Front. Earth Sci. 2017, 11(2): 347–360
DOI 10.1007/s11707-016-0593-0

RESEARCH ARTICLE

Prediction of vertical $PM_{2.5}$ concentrations alongside an elevated expressway by using the neural network hybrid model and generalized additive model

Ya GAO[1], Zhanyong WANG[1], Qing-Chang LU (✉)[1], Chao LIU[2], Zhong-Ren PENG (✉)[1,2], Yue YU[1,3]

1 Center for ITS and UAV Applications Research, State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
2 Department of Urban and Regional Planning, University of Florida, FL 32611-5706, USA
3 Department of Electrical and Systems Engineering, University of Pennsylvania, PA 19104-6314, USA

Abstract A study on vertical variation of $PM_{2.5}$ concentrations was carried out in this paper. Field measurements were conducted at eight different floor heights outside a building alongside a typical elevated expressway in downtown Shanghai, China. Results show that $PM_{2.5}$ concentration decreases significantly with the increase of height from the 3rd to 7th floor or the 8th to 15th floor, and increases suddenly from the 7th to 8th floor which is the same height as the elevated expressway. A non-parametric test indicates that the data of $PM_{2.5}$ concentration is statistically different under the 7th floor and above the 8th floor at the 5% significance level. To investigate the relationships between $PM_{2.5}$ concentration and influencing factors, the Pearson correlation analysis was performed and the results indicate that both traffic and meteorological factors have crucial impacts on the variation of $PM_{2.5}$ concentration, but there is a rather large variation in correlation coefficients under the 7th floor and above the 8th floor. Furthermore, the back propagation neural network based on principal component analysis (PCA-BPNN), as well as generalized additive model (GAM), was applied to predict the vertical $PM_{2.5}$ concentration and examined with the field measurement dataset. Experimental results indicated that both models can obtain accurate predictions, while PCA-BPNN model provides more reliable and accurate predictions as it can reduce the complexity and eliminate data co-linearity. These findings reveal the vertical distribution of $PM_{2.5}$ concentration and the potential of the proposed model to be applicable to predict the vertical trends of air pollution in similar situations.

Received October 23, 2015; accepted March 27, 2016

E-mails: qclu@sjtu.edu.cn (Qing-Chang LU), zrpeng@sjtu.edu.cn (Zhong-Ren PENG)

Keywords vertical variations, principal component analysis, back propagation neural network, generalized additive model, urban elevated expressway

1 Introduction

There appears extensive evidence that vehicles are a major source of the particulate matter, especially the fine fraction of particulate matte ($PM_{2.5}$), which causes adverse health effects (Mazzoleni et al., 2004; Wang et al., 2006; McNabola et al., 2009; Schleicher et al., 2011). Since the beginning of 21st century, many researchers have investigated the horizontal dispersion of vehicular emissions on roadsides, but the vertical variation of particulate matter is still a matter of discussion. Because of the limited space in cities, most buildings in areas mixed with industry, residential housing, and commerce are high-rise towers of close proximity. People living in different floor heights may inhale different amounts of vehicular particulate matter. Hence, it is important to explore the vertical variations of air pollutants such as particulate matter and to investigate how various factors influence variation of air pollutants.

Studies for vertical concentration of particles mainly focus on how particle number concentration, size distribution and mass concentration change with height above road level. Li et al. (2007) measured particle number size distribution in the range of 10 nm to 487 nm at four heights at an asymmetric street. Similar experiments were conducted by Kumar et al. (2008) and Longley et al. (2004). Their findings suggested that particle size distribution (between 5 to 1000 nm) changed significantly as height increases. Weber et al. (2006) and Colls and Micallef (1999) found that the form of PM profiles was

quite analogical with heightabove the ground. In addition, they reported the mass concentration of PM increases with height in the lower part of the canyon and then decreases in the upper part of the canyon. Chan and Kwok (2000) further strengthened the discovery that the decreasing pattern may be exponential. However, few studies addressed the relationship between traffic and meteorological factors and PM vertical variation, or discussed the potential methods for predicting the PM vertical variation on urban roadside.

For a better investigation of PM variations, many approaches, such as simulation models and statistical models, are proposed to estimate the effects of various factors on the prediction of traffic-related pollutant concentration. The simulation model focuses on $k-\varepsilon$ turbulence and pollutant concentration diffusion equations (He et al., 2009), which are mostly based on numerical simulation technology to predict the turbulent flow and pollutant concentration fields in an urban street canyon. Kumar et al. (2009) and Zhang and Batterman (2010) identified the model defection and improved the prediction of near-road air pollutant concentration by comparing simulation models with statistical models. Although their results indicate that the simulation model can be applied to different situations, it may be insufficient under some specific traffic and meteorological conditions due to the special laboratory tests of vehicle emissions measured during typical driving cycles. Compared with frequently used numerical simulation technology, statistical models can incorporate site-specific information. The general statistical models such as linear regression are often used to relate air pollutant concentration with traffic volume, wind speed, temperature, and other independent variables, while the certified non-linearly coupling relationship between predictor variables and air pollutant concentration is inadequately captured by a simple statistical model. The traffic-related particle matters at roadside are characterized by comprehensive linear and non-linear patterns and difficult to predict. Some non-linear statistical models, such as the stochastic model, rely on a prior assumptions of the data distribution and can't eliminate the multi-collinearity issue (Milionis and Davies, 1994). Accordingly, these models are not computationally manageable and are complicated as well (Nagendra and Khare, 2006; Muñoz et al., 2014).

Unlike the above-mentioned approaches, the artificial neural networks (ANNs) and generalized additive model (GAM), have been proven to be robust and sophisticated methods, as they account for non-linear relationships among different field sites (Schlink et al., 2003; Cai et al., 2009). The back-propagation neural network (BPNN) is a solution to the problem of training multi-layer perceptron. The fundamental advantages of the BPNN are the inclusion of a differentiable transfer function at each node of the network and the use of error back-propagation to modify the internal network weights after each training epoch. To overcome the co-linearity among the input variables, principal components analysis (PCA) can be applied before modeling by artificial neural networks (He et al., 2014; Wang et al., 2015a). GAM blends properties of the additive model and generalized linear model and involves a sum of smooth functions of covariates. The only underlying model assumption is that the functions are additive and the components are smooth. Generally, the GAM model includes a parametric and a non-parametric part (Hastie and Tibshirani, 1990). Some researchers have concluded that the GAM and BPNN are both methodologically comparable in the satisfactory performance due to their ability in modeling static non-linearity (Schlink et al., 2003). Based on the field experiment dataset in our study, it is imperative and meaningful to verify whether these two proposed statistical methods (i.e., BPNN and GAM) are suitable for predicting vertical variations of $PM_{2.5}$ concentration on roadsides.

In this study, the experimental dataset was collected to explore the vertical variations of $PM_{2.5}$ concentration alongside a typical elevated expressway in downtown Shanghai, China. In Chinese megacities, the elevated expressway is a typical roadway pattern with a special design of road geometry, vehicle speed and composition which generate serious vehicular emissions and become a conduit for air pollutants dispersion. As an increasing number of elevated expressways are built to alleviate traffic congestion, dwellers alongside the elevated expressway are increasingly threatened by the danger of high morbidity and mortality. Hence, it is important to investigate vertical concentration distribution of $PM_{2.5}$ at these locations. The objects of this study are: (i) to illustrate vertical variation of $PM_{2.5}$ concentration alongside the elevated expressway; (ii) to estimate the effects of different factors on the vertical distribution of $PM_{2.5}$ concentration; (iii) to verify the effectiveness of methods for $PM_{2.5}$ prediction through a comparison of the PCA-BPNN and generalized additive model.

2 Field experiment

2.1 Study area

Experimental data were used to evaluate the capabilities of two prediction models in this research. Sites selected for experiment met the following criteria: 1) road with relative high traffic volume; 2) places without pollutant influence from surrounding such as industrial plants and parking lots. The data used in this paper was collected at the Shangzhong Building, which is located at the southern side of the street and near the Middle-Ring Elevated Expressway above the Shangzhong Road. Both roads have busy traffic in the east-west direction and are major pollutant

sources. The width of Shangzhong Road is 27 m and the length of the selected segment is 156 m. The mean height of the buildings studied is 56 m on the south side and 28 m on the north side (Fig.1(a)).

Our measurements were conducted at the Shangzhong Building, which sits on the southern side of the street without a major North-to-South avenue nearby. We assume that the pollutants come from emissions of vehicles running on Shangzhong Road and the Middle-Ring Elevated Expressway (Fig.1(b)). Figure 1(c) illustrates the different monitoring locations at the Shangzhong Building. The 3rd, 5th, and 7th floor are selected to detect the pollutant dispersion under the elevated expressway. The data collected at the 8th, 9th, and 11th floor can measure the concentration alongside the elevated expressway because the height of the elevated expressway is 20 m. The other points (13th and 15th) were selected to test the pollutant dispersion patterns in the upper part of street canyon.

