ANALYSIS AND PREDICTION OF PM2.5 CONCENTRATION IN GUILIN CITY

ZHOU Zhi-hao^{1,2,*}, XIE Shao-feng^{1,2}, LI Guo-hong^{1,2}, ZHAO yun^{1,2}, ZHANG wei^{1,2}

¹Guangxi Key Laboratory of Spatial Information and Geomatics, 541006 Guilin, China ²College of Geomatics and Geoinformation, Guilin University of Technology, 541006 Guilin, China;

KEY WORDS: PM2.5, ZTD, Meteorological factors, BP neural network, Gray correlation

ABSTRACT:

Atmospheric particulate matter is one of the most harmful pollutants in the atmospheric environment, and has an important impact on climate change, reduced visibility, environmental hazards (such as acid rain, smoke) and health hazards. PM2.5 is the main factor causing haze weather, reducing visibility and affecting traffic safety. PM2.5 enters the alveoli through the respiratory tract and endangers human health. The basic characteristics of PM2.5 are small size, light weight, long residence time in the atmosphere, and can be transported to a long distance by the atmospheric circulation, causing a wide range of air pollution. In 2018, the proportion of days with excellent and good air quality in Guilin City in January and February was 71.2%, ranking 10th among the 14 cities in the Guangxi Autonomous Region. As a famous tourist city, its air quality should be paid more attention. The main air pollutant in Guilin City is PM_{2.5}, which refers to fine particles with a diameter of 2.5 µm or less in the air, it has great harm to human health and reduces the visibility of the atmosphere. Air quality data, meteorological data, and GPS tropospheric delay data (ZTD: zenith tropospheric delay) were collected to calculate the correlation between PM2.5 concentration and influencing factors by the gray correlation model, and the relationship between the main influencing factors and the variation characteristics of PM2.5 concentration was analyzed. The result shows that PM2.5 has strong correlation with other air pollutants SO2 ,NO2 ,wind direction, relative humidity, ZTD, temperature and wind speed; moderately correlated with rainfall and pressure. The degree of gray correlation between the impact factor and PM2.5 varies with the seasons.Based on this research and analysis, air quality pollutants and meteorological factors are used as input factors and the concentration of PM2.5 values as the output factor of the neural network model. The model is trained by training sample data to establish a neural network model to predict the concentration value of PM2.5 in Guilin. The accuracy of the model is verified by the accuracy index. The summer neural network model has the best precision. The MAE is 4.51 μ g/m³, the RMSE is 5.73 μ g/m³, and the MRE is 15.1%. The correlation coefficient between the predicted value and the measured value reaches 0.908. It shows that the neural network model based on meteorological factors and ZTD has a good predictive effect and has certain guiding significance for the prevention and control of air pollution in Guilin.

1. INTRODUCTION

In the process of human activities, some harmful gases are emitted into the air, which has a serious impact on the atmospheric environment in which people live. Air pollution will become more severe when it is in unsatisfactory weather conditions (JIANG 2015), which has a negative impact on people's living and environment. The Air Quality Index (AQI) is a dimensionless index that quantitatively describes air quality conditions (CAEPI 2012). The air quality assessment indicator consists of 6 major pollutants, including $PM_{2.5}$, PM_{10} , SO₂, NO₂, O₃ and CO, and ranking 6 grades, excellent

Corresponding author: Zhou Zhihao - E-mail: 15387594769@163.com;

(I 0~50), good (II 51~100), mild pollution (III 101 ~150), moderate pollution (IV 151~ 200), heavy pollution (V 201~300), serious pollution (VI 301~500). The primary air pollutant is $PM_{2.5}$, which can reduce visibility and affect human health, in high level (SONG 2013). $PM_{2.5}$, called fine particles (diameter $\leq 2.5 \mu$ m), can be suspended in air for a long time to move. Meanwhile, it has small size, strong activity, carrying toxic and harmful substances, and has great impact on the environment (CAI 2012). In 2018, $PM_{2.5}$ accounted for 71.1% of the non-excellent air quality primary pollutants in Guilin, and PM_{10} was 27.2%. There was a significant linear relationship between $PM_{2.5}$ and PM_{10} , and its correlation coefficient reached 0.904. The correlation coefficient between AQI and $PM_{2.5}$ is 0.980, which indicates that the air status of Guilin City is mainly affected by the $PM_{2.5}$.

