Exploration of spatial and temporal characteristics of PM2.5 concentration in Guangzhou, China using wavelet analysis and modified land use regression model

Fenglei Fan\textsuperscript{a,b} and Runping Liu\textsuperscript{a}

\textsuperscript{a}School of Geography, South China Normal University, Guangzhou, China; \textsuperscript{b}Joint Laboratory of Plateau Surface Remote Sensing, Tibet University, Lhasa, China

1. Introduction

The air quality has always been a big concern and attracts tremendous attention in China’s mega-cities, especially the more international ones, such as Beijing (Zhao et al. 2009), Shanghai (Huang et al. 2013), Hong Kong (Shi et al. 2012), and Guangzhou (Garland et al. 2008). PM2.5, one class of PM, is the designation for particulars that are less than 2.5 µm in aerodynamic diameter, making it small enough to be inhaled. In recent decades, the focus has shifted more toward the effects of long-term exposure to PM2.5 at moderate levels typically seen in inhabited regions (Samet et al. 2000; Leiva et al. 2013). In addition, some studies show that PM2.5 in the city has a trend to exacerbate weather conditions; PM2.5 pollution in the Pearl River Delta (PRD) region is closely related with the continuously frequent occurrence of haze.

Due to the temporal and spatial independence of the pollutants, the characteristics of PM2.5 resolved in one region cannot be replicated to another region (Shi et al. 2012). Some studies in PRD show that the daily concentration of PM is higher in winter than in summer. Some studies also have found that the concentration of PM2.5 is higher in winter than that in other periods both in China (Zhao et al. 2009), middle eastern (Elbayoumi et al. 2013), USA (Lee et al. 2011) and Europe (Marczazan et al. 2001). Other studies have confirmed the truth of significant differences both in temporal and spatial distribution whether in China (Li et al. 2008, 2017a) or in other countries, such as Philadelphia (Burton, Suh, and Koutrakis 1996), Europe (Eeftens et al. 2012). On the other hand, the spatial differences between cities are also analyzed (Yu 2010). It has been observed that PM2.5 concentrations vary to differing degrees in the urban areas examined in the United States and considerable spatial variation is found in other regions, especially in the West (Pinto, Lefohn, and Shadwick 2004). Although there are some existing studies involving PM2.5 in Guangzhou or PRD, they focus mainly on the chemical composition (Wang et al. 2006; Tao et al. 2009). Recently, some important studies on PM2.5 or on other subject, such as land surface temperature (Wu et al. 2015), using Earth observation satellite data are finished, which expands the research scope of PM2.5 or other subject largely. These important studies mostly refers to fusing (Shen et al. 2016) between satellite data and station data to estimate ground-level PM2.5 (Li et al. 2017b).

Wavelet analysis is a powerful method of time series analysis, especially helpful for long time series change because of its ability to provide localized information on a time series in both the time and the frequency domains (Vadrevu and Choi 2011). Wavelet analysis generates an estimate of the local frequency content of a signal by representing the data using a family of wavelet functions that vary in scale.
and position (Chen et al. 2013). Wavelet analysis has been successfully applied in the time series analysis for climate (Kumar and Foufoula 1997), micrometeorology (Sang 2013) and atmospheric environment (Strunin and Hiyama 2005). In this study, Daubechies (DB) wavelets, an important function of wavelet analysis, are used to obtain the representation of time series of PM2.5 concentrations due to its compact support and orthogonality.

Land use regression (LUR) model is a robust technique used to predict concentrations of air pollutants at a given site based on surrounding physical, land use, and traffic characteristics (Jerrett et al. 2005). LUR model usually combines air pollutants concentrations (at a number of sites) and stochastic models using predictor variables obtained through geographic information system (GIS) (Jerrett et al. 2005). The model has been widely applied to simulate the spatial distribution (Hochadel et al. 2006) of PM2.5, NOx (Wang et al. 2012) and Volatile Organic Compounds (VOCs) (Wang et al. 2013). Significant predictor variables in LUR model include: traffic, population, land use and physical geography (Wang et al. 2013).

This paper is aimed to explore temporal and spatial variations of PM2.5 levels in Guangzhou. The objectives include: (1) detailing the time series changes of PM2.5 using wavelet analysis; (2) seeking the characteristics of PM2.5 with a seasonal view; and (3) discussing the spatial distribution of PM2.5 using GIS tools and LUR model.

2. Study area and datasets

2.1. Study area

Guangzhou, located in southern China and in the center of the PRD, is the third largest city of China and comprises 12 districts. A single Thematic Mapper (Landsat 4-5) (TM) or Enhanced Thematic Mapper Plus (Landsat 7) (ETM)+ scene can cover Guangzhou (Figure 1).

