Online-Journals.org

iTDAF IETI Transactions on Data Analysis and Forecasting

iTDAF | eISSN: 2959-0442 | Vol. 1 No. 4 (2023) | OPEN ACCESS

https://doi.org/10.3991/itdaf.v1i4.46311

PAPER

Quantitative Estimation of Urban PM2.5 Pollution Baseline and Meteorological Resource Endowment Using Machine Learning in Chinese Yangtze River Economic Belt

Changhong Ou[1,2], Fei Li[1,2](✉), Jingdong Zhang[1,2], Jinyuan Guo[1,2], Xiyao Chen[1,2], Shaojie Kong[1,2], Pei Jiang[1,2], Mian Wu[1,2], Yazhu Wang[1]

[1]Research Center for Environment and Health, Zhongnan University of Economics and Law, Wuhan, China

[2]School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan, China

lifei@zuel.edu.cn

ABSTRACT

Considering the influence of baseline values, meteorological conditions, and human activities on PM2.5, quantifying them will facilitate the classification, control, and management of pollution. The machine learning model explained the PM2.5-meteorological nonlinear relationship between PM2.5 and meteorological factors in each city across the Yangtze River Economic Belt, China. Meteorological resource endowments (MRE) are used to quantify the variation on PM2.5 concentration caused by meteorological conditions. Contamination baseline (CB) is used to characterize the lowest limit of anthropogenic impact in PM2.5 contamination without meteorological interference. According to the values of MRE and CB, cities in the Yangtze River economic belt can be divided into four categories (Q1-4). The average value of MRE is –0.41 μg/m^3. The average value of CB is 34.05 μg/m^3, which is lower than the Chinese Grade II standard (GB 3095-2012). The additional emissions by humans resulted in an increase of 7 μg/m^3 in concentration, while the meteorological factors led to a decrease of –0.41 μg/m^3. In terms of city classification, Q1 is concentrated in the midstream, and PM2.5 is the most challenging pollutant to control. Q2 is concentrated downstream, with relatively high PM2.5 emissions but favorable meteorological conditions. Q3 is concentrated upstream, and there is surplus environmental capacity even with limited meteorological conditions. Cites in Q4 have the most suitable development potential and exhibit a discrete spatial distribution. The research distinguished various categories of pollution and provided insights into the different characteristics of pollution around the Yangtze River Economic Belt. This information has helped the government classify cities and implement specific policies based on their individual situations.

KEYWORDS

machine learning, PM2.5, meteorological resource endowments, contamination baseline, hierarchical management, Yangtze River economic belt

Ou, C., Li, F., Zhang, J., Guo, J., Chen, X., Kong, S., Jiang, P., Wu, M., Wang, Y. (2023). Quantitative Estimation of Urban PM2.5 Pollution Baseline and Meteorological Resource Endowment Using Machine Learning in Chinese Yangtze River Economic Belt. *IETI Transactions on Data Analysis and Forecasting (iTDAF)*, 1(4), pp. 68–78. https://doi.org/10.3991/itdaf.v1i4.46311

Article submitted 2023-09-28. Revision uploaded 2023-11-01. Final acceptance 2023-11-03.

1 INTRODUCTION

PM2.5 is one of the major pollutants in the atmosphere and already poses significant health risks to the population [1]. Especially in developing countries such as China, PM2.5 causes millions of deaths and results in economic losses due to health-related expenses [2]. In recent years, due to the successful implementation of China's air pollution policy, the overall concentration of PM2.5 has decreased. The scope of pollution has shifted from large to specific areas. Therefore, regional-graded prevention and control plans for China's atmosphere have been proposed. For example, the "2 + 26" cities coordinated their management plans in the Beijing-Tianjin-Hebei urban agglomeration [3]. The measure will address persistent PM2.5 pollution at the regional level.

Before implementing graded management measures for PM2.5 in a specific region, it is necessary to classify the cities within that region. The classification is generally based on the concentration and impact of natural and human activities. Meteorology is an important natural factor considered to have a significant influence on the trend of PM2.5 concentrations, exhibiting strong spatial-temporal heterogeneity [4]. In some studies, the meteorological factor accounted for 10% of the total concentration and even reached 50% in the most polluted cities [5, 6]. Therefore, there is an urgent need for a relatively convenient method to eliminate the meteorological impact and establish the true emission trend of PM2.5. Weather normalization (WN) is a comprehensive algorithm based on machine learning [7–9]. It effectively explains the relationship between PM2.5 and meteorological conditions and removes the meteorological component [10–12]. Meteorological resource endowment (MRE) was used to quantify the specific concentrations of PM2.5 influenced by meteorological factors. It can be understood as the difference in PM2.5 concentration between the observed values (OV) and the predicted values (PV) by weather normalization.

