



Article

Application of a Machine Learning Method for Prediction of Urban Neighborhood-Scale Air Pollution

Ka-Ming Wai * and Peter K. N. Yu *

Department of Physics, City University of Hong Kong, Hong Kong SAR, China

* Correspondence: bhkwmwai@cityu.edu.hk (K.-M.W.); peter.yu@cityu.edu.hk (P.K.N.Y.)

Abstract: Urban air pollution has aroused growing attention due to its associated adverse health effects. A model which could promptly predict urban air quality with considerable accuracy is, therefore, important and will benefit the development of smart cities. However, only a computational fluid dynamics (CFD) model could better resolve the dispersion behavior within an urban canyon layer. A machine learning (ML) model using the Artificial Neural Network (ANN) approach was formulated in the current study to investigate vehicle-derived airborne particulate (PM₁₀) dispersion within a compact high-rise-built environment. Various measured meteorological parameters and PM₁₀ concentrations were adopted as the model inputs to train the ANN model. A building-resolved CFD model under the same environmental settings was also set up to compare its model performance with the ANN model. Our results showed that the ANN model exhibited promising performance ($r = 0.82$, fractional bias = 0.002) when comparing the > 1000 h PM₁₀ measurements. When comparing the diurnal hourly measured PM₁₀ variations in a clear-sky day, both the ANN and CFD models performed well ($r > 0.8$). The good performance of the CFD model relied on the knowledge of the in situ diurnal traffic profile, the adoption of suitable mobile source emission factor(s) (e.g., from MOBILE 6 and COPERT4), and the use of urban thermal and dynamical variables to capture PM₁₀ variations in both neutral and unstable atmospheric conditions. These requirements/constraints make it impractical for daily operation. On the contrary, the ML (ANN) model adopted here is free from these constraints and is fast (less than 0.1% computational time relative to the CFD model). These results demonstrate that the ANN model is a superior option for a smart city application.



Citation: Wai, K.-M.; Yu, P.K.N. Application of a Machine Learning Method for Prediction of Urban Neighborhood-Scale Air Pollution. *Int. J. Environ. Res. Public Health* **2023**, *20*, 2412. <https://doi.org/10.3390/ijerph20032412>

Academic Editors: Tao Liu and Yan Yan

Received: 13 December 2022

Revised: 20 January 2023

Accepted: 24 January 2023

Published: 29 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: urban environment; air quality model; machine learning; ENVI-met model; smart city

1. Introduction

The epidemiologic evidence of particulate pollution-induced health effects is well documented [1,2]. A total economic loss of USD 2.4 billion per year was estimated from PM₁₀-induced premature death and chronic respiratory diseases in the Pearl River Delta of southern China [3]. Road-side vehicular emissions are the main source of atmospheric particulates in the ambient urban air of cities that are not directly influenced by industrial emissions [4,5]. Hong Kong, a megacity in southern China, suffers from a similar air quality problem [6]. In view of this, a simulation model for urban air pollution, which can produce rapid and robust results, is of urgent need for practical use. It would benefit not only Hong Kong but also other megacities around the world. For instance, more than 50% of people live in cities in China [7]. Technological advances in urban air quality management in the context of simulations and monitoring are also essential for smart city development [8].

The plume dispersion in the urban canopy layer (UCL) is unique compared to that in the free atmosphere well above the UCL. The UCL is featured with building-induced flows, such as wake recirculation, channeling and branching in intersections. In addition, the heterogeneity in building heights could result in the asymmetries of the vertical plume structure and, in turn, a shift of the effective source height [9]. The Gaussian dispersion model offers a simplified representation of downwind concentration spread from the

emission sources. Popular models of this type, which parameterize the urban effects, are the US EPA's model AERMOD [10,11] and the UK's ADMS-urban [12,13]. A CFD model is capable of better resolving building-influenced wind and turbulence mixing in the built environment, which governs the pollutant dispersion in the urban canopy layer, and is thus a more accurate method. However, both the computational resource and time for a CFD model are demanding [14–17]. In addition, atmospheric stratification, which governs the vertical motions of fluid particles, is one of the challenges in CFD simulations. Currently, many studies only focus on neutral flows because of their numerical simplification.

More recently, ML technique has been used in predicting regional-scale air pollution. A few studies reported better performance for ML models in regional-scale air quality prediction compared to conventional physiochemical numerical air quality models [18]. Various ML algorithms have been used in air pollution prediction, namely, ANN (Artificial Neural Network; [19]), LASSO regression (Least Absolute Shrinkage and Selection Operator regression; [20]), LSTM (Long Short-Term Memory; [21]), kNN (k-Nearest Neighbor; [22]), RF (Random Forest; [23]), and SVM (Support Vector Machines; [24]). Bozdog et al. [19] reported that ANN algorithm [among other algorithms (LASSO, SVR, RF, kNN)] produces the best results ($r^2 = 0.58$; RMSE = 20.8, MAE = 14.4) when performing a spatial prediction of PM₁₀ concentration in Turkey. Studies have shown that meteorological characteristics could play an important role in the prediction of air pollutants [21,25]. Ma and Zhang [26] commented that using some traditional algorithms, such as radical basis function, back propagation neural network and SVM model, requires too many inputs, but the prediction results are not in good agreement with the measurements. Nevertheless, an application of a ML model on neighborhood-scale air pollution dispersion within the UCL of a compact city is rarely found in the literature.

