the implications of our CNN model in other cities, aiming to provide efficient data support for urban air quality management.

2. Methodology

2.1. LES-based prediction models

In this study, we use the PALM model (version 4.0, revision 3689) to simulate the dispersion of traffic-related emissions in the urban area of Nanjing, China. The PALM model, developed by the PALM group at Leibniz University Hannover (Maronga et al., 2015), is grounded on the incompressible Navier-Stokes equations in their Boussinesq-approximated form. The model is explicitly developed as a turbulence-resolving Large Eddy Simulation (LES) model system, with an optimization for high-performance computing. For a detailed understanding of the PALM model and its initial configurations, please refer to our previous work (Zhang et al., 2021), which provides a comprehensive explanation of the PALM model implementation. In this study, carbon monoxide (CO) is chosen as a representative pollutant due to its relatively long lifespan, which can extend from several months to years.

The study area is located at 32.07° N and 118.72° E (Fig. 1). The grid resolution of the model is set at 10 m, with a grid size defined as 960 × 960 × 48. To mitigate the interference induced by building structures, we have incrementally stretched the grid vertically by a factor of 1.1
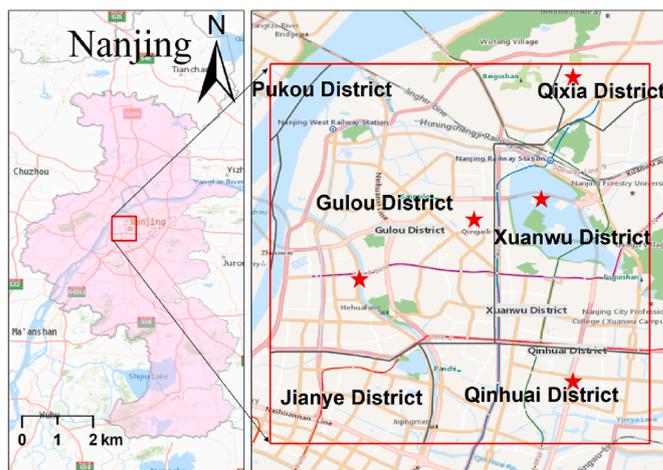

Fig. 1. Model domain with the locations of CO monitoring sites. Red stars are the locations of stationary stations belonging to the national air quality measurement network of China. Map credit: ESRI 2020.

across 48 vertical layers reaching ~1000 m a.s.l., which is more than 3 times of the highest building within the model domain. Fig. S1 displays the distribution of mean CO concentrations in Nanjing urban area under two emission scenarios, as simulated by the PALM model. Given the high computational demand of the model, we confined our simulation to eight wind directions, each 45° apart (N, NE, E, ES, S, SW, W, and NW). We also select 3, 6.5, and 10 m/s to represent low, median, and high wind conditions, respectively (He et al., 2018), resulting in a total of 48 scenarios. Each scenario is run for 3 h with a time step of 6 s. In each simulation, we preserve a consistent state of wind forcing and emissions. This methodology required a period of convergence for the solution, indicating that the simulations were conducted within static boundary conditions. The preliminary 2 h of every model run were allocated to the stabilization of the model's turbulence. We use the hourly average of the third hour for our analysis, which represents the steady-state. All data have been utilized for training and prediction, allowing the model to better learn the patterns within the dataset. To assess the model's generalization capabilities on unseen data, we will validate it using a completely new dataset that differs from previous scenarios.

2.2. CNN input data

We select three types of data as input variables for our model, including building heights, topography, and CO emission rates (Fig. S3). The topography of Nanjing is the sum of two components: baseline elevation and building heights. The emission data is calculated using the "standard road length" method, which assigns total traffic emissions to individual roads based on different road types and traffic flows (Zheng et al., 2009). Leveraging the CO traffic emissions data during peak and non-peak hours, in conjunction with traffic flow data in Nanjing urban area, we derive the high and low emission rate data for CO. Simultaneously, these road emissions data are interpolated into grid points that align with the PALM model parameters. For detail information on the specific data sources and the methodologies employed for processing these input data, kindly refer to Zhang et al. (2021).

2.3. CNN structure

Fig. 2 presents an in-depth visualization of our CNN model's workflow. Before performing convolutional operations, we normalize the input data by rescaling it to a uniform range of 0–1, which serves to mitigate the impact of different scales and dimensions between variables on the model's performance. Additionally, we employ a lambda function to segment the input data into emissions, buildings, wind speed, and
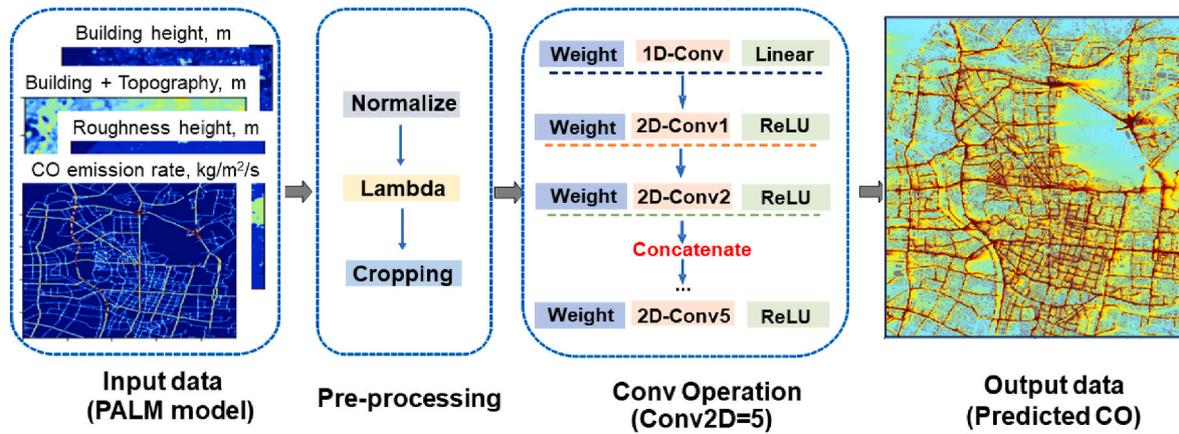
Fig. 2. Workflow of the CNN model.

