

ASSESSMENT OF EXPOSURE TO AMBIENT PM_{2.5} BASED ON GAP-FILLED AEROSOL OPTICAL DEPTH AT URBAN SCALE

Q. Zhan^{1,2,3*}, Z. Fan^{1,2,3}, Y. Xiao^{1,2,3}

¹ School of Urban Design, Wuhan University, Donghunan Street, Wuhan – (qmzhan, zhiyufan)@whu.edu.cn, yhxiao1963@qq.com

² Research Center for Digital City, 8 Donghu South Road, Wuhan 430072, China

³ Collaborative Innovation Centre of Geospatial Technology, 129 Luoyu Road, Wuhan 430079, China -

Commission III, WG III/8

KEY WORDS: Human Exposure, PM_{2.5}, Gap-filled Aerosol Optical Depth, Gradient Boost Decision Tree, Beijing, Urban scale

ABSTRACT:

Air pollution has been a crucial issue affecting human health and has drawn more and more attention in the world. The assessment of exposure to PM_{2.5} of urban residents based on remote sensing is challenging because of the data deficiency in aerosol optical depth (AOD) and the low spatial resolution. This article is devoted to adopt an approach with 2 gradient boosting decision tree (GBDT) models to fill the gaps in AOD and derive continuous PM_{2.5} distribution at urban scale. Then, the assessment of exposure to PM_{2.5} in Beijing was conducted. First, Simplified High Resolution MODIS Aerosol Retrieval Algorithm (SARA) was employed to obtain daily AOD from September 2016 to February 2017 at 500m resolution. Then we used the first GBDT to derive the gap-filled SARA AOD and the second GBDT to estimate PM_{2.5} spatial distribution based on multi-source data. Furthermore, population weighted exposure (PWE) levels of PM_{2.5} and population proportion exposed to PM_{2.5} concentration were estimated by PM_{2.5} distribution and population density data. The result demonstrates that both two GBDT models performed well with cross validation (CV) R² of 0.86 and 0.85 on AOD gap-filling and PM_{2.5} concentration estimation respectively. The areas with high PM_{2.5} concentration are mainly distributed in the east and south of the city but the areas with higher PM_{2.5} exposure are mainly distributed in the urban centre. Moreover, it is found that over 80% people in Beijing are affected by PM_{2.5} pollution in autumn and winter. Overall, the approach this article applied and the analysis results are very useful for epidemiological investigation and air pollution control policy formulation.

1. INTRODUCTION

As the rapid development of urbanization and economy in recent decades, air pollution in urban areas is getting worse and worse. Due to more and more people aggregate to cities, more than 90 percent people in the world are exposed to air pollution (Shi et al., 2019). In urban areas, fine particulate matter (with aerodynamic diameters less than 2.5 μm, PM_{2.5}) has been viewed as one of the most serious air pollutants due to its convenient transmission and easy inhalation by human (Sun et al., 2016). Exposure to ambient PM_{2.5} has an adverse impact on human health with increasing the risk of cardiovascular problems, respiratory problems, stroke and many other diseases (Brauer et al., 2015; Dominici et al., 2006; Huang et al., 2018; Thurston et al., 2015). In addition, the premature death of 8.9 million people had relationship with PM_{2.5} in 2015 (Burnett et al., 2018). Therefore, it is of a great significance to estimate PM_{2.5} concentration and investigate PM_{2.5} exposure especially in urban areas.

Satellite-based PM_{2.5} concentration estimations are applied more and more widely. AOD data that remote sensing satellites offer has been proved to be closely correlated with PM_{2.5} concentration (Li et al., 2005; Lin et al., 2015). Hence, it is available to obtain widely spread PM_{2.5} distribution through AOD products. For example, Hu et al (2017) derived spatial PM_{2.5} distribution of the conterminous United States based on moderate-resolution imaging spectroradiometer (MODIS) AOD products (MYD04_L2) of 10km resolution. Wei et al (2019) applied multi-angle implementation of atmospheric correction (MAIAC) AOD data obtain the PM_{2.5} distribution of China with high accuracy. Whereas, there are numerous data deficient gaps in satellite-based AOD products due to clouds (Yu et al., 2015). It is also a big challenge for

accurate PM_{2.5} estimation because the missing of AOD will cause the data deficiency of PM_{2.5} samples and induce biases to the result furthermore. Some studies used some simple methods such as spatial interpolation and imputation to fill the gaps in AOD (Kloog et al., 2011; Liang et al., 2018). Although these methods are convenient to be applied, the results were poor when there is too much missing neighbourhood data. Thus, complicated and robust methods like machine learning may be considered in AOD gap-filling.