The study goal here is to investigate if a recently developed ML technique is feasible to build a fast and relatively accurate model to predict neighborhood-scale PM₁₀ concentration levels in a compact-city environment. The performance of the ML model was compared with the PM₁₀ measurements and then with a CFD model, which is known to provide more accurate results in predicting the PM₁₀ levels in the UCL. Prior to the simulations, the ML model was formulated by a dataset of past PM₁₀ monitoring data. The CFD model was set up with the environmental settings (e.g., building configurations) in the study area.

2. Materials and Methods

This study was conducted within a densely populated urban environment of Hong Kong (22.30° N, 114.17° E). The study site (Figure 1) is featured with a road-side air quality monitoring station operated by the Hong Kong Environmental Protection Department (EPD), two major roads, and sparse vegetation, and it is surrounded by buildings with different heights (5–26 stories). The subsequent section details the ANN and CFD models used here.

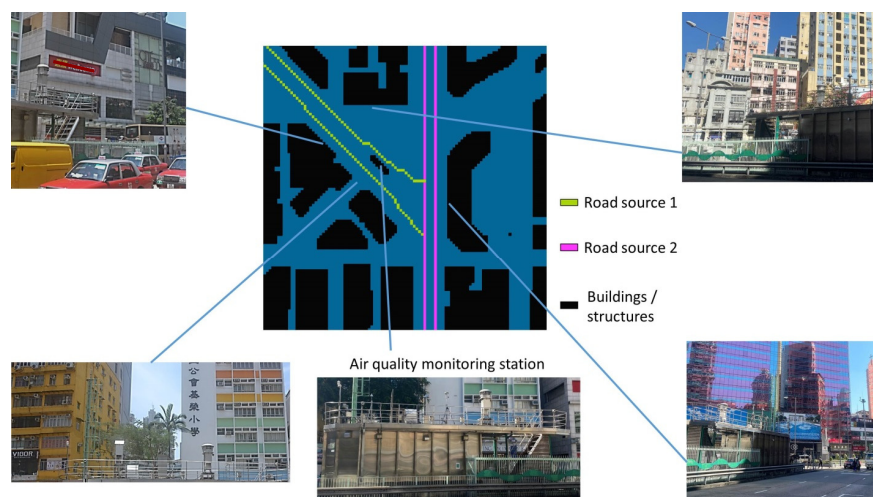


Figure 1. The site environment. The CFD model domain (center) and snapshots around the site are shown.

2.1. Artificial Neural Network (ANN) Model

The ANN (an ML algorithm) model was formulated to predict the neighborhood-scale PM₁₀ dispersion within the UCL of the study site. The model mimics natural neurons in animal brains. The details of the model have been discussed elsewhere, e.g., [27]. Briefly, the ANN model consists of interconnected neurons at the input, hidden, and output layers (Figure 2). Input values are collected in the input layer and then sent to different neurons (or processing units), which constitute the hidden layer. Output variables are eventually obtained at the output layer after the data are processed. Each neuron in the hidden layer computes a weighted sum of the inputs. The weight is subjected to change during the ANN training in order to provide its best estimate to the output. The selection of a proper number of hidden layer is important for the model construction. Although adding more hidden layer might improve the model's performance, it is noted that more complexity of the training process is imposed [28]. Therefore, one hidden layer was used here. The number of neurons in the hidden layer was determined by $N_{\text{hidden}} = 2 N_{\text{input}} + 1$, where N_{hidden} and N_{input} are the number of neurons in the hidden and input layers, respectively [29]. To avoid model instability, all input parameters were scaled from 0 to 1. The feed-forward neural network was used, which was successfully adopted in other pollution transport studies, e.g., [30]. It is called the feed-forward network since data flow within the network from one layer to the next one without any return path. A hyperbolic tangent sigmoid transfer function for the neurons in the hidden layer was adopted to reduce the computational time required during the training process. The efficient Levenberg–Marquardt algorithm for training was used, such that the model achieved a mean squared error (MSE) < 0.004. The model was constructed by the MATLAB software (The MathWorks, USA). Table 1 details the model settings.

Table 1. The ANN model settings.