wind direction. In past experiments, we faced difficulties in training the CNN model to comprehend the importance of wind direction. Specifically, we struggled to simulate the characteristics of plume formation in the downwind direction accurately. Therefore, we pre-rotate all the input data to a west wind to ensure consistent and effective training. For example, the input data such as building height and emissions are rotated 90° anticlockwise for a north wind scenario, 90° clockwise for a south wind, etc. Nevertheless, rotation may lead to the loss of some features, such as the emergence of blank spaces at the edges of the rotated images. To mitigate this effect, we employ edge padding and cropping to reduce the loss of information.

We incorporate five convolutional layers with 64 filters each to extract intricate features. To avoid overfitting, we apply L2 regularization (Lambda = 1e-4) to the kernel weights. We use the Adam optimizer to efficiently adjust the learning rate to optimize weight updates based on the data. Feature extraction is achieved through multiple several convolutional layers with varying kernel sizes (LeCun et al., 2015). In the first layer, we employ a convolutional kernel of size 21 × 21 and 64 convolutional filters (Figs. S3a and 3b). This is succeeded by batch normalization and a Rectified Linear Unit (ReLU) activation function. The second layer uses a 41 × 41 convolutional kernels (filters = 64), following the same steps as the first layer. In the third layer, the convolutional kernel size is increased to 61 × 61, while maintaining 64 convolutional filters. Finally, the output features derived from each convolutional layer are amalgamated, yielding a comprehensive feature map that encapsulates a broad spectrum of extracted information (Fig. 2). In addition to adjusting certain meta-parameters like the number of convolutional layers, we have also explored adjustments to various factors, including the size of the convolution window and the number of channels. We are currently using the optimal configurations determined through these experiments (Table S1).

We deploy our CNN model forecast framework on a personal computing platform with eight virtual central processing unit (CPU) cores (Intel Core i58250U CPU at 1.6 GHz). The CNN model is run and its results are processed using Python 3.8.8. The Keras and TensorFlow packages are installed for this operation. Also, the Scikit-learn library (version 0.20 or higher) and TensorFlow (version 2.0 or higher) are required for this procedure.

2.4. Convolutional layers

Convolutional layers are an essential component of CNNs models, serving the purpose of extracting features from input data, such as images (LeCun et al., 2015). The convolution process calculates the dot product of a moving filter and the input data, resulting in a feature map that captures spatial relationships among the input features. Weight sharing is a method used in CNNs to decrease the model's parameter count (Fang et al., 2017). This method uses the same weights across numerous input patches, preventing overfitting of the training data (Sharma et al., 2019). The weights of our CNN model are initialized by the Xavier method, which scales the weights based on the number of units connected to a node (Glorot and Bengio, 2010). This approach aims to maintain consistent variance in the inputs and outputs, thereby preventing the vanishing or exploding gradients phenomenon during training. Gradually, each convolutional filter becomes attuned to specific visual features such as edges, corners, or color patches. These features may combine into more complex patterns in the deeper layers of the network, enabling the CNN to perform sophisticated image recognition tasks.

The convolution layers in our CNN model apply a kernel across the input using 1-pixel strides. The convolution operation, executed without padding, guarantees that the kernel stays within the input boundaries. Through this process, each input segment is multiplied by the kernel's weights. Given an input matrix '*I*', a 'k × k' filter matrix '*F*', and a bias '*b*', the computation for each element '*E*' in the feature map can be expressed as follows (Palsson et al., 2017):

$$E = \sum (I[i,j] \times F[i,j]) + b \tag{1}$$

where *I* [*i, j*] denotes the element at the *i*th row and *j*th column in the local region of the input. *F* [*i, j*] refers to the element at the *i*th row and *j*th column in the filter matrix *F*. The term (*I* [*i, j*] × *F* [*i, j*]) calculates the sum of the element-wise multiplication between the local input region and the filters. The bias term *b* is added to the convolution results.

2.5. Cross-validation

We employ two types of cross-validation methods to assess the performance of our model, including 10-fold cross-validation and out-of-sample cross-validation. K-fold cross-validation is widely used for validation the model performance as it effectively mitigates overfitting in comparison with other cross-validation techniques (Lu et al., 2021a, b). For our study, the entire dataset is randomly divided into 10 subsets, with one subset serving as the testing set and the remaining nine as the training set. This process is repeated 10 times, with a different subset chosen as the test set each time. By evaluating the model multiple times on different data subsets, we reduce the variability in performance estimates due to different data partitioning strategies. The coefficient of determination (R^2) and root mean squared error (RMSE) are selected as the primary metrics for assessing prediction accuracy. The second approach, out-of-sample cross-validation, is implemented using three entirely new wind scenarios (157.5°, 2 m/s; 247.5°, 4.75 m/s; and, 67.5°, 8 m/s) to evaluate the model's predictive capabilities and generalization. The R^2 and RMSE are also used as predictive indicators.

