Environmental Research 157 (2017) 17–29

Contents lists available at ScienceDirect

Environmental Research

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

Incorporating wind availability into land use regression modelling of air quality in mountainous high-density urban environment

Yuan Shi[a,*,1], Kevin Ka-Lun Lau[b,c,d], Edward Ng[a,b,c]

[a] School of Architecture, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong S.A.R., China
[b] Institute of Environment, Energy and Sustainability (IEES), The Chinese University of Hong Kong, Shatin, N.T., Hong Kong S.A.R., China
[c] Institute Of Future Cities (IOFC), The Chinese University of Hong Kong, Shatin, N.T., Hong Kong S.A.R., China
[d] CUHK Jockey Club Institute of Ageing, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong S.A.R., China

ARTICLE INFO

Keywords:
Air pollution modelling
Land use regression
Mountainous high-density city
Wind availability
Urban surface geomorphometry

ABSTRACT

Urban air quality serves as an important function of the quality of urban life. Land use regression (LUR) modelling of air quality is essential for conducting health impacts assessment but more challenging in mountainous high-density urban scenario due to the complexities of the urban environment. In this study, a total of 21 LUR models are developed for seven kinds of air pollutants (gaseous air pollutants CO, NO_2, NO_x, O_3, SO_2 and particulate air pollutants $PM_{2.5}$, PM_{10}) with reference to three different time periods (summertime, wintertime and annual average of 5-year long-term hourly monitoring data from local air quality monitoring network) in Hong Kong. Under the mountainous high-density urban scenario, we improved the traditional LUR modelling method by incorporating wind availability information into LUR modelling based on surface geomorphometrical analysis. As a result, 269 independent variables were examined to develop the LUR models by using the "ADDRESS" independent variable selection method and stepwise multiple linear regression (MLR). Cross validation has been performed for each resultant model. The results show that wind-related variables are included in most of the resultant models as statistically significant independent variables. Compared with the traditional method, a maximum increase of 20% was achieved in the prediction performance of annual averaged NO_2 concentration level by incorporating wind-related variables into LUR model development.

1. Introduction

Urban air quality serves an important function in urban living quality. People living in cities, especially those in megacities, are facing severe health threats resulted from urban air pollution issues (Gurjar et al., 2010; Zhu et al., 2012). As a robust and efficient technique to estimate air pollution concentration level, land use regression (LUR) has been widely adopted to map the spatial distribution of outdoor air pollution (Hoek et al., 2008) and assess the long-term human health exposure (Ryan and LeMasters, 2007). Since its first application in the investigation of intra-urban traffic-related air pollution in European cities (Briggs et al., 1997), LUR models have been developed for many cities and regions, such as Europe (Beelen et al., 2013; Vienneau et al., 2009), North America (Hystad et al., 2011; Novotny et al., 2011), Asia (Kashima et al., 2009) and Australia (Knibbs et al., 2014). Using real monitored data, LUR estimates the ambient air pollution concentration at unmonitored locations based on the surrounding land use, population and traffic conditions with empirical regression modelling techniques. Compared with other methods of air pollution concentration modelling, the main advantage of LUR is that the mapping of small-scale variability is able to provide a more accurate evaluation of the health risks in unmonitored sites when dealing with the difficulty of epidemiological studies on the health impacts of human exposure to outdoor air pollution (Briggs et al., 1997).

The model performance of most LUR cases are reasonably good in many air pollution exposure studies (Ryan and LeMasters, 2007). However, its performance in high-density mountainous areas remains unknown, because so far most LUR case cites/regions have either flat terrain or relatively low-density urban development or both. In areas with mountainous topography, the complex surface morphology strongly perturbs the boundary layer wind field (Finardi et al., 1998; Raupach and Finnigan, 1997). The high-density building environment significantly alters the aerodynamic roughness of land surface (Grimmond, 1998; Kastner-Klein and Rotach, 2004), and consequently, changes the sub-layer wind flows and the dynamic potential of atmospheric pollutant dispersion (Bottema, 1997). It has been demonstrated

* 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.

http://dx.doi.org/10.1016/j.envres.2017.05.007
Received 25 October 2016; Received in revised form 24 April 2017; Accepted 7 May 2017
0013-9351/

Nomenclature

Symbols and abbreviations

3D	Three-dimensional
ADDRESS	A Distance Decay REgression Selection Strategy
AICc	Akaike information criterion
AQMN	Air quality monitoring network
AQMS	Air quality monitoring station
AWS	Automatic weather station
CO	Carbon monoxide
DSM	Digital surface model
GIS	Geographical information system
H/W	The aspect ratio of street canyon
HKEPD	Hong Kong Environmental Protection Department
HKO	Hong Kong Observatory
LOOCV	Leave-one-out cross validation
LUR	Land use regression
MM5	National Center for Atmospheric Research Mesoscale Model, version 5
NO_2	Nitrogen dioxide
NO_X	Nitrogen oxides
O_3	Ozone
PlanD	Hong Kong Planning Department
PM_{10}	Respirable particulate matter
$PM_{2.5}$	Fine particulate matter
PRD	Pearl River Delta
RMSE	Root-mean-square error
SA	Source area
SO_2	Sulfur dioxide
SVF	Sky view factor
VIF	Variance inflation factor
z_0	Roughness length
α	Slope aspect
$\alpha_{r(\theta)}$	The angle between the slope aspect α of a certain location and wind direction θ
β	Slope angle
$\theta, Wdir$	Wind direction (0–$360°$)
λ_F, FAI	Frontal area index
λ_P	Building coverage ratio
φ	Horizon angle
Φ	Azimuth direction
A_F	The total frontal area of all buildings in an urban lot along with the a certain wind direction
A_P	Building footprint area
A_T	The area of a certain urban lot
C_{Dh}	Drag coefficient
d	The radius of the hemisphere circle for SVF calculation
K	Kármán's constant
$P_{(\theta)}$	The probability of wind direction θ.
R	Correlation coefficient
v	Wind speed (m/s)

that wind plays an essential role on the movement, concentration and dispersion of air pollution (Cogliani, 2001; Pasquill, 1971; Seaman, 2000). In mountainous and high-density Hong Kong, the interaction between the hilly topography, high-density building morphology and wind fields are very complex, which makes the wind availability vary vastly between different locations (Tong et al., 2005) and leads to significant spatial variations of air pollution concentrations (Wang et al., 2001). Therefore, it is essential to take wind conditions into account, while modelling air pollution using LUR.

However, few efforts have been made to take wind information and meteorological variables into consideration in the LUR modelling of air pollution concentrations (Arain et al., 2007; Su et al., 2008). Wind

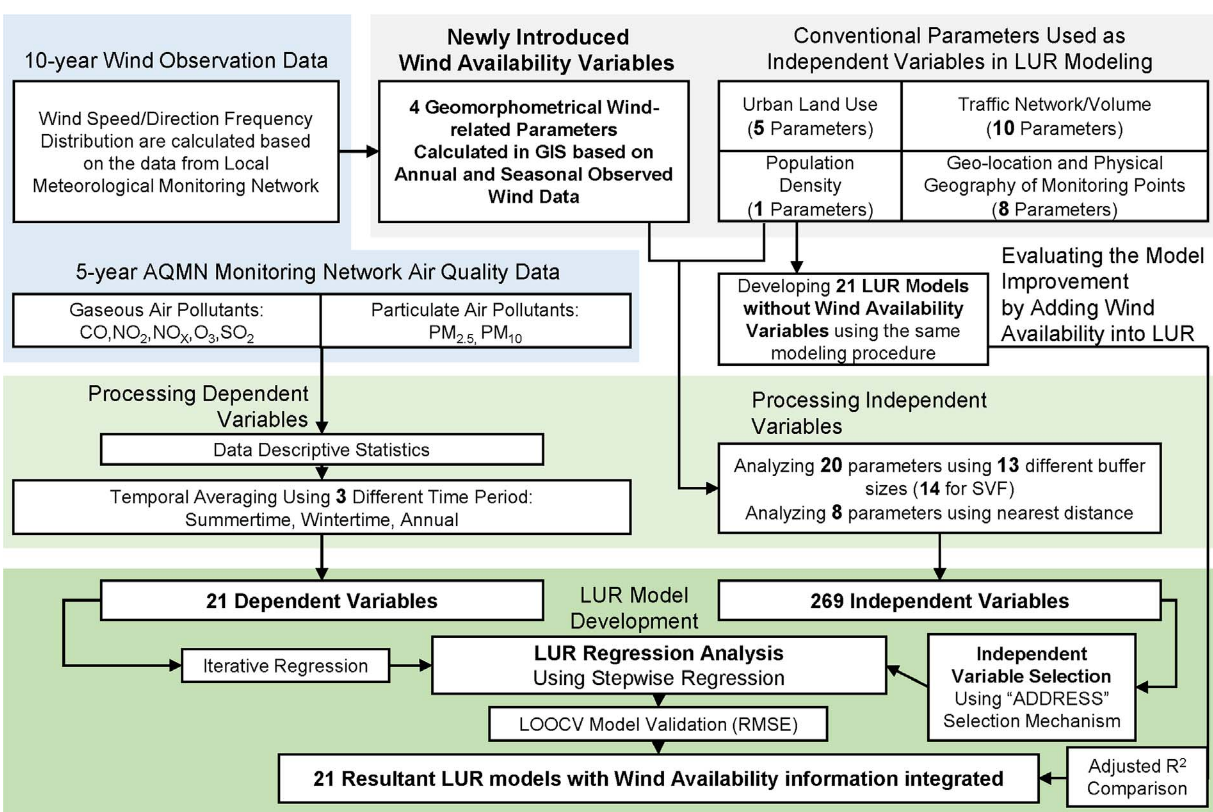

Fig. 1. The workflow chart of this present study.

fields simulated by a regional weather forecasting model were used in a study which used LUR model to predict NO_2 concentrations for the links between health and exposure before (Arain et al., 2007). The study results show that the use of wind variables can significantly improve the LUR model performance but also indicate that the complex wind field on a local scale cannot be fully reproduced by only using the regional weather forecasting model. An attempt to simulate the wind field within the urban boundary layer over the complex terrain using the National Center for Atmospheric Research Mesoscale Model, version 5 (MM5), of Hong Kong has been made with a resolution of 500 m (Tong et al., 2005). However, the modelling output of the above study only lasts a short period of three days, therefore, still didn't provide sufficient wind information for the LUR modelling in this present study.