2.2 Field experiment design

All of the sensors were deployed at different heights of the Shangzhong Building, 0.5 m away from the building wall on the north side and 2.5 m away from the sidewalk. Two sets of portable monitors were used to detect pollutant at the different monitoring heights. $PM_{2.5}$ concentration was captured by a TSI Sidepak AM510 detector, which is a portable device designed to monitor and record fine PM concentrations on second-by-second scale. Traffic volumes from Shangzhong Road and Middle Ring Elevated Expressway were separately recorded by two video cameras and then counted on a 5-min interval by volume counting software considering that it is a stable fine-time scale to reduce sampling randomness. Meteorological factors such as relative outdoor humidity, temperature, wind direction, wind speed, and sun radiation were measured on the roof of the Shangzhong Building by Davis Vantage Weather Station (Li et al., 2007; Kumar

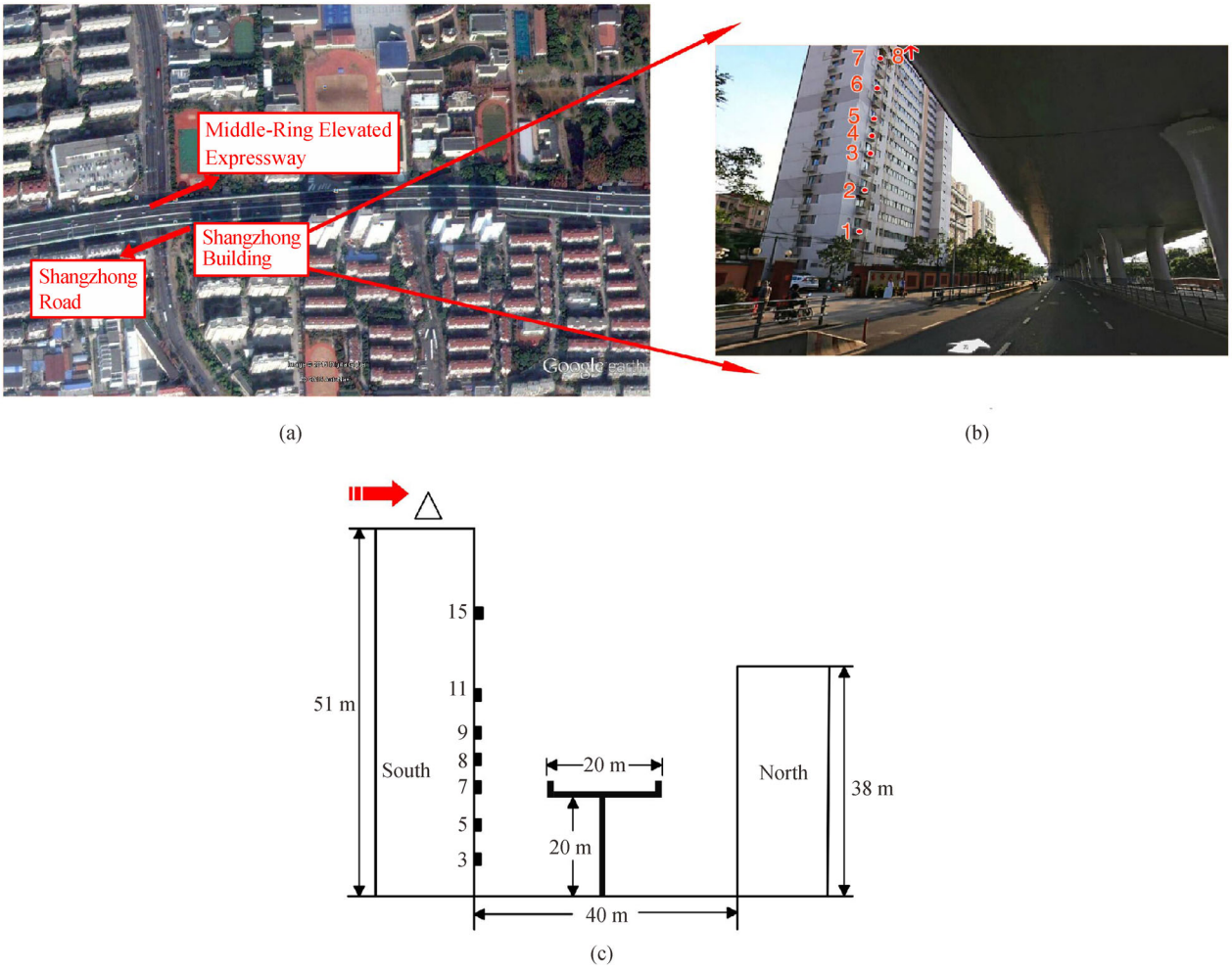

Fig. 1 Schematic of field measurement. (a) Field measurement site. (b) Shangzhong Building. (c) The distribution of monitoring sites in vertical section and within the street canyon.

et al., 2008). Meteorological factors were recorded every minute during the experiment. As said in our previous studies (Wang et al., 2015a,b; Wang et al., 2016), all the $PM_{2.5}$ detectors were calibrated before leaving factory and further estimated with standard methods at outdoor locations in Shanghai, China prior to this field study.

The experiment was conducted on 14th February and 23th February with two periods of each day, i.e., morning (7:00 a.m. to 12:00 p.m.) and afternoon (1:00 p.m. to 6:00 p.m.). $PM_{2.5}$ concentration was measured at each height as Li et al. (2007) and Kumar et al. (2008). It is noted that we only have two sets of portable instruments in hand to simultaneously measure two heights. To acquire a representative data set at each sampling height, it took about 15 min complete two sampling on two heights (i.e. 15 min per height) and about 1 h to complete one set at all the eight heights. Sampling was done in ten sets of the measurements every day. The first 3 minutes' data from the $PM_{2.5}$ detector was deleted to guarantee the accuracy of the concentration at each height. That is because the sensors need a buffer time to get the same concentrations as that of the environment. Except for vehicle emissions, the impacts of resident, industrial, and restaurant emissions cannot be completely ignored. To approximate other sources' impacts, we collected the background concentrations as an input factor prior to prediction. The background concentrations of $PM_{2.5}$ were extracted from the hourly monitored data of Shanghai Environmental Monitoring Centre. The detectors at the Shanghai Environmental Monitoring Centre neglected the impacts of some factors such as arterial roads, parking lots, and industrial plants, and such monitors were deployed at a university which is approximately 3.5 km away from the Shangzhong Building.

2.3 Traffic and meteorological conditions

Statistical descriptions of traffic and Meteorological parameters from the field observations are presented in Table 1. Traffic flow system presents significant impacts on social economic and urban environment (Tang et al., 2015a; Zhang and Zhu, 2015). And there are too many traffic-related factors such as driving behavior and road condition contributing to fuel consumption and emissions (Tang et al., 2015b, c). In this research, traffic-related factors focus on traffic flow and vehicle type. Vehicles are divided into 3 categories, i.e., small cars, buses, and trucks. Small cars are considered as light-duty vehicles, while buses and trucks are considered as heavy-duty ones. The average traffic volumes is 1000−1600 vehicles per hour which are composed of 94% of the light-duty vehicles and 6% of the heavy-duty vehicles.

The variation of wind speeds and directions is shown in Fig. 2. The wind goes across the street canyon with a mean speed of 1.322 $m \cdot s^{-1}$ during the experimental period. As shown in Fig. 1(c), the monitoring sites are at the leeward side of the street canyon.

3 Methodology

3.1 Wilcoxon rank-sum test

In statistics, the Wilcoxon rank-sum test is a non-parametric test to assess whether two samples of observations come from the same distribution. (Ng et al., 2007). Here, it is applied to test the difference between the data measured under the 7th floor (under elevated expressway) and above the 8th floor (above elevated expressway) accordingly. As described by He and Lu (2012), two sets of samples need to be combined into one set of samples and then sorted in increasing order. The sum of ranks is defined by the following equation:

$$W = \sum_{i=1}^{N} R_i - \frac{N(N+1)}{4}, \quad (1)$$

where, N is the sample size, and R_i represents the rank sum of each group. If two populations have the same distribution, the sum of the two sample ranks should be close to the same value. In addition, the p-value is also calculated to test the null hypothesis that the two samples of observations come from the same population in terms of confidence levels. In this study, Wilcoxon rank-sum test were implemented with the software SAS.