The anthropogenic emission of PM2.5 can be analyzed by the concentration value after eliminating the influence of meteorological factors. The contamination baseline (CB) is characterized as the lowest limit of long-term anthropogenic impact on soil and atmospheric contamination [13–15]. The cumulative frequency curve method is the most commonly used approach for determining CB. This method can determine the baseline value of almost any substance [16]. In this paper, CB refers to the value of PM2.5 concentrations in each city, taking into account long-term human activities and meteorological conditions. Therefore, integrating the PV using the cumulative frequency method would yield CB [17]. After analyzing the two impacts of PM2.5, cities in a region will be divided into multiple categories, representing the different combined effects of anthropogenic emissions and meteorological conditions. These classifications provide a reliable data basis for regional grade management.

The Yangtze River Economic Belt is a significant national strategic development region in China, with a population and gross domestic product (GDP) that exceed 40% of the country's total. Spanning three major regions in East and West China, there are significant differences in meteorological resource endowments (MRE) and urban development levels among various cities [18]. This paper confirmed the presence of CB and MRE in each city located around the Yangtze River Economic Belt. Based on the two indicators mentioned above, cities along the Yangtze River Economic Belt are classified into four categories. This classification method has been practiced in the literature [19]. In addition to supporting government management, the resulting dataset can play an important role in studying the relationship between environment and health, weather and pollution, and environmental inequalities in different cities.

2 MATERIALS AND METHODS

2.1 Research area and data sources

The Yangtze River Economic Belt involves nine provincial administrative units, namely Jiangsu, Zhejiang, Anhui, Jiangxi, Hunan, Hubei, Sichuan, Yunnan, and Guizhou, as well as Shanghai and Chongqing (see Figure 1). It also includes 130 prefecture-level administrative regions, which consist of prefecture-level cities, regions, and autonomous regions. Based on their geographical locations, they can be divided into three urban agglomerations in the upstream, midstream, and downstream areas. After excluding the missing data values, 125 cities participated in this study.

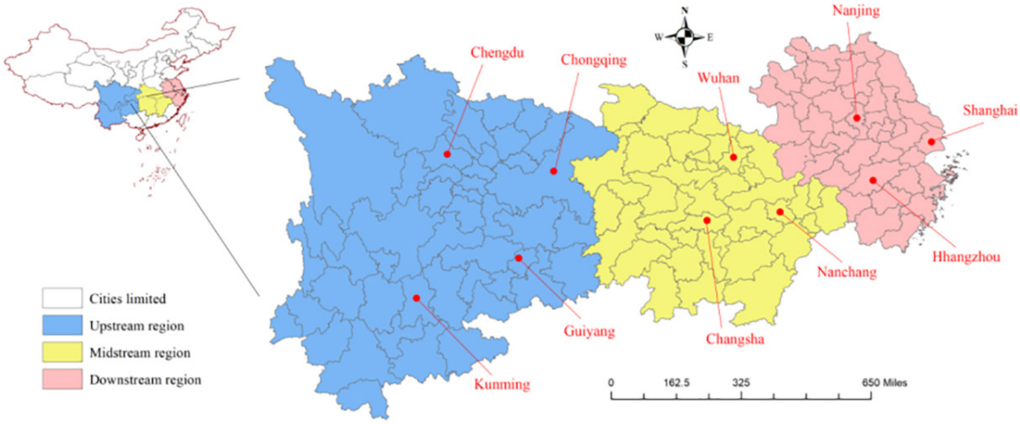

Fig. 1. The location of division of the Yangtze River Economic Belt in China and the up, mid and down streams

Daily PM2.5 data in various prefecture-level cities is provided by the Ministry of Ecology and Environment of China (http://www.mee.gov.cn/). Daily meteorological data were extracted from the China National Meteorological Administration website (http://www.cma.gov.cn/). The input dataset is presented in Table 1. After removing anomalous and missing values, a total of 272,738 records were obtained for 125 cities with available data, including PM2.5 and meteorological indicators. The data covers a time period of 2191 days, from 2015 to 2020, with an average of 2182 records per city.

Table 1. The abbreviation and meaning of the inputting variables of WN

Abbreviation	Meaning	Unit
PM2.5		μg/m^3
Day_julian	Number of days in a year	day
Weekday	Number of days in a week	day
Apm	Air pressure	Pa
Rhm	Relative Humidity	%
Rzm	Daylight hours	h
Tm	Temperatures	℃
Wsm	Wind speed	m/s
Wdm	Wind direction (1~16 bearing numbers)	–
Wsmax	Maximum wind speed	m/s

2.2 Creation of weather normalization models (WN)

The basic structure of WN is a random forest in machine learning that can predict PM2.5 features based on input data from observed PM2.5 and meteorological conditions. Random forest (RF) is composed of a specific number of individual decision tree models. Each tree randomly extracts observations and samples its prediction features. It then replaces the training dataset, randomly selects the prediction features to achieve the best split for each node, and finally obtains the predicted values [20]. Random forest only requires a small number of tuning parameters from the user [21]. Grange (2020) set the number of independent or explanatory variables used to grow a tree at three and the minimum node size at five. The user-adjustable parameter is the number of predicted features, which was set to 1000, following Mallet [22].