Another method of incorporating wind information into air pollution LUR modelling is to integrate a source area (SA) air dispersion box model into LUR models. Box model is the simplest air dispersion model type which approximates the study domain as a defined 3D space with definite sides, top and bottom, in which pollutants are emitted and relevant physical processes/chemical reactions can be monitored (RJ Allen, 1975). In previous studies (Ainslie et al., 2008; Su et al., 2008), a homogeneous meteorological condition was set for each 3D wedge-shaped source area in the LUR study area, so the wind information can be integrated into the LUR modelling. However, it remains unknown if such simplification of meteorological conditions and uniformity throughout in the box can well represent the influence of complex and heterogeneous local terrains on wind conditions. Therefore, for the LUR development in Hong Kong, other methods should be tested to deal with the mountainous and high-density city scenario.

In this paper, we estimate and incorporate the fine scale spatial information of wind availability into LUR modelling to predict air pollution concentration in a mountainous high-density city scenario by using Hong Kong as a case. We innovatively adopt the interdisciplinary knowledge of surface geomorphometry to estimate local wind availability to improve the LUR model prediction accuracy of air pollution concentration.

2. Materials and methods

In this present study, local wind conditions were investigated using the meteorological records during an extended period observed by the network of 36 automatic weather stations (AWSs) operated by the Hong Kong Observatory (HKO). With the information of local wind conditions, the fine-scale spatial distribution of wind availability represented by several geomorphometrical parameters was calculated and mapped in the geographical information system (GIS) as high-resolution raster data layers. Data derived from these data layers of wind availability are integrated into the LUR modelling process as independent variables to improve the model's performance. Fig. 1 briefly illustrates the workflow of this present study.

2.1. Study area and context

Hong Kong - a mountainous high-density megacity - is selected as the study area in this paper. Hong Kong is situated at 22° 150' N, 114° 100' E and has an area of approximately 1100 km^2. It has a population of over seven million, which makes it one of the largest megacities in Asia and also around the world. The climate of Hong Kong is subtropical maritime type according to the Koppen Climate Classification, which features hot and humid summers, and warm winters of seasonal mean temperature of 23.3 °C and 18.2 °C respectively (HKO, 2015). The annual mean rainfall measures 2398.5 mm. The annual prevailing wind of Hong Kong is basically easterly with the annual mean wind speed in high-density urban area measuring 11.0 km/h and in open rural area 23.3 km/h. It should be emphasized

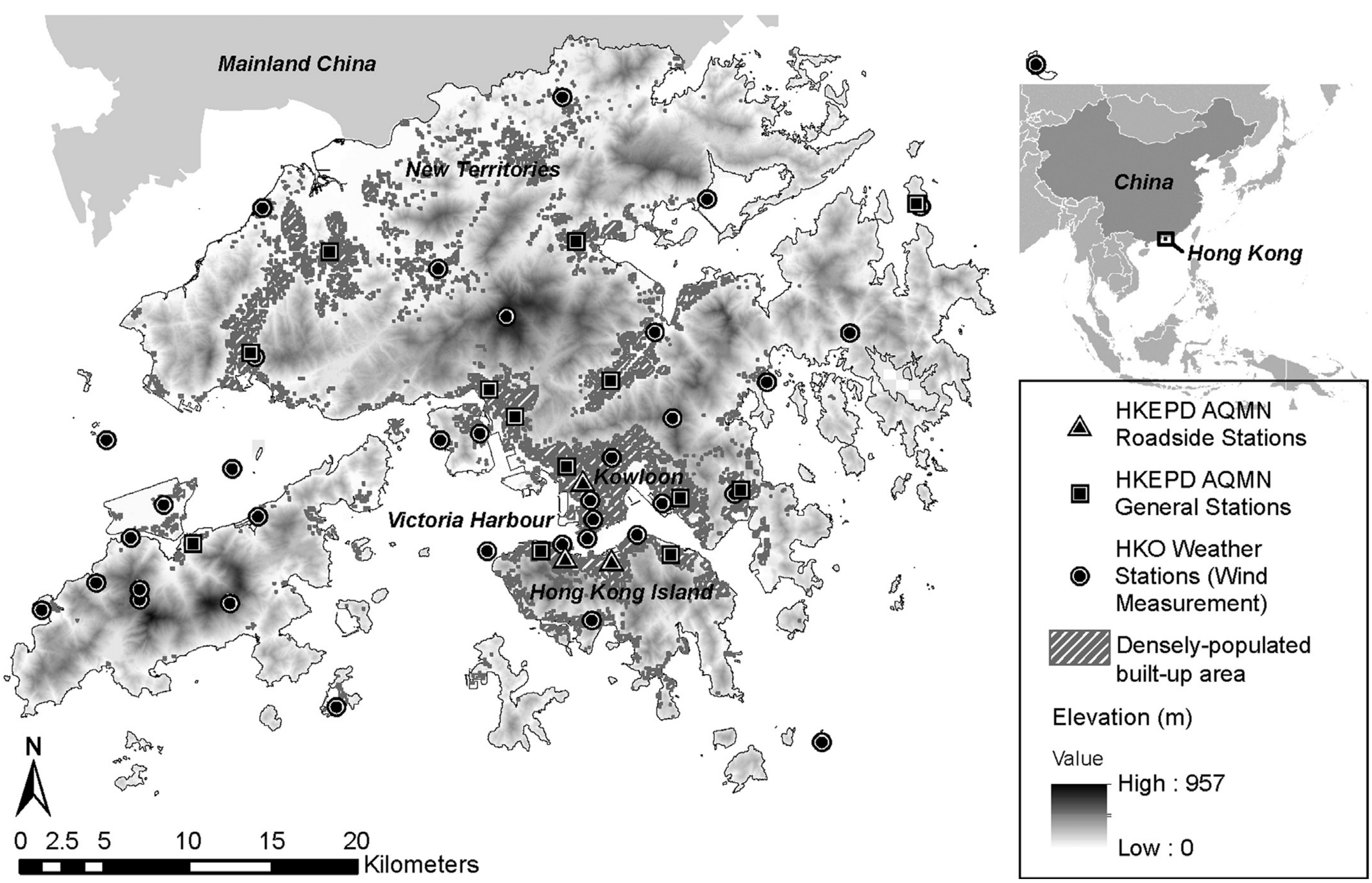

Fig. 2. Study area and locations of the 36 AWSs of HKO and 15 air quality monitoring stations of HKEPD across Hong Kong.

that Hong Kong experiences very complex wind environments as a result of the combination of its mountainous topography, high-density urban morphological characteristics, land and sea breeze circulation, and the alternation of the subtropical monsoon (Chin et al., 1986; Yan, 2007).

2.2. Data collections

2.2.1. Observed wind information

The 5-year (2011–2015) long-term meteorological records of hourly wind speed (*v*, m/s) and direction (*Wdir*, 0–360°) observed by a network of 36 AWSs (Fig. 2) were collected from the HKO. The metadata of these collected wind data were also collected to describe the location, surroundings and the height of the anemometer of each AWS. Annual and seasonal probability distributions of *v* and *Wdir* of the location of each AWS were calculated using observed wind data to provide the basic information of estimating the spatial variation of wind availability in Hong Kong (Fig. 3).

2.2.2. Long-term monitoring data of air pollution concentrations

Long-term hourly air pollution data of high temporal resolution recorded in the recent five years (from 2011 to 2015) by the local air quality monitoring network (AQMN) of the Hong Kong Environmental Protection Department (HKEPD) were collected to develop the LUR models for different air pollutants. Data of the hourly averaged concentrations of gaseous air pollutants CO (carbon monoxide), NO_2 (nitrogen dioxide), NO_X (nitrogen oxides), O_3 (ozone), SO_2 (sulfur dioxide) and particulate air pollutants $PM_{2.5}$ [particulate matters with aerodynamic diameter less than or equal to 2.5 µm, also known as fine particulate matter (FSP)], PM_{10} [particulate matters with aerodynamic diameter less than or equal to 10 µm, also known as respirable particulate matter (RSP)] were collected from the 15 air quality monitoring stations (AQMSs) of AQMN (Fig. 2). The annual and seasonal averages (summer time - May to Aug, wintertime - Dec to Feb) of different air pollutants are calculated with AQMN data. They are then used as dependent variables to develop the LUR models (Fig. 4).

2.3. Estimating the spatial distribution of wind availability in Hong Kong

Air pollution concentration is significantly determined by the local wind availability, which is highly related to the surface topography. Urban climate research shows that wind condition is closely related to the physical interactions between the land surface and the atmosphere (Arnfield, 2003). Wind field is largely determined by the heterogeneous spatial distribution of the air pressure as a result of the irregular surface topography and various surface roughness properties (Landsberg, 1981). For example, a mountain usually traps air pollutants on its leeward side. Wind conditions are also significantly affected by the urban surface roughness in urban built-up areas (Bottema, 1997). Therefore, it is essential to determine the surface topographical characteristics and spatial distribution of the surface roughness to estimate the wind availability of different localities. Hence, it is necessary to consider the land surface topographical/morphological parameters as independent variables in LUR modelling in the case of Hong Kong to improve the prediction accuracy.

