Vol. 14 (2026): Proceedings of the 2nd International Conference on Artificial Intelligence, Modern Engineering and Environmental Sustainability (IC-AIMEES 2026)
E-ISSN: 3105-0522 P-ISSN：3105-0514
DOI: https://doi.org/10.70267/ ic-aimees.2026020107

Comprehensive Influence of Meteorology on Air Quality

Haoran Shuai*

School of Mathematical Sciences, Xiamen University, Xiamen, Fujian, 361000, China

**Corresponding author: Haoran Shuai*

Abstract

In view of the increasingly serious air pollution problem in China in recent years, especially the significant threat of PM2.5 to public health, this study focuses on the key driving factor of meteorological conditions, aiming to reveal the specific impact path of PM2.5 concentration in Beijing, and provide a solid theoretical basis and practical guidance for scientific formulation of provention and control strategies and effective response to climate change. Spearman correlation analysis, multiple linear regression and stepwise regression were used in the empirical analysis. The results showed that the air temperature was the most important meteorological factor affecting the concentration of PM PM. There was a significant negative correlation between the air temperature and the concentration of PM Yu. (Standardized coefficient Beta = -0. 691, $p < 0.05$), and the concentration of PM °C. Decreased by 0. 975 μg/m ³ when the air temperature increased by 1 °C. In addition, it is found that there is a serious multicollinearity problem among the meteorological factors, which may be the main reason why other factors except temperature are not significant in the regression model. The final regression model was significant ($F = 4.941$, $p = 0.002$). This study provides a quantitative basis for understanding the impact of meteorological conditions on air quality, and suggests that temperature should be used as a key indicator for short-term pollution warning, while providing a methodological reference for subsequent studies to overcome the multicollinearity problem.

Keywords

PMPM., meteorological factors, correlation analysis, multiple linear regression, stepwise regression

1. Introduction

Air Quality Is Related to People's Health and Urban Beauty. With China's Economic Growth, Air Quality Problems Are Becoming More and More Worthy of Attention. The Study Found That Using Big Data Analysis Method, Urban Greening Can Improve the Air, Plant Species Affect the Effect [1]. According to the Geographical Breakpoint Regression, Heating Is Still the Main Cause of Pollution. The Concentration of PM2.5 in the Heating Area Increased by 28.28%, but the Clean Heating Policy Reduced the AQI of the Pilot Cities by 10 Units [2]. Using Fixed Effect Model and Spatial Econometric Model, It Is Concluded That the Construction of Digital Countryside Purifies the Air, Especially in the East, and Can Also Drive the Surrounding Areas [3]. Using Sensitivity Tests, Climate Change Exacerbates Ozone Pollution. In the Summer of 2021, the Ozone Concentration in the Central and Eastern Regions Will Increase by 19.8 Mg/m ³ Due to Plant Emissions, and Climate Warming Will Make the Situation More Serious [4]. Using Pearson Analysis and Significance Test, Land Use Also Has an Impact on Air Quality, and Forest Land and Cement Land Affect Air Composition [5]. The Conclusions Drawn from the Above Literature Focus on the Impact of Human Activities on Air Quality.

In Meteorology, Most of the Current Research Only Looks at the Impact of a Single Weather Factor (such as Wind Speed) on the Air. However, There Are Some Gaps in the Current Research on How Multiple Weather Factors Work Together and the Complex Relationship Between Weather and Pollutants. The Discussion of These Problems Is Helpful to Supplement the Study of the Relationship Between Natural Meteorology and Air Quality, So as to Better Face the Problem of Air Quality.

Based on Air Quality Data and Meteorological Data from 2018 to 2022 (taking Beijing as an Example), Spearman Correlation Analysis Was Used to Identify Key Meteorological Factors, Multiple Linear Regression and Stepwise Regression Were Used to Establish Models, and the Comprehensive Effects of Precipitation, Temperature, Wind Speed and Other Factors on Pollutants Were Systematically Quantified. So as to Provide a Scientific Basis for Precise Pollution Control and Dynamic Control Strategies.