Studies have shown that meteorological conditions have a significant impact on PM2.5. Air quality pollutants are affected by meteorological conditions and topographic conditions such as wind direction, wind speed, air temperature distribution, and atmospheric stability. For example, these scholars found that meteorological factors such as temperature, relative humidity and precipitation have a close correlation with the daily mean value of PM2.5 concentration in a certain place in Zhejiang Province (DANG 2019). The concentration of PM2.5 in Shaoguan City has a negative correlation with temperature, humidity and wind speed (LIANG 2019). Some scholars pointed out that GPS technology can be applied to the monitoring of PM2.5. A correlation study between GNSS ZTD and PM2.5 concentration was carried, based on wavelet transform method, and the results show that: North China, The correlation between GNSS ZTD and PM25 concentrations in Central China, Northwest China and Northeast China was significantly positively correlated (WANG 2018). Therefore, this paper combines meteorological data, tropospheric delay products, and air pollution data to analyze and predict the PM2.5 concentration characteristics.

2. RESEARCH DATA

Guilin, the earliest cultural and educational center in Guangxi, is one of the best tourist cities in China. It is located in the northeastern part of Guangxi Autonomous Region, between the east longitude 109°36′50″-111°29′30″, north latitude 24°15′23″-26°23′30″. Guilin lie in the southwestern part of the Nanning Mountains and in the middle of the northern end of the Guilin-Yangshuo karst basin. It lie in the "Hunan-Guizhou province Road".The west, north and southeast are high and the middle is low, with medium or low mountains. It has a tropical monsoon climate, with mild temperature, abundant rainfall, sufficient sunshine, rich heat. Guilin is long in summer and short in winter, distinct season, and its rainy season and hot season are the same.

The experimental data mainly includes air quality data, meteorological data, and GPS tropospheric data. The relationship between $PM_{2.5}$ concentration and meteorological factors and tropospheric data was analyzed by gray correlation method, and then the neural network model was established to predict the concentration of $PM_{2.5}$ in Guilin.

①Air quality data: The data center of the Environmental Protection Department of Guangxi provides historical data of various pollutant concentration values hour by hour;

⁽²⁾ Meteorological data: Meteorological hourly data such as air pressure, temperature, rainfall, wind speed, etc. are provided by the National Meteorological Data Center;

③ ZTD data: The China Earthquake Administration GNSS data product service platform



Fig.1 Distribution of Guilin Air Monitoring Station

3. ANALYSIS OF PM_{2.5} CONCENTRATION

3.1 CHARACTERISTICS OF DAILY AND SEASONAL PM_{2.5} VARIATION

The PM_{2.5} concentration level is affected by pollution source discharge and meteorological conditions, and there are obvious seasonal changes. Therefore, the 2018 PM_{2.5} data of Guilin City is plotted as a 24-hour daily average value by season. From Fig. 1(a), the daily variation of the PM_{2.5} concentration value in Guilin City is roughly V-shaped, reaching the maximum value of the day at 24 o'clock in the middle of the night, and the trough value appears from 15:00 pm to 16:00 pm, and there is a section. The easing period is from 3:00 am

to 10:00 am. The summer change in Fig. 1(b) shows the characteristics of double-peak and double-valley. The peaks appear at 9:00 am and 21:00 pm, and the troughs are at 3:00 am and 17:00 pm, and the change in $PM_{2.5}$ concentration in the second half of the day is required. More intense than the previous half. The change in Fig. 1(c) in autumn is similar to that in summer. The peak appears at 8:00 am and 20:00 pm, and the trough is at 1600 pm. In winter, Figure 1(d) shows the peak value of the daily variation of the $PM_{2.5}$ concentration value at 1 o'clock in the middle of the night, and continues to drop to 8:00 am, and a small recovery phase from 8:00 am to 10:00 am, and then decreases to 17:00 pm in the afternoon. The lowest value.