3. Results and discussion

3.1. Spatial forecast performance

Fig. 3 presents the CO spatial distributions predicted by our forecast model. We divide the simulation results into two general scenarios, peak hours emissions and non-peak hours emissions, with the PALM and CNN results displayed on the left and right respectively. Within the same emission scenario, the simulated CO concentrations are significantly lower for high wind speed pairs compared to low wind speeds. As depicted in Fig. 3, both the PALM and CNN models predict elevated CO concentrations on roads compared to off-roads, with major roads exhibiting higher levels than secondary roads (Wang et al., 2021). However, the variation in CO distribution among the eight wind directions is minimal (Zhang et al., 2021). The high correlation coefficient ($R^2 = 0.96$) between mean CO concentrations predicted by both the PALM model and CNN models indicates a high degree of accuracy in urban pollutant prediction offered by the CNN model (Fig. 4). Fig. S4 presents correlation heatmaps that illustrate the relationship between the predicted CO concentration data points from the PALM and CNN models across 48 scenarios. The heatmaps reveal strong correlations with r values ranging from 0.90 to 0.93 and low RMSE [0.04–0.19]. These findings highlight the high accuracy of our CNN model.

Several factors contribute to this enhanced performance. Firstly, compared to traditional methods, CNNs possess superior capabilities in automatic feature extraction, handing high-dimensional data, parameter sharing, and translation invariance (Alzubaidi et al., 2021). Secondly, by integrating high-resolution prediction inputs such as building height, topography, and emissions, and PALM outputs, our model can
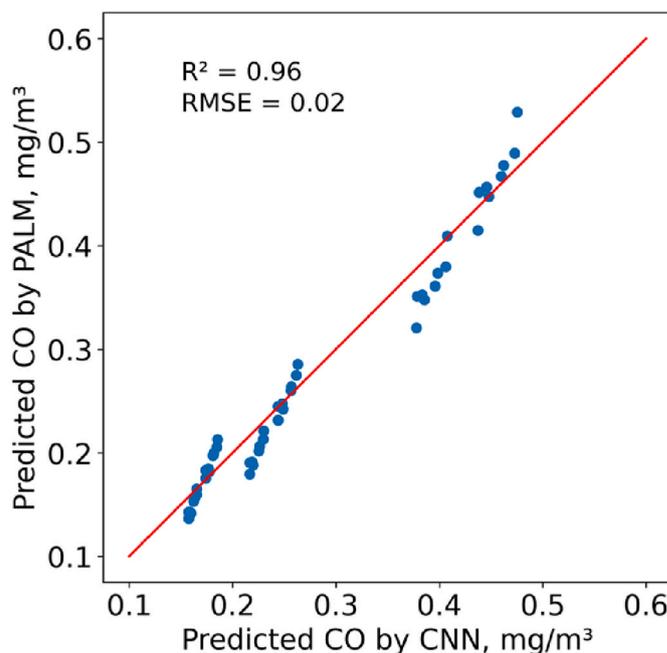

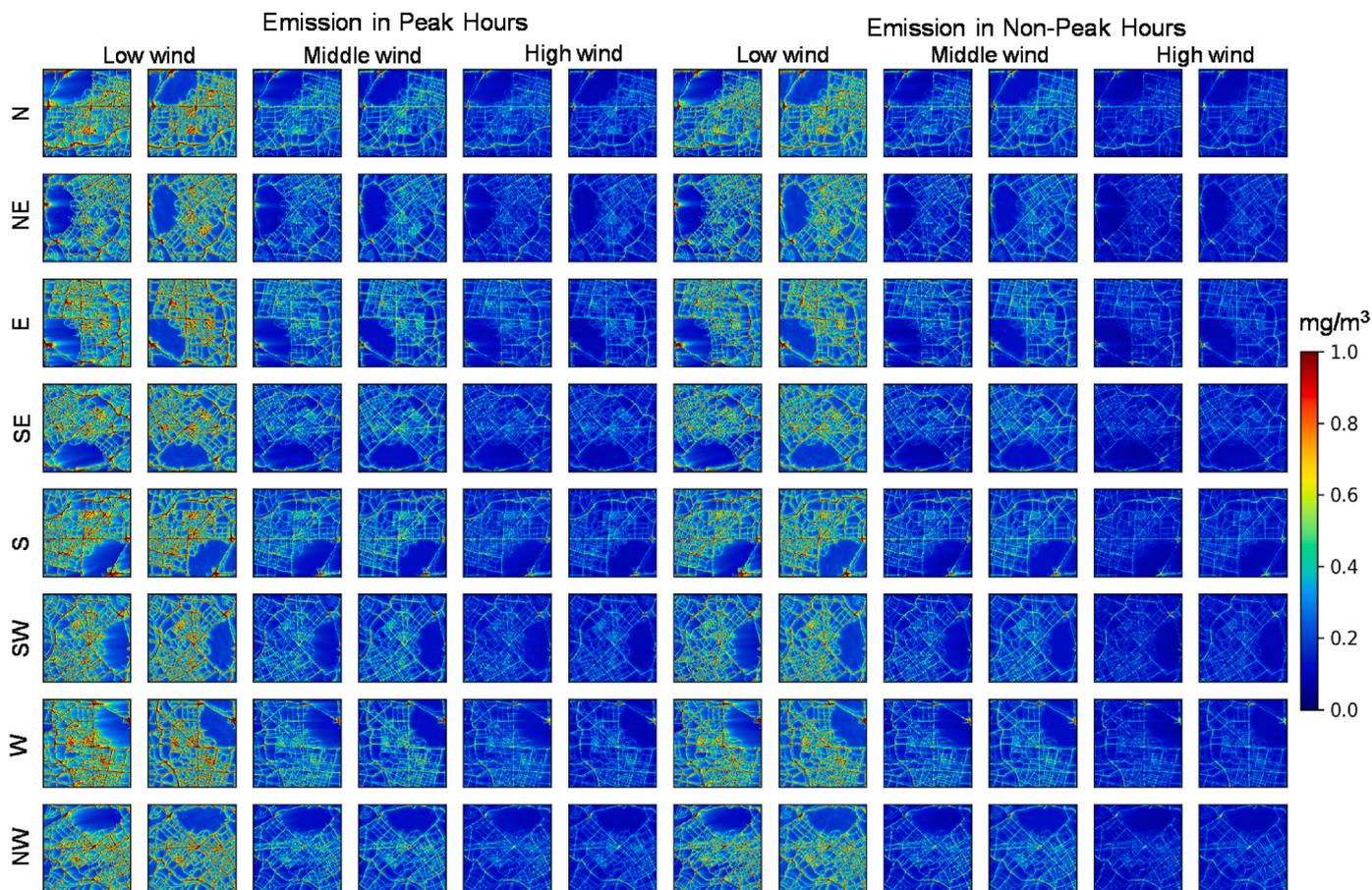**Fig. 4.** Correlation coefficient of mean CO concentrations predicted by the PALM and CNN models.