2. Data and Research Methods

2.1 Data Source and Description

The data needed for the analysis of this study mainly come from two authoritative departments. The air quality data is taken from China National Environmental Monitoring Center (CNEMC) [6]. The daily concentration data (unit: μg/m ³) of PM in Beijing from January 1, 2018 to December 31, 2022 is selected, and the monthly average value is taken. Meteorological data are from China Meteorological Association (CMA) [7], including daily average wind speed (m/s), average temperature (℃), relative humidity (%) and daily cumulative precipitation (mm) in the same period (Table 1).All data are open observations with integrity and reliability, which can effectively support the study of the relationship between meteorological factors and air quality.

2.2 Index Selection and Description

In This Study, the Average Daily Average Concentration of PM PM Was Selected as the Dependent Variable to Characterize the Air Quality. The Independent Variables Are Wind Speed, Air Temperature, Humidity and Precipitation, in Which Wind Speed Reflects the Atmospheric Diffusion Capacity, Air Temperature Affects the Photochemical Reaction and Emission Conditions, Humidity Is Closely Related to the Formation of Secondary Particles, and Precipitation Represents the Scavenging Effect of Wet Deposition. These Indicators Have Been Widely Confirmed to Have an Important Impact on Air Quality in Existing Studies, and Have Clear Physical Significance and Statistical Operability. All Indicators Are Averaged Over the Daily Mean of Each Month to Maintain Consistency of Time Scale.

2.3 Method Introduction

Data Analysis Methods Based on Statistics Were Used in the Study. Firstly, Descriptive Statistics of the Data Is Carried Out to Calculate the Mean, Standard Deviation and Extreme Value of Each Variable in Order to Grasp the Characteristics of Data Distribution. Then Spearman Correlation Coefficient Was Used to Analyze the Linear Correlation Between PM PM and Meteorological Factors. Finally, the Multiple Linear Regression Model and the Stepwise Regression Model Were Established, with the Concentration of PM as the Dependent Variable and the Four Meteorological Elements as the Independent Variables, and the Enter Method Was Used for Regression Analysis. The Multiple Linear Regression Model Is of the Form:

$$PM_{2.5} = \beta_0 + \beta_1 \cdot WS + \beta_2 \cdot T + \beta_3 \cdot RH + \beta_4 \cdot P + \varepsilon \tag{1}$$

Multicollinearity Was Tested by Variance Inflation Factor (VIF) to Ensure Model Robustness. All Analyses Were Done Using SPSS Software with a Significance Level Set at A = 0.05 [8].

Table 1: Definition of variables

Variable symbol	Meaning	Unit
$PM_{2.5}$	Mass concentration of fine particles	μg/m³

β_0	Model intercept	μg/m³
$\beta_1, \beta_2, \beta_3, \beta_4$	Partial regression coefficient of each variable	(μg/m³) / unit
WS	Wind speed	m/s
t	Temperature	°C
RH	Relative humidity	%
P	Precipitation	mm
ε	Random error term	μg/m³

3. Results and Discussion

3.1 Results of Descriptive Statistical Analysis

In this study, descriptive statistical analysis of air quality and meteorological data in Beijing was carried out, and the sample size was 60 observations. As shown in Table 2, the mean concentration of PM PM. Was 37.6 μg/m ³, with a concentration range of 16 to 82 μg/m ³ and a standard deviation of 14.304, indicating some volatility in the data. The kurtosis value was 0.565 and the skewness was 0.739, which showed that the data distribution was right-skewed and had spike characteristics.