Due to large energy consumption caused by urbanization, Chinese PM_{2.5} pollution is particularly serious in the world (Van Donkelaar et al., 2014). Through the efforts on environmental policies of Chinese government, the overall average PM_{2.5} concentration in China is 45 μg/m³ in 2017 which decreased from 67 μg/m³ in 2013 (Xue et al., 2019). However, compared with the “good” level national standard in GB3095-2012 (China's National Ambient Air Quality Standard, CNAAQs) of 35 μg/m³, China still has a long way to go (China, 2012). In addition, it is noted that the exposure to PM_{2.5}, i.e. population weighted exposure (PWE) with considering the distribution of population can better depict the negative effects of PM_{2.5} on human health than using PM_{2.5} concentration. Donkelaar et al (2014) combined three satellite-derived PM_{2.5} sources to evaluate PM_{2.5} exposure world widely. He and Huang (2018) used geographically and temporally weighted regression (GTWR) to derive PM_{2.5} distribution and calculate the exposure to PM_{2.5}. It was found that over 92% people in China are facing the threat of PM_{2.5} pollution. The PM_{2.5} exposure in Australia was estimated and only 17% population lived in the areas with PM_{2.5} concentration higher than 8 μg/m³ (Knibbs et al., 2018). However, the information these studies provided is only adaptive at large scale. There were few studies conducted at

urban scale due to the limitation of spatial resolution. The resolution of PM_{2.5} data most studies adopted is higher than 1km but the studies in urban-scale requires data with finer resolution.

Our work is aimed to use the SARA AOD products with 500m resolution to assess the exposure to PM_{2.5} with taking Beijing as the example. GBDT, a good machine learning method, was applied to fill the gaps in SARA AOD and estimate PM_{2.5} concentration. Then, the PWE to PM_{2.5} and population proportion exposed to PM_{2.5} concentration was calculated further based on population density.

2. MATERIALS AND METHODS

2.1 Study Area

Beijing (as Figure 1 shows) is the capital of China, located in the North China Plain (NCP) with total area of 16410.54 km² and population of 21.54 million. Due to the rapid economic development and urbanization, Beijing has been the area with the most serious air pollution burden in China (Wang et al., 2018).

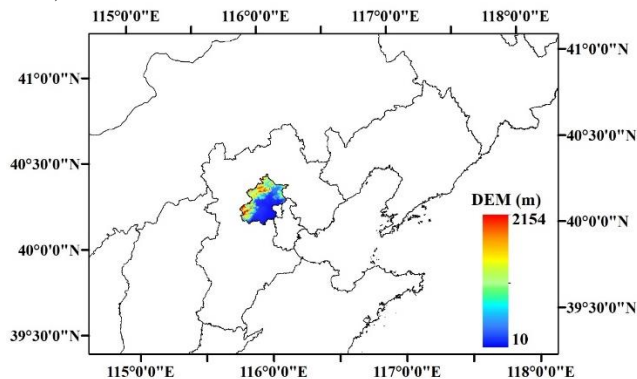


Figure 1. The location of Beijing.

2.2 SARA AOD

Simplified Aerosol Retrieval Algorithm (SARA) is a stable AOD retrieval algorithm with good reliability under complex atmospheric environment (Bilal et al., 2013; Bilal et al., 2014). More importantly, the spatial resolution of SARA AOD is 500m, which is much higher than other AOD products and more adaptive for urban scale. SARA have three assumptions: (1) The surface is Lambertian. (2) Single scattering approximation. (3) The single scattering albedo and asymmetric factor do not vary spatially over the region on day of retrieval (Bilal et al., 2013). In addition, traditional retrieval lookup table is not required in the retrieval process so it is convenient to implement this AOD retrieve method. It directly obtains AOD results based on AOD ground observation data, topography and angle data, Top of Atmosphere (TOA) radiance data and surface reflectance data. Here, we used ground AOD data from Aerosol Robotic Network (AERONET, <https://aeronet.gsfc.nasa.gov/>) and downloaded TOA product MOD02HKM, topography and angle product MOD03 and surface reflectance product MOD09 from the NASA website (<https://ladsweb.nascom.nasa.gov/>).