Fig. 3. Spatial distribution of CO concentration predicted by the PALM (left) and CNN models (right) under two emission scenarios (peak and non-peak hours emissions), three wind speeds (high, medium, and low), and eight wind directions (N, NE, E, ES, S, SW, W, and NW). We also rotate and center cropping the inputs and outputs of the PALM and CNN model.

efficiently capture the real urban environments and the nonlinearity of atmospheric physics laws like atmospheric turbulence (Wright et al., 2022).

3.2. Model evaluation

Our CNN model outperforms the linear model across three variable inputs with 10-fold cross-validation. The random 10-fold CV R^2 of our CNN model is 0.86 ± 0.01, and the CV RMSE is 0.41 ± 0.20 mg/m³. This indicates a strong correlation of CO concentration predicted by the PALM and CNN models, demonstrating the CNN model's predictive accuracy. Moreover, we observe that constraining the prediction variables, such as reducing wind direction, wind speed, or emission scenario, resulted in a marginal decrease of 0.01 in the R^2 score, while the RMSE values increased by 0.01 mg/m³. These findings demonstrate the importance of these predictor variables in our model's predictive accuracy.

We also simulate three new wind scenarios that are completely different from the previous training data. The heatmaps presented in Fig. 5 and Fig. S5 depict the correlation between the predicted CO concentrations from the PALM and CNN models under three different wind fields. The subsequent results reveal R^2 and RMSE values for wind speeds of 2 m/s ($R^2 = 0.803$, RMSE = 0.534), 4.75 m/s ($R^2 = 0.833$, RMSE = 0.163), and 8 m/s ($R^2 = 0.688$, RMSE = 0.153). These results suggest that our CNN model outperforms other machine learning models in conditions of lower wind speed, with R^2 values surpassing those reported by Chen et al. (2023) ($R^2 = 0.72$–0.78) and Lu et al. (2021a) ($R^2 = 0.66$–0.80).

However, our CNN model does exhibit certain limitations, such as the mislabeling of the plume on Xuanwu Lake. Given the lake's relatively flat surface, this mislabeling appears to be confined to a small, localized area. Nonetheless, the model accurately represents the primary characteristics of architectural occlusion across the majority of the map. Despite the map retaining the fundamental elements of occluded buildings, the model falls short in comprehending some of the hyper-localized details. The performance of the CNN model can be further enhanced by expanding the sample size and implementing analytical techniques such as feature engineering (Xing et al., 2020).

3.3. Computational efficiency analysis

Table 1 compares the difference in runtime and memory usage between the PALM and CNN models. The PALM model exhibits an average runtime of 16.7 h, in stark contrast to the CNN model, which necessitates a mere 120 s (over 10^3 time faster than the PALM model). For memory consumption, the PALM model requires, on average, 19285.7 MB, compared to the CNN model's modest requirement of 917 MB, which constitutes approximately 4.75 % of the PALM model's memory footprint.

Research conducted by Jurado et al. (2022) has explored the synergy of computational fluid dynamics (CFD) and deep learning models, specifically multiResUnet, to simulate a 150 $\times$ 150 m² area. Findings indicate that deep learning-based predictions significantly surpass the CFD model in computational efficiency. In contrast with the CFD approach, our PALM model harnesses parallel computing to efficiently manage larger-scale computations and addresses more intricate resolution schemes, enhancing operational efficiency. While the multiResUnet model, when integrated with CFD, achieves high-accuracy predictions, its performance in predicting high-resolution pollutants at larger scales requires further validation. Our CNN model, which builds upon the PALM framework, promises to deliver faster and more precise predictions for large-scale, high-resolution scenarios.