	Parameters
Input layer	Number of neurons: 11
	Background wind speed
	Background wind direction
	Background air temperature
	Background PM ₁₀ concentration
	Atmospheric pressure
	Rainfall
	Canyon wind speed
	Canyon wind direction
	Canyon air temperature
Hidden Layer	Dates of a week
Output Layer	Weekday/weekend
Transfer function for hidden layer	Number of neurons: $N_{\text{hidden}} = 2N_{\text{input}} + 1$
Transfer function for output layer	Number of neurons: 1
Training method	Tangent Sigmoid
	Linear
	Goal: minimum MSE
Dataset	Epoch: 1000 times
	Algorithm: Levenberg–Marquardt
	Total size: 8616
	Data for training: 70%
	Data for validation: 15%
	Data for testing: 15%

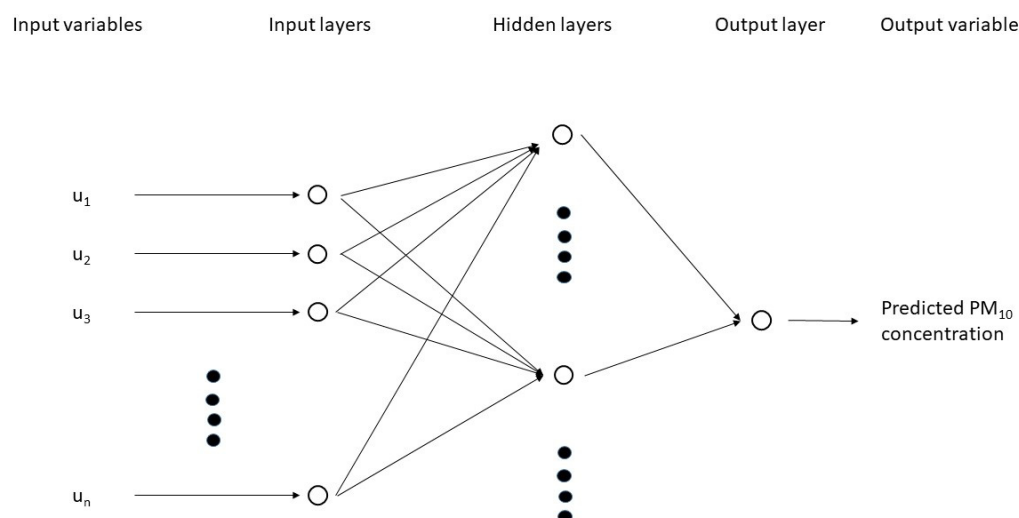


Figure 2. A schematic representation of the feed-forward neural network (FFNN).

2.2. Computational Fluid Dynamics (CFD) Model

The ENVI-met model (version 5.0) was used to simulate the PM₁₀ dispersion in the UCL of the study area. It is a 3-dimensional, microscale, non-hydrostatic computational fluid dynamics (CFD) model and uses the RANS (Reynolds-Averaged Navier–Stokes) equations to simulate surface–plant–air interactions. The Boussinesq approximation was adopted for the thermal-forced vertical motion. The model description is detailed in Bruse and Fleer [31]. It has been used to study the atmospheric dispersion of air pollutants included in urban environments [32,33]. Particle sedimentation due to gravity and particle deposition to different surfaces by considering the aerodynamic and sub-layer surface resistances [34] were simulated. The simulation domain covered an area of 100 m × 100 m. A horizontal grid resolution was set as 2 m with 6 nesting grids at each border to avoid the edge effects. For vertical grids, the grid size varied from 20 cm in the first 1 m to a telescoping factor of 20% after a height of 1 m above ground.

The hourly wind speed, wind direction, and air temperature measured at the EPD's air quality monitoring station (AQMS) were adopted as the model inputs [or the inflow boundary condition (BC)]. The hourly measured relative humidity was obtained from a nearby weather station. The wind speed at 10 m above ground, as required by the model, was derived by the following power-law equation:

$$\frac{U_z}{U_{ref}} = \left(\frac{z}{z_{ref}} \right)^\alpha \quad (1)$$

while taking a roughness length α of 0.1 [35]. The BC for PM₁₀ was set to 0 µg m^{−3}, since PM₁₀ enhancement due to traffic was modeled. Other values for the BC for PM₁₀ were considered not appropriate since accurate BC values from measurements are not available. The resultant PM₁₀ levels reported here were the CFD-predicted PM₁₀ enhancement plus the measured background concentrations. A 24 h simulation was performed from 9:00 a.m. on 30 November to 8:00 a.m. on 1 December 2009. It was about the middle testing period of the ANN simulation performed above. A model spin-up of 6 h was used prior to the adoption of the CFD model outputs to avoid the influence from model initialization.