Weather normalization can predict the concentration of air pollutants at a specific measurement time point based on randomly selected meteorological conditions. In the input dataset of WN, PM2.5 observations and various meteorological factors are located in the same table. WN would add temporal variables for each PM2.5 data. Temporal and meteorological data would be considered control variables. The dataset was added to a predictive RF model in order to analyze the trend of observed PM2.5 over a period of five years and gain insights into the formation pathways of PM2.5. RF models can derive nonlinear relationships between PM2.5 and control variables in this dataset. After that, WN would output the weather-normalized levels of pollutants. For a specific day, the model randomly selects time variables and weather parameters from the predicted feature data subset for any day during the entire study period. Subsequently, they were inputted into the RF model to predict PM2.5 at that specific time point. The aforementioned behavior was repeated 1000 times, and the predictions were averaged to represent the "average" meteorological conditions. This average was then considered the weather-normalized trend [20]. In addition, WN can describe the relationship between air pollutants and predictive index characteristics, as well as the importance and partial dependency of each meteorological factor. WN is a collection of multiple computing processes that can be completed with just a few commands, making it widely applicable worldwide.

Weather normalization does not directly offer the ability to determine error or uncertainty estimates [12]. The input dataset will be divided into two parts: a training set (80%) and a test set (20%). The test set was used to verify whether the model was overfitting. It measured the difference between the predicted value and the actual PM2.5 using the variables in the test set [9]. The result is characterized by R^2.

2.3 Creation of MRE, CB and city classification

After obtaining daily PM2.5 PV data, the annual mean difference between the PM2.5 OV and PV, obtained by removing the fluctuating variation of meteorology, is referred to as MRE. The formula used to calculate the MRE is as follows:

$$\mathrm{MRE} = \frac{\sum_{i=1}^{n} (\mathrm{OVi} - \mathrm{PVi})}{n} \tag{1}$$

Where "n" represents the total number of days involved in the calculation. Therefore, it is possible to calculate the MRE for a specific year or a span of 6 years using a range of n. OVi and PVi denote the OV and PV on the i-th day. When the value of it is less than 0, MRE has a positive effect on reducing PM2.5 concentration. Conversely, when the value of MRE is greater than 0, it has a negative effect.

The cumulative frequency curve method is used for determining CB [23]. Referring to the calculation method [24, 25], we plotted the distribution curve of PM2.5 concentration. The X-axis represents the concentration, while the Y-axis represents the decimal coordinates of cumulative frequency. The regression coefficients of the curve, along with their significance levels, were utilized to determine the inflection points for the limit conditions. Below the inflection point, there is either the upper limit of background values or the lower limit of human activities. The baseline value is the mean value of the content of samples smaller than the inflection point. Since a certain concentration of PM2.5 is inevitably generated during human activities, the current study includes both naturally occurring PM2.5 in each city and the minimum emissions generated from basic human activities. In this research, firstly, the daily PM2.5 PV for each city over a six-year period was arranged and segmented. The cumulative frequency of each segment was calculated. The interval with the highest regression coefficient of determination (R^2) > 95% of the relative cumulative frequency curve was then selected [16]. The average value of all the recorded data within this interval was then calculated as the contamination baseline.

A two-dimensional space can be constructed using the MRE and CB of each city. The critical values of MRE and CB were used as the basis for space construction. The critical value of CB can be set at 35 μg/m^3, which aligns with the Chinese Grade II standard (GB 3095-2012). Since there is no standard value for MRE and the average MRE of all cities over a six-year period is less than 0, we have taken the average value of MRE (–0.41 μg/m^3) as the critical value. The MRE higher than the average value has a relatively poor effect on reducing PM2.5 concentration, while the MRE lower than the average value has a relatively good effect. According to the two-dimensional space, cities in the Yangtze River Economic Belt will be divided into four categories: Q1: CB > 35 μg/m^3, MRE > –0.41 μg/m^3; Q2: CB > 35 μg/m^3, MRE > –0.41 μg/m^3; Q3: CB > 35 μg/m^3, MRE > –0.41 μg/m^3; Q4: CB > 35μg/m^3, MRE > –0.41 μg/m^3. The four categories represent the combined effects of anthropogenic emissions and meteorological conditions on urban PM2.5.

