



Identifying critical building morphological design factors of street-level air pollution dispersion in high-density built environment using mobile monitoring



Yuan Shi^{a,*}, Xiaolin Xie^b, Jimmy Chi-Hung Fung^{b,e}, Edward Ng^{a,c,d}

^a School of Architecture, The Chinese University of Hong Kong, Shatin, NT, Hong Kong Special Administrative Region, China

^b Division of Environment and Sustainability, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong Special Administrative Region, China

^c Institute of Environment, Energy and Sustainability (IEES), The Chinese University of Hong Kong, Shatin, NT, Hong Kong Special Administrative Region, China

^d Institute of Future Cities (IOFC), The Chinese University of Hong Kong, Shatin, NT, Hong Kong Special Administrative Region, China

^e Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong Special Administrative Region, China

ARTICLE INFO

Keywords:

Air pollution dispersion
Mobile monitoring
Building morphology
Planning optimization

ABSTRACT

In high-density cities, optimization of their compact urban forms is important for the enhancement of pollution dispersion, improvement of the air quality, and healthy urban living. This study aims to identify critical building morphological design factors and provide a scientific basis for urban planning optimization. Through a long-term mobile monitoring campaign, a four-month (spanning across summer and winter seasons) spatiotemporal street-level PM_{2.5} dataset was acquired. On top of that, the small-scale spatial variability of PM_{2.5} in the high-density downtown area of Hong Kong was mapped. Seventeen building morphological factors were also calculated for the monitoring area using geographical information system (GIS). Multivariate statistical analysis was then conducted to correlate the PM_{2.5} data and morphological data. The results indicate that the building morphology of the high-density environment of Hong Kong explains up to 37% of the spatial variability in the mobile monitored PM_{2.5}. The building morphological factors with the highest correlation to PM_{2.5} concentration are building volume density, building coverage ratio, podium layer frontal area index and building height variability. The quantitative correlation between PM_{2.5} and morphological factors can be adopted to develop scientifically robust and straightforward optimization strategies for planners. This will allow considerations of pollution dispersion to be incorporated in planning practices at an early stage.

1. Introduction

Air pollution has been identified as a major problem in high-density cities in Asia [1]. Urbanization physically changes the natural landscape into a highly artificial built environment [2]. In a high-density city environment, closely packed building groups weaken air flows and consequently limit the dispersion of pollutants [3,4]. Therefore, street-level air pollution has become a severe environmental issue in high-density cities, such as Hong Kong [5]. The PM_{2.5} concentration level monitored by roadside stations shows that the air quality of Hong Kong does not fulfill the requirements of either the local air quality objectives or other international air quality standards [6]. In Hong Kong, many public health investigations have shown that air pollution are strongly connected to adverse health outcomes. For every 10 µg/m³ increase in the daily average concentration level of PM_{2.5}, there will be

approximately 2% more hospitalization and 2% increase in the mortality due to respiratory diseases alone [7,8]. Under such circumstances, the Environment Bureau of Hong Kong released “A Clean Air Plan” for Hong Kong in 2013, with the reduction of roadside air pollution as a major focus [9].

Enhancing the rate of pollution dispersion is an effective way to reduce its concentration [10]. A properly planned/designed urban morphology will significantly improve pollution dispersion [11], and thereby reduce the health risk of exposure. Under such context, academic research and the planning practice are increasingly focusing on enhancing pollution dispersion in cities [12]. A wide range of techniques has been used to monitor or model street-level pollutant concentrations and human exposure in the built environment [13,14]. Most current methods on pollution dispersion in an urban environment are based on complex numerical simulations [15–17]. They are advanced

* Corresponding author.

E-mail addresses: shiyuan@cuhk.edu.hk, shiyuan.arch.cuhk@gmail.com (Y. Shi).

¹ Postal addresses: Rm505, AIT Building, School of Architecture, The Chinese University of Hong Kong, Shatin, NT, Hong Kong SAR, China.

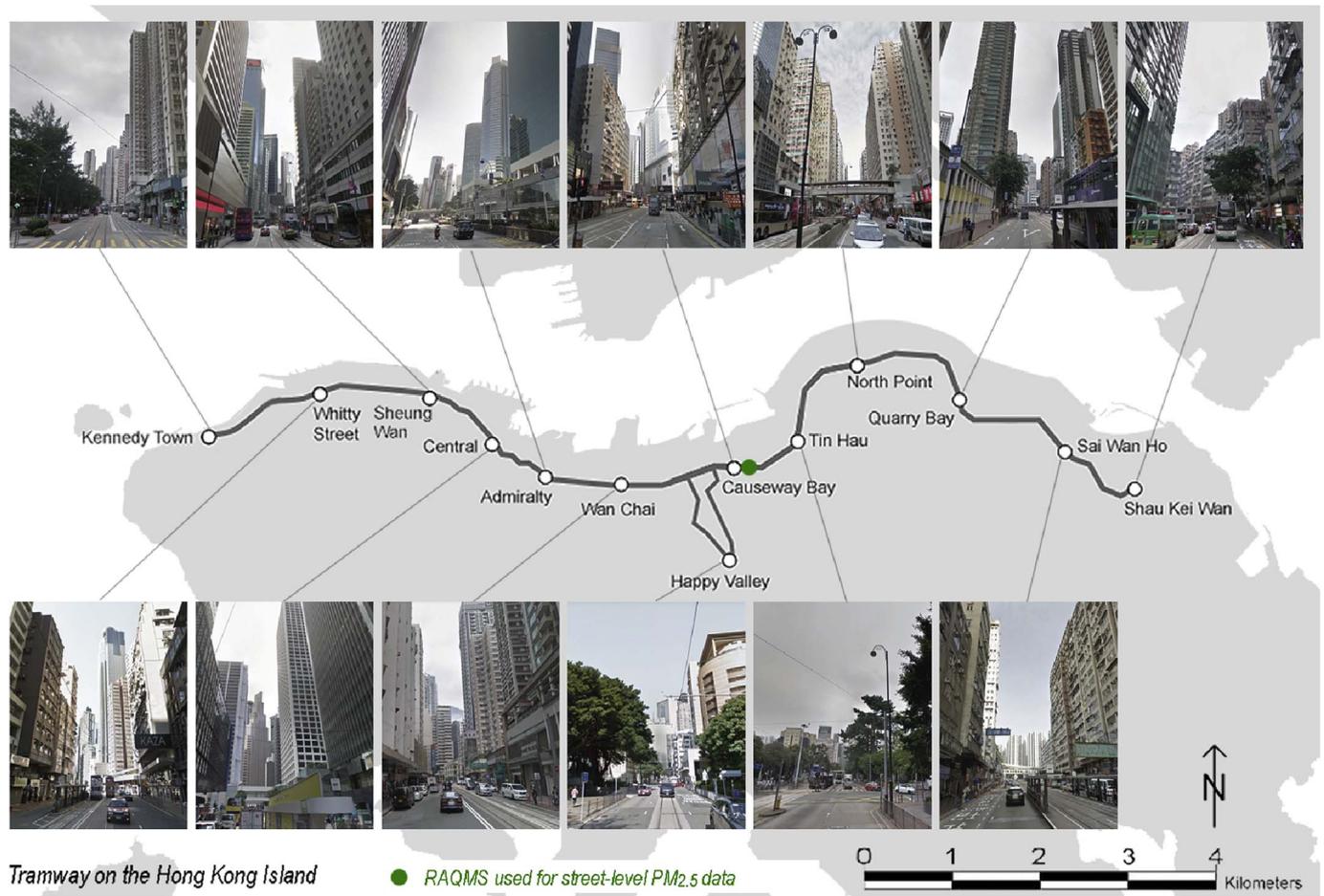


Fig. 1. Tramway on the Hong Kong Island and the building morphology of along the tram route.

and accurate, but too complicated and time-consuming to help planners and practitioners optimize the planning scheme at an early stage efficiently. For example, in the practical planning process of Hong Kong, planners need straightforward information of reasonable accuracy and quick methods at the initial strategic planning stage of urban renewal and new development areas (NDAs) projects.

It has been indicated that the densely built urban form of Hong Kong is not optimized for pollution dispersion [18,19]. It blocks ventilation and consequently retards the dispersion [3,19]. The tall building clusters and narrow roads result in deep street canyons with intensive traffic flows and high pollutant emission intensity. Besides traffic-related air pollution, many non-vehicular $PM_{2.5}$ pollution sources at the roadside [20] (such as shops, bus stops, parking entrance, cargo areas of shopping malls, and ventilation discharge outlets of restaurants/commercial cooking [21,22]) also contribute to the problem. They all emit an enormously high intensity of $PM_{2.5}$ and are a significant contribution to the street-level air pollution. However, pollution dispersion as a dimension of air pollution mitigation is not commonly considered in the daily urban planning/design practice of Hong Kong due to the lack of easy-to-use design method and practical guidance. Therefore, it is important to obtain a scientifically robust but more straightforward understanding of how to optimize urban planning for better pollution dispersion in the high-density urban context. This study focuses on quantitatively investigating the dispersion capability of different morphological configurations along the street canyons and identifying critical building morphological design factors for the development of practical planning optimization strategies. This will allow considerations of pollution dispersion to be incorporated at an early stage in the planning practice. Considering the above, $PM_{2.5}$

(particulate matters with an aerodynamic diameter $< 2.5 \mu m$, a commonly-used proxy to investigate pollution dispersion [23]), was used as a comprehensive marker to quantitatively represent the dispersion capability (of both traffic-related and non-traffic air pollution) along the street canyons.