SARA AOD is also deeply affected by clouds like other AOD products. The existing of clouds disturbs the acquaintance of correct data. We used MOD35, the cloud mask product of MODIS, to remove the pixels covered by clouds. However, this approach conducted many gaps in SARA AOD so we

would use a GBDT to fill them and obtain the full-covered SARA AOD.

2.3 PM_{2.5} Concentration

Ground-level PM_{2.5} observation data used in this article was downloaded from Ministry of Environmental Protection of China (MEPC). The hourly PM_{2.5} data spanned from September, 2016 to February, 2017 (basically the autumn and winter in Beijing) in 35 ground monitoring sites of Beijing was collected and calculated further for daily average PM_{2.5} concentrations. Among the 35 sites, there are 17 sites located in core urban area (Haidian district, Dongcheng district, Xicheng district, Chaoyang district, Shijingshan district and Fengtai district) and 18 sites evenly distributed in suburb areas.

2.4 Auxiliary Data

Meteorological field data, digital elevation model (DEM) data and Normalized Difference Vegetation Index (NDVI) were applied for the SARA AOD gap-filling and PM_{2.5} concentrations estimating. Meteorological field data including temperature at 2m (TEM), relative humidity (RH), surface pressure (SP), planetary boundary layer height (PBLH), eastward wind at 10m above displacement height (U10M) and northward wind at 10m above displacement height (V10M) were downloaded from European Centre for Medium-Range Weather Forecasts (ECMWF, <http://www.ecmwf.int/>). The daily meteorological field were averaged based on the data at 2:00, 8:00, 14:00 and 20:00. MODIS NDVI 16-day products (MOD13A1) were obtained from LAADS website (<https://ladsweb.nascom.nasa.gov/>). The DEM data used in this article was downloaded from Shuttle Radar Topographic Mission (SRTM) (<http://srtm.csi.cgiar.org/>). Each auxiliary raster data was re-projected and resampled to a grid with 500m followed by SARA AOD.

2.5 Population Density

For assessing the exposure to PM_{2.5}, population density raster data with 1km was obtained from the Data Center for Resources and Environmental Sciences (DCRES) (<http://www.resdc.cn>). The population density raster product was derived based on land use data, night light remote sensing data and the distribution of residential density. It was also re-projected and resampled to the defined 500m grid mentioned in 2.4.

2.6 Gradient Boost Decision Tree (GBDT)

GBDT is one of the boosting learning methods in machine learning family. It is an ensemble learning method containing a series of based learners. The parameters of GBDT is complicated compared with other ensemble learning methods which ensure its outstanding performance on processing multi-dimensioned and massive data (Friedman, 2001; Reid et al., 2015). Thus, GBDT is a good option for AOD missing data recovering and PM_{2.5} concentration estimation.

In terms of model structure, the form of based learner in GBDT is decision tree which is similar to Random Forest (RF). However, unlike RF, these based learners are not independent due to each one is generated based on the residues of the previous one. The predicted result is the weighted addition of all based decision trees. The form is as (1) shows. Here, T_i is the additive result of all previous $i - 1$ trees. H_i is the decision tree trained based on the former learner. W_i is the weight of H_i .

$$T_i = T_{i-1} + W_i H_i \quad (1)$$

We constructed 2 GBDT models for AOD gap-fill and PM2.5 estimation respectively. In the model for AOD gap-fill (call it model 1), all auxiliary data mentioned in 2.4 and the day of year (DOY) were used as the independent variables. Compared with model 1, the model for PM2.5 estimation (call it model 2) added the full-covered AOD obtained from model 1 as the major independent variable. Some key parameters in both two models were selected, which is shown in Table 1. GBDT can evaluate the importance of all predictors by its variable importance measure. Due to there were 2 GBDT models for AOD gap-fill and PM2.5 estimation, the variable importance measures (VIM) of the two models were calculated separately. VIM was evaluated by the variable using frequency in model construction, which is to say the more the variable are adopted, the more its contribution to the model. A sample based 10-fold CV was adopted for validating the performance of two GBDT models. Three measurements including R square (R^2), mean predictive error (MPE), root mean square error (RMSE) are used for the evaluation. Both the construction of GBDT and the CV evaluation were done based on the Python 3.7 environment.