2.2. PM2.5 data (1 June 2012 to 31 May 2013)

Daily concentrations (1 June 2012 to 31 May 2013) of PM2.5 are obtained from the environmental monitoring center of the environmental protection administration of Guangdong. Those 13 ground-monitoring sites located within the Guangzhou metropolitan area (Liwanxichun: LWXC; Gongyuanqian: GYQ; Haizhubaogang: HZBG; Tianhecheng: THC; Haizhucisa: HZCS; Huangpuadasatou: HPDST; Panyushiqiao: PYSQ; Huadongchenggu: HDCQ; Luogangzhenlong: LGZL; Luhu: LH; Tianhu: TH; Zudong: ZD; Modiesha: MDS) are distributed in nine districts (Figure 1).

2.3. Impervious surface area

Impervious surface area is used to rebuild the LUR model. Impervious surface area can be obtained using linear spectral mixture analysis method (LSMA). Impervious surface area is the artificial construction that prevents water entering into the soil (Weng and Lu 2008). The roof, square and the land for transportation are the main components of the impervious surface area. In this study, impervious surface area is extracted based on Landsat ETM+ image (11 November 2010) (Figure 2). At the same time, four buffers of each monitoring site in different radiuses (1, 2, 3 and 4 km) are built and the areas of impervious surface in each buffer are calculated.

2.4. Vector traffic network and population density

Road network information is important data for the LUR model, which is provided by Guangzhou Urban Planning Department with vector layout (can be
handled by GIS software). Attributes of the roads include road name, length, district, code number and road classification. The vector traffic network is used to calculate the total length of all roads within several buffers (1, 2, 3 and 4 km). The length of all roads within each buffer of 13 monitoring sites is calculated and summed using GIS tools. The population densities data are obtained from the Guangzhou Statistic Yearbook (2013) and are applied as the population variable of the LUR model.

2.5. Wind index

Wind index is another important variable of LUR model. According to the meteorological statistics of 1995–2005, Guangzhou is a monsoonal climate city with two prevailing winds. The prevailing wind is demonstrated in Figure 3. Northerly (0°) and south-easterly (135°) are two prevailing wind directions. To account for the location of monitoring site relative to the wind direction and the nearest traffic source, the Euclidean direction from the nearest source (0°–360°) is derived. The wind index can be calculated using following Equations (1) and (2).

\[
\text{windindex} = \frac{1 - \cos(\theta - 135°)}{2}
\]

\[
\text{windindex} = \frac{1 - \cos(\theta - 0°)}{2}
\]

where \(\theta\) is the Euclidean direction from the nearest traffic source to the monitoring site in degrees east of north. The value of wind index ranges from 0 to 1.0 means the monitoring site is on the upwind of its nearest road and 1 means on the downwind.

3. Methodology

3.1. Wavelet analysis and Daubechies (DB) wavelet

Wavelet analysis is becoming a common tool for analyzing localized variations of power within a time series. Through this approach, one can track how the different scales related to the periodic components of the signal change overtime (Torrence and Compo 1997). There are various types of wavelet transforms and they are generally grouped as continuous or discrete (Torrence and Compo 1997). Based on the discrete wavelet transform (DWT), the signal is decomposed into a set of approximations and details, and the approximation is the low-frequency component of the signal and the detail is the high-frequency components of the signal (Lee and Tarng 2000). Observed PM2.5 concentration series in nature are usually discrete signals; DWT is often conducted as:

\[
W_j(t, k) = \sum_{-\infty}^{+\infty} f(t)\psi_j^*(t)dt
\]

\[
\psi_{j, k}(t) = a_0^{-j/2}\psi(a_0^{-j}t - b_0k)
\]

where \(a_0\) and \(b_0\) are constants, integer \(j\) is decomposition level, and \(k\) is time translation factor. In order to understand the change rule of signal, the signal will be decomposed and then be reconstructed using DWT at a certain level.

DB wavelet is one of the important wavelets because of the compact support and the perfect performance of time-frequency analysis; moreover, it can
describe the details of the signal conveniently and accurately so that it is successfully applied for signal processing (Ismail and Asfour 1999). Hence, DB wavelet is applied in this research to find the law of daily PM2.5 concentration.

3.2. LSMA and impervious surface area
LSMA is utilized to obtain impervious surface area in this paper, which is put into the LUR model as an alternative variable for land use data. The major assumption of the LSMA is that the spectral signature of a given pixel is proportion-weighted and linear combination of the end-member spectra (Roberts, Smith, and Adams 1993). An end-member is a “pure” land cover type or surface material that is assumed to have a unique spectral signature. Mathematically, the formula of the LSMA is:

$$R_i = \sum_{k=1}^{n} f_k R_{ik} + E_{Ri} \tag{4}$$

where $R_i$ is the reflectance value of band $i$; $f_k$ is the fraction of end-member $k$; $R_{ik}$ is end-member $k$'s reflectance in band $i$; $E_{Ri}$ is the un-modeled residual. Associated with determining $R_i$, $\sum_{k=1}^{n} f_k = 1$ and $f_k \geq 0$ are required.

