

QUANTIFYING THE RELATIONSHIP BETWEEN NATURAL AND SOCIOECONOMIC FACTORS AND WITH FINE PARTICULATE MATTER (PM_{2.5}) POLLUTION BY INTEGRATING REMOTE SENSING AND GEOSPATIAL BIG DATA

Chen Yang^{1,2}, Qingming Zhan^{1,2,*}, Jing Zhang¹, Huimin Liu^{1,2}, Zhiyu Fan^{1,2}

¹ School of Urban Design, Wuhan University, 8 Donghu South Road, Wuhan 430072, China – yangchen_cug@foxmail.com

² Collaborative Innovation Centre of Geospatial Technology, 129 Luoyu Road, Wuhan 430079, China - yangchen_cug@foxmail.com

Commission III, WG III/8

KEY WORDS: Fine particulate matter (PM_{2.5}), Spatial-temporal CoKriging, Random Forest, Volunteer Geographic Information (VGI)

ABSTRACT:

PM_{2.5} pollution is an environmental issue results from various natural and socioeconomic factors, frequently witnessed in the spring and winter across mainland China. However, the dominant influence of natural and socioeconomic factors within a city on PM_{2.5} is not extensively studied yet. In this study, the Random Forest Regression (RFR) is utilized to quantify the relationships between PM_{2.5} and potential factors within Wuhan city on a typical day turn from winter to spring. Technically, the 24-hour average PM_{2.5} concentration in downtown area on February 17th 2017 are collected at 9 sites. In the meantime, we retrieve simultaneous aerosol depth optical depth (AOD) from the Moderate Resolution Imaging Spectroradiometer (MODIS). The ground measured PM_{2.5} and AOD are coupled for the retrieval of near-surface PM_{2.5} concentration by Spatial-temporal CoKriging (STCK) with Normalized Vegetation Index (NDVI), Modified Normalized Water Index (MNDWI), Normalized Building Index (NDBI) from Landsat-8 and DEM from Shuttle Radar Topography Mission (SRTM). As the geospatial big data booms, the Internet-collected volunteered geographic information (VGI), representing the urban form and function, are integrating for the regression to obtain the spatial variables importance measures (VIMs) by RFR both in centre and sub-urban region of Wuhan. The results reveal that terrain characteristics and the density of industrial enterprises have obvious relationships with the accumulation of PM_{2.5} while the density of roads also contributes to this.

1. INTRODUCTION

The harm of fine particulate matter (PM_{2.5}) to public health has drawn across China in recent years (Ma et al., 2016; Wang et al., 2015). As an environmental issue, PM_{2.5} is widely recognized as related to natural and socioeconomic factors (Chen et al., 2016; Tan et al., 2017). China possesses vast territory and changeable natural elements across its mainland. There are various interactions between natural and socioeconomic factors across the country (Ma et al., 2016). Industries, fossil fuel consumption and vehicles are explored to be the main cause of PM_{2.5} accumulation of PM_{2.5} worldwide (Perrone et al., 2012; Zalakeviciute et al., 2018). Socioeconomic factors are directly related to the emergence and spatial-temporal variations of PM_{2.5} and natural elements have influence on the diffusion and accumulation of PM_{2.5} (Fanizza et al., 2018; Kioumourtzoglou et al., 2016).

Cities are not only closely related to human life, but also the interaction between man and nature in the city is more complicated. However, the spatial-temporal variations and driving factors of PM_{2.5} in urban scale have not been thoroughly explored (Yang et al., 2018a). Many scientific research institutions have made enormous efforts to monitor the environmental issues within the city. China Atmosphere Watch Network (CAWN) and Aerosol Robotic Network (AERONET) are established under this background.

Most of the existing PM_{2.5} monitoring data are scattered points provided by monitoring stations such as CAWN and AERONET. In order to obtain seam-less and continuous PM_{2.5} concentration at surface-level, many previous studies utilize the relationship between satellite observed atmospheric optical depth (AOD) and

ground-based PM_{2.5} observations (Guo et al., 2014; Lin et al., 2015). These methods retrieve surface-level PM_{2.5} concentration using AOD-PM_{2.5} relationship can be classified into two categories: empirically and semi-empirically observation based methods (Lin et al., 2015). The empirically observation based methods rely on the statistical regressions of AOD-PM_{2.5} relationships. These methods usually adopt polynomial linear regression (Li et al., 2011), artificial neural network (ANN) (Wu et al., 2012), and non-linear regression (Benas et al., 2013). The semi-empirically observation based methods integrate the AOD-PM_{2.5} relationship with atmospheric parameters (Boyounk et al., 2010; Yang et al., 2018b).