3 RESULTS AND DISCUSSION

The R^2-training set and R^2-test set were used to evaluate the fit and overfitting of the WN model. The R^2-training set ranged from 0.8 to 0.96, and the R^2-test set ranged from 0.38 to 0.77. The R^2-test set in the most cities is higher than 0.5 and shows no apparent correlation with the R^2-training set. The relationship between PM2.5 and meteorological factors in most cities is well explained by the RF model. The midstream cities had a relatively high degree of fit in both the training and test sets. In these cities, PM2.5 had obvious anthropogenic sources, and the changeable weather had a significant impact on aggregation and diffusion [11]. On the other hand, the locations of meteorological and PM2.5 monitoring stations may occasionally result in extreme values that can impact the outcomes of machine learning [12]. When calculating CP, the selection of the confidence interval for cumulative frequency helps to further avoid the influence of outliers and ensure the availability of data.

The statistics of OV, PV, MRE, and CB are shown in Figure 2. In all cities, the PM2.5 levels were lower than the OV in the winter and higher in the summer. However, the winter values still remained slightly higher overall. The 6-year annual average value of PV is 41.05 μg/m^3, and of MRE is –0.41 μg/m^3. The average value of CB is 34.05 μg/m^3. CB in 72 cities is lower than the Chinese Grade II standard (GB 3095-2012). The additional emissions by humans resulted in an increase of 7μg/m^3 in concentration, while the meteorological factors caused a decrease of –0.41 μg/m^3. CB is close to PV in 2020. One reason could be the ongoing decline in emissions. PV declined from 48.14 μg/m^3 to 41.05 μg/m^3 over a span of 6 years. The other reason might be related to the COVID lockdown in the first half of 2020. In Liu et al. (2022), PM2.5 was detected during the lockdown period of the COVID-19 epidemic, with fewer anthropogenic emission sources. MRE is in a fluctuating state. From 2015 to 2019, the MRE ranged from –0.95 μg/m^3 to 0.78 μg/m^3. Whereas in 2020, MRE descended to –3.25 μg/m^3. The relatively low PM2.5 concentration in 2020 is due to a combination of reduced anthropogenic emissions and favorable meteorological conditions that facilitate settling or dispersion.

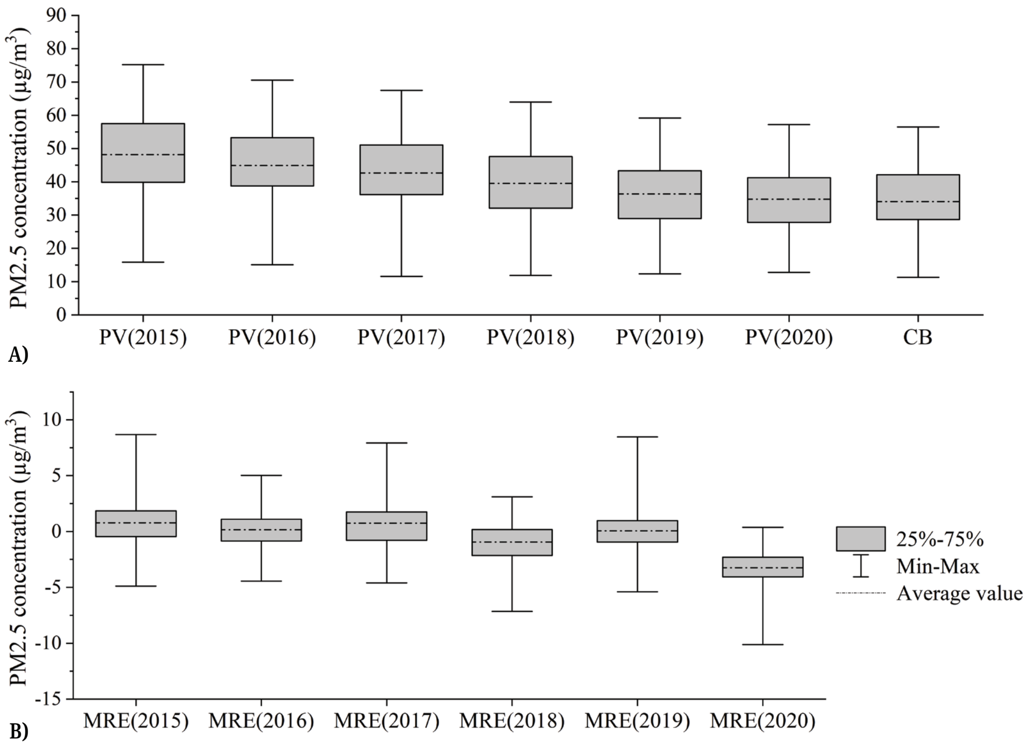

Fig. 2. Statistics of each year PV (A), each year MRE (B) and CB (A)