Daily traffic was obtained from the annual average data reported by the government's Transport Department at the roads of concern in 2009 [36]. The model's default diurnal profile of traffic for an urban road was assumed, with peak hourly daytime traffic flow contributing about 7%. The traffic data at the two major roads [Nathan Road (17,000 vehicles per day) and Lai Chi Kok Road (7000 vehicles per day)] near the AQMS was input into the model. The roads were the only major sources of PM₁₀ concentrations measured at the AQMS and were modeled as the line sources. The source height was 0.3 m above

the ground. An average emission factor of $105 \mu\text{g veh}^{-1} \text{m}^{-1}$ for PM_{10} [37], which was obtained from observations at different sites, was used.

The model settings are summarized in Table 2.

Table 2. The CFD model settings.

Parameters	Remarks/Values ¹
Meteorological conditions (wind speed, wind direction, relative humidity, and air temperature)	Hourly local measurements
Boundary condition for PM_{10}	$0 \mu\text{g m}^{-3}$, since PM_{10} enhancement due to traffic was modeled
Pollution source	Line sources with a height of 0.3 m above the ground
Source emission factor for PM_{10}	$105 \mu\text{g veh}^{-1} \text{m}^{-1}$

¹ See text for details.

3. Results and Discussion

3.1. Results of the ANN Model

Figure 3 shows the temporal variation in PM_{10} as predicted by the ANN model during the testing period. The model demonstrates a good performance ($r = 0.82$, $\text{FB} = 0.002$, $\text{RSME} = 15.4$, $\text{MAE} = 11.6$) and captures the diurnal cycles, the general trend from November to December, and some episodic levels (e.g., on 2 November and 1–4 December).

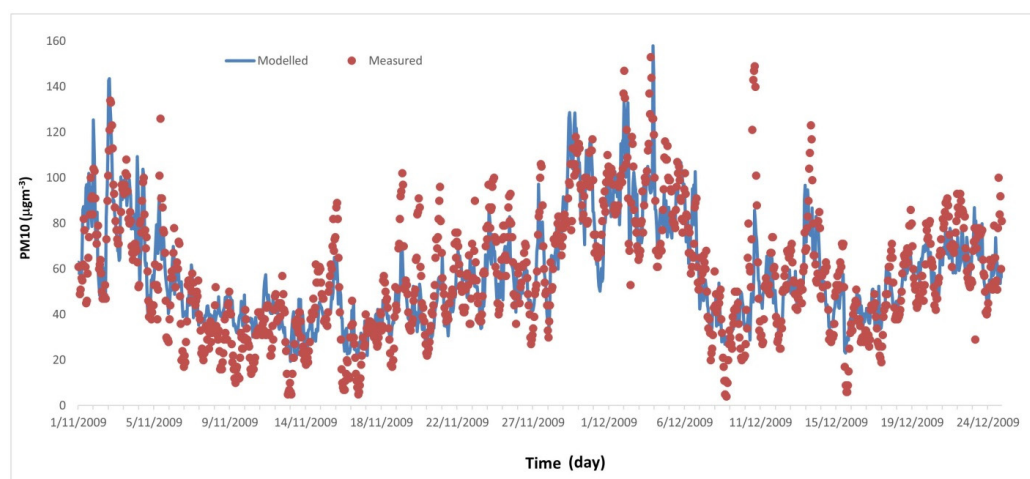


Figure 3. Comparison between the ANN modeling results and the measurements. Temporal variation in PM_{10} as predicted by the ANN model. The measured data are shown as circles.

A series of sensitive tests for the ANN model was performed to determine whether a single input parameter or a combination of them governed the model performance. Prior to the tests, a principal component analysis (PCA) was performed. The PCA results showed that the first four principal components (PCs) accounted for 74% of the total variance (Supplementary Material Table S1). One of the PCs (PC3) showed high loadings (>0.9) with the background PM_{10} and the predicted PM_{10} , suggesting a strong association between them. The ANN model construction using only the background PM_{10} as the input parameter could achieve a relatively good model performance ($r = 0.77$), when compared to the observations. At this point, the result of the PCA was consistent with that of the ANN model's sensitive test. However, an additional sensitive test by constructing an ANN model using in-canyon wind speed and in-canyon air temperature (essential parameters in the CFD simulation) showed a very poor model performance with $r = 0.25$. The poor performance might be attributed to the omission of the background PM_{10} levels. Nevertheless, our results suggested that the ML model could perform reasonably well

even without the knowledge of traffic data. Such simplification has a major benefit to the practical model application in a smart city, which is discussed in the subsequent sections.