Since the accumulation of PM_{2.5} is extensively studied to be affected by natural and socioeconomic factors. The mapping and analysis of surface-level PM_{2.5} and its factors depend on the significance and representativeness of the related factors in the retrieval and regression. The natural factors within the city, such as greenness, bareness and wetness, can be retrieved from the remotely sensed images by multiple space-borne sensors with high spatial/temporal resolution and great visual scene. However, the economic factors within the city are too various and volatile to be extracted accurately and duly (Yao et al., 2017). Conventional statistical and monitoring methods are not qualified enough to meet the rigorous requirements of data volume and temporal resolution for implementing spatial-temporal pattern analysis of socioeconomic activities within cities. As the geospatial big data booms, the Internet-collected volunteered geographic information (VGI) such as Point of Interests (POI) and real time population thermodynamic map (PTM) can now be used to settle the dilemma (Elwood, 2008; Zhang et al., 2018). Previous studies show that POI and PTM are related to

socioeconomic activities at multi-scale and effectively represent the urban form and urban function (Yao et al., 2017; Zhang et al., 2018).

In this study, the fine scale surface PM2.5 concentration is retrieved using AOD from the Moderate Resolution Imaging Spectroradiometer (MODIS) coupling with Normalized Vegetation Index (NDVI), Modified Normalized Water Index (MNDWI), Normalized Building Index (NDBI) from Landsat-8 and DEM from Shuttle Radar Topography Mission (SRTM). The POIs and PTM collected over the same period and some other factors are utilized to explore the main driving factors in central and sub-urban areas within Wuhan city by RFR.

2. MATERIALS AND METHODS

2.1 Study area and data

Wuhan, China, is selected as the case area. Wuhan is a mega city with most populous and largest economic volume of central China. The extent of the study area covers $75 \times 75 \text{ km}^2$ and possess heterogeneous land cover composition. The urban area surrounded by the 2nd Ring Road is the centre of Wuhan, which has a high building density and land use efficiency. Land use types in this area are mainly residential land use, commercial land use and public service land use. The exterior area of the second ring road and neighbourhood area the third ring road is mainly the sub-urban and expanded urban areas. The land use types are more diversified.

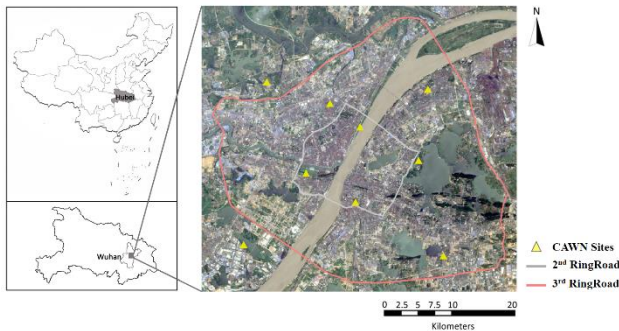


Figure 1. Study area

The AOD product (MOD04_3K) with 3000 meters of MODIS (<https://ladsweb.modaps.eosdis.nasa.gov/search/>) and PM2.5 concentration at 9 sites within downtown area of Wuhan are coupled in the retrieval of surface-level PM2.5. The detailed information of these sites are reported in Table 1.

Station Name	Station Number	X Coordinate	Y Coordinate
Donghu Liyuan	1325A	114.3672	30.5719
Hanyang Yuehu	1326A	114.2511	30.5514
Hankou Huaqiao	1327A	114.2836	30.6197
Wuchang Ziyang	1328A	114.3006	30.5494
Qingshan Ganghua	1329A	114.4272	30.6103
Zhuankou Xinqu	1330A	114.1525	30.4753
Hankou Jiangtan	1331A	114.3008	30.5947

Table 1. Detailed information of nine stations within Wuhan

The natural constitution is represented by NDVI, NDBI, MNDWI and Albedo. These indicators are calculated using the LIT product of Landsat-8 Operational Land Imager (OLI). The POIs and PTM are collected from the Baidu Map (<http://lbsyun.baidu.com>) and spatialized on ArcGIS 10.2.