Figure 3A and B show the spatial distribution of the annual average PM2.5 MRE and CB. From the spatial distribution of MRE, it can be observed that meteorological conditions in the downstream region have a relatively effective impact on reducing PM2.5 concentrations. The meteorological conditions in the midstream region are relatively weak, and there is a concentration of cities with high MRE values (Ezhou: –0.11; Zhuzhou: –0.16; Yueyang: –0.06; Jingzhou: 0.07). The meteorological conditions in the upstream region are generally weaker than those in the downstream but better than those in the midstream. Some cities with higher MRE are dispersed, such as Lijiang (–0.10) and Liupanshui (0.32). CB exhibits clear spatial aggregation. The cities with high CB are clustered in the northeast of the Yangtze River Economic Belt and the Chengdu-Chongqing Plain. These cities have relatively high anthropogenic emissions or are chronically affected by pollutants from the north [26]. The cities with a relatively low CB are concentrated in the upstream area, which is associated with relatively low intensity of human activities. Areas near the coast with high human activity intensity but low average CB may have a lower MRE.

The spatial distribution of the four types of cities in the Yangtze River Economic Belt is shown in Figure 3C. The number of Q1 cities is 22, mainly distributed in the midstream of the Yangtze River and the Chengdu-Chongqing city clusters. The number of Q2 cities is 31, most of which are distributed in the downstream region. Even though meteorological conditions have a relatively positive impact on air quality in Q2 cities, the levels of PM2.5 remain high in these areas, primarily due to man-made emissions. There are 70 cities in the Q3 and Q4 categories whose CB is less than 35 μg/m^3. The number of Q3 cities is 50. They are mainly distributed in the mid- and upstream areas. Even if the MRE is relatively weak, the observed PM2.5 concentration can still easily reach the standard value. There are 20 Q4 cities evenly distributed throughout the Yangtze River Economic Belt, including Shanghai and other major cities. Such cities have both better meteorological conditions and more reasonable anthropogenic emissions.