Wu and Murray (2003) built a model to extract the impervious surface and this idea is accepted widely. The model is shown as follow:

$$R_{imp,b} = f_{low}R_{low,b} + f_{high}R_{high,b} + e_b \tag{5}$$

where $R_{imp,b}$, $R_{low,b}$ and $R_{high,b}$ are reflectivity of impervious surface, low albedo and high albedo of band $b$. Subsequently, $f_{low}$ and $f_{high}$ are density of low albedo and high albedo, respectively. $e_b$ is the value of what remains.

3.3. Modified LUR model using impervious surface area
We developed a modified LUR model using impervious surface area, in which the land use data is replaced by impervious surface area. The detail process can be found in Figure 4.

Variables of this modified model are chosen because of these following reasons: the length of roads implies the traffic condition; impervious surface area implies an air pollutant source; population density and meteorological conditions (wind index) can affect the spatial distribution of air pollutant concentration. The potential predictors are selected for model building and multiple linear regression models are created using stepwise regression in Statistical

![Figure 4. Simulation flow of spatial distribution for PM2.5 concentration using modified LUR model.](image-url)
Product and Service Solutions software tool (SPSS) by involving the most significant predictors.

The measured concentrations are divided into two parts in random selection: data from nine monitoring sites are used to establish the multiple linear regression equations while the other four are selected as validating sites. We use the hold-out-validation for the model evaluation; the evaluation approach applies the model developed for modeling sites to estimate concentrations at the validating sites. Finally, we use the leave-one-out cross-validation (LOOCV) to evaluate the LUR modes once more. LOOCV, which successively leaves out one site from the training dataset and estimates models based on the remaining N–1 sites, is also used for LUR model evaluation in order to test certainty of multiple linear regression models in this research. LOOCV is simply K-fold cross-validation with K set to n, the number of instance in the dataset. This means that the test set only consists of a single instance, which will be classified either correctly or incorrectly.

4. Results and discussions

4.1. Descriptive statistical analysis

Descriptive statistical analysis is fundamental and studies aiming at describing the patterns and general trends in a dataset are necessary. The indicators used in this paper are: central tendency (mean, median and mode), dispersion (maximum, minimum, range, coefficient of variation, standard deviation, inferior quartile, superior quartile) and distribution (kurtosis, skewness). All details of those indicators are shown in Table 1. A boxplot of monthly PM2.5 concentration is used to show the performance of monthly PM2.5 concentrations (Figure 5). In all observation days, PM2.5 concentration values of 71 days are less than 35 μg/m³ (National Primary Standard). One hundred and seven days meet National Secondary Standard (35–75 μg/m³) and 43 days are under slightly polluted conditions with concentrations greater than 75 μg/m³.

The minimum and maximum average concentrations are 36.12 and 58.81 μg/m³, which take place in summer and winter, respectively. During the winter, the median of PM2.5 concentration is 55.06 μg/m³. The greatest frequency of PM2.5 concentration in each season is 27.00, 37.25, 37.00 and 97.15 μg/m³. The value in winter is much larger than that in other seasons. The smallest standard deviation of the date is 16.37 μg/m³, occurring in summer, while the other three are approximate, 23.83 μg/m³ (spring), 27.05 μg/m³ (fall) and 26.24 μg/m³ (winter). The coefficient of variation shows low values in all seasons with an annual average of 0.49, which implies a low dispersion of measured PM2.5 concentration. The minimum PM2.5 concentration of all measured

| Index                        | Spring | Summer | Fall  | Winter | Annual |
|------------------------------|--------|--------|-------|--------|--------|
| Mean                         | 52.30  | 36.12  | 57.62 | 58.81  | 51.37  |
| Median                       | 45.92  | 33.86  | 51.35 | 55.06  | 45.92  |
| Mode                         | 27.00  | 37.25  | 37.00 | 97.15  | 37.25  |
| Standard deviation           | 23.83  | 16.37  | 27.05 | 26.24  | 25.12  |
| Kurtosis                     | 0.37   | −0.29  | −0.74 | −0.89  | −0.18  |
| Skewness                     | 0.97   | 0.56   | 0.41  | 0.36   | 0.73   |
| Coefficient of variation     | 0.46   | 0.45   | 0.47  | 0.45   | 0.49   |
| Minimum                      | 17.85  | 6.92   | 18.69 | 18.31  | 6.92   |
| Maximum                      | 116.23 | 73.38  | 128.31| 112.00 | 128.31 |
| Range                        | 98.38  | 66.46  | 109.62| 93.69  | 121.39 |
| Inferior quartile            | 32.85  | 23.36  | 33.30 | 40.32  | 32.33  |
| Superior quartile            | 63.31  | 33.88  | 80.76 | 82.60  | 66.97  |