Parameters	Model 1	Model 2
Number of trees	300	200
Max depth	3	4
Learning rate	0.1	0.08
Alpha	0.9	0.9

Table 1. Some key parameters selection in two models

2.7 Ambient PM2.5 Population Weighted Exposure Assessment

For investigating the impact of PM2.5 on urban residents, ambient PM2.5 population exposure was assessed based on the PM2.5 estimation results. The total PWE was calculated based on (2)(Liu et al., 2017; Zhan et al., 2017). Here, P_i and C_i are the population and the PM2.5 concentration in grid i respectively.

$$PWE_{total} = \frac{\sum_{i=1}^N P_i C_i}{\sum_{i=1}^N P_i} \quad (2)$$

Furthermore, we calculated the PWE of each grid as (3) shows.

$$PWE_i = \frac{P_i C_i}{\sum_{i=1}^N P_i} \quad (3)$$

3. RESULT AND DISCUSSION

3.1 Model Performance

The CV performance of two GBDT models are shown in Table 2 and Table 3. For model 1, the result shows a good performance with CV R^2 of 0.86, MPE of 0.10 and RMSE of 0.12. The performance in October is the best with CV R^2 of 0.91 while that in November is the worst with CV R^2 of 0.77. For model 2, the performance is also good with CV R^2 of 0.85, MPE of 22.47 $\mu\text{g}/\text{m}^3$ and RMSE of 30.94 $\mu\text{g}/\text{m}^3$. CV R^2 , calculated based on test sets, is an important indicator related to overfitting with higher value indicating weaker overfitting relevantly. Overall, the performance of both models are good despite some tolerated overfitting with CV R^2 lower than 0.8 in few months. The best performance and the worst performance are in January and November respectively. The reason for the different performance among months may be

related to the sample size. The good performance of two models suggests that GBDT can capture the characteristics of both AOD and PM2.5 at urban scale. Zhang et.al applied RF in MODIS 3km AOD gap filling and PM2.5 estimation using Sichuan basin as the study area with CV R^2 of 0.95 and 0.86 respectively(Zhang et al., 2018). Though the accuracy of two GBDT models in our work is a little lower in contrast, it is acceptable because of the differences in study area and period range.

Table 4 shows the VIMs of two models. PBLH plays a very significant role in both two models. That is because AOD is the attribution of aerosol in the vertical structure and PBLH determines the range of vertical aerosol distribution. In addition, PBLH affects the ratio of PM2.5 to AOD with a negative relationship so it is also important for PM2.5 estimation models(Sun et al., 2018). DOY is another important factor which is also demonstrated in some other studies (Guo et al., 2017; Zhao et al., 2019; Zheng et al., 2017). Both AOD and PM2.5 vary greatly temporally and sometimes there is a big difference between two continual days. In addition, there is a huge difference on performance between winter and autumn. On one hand, in the winter (includes December, January and February) of Beijing, aerosol accumulates in the low height which can aggregate the PM2.5 pollution. On the other hand, due to the requirement of heating, more energy consuming are conducted in winter which also produces many pollutants. In addition, some scholars have found that other pollutants are also helpful for the improvement of model accuracy(Wang et al., 2020). It is worth trying to apply other pollutants such as SO_2 and NO_2 as the independent variables in our future research.

Model 1	R^2	MPE	RMSE
September	0.86	0.10	0.17
October	0.91	0.10	0.14
November	0.77	0.08	0.12
December	0.85	0.07	0.10
January	0.83	0.07	0.10
February	0.81	0.05	0.08
Overall	0.86	0.10	0.12

Table 2. CV performance of the first GBDT for AOD gap-filling

Model 2	R^2	MPE ($\mu\text{g}/\text{m}^3$)	RMSE ($\mu\text{g}/\text{m}^3$)
September	0.77	15.86	22.76
October	0.78	22.08	29.60
November	0.73	28.01	37.08
December	0.88	26.23	36.77
January	0.89	23.83	32.32
February	0.80	21.54	32.93
Overall	0.85	22.47	30.94

Table 3. CV performance of the second GBDT for PM2.5 estimation

Variable	Model 1	Model 2
AOD	-	0.23
TEM	0.17	0.06
RH	0.07	0.08
SP	0.02	0.06
U10M	0.04	0.02
V10M	0.05	0.06
PBLH	0.24	0.36
NDVI	0.01	0.01
DEM	0.03	0.01
DOY	0.36	0.18

Table 4. VIM of two models

3.2 Gap-filled SARA AOD

Figure 2 shows the differences between the average cloud-removal AOD and the average gap-filling AOD. It can be obviously found that both of them approximately share the same value range from 0.16 to 0.57 while the average gap-filled AOD is smoother. The reason for this phenomenon is the random disturbance of clouds which cause the non-continuous change of average AOD in space. Without AOD gap-filling process, these gaps will not only cause the unavailability of daily estimation but also eventually induce some biases to the average estimation which inevitably affects the accuracy of human health investigations such as PWE level assessment. In terms of space, AOD values are low in western and northern mountains and high in eastern and southern flatten plains. Other than terrain factor, socioeconomic factors also play important roles affecting AOD distribution. For example, most populations and industries are concentrated on the southern and eastern areas conducting much emissions.