To resolve the effects of building morphological factors on pollution dispersion, information of small-scale spatial variability of street-level air pollution needs to be observed at a very fine spatial scale. In Hong Kong, the heterogeneous building morphology and complicated traffic network make the street-level air quality vary vastly between different locations. Therefore, small-scale spatial variability of air pollution is impossible to be effectively observed using data from the only a couple of fixed roadside air quality monitoring stations (RAQMS) in Hong Kong. Mobile monitoring as a cost-effective way to cover larger study areas has been gaining popularity in air pollution research [24–26] due to its advantage of fine spatial coverage. The method uses a vehicle as a platform and its feasibility has been tested in a pilot study of mapping the spatial distribution of street-level $PM_{2.5}$ in the downtown area of Hong Kong [19]. However, the two-week dataset measured by that study possibly contains uncertainties, as the monitoring time at each position is limited. As a consistent mode of public transport of Hong Kong, trams continuously run along some fixed routes in the high-density downtown area over a long period of time. Thus, a much larger dataset can be obtained than the vehicle-based monitoring platform. It has been indicated that increasing the size of mobile monitoring dataset can greatly decrease the uncertainties in the mapping of the spatial distribution of air pollution [27]. Hence, by monitoring the air quality continuously for a long period of time on a tram, the abovementioned limitations can be overcome and the robustness of the monitoring

results can be improved.

2. Materials and methods

In this present study, a spatiotemporal street-level $PM_{2.5}$ dataset was acquired by long-term mobile monitoring using the tram. It was to resolve the small-scale spatial variability of $PM_{2.5}$ in the high-density downtown area of Hong Kong (the northern part of Hong Kong Island). A set of building morphological design factors was calculated for the areas along the tram routes using GIS. Multivariate statistical analysis was performed to investigate the correlation between the $PM_{2.5}$ spatial data and morphological data and identify critical building morphological factors of pollution dispersion in the urban context of Hong Kong.

2.1. Mobile monitoring of street-level $PM_{2.5}$ using tram

2.1.1. Measurement routes and campaign

Street-level $PM_{2.5}$ measurement was made when a tram was in service according to its normal day schedule. The fixed routes on the northern side of the Hong Kong Island contain a RAQMS of the Hong Kong Environmental Protection Department (Fig. 1), and the $PM_{2.5}$ data from these stations were used to compare and calibrate the street-level $PM_{2.5}$ measurements when the tram passed by. The measurement campaign started in August 2013 and continues to the day of writing. Measurement was made when the tram was in its normal business service in town.

2.1.2. Instrumental setup

Collaborating with staff of Hong Kong Tramways Company, a $PM_{2.5}$ measurement unit was assembled and installed on a tram vehicle (Fig. 2). The measurement unit is composed of an optical aerosol monitor (DustTrak DRX, TSI) with an auto-zero module for $PM_{2.5}$ measurement and a GPS to locate the tram as it moved. Auto-zeroing of DustTrak was performed every 3 h to minimize the impact of instrument drifting to the measurement. A preprogrammed data logger was used to control the operation of the system and archive the high time-resolution $PM_{2.5}$ and GPS data which were obtained at a frequency of 1 Hz. The whole system was contained in a metal-casing installed underneath a seat at the rear end of the upper deck (Fig. 2-a). Powered by the tram's DC supply and connected to DustTrak through a conductive tubing, the system sampled ambient air from a water-drain hole on the upper deck (3 m from ground, referring to Fig. 2-b).

2.1.3. Data quality control

The DustTrak was checked and compared with the regular $PM_{2.5}$ measurements made at the HKUST Air Quality Research Supersite with filter-based method and/or on-line FEM (Federal Equivalent Method) instrument before it was deployed. Regular system maintenance/checking and data download ($PM_{2.5}$ /GPS data together with the

control/performance parameters of the instruments) were performed when the tram was back to depot for servicing, normally at a 8-working-day interval. In the beginning trial phase, measurement was conducted every day from 10:00 to 19:00 with auto-zeroing every 3 h to test out the system. In the normal operational phase, measurement was performed every day after depot servicing from 7:00 to 19:00 with auto-zeroing every 3 h to cover all the normal day hours when the tram was in service. Therefore, the mean pollution concentration value measured at each location robustly reflects the long-term average level. The mobile monitored dataset can depict the small-scale spatial variability of $PM_{2.5}$ without any uncertainties introduced by temporal variations.

To be more specific, when the tram moves from the west end of Hong Kong Island to the east end, the whole journey takes approximately 1 h. Within that hour, the background concentration of $PM_{2.5}$ normally varies very little (typically less than $5 \mu\text{g}/\text{m}^3$, based on the data from background monitoring stations of Hong Kong). However, the measured $PM_{2.5}$ concentration along the tram route during one particular journey can vary by as much as $40 \mu\text{g}/\text{m}^3$ or even more in a relatively short distance of less than 100 m (based on the mobile measurement data, Fig. 5). The extent of spatial variations of concentration reflects that they are not caused by the background variation. The GPS data were checked/validated against the monthly time-location (stop location) of the tram provided by Hong Kong Tramways Limited.

The linear regression assumption is common and has been used in several previous studies using DustTrak in Hong Kong [28–30]. In one of them [28], the collocating data from DustTrak (5min) and Reference method (1 h) at Causeway Bay RAQMS were used. The results show a good correlation of $R = 0.91$. In this study, we used the observations from the Causeway Bay RAQMS to calibrate the tram measurements. The tram measurements represent the polluted situation at the lower level of the street canyon. The $PM_{2.5}$ inlet of the Causeway Bay RAQMS is installed 3 m above ground, around the same height as the instrument on the tram. To calibrate the tram measurements with the observations from Causeway Bay RAQMS, the tram data collected between the stations of 51E Percival Street and 53E Paterson Street were extracted. The raw tram data were measured second by second, while the provided observations from Causeway Bay RAQMS were hourly data. The monthly averages within the study period of both the tram data and RAQMS observations were then calculated for calibration purpose. Monthly averages were used instead of hourly averages because the quantity of data collected by the tram in an hour was too small to be representative for calibration purpose.

Firstly we divided the monthly average of the tram-based $PM_{2.5}$ data by the monthly RAQMS observations to obtain a factor for each month of the study period. The result varies from 1.43 to 2.83. Then we applied the optimal method to search for the best factor with the least difference between the observations and tram data after calibration

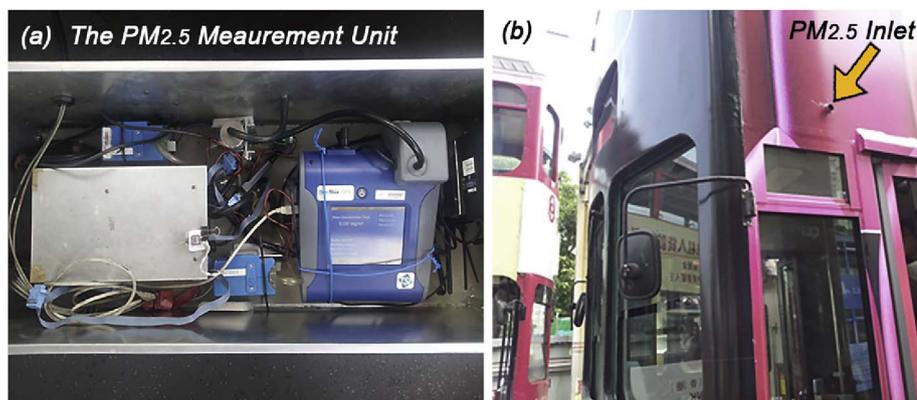


Fig. 2. (a) Instruments in the casing on tram – the data logger on the left and DustTrak DRX on the right. The GPS is attached on the outer wall of the casing. (b) The $PM_{2.5}$ inlet on the tram.