The analysis of meteorological elements shows that the average surface pressure is 1015.362 hPa, ranging from 994.9 to 1039.1 hPa, and the coefficient of variation is only 0.011, indicating that the pressure data is relatively stable. The average precipitation is 1. 387 mm, but its standard deviation is 1. 335, and the coefficient of variation is 0. 962, indicating that the distribution of precipitation is extremely uneven and there is an obvious fluctuation. The average wind speed is 3.89 m/s, with a small variation range (3.21-4.53 m/s) and a coefficient of variation of 0.085, indicating that the wind speed data is relatively stable. The average temperature is 12.067 ° C, ranging from -5.6 ° C to 26.35 ° C, with a standard deviation of 10.138 and a coefficient of variation of 0.84, reflecting the obvious characteristics of seasonal temperature difference.

Table 2: Descriptive Analysis

Variable name	Sample size	Maximum value	Minimum value	Average value	Standard deviation	Median	Variance
PM2.5	60	82	16	37.6	14.304	37	204.617
Surface air pressure	60	1039.1	994.9	1015.362	10.837	1013.15	117.446
Precipitation	60	4.31	0	1.387	1.335	1.005	1.782
Wind speed	60	4.53	3.21	3.89	0.329	3.92	0.108
Temperature	60	26.35	-5.6	12.067	10.138	12.88	102.782

3.2 Spearman Correlation Analysis Results

The results of Spearman correlation analysis (Table 3) showed that there was a significant correlation between the variables ($p < 0.01$). The correlation between the concentration of PM PM and the meteorological elements showed that there was a significant positive correlation with the surface pressure and wind speed, and a significant negative correlation with the precipitation and temperature.

The correlation analysis of meteorological elements shows that the surface pressure and rainfall are significantly positively correlated with wind speed, and strongly negatively correlated with precipitation and temperature. There is a significant negative correlation between precipitation and wind speed, and a strong positive correlation between precipitation and temperature. There is a significant negative correlation between wind speed and air temperature.

Table 3: Spearman Correlation Analysis Results

Variables	$PM_{2.5}$	Surface air pressure	Precipitation	Wind speed	Temperature
$PM_{2.5}$	1	0.392**	-0.463**	0.35**	-0.494**
Surface air pressure	0.392**	1	-0.83**	0.676**	-0.869**
Precipitation	-0.463**	-0.83**	1	-0.688**	0.977**
Wind speed	0.35**	0.676**	-0.688**	1	-0.71**
Temperature	-0.494**	-0.869**	0.977**	-0.71**	1

*Note: * * means p < 0.01*

These correlation results indicate that the increase of air temperature and precipitation is helpful to reduce the concentration of PM, and the increase of surface air pressure and wind speed may be related to the accumulation of pollutants. There was a strong correlation between meteorological elements, especially between precipitation and temperature, which showed a nearly complete positive correlation (R = 0. 977), indicating that multicollinearity should be paid attention to in the subsequent regression analysis.

3.3 Results of Multiple Linear Regression Analysis

Table 4 Results of Multiple Regression Analysis

	Non-normalized coefficient		Normalization factor	t	P	VIF	R^2	Adjust R^2	F
	B	Standard error	Beta						
Constant	428.269	308.93	-	1.386	0.171	-	0.264	0.211	F=4.941 P=0.002***
Surface air pressure	-0.375	0.305	-0.284	-1.23	0.224	3.988			
Precipitation	-0.446	3.136	-0.042	-0.142	0.887	6.403			
Wind speed	0.587	8.337	0.014	0.07	0.944	2.755			
Temperature	-0.975	0.443	-0.691	-2.2	0.032**	7.377			
Dependent variable: PM2.5									

*Note: * * *, * * and * represent the significance levels of 1%, 5% and 10% respectively*

Table 4 shows that the regression model with PM PM. Concentration as the dependent variable and surface air pressure, precipitation, wind speed and air temperature as the independent variables is statistically significant (F = 4.941, p = 0.002). After model adjustment, R^2 was 0.211, indicating that the independent variable could explain 21.1% of the variation in PM PM. Concentration.