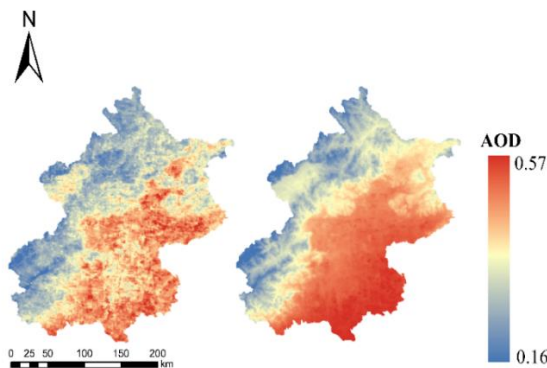


Figure 2. Average AOD in autumn and winter of Beijing (The left one is the cloud-removal SARA AOD and the right one is the gap-filled SARA AOD).

3.3 Spatial Distribution of PM2.5 Concentrations and Population Density Weighted Exposure

PM2.5 concentration spatial distribution was retrieved based on the second GBDT model with using the full-covered AOD as the major predictor. The average PM2.5 spatial distribution is shown in Figure 3. The spatial distribution pattern of average PM2.5 demonstrates a decreasing trend from southeast to northwest which is similar to AOD. This also proves the strong relationship between AOD and PM2.5. The range of average PM2.5 in autumn and winter of Beijing is from 47 $\mu\text{g}/\text{m}^3$ to 144 $\mu\text{g}/\text{m}^3$. According to CNAAQs, when PM2.5 concentration is higher than 75 $\mu\text{g}/\text{m}^3$, the environment quality can be defined as ‘pollution’. Thus, it can be seen that the PM2.5 pollution issue in autumn and winter of Beijing is extremely serious with most areas of PM2.5 concentration higher than that level.

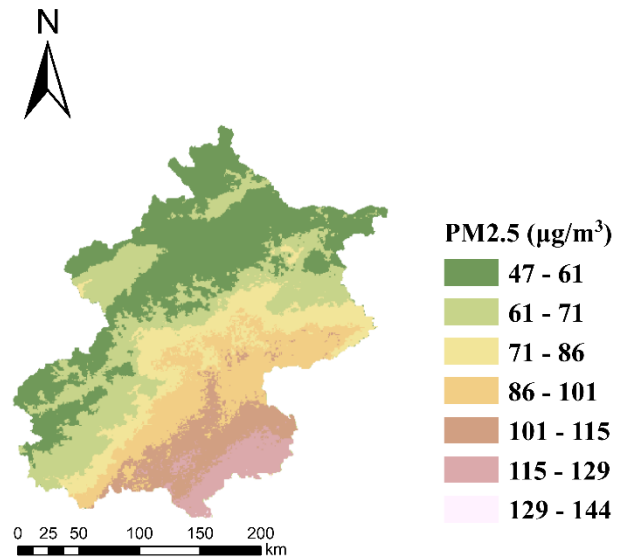


Figure 3. Estimated average PM2.5 concentration distribution in autumn and winter of Beijing.

Furthermore, we obtained the PWE combined with population density data and PM2.5 average distribution. We used Natural Breaks (Jenks) method divide it into 7 levels (higher levels, heavier exposure to PM2.5) as Figure 4 shows. From the perspectives of spatial PWE distribution, it demonstrates that the high level PWEs are mainly distributed in urban central areas. This is different from the distribution pattern of PM2.5 concentration that southeast areas of Beijing suffer heavier PM2.5 pollution. It suggests that Beijing government should pay more attention to the pollution transmission from southeast areas to urban central areas. Total PWE of the whole period and six months were calculated as Table 5 shows. Among all six months, residents in Beijing are most affected by PM2.5 pollution in December with PWE of 124 $\mu\text{g}/\text{m}^3$. Moreover, there is no month’s average PWE reaches the “good” air quality in CNAAQs (less than 35 $\mu\text{g}/\text{m}^3$). Figure 5 shows the population proportion in Beijing exposed to PM2.5 concentration over the six months. On the whole, more than 80% people of Beijing suffered from PM2.5 pollution (higher than 75 $\mu\text{g}/\text{m}^3$). In addition, it is noted that there are 16% and 6% people living in the areas with PM2.5 concentration exceeding 150 $\mu\text{g}/\text{m}^3$ (the standard of “serious pollution” in CNAAQs) in December and January respectively. Thus, the series of analysis results above underline the importance and the urgency of air pollution reduction in Beijing.