Table 1 Descriptions of Traffic and Meteorological Parameters

	Mean	SD	Max	Min
$Traffic_G$/(vehicles·h^{-1})	1273	456	1875	778
$Traffic_F$/(vehicles·h^{-1})	4375	2899	5786	3989
Heavy-duty vehicles rate/%		5.13%($Traffic_G$)	0.83%($Traffic_F$)	
Temperatures/°C	10.68	0.952	11.80	9.10
Wind speed/($m \cdot s^{-1}$)	1.322	0.883	3.35	0.00
Relative outdoor humidity/%	54.462	3.335	60.50	49.00
Sun radiation	683.576	138.665	781.154	480.413

Note: $Traffic_G$ and $Traffic_F$ indicate traffic volumes from Shangzhong road and Middle-Ring Elevated Expressway.

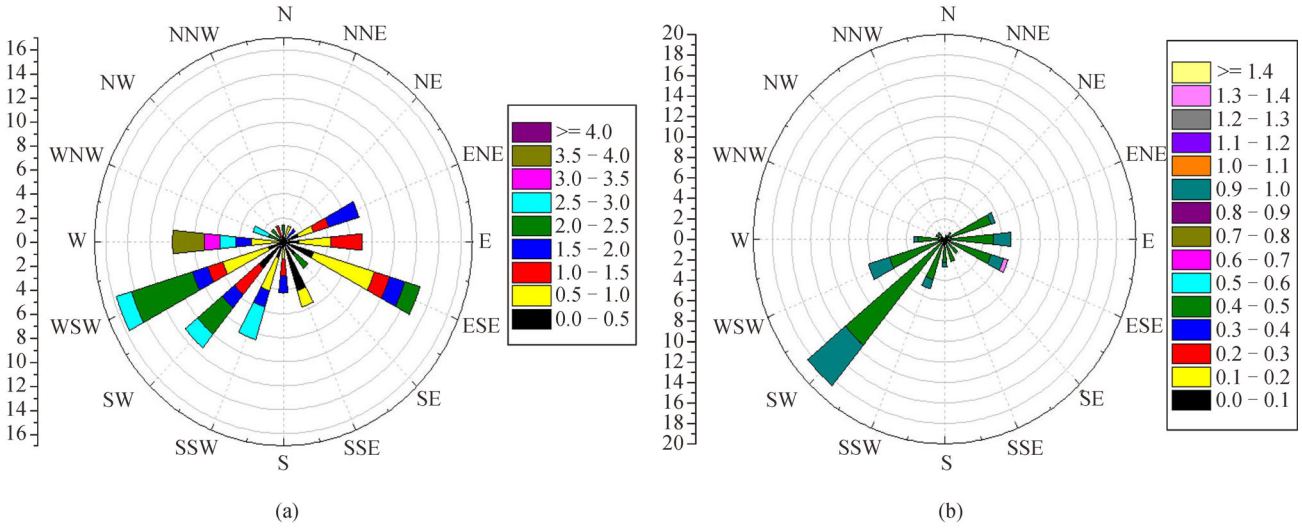

Fig. 2 Wind rose of wind speed and direction measured in the experiment period: (a) February 14th. (b) February 23th.

3.2 Correlation coefficient and PCA

Generally, factors that have a significant impact on air pollutants in a street canyon are determined and classified into four categories: traffic-related, weather-related, location-related, and background-related (Zhang and Peng, 2014). According to this classification, seven factors, including traffic volume, temperature, relative outdoor humidity, wind speed, wind direction, and height have been selected in our study.

Essentially, correlation coefficient analysis is a measure that determines the degree to which two variables' movement are associated. PCA is a multivariate technique that transforms the original set of inter-correlated variables into a new set of an equal number of independent uncorrelated variables or principal component (PC) that are in linear combinations of the original variables (Sousa et al., 2007). It can extract the important information from a data table, as its first PC explains most of the variance and each subsequent PC accounts for the largest proportion of variability that has not been accounted for by its predecessor(s) (Abdul-Wahab et al., 2005). Rotated factor loadings represent the influence of each variable in a specific PC were calculated by varimax rotation. The higher rotated factor loadings, the more that variable contributes to the variation. The PCs with approximate 85% cumulative amounts of variance are already enough to contain information of original variables.

The main purpose of this analysis lies in the following two aspects: 1) to analyze the correlations between the mass concentration of $PM_{2.5}$, height, meteorological conditions, and traffic volume; 2) to generate PCs as input variables to reduce the dimension of the dataset with minimal loss of information.

3.3 Statistical models

The variations of fine particulate matter concentrations in street canyon are mainly originated from traffic volume and are strongly governed by certain meteorological conditions. Obviously, the relationships among these variables are all nonlinear, and as such it is difficult to apply linear models to examine them. Fortunately, the artificial neural network model and generalized additive model have been recently developed and verified as cost-effective approaches for analyzing nonlinear environmental problems (Aldrin and Haff, 2005; Carslaw et al., 2007; He et al. 2014; Wang et al., 2015a, b). With these considerations, these two models are separately adopted in this paper and their performances are compared.

3.3.1 Back propagation neural network based on PCA

The back propagation algorithm (BP) is a well-known method of training a multilayer feed-forward artificial neural network, which consists of three distinctive layers (input layer, hidden layer and output layer). In this article, variables from traffic conditions, metrological conditions, local conditions, and background concentrations were transformed into PCs through PCA method and the input variables of back propagation neural network are selected from these PCs. The architecture of a PCA-BPNN model is shown in Fig. 3.

In order to avoid the asymptotic effect, the input and output data are normalized between 0.1 and 0.9 with the following equation (Wang et al., 2015a,b):

$$X_{\text{norm}} = 0.1 + (X_i - X_{\min})/(X_{\max} - X_{\min})*(0.9 - 0.1), \quad (2)$$

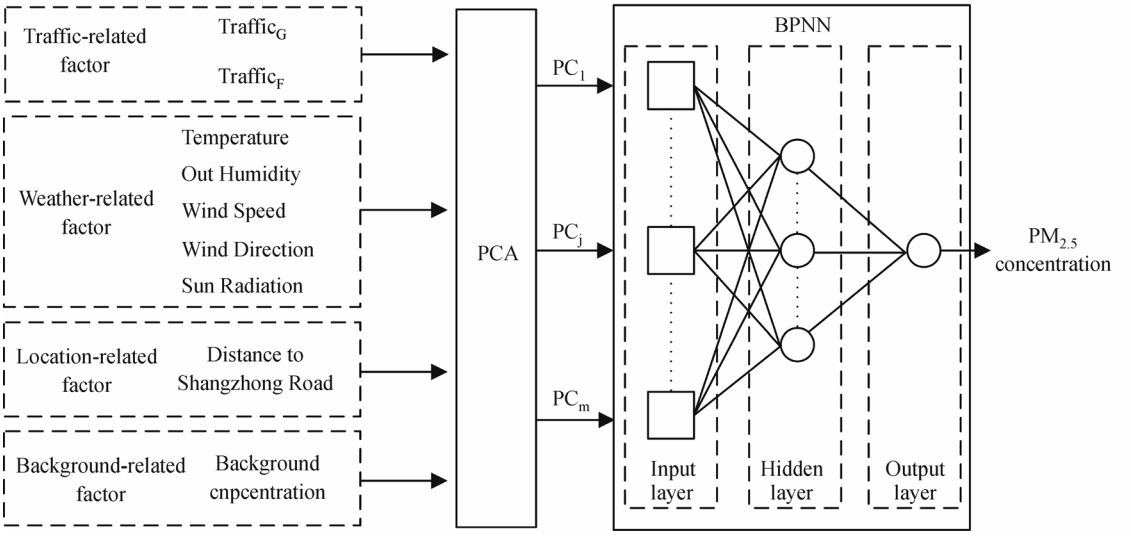

Fig. 3 Architecture of a PCA-BPNN model.

where X_i refers to the increment pollution concentrations, $X_{\max}$ and $X_{\min}$ are maximum and minimum value of X_i. The outputs are transformed back to real values after prediction.