Figure 5. Boxplot of monthly PM2.5 concentration data.
records is 6.92 μg/m³, occurring in summer, while the maximum one with a value of 128.31 μg/m³ appears in fall. During the summer, 25% of the measurements are under the value of 23.36 μg/m³ and 75% are under 33.88 μg/m³. In addition, it has been observed that 25% of measured records in winter are smaller than 40.32 μg/m³, whereas 75% are under 82.86 μg/m³. The coefficients of kurtosis are negative with low absolute values of −0.29, −0.74, −0.89 and −0.18. However, the value of spring (0.37) is positive, thus implying that measured records are more decentralized than normal distribution. At the same time, the coefficient of skewness is close to zero, namely, the data tends toward symmetrical distribution and demonstrates no significant seasonal differences in regard to the skewness of the data.

4.2. Temporal variation (daily average concentration of PM2.5)

For understanding the variation of PM2.5 concentration during the period of June 2012 to May 2013, DB wavelet of order 4 (DB4) is applied to decompose the time series of PM2.5 daily concentration into three levels. The original signal is decomposed into a hierarchy set of orthogonal approximation and detail functions. Finally, diurnal variation of PM2.5 average concentration is clearly displayed (Figure 6).

As shown in Figure 6(b): (1) July of 2012 (183–213 days) is the month with lowest daily concentrations of PM2.5. The highest concentrations occur in October of 2012 (275–305 days). (2) There is a longer period with high PM2.5 levels during December 2012 (336–365 days) and January 2013 (366–396 days). In detail, it significantly decreases in July 2012 and meets the trough at last third of the month; the lowest average concentration of PM2.5 is in July with 25.92 μg/m³. (3) Simultaneously, only the average concentration of PM2.5 is lower than the 24-h average concentration of national primary standard (35 μg/m³). (4) A sharp increase appears from the beginning of September to the middle of October and touches the peak. Consequently, the PM2.5 of October becomes the highest level (74.44 μg/m³). Compared to the primary standard value, it is more than double. However, there is a rapid decline during November (306–335 days). Then, at the end of December 2012 and beginning of January 2013, it increases again. (5) Another two rapid declines appear in February (397–424 days) and March 2013 (425–455 days), with a sharp increasing during the period. The phenomenon of continuous decline appears during the period between April (456–485 days) and May (486–516 days).

4.3. Seasonal variation of PM2.5

Seasonal variation for PM2.5 concentration of Guangzhou is shown as follows: winter (63.22 μg/m³) >
fall (59.59 μg/m³) > spring (52.63 μg/m³) > summer (37.65 μg/m³). It is nearly twofold higher in winter than in summer. In comparison to the low temperatures and more stable atmospheric stratification on spring mornings, the factors of high temperatures, unstable atmospheric and strong air convections in summer contribute to diffusion and migration of PM2.5. Furthermore, abundant rainfall in summer can purify the air by removing PM2.5 in the way of wet deposition.

To demonstrate the seasonal diversification, PM2.5 concentrations for each season were further studied with the finding that the change curve of summer is decidedly different from the others (Figure 7). It is apparent that most of the daily concentration is lower than 35 μg/m³ (in green) in summer. The highest and lowest concentrations of summer are 67.92 and 6.92 μg/m³, respectively. Secondly, in addition to two peaks, it is quite smooth with a similar serrated trend in the range of 35–75 μg/m³ (the yellow part) for spring. Lastly, the fluctuation in fall and winter is worth noting because many of the concentrations are so high (>75 μg/m³, the orange) that will make immediate impact on air quality.

4.4. Mutation characteristics of PM2.5 concentration

Identifying the mutation characteristics of PM2.5 concentration is significant to protecting environment. DB1 wavelet is considered because of its good consistency. Firstly, the time series of PM2.5 concentration are decomposed at the scale of three. Secondly, after the reconstructions of the first and the second high-frequency coefficients, signal charts are shown in Figure 8. There are seven mutated points, which are located on the 264, 289, 331, 356, 360, 433, 434 days, respectively (20 September 2012: 94.15 μg/m³, 15 October 2012: 128.31 μg/m³, 26 November 2012: 87.17 μg/m³, 21 December 2012: 109.08 μg/m³, 25 December 2012: 97.08 μg/m³, 8 March 2013: 116.23 μg/m³, 9 March 2013: 110.08 μg/m³). All of those days are polluted at a certain level.