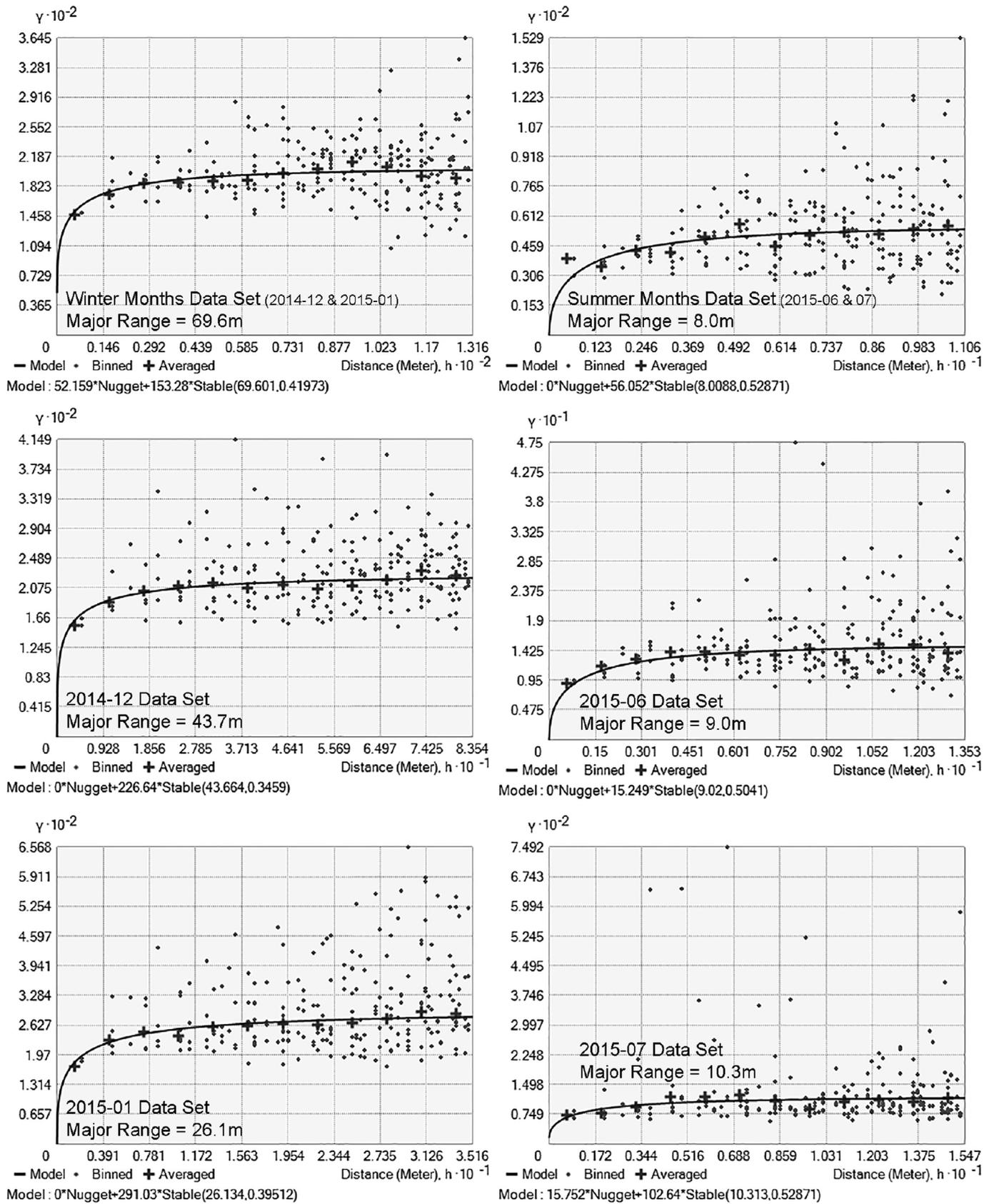


Fig. 3. The finalized empirical semivariogram models and the corresponding major ranges of the PM_{2.5} datasets.

(i.e., the tram data over the fixed factor) based on the monthly tram measurement and observations. A value of 1.91 was determined to be the optimal calibration factor, basically consistent with a prior study in Hong Kong using Mong Kok RAQMS as the calibration reference [19]. The above method is illustrated by the following equations:

$$CF_i = \frac{PM_{2.5i, Tram}}{PM_{2.5i, RAQMS}} \tag{1}$$

where the CF_i is the calibration factor of the i th month. $PM_{2.5i, Tram}$ and $PM_{2.5i, RAQMS}$ are the monthly averaged $PM_{2.5}$ concentration of the i th month by tram and RAQMS respectively. Find F_j , so that $D_j = \min(D_1, D_2, \dots, D_{1000})$, $j = 1, 2, \dots, 1000$.

$$F_j = \frac{j}{100} \tag{2}$$

$$D_j = \sum_{i=1}^n \left(\frac{PM_{2.5i, Tram}}{F_j} - PM_{2.5i, RAQMS} \right)^2 \tag{3}$$

where n is the total number of the monthly averaged value used for the calibration. F_j is the j th calibration factor. D_j is the square of difference for the j th calibration factor.

2.2. Tram-based $PM_{2.5}$ data processing

2.2.1. Collating data in GIS

The raw dataset collected by the $PM_{2.5}$ measurement unit mainly contains two parts – $PM_{2.5}$ concentration level measured every second and the synchronously recorded geographical locations (GPS data). These two parts of data were collated according to the time stamps and imported into GIS with the HK1980 coordinate system for further processing and analysis. In this present paper, a four-month dataset—2014-12, 2015-01, 2015-06 and 2015-07—was extracted and analyzed. Considering the differences in the dominant air pollution modes between seasons [31] and to analyze the resulting variation, we also divided the data into two seasonal datasets (summertime and wintertime).

2.2.2. Determining a suitable spatial scale for data aggregation

A properly determined spatial scale for data aggregation is important for the reduction of uncertainties in geographical analysis, especially when it involves a large dataset like the one in this study (which is a spatiotemporal dataset with 1-s temporal resolution for four months). In air pollution mapping studies, over-aggregated data introduce bias in regression analysis [32] and can possibly lead to over-estimation of the correlation coefficient in the regression analysis. Following the method used by Lightowlers, Nelson, Setton and Keller [33], the semivariogram method was adopted to determine the spatial scale for the data aggregation of tram-based $PM_{2.5}$ data in this present study. The semivariogram modelling has been adopted to inform the appropriate spatial scale for many spatial analysis methods (e.g. hotspot analysis, kriging/cokriging interpolation) in air pollution and health studies. A search radius or the neighbourhood kernel is usually defined as a parameter [34]. An optimized semivariogram function is essential for the determination of this parameter to avoid misleading conclusions associated with inappropriate spatial aggregation [35]. The semivariogram model used in geographical data analysis is defined as a function of distance as shown in the following equation:

$$\hat{\gamma} = \frac{1}{2n(d)} \sum_{s_i-s_j=d} (z_i - z_j)^2 \tag{4}$$

where $\hat{\gamma}$ is the semivariogram. The spatial points s_i and s_j are paired in a semivariogram modelling. z_i and z_j are the measured data of s_i and s_j . d is the distance between s_i and s_j . $n(d)$ is the amount of the pairs of all spatial points [34]. In a dataset of a group of spatially-distributed data points, the semivariogram value keeps increasing with the distance

until a limit defined as the sill ($\sigma(0)$). By calculating the semivariogram and developing the best fitting semivariogram function, the range (r), as a parameter of the empirical semivariogram model, can be determined based on the following equation (there are different semivariogram model types, such as the spherical, exponential and Gaussian model. In this present study, as shown in Fig. 3, all optimized models have the stable semivariogram model type. Therefore, the function of the stable semivariogram is shown here as an example of how an appropriate spatial scale was determined basing on the semivariogram model).

$$\hat{\gamma}(h; \theta) = \sigma(0) \left[1 - \exp\left(-\frac{3h^\omega}{r^\omega}\right) \right] \quad 0 < \omega \leq 2\theta = [\sigma(0) r \omega] \tag{5}$$

where h is the lag distance of the corresponding $\hat{\gamma}(h; \theta)$. ω and θ are model parameters. A more detailed procedure of semivariogram modelling has been described in a prior geostatistical study [36]. The corresponding h at which 95% of $\sigma(0)$ reached is determined as the major range (also named semivariogram range). The geographical meaning of a semivariogram model is that the data points within the major range are spatially correlated, while the data points beyond the major range are independent of each other. Semivariogram modelling is commonly adopted to deal with the spatial dependence/autocorrelation issues of spatially distributed observation points [37]. It has been used as a method in the determination of the spatial resolution of air quality mapping [19,38].

The ArcGIS software was used as a tool for all geo-spatial analysis in this study and the instruction of the semivariogram algorithm in the following literature was referred to for a reliable modelling [39,40]. As a result, six stable type semivariogram models were modelled for the datasets of four months (2014-12, 2015-01, 2015-06 and 2015-07). As mentioned, the data were also divided into two seasonal datasets (summer and winter) for semivariogram modelling. The resultant empirical semivariogram models and their major ranges are shown in Fig. 3. The results show that the data measured during wintertime have larger major ranges (from 26.1 m to 69.6 m with an average level of 46.5 m) than summertime data. This result indicates that summertime datasets provide spatial information at a finer spatial scale (ranges from 8.0 m to 10.3 m with an average level of 9.1 m) and can profile very short-range variations. This is possibly due to the lower impact of regional pollution during summertime. The variability in locally emitted air pollutants becomes clearer to be observed as a result.

2.2.3. Spatial aggregation of the $PM_{2.5}$ data

Considering the differences in the spatial independence between the summertime and wintertime datasets, different spatial scales were used for data aggregation. Based on the results of the semivariogram modelling, a group of points was firstly created on the tram route using a fixed spatial interval (10 m for summertime data and 50 m for wintertime data, Fig. 4). All observations within a search radius (radii) of each point (radii = 5 m for summertime data and radii = 25 m for wintertime data) were then aggregated to the corresponding point using the mean concentration value. Fig. 5 shows the seasonal average spatial distribution of the street-level $PM_{2.5}$ concentration based on the aggregated dataset.

2.3. Analyzing the building morphology along the tramway

2.3.1. Calculating the building morphological factors

To depict the current building morphological design features in the monitored area, a total of 17 building morphological factors was calculated. They include the mean and standard deviation of building height (h)/ground coverage ratio (λ_p), building volume density (BVD), sky view factor (of the entire hemispherical sky view and its eight sectors respectively, Ψ_{sky}) and three layers of frontal area index (λ_F , the values of the 16 wind directions and the weighted value based on the

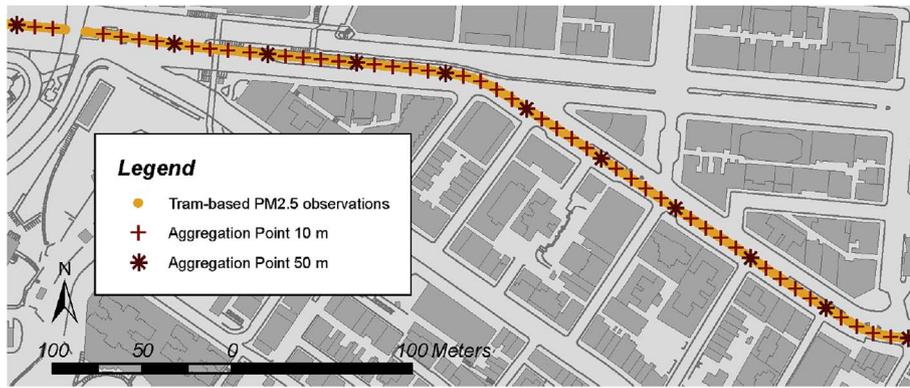


Fig. 4. The aggregation points along the tramway route generated in GIS.

probability of each direction for the two seasons) (Fig. 6 and Table 1). A previous study found that the total frontal area index of all layers was related to the concentration level of many air pollutants [41]. In this study, we further divided this factor into three layers at different heights (the podium layer between 0 and 15 m, the building tower layer between 15 and 60 m and the total layer between 0 and 500 m) to cater for the typical building structure of Hong Kong [42].