3.4. Feature map

Upon analyzing the weight contributions in Fig. 6, our CNN model reveals significant interpretive patterns for topography, particularly building height. For example, matrices [1, 1], [2, 2], and [7, 5] display localized neighborhoods, indicating that CO concentration at a location is more influenced by nearby building height rather than distant ones. This emphasizes the barrier effect that buildings have on CO dispersion (Hajra et al., 2011). Additionally, matrix [1, 5] and [7, 1] show higher weights on the left side, while [1, 6] and [5, 2] exhibit the opposite trend. This pattern can be attributed to the convective effect of westerly winds, which blow stronger on the windward side of buildings than on the leeward side. When building height differences exit, flow patterns form, accumulating pollutants on the leeward side (Fu et al., 2017). Moreover, directional patterns in matrices [8, 1], [6, 2], [6, 5], and [2, 6] suggest that CO concentration is also influenced by building and road orientation. All of these influencing factors and features align with the findings from our PALM model (Zhang et al., 2021). However, some features are not well explained. For instance, the topography weight contributions in matrices [2, 3], [1, 3], and [3, 7] do not exhibit a discernible pattern. This lack of clear patterns implies that CO concentrations may depend on complex combinations of various factors, including emissions, building height, terrain, wind direction, and wind
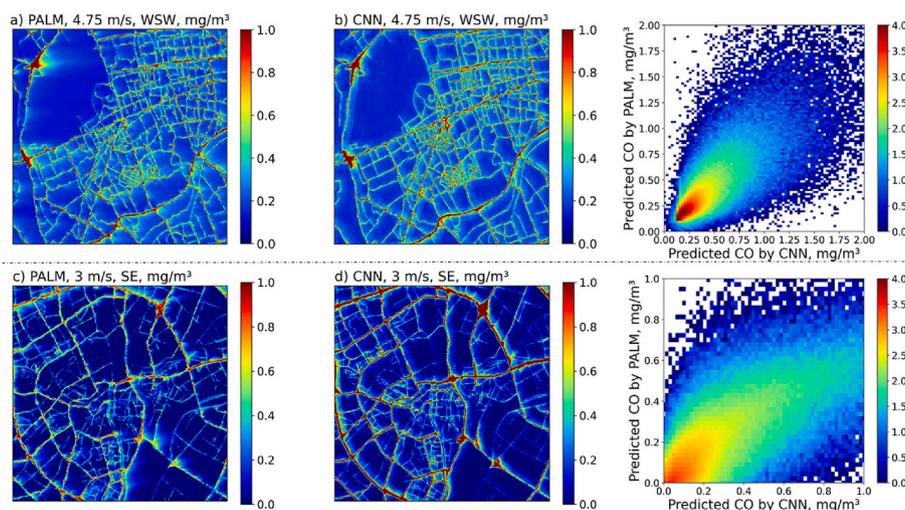

Fig. 5. Comparison of predicted CO concentration distributions by the PALM and CNN models in Nanjing (a and b) and Dongguan (c and d). The correlation heatmaps on the right illustrates the relationship between all data points of CO concentrations predicted by both the PALM and CNN models under these new scenarios.

Table 1
Comparative analysis of computational efficiency across various models.

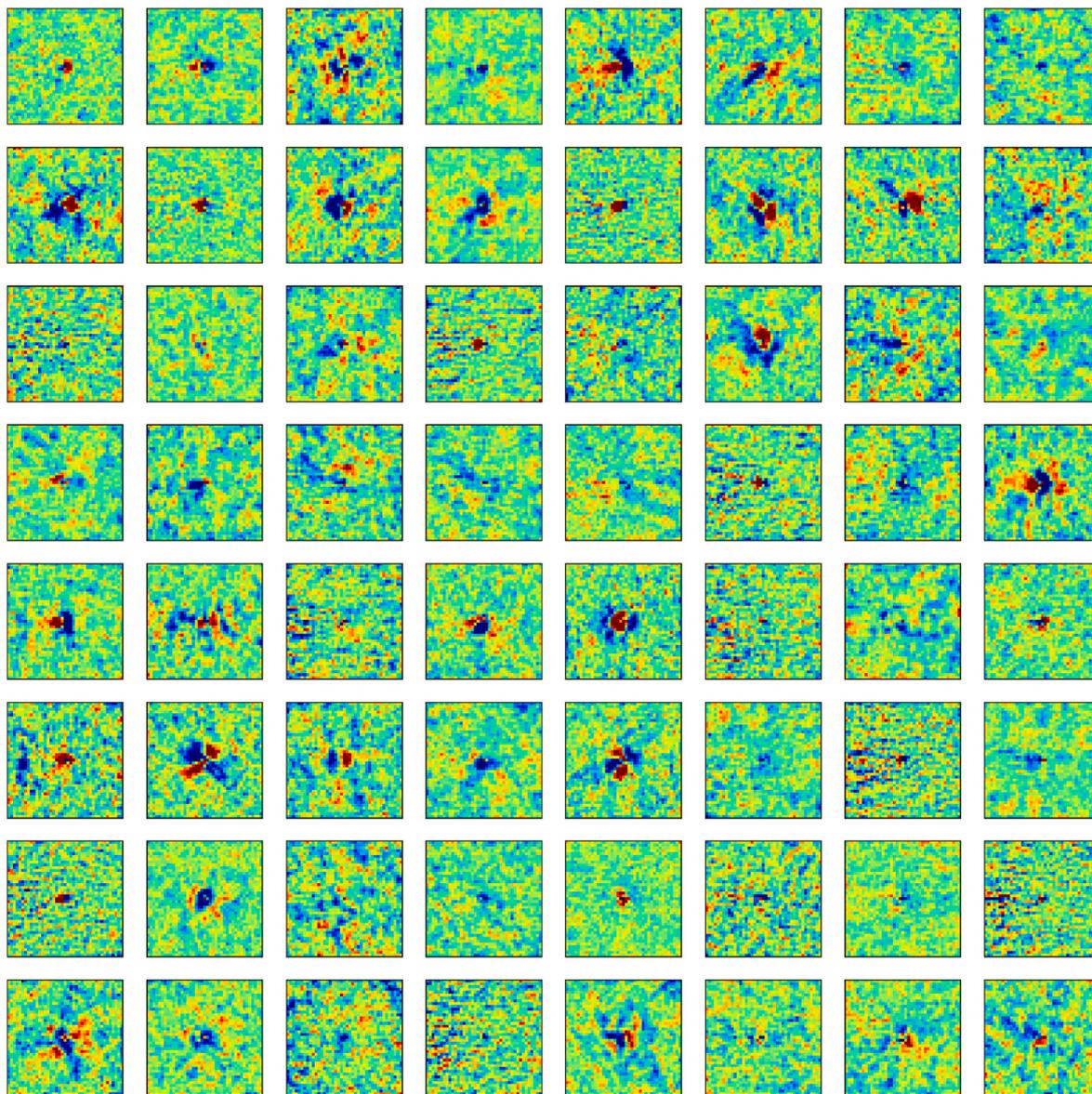| Models		Runtime, s	Memory, MB	Resolution, m	Domain	CPU	Reference
PALM	mean	60224.71	19285.69	10	960*960	120	This study
	STD	25753.83	608.24				
	median	57251.50	19316.38				
CNN		120	917.10	10	960*960	8	Jurado et al. (2022)
CFD		14400	–	1	200*200	48	Reiminger et al. (2020)
multiResUnet		22	–	1	150*150	–	Jurado et al. (2022)