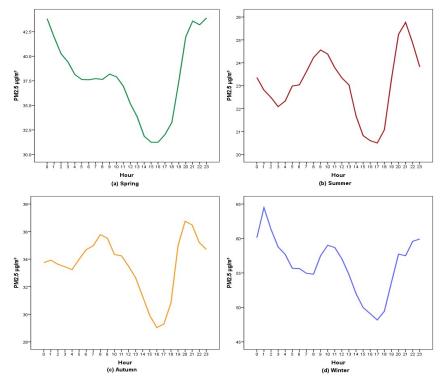


Fig. 2 Seasonal diurnal variation of PM_{2.5} concentration

Seasonal mean values of $PM_{2.5}$ concentration in Guilin were 37.74 µg/m³ (spring Mar~May), 23.07 µg/m³ (summer Jun~Aug), 33.58 µg/m³ (autumn Sept~Nov), 56.16 µg/m³ (winter Dec~Feb), showing winter>spring>autumn>summer. It can also be seen from Fig. 2 that the fluctuations in winter are large and the fluctuations in summer are small, which indicates seasonal variation characteristics. In particular, the

air quality on February 16 has been seriously polluted. The pollutants PM_{10} and $PM_{2.5}$ are seriously over-standard, and the concentration is abnormally high. At 1:00 am, the value of $PM_{2.5}$ concentration reaches 504 µg/m³. It can be found that it was due to the frequent human activities of the Lunar New Year on February 16th. There were much fireworks and firecrackers in the Spring Festival, and the temperature

increased during the Spring Festival, and the surface wind speed was small. Atmospheric conditions are not conducive to the spread of air pollutants. This explains to a certain extent that the concentration of pollutant PM2.5 is affected by the

Spring Festival. After counting the distribution of air quality levels in different seasons, draw it into Table 1. From Tab.1, it

can seen that the air quality in summer is the best, and most of the time is in excellent air quality. The air quality situation of winter is the worst, and grade I accounting for only 37.3% of the whole quarter, is far less than other season's, and exists heavy and serious pollution.

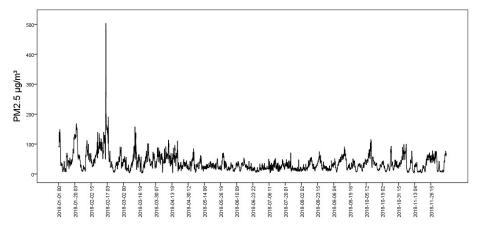


Fig.3 Variation of PM_{2.5} concentration in Guilin city in 2018

Tab.1 Air quality level distribution

Season	Ι	Π	III	IV	V	VI
Spring	46.9%	45.8%	6.8%	0.5%	0.1%	
Summer	78.6%	21.3%	0.1%			
Autumn	51.0%	45.2%	3.8%			
Winter	37.3%	37.7%	13.7%	8.0%	3.1%	0.2%

(Note : I excellent air quality; II good; III mild pollution; IV moderate pollution; V heavy pollution; VI serious pollution;)

3.2 GRAY CORRELATION ANALYSIS BETWEEN PM_{2.5} METEOROLOGICAL FACTORS, ZTD

The data of $PM_{2.5}$, meteorological factors and tropospheric delay in different seasons are plotted in Table 2. Table 2 shows that Guilin has high temperature and rain in summer, low air pressure and large tropospheric delay; low temperature and low rainfall in winter, high pressure and low total tropospheric delay. The concentration of $PM_{2.5}$ in atmospheric pollutants is closely related to meteorological factors. The gray correlation degree between $PM_{2.5}$ concentration value and meteorological factors and other pollutant concentration values in different seasons is calculated as Table 3. The gray correlation degree is used to describe the measure of the correlation between

different elements over time or different objects. In the process of system development, if the trend of the two factors changes, that is, the degree of synchronization changes is high, then it can be said that The degree of association is higher; on the contrary, it is lower. It can be seen from Tab.3 that the gray correlation degree of pollutants SO₂, NO₂ and PM_{2.5} is strong and varies with the seasons. NO₂, surface temperature and relative humidity have the most significant influence on PM_{2.5} concentration in spring; NO₂, pressure, ZTD and surface temperature have great influence on summer PM_{2.5} concentration value; PM_{2.5} concentration value and various factors in autumn The correlation degree shows less obvious effect, mainly relative humidity and air pressure; winter PM_{2.5} has obvious correlation with SO₂ and ZTD. The effect of temperature on $PM_{2.5}$ concentration is mainly reflected in the fact that when the temperature is high, the vertical turbulence of the atmosphere is intensified, which is conducive to the diffusion of pollutants. When the inversion layer appears in the troposphere, it will hinder the diffusion movement of $PM_{2.5}$ in the vertical direction, so that the pollutants accumulate in the atmosphere of the near-surface layer and thus the degree of air pollution is aggravated. Therefore, the influence of