2.3.2. Neighboring analysis of the building morphological factors for data aggregation points

The street-level PM_{2.5} concentration at each monitoring point is influenced by the building morphological condition in its surrounding area. The neighboring analysis is composed of two steps: (1) creating buffers, (2) sensitivity test of critical buffer identification. First, the

buffering analysis method was used in this study. As mentioned in section 2.2.3, the PM_{2.5} observations were aggregated into a group of points on the tram route based on the spatial aggregation scales determined by the semivariogram modelling (Section 2.2.2). A series of buffers (with radii of 50 m, 100 m, 200 m, 300 m, 400 m and 500 m) was created around each data aggregation point (Fig. 7).

When using building morphological factors as the predictor variables to explain the variation in street-level PM_{2.5} observations, the critical buffer widths of different building morphological factors may vary due to the physical basis of pollution dispersion. Geographically, a building morphological feature measured by a specific factor within its critical buffers explains the variation of pollution to the greatest extent. Therefore, sensitivity test was conducted for each building morphological factor to determine its critical buffer width in explaining the PM_{2.5}

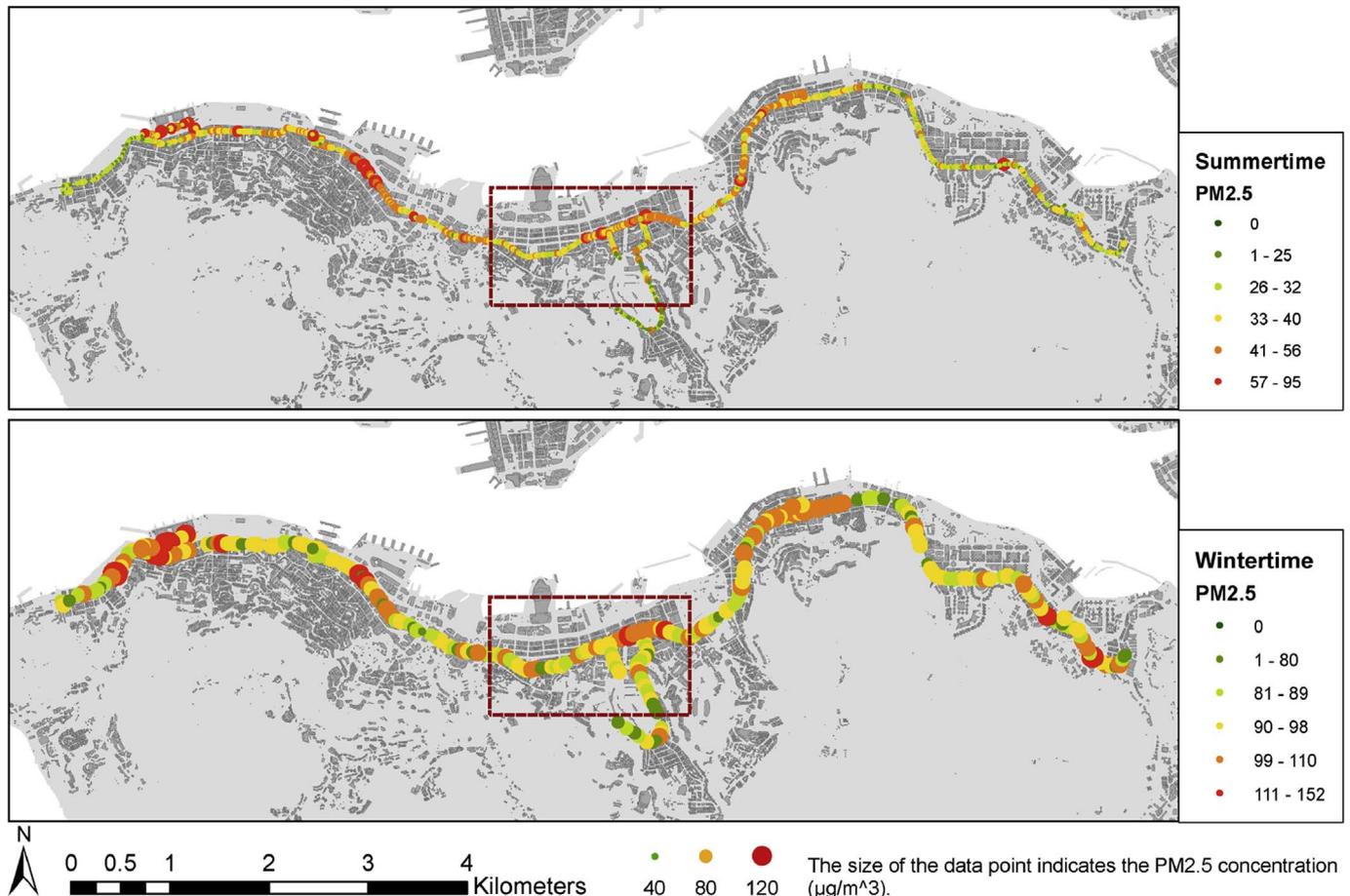


Fig. 5. The spatial plot of the seasonal averaged street-level PM_{2.5} concentration along the measurement routes. The spatial variation of the measured PM_{2.5} concentration and building morphological factors in the range of the inset boundary (dashed box) is plotted in Fig. 9.

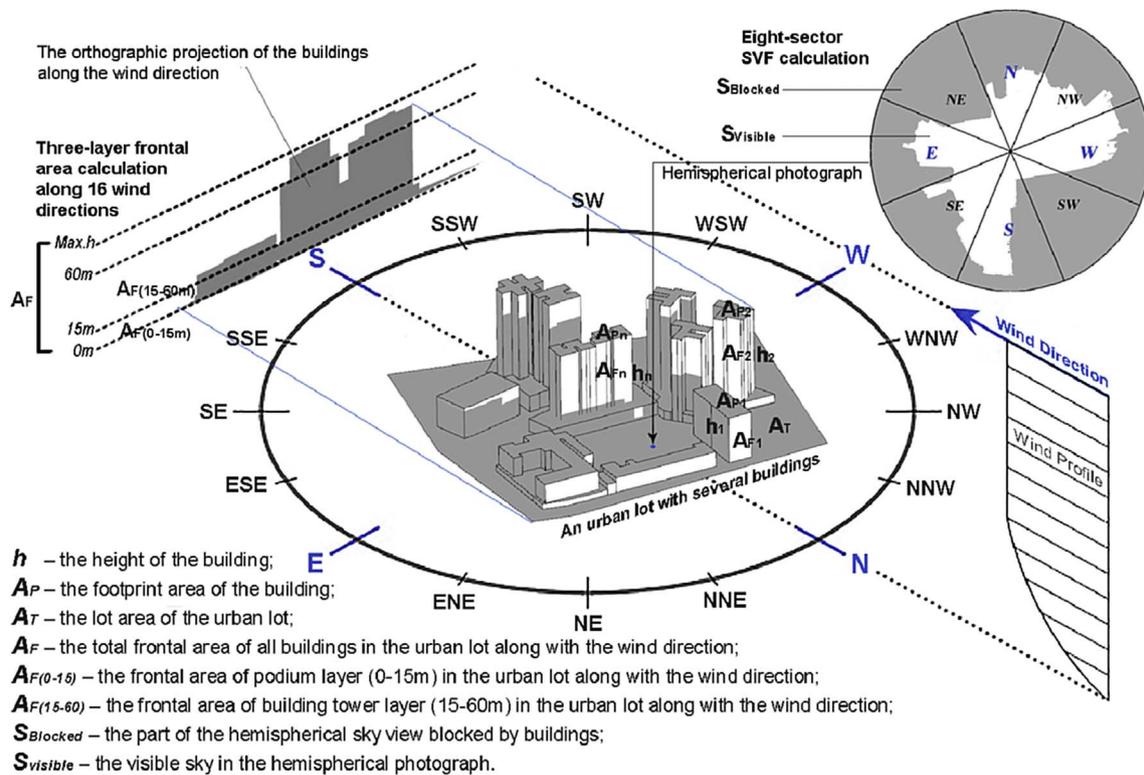


Fig. 6. A schematic diagram of calculating the building morphological factors (using an example of a street block in the North Point, a street-level PM_{2.5} concentration hotspot).

Table 1
The equations used in the calculation of building morphological factors (improved from [19]).

Building Morphological Factor	Unit	Equation of Calculation	Theoretical Meaning
Mean of building height	m	$\bar{h} = \frac{1}{n} \sum_{i=1}^n h_i$	(6) Vertical building development intensity.
σ of building height	m	$\sigma_h = \sqrt{\frac{1}{n} \sum_{i=1}^n (h_i - \bar{h})^2}$	(7) Diversity of building height within a specific area.
Building coverage ratio	% ^a	$\lambda_p = (\sum_{i=1}^n A_{p_i}) / A_T$	(8) Building ground coverage intensity.
σ of the ground coverage ratio of all building clusters in a specific area	%	$\sigma_{\lambda_p} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\lambda_{p_i} - \bar{\lambda}_p)^2}$	(9) Diversity of building coverage within a specific area.
Building volume density	%	Total building volume of each lot is: $V = \sum_{i=1}^n A_{p_i} h_i$ V_{max} is the highest V among all j lots whole city. The building volume density of lot j is: $BVD_j = V_j / V_{max}$	(10) BVD is a percentage value for reflecting the spatial distribution of the building density in a study area. (11)
Sky view factor (SVF) ^b	[0-1]	A detailed formula by Dozier and Frew [43] $\Psi_{sky} = \frac{1}{2\pi} \int_0^{2\pi} [\cos \beta \cos^2 \varphi + \sin \beta \cdot \cos(\Phi - \alpha) \cdot (90 - \varphi - \sin \varphi \cos \varphi)] d\Phi$	(12) A measure of the openness to the sky of a given location, Please see reference [44] for a more detailed description. In this study, the hemispherical sky view was also divided into eight sectors for sector-SVF calculation.
Frontal area index ^c – Total (~500 m)	C	$\lambda_F = A_F / A_T$	(13) A wind direction – dependent measure of the horizontal permeability.
Frontal area index – Podium Layer (0-15 m)	C	$\lambda_{F(0-15m)} = A_{F(0-15m)} / A_T$	(14) The horizontal permeability at the podium layer of Hong Kong.
Frontal area index – Building Tower Layer (15-60 m)	C	$\lambda_{F(15-60m)} = A_{F(15-60m)} / A_T$	(15) The horizontal permeability at the building layer of Hong Kong.