3.2. Results of CFD Model

Figure 4a shows a typical traffic-induced PM_{10} horizontal distribution within the study area as predicted by the CFD model during peak hours. Higher PM_{10} levels near the road sources are clearly depicted under the influence of a weak, northeasterly wind ($<0.5 \text{ ms}^{-1}$). When compared to the area near Lai Chi Kok Road, the PM_{10} concentrations near Nathan Road are higher because of the higher traffic flow. Specifically, in the morning of 29 November, under the influence of a weak, northerly/northeasterly wind, the monitoring station and nearby areas were at downwind of Nathan Road (Figure 1) and, thus, had relatively high PM_{10} levels (Figure 4b) due to the impact of vehicular pollution plume. At earlier noontime, however, the decreasing PM_{10} levels at the monitoring station and nearby areas were mainly due to the change in wind direction (i.e., southwesterly at noontime) and enhanced vertical mixing with relatively clean air aloft. The CFD results showed that the PM_{10} enhancements due to road traffic during nighttime were very small at most of the areas within the domain ($<2 \mu\text{g m}^{-3}$) because of the low traffic flow. A detailed discussion of the pollution dispersion is not the aim of the current work. Figure 4b shows the diurnal variation in the measured PM_{10} concentrations, as well as the calculated PM_{10} concentrations by the CFD model and ANN model. Daytime-measured PM_{10} concentrations are higher than those at nighttime because of the lower traffic flow at nighttime. The lower measured concentration near noontime is attributed to stronger solar heating that promotes the vertical mixing of pollutants, given a relatively small variation in the daytime traffic flow. In general, the CFD model performs well (Table 3) and captures the temporal variation in the measured PM_{10} levels. Its good performance is likely due to the diurnal profile of traffic assumed in the model, hourly wind speed and direction as the input model boundary conditions, and simulated vertical mixing in the unstable atmosphere near noontime.

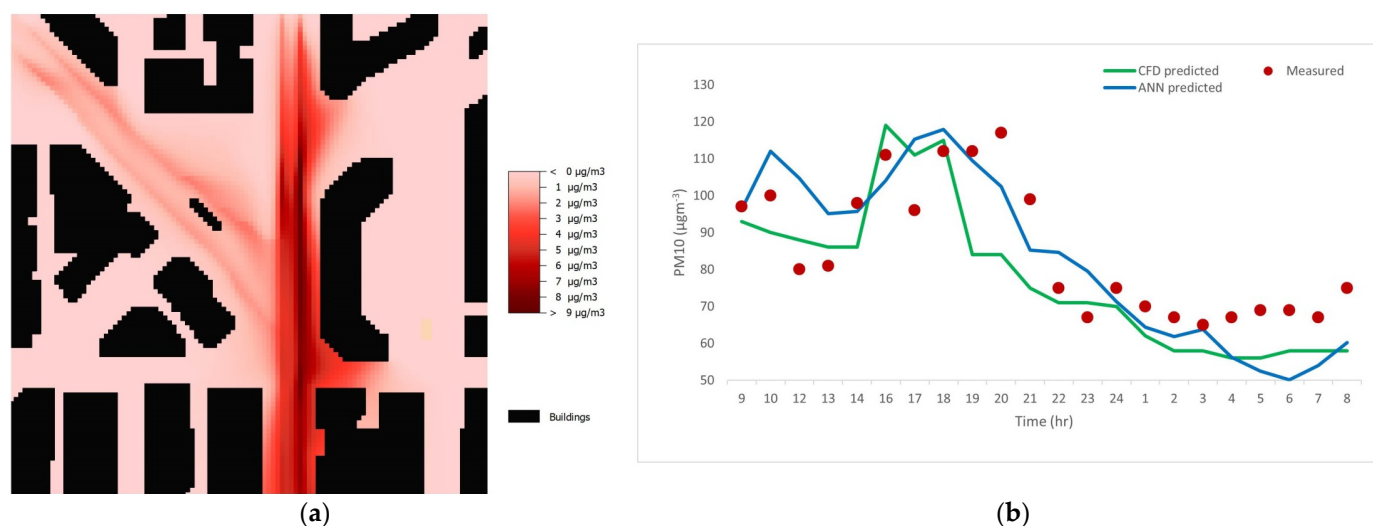


Figure 4. The CFD modeling results and comparison with the ANN modeling results and measurements. (a) Sample-predicted distribution of PM_{10} enhancement due to traffic (1.5 m above ground) by the CFD model during peak hours. (b) Comparison of the predicted PM_{10} diurnal variations by the CFD model and the ANN model. The measured data are shown as circles.

Table 3. Summary of model performance.

Model	R	FB	RMSE	MAE
ANN	0.84	0.02	12.2	10.4
CFD	0.81	0.09	13.7	11.3

The discrepancy in the CFD results for the prediction in the evening hours (1700–1900; Figure 4b) might be attributed to the considerable deviation in traffic flow between the real-time situation and the model's default profile. Except during 12:00–18:00, the CFD model shows an under-estimation of the measurements most of the time. This under-estimation has been reported elsewhere. Deng et al. [32] pointed out that the under-estimation was profound, especially during days with elevated particulate levels, although the model depicted similar temporal pattern in the measured pollution levels. For a pollution dispersion study from a motorway, De Maerschalc et al. [38] demonstrated a good agreement between the measurements and the modeling results for NO₂, but not for particulate levels.