temperature on the variation of $PM_{2.5}$ concentration value is significant. The average temperature in winter in Tab.2 is 9.7 °C, which is the most prone to inversion layer phenomenon, resulting in high pollution of air. The average summer temperature is the highest, reaches 28.4 °C. Therefore, the intensity of atmospheric turbulence is large, and the atmospheric pollutants have good diffusion conditions, which are mostly excellent and good air quality.

Season	Temperature/ (°C)	Wind speed/ (ms ⁻¹⁾	Relative humidity/(%)	Rain/ (mmh ⁻¹)	Air pressure /(hpa)	ZTD/(mm)
Spring	21.8	1.81	74.8	0.34	993.6	2510.5
Summer	28.4	1.73	74.3	0.27	984.5	2613.1
Autumn	21.0	2.07	67.9	0.17	997.1	2511.2
Winter	9.7	2.47	64.5	0.05	1002.5	2425.1

Tab.2 Meteorological factors, tropospheric delay seasonal mean value

Tab.3 Gray correlation between PM2.5 and meteorological factors and tropospheric delay

Season	SO_2	NO ₂	СО	Wind	Tempera-	Relative	Rain	Air	ZTD
	50_2	NO ₂	0	speed	ture	humidity		pressure	
Spring	0.784	0.904	0.501	0.532	0.878	0.852	0.523	0.631	0.535
Summer	0.777	0.960	0.512	0.591	0.903	0.728	0.517	0.912	0.971
Autumn	0.558	0.608	0.503	0.550	0.646	0.769	0.505	0.735	0.549
Winter	0.931	0.643	0.504	0.517	0.530	0.662	0.500	0.513	0.829

4. RESULTS AND ANALYSIS

From the gray correlation scale, $PM_{2.5}$ has a high correlation with pollutants SO₂, NO₂, surface temperature, relative humidity and gas pressure. Based on the analysis of the correlation and trend graph between pollutants, meteorological factors and $PM_{2.5}$, they are modeled by neural networks and the correlation model between them is established. Through the gray correlation degree between $PM_{2.5}$ and other factors, multiple independent variables were selected to predict the $PM_{2.5}$ concentration value. For the $PM_{2.5}$ concentration values in different seasons as the dependent variable, the factor with the large gray correlation degree was used as the independent variable, and the concentration value of $PM_{2.5}$ in Guilin was forecasted for 7 days. Figure 3(a) is the comparison between the predicted value and the measured value of the spring model. The solid blue line is the measured $PM_{2.5}$ concentration value, and the green dotted line is the predicted value of the model. It can be seen that the overall trend of the two is consistent, and The fluctuation frequency of the predicted value is larger than the measured value, and the value is also larger than the measured value. The measured value is generally stable for a certain period of time, that is, the sensitivity to meteorological factors is not large, and will not immediately increase or decrease with the change of a certain meteorological factor. The predicted value in Fig. 3(b) of the summer model is consistent with the trend of the measured value as a whole, but there is a big difference between the two on August 13th, and the concentration value of the measured line fluctuates around 50 μ g/m³. The predicted value line fluctuates around 40 μ g/m³, and the difference between the two values is up to 15 μ g/m³. In Fig. 3(c), the difference between the predicted value and the measured value of the autumn model is the largest at 15 o'clock on November 13, the predicted value is obviously larger, and the measured value is smaller. The predicted value of PM_{2.5} is higher than the measured value because the concentration value of the

independent variable O_3 is at an abnormally high concentration value of 123 µg/m³ during this period, and the adjacent time period is only 40-50 µg/m³. May be affected by artificial emission source. The measured value of the winter model in Figure 3(d) is obviously higher than the predicted value on February 16th, mainly because during the period of the Lunar New Year, human activities are more frequent, and fireworks and firecrackers are more frequently emitted, resulting in PM_{2.5} concentration. The value increased sharply, indicating that the PM_{2.5} concentration value was affected by human activities.