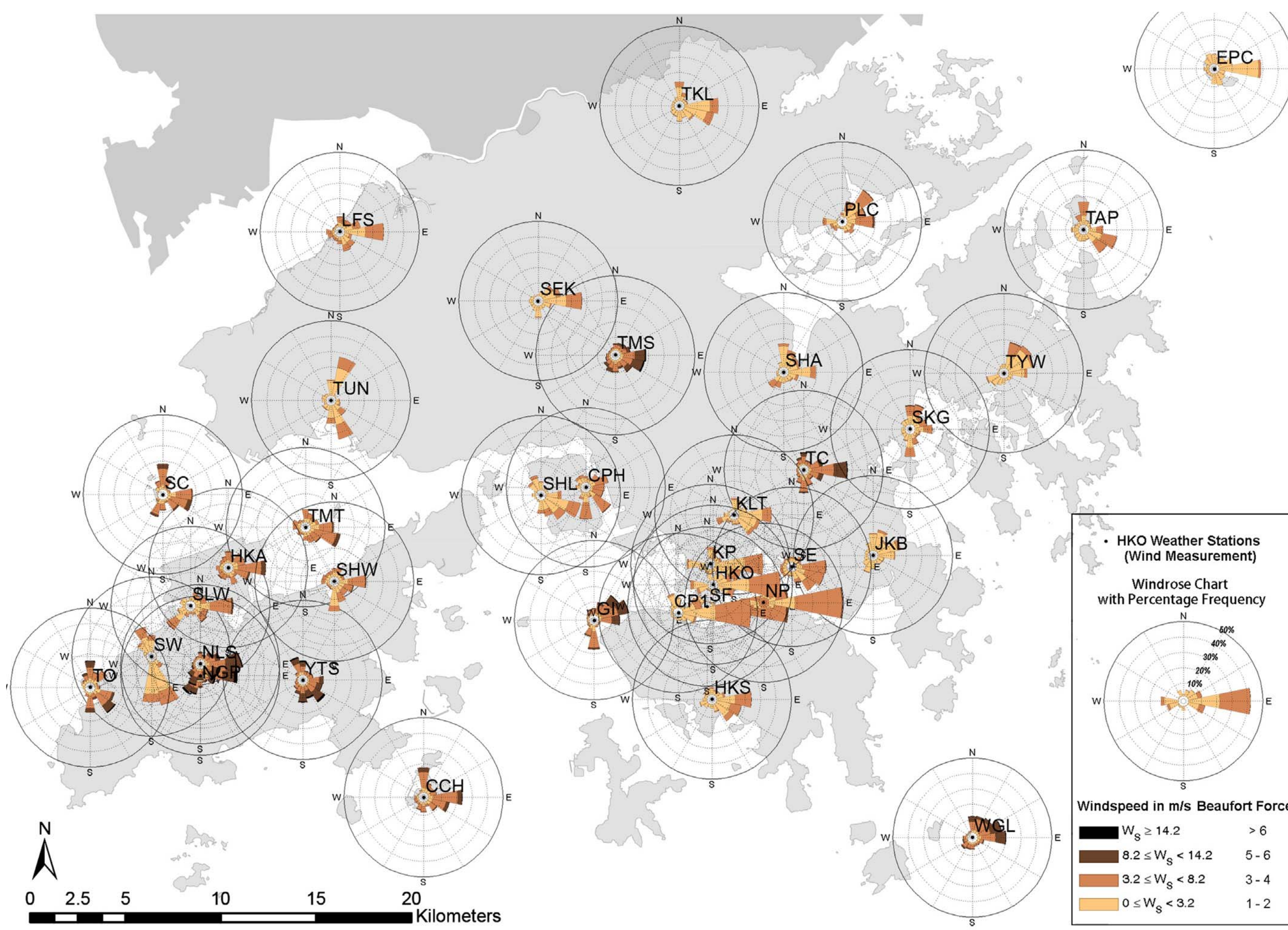

Fig. 3. Annual 16-direction wind roses map of the spatial distribution of wind conditions in Hong Kong based on 5-year hourly wind data from HKO stations.

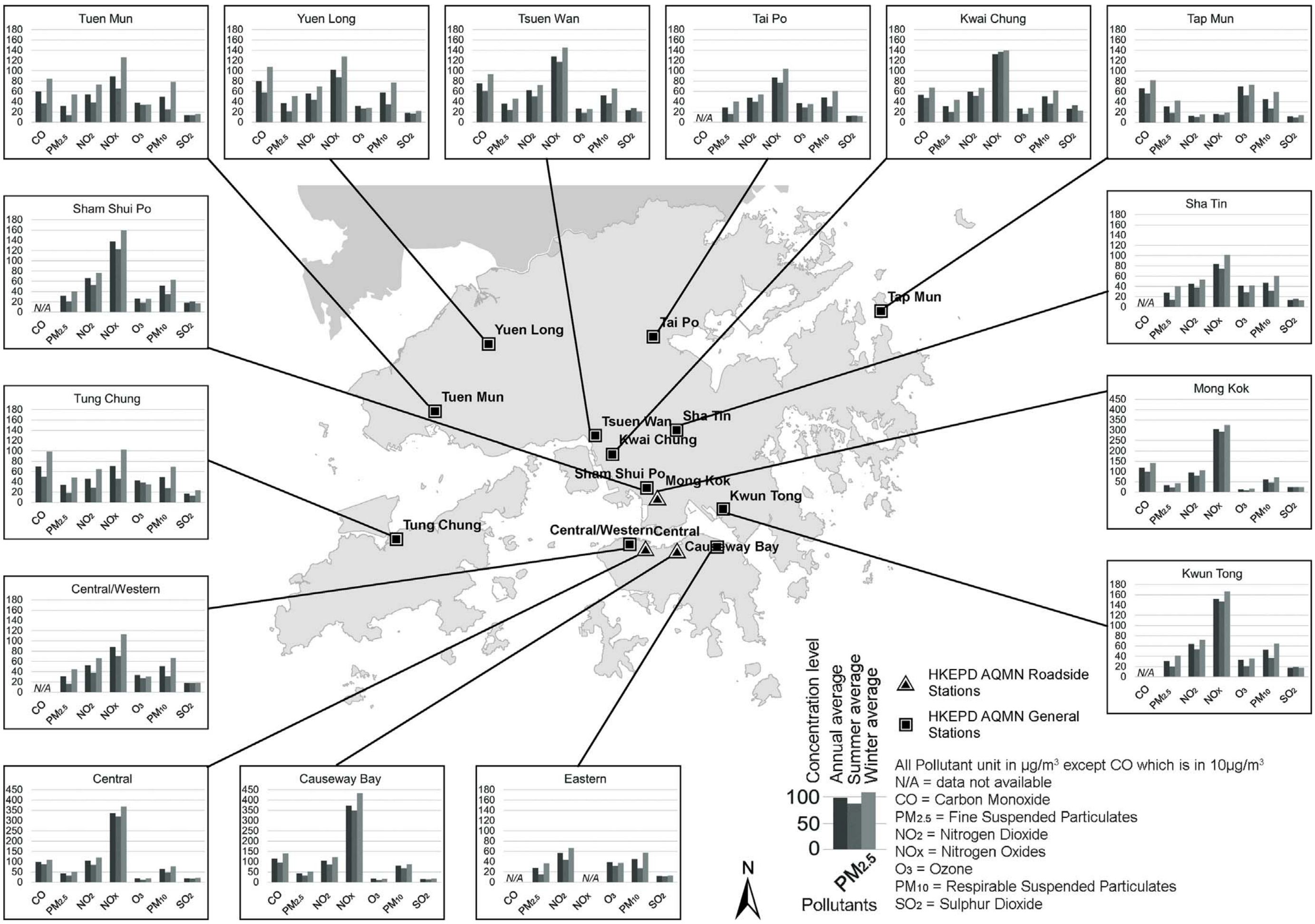

Fig. 4. Annual and seasonal average concentration level of air pollution in Hong Kong based on the long-term monitoring data from AQMSs.

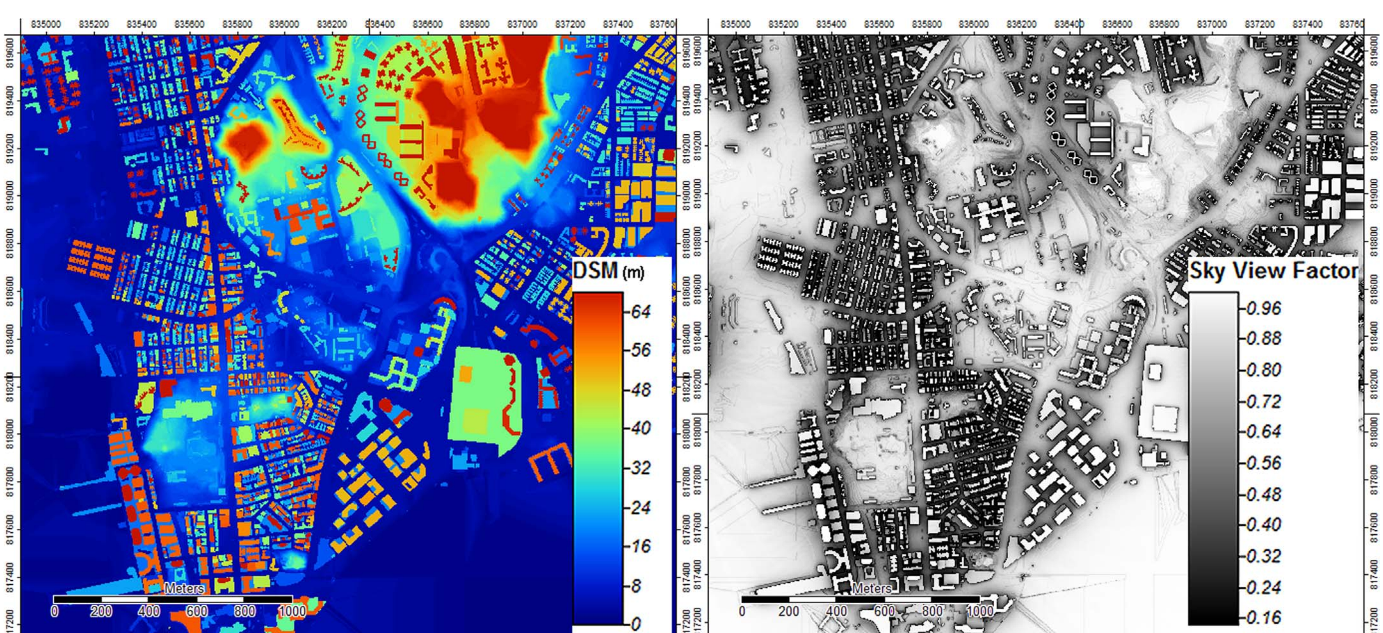

Fig. 5. Input DSM of Hong Kong (Tsim Sha Tsui downtown area) and output calculated SVF map.