Fig. 6. The weight contribution from the first layer of the CNN model for topography. The matrix, denoted as F [i, j], represents the diverse features that each filter learns as it convolves across the entire study area, with 'i' rows and 'j' columns. The range for 'i' and 'j' is from 1 to 8, with the upper left cell being [1, 1]. The weights are normalized to a specific range [-0.025, 0.025]. Typically, red indicates higher positive weight contribution values, while blue denotes higher negative values.

speed (Liu et al., 2021b).

Fig. S6 displays the contribution matrix of the first layer CNN weights for emissions. Notably, matrices [2, 2], [1, 2], [2, 3], [6, 5], and [8, 8] exhibit a distinct center effect, indicating that emissions in the vicinity of a point have a greater impact on CO levels compared to those farther away (Wang et al., 2021; Biggart et al., 2020). Furthermore, matrices [5, 4] and [4, 6] effectively capture the absence of traffic emissions in Xuanwu Lake, highlighting the robustness of our CNN model in extracting this feature. Additionally, the left weights in matrices [2, 5] and [2, 8] are noticeably higher than the right side. This suggests that wind direction influences the transport of emission, consequently impacting the distribution of CO. The emission matrix also helps identify roads with elevated emissions, such as Huju Road, which is extracted by the weights in matrix [5, 2]. This is also consistent with the results obtained from our previous observations (Wang et al., 2021).

To better understand the influence of different features on prediction

outcomes and to discern the saliency of individual features, we employed saliency map in our analysis (Simonyan et al., 2014). Fig. S7 presents the significance plots of various input features under a selected scenario (247.5°, 4.75 m/s). The figure elucidates the differential impacts of topography, emissions, and wind characteristics on the model's CO predictions. The saliency maps for topography and wind features suggest a stronger influence on the model's predictions in on-road situations compared to off-road ones. However, the impact of emission features on predictions appears more intricate, being affected by a multitude of factors such as roadside vegetation, building heights (Zhang et al., 2021). In summary, the CNN model utilizes weight contribution matrices and saliency map to identify and analyze patterns or features within the PALM model results, as well as the input data. The weight patterns reveal that higher accuracy of the model can be achieved by aligning all wind directions to a single orientation and incorporating topography or emission parameters. Furthermore, the weights demonstrate that our CNN model can effectively learn features related to physic-laws within the PALM model on a fine scale, and can predict the spatial distribution of pollutant concentrations comprehensively (Liu et al., 2022a, b).

3.5. Model application

We have implemented the pretrained CNN model in Dongguan, Guangdong province. Fig. S8 shows the model domain. The model resolution remains the same as that in Nanjing (10 m), with a grid size of $960 \times 960 \times 48$. We use the PALM outputs and three key predictive variables (topography, building heights, and emissions) in Dongguan as inputs data (Fig. S9). These data are gathered and processed following the same methodology as that employed in Nanjing.

Comparing the distribution of predicted CO concentrations by the PALM and CNN models in Dongguan, our study reveals a strong correlation coefficient (R value) (0.675–0.804), and low RMSE (0.273–0.331 mg/m^3) (Fig. 5c–d and Figs. S10–S11). These results show that the prediction performance of the CNN model for CO concentration in Dongguan is stronger in the northwest and southeast directions (R = 0.804, RMSE $\approx$ 0.273 mg/m^3) than other wind directions. Furthermore, the CNN model successfully identifies some key features, such as higher CO concentrations on major roads than on secondary roads. The modeled CO concentrations on-roads are also higher than those in off-road areas, indicating that they are mainly affected by traffic emissions (Zhang et al., 2021). The CNN model's overestimation of pollutant concentrations on the roads of Dongguan city can be primarily attributed to the input variables, with a significant emphasis on emissions data. Dongguan experiences higher vehicle ownership and traffic flow compared to Nanjing, resulting in increased traffic-related emissions and thus higher simulated pollutant levels. Furthermore, the model's predictions are influenced by factors such as surface roughness and the height of urban structures, which affect pollutant dispersion and contribute to the observed overestimation. To improve the model's performance, future research should consider fine-tuning the parameters and expanding the model's application to a variety of urban environments. This approach is expected to yield more accurate predictions across diverse urban scenarios.