The activation of each unit is determined by a simple and widely-recognized sigmoid transfer function (Moseholm et al., 1996):

$$y_j = f(net) = \frac{1}{1+e^{-net_j}}, \tag{3}$$

where, net_j is the state of the j neuron in the hidden layer, and

$$net_j = \sum_i w_{ji}x_i - \theta_j, \tag{4}$$

where, w_{ji} is the weight from unit i to$j$, and θ_j is the bias invariant for unit$j$. In order to minimize the total errors in BP training method based on delta rule, the modified increment Δw_{ji} of the weight could be obtained:

$$\Delta w_{ji}(n+1) = \eta\sum_i \delta_{ij}y_j + \alpha\Delta\omega_{ji}(n), \tag{5}$$

where, n is the presentation number, η is the learning rate that controlled the convergence speed to the minimum of errors and δ_{ij} is the error signal for unit j, and α is the momentum factor. In this paper, according to the previous research by Cai et al. (2009), we set $\eta = 0.3$. As a result, the input and output of the hidden layer and the output layer are calculated by the following equation:

$$hi_h(k) = \sum_{i=1}^{n} w_{ih}x_i(k) - b_h,\ ho_h(k) = f(hi_h(k))$$
$$= 1/(1+\exp(-\sum_{\mathrm{i}=1}^{n} w_{ih}x_i(k) - b_n))$$

$$yi_o(k) = \sum_{h=1}^{p} w_{ho}ho_h(k) - b_0,\ yo_0(k) = f(yi_o(k))$$
$$= 1/(1+\exp(-\sum_{\mathrm{h}=1}^{n} w_{ho}ho_h(k) - b_n)), \tag{6}$$

where, $hi_h(k)$ and $ho_h(k)$ are the input and output of the hidden layer, $yi_o(k)$ and $yo_0(k)$ are the input and output of the out layer.

The models' performance in development and validation steps was evaluated with the following statistical parameters: R (correlation coefficient) was used to measure the linear relation between observed and predicted values, while root mean squared error (RMSE) test residual errors to measure how close the predicted and observed values are. Mean bias error (MBE) indicates if the observed concentrations are over- or under-estimated. The value of index of agreement (IA) indicate the degree of error free for the prediction with the following equation (Gardner and Dorling, 2000; Chaloulakou et al., 2003),

$$IA = 1 - \frac{\left[\sum_{i=1}^{n} |\hat{Y} - Y_i|^2\right]}{\left[\sum_{i=1}^{n} \left(|\hat{Y} - \overline{Y}_i| + |Y - \overline{Y}_i|\right)^2\right]}. \tag{7}$$

3.3.2 GAM

The generalized additive model with no assumptions about the parametric relationship between variables, makes it very attractive when there exists complicated nonlinearity in the multivariate case. The model can be described as (Hastie and Tibshirani, 1990),

$$Y_i = A_i + \sum_{j=1}^{n} S_j(x_{ij}) + \varepsilon_i,\ \varepsilon_i N(0,\sigma^2), \quad (8)$$

where Y_i is the concentration of the time series, $S_j(x_{ij})$ is a smooth function of covariatex_j, n is the number of covariates, ε is the residual.

Some predictor variables were used in the model including temperature, out humidity, wind speed, wind direction, sun radiation, height, and traffic volumes from Shangzhong road ($Traffic_G$) and Middle-Ring Elevated Expressway ($Traffic_F$). In this model, the W_s (wind speed) and W_d (wind direction) interact with each other. Hence, Ws and Wd are modeled as an interaction term of the wind components, i.e., $s(u,v)$ (Carslaw et al., 2007). Our basic model is a generalized additive model with Gaussian response as described above. In order to assess the specific contributions of various factors to the pollutant, models having a explicit response-predictor relationship in each case could be formulated as follows,

$$\begin{aligned}\log(PM_{2.5}) = {} & s_1(u,v) + s_2(\mathrm{Temp}) + s_3(\mathrm{Hum}) + s_4(\mathrm{Sun.Rad}) \\ & + s_5(\mathrm{Height}) + s_6(\mathrm{Traffic_G}) + s_7(\mathrm{Traffic_F}) + \varepsilon \\ u = {} & W_s \bullet \sin e(W_d) \\ v = {} & W_s \bullet \cos ine(W_d), \end{aligned} \quad (9)$$

Where, $s(u,v)$ is the bivariate smooth function of wind components u and v, while s(Temp), s(Hum), s(Sun.Rad), s(Height), $s(\mathrm{Traffic_G})$, $s(\mathrm{Traffic_F})$ are the smooth function of temperature, relative humidity, sun radiation, height, traffic volume from Shangzhong road and traffic volume from Middle-Ring Elevated Expressway respectively. The penalized regression splines, which while optimizing the fit, penalize roughness, were used to integrate model selection and automatic smoothing parameter selection (Wood and Augustin, 2002).

4 Results and discussion

4.1 Vertical variation of $PM_{2.5}$ concentration

During the field experiment, 2400 groups of 1-min samples of $PM_{2.5}$ concentrations and meteorological factors and 358 groups of 5-min traffic records were collected. Subsequently, five continuous 1-min $PM_{2.5}$ concentration and meteorology samples were regularized into one 5-min average in order to match the time step of traffic series. Data of typical sunny day (14th) was selected for further statistical analysis. Hence, a total of 175 groups of 5-min samples of all variables are available.

Figure 4 presents the vertical variation of $PM_{2.5}$ concentration with height. From the figure, it is obvious that a sudden increment from the 7th floor to the 8th floor exists. Due to the division of Middle Ring Elevated Expressway, two vehicular emission sources seldom interact with each other and vehicular pollutants do not diffuse easily near the 8th floor which is the same height as the elevated expressway with a 2 m noise-proof wall. Hence this height which can represent the exposure of the entire local population should be the applicable height for pollutant measurements to investigate the horizontal vehicular pollutant dispersion. The same phenomenon has been proven by a fluid mechanics experiment in China (Wang and Huang, 2002). Middle Ring Elevated Expressway acting like a "hat" separates the upper layer (higher than the 8th floor, hereafter named Case II) from the down layer (lower than the 7th floor, hereafter named Case I). So the pollutant can only diffuse through the gap between the elevated expressway and the surrounding buildings. What's more, the pollutant produced in the elevated expressway are likely to be rolled to the ground by the airflow whirlpool, resulting in a more serious pollution at the lower layer.

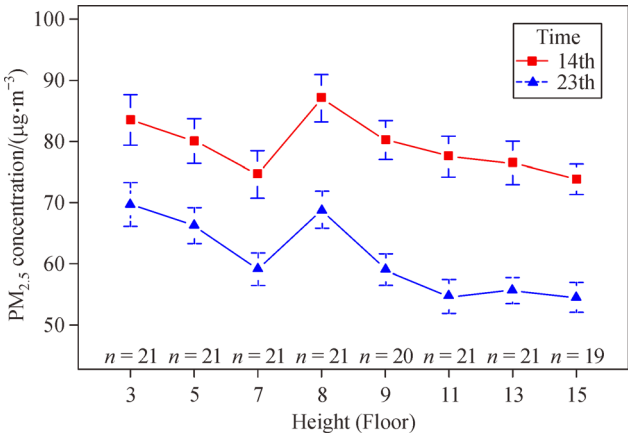

Fig. 4 Vertical variations of $PM_{2.5}$ concentration with height.

In Case I and Case II, the concentration of $PM_{2.5}$ decreases with the increases of height. The minimum value of the concentration appears at the Middle Ring Elevated Expressway and rooftop level, respectively. And in Case I, the goodness of fit for exponential curve of $PM_{2.5}$ is 0.980. This is consistent with previous findings by Li et al. (2007) and Chan and Kwok (2000) who found $PM_{2.5}$ concentration exponentially decrease with height. But the fitting curve of $PM_{2.5}$ in Case II are unstable in our study.

The Wilcox rank-sum test is employed to assess the difference between two sets of observations. As shown in Table 2, all p-values are smaller than 0.05. Hence, it can be deduced that vertical variance of $PM_{2.5}$ concentration is different under the 7th floor (Case I) and above the 8th floor (Case II).

4.2 Pearson correlation and principal component analysis on influencing factors

To better identify the influence of each original variable on

Table 2 Results of Wilcox-in Rank-Sum test and correlation coefficient analysis.