2.2 Spatial-temporal CoKriging (STCK)

Technically, to obtain the spatially-temporally continuous and typical latent pattern of the surface-level PM2.5 concentration in Wuhan within one day, the Spatial-temporal CoKriging (STCK) is adopted for its proper consideration of spatial-temporal relationship between variances and dependence factors (Liu et al., 2015). STCK relies on the basic assumption of spatial-temporal stationary. This assumption indicates that the sampling data of many natural phenomena have a certain correlation in space or time domain. The prediction process of STCK can be represented as:

$$\widehat{Z}_F(x_0, t_k) = \sum_{j=1}^n a_{i1} Z_V(x_i, t_i) + \dots + \sum_{j=1}^n a_{ik} Z_V(x_i, t_k) + \frac{\sum_{i=1}^m \beta_i Z_F(x_j, t_0)}{\sum_{i=1}^m \beta_i} \quad (1)$$

where $\widehat{Z}_F(x_0, t_k)$ is the predicted pixel of fine resolution F at time t_k for each sliding-window, $Z_V(x_i, t_k)$ is the pixel at coarse resolution V at time t_k , $Z_F(x_j, t_0)$ is the pixel at fine resolution V at time t_0 . For the 3×3 sliding window, $m=9$. The optimization of weights in time domain can be realized by solving the following equation:

$$w = C^{-1}B \quad (2)$$

where C^{-1} is the inverse matrix of covariance matrix C , $B = [C_F^0(Z_{0M}), C_V^1(Z_{0N}), \dots, C_V^T(Z_{0N}), 0_1, 0_2, \dots, 0_{t-1}, 1]^T$, and C can be represented as:

$$C = \begin{bmatrix} C_{FF}^{00} & C_{VF}^{01} & \dots & C_{VF}^{0T} & 1 & 0 \\ C_{FV}^{10} & C_{VV}^{11} & & C_{VV}^{1T} & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ C_{FV}^{T0} & C_{VV}^{T1} & & C_{VV}^{TT} & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & & 1 & 0 & 0 \end{bmatrix} \quad (3)$$

where C_{FF}^{00} is the sub covariance matrix with size of $m \times m$ for fine resolution image ($Z_F(x_i, t_0)$, $Z_F(x_j, t_0)$). STCK is scalable. When the number of independent variables increase, STCK is still applicable.

2.3 Random Forests Regression (RFR)

RFR is one of the most suitable machine-learning method for exploring the spatial-temporal variation and driving factors of atmospheric issues, such as PM2.5 concentration (Hu et al., 2017). The foregoing study also indicates that RFR is able to solve multicollinearity and dimensionality of the input data and can properly settle the overfitting problem. RFR is a non-linear statistical regression method that is composed of and averages multiple randomised and de-correlated decision trees (Hutengs and Vohland, 2016). RFR introduces randomness into the construction of each de-correlated decision tree by bootstrap sampling. For each tree, the following process is repeated:

1. m of all p variables of the original dataset are selected randomly,
2. find the variable which is able to split dataset optimal and compute the corresponding split point,
3. split input dataset at this split point.

Subsequently, RFR computes such trees and averages them to obtain predictions. Out-of-bag (OOB) (Hutengs and Vohland, 2016) error estimation and spatial variables importance measure (VIM) (Zhang et al., 2018) are two key parameters of RFR. In each split, the number m of variables selected by minimising the

OOB error of predictions. The growth of each decision tree stops when the OOB error stabilized.

3. IMPLEMENTATION AND RESULTS

3.1 Surface-level PM_{2.5} concentration retrieval

The hourly ground-observed PM_{2.5} concentration data is discrete spatially and temporally, so they can not reflect the typical pattern of PM_{2.5} concentration and spatial disparity within one day throughout the region. In order to obtain the continuous and seamless pattern of surface-level PM_{2.5} concentration, STCK are adopted using AOD data from MODIS and spectral indices from Landsat-8 OLI.

MOD04_3K is the level 2 aerosol product with 3000 meters resolution of MODIS representing the ambient aerosol optical properties, aerosol mass concentration and data quality assurance. MOD04_3K product in Collection 6 retrieving AOD using Dark Target algorithm (Sayer et al., 2014). In this study, we collected the MOD04_3K data on February 17th 2017 (Figure 2).