Combining the feature maps, we find that the different patterns obtained by the trained CNN model are also suitable for Dongguan. For instance, the emission epicenter significantly influences the CO distribution more than locations peripheral to the center (Fig. S6). This suggests that our model has the potential for broad application across diverse urban environments. However, the effectiveness of our model may be limited in cities with significantly variations in prediction variables. To enhance the migration capability and prediction accuracy, it is essential to optimize and adjust the model parameters by incorporating data from other cities. Our CNN model can also be extended to predict the high-resolution distribution of other pollutants in the future, such as NO$_2$ and PM$_{2.5}$ (Wolf et al., 2020). Furthermore, the model could

potentially be integrated with other predictive models, such as the Long Short-Term Memory (LSTM) model, to provide real-time air quality estimates (Lu et al., 2021b).

4. Conclusion

Our research demonstrates that by integrating the outputs of the PALM model with CNN model, we can enhance the accuracy and performance of urban air quality prediction. Our CNN model not only offers high computational speed (less than 0.2% computational time relative to the PALM model) and cost-effectiveness, but also preserves the fluid dynamic principles of the PALM model, thereby improving interpretability.

The CNN model has some uncertainties. For example, input data quality can impact learning ability and increase prediction uncertainty. Since the CNN learns from PALM outputs, errors in PALM propagate to the CNN. Uncertainties may also arise from the CNN's architecture and parameters like kernel size and activation functions. Techniques like data augmentation, regularization, and transfer learning, could be employed to tackle the limitations of CNN model. In conclusion, this study is among the first to propose high-resolution CNN model that combines PALM results and urban morphology data for air quality prediction. With the increasing availability of high-resolution data on urban architecture and traffic distribution, this model framework can be extended to other regions and air pollutants, such as nitrogen dioxide and ozone, using the same data sources (topography). The highly detailed concentration maps simulated across major urban regions will prove instrumental for smart city systems. Moreover, this air quality prediction can provide invaluable data for scientific research, assist in policy formation, raise public awareness, and support emergency response efforts to combat the harmful effects of air pollution.

CRediT authorship contribution statement

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

Data availability

Data will be made available on request.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2024.123371.

References

Alzubaidi, L., Zhang, J., Humaidi, A.J., Dujaili, A., Duan, Y., Shamma, O., Santamaría, J., Fadhel, M.A., Amidie, M., Farhan, L., 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J. Big Data. 8, 53.

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.

Biggart, M., Stocker, J., Doherty, R.M., Wild, O., Hollaway, M., Carruthers, D., Li, J., Zhang, Q., Wu, R., Kotthaus, S., Grimmond, S., Squires, F.A., Lee, J., Shi, Z., 2020. Street-scale air quality modelling for Beijing during a winter 2016 measurement campaign. Atmos. Chem. Phys. 20, 2755–2780.

Bozdag, A., Dokuz, Y., Gökçek, Ö.B., 2020. Spatial prediction of PM_{10} concentration using machine learning algorithms in Ankara, Turkey. Environ. Pollut. 263, 114635.

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.

Fang, J., Zhou, Y., Yu, Y., Du, S., 2017. Fine-grained vehicle model recognition using a coarse-to-fine convolutional neural network architecture. IEEE Trans. Intell. Transport. Syst. 18 (7), 1782–1792.

Feng, R., Zheng, H.J., Gao, H., Zhang, A.R., Huang, C., Zhang, J.X., Luo, K., Fan, J.R., 2019. Recurrent neural network and random forest for analysis and accurate forecast of atmospheric pollutants: a case study in Hangzhou, China. J. Clean. Prod. 231, 1005–1015.

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.

GBD 2019 Risk Factors Collaborators, 2019. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 396, 1223–1249.

Ghorani-Azam, A., Riahi-Zanjani, B., Balali-Mood, M., 2016. Effects of air pollution on human health and practical measures for prevention in Iran. J. Res. Med. Sci. Off. J. Isfahan Univ. Med. Sci. 21–65.

Glorot, X., Bengio, Y., 2010. Understanding the Difficulty of Training Deep Feedforward Neural Networks. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, vol. 9, pp. 249–256.

Hajra, B., Stathopoulos, T., Bahloul, A., 2011. The effect of upstream buildings on near-field pollutant dispersion in the built environment. Atmos. Environ. 45 (28), 4930–4940.

He, J., Lu, C., Xie, S., Jiang, Y., Han, Z., 2018. Assessing the performance of wind profile radar in Nanjing and its application. J. Meteorol. Sci. 38, 406–415.

Hu, X., Belle, J.H., Meng, X., Wildani, A., Waller, L.A., Strickland, M.J., Liu, Y., 2017. Estimating $PM_{2.5}$ concentrations in the conterminous United States using the random forest approach. Environ. Sci. Technol. 51 (12), 6936–6944.

Jiang, L., Xia, Y., Wang, L., Chen, X., Ye, J., Hou, T., Wang, L., Zhang, Y., Li, M., Li, Z., Song, Z., Jiang, Y., Liu, W., Li, P., Rosenfeld, D., Seinfeld, J.H., Yu, S., 2021. Hyperfine-resolution mapping of on-road vehicle emissions with comprehensive traffic monitoring and an intelligent transportation system. Atmos. Chem. Phys. 21, 16985–17002.

Jurado, X., Reiminger, N., Benmoussa, M., Vazquez, J., Wemmert, C., 2022. Deep learning methods evaluation to predict air quality based on Computational Fluid Dynamics. Expert Syst. Appl. 203, 117294.

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.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.