While both the ANN and CFD models performed similarly in the PM₁₀ predictions studied above (Table 3), the computational time for the ANN model was less than 0.1% of the CFD model. Simulating a one-day hourly PM₁₀ variation by the ENVI-met required more than 30 wall-clock hours in parallel processing mode for a computer with four cores, while a ~50-day hourly PM₁₀ simulation by the ANN model required less than 30 wall-clock minutes using the same computer. To resolve the demanding computational resource and lengthy time required by CFD simulations, a plausible solution might be a fast-mathematical model with simplified equations for air quality predictions. However, it is well known that a simplified dispersion equation, such as a Gaussian-type equation, performs poorly in dispersion calculations in complicated built environments.

One of the major limitations for the CFD model (and other conventional physiochemical models) in simulating street-canyon air quality is the requirement of real-time traffic counting. Another limitation is that, in reality, it is very difficult to accurately obtain vehicular emission information for all vehicles on the roads in a simulation period. For instance, there are large uncertainties in the vehicular emissions adopted in the model when compared with reality. Actual information, such as emission standards (from EURO-III to EURO-VI), and additional mitigation measures (e.g., diesel particulate filter) fitted at the tailpipe for each vehicle are very difficult (if not impossible) to obtain during routine monitoring. Some studies adopted a vehicular emission model (e.g., Mobile 6 and COPERT4) to better mimic the variation in road traffic emissions and then to feed the information into an air dispersion model, including a CFD model [39–41]. However, this kind of model requires many inputs, such as fuel consumption, fleet configuration, trip length, distribution of vehicle miles traveled by road types, average speed distribution by road types, annual mileage, which are not available in many areas/countries; thus, large uncertainty in the simulated traffic emissions and, in turn, the air quality simulations results. This poses a challenge in using a vehicular emission model to obtain relatively accurate results for practical use in an urban environment.

Besides that, the good performance of the CFD model is likely due to the diurnal profile of traffic adopted and the use of hourly wind speed and direction as the model boundary conditions. On the contrary, many research efforts available in the literature, for the purpose of scenario simplification, adopted constant emissions and boundary conditions (e.g., for wind), without considering unstable atmospheric conditions. It demonstrates that the practical use of a CFD modeling technique as an air quality management tool for the urban neighborhood-scale air pollution problem is, in general, very difficult.

4. Conclusions

In this study, the ANN approach, as an ML algorithm, was used to make PM₁₀ predictions near road traffic emissions in the UCL. The performance of the ANN model

was further compared with the CFD model. Both the ANN and CFD models performed similarly when their predictions were compared with the measurements. However, the ANN model is much faster and requires less computational resources and fewer input parameters. The last factor might be critical in the context of air quality management for a smart city. For instance, acquisition of accurate real-time vehicle emission factors is difficult for CFD simulations, but traffic flow and emission factors are not required for the ANN model simulations based on the finding of the current study. These issues have been discussed in more details. Nevertheless, one of the strengths of the CFD model is that it provides the spatial dynamics of urban air pollution, which is difficult to obtain with the currently formulated ANN model.

The ANN model adopted in our study demonstrates its usefulness in air quality predictions, especially as a useful tool for smart city applications. It provides acceptable results in both neutral and unstable atmospheric conditions, whereas additional complicated model settings/assumptions are required for the CFD model to simulate the conditions in an urban environment. Nevertheless, the ANN model, like other ML models, is a so-called “black-box”, which has limited contribution to knowledge development of physical processes and interaction of the driving mechanisms related to dispersion within urban street canyons. This issue may need further research in future.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/ijerph20032412/s1>, Table S1: Principal component analysis for the association between input parameters.