Period	PWE ($\mu\text{g}/\text{m}^3$)
September	54
October	83
November	101
December	124
January	112
February	71
Overall	91

Table 5. Total PWE of the whole period and all month.

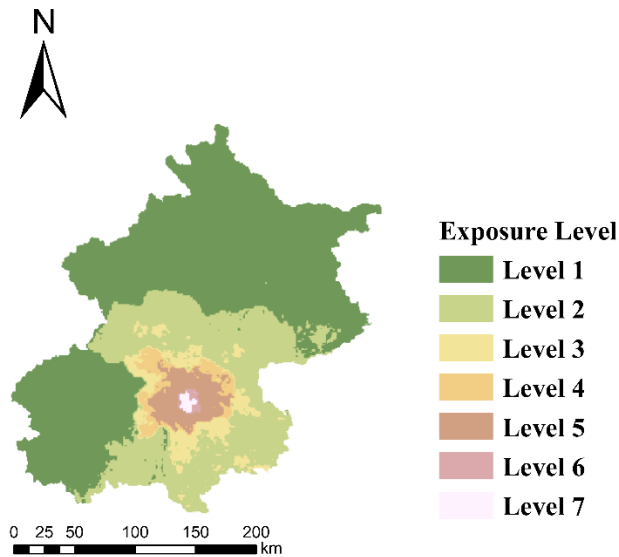


Figure 4. PWE level spatial distribution in autumn and winter of Beijing.

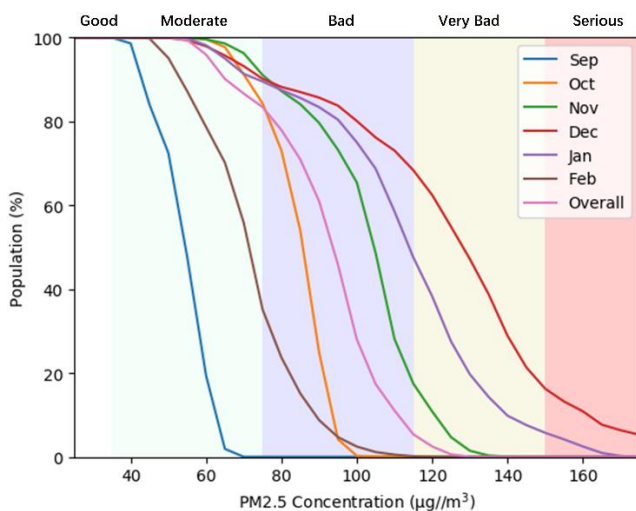


Figure 5. Beijing population proportion exposed to PM2.5 concentration in different months.

4. CONCLUSION

To summary, this article applied two GBDT models to derive the gap-filled SARA AOD and PM2.5 concentration distribution with 500m resolution. Based on the result of PM2.5 estimation, PWE and population proportion exposed to PM2.5 in Beijing are derived. The performance is good with CV R² of 0.86 and 0.85, MPE of 0.10 and 22.47 µg/m³ and RMSE of 0.12 and 30.94 µg/m³ for AOD gap-filling (model 1) and PM2.5 concentration estimation (model 2) respectively. Through the results of exposure to PM2.5, it is found that high levels of PWE mainly assemble in urban central areas. Overall, over 80% people in Beijing live in the areas with PM2.5 concentration higher than 75 µg/m³ which meets the standard of “pollution” in CNAQS and the pollution situation is serious in December and January. This work also has some disadvantages. First, the population density data doesn’t vary temporally thus the dynamic changes of exposure to PM2.5

may not be captured precisely. Second, the PM2.5 observations in our study are relatively low which causes a little overfitting in few months. Thus, we will extend the study period and obtain more observations in the future. In addition, the PM2.5 estimation model’s accuracy may be improved by introducing other pollutants as independent variables. In general, the approach and the information this article provide are valuable for air pollution abatement in the future.