In terms of the contribution of each variable, the effect of air temperature on the concentration of PM PM was the most significant (B = -0.975, p = 0.032), and the normalization coefficient Beta was -0.691, indicating that the concentration of PM, decreased by 0.975 $\mu g/m^3$ on average when the air temperature increased by 1 ° C. Although the surface air pressure did not reach a significant level (B = -0.375, p = 0.224), its standardized coefficient was -0.284, showing a certain negative trend. The effects of precipitation and wind speed on PM PM concentration were not statistically significant (p > 0.05).

The multicollinearity diagnosis showed that the variance inflation factor (VIF) of each variable was greater than 2. 5, and the VIF values of precipitation and temperature were 6. 403 and 7. 377, respectively, indicating that there was a certain degree of multicollinearity between independent variables, which may affect the estimation accuracy of the coefficient.

Regression analysis verified the dominant effect of air temperature on PM PM. Concentration, which was consistent with the conclusion of correlation analysis. For every 1 ° C increase in temperature, the concentration of PM PM. Is expected to decrease by 0.975 $\mu g/m^3$. This result is consistent with the laws of atmospheric physics-high temperatures usually enhance vertical convection and promote the diffusion of pollutants; however, there are other factors that contribute to this result, such as air pollution caused by heating in winter. However, other meteorological elements are not significant, which may be related to multicollinearity. The correlation coefficient between temperature and precipitation is 0.977, which makes it difficult for the regression model to distinguish the independent effects of temperature and precipitation.

3.4 Stepwise Regression Results

Table 5: Stepwise Regression Analysis Results

Linear regression analysis n = 60								
	Non-normalized coefficient		Normalization factor	P	VIF	R^2	Adjust R^2	F
	B	Standard error	Beta					
Constant	46.005	2.531	0	0.000***	-	0.244	0.231	F=18.687,P=0.000***
Temperature	-0.696	0.161	-0.494	0.000***	1			
Dependent variable: PM2.5								

*Note: * * *, * * and * represent the significance levels of 1%, 5% and 10% respectively*

The purpose of this stepwise regression analysis is to select the most explanatory variables for PM2.5 concentration from the four meteorological factors of surface pressure, precipitation, wind speed and temperature, and to construct the optimal model. Table 5 shows that only the "air temperature" variable was retained in the final model because of its statistically significant contribution, while surface pressure, precipitation, and wind speed were all automatically rejected by the model because their contributions were not significant.

The final linear regression model shows that air temperature is the key factor to predict the concentration of PM2.5. There is a negative correlation between temperature and PM2.5 concentration, that is, PM2.5 concentration is expected to decrease by about 0.696 units for every unit increase in temperature.

From the statistical test, the overall effect of the model is significant. The P value of the F test was less than 0.001, indicating that the model was highly statistically significant as a whole, and the null hypothesis that there was no linear relationship between the variables was rejected. Although the goodness of fit of the model (adjusted R^2 to 0.231) indicates that the single variable of temperature can explain about 23.1% of the change in PM2.5 concentration, suggesting that there are still other factors (such as anthropogenic emissions, regional transport, etc.) that are not included in the model, it is common for a single meteorological factor to achieve such an explanation in environmental studies. In addition, the variance inflation factor (VIF) is 1, confirming that the model does not have multicollinearity problems and is well constructed.

4. Discussion

In this study, the influence mechanism of meteorological factors on PM PM concentration was revealed through systematic statistical analysis. The results showed that air temperature was the key meteorological factor affecting the concentration of PM PM, and its standardized regression coefficient was -0.691, which was statistically significant. This finding is consistent with the existing theoretical studies and confirms the role of enhanced atmospheric convection in the diffusion of pollutants under high temperature conditions.