^a The resulting percentage values from this calculation were converted to an interval of [0-1] during further multivariate analysis.

^b Calculated for the entire hemispherical sky view and also its eight sectors (9 SVF values for each point).

^c λ_F is a dimensionless quantity. It was calculated at three different height layers along 16 wind directions [42]. The weighted λ_F value based on probability of the 16 wind directions for two seasons are also calculated. Therefore, there are 17 λ_F values for each point.

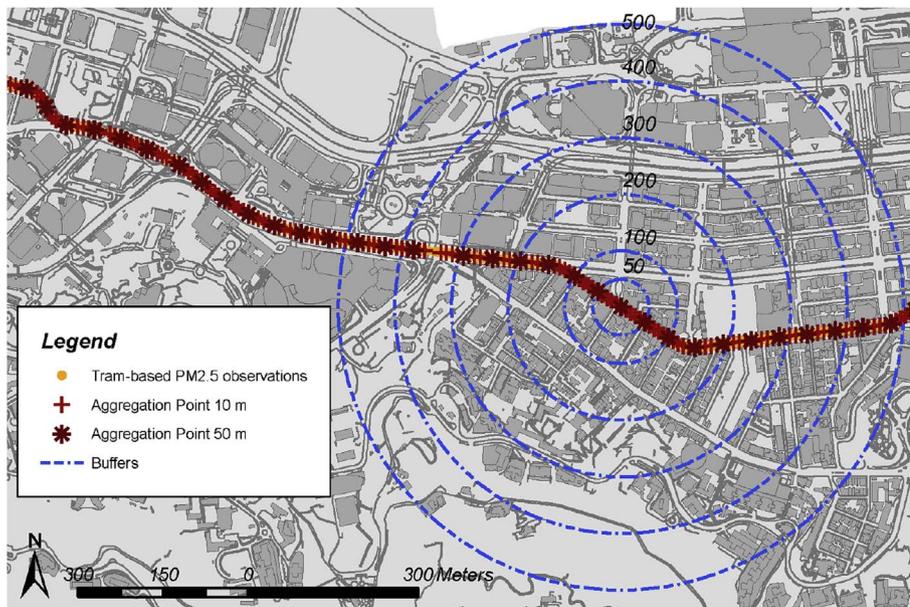


Fig. 7. A series of buffers for the neighboring analysis of the building morphological factors of the surrounding area of data points.

variation. A simple linear regression between the building morphological factors calculated using each buffer and the aggregated PM_{2.5} concentration data were performed. Pearson correlation coefficients (*r*) were calculated for the comparison of buffer widths. Only the buffer-based building morphological factors with the highest $|r|$ were selected as the predictor variables for further correlation analysis.

2.4. Correlating the building morphology with PM_{2.5} concentration

2.4.1. Using stepwise multiple linear regression modelling

The stepwise multiple linear regression (MLR) method was adopted to examine all possible regressions between the aggregated tram-based PM_{2.5} (response variables) and predictor variables within the critical buffers. Regression models were developed by the rules of minimum Bayesian information criterion (BIC), and the forward order is used [45,46]. The formula of an initial MLR model is stated below:

$$PM_{2.5i} = \alpha_1 Var_{1i} + \alpha_2 Var_{2i} + \dots + \alpha_n Var_{ni} + \gamma + \varepsilon \quad (16)$$

where $PM_{2.5i}$ is the PM_{2.5} concentration value at the aggregation point *i* on the tram route. The model includes *n* building morphological factors as the predictor variables. $\alpha_1, \dots, \alpha_n$ are the slopes of values of the building morphological factors $Var_{1i}, \dots, Var_{ni}$ at the aggregation point *i*. γ is the model intercept, and ε is the residual.

The model and all its variables fulfil the significance level of the *p*-value < 0.0001. The model initially developed with the stepwise method was further examined to avoid multicollinearity. Multicollinearity (the situation where predictor variables are highly correlated with each other) in a model leads to limited explanatory capacity and introduces suspicious regressions [47]. In this present study, both the variance inflation factor (VIF) and multivariate correlation analysis were used to detect the underlying correlations among predictor variables, and to ensure that there is no significant multicollinearity among the final predictor variables included in resultant models. Firstly, we examined the VIF of each variable in the initial models. Those with $VIF > 2$ were excluded. Then, we performed multivariate correlation analysis. If significant multicollinearity (correlation of above 0.8) among predictor variables was detected [48], only the variable with a higher simple linear correlation to the response variable was preserved for regression models. The final model was adjusted to ensure there is no multicollinearity issue. The correlation coefficient ($\overline{R^2}$) was used to evaluate the model performance.

2.4.2. Model validation method

To evaluate the model performance, we conducted leave-one-out cross-validation (LOOCV) to compare the differences between the monitored and estimated concentration. The root-mean-square error (RMSE) and the R^2 from the LOOCV (R^2_{LOOCV}) were used to validate the resultant LUR models:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (PM_{2.5i}^t - PM_{2.5i}^e)^2} \quad (17)$$

$$R^2_{LOOCV} = \frac{\sum_{i=1}^n (PM_{2.5i}^t - \widehat{PM}_{2.5i})^2}{\sum_{i=1}^n (PM_{2.5i}^t - \overline{PM}_{2.5i})^2} \quad (18)$$

where $PM_{2.5i}^t$ is the monitored concentration at the aggregation point *i*. $PM_{2.5i}^e$ is the estimated PM_{2.5} concentration at the aggregation point *i* acquired based on the above MLR modelling. $\widehat{PM}_{2.5i}$ is the average value of the $PM_{2.5i}^t$. *n* is total amount of aggregation points in the dataset.

3. Results

The critical buffer width was firstly identified by the sensitivity test for each building morphological factor for summertime and wintertime. The results of the sensitivity test are shown in Table 2. It can be observed that the critical buffers of most morphological factors remain unchanged between summer and winter. The consistency of critical buffers between seasons implies that the influence of urban morphological features on street-level air quality remains significant regardless of the seasonal changes in air pollution modes [31].

Using spatially aggregated seasonal PM_{2.5} concentrations as the response variables and all selected morphological factors (Table 2) as the predictor variables, we developed separate correlation models for summertime and wintertime (Table 3 and Fig. 8). The adjusted R^2 (*Adj R*²) values of the resultant model of the 10 m-spatially aggregated summertime PM_{2.5} concentration is 0.368. The *Adj R*² of the model of the 50 m-spatially aggregated wintertime PM_{2.5} concentration is 0.306.

Table 2
Results of the sensitivity test of the critical buffer (unit: m) of each building morphological factor.

Morphological Factors	r Summer	Buffer, Summer	r Winter	Buffer, Winter
λ_P	0.523	200	0.416	200
σ_{λ_P}	0.419	500	0.346	500
\bar{h}	0.401	200	0.109	500
σ_h	0.228	500	-0.034	500
<i>BVD</i>	0.547	200	0.270	300
λ_F	0.474	200	0.130	200
$\lambda_{F,N}$	0.538	200	0.395	200
$\lambda_{F,NNE}$	0.527	200	0.382	200
$\lambda_{F,NE}$	0.504	200	0.302	200
$\lambda_{F,ENE}$	0.455	200	0.143	200
$\lambda_{F,E}$	0.508	200	0.275	200
$\lambda_{F,ESE}$	0.511	200	0.416	200
$\lambda_{F,SE}$	0.507	200	0.418	200
$\lambda_{F,SSE}$	0.536	200	0.395	200
$\lambda_{F,S}$	0.538	200	0.395	200
$\lambda_{F,SSW}$	0.527	200	0.382	200
$\lambda_{F,SW}$	0.504	200	0.302	200
$\lambda_{F,WSW}$	0.455	200	0.143	200
$\lambda_{F,W}$	0.508	200	0.275	200
$\lambda_{F,WNW}$	0.511	200	0.416	200
$\lambda_{F,NW}$	0.507	200	0.418	200
$\lambda_{F,NNW}$	0.536	200	0.395	200
$\lambda_F(0-15m)$	0.396	200	0.125	200
$\lambda_F(0-15m),N$	0.469	300	0.397	200
$\lambda_F(0-15m),NNE$	0.448	300	0.382	200
$\lambda_F(0-15m),NE$	0.409	300	0.291	200
$\lambda_F(0-15m),ENE$	0.409	300	0.291	200
$\lambda_F(0-15m),E$	0.436	300	0.291	200
$\lambda_F(0-15m),ESE$	0.433	300	0.425	200
$\lambda_F(0-15m),SE$	0.428	300	0.420	200
$\lambda_F(0-15m),SSE$	0.459	300	0.402	200
$\lambda_F(0-15m),S$	0.469	300	0.397	200
$\lambda_F(0-15m),SSW$	0.448	300	0.382	200
$\lambda_F(0-15m),SW$	0.409	300	0.291	200
$\lambda_F(0-15m),WSW$	0.409	300	0.291	200
$\lambda_F(0-15m),W$	0.436	300	0.291	200
$\lambda_F(0-15m),WNW$	0.433	300	0.425	200
$\lambda_F(0-15m),NW$	0.428	300	0.420	200
$\lambda_F(0-15m),NNW$	0.459	300	0.402	200
$\lambda_F(15-60m)$	0.476	200	0.118	200
$\lambda_F(15-60m),N$	0.542	200	0.393	200
$\lambda_F(15-60m),NNE$	0.532	200	0.387	200
$\lambda_F(15-60m),NE$	0.505	200	0.306	200
$\lambda_F(15-60m),ENE$	0.441	200	0.136	200
$\lambda_F(15-60m),E$	0.492	200	0.251	200
$\lambda_F(15-60m),ESE$	0.506	200	0.392	200
$\lambda_F(15-60m),SE$	0.507	200	0.405	200
$\lambda_F(15-60m),SSE$	0.534	200	0.385	200
$\lambda_F(15-60m),S$	0.542	200	0.393	200
$\lambda_F(15-60m),SSW$	0.532	200	0.387	200
$\lambda_F(15-60m),SW$	0.505	200	0.306	200
$\lambda_F(15-60m),WSW$	0.441	200	0.136	200
$\lambda_F(15-60m),W$	0.492	200	0.251	200
$\lambda_F(15-60m),WNW$	0.506	200	0.392	200
$\lambda_F(15-60m),NW$	0.507	200	0.405	200
$\lambda_F(15-60m),NNW$	0.534	200	0.385	200
Ψ_{sky}	-0.318	0	-0.154	50
$\Psi_{sky,N}$	-0.250	0	-0.154	0
$\Psi_{sky,NE}$	-0.110	0	-0.025	0
$\Psi_{sky,E}$	-0.059	0	-0.064	0
$\Psi_{sky,SE}$	-0.192	0	-0.172	0
$\Psi_{sky,S}$	-0.279	0	-0.161	0
$\Psi_{sky,SW}$	-0.265	0	-0.066	0
$\Psi_{sky,W}$	-0.281	0	-0.122	0
$\Psi_{sky,NW}$	-0.306	0	-0.250	50