4.5. Spatial distribution of PM2.5 levels

To evaluate the effect of variables on different seasons, the LUR model is applied separately, instead of an overall LUR model in this study. The most frequent wind in winter is the northerly and in summer, the southeasterly wind. In order to select potential predictor variables, the correlations between predictor variables and dependent variable are calculated using unit linear regression function in SPSS software (Table 2). The correlation with southeasterly wind is stronger than that with northerly, so the southeasterly wind may be regarded as the most frequent wind for spring and fall. In winter, the strongest correlation with PM2.5
concentration is the total length of road within the buffer of 2 km ($R^2 = 0.60$). In spring, the strongest correlation variable is the total length of road within the buffer of 1 km ($R^2 = 0.37$). Traffic emissions play an important role on PM2.5 concentration. Differing from winter and spring, PM2.5 concentrations of summer and fall have weak correlation with all variables. It may suggest that other factors are likely to influence PM2.5 levels for the two seasons. As a result, the LUR model is able to explain the spatial variation of PM2.5 concentration in each area of Guangzhou, both in winter and spring.

It is necessary to standardize those variables. The formula for standardization is introduced in Equation (6).

$$A = \frac{(X - X_{\text{min}})}{(X_{\text{max}} - X_{\text{min}})}$$  \hspace{1cm} (6)

Several multiple linear regression equations for winter and spring are established, respectively, using stepwise regression for selecting the best one (Equations (7)–(11) and Table 3). The test indicates that Equation (7) ($R^2 = 0.75$) is better and could be used to obtain insight into the distribution of PM2.5 concentration for winter; Equation (11) ($R^2 = 0.85$) could be used for spring. The error rate of 17.91% for Equation (10) is higher than that for Equation (11) with 9.31%. It suggests that wind direction plays an important role in affecting the PM2.5 concentration in spring. Therefore, Equation (11) will be the best to simulate the spatial distribution.

$$Y = 22.62x_5 - 27.32x_6 + 63.15x_7 - 19.16x_8 - 8.58x_9 + 48.27$$  \hspace{1cm} (7)
Table 3. Comparison of predicted and measured PM2.5.

| Equation   | PM2.5            | Sample 1 | Sample 2 | Sample 3 | Sample 4 |
|------------|------------------|----------|----------|----------|----------|
| Equation (7) | Predicted value  | 54.90    | 61.19    | 65.54    | 56.98    |
|            | Error rate (%)    | 9.77     | 5.36     | 6.01     | 4.56     |
| Equation (8) | Predicted value  | 50.56    | 54.86    | 62.96    | 51.13    |
|            | Error rate (%)    | 16.90    | 15.14    | 9.17     | 14.36    |
| Equation (9) | Predicted value  | 105.76   | 39.80    | 92.20    | 79.17    |
|            | Error rate (%)    | 73.83    | 38.50    | 32.20    | 32.60    |
| Equation (10) | Predicted value | 52.01    | 61.52    | 58.89    | 49.32    |
|            | Error rate (%)    | 16.80    | 20.61    | 2.05     | 32.16    |
| Equation (11) | Predicted value | 46.03    | 54.84    | 54.94    | 45.28    |
|            | Error rate (%)    | 11.49    | 10.86    | 6.70     | 8.20     |

* Error rate = |predicted PM2.5 − measured PM2.5| /measured PM2.5 × 100%.

\[ Y = -23.67x_3 + 28.4x_5 - 9.9x_6 + 25.24x_7 + 4.55x_8 - 4.53x_9 + 49.14 \]  
Equation (8)

\[ Y = -30.54x_3 - 40.17x_5 + 262x_6 - 368x_7 + 75.2x_8 + 110.9x_9 - 8720.7x_{10} + 7458 \]  
Equation (9)

\[ Y = 50.38x_5 - 111.24x_6 + 204.19x_7 - 115x_8 + 37.71 \]  
Equation (10)

\[ Y = 25.61x_5 - 33.22x_6 + 88.72x_7 - 60.05x_8 - 20.64x_{10} + 38.99 \]  
Equation (11)

In this paper, LOOCV is applied to test Equations (7) to (11). We divide the 13 sites into two datasets: winter datasets (nine monitoring sites) and spring datasets (four monitoring sites). For each dataset, N experiments will be performed, and each experiment use N-1 examples for training and the remaining example for testing. The true error is estimated as the average error rate on test examples. According to LOOCV, R^2 is 0.71 (winter) and 0.87 (spring), respectively. Meanwhile, the average RMSE is 0.038 (winter) and 0.042 (spring). The verification results of LOOCV on Equations (7) and (11) demonstrate that these two LUR models have good stability and predictability.

Due to the small number of monitoring sites, we test a grid of 5 km × 5 km instead of higher spatial resolution grid by the mean of ArcGIS. PM2.5 concentrations are calculated at each intersection (the intersect points of rows and columns in 5 km × 5 km grid) using multiple linear regression equations. Then the seasonal spatial concentration distributions of PM2.5 are interpolated by the ordinary Kriging approach. The simulating distributions are described in Figure 9.