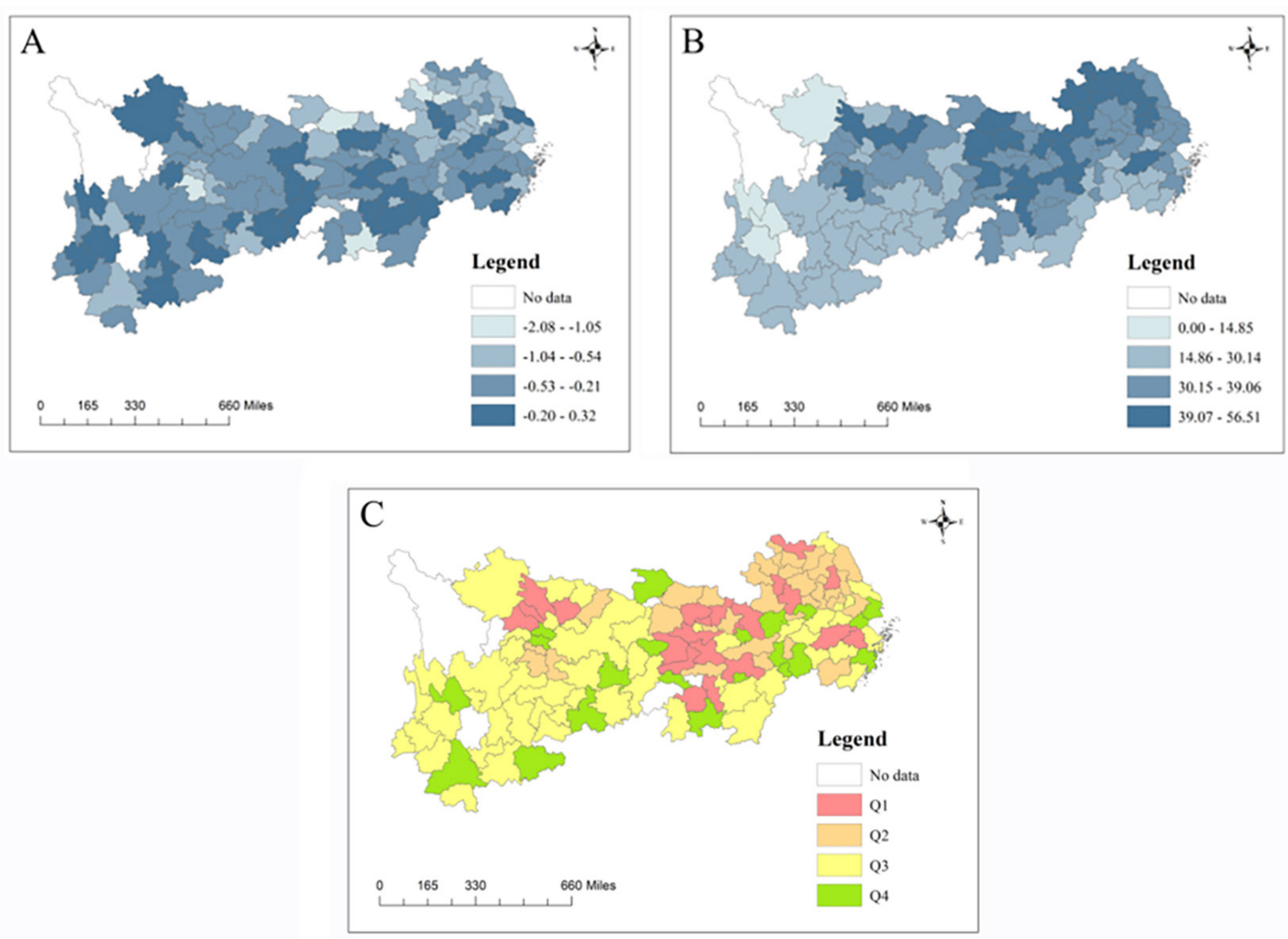

Fig. 3. Spatial distribution of MRE (A, μg/m^3), CB (B, μg/m^3), classification (C) in each city. Q1: CB >35 μg/m^3, MRE > –0.41 μg/m^3; Q2: CB >35 μg/m^3, MRE > –0.41 μg/m^3; Q3: CB >35 μg/m^3, MRE > –0.41 μg/m^3; Q4: CB >35 μg/m^3, MRE > –0.41 μg/m^3

4 CONCLUSION

Based on machine learning and the cumulative frequency curve method, this research determined the MRE and CB values of PM2.5 in cities located in the Yangtze River Economic Belt. Cities are classified by setting thresholds for MRE and CB. The average value of CB is 34.05 $\mu g/m^3$, and the average value of MRE is –0.41 $\mu g/m^3$. The anthropogenic emissions have reached Chinese Grade II standards in most cities. The concentration of CB was higher in the middle and downstream areas of the Yangtze River and the Chengdu-Chongqing urban agglomeration. MRE has a slight decreasing effect on PM2.5, and the effect is greater downstream > upstream > midstream. Among the four types of cities, Q1 (with relatively high emissions and poor meteorological conditions) is concentrated in the midstream, Q2 (with relatively high emissions and good meteorological conditions) is concentrated in the downstream, Q3 (with relatively low emissions and poor meteorological conditions) is concentrated in the upstream, and Q4 (with relatively low emissions and good meteorological conditions) is distributed discretely in space. The number of cities in Q3 is the largest. According to the classification results, the following suggestions can be made for controlling PM2.5 levels: (1) Controlling PM2.5 in Q1 cities is the most challenging. Transferring a few high-emission industries to Q3/4 cities can help reduce pollution. (2) Q2 cities should capitalize on favorable meteorological conditions to effectively reduce pollution through emission control and industrial upgrading. In addition, the north and south coastal areas can collaborate to balance emissions. (3) Q3 cities can accept a portion of industrial transfer, but they need to pay attention to their relatively fragile environmental capacity. (4) Q4 cities suitable for economic development. Cooperation between cities can strive to promote green and high-quality development.

5 ACKNOWLEDGMENT

This study was supported by Hubei Provincial Outstanding Young Science and Technology Innovation Team Project (T2021032) and the Fundamental Research Funds for the Central Universities, Zhongnan University of Economics and Law (2722023EZ009, 202351416).

6 REFERENCES

[1] S. L. Feng, D. Gao, F. Liao, *et al.*, "The health effects of ambient PM 2.5 and potential mechanisms," *Ecotoxicology and Environmental Safety*, vol. 128, pp. 67–74, 2016. https://doi.org/10.1016/j.ecoenv.2016.01.030

[2] J. Y. Guo, F. Li, Z. G. Qu, *et al.*, "Quantitative evaluation of PM2.5-related health economic losses and analysis of their driving factors in Chinese cities," *Frontiers in Environmental Science*, vol. 10, p. 951505, 2022. https://doi.org/10.3389/fenvs.2022.951505

[3] X. M. Luo, K. Sun, L. Li, *et al.*, "Impacts of urbanization process on PM2.5 pollution in '2+26' cities," *Journal of Cleaner Production*, vol. 284, p. 124761, 2021. https://doi.org/10.1016/j.jclepro.2020.124761

[4] X. Y. Chen, F. Li, J. D. Zhang, *et al.*, "Spatio-temporal mapping and multiple driving forces identifying of PM2.5 variation and its joint management strategies across China," *Journal of Cleaner Production*, vol. 250, p. 119534, 2020. https://doi.org/10.1016/j.jclepro.2019.119534