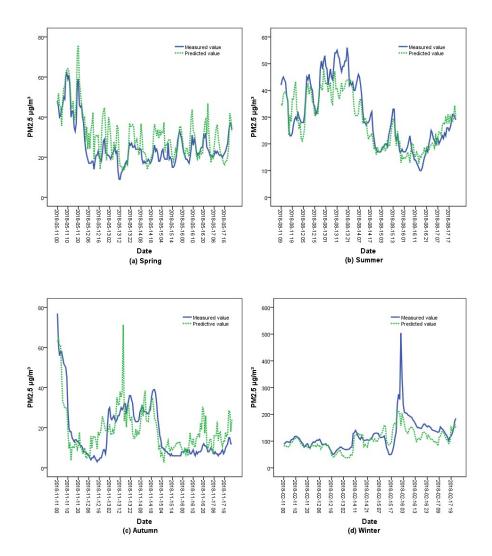


Fig.3 Comparison of measured and predicted values, based on BP model

The accuracy prediction of the seasonal prediction model uses three kinds of test indicators: mean absolute error (MAE), root mean square error (RMSE), and average relative error (MRE). MAE refers to the absolute deviation between the predicted value and the measured value. MAE can effectively avoid the problem that the deviation values cancel each other out, so it can accurately reflect the actual accuracy of the predicted value. RMSE is the square root of the ratio of the square of the predicted value to the true value and the number of observations. The RMSE is used to measure the deviation between the observed value and the true value, which can well reflect the precision of the predicted value. MRE refers to the average value of the relative error. The relative error represents the ratio of the absolute error to the measured value. It is usually expressed as a percentage. The MRE can well reflect the credibility of the interpolation model. At the same time, the correlation coefficient between the predicted value and the measured value is calculated, and the accuracy of the statistical neural network model is calculated. The statistical results are shown in Table 5. Among the accuracy indicators, the summer MAE, RMSE, and MRE are the smallest, and the correlation coefficient is the highest, indicating that the summer prediction model is higher than other seasons. The average relative error in autumn is the largest. After the study, it is found that O₃ is in an abnormally high concentration range during this period, which leads to the predicted value of PM2.5 being much larger than the measured value, and the concentration of PM2.5 itself is not large during this period. In the end, the MRE of the autumn prediction model is the largest in the four seasons. O3's abnormal fluctuations are related to human activities, and the main force causing ozone pollution is "human emitting sources": primary pollutants emitted from coal, motor vehicle exhaust, petrochemicals, etc. At the same time, under the irradiation of ultraviolet rays, nitrogen oxides will generate ozone and other pollutants, which is also called "secondary photochemical reaction". The MAE and RMSE of the winter model were significantly higher than other seasons, and the analysis was found to be affected by the Spring Festival holiday. The PM2.5 concentration values on February 15 and 16 continued to be in the high concentration range, reaching 504 μ g/m³ at 0:00 on the 16th, which was a serious pollution category. Although the concentration values of other contaminants also increased during this time period, the performance was not as pronounced as PM2.5, resulting in a predicted PM2.5 concentration value that was much lower than the measured value. The correlation coefficient between the predicted and measured values of the winter model is also at the lowest level, and the overall accuracy is worse than that of the spring, summer and autumn models.

Season	MAE (µg/m ³)	RMSE (µg/m ³)	MRE (%)	Correlation coefficient
Spring	6.41	8.22	28.8	0.848
Summer	4.51	5.73	15.1	0.908
Autumn	7.14	9.02	63.3	0.768
Winter	26.62	42.7	21.5	0.718

Tab.5 Neural network model accuracy verification

5. CONCLUSION

From the above research and analysis, it can be seen that the variation of $PM_{2.5}$ concentration value and meteorological factors in Guilin City show obvious seasonal distribution

characteristics. The seasonal characteristics of $PM_{2.5}$ are winter (56.16 µg/m³)> spring (37.74 µg/m³). > Autumn (33.58 µg/m³)> Summer (23.07 µg/m³). The daily variation of $PM_{2.5}$

concentration in winter is roughly V-shaped, and the daily variation of $PM_{2.5}$ in summer is characterized by double-peak and double-valley. The results of gray correlation analysis of $PM_{2.5}$, meteorological factors and tropospheric delay show that