Author Contributions: Conceptualization, methodology, validation, formal analysis, investigation, writing—original draft preparation, K.-M.W.; writing—review and editing, P.K.N.Y. and K.-M.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Please refer supplementary data of the study in Supplementary Material.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pope, C.A., III. Review: Epidemiological Basis for Particulate Air Pollution Health Standards. *Aerosol Sci. Technol.* **2000**, *32*, 4–14. [CrossRef]
2. WHO. Health Effects of Particulate Matter: Policy Implications for Countries in Eastern Europe, Caucasus and Central ASIA. 2013. Available online: https://www.euro.who.int/__data/assets/pdf_file/0006/189051/Health-effects-of-particulate-matter-final-Eng.pdf (accessed on 17 July 2022).
3. Huang, D.; Xu, J.; Zhang, S. Valuing the health risks of particulate air pollution in the Pearl River Delta, China. *Environ. Sci. Policy* **2012**, *15*, 38–47. [CrossRef]
4. Morawska, L.; Thomas, S.; Gilbert, D.; Greenaway, C.; Rijnders, E. A study of the horizontal and vertical profile of submicrometer particles in relation to a busy road. *Atmos. Environ.* **1999**, *33*, 1261–1274. [CrossRef]
5. Wai, K.M.; Tanner, P.A. Relationship between ionic composition in PM₁₀ and the synoptic-scale and mesoscale weather conditions in a south China coastal city: A 4-year study. *J. Geophys. Res.* **2005**, *110*, D18210. [CrossRef]
6. GovHK, Air Quality in Hong Kong. Available online: <https://www.gov.hk/en/residents/environment/air/airquality.htm> (accessed on 22 July 2022).
7. Hillman, B.; Unger, J. Editorial—The Urbanisation of Rural China. *China Perspect.* **2013**, *3*, 3. [CrossRef]
8. Croitoru, C.; Nastase, I. A state of the art regarding urban air quality prediction models. In *E3S Web of Conferences*; EDP Sciences: Les Ulis, France, 2018; p. 01010.
9. Hanna, S.R.; Brown, M.J.; Camelli, F.E.; Chan, S.T.; Coirier, W.J.; Hansen, O.R.; Huber, A.H.; Kim, S.; Reynolds, R.M. Detailed simulations of atmospheric flow and dispersion in downtown Manhattan: An application of five computational fluid dynamics models. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 1713–1726. [CrossRef]
10. Cimorelli, A.J.; Perry, S.G.; Venkatram, A.; Weil, J.C.; Paine, R.J.; Wilson, R.B.; Lee, R.F.; Peters, W.D.; Brode, R.W. AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization. *J. Appl. Meteorol.* **2005**, *44*, 682–693. [CrossRef]