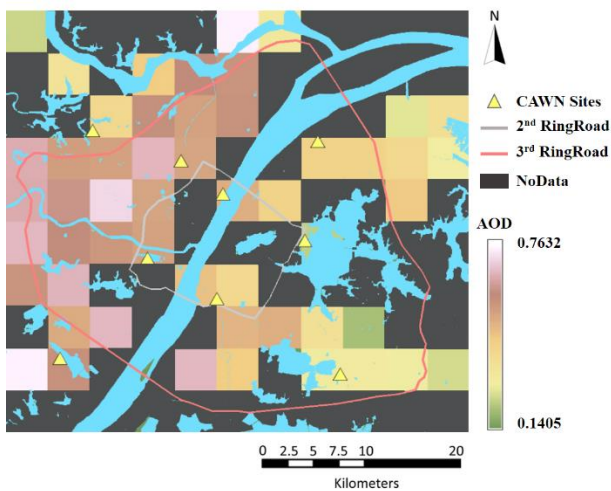


Figure 2. MOD04_3K data on February 17th 2017 of Wuhan

In this study, NDVI, MNDWI, NDBI and Albedo retrieved from Landsat-8 with 30 meters resolution are resampled into 120 meters resolution, which is optimized by semi-variogram. Then surface-level PM_{2.5} concentration with 120 meters resolution is obtained by STCK (Figure 3). The accuracy assessment of STCK reveals that the Root Mean Square Error (RMSE) of STCK is 2.482, reduction of RMSE (RRMSE) compared with CoKriging (CK) is 19.72%. In other words, the accuracy of spatial-temporal interpolation conducted by STCK is 19.72% higher than conducted by CK.

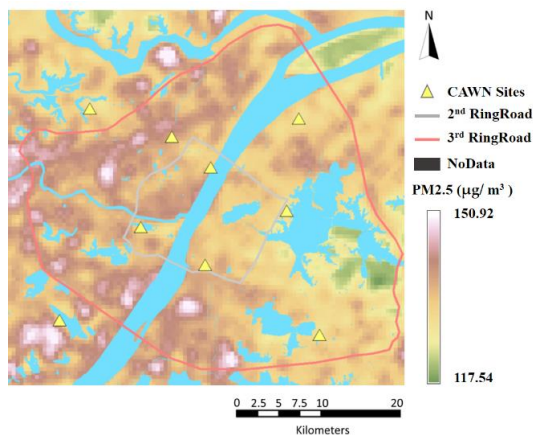


Figure 3. Surface-level PM_{2.5} concentration with 120 meters resolution of Wuhan

Figure 3 reveals that PM_{2.5} concentration in the centre of Wuhan, enclosed by the 2nd ring road, is relatively heterogeneous. The accumulation of PM_{2.5} in Hankou is slightly obvious than that in Wuchang and Hanyang. The areas with high concentration of PM_{2.5} are mainly concentrated in the areas scattered located along the 3rd ring road in Hankou and Hanyang. Especially, some main industrial districts in Wuhan, covering:

1. Dongfeng Peugeot Citroren Automobile Company (DPCA) industrial park located in the South-West corner of Figure 3,
2. Jinyintan industrial park at the North boundary of Figure 3,
3. Baishazhou area where exists many open-air construction sites and logistics parks,
4. Wuchangnan railway station performs as the one of the most important freight marshalling station in the Jiangguang railway.

It can be seen that the distribution of industrial parks in Wuhan has a significant impact on regional PM_{2.5} accumulation. In addition, in the process of urban development and renewal, open-air construction sites also result in locally higher PM_{2.5} concentrations. However, it is noteworthy that Qingshan District, which owns Wuhan Iron and Steel Works, the largest iron and steel plant in central China, does not have a high concentration of PM_{2.5}. To the best of the knowledge of the authors, Qingshan industrial district was selected as one of the second batch of Pilot Demonstrations of Circular Economy in China in 2007. Since then, various measures have been promoted on technological upgrading and energy recycling aiming at weakening the negative effects of iron and steel production on environment (<http://www.qingshan.gov.cn>).