	$PM_{2.5}$	Height	Temp	Hum	W_s	W_d	Sun.Rad	$Traffic_G$	$Traffic_F$
$PM_{2.5}$		**−0.22***	**−0.78****	**0.79****	**−0.24***	**−0.13**	**−0.69****	**0.66****	**0.26***
Height	**−0.49****		0.43**	−0.14	0.13	0.11	0.16	0.25*	0.18
Temp	**0.72****	−0.72**		−0.92**	0.09	0.44**	0.79**	0.57**	0.64**
Hum	**0.68****	0.57**	−0.79**		−0.21	−0.49**	−0.23*	−0.47**	−0.12
W_s	**−0.33****	−0.38**	−0.26**	−0.1		0.12	0.33**	−0.03	0.34**
W_d	**0.11**	0.04	0.32**	−0.213*	−0.35**		0.44**	0.09	0.51**
Sun.Rad	**−0.25***	0.35**	0.85**	−0.12	−0.15	0.45**		0.36**	0.70**
$Traffic_G$	**0.18**	0.18	−0.01	0.104	−0.02	−0.06	−0.056		0.25*
$Traffic_F$	**0.43****	−0.28*	−0.23*	0.118	0.071	0.44**	0.832**	0.101	
Wilcoxin				1757.5					
P-value				0.007					

Note: (1) ** There is a significant relation between two variables on the level of 0.01; * There is a significant relation between two variables on the level of 0.05. (2) The upper right panel is Case I; The lower left panel is Case II. (3) $Traffic_G$ and $Traffic_F$ indicate traffic volumes from Shangzhong road and Middle-Ring Elevated Expressway. (4) Hum, Sun.Rad, W_s, and W_d mean relative outdoor humidity, sun radiation, wind speed and wind direction, respectively.

the vertical distribution of $PM_{2.5}$ concentration, the correlation analysis was carried out first (Table 2). Correlation coefficients in bold font were calculated to quantitatively investigate the relationship between $PM_{2.5}$ concentration and influencing variables. The traffic volume remained a crucial impact on the $PM_{2.5}$ concentration. However, the correlations between $PM_{2.5}$ concentration and traffic volumes in both cases were generally different. For example, $PM_{2.5}$ concentration is significantly positively correlated with traffic volumes from Shangzhong Road ($Traffic_G$) in Case I and significantly positively correlated with traffic volumes from Middle Ring Elevated Expressway ($Traffic_F$) in Case II, which was also found in CFD simulation experiment in China (Zhang et al., 2012).

Meteorological factors such as temperature and relative outdoor humidity play important roles in both cases (bold font in Table 2). In addition, the significant positive or negative correlations should be a reflection of the physical response mechanism. For example, high humidity often appears in windless, cloudy, and weak sunshine days, which are favorable for the accumulation and chemical reaction of pollutants. The significant negative correlations with wind speed indicated that air pollutant concentrations decrease with the increasing wind speed, which are conducive to the dispersion of air pollutants. Notably, wind speeds start to play a more important role in Case II compared to the situations in Case I. The cause of this phenomenon is the "hat" that barricades the exchange of air flow from bottom to top of the street canyon.

From the Table 2, it can also be found that the corresponding dependent factors have a close relationship with each other at the 5% significance level, indicating the existence of multi-collinearity among these variables. The collinearity between these variables makes the $PM_{2.5}$ prediction inaccurate and complex. With this consideration, PCA was applied before the BPNN model to eliminate the collinearity between the input variables and reduce the complexity. The result of PCA is discussed below.

The eigenvalues and cumulative amounts of 8 original variables were calculated through the PCA and are shown in Fig. 5(a). In Case I, three PCs are identified accounting for 89.28% of the variances. PC1 is applied for measuring the relevance among the temperature, relative outdoor humidity, and sun radiation. In Case II, 4 PCs are responsible for 91.07% of the total variance. PC1 shows a significant loading for temperature, relative outdoor humidity, sun radiation, and traffic volume from the Middle Ring Elevated Expressway. For both cases, PC1 has high loadings in the case of temperature, relative outdoor humidity, and sun radiation, suggesting that these three variables are likely to have significant impacts on air pollutant concentrations. In addition, communalities of original variables are shown in Fig. 5(b). It reveals that almost all the factors have a varied communality with different strength between the two cases. In the table, it can be observed that all variables present communalities that can explain at least half of their own variance in Case I. But wind direction and traffic volume from Shangzhong Road with communalities below 0.5 in Case II can be dismissed, as they do not have any explanation power for their variances.

4.3 Evaluation of the PCA-BPNN and GAM models on $PM_{2.5}$ prediction

4.3.1 Performances of the PCA-BPNN model

Eight original factors that influenced the concentration and dispersion of $PM_{2.5}$ were incorporated into the PCA, as shown in Fig. 5. These factors were classified into three categories: location-related, meteorological-related, and

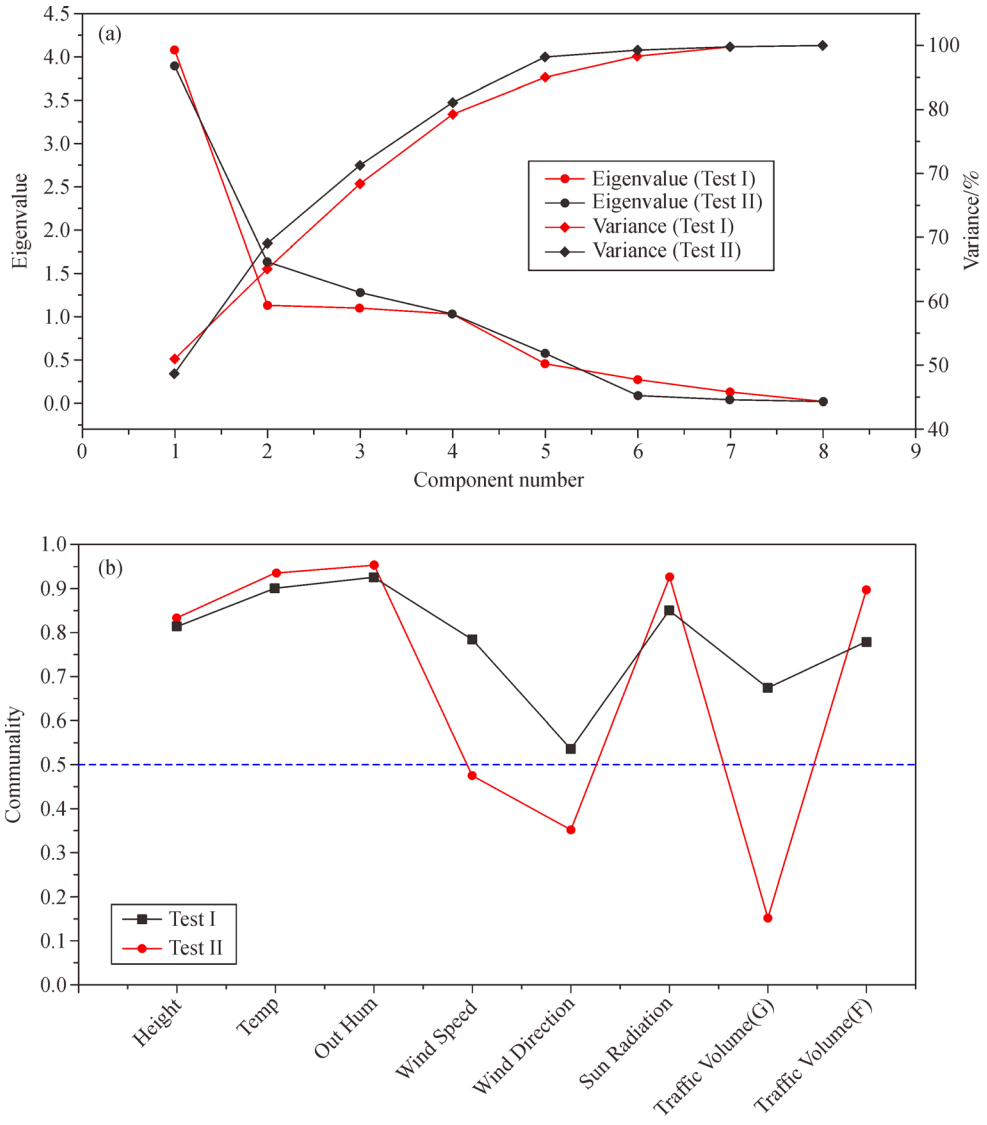

Fig. 5 PCA results of original variables for different cases. (a) Scree plot and respective cumulative variance (%). (b) Communalities of original variables.

traffic related. In the modeling of Case I (69 samples) and Case II (109 samples), 75% and 25% of the samples were assigned for training and testing, respectively.

Neural network is sensitive to the number of neurons in the hidden layers. The number of neurons in this paper was tested and validated for Case I and Case II, respectively. From 3 neurons to 20 neurons, MRE, MAE, and MSE were used to evaluate the estimation results. Eventually, the optimum number of hidden neurons was determined as 5 for Case I and 13 for Case II. The expected mean squared error, maximum learning iteration, and training rate were 0.001, 50000, and 0.08 respectively. With one output (i.e., $PM_{2.5}$), three PC inputs in Case I and four PC inputs in Case II were allocated for training and testing. Then, BPNN was built for Case I and Case II, respectively.