[5] K. Gui, H. Z. Che, Y. Q. Wang, *et al.*, "Satellite-derived PM2.5 concentration trends over Eastern China from 1998 to 2016: Relationships to emissions and meteorological parameters," *Environmental Pollution*, vol. 247, pp. 1125–1133, 2019. https://doi.org/10.1016/j.envpol.2019.01.056

[6] X. Z. Zhang, X. D. Xu, Y. H. Ding, *et al.*, "Impact of changes in meteorological conditions on the decrease of PM2.5 mass concentration in key regions in China from 2013 to 2017," *Chinese Science: Earth Science*, vol. 50, no. 4, pp. 483–500, 2020.

[7] Y. M. Liu, Y. Y. Hong, Q. Fan, *et al.*, "Source-receptor relationships for PM2.5 during typical pollution episodes in the Pearl River Delta city cluster, China," *Science of the Total Environment*, vol. 596, pp. 194–206, 2017. https://doi.org/10.1016/j.scitotenv.2017.03.255

[8] J. Luo, P. J. Du, A. Samat, *et al.*, "Spatiotemporal pattern of PM2.5 concentrations in Mainland China and analysis of its influencing factors using geographically weighted regression," *Science Report*, vol. 7, p. 40607, 2017. https://doi.org/10.1038/srep40607

[9] S. K. Grange and D. C. Carslaw, "Using meteorological normalisation to detect interventions in air quality time series," *Science of the Total Environment*, vol. 653, pp. 578–588, 2019. https://doi.org/10.1016/j.scitotenv.2018.10.344

[10] Z. X. Zou, C. X. Cheng, and S. Shen, "The complex nonlinear coupling causal patterns between PM2.5 and meteorological factors in Tibetan Plateau: A case study in Xining," *IEEE Access*, vol. 9, pp. 150373–150382, 2021. https://doi.org/10.1109/ACCESS.2021.3123455

[11] H. W. Liu, F. G. Yue, and Z. Q. Xie, "Quantify the role of anthropogenic emission and meteorology on air pollution using machine learning approach: A case study of PM2.5 during the COVID-19 outbreak in Hubei Province, China," *Environmental Pollution*, vol. 300, p. 118932, 2022. https://doi.org/10.1016/j.envpol.2022.118932

[12] S. K. Grange, D. C. Carslaw, A. C. Lewis, *et al.*, "Random forest meteorological normalisation models for Swiss PM10 trend analysis," *Atmospheric Chemistry and Physics*, vol. 18, no. 9, pp. 6223–6239, 2018. https://doi.org/10.5194/acp-18-6223-2018

[13] H. B. Xian, X. H. Dong, Y. Wang, *et al.*, "Geochemical baseline establishment and pollution assessment of heavy metals in the largest coastal lagoon (Pinqing Lagoon) in China mainland," *Marine Pollution Bulletin*, vol. 177, p. 113459, 2022. https://doi.org/10.1016/j.marpolbul.2022.113459

[14] X. R. Zhao, T. Nasier, Y. Y. Cheng, *et al.*, "Environmental geochemical baseline of heavy metals in soils of the Ili river basin and pollution evaluation," *Environmental Science*, vol. 35, no. 6, pp. 2392–2400, 2014.

[15] L. Fan, J. Wang, Y. Y. Yang, *et al.*, "Baseline investigation on residential PM2.5 pollution of general living scenarios-12 cities, China, 2018," *China CDC Weekly*, vol. 2, no. 32, pp. 609–613, 2020. https://doi.org/10.46234/ccdcw2020.165

[16] D. D. Jin, S. J. Kong, C. H. Ou, *et al.*, "The provincial baseline of PM2.5 in China and its hierarchical management strategy," *The Provincial Baseline of PM2.5 in China and Its Hierarchical Management Strategy*, vol. 10, p. 908760, 2022. https://doi.org/10.3389/fpubh.2022.908760

[17] C. Emery, Z. Liu, A. G. Russell, *et al.*, "Recommendations on statistics and benchmarks to assess photochemical model performance," *Journal of the Air & Waste Management Association*, vol. 67, no. 5, pp. 582–598, 2017. https://doi.org/10.1080/10962247.2016.1265027

[18] X. J. Liu, S. Y. Xia, Yu. Yang, *et al.*, "Spatio-temporal dynamics and impacts of socioeconomic and natural conditions on PM2.5 in the Yangtze River Economic Belt," *Environmental Pollution*, vol. 263(Pt A), p. 114569, 2020. https://doi.org/10.1016/j.envpol.2020.114569

[19] Z. Y. Chen, D. L. Chen, D. F. Zhao, *et al.*, "Influence of meteorological conditions on PM2.5 concentrations across China: A review of methodology and mechanism," *Environment International*, vol. 139(C), no. 21, 2020. https://doi.org/10.1016/j.envint.2020.105558