By using geomorphometrical methods, land surface parameters were calculated and used to estimate the openness of the local topography to the background wind. Geomorphometrical methods are able to indirectly depict the long-term local wind availability at a reasonable accuracy and fine spatial resolution without involving intensive computational resources and massive time consumption (Böhner and Antonić, 2009). Four land surface topographical/morphological parameters that have been widely used in topo-climatological studies – windward/leeward index, sky view factor (SVF), frontal area index (FAI, λ_F) and roughness length (z_0)— are selected and calculated at a series of LUR buffers as the independent variables to represent the wind availability in LUR modelling. The windward/leeward index reflects the influence of the natural topography on the wind availability at a relatively large-scale. SVF and FAI are good indicators of the intra-urban wind condition with respect to the local-scale building morphology. z_0 is the roughness length that comprehensively depict the surface aerodynamic properties.

The windward/leeward index is a land surface topographical parameter that describes the spatial relationship between the land surface slope aspect and a specific wind direction (Böhner and Antonić, 2009). The angle between the slope aspect α of a certain location and wind direction θ is $\alpha_{r(\theta)}$ which ranges from 0° (fully windward) to 180° (fully leeward). The wind direction-weighted windward/leeward index (ranges from 1 - windward positions to −1 - leeward positions) of a certain location can be calculated with its corresponding wind direction probability $P_{(\theta)}$ from 16 main directions θ:

$$Windward/leewardindex = \sum_{\theta=1}^{16} \cos\alpha_{r(\theta)} \bullet P_{(\theta)}$$

SVF is a topographical/morphological parameter that describes the openness of a specific location at land surface to the open sky hemisphere (Watson and Johnson, 1987). SVF ranges from 0 (which represents a theoretical circumstance that the location is completely blocked by the surrounding topography) to 1 (which indicates a complete openness to the sky hemisphere). SVF is usually used to explain the solar radiation and air temperature variations. Moreover, it is also one of the most important indicators related to wind availability and pollution dispersion and other environmental factors (Eliasson et al., 2006; Rafieian et al., 2014; Ratti and Richens, 1999). It has been found by two previous LUR studies that the prediction accuracy of urban air quality improves when SVF is included in LUR modelling (Eeftens et al., 2013; Tang et al., 2013). With a 2 m-resolution digital surface model (DSM) of Hong Kong generated from Hong Kong Planning Department (PlanD) data sources, the SVF were calculated (Fig. 5) by using the equation proposed by Dozier and Frew (1990):

$$SVF = \frac{1}{2\pi}\int_0^{2\pi} [cos\beta\ cos^2\varphi + sin\beta \bullet cos(\Phi-\alpha) \bullet (90-\varphi - sin\varphi\ cos\varphi)]d\Phi$$

where SVF is calculated for each point location of the DSM with slope aspect α and angle β based on the horizon angles φ in azimuth directions Φ of the hemisphere circle with the radius d.

Surface roughness is an essential factor that has been widely used in boundary layer climate research to describe the wind behaviour over the land surface (Grimmond and Oke, 1999). The roughness properties of land surface affect the surface drag, turbulence intensity, wind velocity and vertical wind profile (Landsberg, 1981). It means the wind availability of a locality is highly dependent on the local natural topography and urban morphology, because roughness varies in different areas due to the differences in topographical and morphological characteristics (Grimmond and Oke, 1999; Raupach and Finnigan, 1997). It has been confirmed that surface roughness is essential for the air quality modelling (Barnes et al., 2014; Bottema, 1997). The complex terrain and high density of Hong Kong mean the city's surface roughness properties vary vastly, and so does the wind environment. The spatial distribution of air pollution will be affected accordingly.

Surface roughness can be estimated using the Davenport roughness classification (Wieringa, 2001), through field observation using the anemometric methods (Grimmond, 1998), or from surface morphometric analysis using geomorphometrical methods (Grimmond and Oke, 1999). Although the Davenport classification has been widely adopted to estimate surface roughness for low-density and mid-density city scenarios, previous research indicates that this method is not suitable for estimating the spatial distribution of surface roughness in high-density city scenario (Ng et al., 2011) because high-density urban areas usually belong to the "chaotic" surface classification with the surface roughness length value $z_0 > 2$. That means no more detailed classification can be made to estimate the intra-urban variation within high-density urban areas. Anemometric methods have been adopted to measure the atmospheric turbulence both in urban areas (Brook, 1972) and in suburban areas (Duchene-Marullaz, 1975). However, it is not a practical way to estimate surface roughness due to the need for high

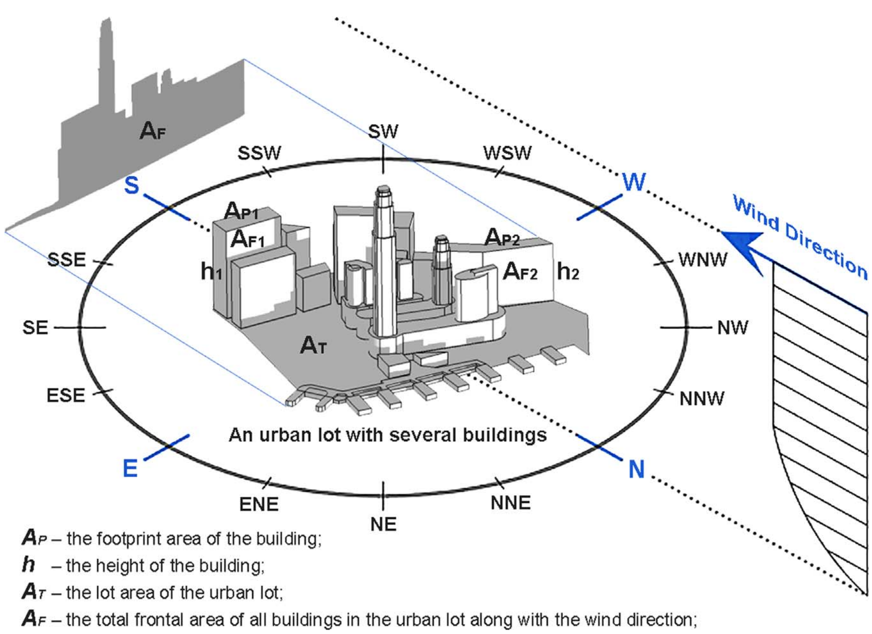

Fig. 6. Demonstration of FAI Calculation using 16 wind directions.

quality field measurement and the lack of clear optimal methods (Grimmond, 1998). Therefore, the geomorphometrical method has become the most popular tool in recent years, and is used in this study to describe the spatial distribution of surface roughness as independent variables to be incorporated into LUR models.

The geomorphometrical methods estimate the surface roughness characteristics by using empirical modelling based on the morphometric characteristics of land surface (Bottema, 1996; Macdonald et al., 1998; Raupach, 1992). The basic calculation method of z_0 was mentioned by Bottema (1996) for regular building clusters:

$$z_0 = [h - h\bullet(\lambda_P)^{0.6}]\,exp\left(-\frac{K}{\sqrt{0.5\bullet C_{Dh}\bullet\lambda_F}}\right)$$

where h is building height, λ_P and λ_F are the building coverage ratio and building FAI of a building lot. C_{Dh} is drag coefficient (constant of 0.8) and K is the Kármán's constant (considered as 0.4). The calculation of λ_P and λ_F are shown in following equations:

$$\lambda_P = (\textstyle\sum_{i=1}^{n} A_{Pi})/A_T$$

$$\lambda_F = (\textstyle\sum_{i=1}^{n} A_{Fi})/A_T$$

where n is the total number of building in the urban lot (Fig. 6). Therefore, the complete z_0 equation is as follow:

$$z_0 = \left\{h - h\bullet\left[(\textstyle\sum_{i=1}^{n} A_{Pi})/A_T\right]^{0.6}\right\} exp\left[-\frac{K}{\sqrt{0.5\bullet C_{Dh}\bullet(\sum_{i=1}^{n} A_{Fi})/A_T}}\right]$$

In Hong Kong, particularly in some high-density downtown areas, the building geometry and layout are extremely irregular due to the mountainous topography. Therefore, the improved calculation method developed by Gál and Unger (2009) is adopted to deal with the irregular building geometries and layouts. This method uses the Voronoi diagram for the partitioning of building lot polygons so that the relevant roughness parameters can be calculated for each lots (Fig. 7). The λ_F of each building lot depends on the wind flow direction. In this present study, λ_F was calculated using 16 wind equiangular directions (Fig. 6). Then, the weighted average of λ_F was calculated according to the annual and seasonal wind direction probability distribution mentioned in Section 2.2.1, so the seasonal variation of wind availability could be incorporated into the LUR modelling:

$$\bar{\lambda}_F = \sum_{\theta=1}^{16} \lambda_{F(\theta)}\bullet P_{(\theta)}$$

2.4. Independent/dependent variables of LUR modelling

As suggested in many previous LUR studies (Hoek et al., 2008), independent variables were prepared for LUR modelling according to the following four classifications: (1) traffic network/volume, (2) urban land use, (3) population density, (4) geo-location and physical geography of monitoring points. As mentioned, four other parameters related to the wind availability were also prepared as independent variables to depict the influence of the wind environment on LUR model performance. This present study calculated 20 parameters using 13 different buffers (14 for SVF) and 8 parameters using nearest distance analysis for each sampling point to check all potential independent variables for the LUR models for air pollution of Hong Kong (Table 1). In total, 269 potential independent variables were prepared and analysed in this study.

Annual, summertime and wintertime average long-term concentration of seven air pollutants (CO, NO_2, NO_X, O_3, SO_2, $PM_{2.5}$ and PM_{10}) were calculated and used as the dependent variables for LUR modelling. As such there were 21 resultant LUR models produced by this present study. The reason for separately modelling the air pollution concentration for different seasons is that the seasonal variation of the meteorological conditions of Hong Kong is considerable. In Hong Kong, because of the seasonal variation of monsoons, the prevailing synoptic wind directions are northerly and northeasterly in wintertime, mainly easterly during spring and autumn seasons, and southerly wind dominates the summer months (Guo et al., 2007; Sin et al., 2002; Yan, 2007). Consequently, the overall air pollution concentration level is significantly affected by the long-distance transportation of pollution from sources in the Pearl River Delta (PRD) region in mainland China during the wintertime (Kwok et al., 2010; Yuan et al., 2006), while in summertime, it is dominated by local emission sources (Lau et al., 2007). Separately conducting LUR modelling for different seasonal periods provides a clear estimation of the contribution of local sources, identifies the local determinants of the spatial distribution of air pollution in Hong Kong and also helps to evaluate the regional impacts from surrounding areas of Hong Kong.