Based on the subsets of Case I and Case II, the best models were trained for predicting $PM_{2.5}$ concentration and the statistical results are shown in Table 3. PCA-BPNN has a better goodness of fit with lower RMSE, higher IA and R, although the positive MBE value indicates a slight over-prediction. In addition, the performances during the training and testing progresses are quite close, which

Table 3 Performance of PAC-BPNN model for prediction.

		Case I		Case II	
		Training	Testing	Training	Testing
$PM_{2.5}$	R	0.97	0.89	0.96	0.97
	IA	0.99	0.96	0.99	0.99
	RMSE	1.36	2.29	1.61	1.94
	MBE	−0.03	0.55	0.06	0.11

indicate that the models have a good generalization capacity. Figure 6 presents the observed versus training and testing $PM_{2.5}$ concentration in Case I and Case II. From the figure, it can be seen that the predictions in Case II are better than those of Case I which are in agreement with the observed values.

4.3.2 Performances of the GAM model

Compared with the neural-network model, the generalized additive model not only can describe non-linearity and interactions, but also can sort out and quantify the separate effect of each predictor variable (Aldrin et al., 2005). In this paper, the final models that use appropriate meteorological and traffic covariate derived for each pollutant in different tests are given by Eq. (9).

There are many factors related to variations of traffic emission and meteorology that influence $PM_{2.5}$ concentrations. Thus it is of significance to know how the factors affect the prediction of concentration trend. According to the approximate significance of smooth terms in Table 4, the impacts of $Traffic_G$ and $Traffic_F$ in both cases are different in that $Traffic_G$ is a major source in Case I ($PTraffic_G < 0.01$, $PTraffic_F = 0.0635$) and $Traffic_F$ becomes a more important contribution to $PM_{2.5}$ concentrations in Case II ($PTraffic_F < 0.01$, $PTraffic_G = 0.0413$). At the significance level of 0.01, the *s*(Height) and *s*(Hum) are found to be distinct factors that influence the $PM_{2.5}$

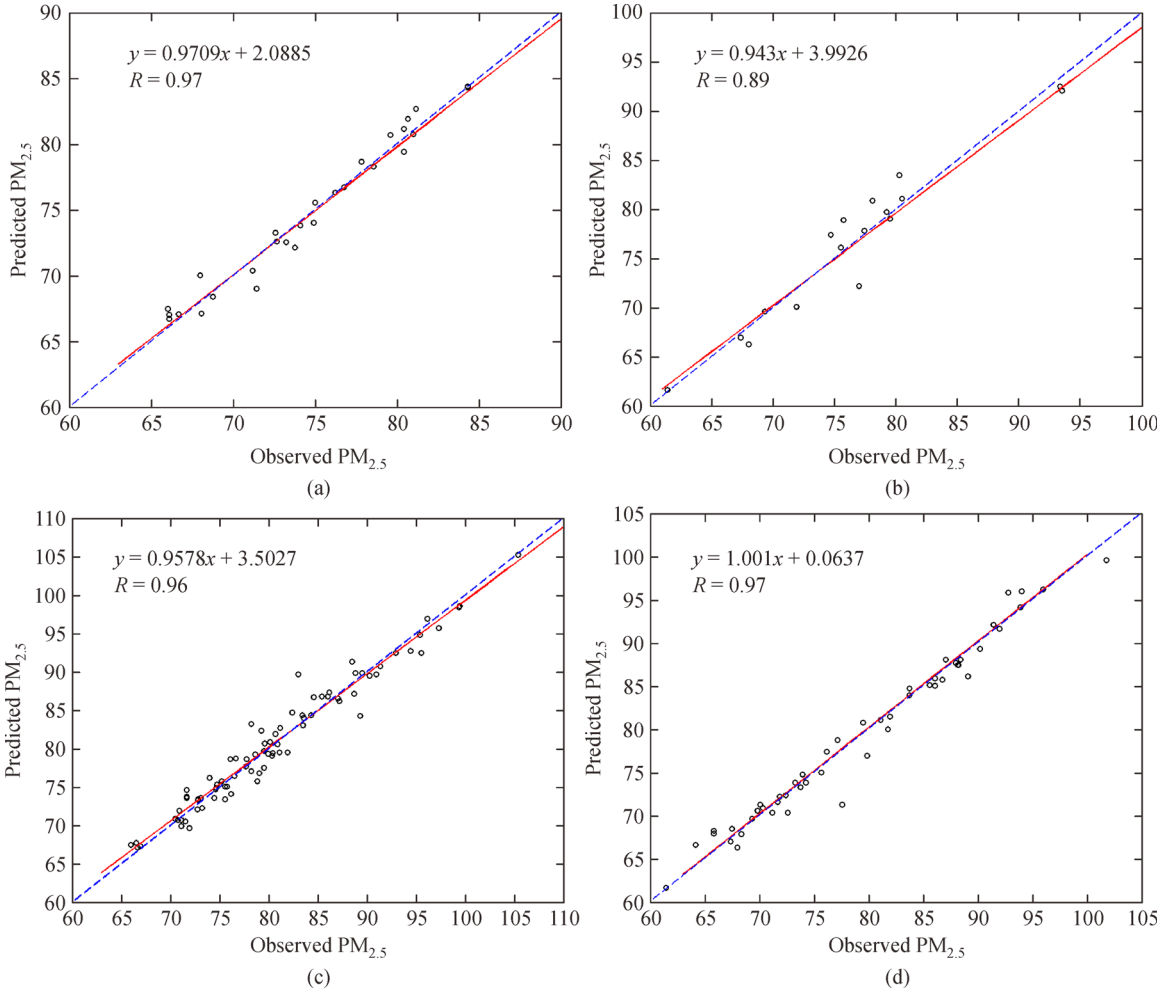

Fig. 6 Observed versus training and testing $PM_{2.5}$ concentrations (all concentration in $\mu g/m^3$). (a) training Case I. (b) testing Case I. (c) training Case II. (d) testing Case II.

Table 4 Significance of smooth terms in Case I and Case II.

	s(Height)	*s*(u,v)	*s*(Temp)	*s*(Hum)	*s*(Sun.Rad)	*s*($Traffic_G$)	*s*($Traffic_F$)
Case I	5.40e−0.5**	0.0215*	0.0633*	1.83e−0.5**	0.2902	0.00013**	0.1762
Case II	1.96e−0.7**	3.57e−0.8**	0.4438	< 2e−1.6**	0.6350	0.6613	< 2e−1.6**

Note: (1) ** There is a significant relation between two variables on the level of 0.01; * There is a significant relation between two variables on the level of 0.05. (2) *s* (Hum), *s*(Sun.Rad) and *s*(*u*, *v*) mean relatively outdoor humidity ,sun radiation and wind components, respectively.

concentrations in both cases. The $s(u, v)$ that depicts the distribution of $PM_{2.5}$ concentrations was more important in Case II ($P < 0.01$) than in Case I ($P = 0.0215$). It can be explained that the presence of an elevated expressway changed the flow structure within the street canyon and the dispersion characteristic of vehicle emission. Most of the results were consistent with the correlation coefficient analysis.

The estimated smooth non-linear curves are shown in Fig. 7. Some increasing factors (Traffic Volume, Relative outdoor Humidity, Temperature, Sun Radiation) bear a substantial relationship to the increase of $PM_{2.5}$. Apart from this main trend, there is a rather large variation in the estimated effects in Case II. From the bivariate surface smooth plot for $s(u, v)$, we can find in Fig. 6(b) that the increasing concentrations which were induced by the increasing wind speed is opposite to how the wind speed behaves under most situations, where an increasing plume dilution can be expected. This phenomenon pointed out the complex wind flow pattern on the street canyon with an elevated expressway. The performance of GAM prediction together with the PCA-BPNN model is discussed below.

4.3.3 Performance comparison of the PCA-BPNN and GAM models

The predicting results of the $PM_{2.5}$ concentrations by using GAM models in both cases are exhibited in Table 5. The R values between modeling and observed for $PM_{2.5}$ on log-scale are 0.818 in Case I and 0.888 in Case II respectively, which indicate a satisfactory goodness of fit. It also can be seen that GAM model explains a large amount of the daily variation of $PM_{2.5}$ concentration. Fig. 8 depicts the observed and predicted $PM_{2.5}$ concentrations based on samples from eight different heights (floors) using two methods proposed in this study. From the figure, it is clear that the PCA-BPNN model provides more reliable and accurate prediction of $PM_{2.5}$ concentration than the GAM model (RPCA-BPNN = 0.89 vs RGAM = 0.82, RPCA-BPNN = 0.97 vs RGAM = 0.89). This is probably due to the fact that the PCA-BPNN model combines the merits of PCA to overcome the co-linearity between the input variables. GAM is also suitable for handling nonlinear associations, but it is difficult to eliminate multi-collinearity that generally leads to the incorrect identification of the most important prediction. Conclusively, the PCA-BPNN model is a reliable alternative for evaluating the $PM_{2.5}$ vertical concentration with a reasonable accuracy in the prediction.