11. Jittra, N.; Pinthong, N.; Thepanondh, S. Performance Evaluation of AERMOD and CALPUFF Air Dispersion Models in Industrial Complex Area. *Air Soil Water Res.* **2015**, *8*, 87–95. [\[CrossRef\]](#)
12. Nelson, M.; Addepalli, B.; Hornsby, F.; Gowardhan, A.; Pardyjak, E.; Brown, M. Improvements to a fast-response urban wind model. In Proceedings of the 15th Joint Conference on the Applications of Air Pollution Meteorology with the A&WMA, New Orleans, LA, USA, 15–19 November 2008.
13. Cao, X.; Tian, Y.; Shen, Y.; Wu, T.; Li, R.; Liu, X.; Yeerken, A.; Cui, Y.; Xue, Y.; Lian, A. Emission Variations of Primary Air Pollutants from Highway Vehicles and Implications during the COVID-19 Pandemic in Beijing, China. *Int. J. Environ. Res. Public Health* **2021**, *18*, 4019. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Chu, A.K.M.; Kwok, R.C.W.; Yu, K.N. Study of pollution dispersion in urban areas using Computational Fluid Dynamics (CFD) and Geographic Information System (GIS). *Environ. Model. Soft.* **2005**, *20*, 273–277. [\[CrossRef\]](#)
15. Houda, S.; Belarbi, R.; Zemmouri, N. A CFD Comsol model for simulating complex urban flow. *Energy Procedia* **2017**, *139*, 373–378. [\[CrossRef\]](#)
16. Wai, K.-M.; Yuan, C.; Lai, A.; Yu, P.K. Relationship between pedestrian-level outdoor thermal comfort and building morphology in a high-density city. *Sci. Total Environ.* **2020**, *708*, 134516. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Aflaki, A.; Esfandiari, M.; Mohammadi, S. A Review of Numerical Simulation as a Precedence Method for Prediction and Evaluation of Building Ventilation Performance. *Sustainability* **2021**, *13*, 12721. [\[CrossRef\]](#)
18. Feng, R.; Zheng, H.J.; Gao, H.; Zhang, A.R.; Huang, C.; Zhang, J.X.; Luo, K.; Fan, J.R. Recurrent Neural Network and random forest for analysis and accurate forecast of atmospheric pollutants: A case study in Hangzhou, China. *J. Clean. Prod.* **2019**, *231*, 1005–1015. [\[CrossRef\]](#)
19. Bozdağ, A.; Dokuz, Y.; Gökçek, Q.B. Spatial prediction of PM₁₀ concentration using machine learning algorithms in Ankara, Turkey. *Environ. Poll.* **2020**, *263*, 114635. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Xu, G.; Ren, X.; Xiong, K. Analysis of the driving factors of PM_{2.5} concentration in the air: A case study of the Yangtze River Delta, China. *Ecol. Indic.* **2020**, *110*, 105889. [\[CrossRef\]](#)
21. Krishan, M.; Jha, S.; Das, J.; Singh, A.; Goyal, M.K.; Sekar, C. Air quality modelling using long short-term memory (LSTM) over NCT-Delhi, India. *Air Qual. Atmos. Health* **2019**, *12*, 899–908. [\[CrossRef\]](#)
22. Qin, Z.; Chen, C.; Guo, X. Prediction of Air Quality Based on KNN-LSTM. *J. Phys. Conf. Ser.* **2019**, *1237*, 042030. [\[CrossRef\]](#)
23. Wang, Y.; Du, Y.; Wang, J.; Li, T. Calibration of a low-cost PM_{2.5} monitor using a random forest model. *Environ. Int.* **2019**, *133A*, 105161.
24. Saxena, A.; Shekhawat, S. Ambient air quality classification by grey wolf optimizer based support vector machine. *J. Environ. Public Health* **2017**, *2017*, 3131083. [\[CrossRef\]](#)
25. Bai, Y.; Li, Y.; Wang, X.; Xie, J.; Li, C. Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions. *Atmos. Poll. Res.* **2016**, *7*, 557–566. [\[CrossRef\]](#)
26. Ma, D.; Zhang, Z. Contaminant dispersion prediction and source estimation with integrated Gaussian-machine learning network model for point source emission in atmosphere. *J. Hazard. Mater.* **2016**, *311*, 237–245. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Bishop, C.M. *Neural Networks for Pattern Recognition*; Clarendon Press: Oxford, UK, 1995.
28. Yang, J. Intelligent Data Mining Using Artificial Neural Networks and Genetic Algorithms: Techniques and Applications. 2010. Available online: http://wrap.warwick.ac.uk/3831/1/WRAP_THESIS_Yang_2010.pdf (accessed on 17 July 2022).
29. Nielsen, R. The backpropagation neural network. *Int. Jt. Conf. Neural Netw.* **1989**, *1*, 593–605.
30. Azid, A.; Juahir, H.; Latif, M.T.; Zain, S.M.; Osman, M.R. Feed-Forward Artificial Neural Network Model for Air Pollutant Index Prediction in the Southern Region of Peninsular Malaysia. *J. Environ. Prot.* **2013**, *4*, 1–10. [\[CrossRef\]](#)
31. Bruse, M.; Fleer, H. Simulating surface–plant–air interactions inside urban environments with a three dimensional numerical model. *Environ. Model. Softw.* **1998**, *13*, 373–384. [\[CrossRef\]](#)
32. Deng, S.; Ma, J.; Zhang, L.; Jia, Z.; Ma, L. Microclimate simulation and model optimization of the effect of roadway green space on atmospheric particulate matter. *Environ. Poll.* **2019**, *246*, 932–944. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Taleghani, M.; Clark, A.; Swan, W. Air pollution in a microclimate; the impact of different green barriers on the dispersion. *Sci. Total Environ.* **2020**, *711*, 134649. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Bruse, M. Particle filtering capacity of urban vegetation: A microscale numerical approach. *Berl. Geogr. Arb.* **2007**, *109*, 61–70.
35. Wai, K.M.; Tan, T.Z.; Morakinyo, T.E.; Chan, T.C.; Lai, A. Reduced effectiveness of tree planting on micro-climate cooling due to ozone pollution—A modeling study, *Sustain. Cities Soc.* **2020**, *52*, 101803. [\[CrossRef\]](#)
36. TD. The annual traffic census—2009, Transport Department of the Hong Kong Special Administrative Region Government. 2009. Available online: https://www.td.gov.hk/en/publications_and_press_releases/publications/free_publications/the_annual_traffic_census_2009/index.html (accessed on 20 January 2023).
37. Ketzel, M.; Omstedt, G.; Johansson, C.; Düring, I.; Pohjola, M.; Oetl, D.; Gidhagen, L.; Wåhlin, P.; Lohmeyer, A.; Haakana, M.; et al. Estimation and validation of PM_{2.5}/PM₁₀ exhaust and non-exhaust emission factors for practical street pollution modelling. *Atmos. Environ.* **2007**, *41*, 9370–9385. [\[CrossRef\]](#)
38. Maerschalc, B.; Janssen, S.; Vankerkom, J.; Mensink, C.; van den Burg, A.; Fortuin, P. CFD simulations of the impact of a line vegetation element along a motorway. In Proceedings of the 12th Conference on Harmonisation Within Atmospheric Dispersion Modelling for Regulatory Purposes (HARMO12), Cavtat, Croatia, 6–9 October 2008.

39. Librando, V.; Tringali, G.; Calastrini, F.; Gualtieri, G. Simulating the production and dispersion of environmental pollutants in aerosol phase in an urban area of great historical and cultural value. *Environ. Monit. Assess.* **2009**, *158*, 479–498. [[CrossRef](#)]
40. Potoglou, D.; Kanaroglou, P.S. Carbon monoxide emissions from passenger vehicles: Predictive mapping with an application to Hamilton, Canada, *Transp. Res. D Transp. Environ.* **2005**, *10*, 97–109. [[CrossRef](#)]
41. McAlpine, J.D.; Ruby, M. Using CFD to Study Air Quality in Urban Microenvironments. In *Environmental Sciences and Environmental Computing*; Zannetti, P., Ed.; The EnviroComp Institute: Fremont, CA, USA, 2004; Volume II, Chapter 1.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.