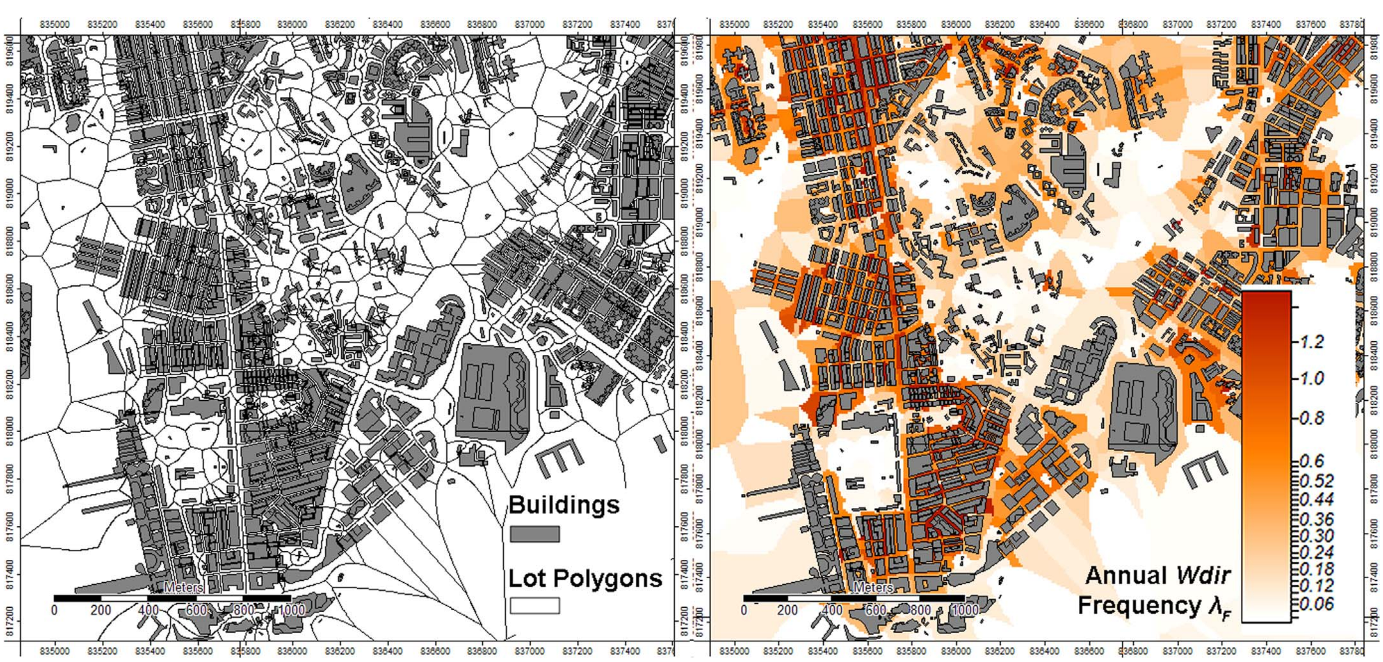

Fig. 7. Generated lot polygons for building and calculated annual Wdir probability distribution weighted average λ_F value.

Table 1
List of all independent variables used in this study for LUR modelling.

Parameters used as independent variables		Units	Analysis methods	Abbreviation
Traffic Network/Volume				
Road network line density	Expressways and trunk road	km/km^2	Buffer[a]	EXP
	Primary road	km/km^2	Buffer	PRI
	Secondary road	km/km^2	Buffer	SEC
	Tertiary road	km/km^2	Buffer	TER
	Ordinary road	km/km^2	Buffer	ORD
Road area ratio (%)		%[b]	Buffer	RDA
Traffic volume	Public transport vehicles	Passenger Car Units (PCUs)	Buffer	PTPCU
	Private and government vehicles	PCUs	Buffer	PGPCU
Count of bus stops		number	Buffer	BUSST
Distance to marines ports & routes		km	Distance	MARINE
Urban Land Use				
Land use area	Residential	m^2	Buffer	RES
	Commercial	m^2	Buffer	COM
	Industrial	m^2	Buffer	IND
	Government	m^2	Buffer	GOV
	Open space	m^2	Buffer	OPN
Population Density				
Population density		person/km^2	Buffer	POP
Geo-location and Physical Geography of Monitoring Points				
Longitude (Δx to the coordinate origin of HK1980 Gird)		m	Distance	LNG.
Latitude (Δy to the coordinate origin of HK1980 Gird)		m	Distance	LAT.
Elevation above the Hong Kong Principal Datum ("mPD")		m	Distance	ELEV.
Distance to waterfront		km	Distance	WATER
Distance to local power plants		m	Distance	POWER
Distance to city parks		km	Distance	CITYPARK
Distance to country parks		km	Distance	COUNTRYPARK
Greening coverage ratio		%	Buffer	GCR
Wind Availability[c]				
Sky view factor (SVF)[d]		[0,1]	Point, Buffer	SVF
Wind effect (Windward/leeward index)	Annual Summertime Wintertime	Dimensionless quantity	Buffer	WE_ANN WE_SUM WE_WIN
Frontal area index (FAI, λ_F)	Annual Summertime Wintertime	Dimensionless quantity	Buffer	FAI_ANN FAI_SUM FAI_WIN
Land surface roughness length (z_0)	Annual Summertime Wintertime	m	Buffer	z_0_ANN z_0_SUM z_0_WIN

[a] A series of LUR buffers that adopted in this present study: 50 m,100 m,200 m,300 m,400 m,500 m,750 m,1000 m,1500 m,2000 m,3000 m,4000 m and 5000 m;
[b] Values of all percentage-based variables (%) have been standardized to [0,1] during LUR model development;
[c] All variables related to the wind availability are calculated based on the weighted probability distribution of wind direction;
[d] Original SVF is a parameter calculated for point locations. Hence, besides mean SVF in different buffers, point SVF values are also retained as an independent variable (defined as 0 m buffer, directly use the abbreviation – SVF without showing buffer size).

2.5. LUR modelling process

2.5.1. Independent variables selection

In this study, instead of manually conducting the exploratory regression modelling, we adopted the widely used stepwise regression technique which is an iterative regression modelling method that automatically selects independent variables by testing statistical significance of regression models (Tabachnick and Fidell, 2001). It has been found that introducing too many independent variables into a stepwise multiple regression modelling process will cause spurious regressions and usually leads to over-fitted multiple regression models (Tu et al., 2005). To avoid this issue, "A Distance Decay REgression Selection Strategy (ADDRESS)" developed by Su et al. (2009) was adopted to preliminary filter all independent variables. Firstly, a series of distance-decay curves were plotted for different buffers using the correlation coefficient (R) between the dependent variables of concentration of seven air pollutants and each LUR parameter. Then, by strictly following the independent variable selection mechanism developed by Su et al. (2009), independent variables at the optimal buffer sizes for the air pollution concentration are selected as the candidate independent variables before further regression modelling.

It is known that collinearity among independent variables causes over-fitting issues and spurious resultant regressions as well (Tu et al., 2005). Therefore, multivariate statistical analysis is used to examine the correlations between the independent variables preliminary selected by using "ADDRESS". Together with variance inflation factor (VIF), the results of the multivariate analysis provide a reference to detect hidden correlations between independent variables in the resultant models to maintain statistical reliability.

2.5.2. Stepwise regression LUR modelling and model validation

Stepwise regression has been used for selecting independent variables for a long time and was developed in the form of an automatic computer program to deal with the complexity of multiple regression statistical studies (Jennrich, 1977; Miller, 2002, 1984) (Section 1.1, Supplemental material). In this study, SAS JMP statistical software package was used in order to find the optimal LUR models as determined by the minimum Akaike information criterion (AICc) (Freund et al., 2003; Sall et al., 2012). The adjusted R^2 values and leave-one-out cross validation (LOOCV) root-mean-square error (RMSE) of generated models were examined to check the model performance (Section 1.2, Supplemental material). A strict set of criteria based on the VIF (independent variables with VIF < 2) was adopted to screen out independent variables with significant collinearity to produce the final models.

3. Results and discussions

3.1. Patterns in distance-decay curves of independent variables

The "ADDRESS" selection mechanism has been described in detail in other research articles (Su et al., 2009). In our study, a total of 21 groups of distance-decay curves (420 curves in total) were plotted to portray the average concentrations over the past five years of seven kinds of air pollutants in three time periods. Each distance-decay curve group has 20 curves correspondingly depicting the correlation between the pollution concentrations and the 20 buffer-based parameters listed in Table 1. An example of distance-decay curve groups is shown in Fig. 8, which shows the dependent variables of annual average NO_2 concentration. High correlations ($R > 0.8$) can be clearly identified between the wind-related variables (SVF, FAI and z_0) and NO_2 level. Similar results are founded in curves for other pollutants as well, reflecting the significance of local wind condition on the estimation of air pollution concentration.