3.2 VIMs of driving factors for PM_{2.5} concentration in Wuhan

RFR is able to measure the contribution weights of the driving factors involved in the regression, i.e. the VIMs (Zhang et al., 2018). VIM of each input variable (Breiman, 2001) is calculated on the basis of OOB as:

1. after the construction of the each decision tree, OOB estimation error is calculated,
2. repeat OOB estimation error calculation on every variables involved,
3. calculate additional OOB estimation error by increasing the value of each variables in each OOB.
4. the quantified difference between OOB estimation error and additional OOB estimation error is the VIM of each variable.

In our study, the natural factors involved covering bareness (represented by NDBI) and geological characteristics (represented by DEM). The socioeconomic factors are industrial land use (IL) distribution, commercial land use (CL) distribution, administration and public services land use (APL) distribution, road density (RD) and population distribution (PD). In addition, the density of restaurants (RED) are taken into consideration to evaluate the effects of cooking fume on PM_{2.5} accumulation. Table 2 illustrates the natural and socioeconomic factors involved in this study.

Variables	Original Resolution	Source	Definition
NDBI	30m	Landsat-8 OLI	Representation of bareness within Wuhan (resampled into 120m resolution)
DEM	30m	SRTM	Characteristics of geological conditions
IL	10m	Baidu Map	Density of industrial POI points for each $120 \times 120 \text{ m}^2$ pixel
CL	10m	Baidu Map	Density of commercial POI points for each $120 \times 120 \text{ m}^2$ pixel
APL	10m	Baidu Map	Density of administration and public services POI points for each $120 \times 120 \text{ m}^2$ pixel
RD	10m	Baidu Map	Density of roads for each $120 \times 120 \text{ m}^2$ pixel
RED	10m	Baidu Map	Density of restaurant POI points for each $120 \times 120 \text{ m}^2$ pixel
PD	10m	Baidu Map	Density of population in every 10 meters grid, and resampled into 120m resolution

Table 2. Detailed information and definition of each variable

All the POIs, roads shapefile and population density are collected using geospatial Big Data technique from Baidu Map (<http://lbsyun.baidu.com>). To be accordance with the resolution of PM2.5 concentration, natural factors are resampled into 120m resolution, and socioeconomic factors are calculated in $120 \times 120 \text{ m}^2$ pixel. The 24-hour average population density of Wuhan on February 17th 2017 is shown in Figure 3.

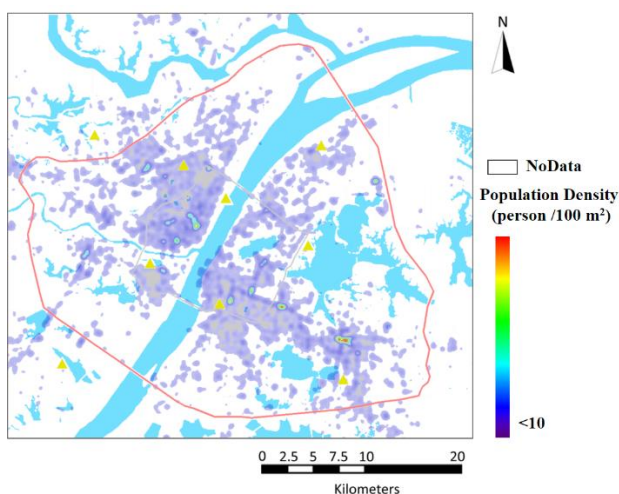


Figure 4. 24-hour average population density of Wuhan collected from Baidu Map on February 17th 2017

In this study, RFR is implemented using the open-source library Sci-kit learn package (<http://scikit-learn.org>) in Python 2.7. The R-Square (R^2) value of RFR using these ten variables is 0.825

and RMSE value is 1.882. In Table 3, the VIMs and cross validated mean absolute predictive error (MAPE) of each spatial variables throughout the whole study area are reported.

Variables	VIMs	MAPE
NDBI	0.13	0.39
DEM	0.22	0.34
IL	0.27	0.32
CL	0.11	0.40
APL	0.08	0.41
RD	0.20	0.35
RED	0.14	0.39
PD	0.08	0.41