Table 3
Resultant MLR models showing the correlation between building morphology and street-level PM_{2.5} concentration in summertime and wintertime respectively.

Correlation in Summertime					
Response Variable	Spatially aggregated summertime tram-based PM _{2.5} data using spatial resolution of 10 m				
R ²	0.369				
Adj R ²	0.368				
RMSE	5.575				
Mean of Response	32.836				
P-value	< 0.0001				
10-fold Cross Validation R ²	0.361				
Predictor Variables	Estimate	Std Error	t Ratio	Prob > t	VIF
Intercept	25.886	0.528	49.06	< 0.0001	n/a
σ_h 500 m	-0.111	0.019	-7.01	< 0.0001	1.624
<i>BVD</i> 200 m	64.789	3.122	20.75	< 0.0001	2.239
$\lambda_F(0-15m)$ 200 m	20.008	3.836	5.22	< 0.0001	1.537
Correlation in Wintertime					
Response Variable	Spatially aggregated wintertime tram-based PM _{2.5} data using spatial resolution of 50 m				
R ²	0.310				
Adj R ²	0.306				
RMSE	5.106				
Mean of Response	91.853				
P-value	< 0.0001				
10-fold Cross Validation R ²	0.304				
Predictor Variables	Estimate	Std Error	t Ratio	Prob > t	VIF
Intercept	82.522	1.234	66.85	< 0.0001	n/a
λ_P 200 m	36.333	2.971	12.23	< 0.0001	1.113
σ_h 500 m	-0.096	0.027	-3.60	0.0004	1.113

4. Discussion

4.1. Interpreting the resultant correlation models

As indicated by the resultant models, in Hong Kong, building morphology explains 37% and 31% of the spatial variability in tram-based street-level PM_{2.5} observations in summer and winter respectively. The building morphological indices with the highest correlations to the tram-based PM_{2.5} concentration (street-level air quality) are building volume density (*BVD* 200 m, positive correlation), building coverage ratio (λ_P 200 m, positive correlation), frontal area index of the podium layer (0-15 m, $\lambda_{F(0-15m)}$ 200 m, positive correlation) and variability in building heights (σ_h 500 m, negative correlation). The resultant models identify the important predictor variables and their corresponding critical buffers. Fig. 9 shows the spatial variation of the aggregated PM_{2.5} concentration data and building morphological factors in different sections along the tram route.

4.1.1. Identifying the predictor variables as important urban morphological design factors for Hong Kong

BVD reflects the land use intensity per unit area. Densely packed building bulks block the airflow and reduce the pollutant dispersion rate. *BVD* is not only an influential factor of dynamic potential of air flow but also an indirect measure of the intensity of anthropogenic activities (for example, the traffic flow in a densely-populated area is usually higher than the low-density ones). In other words, the spatial variability in *BVD* also partially depicts the spatial distribution of pollution sources. λ_F is a morphological factor of the permeability of building shapes with respect to the prevailing wind flow. It has been widely adopted in the assessment of urban ventilation. In this present study, it was further separated into three layers of different heights

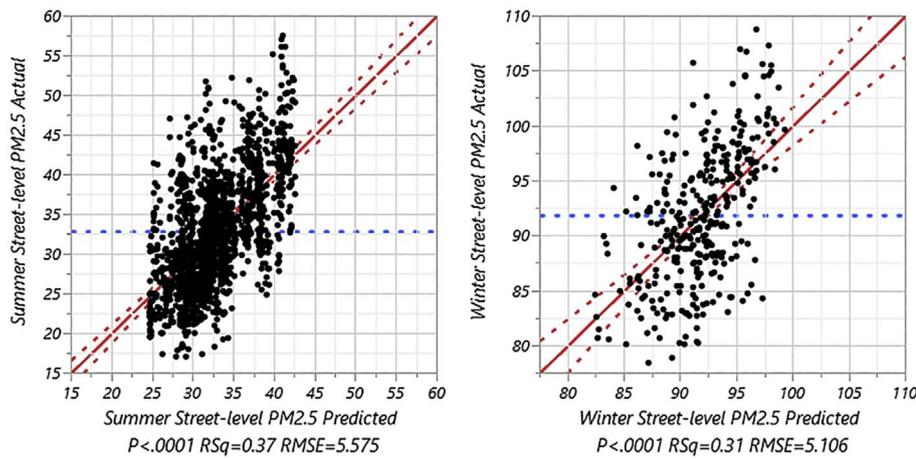
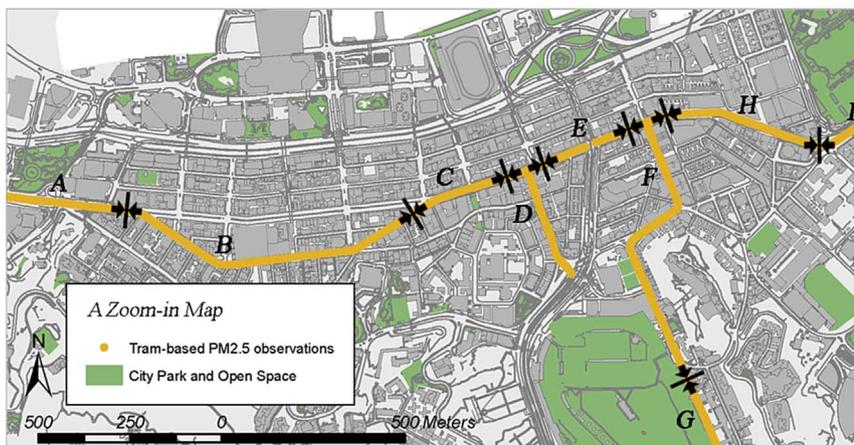
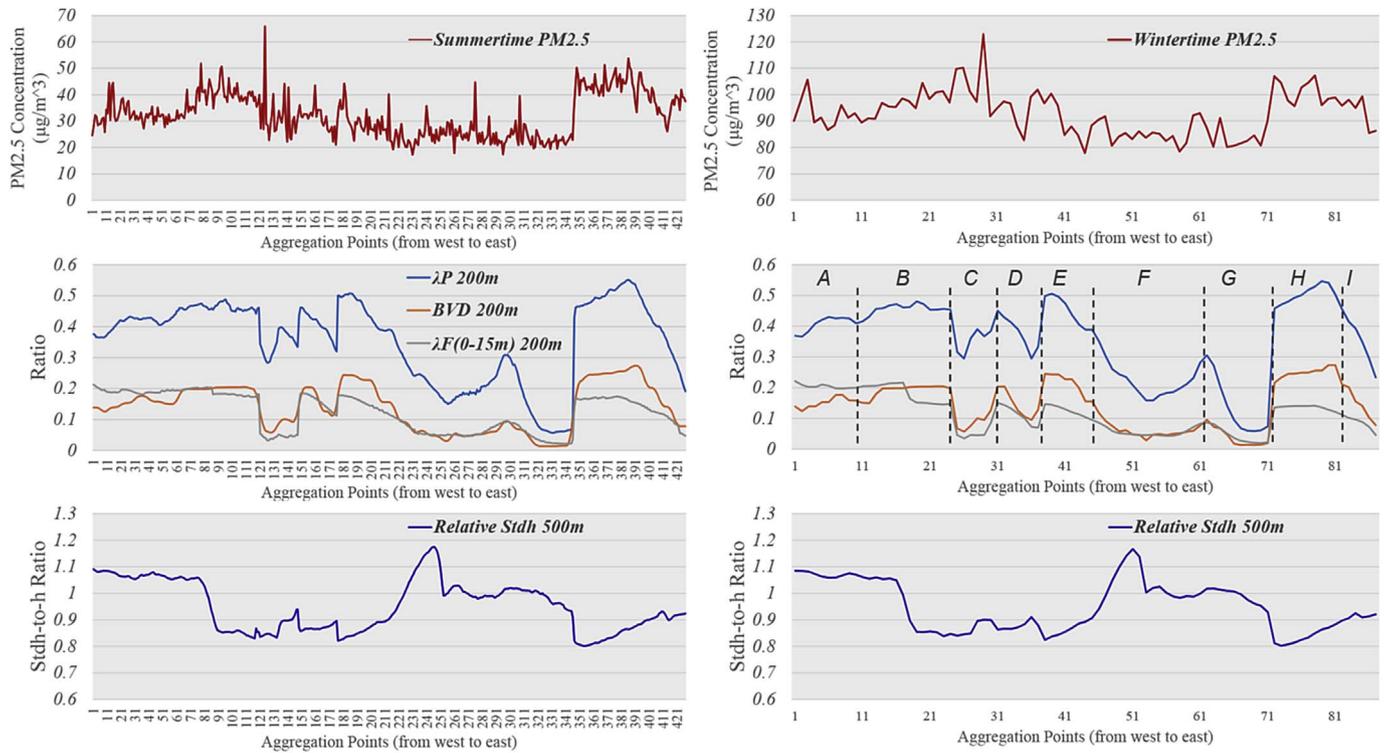


Fig. 8. MLR regression plot of the correlation models. The amount of the wintertime data points are less than summertime because of the different spatial aggregation.