In Figure 9(a), spatial characteristics of two distinct high-value centers get much attention while the surrounding area gradually reduced obviously. Those high-value centers are mainly in the northern and central areas of Guangzhou. A high-value center in the northern part covers the southeast of Huadu, northwest of Baiyun. The central high-value center includes the whole of Yuexiu, north of Liwan, and most of Tianhe and Haizhu. The suburbs of Zengcheng, Conghua and Nansha have better air quality with the low PM2.5 concentration. On the other hand, there is a similar phenomenon happening in spring. One high-value center exists in all of Yuexiu, most of Tianhe, the north of Liwan and south of Baiyun. Conghua, Zengcheng and Nansha are also the region with low PM2.5 concentration. To examine the accuracy of simulation, the results are compared to the measured data. According to the data, monitoring sites of LWXC, LH, GYQ, HD, PY and HZBG are those with high PM2.5 concentration in both of winter and spring. On the contrary, Tianhe is with low PM2.5 concentration. As a whole, the simulation results roughly correspond to the actual values. Without measured data in Zengcheng and Nansha, we cannot determine whether the values of Zengcheng and Nansha are correct or not.

6. Conclusions

In this paper, temporal and spatial variations of PM2.5 average levels are analyzed in Guangzhou. Wavelet analysis is used to judge the day-to-day changes in PM2.5 concentrations, and this approach is confirmed as a reliable method for investigating temporal variations in air pollutant research. Temporal variations will be well expressed and the mutated signal of PM2.5 level can be identified by wavelet analysis. To map the spatial differentiation of PM2.5 levels, the LUR model achieves satisfactory results. To confirm seasonal differences, seasonal multiple linear regression models are naturally considered. We can determine the characteristics of space in seasons. Deciding upon the dependent variable and the independent variables correlation, we find that a weak relationship between impervious surface area and PM2.5 levels exist, whereas there is a strong correlation with traffic, particularly in winter. The measured concentration of PM2.5 and variables are significant when associated just in winter/spring and not in summer/fall. Perhaps additional factors contribute to the air pollutants in the two seasons. Hence, for Guangzhou, the LUR model is only available in winter and spring. Wind direction has immediate impact on PM2.5 levels in spring. Even other independent variables do not affect the spatial...
distribution of PM2.5 in the study domain; it does not suggest those variables has no relationship to PM2.5 but rather the small number of monitoring sites used to build models and test. Further research is necessary because of the existence of deficiencies in this study, such as the sparse monitoring sites. If there are more monitoring sites, the accuracy of the simulation could potentially be higher.

**Funding**

This work is supported by the National Nature Science Foundation of China [grant number: 41201432], the National Science Foundation of Tibet [grant number: 2016ZR-TU-05] and the Foundation for Innovative Research for Young Teachers in Higher Educational Institutions of Tibet [grant number: QCZ2016-07].

**Notes on contributors**

**Fenglei Fan** is a professor at South China Normal University, the Dean of Joint laboratory of plateau surface remote sensing and the Everest Scholar of Tibet University. He received his MS and PhD degrees from the Chinese Academy of Sciences in 2004 and 2006, respectively. He is the author of more than 50 journal papers. His current research interests include urban remote sensing and hyperspectral remote sensing.

**Runping Liu** received her undergraduate degree and master degree from South China Normal University in 2011 and 2015, respectively.

**ORCID**

Fenglei Fan [http://orcid.org/0000-0001-6190-9072]

---

**References**

Burton, R. M., H. H. Suh, and P. Koutrakis. 1996. “Spatial Variation in Particulate Concentrations within Metropolitan Philadelphia.” *Environment Science of Technology* 30 (2): 400–407. doi:10.1021/es950030f.

Chen, Y., R. Shi, S. Shu, and W. Gao. 2013. “Ensemble and Enhanced PM$_{10}$ Concentration Forecast Model Based on Stepwise Regression and Wavelet Analysis.” *Atmospheric Environment* 74: 346–359. doi:10.1016/j.atmosenv.2013.04.002.

Eeftens, M., M. Tsai, C. Ampe, B. Anwander, R. Beelen, T. Bellander, G. Cesaroni, et al. 2012. “Spatial Variation of PM$_{2.5}$, PM$_{10}$, PM$_{2.5}$ Absorbance and PM Coarse Concentrations between and within 20 European Study Areas and the Relationship with NO$_2$-Results of the ESCAPE Project.” *Atmospheric Environment* 62: 303–317. doi:10.1016/j.atmosenv.2012.08.038.