ACKNOWLEDGEMENTS

We sincerely appreciated for the support from the National Natural Science Foundation of China (No. 51878515, 41331175 and 51378399).

REFERENCES

- Bilal, M., Nichol, J.E., Bleiweiss, M.P., Dubois, D., 2013: A Simplified high resolution MODIS Aerosol Retrieval Algorithm (SARA) for use over mixed surfaces. *Remote Sensing of Environment* 136, 135-145.
- Bilal, M., Nichol, J.E., Chan, P.W., 2014: Validation and accuracy assessment of a Simplified Aerosol Retrieval Algorithm (SARA) over Beijing under low and high aerosol loadings and dust storms. *Remote sensing of environment* 153, 50-60.
- Brauer, M., Freedman, G., Frostad, J., Van Donkelaar, A., Martin, R.V., Dentener, F., Dingenen, R.v., Estep, K., Amini, H., Apte, J.S., 2015: Ambient air pollution exposure estimation for the global burden of disease 2013. *Environmental science & technology* 50(1), 79-88.
- Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C.A., Apte, J.S., Brauer, M., Cohen, A., Weichenthal, S., 2018: Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences* 115(38), 9592-9597.
- China, M., 2012. Ambient Air Quality Standards. GB 3095-2012. China Environmental. Science Press, Beijing.
- Dominici, F., Peng, R.D., Bell, M.L., Pham, L., McDermott, A., Zeger, S.L., Samet, J.M., 2006: Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *Jama* 295(10), 1127-1134.
- Friedman, J.H., 2001: Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 29(5), 1189-1232.
- Guo, J., Xia, F., Zhang, Y., Liu, H., Li, J., Lou, M., He, J., Yan, Y., Wang, F., Min, M., 2017: Impact of diurnal variability and meteorological factors on the PM2.5-AOD relationship: Implications for PM2.5 remote sensing. *Environmental pollution* 221, 94-104.
- He, Q., Huang, B., 2018: Satellite-based mapping of daily high-resolution ground PM2.5 in China via space-time regression modeling. *Remote sensing of environment* 206, 72-83.
- Hu, X., Belle, J.H., Meng, X., Wildani, A., Waller, L.A., Strickland, M.J., Liu, Y., 2017: Estimating PM2.5 concentrations in the conterminous United States using the random forest approach. *Environmental science & technology* 51(12), 6936-6944.
- Huang, J., Pan, X., Guo, X., Li, G., 2018: Health impact of China’s Air Pollution Prevention and Control Action Plan: an analysis of national air quality monitoring and mortality data. *The Lancet Planetary Health* 2(7), e313-e323.
- Kloog, I., Koutrakis, P., Coull, B.A., Lee, H.J., Schwartz, J., 2011: Assessing temporally and spatially resolved PM2.5 exposures for epidemiological studies using satellite aerosol