Tram Route Sections:

- A: Admiralty (more open space)
- B: Wan Chai (high-density area)
- C: Wan Chai (lower FAI, more airpaths)
- D: Morrison Hill
- E: Causeway Bay (lower FAI, more airpaths)
- F: Leighton Hill
- G: Happy Valley (greening and open space)
- H: Causeway Bay (high-density area)
- I: Tin Hau (a large park)

Fig. 9. The spatial variation of the aggregated PM_{2.5} concentration data and building morphological factors in different sections along the tram route.

considering the typical building structure of Hong Kong. $\lambda_{F(0-15m)}$ depicts the building morphological permeability at the street-level and thus has an effect on the dispersion rate of air pollutants, especially in a high-density built environment. The inclusion of λ_F in the resultant models (developed using tram-based PM_{2.5} observation in this study) is consistent with the findings in a previous study based on fixed monitoring data from AQMN of HKEPD [41]. Building ground coverage ratio is an alternative indicator of street-level wind availability of $\lambda_{F(0-15m)}$ [42]. The turbulent intensity near the urban surface determines the mixing and dilution of air pollutants. A higher variability in building height (measured as the standard deviation of the building height, σ_h) increases the intensity of turbulence near the urban surface and as a result helps with the dispersion.

4.1.2. Identifying the critical buffers for the morphological factors

The identification of the critical buffer for each building morphological factors is one of the most important contributions of this study. Most prior studies calculated all spatial factors using a fixed grid system with a specific spatial resolution. However, the critical buffer widths of different morphological factors may vary due to the complex physical basis of pollution diffusion and dispersion. As shown in Table 2, the buffer size used in this study is more similar to a typical land use regression approach [49]. Such spatial scale enables the investigation of the neighbourhood-scale building morphological effect within the urban roughness layer. In Hong Kong, it has been proved that neighbourhood scale building morphology within the urban roughness layer has a strong effect on street-level air quality [18]. The buffer is significant for the development of neighbourhood-scale urban design strategies.

As identified by the results of the sensitivity test, the critical buffer width of the BVD, λ_P , $\lambda_{F(0-15m)}$ is 200 m; the critical buffer of σ_h is 500 m. The findings on buffer width can be further interpreted as follows: the street-level air quality (evaluated as PM_{2.5} concentration in this present study) of a specific location in the high-density downtown area of Hong Kong is significantly influenced by the building morphology measured by BVD, λ_P , $\lambda_{F(0-15m)}$ in its surrounding area with a radii of 200 m and σ_h within its surrounding area of 500 m. Alternatively speaking, a building/urban design project strongly affects the street-level air quality of its 200 m-wide surroundings; a distance of 500 m should be defined as the critical range for evaluating and designing the height variability in building clusters. As observed in this study, the critical buffer sizes of the BVD, λ_P , and $\lambda_{F(0-15m)}$ are the same (200 m) while the critical buffer of σ_h is larger (500 m). In a previous similar study in Hong Kong [19], it was found that σ_h has a larger critical buffer than other building morphological factors as well. This phenomenon may be explained by the concept of source area (or ‘footprint’) [50]. A source area refers to the surrounding area (influential buffer) of a sensor location of the measurement with respect to the turbulence. The influential buffer of a screen-level measurement is likely to depend upon the building density. It is thought that this influential buffer has a radius up to approximately 0.5 km [50]. Therefore, it is still reasonable to have a larger buffer of σ_h .

4.2. Estimating the small-scale variability in street-level air quality using building morphology

The estimation of the small-scale spatial variability at the street level in an urban environment serves as a basis for urban environmental planning and policy decision-making, especially for a high-density built environment because the complex building morphology significantly alters street-level air quality. This study has discovered the correlation between the spatial variability of PM_{2.5} concentration and morphological factors, and identified critical design factors. These will enhance the current understanding of the impacts of building design on street-level air quality. For example, as indicated by the resultant correlation models, $\lambda_{F(0-15m)}$ at a 200m buffer is positively correlated to the long-term

average street-level PM_{2.5} concentration both in summertime and wintertime. It means that the PM_{2.5} level of a position within a street canyon is greatly influenced by the morphological permeability of the podium layer within its surroundings with a buffer width of 200 m (a circular area with a diameter of 400 m). It is commonly opined that a high-density urban morphological form with well-developed environmental planning and management policies could be more sustainable because of intensive land use, promotion of public transport mode and efficient use of public resources [51,52]. The findings in this present study can substantially contribute to a more quantitative and scientific basis for the current urban design guidelines in Hong Kong –Chapter 11 of the Hong Kong Planning Standards and Guidelines (HKPSG) [53].

It should be emphasized that this present study will not only be relevant to Hong Kong. As the mobile measurement experimental method is now increasingly used to obtain more accurate spatial estimations of intra-urban air pollution, the findings of this study can be further compared with similar efforts in different regions under different urban contexts. The outputs from this study can be further expanded and applied to other highly urbanized areas in the estimation of street-level air quality.

5. Conclusions

Many previous studies have been conducted to resolve the issues related to air pollution and morphological factors qualitatively or quantitatively, but they were mostly performed at a small spatial scale. Together with some recent efforts (for example [26,27]), this present study is one of the first attempts at dealing with these issues quantitatively at a large spatial scale. The dispersion capability of different morphological configurations along the street canyons was investigated at the level of an urban road network which could not be achieved by conventional methods (such as CFD numerical simulation). The quantitative correlations between PM_{2.5} and morphological factors developed by this present study will allow considerations of pollution dispersion to be incorporated into the urban planning practice. It provides quantitative references and straightforward information of reasonable accuracy to planners at the initial strategic planning stage of urban renewal and new development areas projects in Hong Kong. The findings of this Hong Kong study will serve as a quantitative reference of evidence-based strategy-making of neighbourhood-scale urban designs (e.g. the optimization of the arrangement of buildings, or the spatial layout of urban open space). Moreover, the experimental methods and findings of this study are also readily applicable to investigating the effect of urban morphology on intra-urban air pollution dispersion in other cities.

Author contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. The authors declare no competing financial interest.

Acknowledgment

This research is supported by the General Research Fund (two GRF Projects, RGC Ref No. 14610717 - “Developing urban planning optimization strategies for improving air quality in compact cities using geo-spatial modelling based on in-situ data” and RGC Ref No. 16300715 - “Modelling and Measurement of Micro-scale Variability in Air Pollutant Transport and Dispersion”) from the Research Grants Council (RGC) of Hong Kong. The authors would like to thank Hong Kong Tramways, Limited for providing the tram during the measurement campaign. The authors deeply thank reviewers for their insightful comments, feedbacks and constructive suggestions, recommendations on our research work. The authors also want to appreciate editors for their patient and meticulous work for our manuscript.