Table 3. VIMs of variables throughout the study area

According to Table 3, DEM has the highest VIM value among the natural factors. Topographic factors, such as elevation and slope direction, have a greater impact on the accumulation and dispersion of PM2.5. As we all know, cities located in the basin are more vulnerable to PM2.5 pollution (Hu et al., 2017; Yang et al., 2018a). IL has the relatively highest VIM value compared with the other factors. Industrial land use usually produce more particulate pollutants, nitrogen oxides and sulphides, which are widely recognized as the main chemical source of the PM2.5, than any other land use categories. PM2.5 accumulation also has close relationship with RD for the sulfur dioxide and nitrogen dioxide from auto exhausts. It is remarkable that RED also contributes a lot to PM2.5 concentration in Wuhan. Because the restaurant itself is the source of particulate matter and lampblack. At the same time, social and economic activities around the restaurant are more frequent, resulting in greater population and traffic flow, which also contributes to PM2.5 accumulation. However, the RED is highly co-related with RD and CD. Thus, this phenomenon needs to be further verified by more accurate field observations and chemical analysis.

To explore the spatial disparity of factors in multiple regions under different natural and socioeconomic backgrounds within Wuhan. Simultaneously, RFR is also implemented on these factors in centre (enclosed by 2nd Ring Rod) and sub-urban area (exclude by 2nd Ring Rod) of Wuhan. The VIMs of these factors in centre and sub-urban area of Wuhan are reported in Table 4 and Table 5, respectively.

Variables	VIMs	MAPE
NDBI	0.12	0.39
DEM	0.18	0.36
IL	0.16	0.37
CL	0.15	0.38
APL	0.10	0.41
RD	0.25	0.33
RED	0.20	0.35
PD	0.11	0.42

Table 4. VIMs of variables in the centre of Wuhan

Table 4 reveals that the dominant factor of PM2.5 accumulation in urban central area is RD. An increase of VIMs on socioeconomic factors, except for IL, is witnessed. Specifically, the contributions of CL, RD and RED are increased more obvious. Because the major land use categories in the centre of Wuhan are residential land use, traffic facilities and commercial areas. Hence, the off-gas, kitchen products and some other wastes become the major factors on PM2.5 accumulation under urban background.

Variables	VIMs	MAPE
-----------	------	------

NDBI	0.15	0.38
DEM	0.24	0.33
IL	0.30	0.29
CL	0.09	0.44
APL	0.07	0.47
RD	0.17	0.37
RED	0.11	0.40
PD	0.06	0.45

Table 5. VIMs of variables in sub-urban area of Wuhan

As reported in Table 5, IL, DEM and RD are the top three factors lead to PM_{2.5} pollution in sub-urban area of Wuhan. As mentioned in Section 3.1, the majority of industrial land use located in the sub-urban area along with 3rd Ring Road. And in the areas within 3rd Ring Road of Wuchang and Hankou district (specifically are Baishazhou region and Wangjiadun region), there many open-air logistics yards and building sites which may produce serious particulate emissions. The ground truths of Baishazhou region and Wangjiadun region in Google Earth Pro are shown in Figure 5.

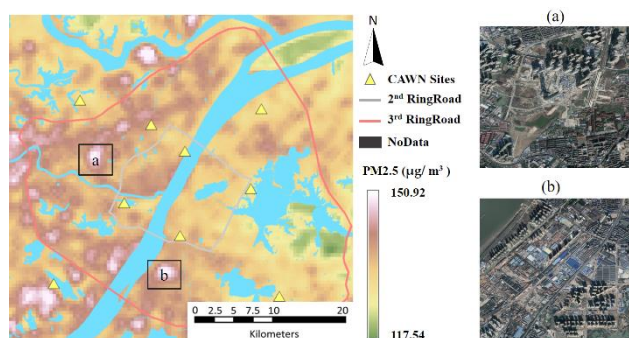


Figure 5. The ground truth of Wangjiadun region (a) and Baishazhou region (b) in Google Earth Pro

4. RESULTS

This study aims at quantifying the effects of natural and socioeconomic factors on the concentration of PM_{2.5} within Wuhan. The spatially and temporally continuous pattern of surface-level PM_{2.5} concentration is produced by STCK integrating 24-hour ground-observed PM_{2.5} data and daily AOD from MODIS. The spatial disparity of surface-level PM_{2.5} accumulation is related to urban function and land use. Contributions of multiple natural and socioeconomic variables, both in center, sub-urban area and whole study area, are quantified using VIMs of RFR. The findings of this study can support better understanding of driving factors for PM_{2.5} accumulation within cities. Although, due to the intrinsic characteristics of RFR, the un-stationary of the relationship between PM_{2.5} concentration and multiple factors is not able to be extensively explored yet. This problem can be solved in future works by using the geo-statistical model considering spatial heterogeneity.