Besides the high correlation between wind-related variables and pollution concentrations, several other common characteristics are observed from these distance-decay curves: Firstly, as functions of the buffer distance, the distance-decay curves for summertime usually have simpler and clearer trends than the wintertime's ones. Most of summertime curve contains only one or two peaks, and didn't change

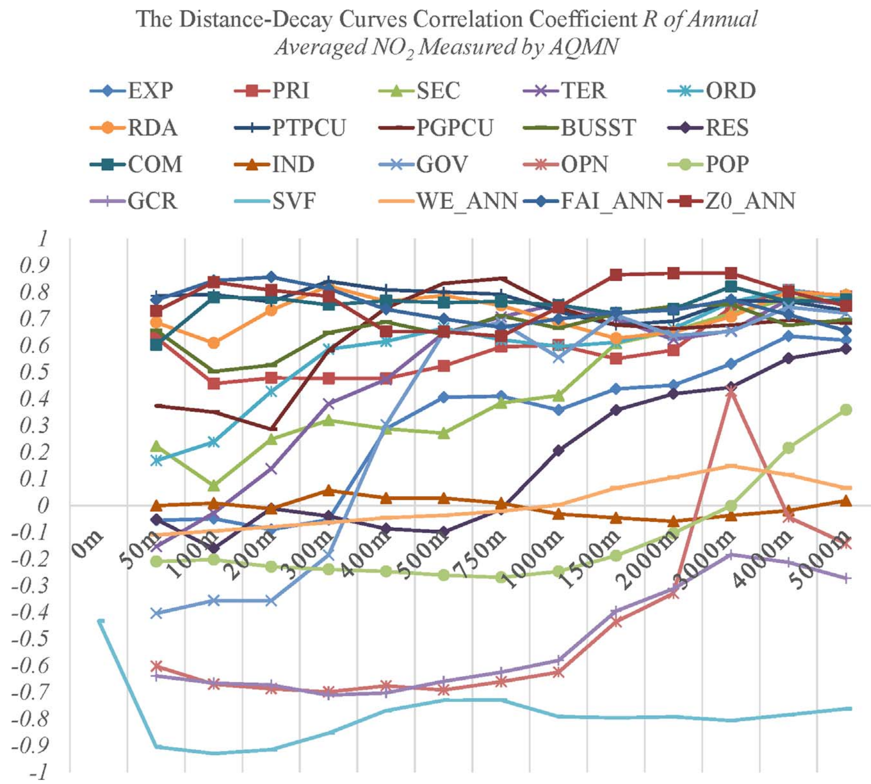

Fig. 8. The Distance-Decay Curves of R between all independent variables (at a series of buffers) and the annual averaged NO_2 concentration level (please also see Fig. S-1 in the supplemental material for more illustration).

between positive and negative. In contrast, wintertime curves are more complex and irregular. Considering that the LUR modelling in this study focuses on using local urban database as independent variables, this observation clearly reflects the seasonal variation of the dominant air pollution modes caused by the seasonal variation of the monsoon between summer (local emission dominant) and winter (combination of both local emission and strong regional impacts from PRD) (Kwok et al., 2010; Yuan et al., 2006). Secondly, the trends of annual distance-decay curves for most pollutants are more similar to the summertime distance-decay curves than the wintertime curves. This indicates the dominance of local emissions in long-term air pollution concentration levels although regional emissions do have considerable contribution to severe air-pollution episodes which usually occur during wintertime (Fung et al., 2005). Thirdly, many curves have a two-peak trend with one peak at a smaller buffer (between 50 m and 400 m) and the other at a much larger buffer (between 2000 m and 5000 m). This finding implies that the same LUR parameter considerably affects the concentration level at two buffers via different mechanisms on different

Table 2
The structure and performance of resultant LUR models include wind availability as independent variables (all independent variables with VIF < 2).

Air Pollutant	Time Period	Resultant LUR Model	Adjusted R^2	LOOCV RMSE	p-value
CO	Annual	45.810 + (2.132e − 7)(COM5000) + 47.628(FAI_ANN0200)	0.844	9.286	0.0016
	Summer	69.319 + (2.334e − 7)(COM5000) − 0.730(ELEV.)	0.873	8.052	0.0009
	Winter	47.068 + 102.949(FAI_WIN0200)	0.766	12.010	0.0012
NO_2	Annual	103.845–84.914(SVF0100) + (9.733e-6)(PGPCU0750)	0.871	8.655	< 0.0001
	Summer	74.109–61.370(SVF0100) + (1.133e-5)(PGPCU0750)	0.843	8.126	< 0.0001
	Winter	102.278–90.457(SVF0100) + 120.943(RDA0300)	0.916	7.751	< 0.0001
NO_X	Annual	60.631–66.239(SVF) + 15.763(ORD4000) + 11.178(z_0_ANN0100)	0.769	51.837	0.0004
	Summer	308.463 + (3.293e − 2)(COUNTRYPARK) + 19.802(TER1000) − 398.601(SVF0100)	0.864	39.023	< 0.0001
	Winter	467.642 + (2.748e − 2)(COUNTRYPARK) − 522.487(SVF0100)	0.872	41.986	< 0.0001
O_3	Annual	13.292-(2.813e − 8)(RES4000) + 50.924(SVF0200)	0.835	5.464	< 0.0001
	Summer	− 58.960-(3.610e − 3)(COUNTRYPARK) + 114.728(SVF1500)	0.874	4.160	< 0.0001
	Winter	56.058–25.470(FAI_WIN3000) − 4.864(z_0_WIN1500)	0.521	9.506	0.0048
SO_2	Annual	9.842-(3.392e − 3) (MARINE) + 6.032(TER2000) + (1.103e − 5)(IND1000)	0.887	1.491	< 0.0001
	Summer	42.412-(7.785e − 5)(POWER) − 33.600(WE_SUM0500) + 1.686(z_0_SUM2000)	0.740	3.280	0.0004
	Winter	40.777-(8.107e − 2)(ELEV.)-(2.514e − 4)(POWER) − 17.575(WE_WIN4000)	0.326	3.328	0.0632
$PM_{2.5}$	Annual	30.222 + (7.101e − 5)(COM0300) + (1.066e − 2)(PTPCU0050)	0.671	2.612	0.0005
	Summer	16.059 + (9.160e − 5)(COM0300) + (4.063e − 3)(PTPCU0100)	0.771	2.714	< 0.0001
	Winter	47.141–1.015(PRI0400) + 1.745(TER1000)-(3.821e-4)(OPN0100)	0.422	3.758	0.0287
PM_{10}	Annual	52.984–9.686(SVF) + (3.703e-2)(PTPCU0050) + 4.955(FAI_ANN0200)	0.854	3.544	< 0.0001
	Summer	32.933–0.151(ELEV.) + (5.336e-2)(PTPCU0050)	0.895	3.524	< 0.0001
	Winter	67.654 + 1.540(BUSST0050)-(5.416e − 4)(OPN0100)	0.634	5.138	0.0016

climatic scales. The smaller optimal buffer suggests the effect of a local phenomenon on small-scale air pollution variation, while the larger optimal buffer possibly reveal a regional impact on large-scale air pollution distribution. The same pattern also came up in Su et al. (2009) s′ LUR study. The high |R| at the larger optimal buffer confirms that the AQMN is appropriate for general air quality monitoring in Hong Kong. The locations of AQMSs are reasonable for the purpose of monitoring the long-term air quality of its corresponding areas. However, the peak in the smaller buffer distance on many LUR parameters decay curves also reflects the local influence on the long-term air quality monitoring results. This also implies that the spatial coverage of AQMN should be further developed in order to catch more detailed information on local variation of air pollution concentrations.

3.2. Resultant LUR models

Finally, a total of 21 LUR models are developed for seven kinds of air pollutants (gaseous air pollutants CO, NO_2, NOx, O_3, SO_2 and particulate air pollutants $PM_{2.5}$, PM_{10}) with reference to three different time periods (summertime, wintertime and annual average of hourly monitoring data). Table 2 shows the 21 resultant LUR models. It is found that wind availability-related independent variables such as SVF, FAI are involved in many of these resultant models, which indicates that wind availability information plays an important role in modelling air pollution in a mountainous high-density city.

3.2.1. Identifying the dominant impact factors of air quality in Hong Kong

Four types of commonly shared independent variables by resultant models are clearly identified - traffic-related, geographical location of monitoring locations, urban land use and wind availability. In Hong Kong, road transport is the main contributor to CO emissions and is also accountable for a large proportion of NOx and PM emissions (Wan et al., 2013). Traffic-related variables include the line density of several road types, road area ratio, vehicle traffic volumes and number of bus stops. These largely represent the emission intensity of the aforementioned air pollutants. The presence of these traffic-related variables in the resultant models confirms transport is one of the determinants of several air pollutants in Hong Kong. Locations of monitoring points prove to provide good estimations of local air pollution as well. The variables of distance from the monitoring locations to marine ports/routes and local power plants are included in the models of SO_2, which is consistent with the emission inventory that has identified marine transport and public electricity generation as two of the determinants of SO_2 emission in Hong Kong (Lau et al., 2007). The variable of distance to the country park in the NO_x models quantitatively depicts the urban-rural contrast of air quality. Like other cities, land use pattern is also a key factor in the concentration level and spatial distribution of air pollution. The spatial pattern of commercial (COM), residential (RES) and industrial (IND) land use areas directly represent the spatial distribution air pollutant emissions from human activities, while the open land use (OPN) areas in Hong Kong are usually vegetation and open space. It has been proved that vegetation plays an important role in air pollution reduction in megacities (Yang et al., 2005). The open land use indicates a much lower urban density which has lower emission intensity and also more effective pollution dispersion. As to the wind availability, as expected, wind-related variables also show in the resultant models as significant independent variables. These wind-related independent variables include SVF, FAI and z_0.

3.2.2. Wind-related variables in the resultant LUR models of air pollutants

Table 2 shows that wind-related variables (SVF, FAI and z_0) are included in 14 of all of the 21 resultant LUR models. SVF measures the openness to sky which represent the vertical permeability of a certain locality to the atmosphere. Previous pollution dispersion studies conclude that the retention time of air pollutants in a street canyon with aspect ratio (H/W) of 2.0 is doubled and tripled respectively, when compared with a street canyon with H/W of 1.0 and 0.5 (Liu et al., 2005), because deep street canyons with high H/W geometrically introduce a low SVF. As a result of this present study, SVF (point based SVF value) and SVF0100 (averaged SVF value calculated using the buffer size of 100 m) shows in all models of NO_2 and NO_X. The significance of SVF in LUR modelling of NO_2, NO_x and PM_{10} in this present study has been proven to be highly consistent with some previous LUR studies (Eeftens et al., 2013; Tang et al., 2013). Increasing FAI reduces the horizontal permeability to the urban ventilation, and impedes pollution dispersion as a result. Both the FAI and z_0 have been used in the detection of urban air path to enhance urban air ventilation (Gál and Sümeghy, 2007; Gál and Unger, 2009). The above findings imply that the spatial pattern of most air pollutants are complex and quite heterogeneous under a mountainous and high-density urban

Table 3
The structure and performance of resultant LUR models that only use conventional independent variables (all independent variables with VIF < 2).