Elbayoumi, M., N. Ramli, N. Yusof, and W. Madhoun. 2013. “Spatial and Seasonal Variation of Particulate Matter (PM$_{10}$ and PM$_{2.5}$) In Middle Eastern Classrooms.” *Atmospheric Environment* 80: 389–397. doi:10.1016/j.atmosenv.2013.07.067.

Garland, R., H. Yang, O. Schmid, D. Rose, A. Nowak, P. Achter, A. Wiedensohler, et al. 2008. “Aerosol Optical Properties in a Rural Environment near the Mega-City Guangzhou, China: Implications for Regional Air Pollution Radiative Forcing and Remote Sensing.” *Atmospheric Chemistry and Physics* 8 (17): 5161–5186. doi:10.5194/acp-8-5161-2008.

Hochadel, M., J. Heinrich, U. Gehring, V. Morgenstern, T. Kuhlbusch, E. Link, H. Wichmann, and U. Krämer. 2006. “Predicting Long-Term Average Concentrations of Traffic-Related Air Pollutants Using GIS-based Information.” *Atmospheric Environment* 40 (3): 542–553. doi:10.1016/j.atmosenv.2005.09.067.

Huang, K., G. Zhuang, Y. Lin, Q. Wang, J. Fu, Q. Fu, T. Liu, and C. Deng. 2013. “How to Improve the Air Quality over Mega-Cities in China? Pollution Characterization and Source
Analysis in Shanghai Before, during and after the 2010 World Expo." Atmospheric Chemistry and Physics (Discuss) 13 (2): 3379–3418. doi:10.5194/acp-13-3379-2013.

Ismail, A., and S. Asfou. 1999. "Discrete Wavelet Transform: A Tool in Smoothing Kinematic Data." Journal of Biomechanics 32 (3): 317–321. doi:10.1016/S0021-9290(98)00171-7.

Jerrett, M., A. Arain, P. Kanaroglou, B. Beckerman, D. Potoglou, T. Sahsuroglu, J. Morrison, and C. Giovis. 2005. "A Review and Evaluation of Interverband Air Pollution Exposure Models." Journal of Exposure Analysis and Environmental Epidemiology 15: 185–204. doi:10.1038/sj.jea.7500388.

Kumar, P., and G. Foufoula. 1997. "Wavelet Analysis for Geophysical Applications." Reviews of Geophysics 35 (4): 385–412. doi:10.1029/97RG00427.

Lee, B., and Y. Tarng. 2000. "Drill Fracture Detection by the Discrete Wavelet Transforms." Journal of Materials Processing Technology 99 (1–3): 250–254. doi:10.1016/S0924-0136(99)00432-X.

Lee, H., J. Gent, B. Leaderer, and P. Koutrakis. 2011. "Spatial and Temporal Variability of Fine Particle Composition and Source Types in Five Cities of Connecticut and Massachusetts." Science of the Total Environment 409 (11): 2133–2142. doi:10.1016/j.scitotenv.2011.02.025.

Leiva, G. M. A., D. A. Santibañez, E. S. Ibarra, C. P. Matus C., and R. Seguel. 2013. "A Five-Year Study of Particulate Matter (PM10) and Cerebrovascular Diseases." Environmental Pollution 181: 1–6. doi:10.1016/j.envpol.2013.05.057.

Li, T., H. Shen, C. Zeng, Q. Yuan, and L. Zhang. 2017a. "Point-Surface Fusion of Station Measurements and Satellite Observations for Mapping PM2.5 Distribution in China: Methods and Assessment." Atmospheric Environment 152: 477–489. doi:10.1016/j.atmosenv.2017.01.004.

Li, T., H. Shen, C. Zeng, Q. Yuan, X. Zhang, and L. Zhang. 2017b. "Estimating Ground-Level PM2.5 by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach." Geophysical Research Letters 44 (23): 11985–11993. doi:10.1002/2017GL075710.

Li, W., Z. Bai, J. Wei, A. Liu, Y. Zhao, and T. Jin. 2008. "Pollution Characteristics of PM2.5 and its Main Components in Tianjin Winter Atmosphere." China Environmental Science 28 (6): 481–486.

Marcazzan, G., S. Vaccaro, G. Valli, and R. Vecchi. 2001. "Characterization of PM10 and PM2.5 Particulate Matter in the Ambient Air of Milan (Italy)." Atmospheric Environment 35 (27): 4639–4650. doi:10.1016/S1352-2310(01)00124-8.

Pinto, J., A. Lefohn, and D. Shadwick. 2004. "Spatial Variability of PM2.5 in Urban Areas in the United States." Journal of the Air & Waste Management Association 54 (4): 440–449. doi:10.1080/10473289.2004.10470919.

Roberts, D., M. Smith, and J. Adams. 1993. "Green Vegetation No-Photosynthetic Vegetation and Soils in AVIRIS Data." Remote Sensing of Environment 44 (2–3): 255–269. doi:10.1016/0034-4257(93)90020-X.