REFERENCES

Benas, N., Beloconi, A., Chrysoulakis, N., 2013. Estimation of urban PM₁₀ concentration, based on MODIS and MERIS/AATSR synergistic observations. *Atmospheric Environment*, Vol. 79, pp. 448-454, doi: <https://doi.org/10.1016/j.atmosenv.2013.07.012> .

Boyouk, N., León, J.-F., Delbarre, H., Podvin, T., Deroo, C., 2010. Impact of the mixing boundary layer on the relationship between PM_{2.5} and aerosol optical thickness. *Atmospheric Environment*, Vol. 44, pp. 271-277, doi: <https://doi.org/10.1016/j.atmosenv.2009.06.053> .

Breiman, L., 2001. Random Forests. *Machine Learning*, Vol. 45, pp. 5-32, doi: <https://doi.org/10.1023/A:1010933404324> .

Chen, Y., Schleicher, N., Cen, K., Liu, X., Yu, Y., Zibat, V., Dietze, V., Fricker, M., Kaminski, U., Chen, Y., Chai, F., Norra, S., 2016. Evaluation of impact factors on PM_{2.5} based on long-term chemical components analyses in the megacity Beijing, China. *Chemosphere*, Vol. 155, pp. 234-242, doi: <https://doi.org/10.1016/j.chemosphere.2016.04.052> .

Elwood, S., 2008. Volunteered geographic information: future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal*, Vol. 72, pp. 173-183, doi: <https://doi.org/10.1007/s10708-008-9186-0> .

Fanizza, C., De Berardis, B., Ietto, F., Soggiu, M.E., Schirò R., Inglessis, M., Ferdinandi, M., Incoronato, F., 2018. Analysis of major pollutants and physico-chemical characteristics of PM_{2.5} at an urban site in Rome. *Science of the total environment*, Vol. 616-617, pp. 1457-1468, doi: <https://doi.org/10.1016/j.scitotenv.2017.10.168> .

Guo, Y., Feng, N., Christopher, S.A., Kang, P., Zhan, F.B., Hong, S., 2014. Satellite remote sensing of fine particulate matter (PM_{2.5}) air quality over Beijing using MODIS. *International Journal of Remote Sensing*, Vol. 35, pp. 6522-6544, doi: <https://doi.org/10.1080/01431161.2014.958245> .

Hu, X., Belle, J.H., Meng, X., Wildani, A., Waller, L.A., Strickland, M.J., Liu, Y., 2017. Estimating PM_{2.5} Concentrations in the Conterminous United States Using the Random Forest Approach. *Environmental science & technology*, Vol. 51, pp. 6936-6944, doi: <https://doi.org/10.1021/acs.est.7b01210> .

Hutengs, C., Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*, Vol. 178, pp. 127-141, doi: <https://doi.org/10.1016/j.rse.2016.03.006> .

Kioumourtoglou, M.-A., Schwartz, J.D., Weisskopf, M.G., Melly, S.J., Wang, Y., Dominici, F., Zanobetti, A., 2016. Long-term PM_{2.5} Exposure and Neurological Hospital Admissions in the Northeastern United States. *Environmental Health Perspectives*, Vol. 124, pp. 23-29, doi: <https://doi.org/10.1289/ehp.1408973> .

Li, C., Hsu, N.C., Tsay, S.-C., 2011. A study on the potential applications of satellite data in air quality monitoring and forecasting. *Atmospheric Environment*, Vol. 45, pp. 3663-3675, doi: <https://doi.org/10.1016/j.atmosenv.2011.04.032> .

Lin, C., Li, Y., Yuan, Z., Lau, A.K.H., Li, C., Fung, J.C.H., 2015. Using satellite remote sensing data to estimate the high-resolution distribution of ground-level PM_{2.5}. *Remote Sensing of Environment*, Vol. 156, pp. 117-128, doi: <https://doi.org/10.1016/j.rse.2014.09.015> .