5 Conclusions

In this study, a field measurement was conducted to characterize the vertical variations of $PM_{2.5}$ concentration alongside a typical elevated expressway in downtown Shanghai, China. Through the data analysis, the $PM_{2.5}$ concentrations below the 7−8th floors present a variation different from that above the 7−8th floors, although both cases follow an exponential distribution. $PM_{2.5}$ concentration is also found to change significantly with the increase of height from the 3th to 15th floor, while there is a sudden increment from the 7th floor to 8th floor which is the same height as the elevated expressway with a 2 m noise-proof wall. This height should be the applicable height for pollutant measurements to investigate the horizontal dispersion of vehicular emissions in similar situations. A non-parametric test further demonstrates that the distribution of $PM_{2.5}$ concentrations is different under the 7th floor (Case I) and above the 8th floor (Case II).

Additionally, three classes of factors that have significant impacts on $PM_{2.5}$ concentrations, i.e., weather-related, traffic-related, and location-related factors, were used to investigate their relationships with $PM_{2.5}$ concentrations. Pearson correlation analysis indicates that traffic-related factor has a crucial impact on $PM_{2.5}$ concentration and meteorological factors also play important roles which is consistent with previous studies. Apart from this main trend, there is a rather large variation in correlation coefficients between two cases.

This work not only focused on interpretation of empirical relationships between $PM_{2.5}$ vertical variation and traffic, height, and a set of meteorological factors, but also discussed alternative methods for forecasting the trend of $PM_{2.5}$ concentration. The vertical distribution of $PM_{2.5}$ concentration consists of complex linear and non-linear patterns and are difficult to predict. BPNN and GAM models that can describe both non-linearity and interaction have been applied to air quality prediction.

The GAM used in this study is additive on the log-scale producing residual distribution that were close to normal. To overcome potential interactions between wind speed and wind direction, the s(wind speed) + s(wind direction) is replaced by an interaction term s(wind speed, wind direction). In addition, GAM was also applied to sort out and quantify the separate effect of each influencing factor. BPNN frequently used in modelling the nonlinear air pollution process has limited accuracy due to the collinearity between the input variables. Hence, before modeling the PCA was implemented to generate PCs as input variables to reduce the data complexity and collinearity. The results showed that PCA-BPNN and GAM can obtain more accurate predictions with higher R, IA, and lower RMSE.

Finally, the performances of BPNN based on PCA and GAM on prediction of $PM_{2.5}$ concentration were compared. The relevant results suggest that the PCA-BPNN model performs better than GAM in predicting the $PM_{2.5}$ concentration. This is probably due to the fact that the PCA-BPNN model combines the merits of PCA to overcome the co-linearity between the input variables.

To an extent, this study has revealed vertical variations

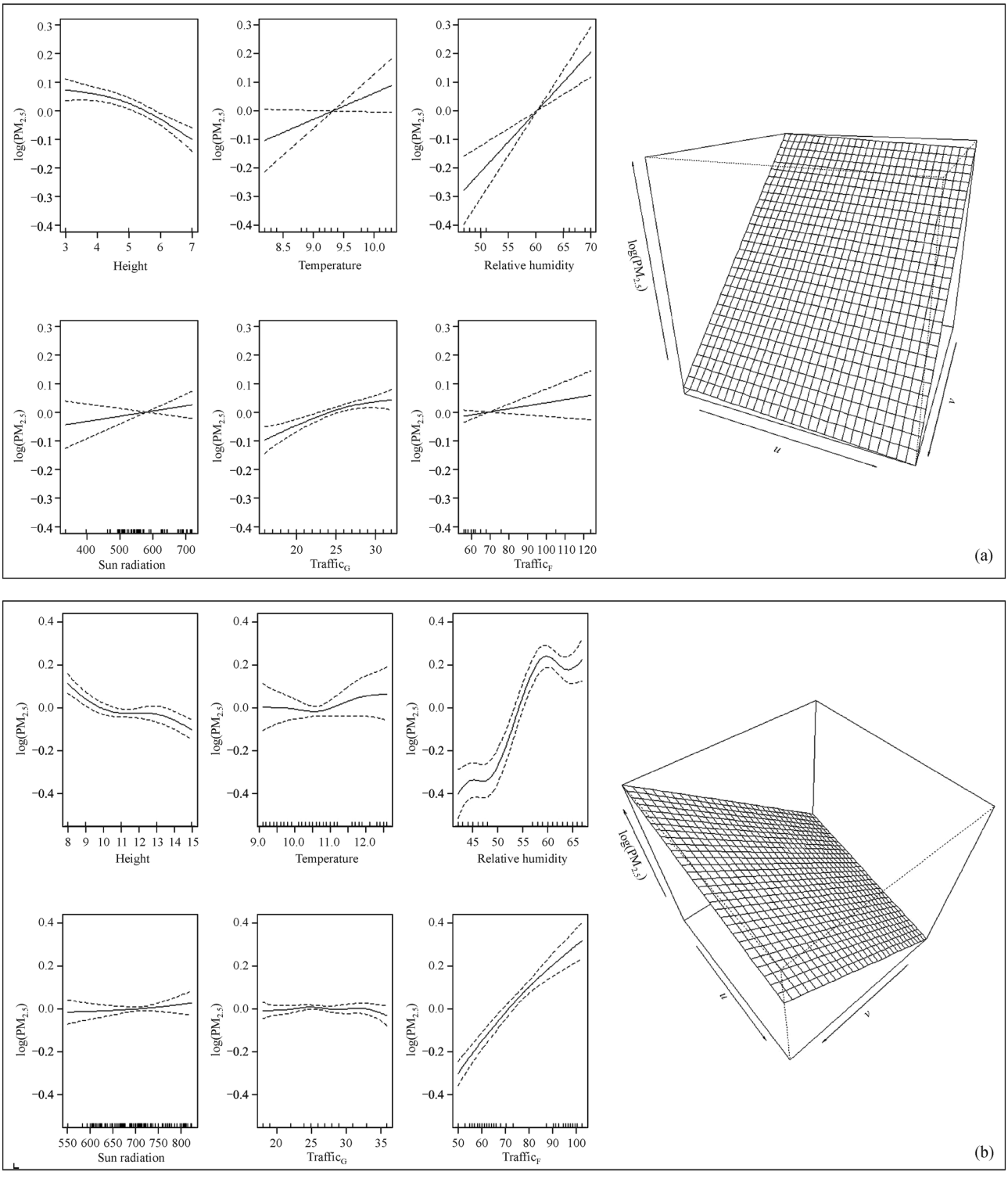

Fig. 7 Estimated effects (on the original scale) of high, temperature, relatively outdoor humidity, sun radiation, traffic volumes from Shangzhong road and Middle-Ring Elevated Expressway and bivariate wind components. (a) Case I. (b) Case II. The dashed lines are the estimated 95% confidence intervals.

Table 5 Performance of GAM model for prediction

		Case I	Case II
$PM_{2.5}$	R	0.82	0.89
	IA	0.99	0.97
	RMSE	5.02	0.06

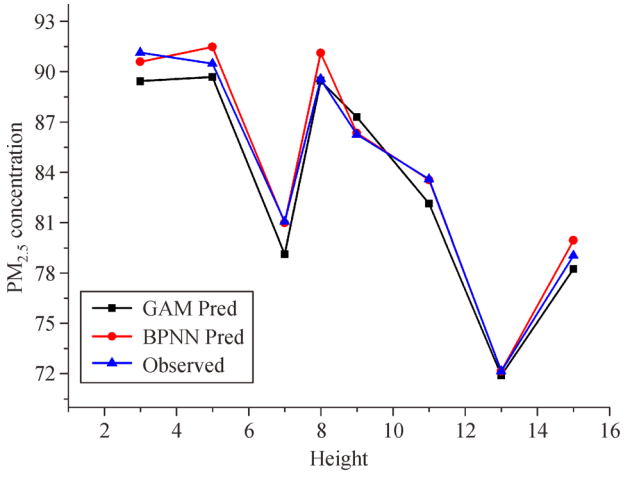

Fig. 8 Comparison of observations and predictions by the models at eight-floor height.

of $PM_{2.5}$ concentration alongside a typical elevated expressway, although it should be emphasized that the results based on the proposed method are limited to a specific case and are encouraged to be applied to other different cases. Further research is recommended to take into account many other pollutants such as ultrafine particles and black carbon, and to input new datasets to further verify these models.