Air Pollutant	Time Period	Resultant LUR Model	Adjusted R^2	LOOCV RMSE	p-value
CO	Annual	63.085 + 0.210(ORD0500) + (3.489e − 7)(COM5000)	0.774	11.149	0.0048
	Summer	69.319 + (2.334e − 7)(COM5000) − 0.730(ELEV.)	0.873	8.052	0.0009
	Winter	83.656 + 1.242(PRI0400) + (2.977e − 7)(COM5000)	0.575	16.187	0.0324
NO_2	Annual	29.191 + (1.915e − 5)(PGPCU0750) + (3.416e − 7)(COM3000)	0.792	10.990	< 0.0001
	Summer	20.148 + (1.815e − 5)(PGPCU0750) + (2.464e − 7)(COM3000)	0.786	9.494	< 0.0001
	Winter	15.220 + 193.179(RDA0300) + (4.716e − 7)(COM3000)	0.911	7.769	< 0.0001
NO_X	Annual	13.294 + 18.552(ORD4000) + (2.315e − 6)(COM3000)	0.753	53.595	0.0002
	Summer	− 27.979 + (3.488e − 2)(COUNTRYPARK) + 30.651(TER1000) + (2.227e − 6)(COM3000)	0.780	49.667	0.0003
	Winter	45.910 + (2.937e − 2)(COUNTRYPARK) + (3.196e − 6)(COM3000)	0.755	59.921	0.0002
O_3	Annual	32.602-(2.032e − 8)(RES4000) + (2.052e − 4)(OPN0300)	0.729	7.000	0.0002
	Summer	23.532-(2.775e − 3)(COUNTRYPARK)-(1.455e − 7)(COM3000) + (5.655e − 5)(OPN0500)	0.695	6.463	0.0010
	Winter	51.854-(4.401e − 8)(RES4000)-(9.974e − 4)(COM0100)	0.459	10.110	0.0100
SO_2	Annual	9.842-(3.392e − 3) (MARINE) + 6.032(TER2000) + (1.103e − 5)(IND1000)	0.887	1.491	< 0.0001
	Summer	9.489 + 3.815(PRI2000) + (2.080e − 5)(IND1000)	0.628	3.920	0.0011
	Winter	18.903 + (1.190e − 4)(POWER)-(1.542e − 4)(OPN0200)	0.288	3.423	0.0518
$PM_{2.5}$	Annual	30.222 + (7.101e − 5)(COM0300) + (1.066e − 2)(PTPCU0050)	0.671	2.612	0.0005
	Summer	16.059 + (9.160e − 5)(COM0300) + (4.063e − 3)(PTPCU0100)	0.771	2.714	< 0.0001
	Winter	47.141–1.015(PRI0400) + 1.745(TER1000)-(3.821e-4)(OPN0100)	0.422	3.758	0.0287
PM_{10}	Annual	47.213 + (3.841e − 2)(PTPCU0050) + (1.311e − 4)(COM0200)	0.808	4.064	< 0.0001
	Summer	32.933–0.151(ELEV.) + (5.336e-2)(PTPCU0050)	0.895	3.524	< 0.0001
	Winter	67.654 + 1.540(BUSST0050)-(5.416e − 4)(OPN0100)	0.634	5.138	0.0016

context. The concentration levels of air pollutants (especially CO, NO_x and PM, which are highly related to road transport emission) at a certain location are very localized and are highly determined by the wind availability within a small buffer (50–200 m). For example, FAI0200 (FAI value calculated using the buffer size of 200 m as the independent variable) was found to be influential to the concentration level of CO. This implies that ventilation is essential to the dispersion of traffic-related air pollutants. In contrast, the windward/leeward index that calculated using the digital elevation model of Hong Kong only shows in SO_2 models (at a much larger buffer size, up to 4000 m). This indicates the spatial pattern of SO_2 has a larger spatial scale than CO, NO_x and PM. This reaffirms that the main sources of SO_2 in Hong Kong are the marine emissions and local electricity generations (Lau et al., 2007). A distribution map of all marine transportation facilities and local power plants is shown in the supplemental material (Fig. S-2).

3.2.3. Evaluating the improvement on model performance by incorporating wind availability into LUR

As described, this study comprehensively incorporates wind information into LUR modelling as the independent variables. This approach has only been adopted in a very small amount of studies. The resultant models shows that the newly introduced independent variables about wind availability are highly significant statistically, and has high prediction power in the resultant LUR models based on multivariate regression analysis (Tabachnick and Fidell, 2007). To evaluate the model performance improvement by integrating wind availability information into LUR modelling as independent variables, LUR models that only using conventional LUR independent variables are also established (Table 3).

By comparing these 21 pairs of resultant models (a side-by-side model comparison table is included in the supplemental material, Table S-1), it is found that 14 of all of the 21 models have achieved an averaged increase of 8% in prediction performance when wind availability-related parameters are added as independent variables into the LUR modelling process. The maximum increase of almost 20% is found in the LUR model of annual average concentration level of NO_2. Such improvement in NO_2 prediction is useful in controlling local traffic pollution emissions and related health risk assessments, because NO_2 is mainly emitted from Hong Kong local road transport. In the annual NO_2 model, the involved wind availability variable – SVF – measures the vertical openness and ventilation capacity of building geometry, which is highly useful and intuitive for the urban planning practice. This is because SVF is high related to a common urban planning index - H/W. In other words, LUR models with wind availability variables are not only useful in air pollution and health risk assessment but also informative for urban environmental planning and decision making.

It should be emphasized that independent variables concerning commercial land use become important when all wind availability variables are excluded from the LUR modelling process. This indicates that commercial land use may be an alternative indicator in the modelling air quality of Hong Kong. The commercial land prices are extremely high in Hong Kong so developers tend to fill up the limited site area with tall buildings to recover the operating costs. Consequently, correlations between commercial land use and building density are high. Commercial land use also attracts a large number of visitors and consumers which leads to more intensive traffic flows, denser road networks, and an increased number of traffic facilities within or around these sites. The multivariate analysis conducted between all independent variables indicates high correlations between commercial land use (COM5000) and road transport (PRI5000 with $R^2 = 0.82$, BUSST5000 with $R^2 = 0.86$) and also urban surface geomorphometry (FAI5000 with $R^2 = 0.58$).

3.3. Limitations and future works

This present study models air quality of Hong Kong by using historical monitoring data from the existing air pollution monitoring network operated by local environmental protection department as dependent variables. Although the data used in this study is a 5-year long-term dataset with hourly-based high temporal resolution, the number of monitoring locations is quite limited (only 15) compared with many previous LUR studies (usually more than 20 monitoring locations) (Hoek et al., 2008) due to the limited spatial coverage of the monitoring network. The concern is that, the complex topography and heterogeneous urban development in Hong Kong make the conditions of air quality vary significantly among different locations, while many high-density and heavily trafficked areas in Hong Kong are not monitored by AQMN. Moreover, most of the general/roadside AQMSs are located in relatively flat regions. This is because almost all the densely-populated built-up areas of Hong Kong are in those flat regions at a relatively lower elevation. It is possible that the air quality estimation in those complex mountains (usually at a higher elevation) is less accurate than the estimation of those densely-populated flat areas. The lack of monitoring data is still a limitation of getting a good estimation in those complex mountains of Hong Kong. As already mentioned in Section 3.1, there is a long-term need for more AQMSs to fill existing monitoring gaps. To compensate this limitation in the short-term for the improvement of LUR modelling, more temporary sampling locations must be carefully selected and monitored in future works to further provide sufficient spatial information and a larger dependent variable dataset.

4. Conclusion

Estimating ambient air pollution concentration is essential to urban planning, policy decision-making and environmental management. On the basis of the resultant LUR models with the identified decisive variables on the concentration level of air pollutants, our findings have clear implications for urban planners, decision makers and public health officials. More importantly, there is a need to incorporate wind-related information into the LUR models that estimate the air quality condition for high-density cities with hilly topography. With reference to the quantitative correlation indicated by these resultant models, planners can work to improve urban air quality by enhancing wind availability through strategies such as reducing the FAI by controlling the building height limits and ground coverage, and decreasing the H/W ratio of urban street canyons. The existing local wind environment should be carefully considered before assigning land use type, making site selection and development plan. To be more specific, high-density or high-emission projects should be avoided at a location with low wind availability. For example, an industrial project should not be placed in the leeward side of a mountain, so that the pollution dispersion will not be hampered by insufficient wind.

The above air quality improvement strategies must be supplemented with strategies that identify and protect the more vulnerable communities (Chau et al., 2002; Ko et al., 2007). It is necessary to combine the LUR results with local public health assessment. Moreover, it should be emphasized that above environmental planning strategies definitely contribute more than just improve air quality. Enhancing the wind availability by implementing above the environmental planning strategies will also reduce the health risks from “heat waves”. This is particularly important to high-density cities like Hong Kong, because the high-density built environment makes it more vulnerable to the extreme weather condition. It has been proven the correlation between hot weather and health impacts are considerable (Chan et al., 2012; Yan, 2000). Improving the wind environment will also improve Hong Kong's resistance against extreme weather condition by enhancing thermal condition (Ng et al., 2011).

The linkage between urban air quality and urban planning is strong (Eliasson, 2000; Marquez and Smith, 1999). Essentially, the surface topography and aerodynamic properties are necessary components in urban planning, decision making and urban design. Therefore, with the increasing severity of air pollution related environmental issues, urban planning and design strategies are playing an increasingly important role in reducing health risks and vulnerability, building resilience and promoting living environmental quality.

Acknowledgement

This research is supported by the Postgraduate Studentship (PGS) 1155005856/PHD/ARK from The Chinese University of Hong Kong. The authors would like to thanks Professor Alexis Lau of the Hong Kong University of Science and Technology, Dr. Derrick Ho and Ms Ada Lee of the Chinese University of Hong Kong for their suggestion and help on this paper.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.envres.2017.05.007.

References

Ainslie, B., et al., 2008. A source area model incorporating simplified atmospheric dispersion and advection at fine scale for population air pollutant exposure assessment. Atmos. Environ. 42, 2394–2404.

Arain, M.A., et al., 2007. The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. Atmos. Environ. 41, 3453–3464.

Arnfield, A.J., 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. Int. J. Climatol. 23, 1–26.

Barnes, M.J., et al., 2014. Spatially-varying surface roughness and ground-level air quality in an operational dispersion model. Environ. Pollut. 185, 44–51.

Beelen, R., et al., 2013. Development of NO_2 and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe–the ESCAPE project. Atmos. Environ. 72, 10–23.

Böhner, J., Antonić, O., 2009. Land-surface parameters specific to topo-climatology. Dev. Soil Sci. 33, 195–226.