Liu, H., Yang, B., Kang, E., 2015. Cokriging method for spatio-temporal assimilation of multi-scale satellite data, 2015 *IEEE*

- International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3314-3316, doi: <https://doi.org/10.1109/IGARSS.2015.7326527> .
- Ma, Z., Hu, X., Sayer, A.M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., Liu, Y., 2016. Satellite-Based Spatiotemporal Trends in PM_{2.5} Concentrations: China, 2004-2013. *Environmental Health Perspectives*, Vol. 124, pp. 184-192, doi: <https://doi.org/10.1289/ehp.1409481> .
- Perrone, M.G., Larsen, B.R., Ferrero, L., Sangiorgi, G., De Gennaro, G., Udisti, R., Zangrando, R., Gambaro, A., Bolzacchini, E., 2012. Sources of high PM_{2.5} concentrations in Milan, Northern Italy: Molecular marker data and CMB modelling. *Science of The Total Environment*, Vol.414, pp. 343-355, doi: <https://doi.org/10.1016/j.scitotenv.2011.11.026> .
- Sayer, A.M., Munchak, L.A., Hsu, N.C., Levy, R.C., Bettenhausen, C., Jeong, M.J., 2014. MODIS Collection 6 aerosol products: Comparison between Aqua's e-Deep Blue, Dark Target, and "merged" data sets, and usage recommendations. *Journal of Geophysical Research-Atmospheres*, Vol. 119, pp. 13965-13989, doi: <https://doi.org/10.1002/2014jd022453> .
- Tan, J., Zhang, L., Zhou, X., Duan, J., Li, Y., Hu, J., He, K., 2017. Chemical characteristics and source apportionment of PM_{2.5} in Lanzhou, China. *Science of the total environment*, Vol. 601-602, pp. 1743-1752, doi: <https://doi.org/10.1016/j.scitotenv.2017.06.050> .
- Wang, P., Cao, J.-j., Shen, Z.-x., Han, Y.-m., Lee, S.-c., Huang, Y., Zhu, C.-s., Wang, Q.-y., Xu, H.-m., Huang, R.-j., 2015. Spatial and seasonal variations of PM_{2.5} mass and species during 2010 in Xi'an, China. *Science of The Total Environment*, Vol. 508, pp. 477-487, doi: <https://doi.org/10.1016/j.scitotenv.2014.11.007> .
- Wu, Y., Guo, J., Zhang, X., Tian, X., Zhang, J., Wang, Y., Duan, J., Li, X., 2012. Synergy of satellite and ground based observations in estimation of particulate matter in eastern China. *Science of The Total Environment*, Vol. 433, pp. 20-30, doi: <https://doi.org/10.1016/j.scitotenv.2012.06.033> .
- Yang, D., Wang, X., Xu, J., Xu, C., Lu, D., Ye, C., Wang, Z., Bai, L., 2018a. Quantifying the influence of natural and socioeconomic factors and their interactive impact on PM_{2.5} pollution in China. *Environmental pollution*, Vol. 241, pp. 475-483, doi: <https://doi.org/10.1016/j.envpol.2018.05.043> .
- Yang, M., Wang, Y., Li, H., Li, T., Nie, X., Cao, F., Yang, F., Wang, Z., Wang, T., Qie, G., Jin, T., Du, L., Wang, W., 2018b. Polycyclic aromatic hydrocarbons (PAHs) associated with PM_{2.5} within boundary layer: Cloud/fog and regional transport. *Science of The Total Environment*, Vol. 627, pp. 613-621, doi: <https://doi.org/10.1016/j.scitotenv.2018.01.014> .
- Yao, Y., Liu, X., Li, X., Zhang, J., Liang, Z., Mai, K., Zhang, Y., 2017. Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data. *International Journal of Geographical Information Science*, Vol. 31, pp. 1220-1244, doi: <https://doi.org/10.1080/13658816.2017.1290252> .
- Zalakeviciute, R., Rybarczyk, Y., López-Villada, J., Diaz Suarez, M.V., 2018. Quantifying decade-long effects of fuel and traffic regulations on urban ambient PM_{2.5} pollution in a mid-size South American city. *Atmospheric Pollution Research*, Vol. 9, pp. 66-75, doi: <https://doi.org/10.1016/j.apr.2017.07.001> .
- Zhang, D., Liu, X., Wu, X., Yao, Y., Wu, X., Chen, Y., 2018. Multiple intra-urban land use simulations and driving factors analysis: a case study in Huicheng, China. *GIScience & Remote Sensing*, pp. 1-27, doi: <https://doi.org/10.1080/15481603.2018.1507074> .