Acknowledgements The authors would like to acknowledge the support from Shanghai Environmental Protection Bureau, Shanghai Environmental Monitoring Center, Science Technology Department of Zhejiang Province (2014C31028), Peking University-Lincoln Institute (DS20120901), and State Key Laboratory of Ocean Engineering of China (GKZD010059). We thank members from Shanghai Environmental Monitoring Center for their assistance in instrumental calibration. We also appreciate members from the Center for ITS and UAV Applications Research at Shanghai Jiao Tong University for their hard work in data collection and processing.

References

Abdul-Wahab S A, Bakheit C S, Al-Alawi S M (2005). Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. Environ Model Softw, 20(10): 1263–1271

Aldrin M, Haff I H (2005). Generalised additive modelling of air pollution, traffic volume and meteorology. Atmos Environ, 39(11): 2145–2155

Cai M, Yin Y, Xie M (2009). Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. Transp Res Part D Transp Environ, 14(1): 32–41

Carslaw D C, Beevers S D, Tate J E (2007). Modelling and assessing trends in traffic-related emissions using a generalised additive modelling approach. Atmos Environ, 41(26): 5289–5299

Chan L Y, Kwok W S (2000). Vertical dispersion of suspended particulates in urban area of Hong Kong. Atmos Environ, 34(26): 4403–4412

Colls J J, Micallef A (1999). Measured and modelled concentrations and vertical profiles of airborne particulate matter within the boundary layer of a street canyon. Sci Total Environ, 235(1−3): 221–233

Chaloulakou A, Saisana M, Spyrellis N (2003). Comparative assessment of neural networks and regression models for forecasting summertime ozone in Athens. Science of the Total Environment, 313 (1): 1–13

Gardner M W, Dorling S R (2000). Statistical surface ozone models: an improved methodology to account for non-linear behaviour. Atmospheric Environment, 34(1): 21–34

Hastie T J, Tibshirani R J (1990). Generalized additive models. London: Chapman and Hall

He H D, Lu W Z (2012). Urban aerosol particulates on Hong Kong roadsides: size distribution and concentration levels with time. Stochastic Environ Res Risk Assess, 26(2): 177–187

He H D, Lu W Z, Xue Y (2014). Prediction of particulate matters at urban intersection by using multilayer perceptron model based on principal components. Stochastic Environ Res Risk Assess, 29(8): 2107–2114

He J, Qi Z, Zhao C, Bao X (2009). Simulations of pollutant dispersion at toll plazas using three-dimensional CFD models. Transp Res Part D Transp Environ, 14(8): 557–566

Kumar P, Fennell P, Langley D, Britter R (2008). Pseudo-simultaneous measurements for the vertical variation of coarse fine and ultrafine particles in an urban street canyon. Atmos Environ, 42(18): 4304–4319

Kumar P, Garmory A, Ketzel M, Berkowicz R, Britter R (2009). Comparative study of measured and modelled number concentrations of nanoparticles in an urban street canyon. Atmos Environ, 43(4): 949–958

Li X, Wang J, Tu X D, Liu W, Huang Z (2007). Vertical variations of particle number concentration and size distribution in a street canyon in Shanghai, China. Sci Total Environ, 378(3): 306–316

Longley I D, Gallagher M W, Dorsey J R, Flynn M (2004). A case-study of fine particle concentrations and fluxes measured in a busy street canyon in Manchester, UK. Atmos Environ, 38(22): 3595–3603

Mazzoleni C, Moosmüller H, Kuhns H D, Keislar R E, Barber P W, Nikolic D, Nussbaum N J, Watson J G (2004). Correlation between automotive CO, HC, NO, and PM emission factors from on-road remote sensing: implications for inspection and maintenance programs. Transp Res Part D Transp Environ, 9(6): 477–496

McNabola A, Broderick B M, Gill L W (2009). The impacts of inter-vehicle spacing on in-vehicle air pollution concentrations in idling urban traffic conditions. Transp Res Part D Transp Environ, 14(8): 567–575

Milionis A E, Davies T D (1994). Box-Jenkins univariate modelling for climatological time series analysis: an application to the monthly activity of temperature inversions. International Journal of Climatol-

ogy, 14(5): 569–579.

Moseholm L, Silva J, Larson T (1996). Forecasting carbon monoxide concentrations near a sheltered intersection using video traffic surveillance and neural networks. Transp Res Part D Transp Environ, 1(1): 15–28

Muñoz E, Martin M L, Turias I J, Jimenez-Come M J, Trujillo F J (2014). Prediction of PM10 and SO_2 exceedances to control air pollution in the Bay of Algeciras, Spain. Stochastic Environ Res Risk Assess, 28(6): 1409–1420

Nagendra S S, Khare M (2006). Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions. Ecol Modell, 190(1−2): 99–115

Ng H K T, Balakrishnan N, Panchapakesan S (2007). Selecting the best population using a test for equality based on minimal Wilcoxon rank-sum precedence statistic. Methodology and Computing in Applied Probability, 9(2): 263–305

Schleicher N J, Norra S, Chai F, Chen Y, Wang S, Cen K, Yu Y, Stüben D (2011). Temporal variability of trace metal mobility of urban particulate matter from Beijing–A contribution to health impact assessments of aerosols. Atmos Environ, 45(39): 7248–7265

Schlink U, Dorling S, Pelikan E, Nunnari G, Cawley G, Junninen H, Greig A, Foxall R, Eben K, Chatterton T, Vondracek J, Richter M, Dostal M, Bertucco L, Kolehmainen M, Doyle M (2003). A rigorous inter-comparison of ground-level ozone predictions. Atmos Environ, 37(23): 3237–3253

Sousa S I V, Martins F G, Alvim-Ferraz M C M, Pereira M C (2007). Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environ Model Softw, 22(1): 97–103

Tang T Q, Huang H J, Shang H Y (2015c). Influences of the driver's bounded rationality on micro driving behavior, fuel consumption and emissions. Transp Res Part D Transp Environ, 41: 423–432

Tang T Q, Yu Q, Yang S C, Ding C (2015a). Impacts of the vehicle's fuel consumption and exhaust emissions on the trip cost allowing late arrival under car-following model. Physica A: Statistical Mechanics and its Applications, 431: 52–62

Tang T Q, Li J G, Yang S C, Shang H Y (2015b). Effects of on-ramp on the fuel consumption of the vehicles on the main road under car-following model. Physica A: Statistical Mechanics and its Applications, 419: 293–300

Wang J S, Chan T L, Ning Z, Leung C W, Cheung C S, Hung W T (2006). Roadside measurement and prediction of CO and PM2.5 dispersion from on-road vehicles in Hong Kong. Transp Res Part D Transp Environ, 11(4): 242–249

Wang J S, Huang Z (2002). Numerical study on impact of urban viaduct on local-scale of atmospheric environment. Shanghai Environmental Sciences, 21(3): 132–135

Wang Z, He H D, Lu F, Lu Q C, Peng Z R (2015a). Hybrid model for prediction of carbon monoxide and fine particulate matter concentrations near a road intersection. Transp Res Rec, 2503: 29–38

Wang Z, Lu F, He H D, Lu Q C, Wang D, Peng Z R (2015b). Fine-scale estimation of carbon monoxide and fine particulate matter concentrations in proximity to a road intersection by using wavelet neural network with genetic algorithm. Atmos Environ, 104: 264–272

Wang Z, Lu Q C, He H D, Wang D, Gao Y, Peng Z R (2016). Investigation of the spatiotemporal variation and influencing factors on fine particulate matter and carbon monoxide concentrations near a road intersection. Front. Earth Sci., doi: 10.1007/s11707-016-0564-5

Weber S, Kuttler W, Weber K (2006). Flow characteristics and particle mass and number concentration variability within a busy urban street canyon. Atmos Environ, 40(39): 7565–7578

Wood S N, Augustin N H (2002). GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecol Modell, 157(2−3): 157–177

Zhang C J, Zeng J R, Wen M, Zhang G L, Fang H P, Li Y (2012). Influence of Viaducts on Dispersion of Air Particles in Street Canyons. Research of Environmental Sciences, 25(2): 159–164.

Zhang D Z, Peng Z R (2014). Near-road fine particulate matter concentration estimation using artificial neural network approach. Int J Environ Sci Technol, 11(8): 2403–2412

Zhang K, Batterman S (2010). Near-road air pollutant concentrations of CO and $PM_{2.5}$: A comparison of MOBILE6.2/CALINE4 and generalized additive models. Atmos Environ, 44(14): 1740–1748

Zhang L D, Zhu W X (2015). Delay-feedback control strategy for reducing emission of traffic flow system. Physica A: Statistical Mechanics and its Applications, 428: 481–492