Correlation analysis and regression analysis were used to identify the influence degree of each factor. In the process of analysis, we found that multicollinearity and sample size constraints are important methodological factors affecting the results of the model. By using the stepwise regression method, we have drawn a reliable research conclusion, but this methodological problem also provides an important reference for subsequent research. It is suggested that other ways can be used to improve the accuracy of the model in the future, such as using the least angle regression (LARS) model and sending the prediction results to the BP neural network for decision-level fusion [9]. Or a digital monitoring system that integrates temperature, humidity, wind speed and wind direction parameters using Internet of Things-based PM_{10} and $PM_{2.5}$ [10].

5. Conclusion

The core conclusion of this study is that air temperature has a significant negative effect on PM PM. Concentration, and PM. Concentration is expected to decrease by 0.696 μg/m ³ for every 1 ° C increase in temperature. This finding verifies the scientificity and feasibility of temperature as a prediction index of PM PM. Pollution.

Based on the results, it is suggested that emission monitoring should be strengthened in the prevention and control of regional air pollution, especially in the low temperature season, and a short-term pollution warning mechanism should be established in combination with meteorological forecast data. At the same time, in order to further enhance the scientific value of the study, the following directions can be explored in the future: expanding the sample size and time span to improve the stability of the model; including anthropogenic emission factors and socio-economic variables to enhance the explanatory power of the model; trying nonlinear methods such as machine learning to capture the complex interactions between factors; Sectional modeling was carried out to study the dominant factors in different seasons. These research results will provide a scientific basis for the establishment of more accurate meteorological early warning models and the formulation of seasonal control strategies.

References

[1] Liu, Z. C. Effect of urban greening on air quality by big data analysis method. Environmental Scienceand Management. 2025, 50(8), pp. 162-166. https://doi.org/10.3969/j.issn.1673-1212.2025.08.034.

[2] Niu, C. Y. and Su, J. D. Research on the impact of northern China's clean heating policy on air quality: Evidence from geographic regression discontinuity design. Environmental Pollution & Control. 2025, pp. 1-16. https://doi.org/10.15985/j.cnki.1001-3865.202503102.

[3] Dong, X., Zhou, N. and Zhao, X. M. Can digital village construction improve regional air quality?: Empirical evidence from Chinese counties Ecological Economy. 2025, 41(7), pp. 186-195.

[4] Wang, X. H., Zhu, J., Li, K., Chen, L., Jin, Q. P. and Liao, H. Impact of biogenic emissions on ozone concentrations during summer under climate change in China. Research of Environmental Sciences. 2025, 38(7), pp. 1474-1484. https://doi.org/10.13198/j.issn.1001-6929.2025.04.17.

[5] Yang, S. J., Ge, Y. F., Yang, C., Mi, X. N., Dong, J. Q., Niu, Y. Y., Yan, Y. L. and Peng, L. Effects of land use and landscape pattern on air quality in north slope area of tianshan mountains. Earth and Environment. 2025, 53(5), pp. 643-653. https://doi.org/10.3724/EE.1672-9250.2025.53.039.

[6] China National Environmental Monitoring Centre. Air Quality Monitoring Report. Available from: https://www.cnemc.cn/jcbg/kqzlzkbg/ (accessed 8 January 2026).

[7] China Meteorological Data Network. Basic Surface Meteorological Observation Data of China. Available from: https://data.cma.cn/data/detail/dataCode/A.0012.0001.html (accessed 8 January 2026).

[8] SPSSPRO. Scientific Platform Serving for Statistics Professional 2021 (Version 1.0.11). Available from: https://www.spsspro.com (accessed 8 January 2026).

[9] Zhang, B., Li, X. and Zhao, Y. Air quality PM2.5 prediction based on multi-model fusion. In Proceedings of the 31st Chinese Control and Decision Conference, Beijing, 2019; pp. 445-450.

[10] Simanullang, R., Soekirno, S. and Larassari, H. A. Design and analysis of air quality monitoring system PM10 and PM2.5 integrated with weather parameters (a case study on Margonda Raya street Depok). Journal of Physics: Conference Series. 2020, 1528(1), p. 012053. https://doi.org/10.1088/1742-6596/1528/1/012053.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access