[20] S. K. Grange, “Rmweather: Tools to conduct meteorological normalisation on air quality data,” 2020. https://cran.r-project.org/web/packages/rmweather/rmweather.pdf
[21] J. Tigges, T. Lakes, and P. Hostert, “Urban vegetation classification: Benefits of multitemporal RapidEye satellite data,” *Remote Sensing of Environment*, vol. 136, pp. 66–75, 2013. https://doi.org/10.1016/j.rse.2013.05.001
[22] M. D. Mallet, “Meteorological normalization of PM10 using machine learning reveals distinct increases of nearby source emissions in the Australian mining town of Moranbah,” *Atmospheric Pollution Research*, vol. 12, pp. 23–35, 2021. https://doi.org/10.1016/j.apr.2020.08.001
[23] H. Xuan, Y. G. Teng, S. J. Ni, *et al.*, “Potential ecological risk assessment on heavy metal in the soil of indexing area based on the geochemical baseline,” *Mineral Rock*, vol. 4, pp. 69–72, 2005.
[24] K. Fan, C. Y. Wei, and X. S. Yang, “Determination and application of geochemical baseline values of soil heavy metals in Qiaokou Town, Changsha City,” *Journal of Environmental Science*, vol. 34, no. 12, pp. 3076–3083, 2014.
[25] S. H. Wang, W. W. Wang, J. Y. Chen, *et al.*, “Geochemical baseline establishment and pollution source determination of heavy metals in lake sediments: A case study in Lihu Lake, China.” *Science of the Total Environment*, vol. 657, pp. 978–986, 2018. https://doi.org/10.1016/j.scitotenv.2018.12.098
[26] C. H. Ou, F. Li, J. D. Zhang, *et al.*, “Multiple driving factors and hierarchical management of PM2.5: Evidence from Chinese central urban agglomerations using machine learning model and GTWR,” *Urban Climate*, vol. 46, p. 101327, 2022. https://doi.org/10.1016/j.uclim.2022.101327

7 AUTHORS

Changhong Ou is member of Research Center for Environment and Health, Zhongnan University of Economics and Law (ZUEL), Wuhan, China. He is Ph.D candidate in ZUEL. His research interests includes Internet of Things, air pollution and smart health management. He has published a total of six papers (E-mail: och@stu.zuel.edu.cn).

Fei Li is member of Research Center for Environment and Health, ZUEL. He is a Professor and Doctoral supervisor, the Ministry of Science and Technology recommended experts in the field of causes and treatment of soil pollution, director of Hubei Environmental Science Society and Hubei System Engineering Society, etc. Main research directions: (1) Intelligent management of regional environmental risks; (2) Ecological environment big data; (3) Environmental health management; (4) System uncertainty control (E-mail: lifei@zuel.edu.cn).

Jingdong Zhang is member of Research Center for Environment and Health, ZUEL. She is a Professor and Doctoral Supervisor, Executive Director of Hubei Environmental Science Society and Wuhan Environmental Science Society, director of Research Center for Environment and Health in ZUEL. Main research directions: (1) Environmental and occupational health risks; (2) Environmental safety management; (3) Environmental pollution control (E-mail: jdzhang@zuel.edu.cn).

Jinyuan Guo is member of Research Center for Environment and Health, ZUEL. She holds a Master’s degree from ZUEL. Her research interests includes air pollution and public health economy. She has published six papers (E-mail: jinyuanguo@stu.zuel.edu.cn).

Xiyao Chen is member of Research Center for Environment and Health, ZUEL. He holds a Master’s degree from ZUEL. His research interests includes PM2.5,

meteorological and emission source analysis. He has published a total of 11 papers (E-mail: 12138042@zju.edu.cn).

Shaojie Kong is member of Research Center for Environment and Health, ZUEL. He holds a master's degree from ZUEL. His research interests inclludes synergistic management of ozone, PM2.5 and carbon emissions. He has published seven papers (E-mail: ksj@stu.zuel.edu.cn).

Pei Jiang is member of Research Center for Environment and Health, ZUEL. She holds a master's degree from ZUEL. Her research interests includes environmental pollution, big data and artificial intelligence (E-mail: jiangpei@zuel.edu.cn).

Mian Wu is member of Research Center for Environment and Health, ZUEL. She is Master-student in ZUEL. Her research interest lies in accurate assessment of heavy metal health risks (E-mail: 202111200011@stu.zuel.edu.cn).

Yazhu Wang is member of Research Center for Environment and Health, ZUEL. She holds a Master's degree in Public Administration in ZUEL (E-mail: yazhuwang@stu.zuel.edu.cn).