Samet, J., F. Dominici, F. Curriero, I. Coursac, and S. Zeger. 2000. "Particulate Air Pollution and Mortality in 20 US Cities: 1987–1994." The New England Journal of Medicine 343 (24): 1742–1749. doi:10.1056/NEJM200012143432401.

Sang, Y. 2013. "A Review on the Applications of Wavelet Transform in Hydrology Time Series Analysis." Atmospheric Research 122: 8–15. doi:10.1016/j.atmosres.2012.11.003.

Shen, H., L. Huang, L. Zhang, P. Wu, and C. Zeng. 2016. "Long-Term and Fine-Scale Satellite Monitoring of the Urban Heat Island Effect by the Fusing of Multi-Temporal and Multi-Sensor Remote Sensed Data: A 26-Year Case Study of the City of Wuhan in China." Remote Sensing of Environment 172 (2016): 109–125. doi:10.1016/j.rse.2015.11.005.

Shi, W., M. Wong, J. Wang, and Y. Zhao. 2012. "Analysis of Airborne Particulate Matter (PM2.5) Over Hong Kong Using Remote Sensing and GIS." Sensors 12 (6): 6825–6836. doi:10.3390/s120606825.

Strunin, M., and T. Hiyama. 2005. "Spectral Structure of Small-Scale Turbulent and Mesoscale Fluxes in the Atmospheric Boundary Layer over a Thermally Inhomogeneous Land Surface." Boundary-Layer Meteorology 117 (3): 479–510. doi:10.1007/s10546-005-2188-2.

Tao, J., K. Ho, L. Chen, L. Zhu, J. Han, and Z. Xu. 2009. "Effect of Chemical Composition of PM2.5 on Visibility in Guangzhou, China, 2007 Spring." Particuleology 7 (1): 68–75. doi:10.1016/j.partic.2008.11.002.

Torrence, C., and G. Compo. 1997. "A Practical Guide to Wavelet Analysis." Bulletin of the American Meteorological Society 79 (1): 61–78. doi:10.1175/1520-0477(1998)079<0061:APGWTW>2.0.CO;2.

Vadrevu, K., and Y. Choi. 2011. "Wavelet Analysis of Airborne CO2 Measurements and Related Meteorological Parameters over Heterogeneous Landscapes." Atmospheric Research 102 (1–2): 77–90. doi:10.1016/j.atmosres.2011.06.008.

Wang, M., R. Beelen, M. Eeftens, K. Meliefste, G. Hoek, and H. Brunekreef. 2012. "Systematic Evaluation of Land Use Regression Models for NO2." Environmental Science and Technology 46 (8): 4481–4489. doi:10.1021/es204183v.

Wang, R., S. Henderson, H. Shih, R. Allen, and M. Brauer. 2013. "Temporal Stability of Land Use Regression Models for Traffic-Related Air Pollution." Atmospheric Environment 64: 312–319. doi:10.1016/j.atmosenv.2012.09.056.

Wang, X., X. Bi, G. Sheng, and J. Fu. 2006. "Chemical Composition and Sources of PM10 and PM2.5 Aerosols in Guangzhou, China." Environmental Monitoring and Assessment 119 (1–3): 425–439. doi:10.1007/s10661-005-9034-3.

Weng, Q., and D. Lu. 2008. "A Sub-Pixel Analysis of Urbanization Effect on Land Surface Temperature and Its Interplay with Impervious Surface and Vegetation Coverage in Indianapolis, United States." International Journal of Applied Earth Observation and Geoinformation 10: 68–83. doi:10.1016/j.ijag.2007.05.002.

Wu, C., and A. Murray. 2003. "Estimating Impervious Surface Distribution by Spectral Mixture Analysis." Remote Sensing of Environment 84: 493–505. doi:10.1016/S0034-4257(02)00136-0.

Wu, P., H. Shen, L. Zhang, and F. Göttische. 2015. "Integrated Fusion of Multi-Scale Polar-Orbiting and Geostationary Satellite Observations for the Mapping of High Spatial and Temporal Resolution Land Surface Temperature." Remote Sensing of Environment 156: 169–181. doi:10.1016/j.rse.2014.09.013.

Yu, T. 2010. "Characterization of Ambient PM2.5 Concentrations." Atmospheric Environment 44: 2902–2912. doi:10.1016/j.atmosenv.2010.04.034.

Zhaow, X., X. Zhang, X. Xu, J. Xu, W. Meng, and W. Pu. 2009. "Seasonal and Diurnal Variations of Ambient PM2.5 Concentration in Urban and Rural Environments in Beijing." Atmospheric Environment 43: 2893–2900. doi:10.1016/j.atmosenv.2009.03.009.