PM_{2.5} has a strong correlation with SO₂ and NO₂, and the gray correlation degree between them fluctuates with the season. PM_{2.5} has a strong correlation with temperature, relative humidity, air pressure and ZTD, but its correlation with rainfall is weak. The change of PM2.5 concentration value in winter in Guilin City is affected by the Spring Festival holiday. The concentrated discharge of fireworks and firecrackers makes the concentration of PM2.5 and SO2 far greater than that during non-Chinese New Year. Based on the analysis of the correlation between PM2.5 and meteorological factors and tropospheric delay, a neural network prediction model is established and its accuracy is verified. The results show that the neural network model combined with meteorological factors and tropospheric delay has a good predictive effect on the PM_{2.5} concentration value in Guilin, and the correlation coefficient between the predicted value and the measured value has reached 0.7 or more. Among them, the summer neural network prediction model has the best precision, the MAE is 4.51 μ g/m³, the RMSE is 5.73 μ g/m³, and the MRE is 15.1%. The correlation coefficient between the predicted value and the measured value reaches 0.908. This study conducted an exploratory analysis of PM2.5 concentration data in Guilin, using gray correlation analysis and neural network to analyze and predict PM2.5 concentration, so as to explore the time distribution characteristics of Guilin PM2.5, providing corresponding support for air pollution prevention and control.

ACKNOWLEDGEMENTS

Foundation support: National Natural Science Foundation of China, No. 41864002, 41704027; Guangxi Natural Science Foundation, No. 2018GXNSFAA281182, 2017GXNSFBA198139; Open Fund of Guangxi Key Laboratory of Spatial Information and Geomatics, No. 15-140-07-11.

The experimental data comes from the Guangxi Environmental Protection Department, the National Meteorological Data Center, and the China Earthquake Administration. I would like to thank you!

REFERENCES

Jiang Yirong, Zhu Rong, Zhu Keyun, et al, 2015. Numerical simulation on the air pollution potential in the severe air pollution episodes in Beijing-Tianjin-Hebei Region [J]. Acta Scientiae Circumstantiae, 35(9): 2681-2692.

Ministry of Environmental Protection, 2012. HJ633—2012 Technical regulation on ambient air quality index (Trial) [S]. Beijing: China Environmental Science Press.

Song Ming, Han Suqin, Zhang Min, et al, 2013. Relationship between visibility and relative humidity, PM_{10} , $PM_{2.5}$ in Tianjin [J]. Journal of Meteorology and Environment, 29(2): 34-41

Cai Gang, 2012. Method for determination of PM_{2.5} in ambient air [J]. Gansu Science and Technology, 28(12): 45-46

Dang Congcong, Jiang Ping, 2019. Effects of Meteorological Conditions on PM_{2.5} Concentration in Atmosphere [J]. Contamination Control & Air-Conditioning Technology, (02):31-37.

LIANG Haifa LUO Yinghua FENG Haiqing, et al, 2019. Spatial and Temporal Distribution of $PM_{2.5}$ in Shaoguan City and Its Relationship with Meteorological Factors [J]. Journal of Shaoguan University, 40(06),76-81.

WANG Yong, REN Dong, HAO Zhenhang, et al, 2018. Study on the correlation between ZTD of CMONOC and PM_{2.5} concentration [J].Science of Surveying and Mapping, 43(05):40-44.

Tai A P K, Mickley L J, Jacob D J, 2010. Correlations between fine particulate matter ($PM_{2.5}$) and meteorological variables in the United States: Implications for the sensitivity of $PM_{2.5}$ to climate change [J]. Atmospheric Environment, 44 (32): 3976-3984.

Xie Shaofeng, Liang Chunli, Liu Zhongliu, et al, 2018. Correlative analysis of haze, troposphere delay and meteorological factors [J]. Science Technology and Engineering, 18(9): 184-188

Mao Ming, Wang Li, Zhang Shuangcheng, et al, 2018. Study on the influencing factors of summer haze in Shanghai [J]. Journal of Geodesy and Geodynamics, 38(7): 714-718

Wang Shan, Xiu Tianyang, Sun Yang, et al, 2014. The changes of mist and haze days and meteorological element during 1960-2012 in Xi'an [J]. Acta Scientiae Circumstantiae, 34(1): 19-26