Bottema, M., 1996. Roughness parameters over regular rough surfaces: experimental requirements and model validation. J. Wind Eng. Ind. Aerodyn. 64, 249–265.

Bottema, M., 1997. Urban roughness modelling in relation to pollutant dispersion. Atmos. Environ. 31, 3059–3075.

Briggs, D.J., et al., 1997. Mapping urban air pollution using GIS: a regression-based approach. Int. J. Geogr. Inf. Sci. 11, 699–718.

Brook, R.R., 1972. The measurement of turbulence in a city environment. J. Appl. Meteorol. 11, 443–450.

Chan, E.Y.Y., et al., 2012. A study of intracity variation of temperature-related mortality and socioeconomic status among the Chinese population in Hong Kong. J. Epidemiol. Community Health 66, 322–327.

Chau, C.K., et al., 2002. Estimating the total exposure to air pollutants for different population age groups in Hong Kong. Environ. Int. 27, 617–630.

Chin, P.C., 1986. Climate and weather. In: Chiu, T.N. (Ed.), A Geography of Hong Kong. Oxford University Press HK, New York.

Cogliani, E., 2001. Air pollution forecast in cities by an air pollution index highly correlated with meteorological variables. Atmos. Environ. 35, 2871–2877.

Dozier, J., Frew, J., 1990. Rapid calculation of terrain parameters for radiation modeling from digital elevation data. IEEE Trans. Geosci. Remote Sens. 28, 963–969.

Duchene-Marullaz, P., Full scale measurement of atmospheric turbulence in suburban area. Fourth international conference on wind effects on buildings and structures, Heathrow, 1975.

Eeftens, M., et al., 2013. Quantifying urban street configuration for improvements in air pollution models. Atmos. Environ. 72, 1–9.

Eliasson, I., 2000. The use of climate knowledge in urban planning. Landsc. Urban Plan. 48, 31–44.

Eliasson, I., et al., 2006. Wind fields and turbulence statistics in an urban street canyon. Atmos. Environ. 40, 1–16.

Finardi, S., et al., 1998. Evaluation of different wind field modeling techniques for wind energy applications over complex topography. J. Wind Eng. Ind. Aerodyn. 74–76, 283–294.

Freund, R.J., et al., 2003. Regression using JMP. J. Wiley, Cary, NC (ISBN 1-59047-160-1).

Fung, J.C.H., et al., 2005. Observational and modeling analysis of a severe air pollution episode in western Hong Kong. J. Geophys. Res.: Atmospheres. 110, D09105.

Gál, T., Sümeghy, Z., 2007. Mapping the roughness parameters in a large urban area for urban climate applications. Acta Climatol. Chorol. 40–41.

Gál, T., Unger, J., 2009. Detection of ventilation paths using high-resolution roughness parameter mapping in a large urban area. Build. Environ. 44, 198–206.

Grimmond, C., Oke, T.R., 1999. Aerodynamic properties of urban areas derived from analysis of surface form. J. Appl. Meteorol. 38, 1262–1292.

Grimmond, C.S.B., 1998. Aerodynamic roughness of urban areas derived from wind observations. Bound. Layer. Meteorol. 89, 1–24.

Guo, H., et al., 2007. C1–C8 volatile organic compounds in the atmosphere of Hong Kong: overview of atmospheric processing and source apportionment. Atmos. Environ. 41, 1456–1472.

Gurjar, B.R., et al., 2010. Air Pollution: Health and Environmental Impacts. CRC Press, Boca Raton, FL (ISBN 978-1-43980-962-4).

HKO, 2015. Monthly Meteorological Normals for Hong Kong, 1981–2010. Hong Kong Observatory, Hong Kong.

Hoek, G., et al., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. 42, 7561–7578.

Hystad, P., et al., 2011. Creating national air pollution models for population exposure assessment in Canada. Environ. Health Perspect. 119, 1123.

Jennrich, R. I., 1977. Stepwise regression. Statistical methods for digital computers. 3, 58–75.

Kashima, S., et al., 2009. Application of land use regression to regulatory air quality data in Japan. Sci. Total Environ. 407, 3055–3062.

Kastner-Klein, P., Rotach, M., 2004. Mean flow and turbulence characteristics in an urban roughness sublayer. Bound. Layer. Meteorol. 111, 55–84.

Knibbs, L.D., et al., 2014. A national satellite-based land-use regression model for air pollution exposure assessment in Australia. Environ. Res. 135, 204–211.

Ko, F., et al., 2007. Effects of air pollution on asthma hospitalization rates in different age groups in Hong Kong. Clin. Exp. Allergy 37, 1312–1319.

Kwok, R.H., et al., 2010. Numerical study on seasonal variations of gaseous pollutants and particulate matters in Hong Kong and Pearl River Delta Region. J. Geophys. Res.: Atmos. (1984–2012) 115.

Landsberg, H.E., 1981. The Urban Climate. Academic Press, London.

Lau, A., et al., 2007. Relative significance of local vs. regional sources: Hong Kong's air pollution. Civ. Exch.

Liu, C.-H., et al., 2005. On the prediction of air and pollutant exchange rates in street canyons of different aspect ratios using large-eddy simulation. Atmos. Environ. 39, 1567–1574.

Macdonald, R.W., et al., 1998. An improved method for the estimation of surface roughness of obstacle arrays. Atmos. Environ. 32, 1857–1864.

Marquez, L.O., Smith, N.C., 1999. A framework for linking urban form and air quality. Environ. Model. Softw. 14, 541–548.

Miller, A., 2002. Subset Selection in Regression. Chapman and Hall/CRC, Boca Raton, FL (ISBN 978-1-58488-171-1).

Miller, A.J., 1984. Selection of subsets of regression variables. J. R. Stat. Soc. Ser. A (General). 147, 389–425.

Ng, E., et al., 2011. Improving the wind environment in high-density cities by understanding urban morphology and surface roughness: a study in Hong Kong. Landsc. Urban Plan. 101, 59–74.

Novotny, E.V., et al., 2011. National satellite-based land-use regression: NO2 in the United States. Environ. Sci. Technol. 45, 4407–4414.

Pasquill, F., 1971. Atmospheric dispersion of pollution. Q. J. R. Meteorol. Soc. 97, 369–395.

R.J. Allen, L.R.B. A.N.L.N., 1975. Air pollution dispersion modeling: application and uncertainty. J. Reg. Anal. Policy 5, 26.

Rafieian, M.. et al., 2014. The necessity of using Sky View Factor in urban planning: A case study of Narmak neighborhood, Tehran. Green Energy for Sustainable Development (ICUE), 2014 International Conference and Utility Exhibition on, pp. 1–5.

Ratti, C., Richens, P., 1999. Urban texture analysis with image processing techniques. In: Augenbroe, G., Eastman, C. (Eds.), Computers in Building. Springer, US, pp. 49–64.

Raupach, M.R., 1992. Drag and drag partition on rough surfaces. Bound. Layer. Meteorol. 60, 375–395.

Raupach, M.R., Finnigan, J.J., 1997. The influence of topography on meteorogical variables and surface-atmosphere interactions. J. Hydrol. 190, 182–213.

Ryan, P.H., LeMasters, G.K., 2007. A review of land-use regression models for characterizing intraurban air pollution exposure. Inhal. Toxicol. 19, 127–133.

Sall, J., et al., 2012. JMP Start Statistics: A Guide To Statistics and Data Analysis using JMP. SAS Institute Inc.,, Cary, NC (ISBN 978-1-61290-204-3).

Seaman, N.L., 2000. Meteorological modeling for air-quality assessments. Atmos. Environ. 34, 2231–2259.

Sin, W.-m., et al., 2002. A study of polychlorinated dibenzo- < i > p < /i > -dioxins and dibenzofurans in the atmosphere of Hong Kong. Chemosphere 47, 647–653.

Su, J.G., et al., 2008. An innovative land use regression model incorporating meteorology for exposure analysis. Sci. Total Environ. 390, 520–529.

Su, J.G., et al., 2009. A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. Sci. Total Environ. 407, 3890–3898.

Tabachnick, B.G., Fidell, L.S., 2001. Using multivariate statistics.

Tabachnick, B.G., Fidell, L.S., 2007. Using Multivariate Statistics. Pearson/Allyn & Bacon, Needham Heights, MA, USA (ISBN 978-0-20545-938-4).

Tang, R., et al., 2013. Using building heights and street configuration to enhance intraurban PM10, NOx, and NO2 Land use regression models. Environ. Sci. Technol. 47, 11643–11650.

Tong, H., et al., 2005. Numerical simulation of the urban boundary layer over the complex terrain of Hong Kong. Atmos. Environ. 39, 3549–3563.

Tu, Y.-K., et al., 2005. Problems of correlations between explanatory variables in multiple regression analyses in the dental literature. Br. Dent. J. 199, 457–461.

Vienneau, D., et al., 2009. A GIS-based method for modelling air pollution exposures across Europe. Sci. Total Environ. 408, 255–266.

Wang, T., et al., 2001. A study of surface ozone and the relation to complex wind flow in Hong Kong. Atmos. Environ. 35, 3203–3215.

Watson, I.D., Johnson, G.T., 1987. Graphical estimation of sky view-factors in urban environments. J. Climatol. 7, 193–197.

Wieringa, J., 2001. New revision of Davenport roughness classification. Proceedings,

3EACWE, Eindhoven, The Netherlands. 285-292.
Yan, Y.Y., 2000. The influence of weather on human mortality in Hong Kong. Soc. Sci. Med. 50, 419–427.
Yan, Y.Y., 2007. Surface wind characteristics and variability in Hong Kong. Weather 62, 312–316.
Yang, J., et al., 2005. The urban forest in Beijing and its role in air pollution reduction. Urban For. Urban Green. 3, 65–78.
Yuan, Z., et al., 2006. Identification and spatiotemporal variations of dominant PM10 sources over Hong Kong. Atmos. Environ. 40, 1803–1815.
Zhu, T., et al., 2012. Impacts of Megacities on Air Pollution and Climate. World Meteorological Organization (WMO), Switzerland.
Wan, Zues, et al., 2013. Hong Kong Emission Inventory Report Air Science Group. Hong Kong Environmental Protection Department, Hong Kong, pp. 2015.