- optical depth measurements. *Atmospheric environment* 45(35), 6267-6275.
- Knibbs, L.D., Van Donkelaar, A., Martin, R.V., Bechle, M.J., Brauer, M., Cohen, D.D., Cowie, C.T., Dirgawati, M., Guo, Y., Hanigan, I.C., 2018: Satellite-based land-use regression for continental-scale long-term ambient PM_{2.5} exposure assessment in Australia. *Environmental science & technology* 52(21), 12445-12455.
- Li, C., Mao, J., Lau, A.K., Yuan, Z., Wang, M., Liu, X., 2005: Application of MODIS satellite products to the air pollution research in Beijing. *Science in China Series D(Earth Sciences)* 48(z2), 209-219.
- Liang, F., Xiao, Q., Wang, Y., Lyapustin, A., Li, G., Gu, D., Pan, X., Liu, Y., 2018: MAIAC-based long-term spatiotemporal trends of PM_{2.5} in Beijing, China. *Science of the Total Environment* 616, 1589-1598.
- Lin, C., Li, Y., Yuan, Z., Lau, A.K., Li, C., Fung, J.C., 2015: Using satellite remote sensing data to estimate the high-resolution distribution of ground-level PM_{2.5}. *Remote Sensing of Environment* 156, 117-128.
- Liu, Z., Xie, M., Tian, K., Gao, P., 2017: GIS-based analysis of population exposure to PM_{2.5} air pollution—A case study of Beijing. *Journal of Environmental Sciences* 59(9), 48-53.
- Reid, C.E., Jerrett, M., Petersen, M.L., Pfister, G.G., Morefield, P.E., Tager, I.B., Raffuse, S.M., Balmes, J.R., 2015: Spatiotemporal prediction of fine particulate matter during the 2008 Northern California wildfires using machine learning. *Environmental science & technology* 49(6), 3887-3896.
- Shi, Y., Ren, C., Lau, K.K.-L., Ng, E., 2019: Investigating the influence of urban land use and landscape pattern on PM_{2.5} spatial variation using mobile monitoring and WUDAPT. *Landscape and Urban Planning* 189, 15-26.
- Sun, L., Wei, J., Duan, D., Guo, Y., Yang, D., Jia, C., Mi, X., 2016: Impact of Land-Use and Land-Cover Change on urban air quality in representative cities of China. *Journal of Atmospheric and Solar-Terrestrial Physics* 142, 43-54.
- Sun, T., Che, H., Qi, B., Wang, Y., Dong, Y., Xia, X., Wang, H., Gui, K., Zheng, Y., Zhao, H., 2018: Aerosol optical characteristics and their vertical distributions under enhanced haze pollution events: effect of the regional transport of different aerosol types over eastern China. *Atmospheric Chemistry and Physics* 18, 2949-2971.
- Thurston, G.D., Burnett, R.T., Turner, M.C., Shi, Y., Krewski, D., Lall, R., Ito, K., Jerrett, M., Gapstur, S.M., Diver, W.R., 2015: Ischemic heart disease mortality and long-term exposure to source-related components of US fine particle air pollution. *Environmental health perspectives* 124(6), 785-794.
- Van Donkelaar, A., Martin, R.V., Brauer, M., Boys, B.L., 2014: Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. *Environmental health perspectives* 123(2), 135-143.
- Wang, X., Tian, G., Yang, D., Zhang, W., Lu, D., Liu, Z., 2018: Responses of PM_{2.5} pollution to urbanization in China. *Energy policy* 123, 602-610.
- Wang, X., Sun, W., Zheng, K., Ren, X., Han, P., 2020. Estimating hourly PM_{2.5} concentrations using MODIS 3 km AOD and an improved spatiotemporal model over Beijing-Tianjin-Hebei, China. *Atmospheric Environment* 222, 117089.
- Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., Cribb, M., 2019: Estimating 1-km-resolution PM_{2.5} concentrations across China using the space-time random forest approach. *Remote Sensing of Environment* 231, 111221.
- Xue, T., Liu, J., Zhang, Q., Geng, G., Zheng, Y., Tong, D., Liu, Z., Guan, D., Bo, Y., Zhu, T., 2019: Rapid improvement of PM_{2.5} pollution and associated health benefits in China during 2013–2017. *Science China Earth Sciences* 62(12), 1847-1856.
- Yu, C., Di Girolamo, L., Chen, L., Zhang, X., Liu, Y., 2015: Statistical evaluation of the feasibility of satellite-retrieved cloud parameters as indicators of PM_{2.5} levels. *Journal of Exposure Science and Environmental Epidemiology* 25(5), 457.
- Zhan, Y., Luo, Y., Deng, X., Chen, H., Grieneisen, M.L., Shen, X., Zhu, L., Zhang, M., 2017: Spatiotemporal prediction of continuous daily PM_{2.5} concentrations across China using a spatially explicit machine learning algorithm. *Atmospheric environment* 155, 129-139.
- Zhang, R., Di, B., Luo, Y., Deng, X., Grieneisen, M.L., Wang, Z., Yao, G., Zhan, Y., 2018: A nonparametric approach to filling gaps in satellite-retrieved aerosol optical depth for estimating ambient PM_{2.5} levels. *Environmental pollution* 243, 998-1007.
- Zhao, C., Liu, Z., Wang, Q., Ban, J., Chen, N.X., Li, T., 2019: High-resolution daily AOD estimated to full coverage using the random forest model approach in the Beijing-Tianjin-Hebei region. *Atmospheric environment* 203, 70-78.
- Zheng, C., Zhao, C., Zhu, Y., Wang, Y., Shi, X., Wu, X., Chen, T., Wu, F., Qiu, Y., 2017: Analysis of influential factors for the relationship between PM_{2.5} and AOD in Beijing. *Atmospheric Chemistry and Physics* 17(21), 13473-13489.