References

- [1] D. Schwela, G. Haq, C. Huizenga, W.J. Han, H. Fabian, *Urban Air Pollution in Asian Cities: Status, Challenges and Management*, Taylor & Francis, 2012.
- [2] H.E. Landsberg, *The Urban Climate*, Academic press, London, 1981.
- [3] E. Ng, Policies and technical guidelines for urban planning of high-density cities—air ventilation assessment (AVA) of Hong Kong, *Build. Environ.* 44 (7) (2009) 1478–1488.
- [4] H.J.S. Fernando, S.M. Lee, J. Anderson, M. Princevac, E. Pardyjak, S. Grossman-Clarke, Urban fluid mechanics: air circulation and contaminant dispersion in cities, *Environ. Fluid Mech.* 1 (1) (2001) 107–164.
- [5] HKEPD, *An Overview on Air Quality and Air Pollution Control in Hong Kong*, (2005) Accessed 20 April 2017 http://www.epd.gov.hk/epd/english/environmentinhk/air/air_maincontent.html.
- [6] V. Brajer, R.W. Mead, F. Xiao, Valuing the health impacts of air pollution in Hong Kong, *J. Asian Econ.* 17 (1) (2006) 85–102.
- [7] T.W. Wong, W.S. Tam, T.S. Yu, A.H.S. Wong, Associations between daily mortalities from respiratory and cardiovascular diseases and air pollution in Hong Kong, China, *Occup. Environ. Med.* 59 (1) (2002) 30–35.
- [8] G.W.K. Wong, F.W.S. Ko, T.S. Lau, S.T. Li, D. Hui, S.W. Pang, R. Leung, T.F. Fok, C.K.W. Lai, Temporal relationship between air pollution and hospital admissions for asthmatic children in Hong Kong, *Clin. Exp. Allergy* 31 (4) (2001) 565–569.
- [9] A. ENB, *Clean Air Plan for Hong Kong*, Environment Bureau, Hong Kong, 2013.
- [10] P. Zannetti, *Air Pollution Modeling: Theories, Computational Methods, and Available Software*, Computational mechanics publications, New York, 1990.
- [11] A. Robins, R. Macdonald, *Review of Flow and Dispersion in the Vicinity of Groups of Buildings vol 99*, Report of the Fluid Dynamics and Thermodynamics Group & Environmental Flow Research Centre, 1999 ME-FD.
- [12] D. Carruthers, Dispersion in cities, in: J.G. Ayres, R.M. Harrison, G.L. Nichols, R.L.M. CBE (Eds.), *Environmental Medicine*, CRC Press, Boca Raton, Florida, USA, 2010, p. 320.
- [13] S. Vardoulakis, B.E.A. Fisher, K. Pericleous, N. Gonzalez-Flesca, Modelling air quality in street canyons: a review, *Atmos. Environ.* 37 (2) (2003) 155–182.
- [14] J. Heinrich, U. Gehring, J. Cyrys, M. Brauer, G. Hoek, P. Fischer, T. Bellander, B. Brunekreef, Exposure to traffic related air pollutants: self reported traffic intensity versus GIS modelled exposure, *Occup. Environ. Med.* 62 (8) (2005) 517–523.
- [15] CERC, *ADMS 5 World Leading Software for Modelling Industrial Air Pollution*, (2016) Accessed April 15 2016 <http://www.cerc.co.uk/environmental-software/ADMS-model.html>.
- [16] ANSYS, *ANSYS Computational Fluid Dynamics (CFD) Simulation Software*, (2014) <http://www.ansys.com/Products/Simulation+Technology/Fluid+Dynamics>.
- [17] K.E. Kakosimos, O. Hertel, M. Ketzler, R. Berkowicz, Operational Street Pollution Model (OSPM) – a review of performed application and validation studies, and future prospects, *Environ. Chem.* 7 (6) (2010) 485–503.
- [18] C. Yuan, E. Ng, L.K. Norford, Improving air quality in high-density cities by understanding the relationship between air pollutant dispersion and urban morphologies, *Build. Environ.* 71 (0) (2014) 245–258.
- [19] Y. Shi, K.K.-L. Lau, E. Ng, Developing street-level PM2.5 and PM10 land use regression models in high-density Hong Kong with urban morphological factors, *Environ. Sci. Technol.* 50 (15) (2016) 8178–8187.
- [20] Y. Shi, E. Ng, Fine-scale spatial variability of pedestrian-level particulate matters in compact urban commercial districts in Hong Kong, *Int. J. Environ. Res. Public Health* 14 (9) (2017) 1008.
- [21] L.-Y. He, M. Hu, X.-F. Huang, B.-D. Yu, Y.-H. Zhang, D.-Q. Liu, Measurement of emissions of fine particulate organic matter from Chinese cooking, *Atmos. Environ.* 38 (38) (2004) 6557–6564.
- [22] S.C. Lee, W.-M. Li, L. Yin Chan, Indoor air quality at restaurants with different styles of cooking in metropolitan Hong Kong, *Sci. Total Environ.* 279 (1–3) (2001) 181–193.
- [23] L.Y. Chan, W.S. Kwok, Vertical dispersion of suspended particulates in urban area of Hong Kong, *Atmos. Environ.* 34 (26) (2000) 4403–4412.
- [24] J. Peters, M. Van Poppel, J. Theunis, *Air Quality Mapping in Urban Environments Using Mobile Measurements*, Sensing a Changing World 2012, (2012) Wageningen, The Netherlands.
- [25] S. Hankey, J.D. Marshall, Land use regression models of on-road particulate air pollution (particle number, black carbon, PM2.5, particle size) using mobile monitoring, *Environ. Sci. Technol.* 49 (15) (2015) 9194–9202.
- [26] W.J. Farrell, L. Deville Cavellin, S. Weichenthal, M. Goldberg, M. Hatzopoulou, Capturing the urban canyon effect on particle number concentrations across a large road network using spatial analysis tools, *Build Environ* 92 (Supplement C) (2015) 328–334.
- [27] M. Hatzopoulou, M.F. Valois, I. Levy, C. Mihele, G. Lu, S. Bagg, L. Minet, J. Brook, Robustness of land-use regression models developed from mobile air pollutant measurements, *Environ. Sci. Technol.* 51 (7) (2017) 3938–3947.
- [28] W.W. Che, H.C. Frey, A.K.H. Lau, Sequential measurement of intermodal variability in public transportation PM2.5 and CO exposure concentrations, *Environ. Sci. Technol.* 50 (16) (2016) 8760–8769.
- [29] Z. Li, W. Che, H.C. Frey, A.K.H. Lau, C. Lin, Characterization of PM2.5 exposure concentration in transport microenvironments using portable monitors, *Environ. Pollut.* 228 (Supplement C) (2017) 433–442.
- [30] Z. Li, W. Che, H.C. Frey, A.K.H. Lau, Factors affecting variability in PM2.5 exposure concentrations in a metro system, *Environ. Res.* 160 (Supplement C) (2018) 20–26.
- [31] A. Lau, A. Lo, J. Gray, Z. Yuan, C. Loh, Relative significance of local vs. regional sources: Hong Kong's air pollution, (2007) Civic Exchange.
- [32] W.A.V. Clark, K.L. Avery, The effects of data aggregation in statistical analysis, *Geogr. Anal.* 8 (4) (1976) 428–438.
- [33] C. Lightowlers, T. Nelson, E. Setton, C.P. Keller, Determining the spatial scale for analysing mobile measurements of air pollution, *Atmos. Environ.* 42 (23) (2008) 5933–5937.
- [34] D. O'Sullivan, D. Unwin, *Geographic Information Analysis*, Wiley, New York, 2014.
- [35] M. Jerrett, R.T. Burnett, R. Ma, C.A. Pope III, D. Krewski, K.B. Newbold, G. Thurston, Y. Shi, N. Finkelstein, E.E. Calle, Spatial analysis of air pollution and mortality in Los Angeles, *Epidemiology* 16 (6) (2005) 727–736.
- [36] R.A. Olea, A six-step practical approach to semivariogram modeling, *Stochastic Environmental Research and Risk Assessment* 20 (5) (2006) 307–318.
- [37] P.A. Burrough, R.A. McDonnell, R. McDonnell, C.D. Lloyd, *Principles of Geographical Information Systems*, OUP Oxford, 2015.
- [38] J.E. Diem, A critical examination of ozone mapping from a spatial-scale perspective, *Environ. Pollut.* 125 (3) (2003) 369–383.
- [39] K. Johnston, *ArcGIS 9: using ArcGIS geostatistical analyst*, esri Press, 2004.
- [40] ESRI, *Choosing a Lag Size*, (2015) <http://desktop.arcgis.com/en/desktop/latest/guide-books/extensions/geostatistical-analyst/choosing-a-lag-size.htm>.
- [41] Y. Shi, K.K.-L. Lau, E. Ng, Incorporating wind availability into land use regression modelling of air quality in mountainous high-density urban environment, *Environ. Res.* 157 (2017) 17–29.
- [42] E. Ng, C. Yuan, L. Chen, C. Ren, J.C.H. Fung, Improving the wind environment in high-density cities by understanding urban morphology and surface roughness: a study in Hong Kong, *Landsc. Urban Plan.* 101 (1) (2011) 59–74.
- [43] J. Dozier, J. Frew, Rapid calculation of terrain parameters for radiation modeling from digital elevation data, *IEEE Trans. Geoscience Remote Sens.* 28 (5) (1990) 963–969.
- [44] L. Chen, E. Ng, X. An, C. Ren, M. Lee, U. Wang, Z. He, Sky view factor analysis of street canyons and its implications for daytime intra-urban air temperature differentials in high-rise, high-density urban areas of Hong Kong: a GIS-based simulation approach, *Int. J. Climatol.* 32 (1) (2012) 121–136.
- [45] R.J. Freund, R.C. Littell, L. Creighton, *Regression Using JMP*, SAS Institute Inc. and J. Wiley, Cary, NC, USA., 2003.
- [46] J. Sall, A. Lehman, M.L. Stephens, L. Creighton, *JMP Start Statistics: a Guide to Statistics and Data Analysis Using JMP*, SAS Institute, Cary, NC, USA., 2012.
- [47] G.R. Franke, *Multicollinearity*, Wiley International Encyclopedia of Marketing, John Wiley & Sons, Ltd, 2010.
- [48] C.H. Mason, W.D. Perreault Jr., Collinearity, power, and interpretation of multiple regression analysis, *Journal of Marketing research* (1991) 268–280.
- [49] G. Hoek, R. Beelen, K. de Hoogh, D. Vienneau, J. Gulliver, P. Fischer, D. Briggs, A review of land-use regression models to assess spatial variation of outdoor air pollution, *Atmos. Environ.* 42 (33) (2008) 7561–7578.
- [50] T.R. Oke, *Instruments and Observing Methods: Report No. 81: Initial Guidance to Obtain Representative Meteorological Observations at Urban Sites*, World Meteorological Organization, Geneva, 2004 World Meteorological Organization, WMO/TD (1250).
- [51] Y. Yin, S. Mizokami, T. Maruyama, An analysis of the influence of urban form on energy consumption by individual consumption behaviors from a microeconomic viewpoint, *Energy Policy* 61 (2013) 909–919.
- [52] M. Betanzo, Pros and cons of high density urban environments, *Build* (2007) 39–40. April/May.
- [53] *PlanD, HONG KONG PLANNING STANDARDS and GUIDELINES (HKPSG)*, (2005) http://www.pland.gov.hk/pland_en/tech_doc/hkpsg/full/index.htm.