Letzel, M.O., Krane, M., Raasch, S., 2008. High resolution urban large-eddy simulation studies from street canyon to neighbourhood scale. Atmos. Environ. 42 (38), 8770–8784.

Li, C., Wang, Z., Li, B., Peng, Z.R., Fu, Q., 2019. Investigating the relationship between air pollution variation and urban form. Build. Environ. 147, 559–568.

Li, L., Wang, J., Franklin, M., Yin, Q., Wu, J., Camps-Valls, G., Zhu, Z., Wang, C., Ge, Y., Reichstein, M., 2023. Improving air quality assessment using physics-inspired deep graph learning. npj Clim. Atmos. Sci. 6, 152.

Liu, C., Lin, T., Tuan, K., Chiueh, P.T., 2022a. Spatio-temporal prediction and factor identification of urban air quality using support vector machine. Urban Clim. 41, 101055.

Liu, M., Chen, H., Wei, D., Wu, Y., Li, C., 2021b. 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.

Liu, X., Lu, D., Zhang, A., Liu, Q., Jiang, G., 2022b. Data-driven machine learning in environmental pollution: gains and problems. Environ. Sci. Technol. 56 (4), 2124–2133.

Liu, X., Zhang, X., Schnelle-Kreis, J., Jakobi, G., Cao, X., Cyrys, J., Yang, L., Schloter-Hai, B., Abbaszade, G., Orasche, J., Khedr, M., Kowalski, M., Hank, M., Zimmermann, R., 2021a. Spatiotemporal characteristics and driving factors of black

carbon in Augsburg, Germany: combination of mobile monitoring and street view images. Environ. Sci. Technol. 55 (1), 160–168.

Lu, T., Marshall, J.D., Zhang, W., Hystad, P., Kim, S.Y., Bechle, M.J., Demuzere, M., Hankey, S., 2021a. National empirical models of air pollution using microscale measures of the urban environment. Environ. Sci. Technol. 55 (22), 15519–15530.

Lu, X.C., Sha, Y.H., Li, Z.N., Huang, Y.Q., Chen, W.Y., Chen, D.H., Shen, J., Chen, Y., Fung, J.C.H., 2021b. Development and application of a hybrid long-short term memory – three-dimensional variational technique for the improvement of $PM_{2.5}$ forecasting. Sci. Total Environ. 770 (10), 144221.

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.

Palsson, F., Sveinsson, J.R., Ulfarsson, M.O., 2017. Multispectral and hyperspectral image fusion using a 3-D-convolutional neural network. Geosci. Rem. Sens. Lett. IEEE 14 (5), 639–643.

Reiminger, N., Vazquez, J., Blond, N., Dufresne, M., Wertel, J., 2020. CFD evaluation of mean pollutant concentration variations in step-down street canyons. J. Wind Eng. Ind. Aerod. 196, 104032.

Rood, A.S., 2014. Performance evaluation of AERMOD, CALPUFF, and legacy air dispersion models using the winter validation tracer study dataset. Atmos. Environ. 89, 707–720.

Sharma, A., Vans, E., Shigemizu, D., Boroevich, K.A., Tsunoda, T., 2019. DeepInsight: a methodology to transform a non-image data to an image for convolution neural network architecture. Sci. Rep. 9, 11399.

Simonyan, K., Vedaldi, A., Zisserman, A., 2014. Deep inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps arXiv preprint arXiv: 1312.6034.

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.

Tian, G., Qiao, Z., Xu, X., 2014. Characteristics of particulate matter (PM_{10}) and its relationship with meteorological factors during 2001–2012 in Beijing. Environ. Pollut. 192, 266–274.

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.

Wei, J., Liu, S., Li, Z., Liu, C., Qin, K., Liu, X., Pinker, R.T., Dickerson, R.R., Lin, J., Boersma, K.F., Sun, L., Li, R., Xue, W., Cui, Y., Zhang, C., Wang, J., 2022. Ground-level NO_2 surveillance from space across China for high resolution using interpretable spatiotemporally weighted artificial intelligence. Environ. Sci. Technol. 56 (14), 9988–9998.

WHO, 2023. Ambient Air Pollution: Training for Health Care Providers, third ed. World Health Organization https://iris.who.int/handle/10665/374515.

Wolf, T., Pettersson, L.H., Esau, I., 2020. A very high-resolution assessment and modelling of urban air quality. Atmos. Chem. Phys. 20, 625–647.

Wright, L.G., Onodera, T., Stein, M.M., Wang, T., Schachter, D.T., Hu, Z., McMahon, P., 2022. Deep physical neural networks trained with backpropagation. Nature 601, 549–555.

Wu, X., Yang, D., Wu, R., Gu, J., Wen, Y., Zhang, S., Wu, R., Wang, R., Xu, H., Zhang, K. M., Wu, Y., Hao, J., 2022. High-resolution mapping of regional traffic emissions using land-use machine learning models. Atmos. Chem. Phys. 22, 1939–1950.

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.

Yang, J., Yan, R., Nong, M., Liao, J., Li, F., Sun, W., 2021. PM2.5 concentrations forecasting in Beijing through deep learning with different inputs, model structures and forecast time. Atmos. Pollut. Res. 12, 101168.

Zhang, L., Li, D., Guo, Q., 2020. Deep Learning from spatio-temporal data using orthogonal regularizaion residual CNN for air prediction. IEEE Access 8, 66037–66047.

Zhang, W., Qian, C., Zeng, Y.X., 2014. Air pollution: a smoking gun for cancer. Chin. J. Cancer 33 (4), 173.

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.

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 Manag. Assoc. 59, 1405–1416.

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.