Proceedings of CONF-MPCS 2026 Symposium: Theoretic Physics and Plasma Physics
DOI: 10.54254/2753-8818/2026.DL35114

Meteorological Factors of PM2.5 and O_3 Combined Pollution in Beijing Based on Pearson Correlation Analysis and Multiple Linear Regression

Jiachen Kang

School of Mathematical Sciences, Inner Mongolia Normal University, Hohhot, China
20241101867@mails.imnu.edu.cn

Abstract. PM2.5 and O_3 often appear simultaneously in the urban atmospheric environment, forming typical composite pollution. Various meteorological elements will have different regulation effects on the concentration of pollutants. In order to objectively quantify the linear relationship between multiple variables, it is necessary to rely on standardized mathematical statistics to carry out systematic calculations. In this paper, the daily air quality and meteorological synchronous observation data of Beijing are selected. The daily average concentration of PM2.5 and O_3 is taken as two sets of dependent variables, and the five meteorological indexes of temperature, air pressure, wind speed, precipitation and relative humidity are taken as independent variables. Pearson correlation analysis is used to judge the correlation degree of variables, and a multivariate linear regression model is constructed to quantify the effect of factors. The collinearity diagnosis and overall F test are carried out simultaneously to avoid the deviation caused by multicollinearity and ensure that the statistical calculation results are statistically significant. The output results of the model show that the overall regression model of PM2.5 is significantly effective. Wind speed, temperature, precipitation and relative humidity have obvious linear effects on PM2.5, and air pressure has no significant effect. The influence direction and intensity of meteorological factors on PM2.5 and O_3 are significantly different, and the two types of pollutants show the opposite change characteristics. The study not only clarifies the meteorological driving law of atmospheric compound pollution, but also provides a feasible mathematical implementation scheme for the quantitative analysis of multivariate environmental problems.

Keywords: PM2.5, O_3, meteorological factors, multiple linear regression, synergy

1. Introduction

As industrialization and urbanization continue to advance, the state of atmospheric environmental quality has drawn widespread concern across all sectors of society [1]. In recent years, countries around the world have ramped up efforts on atmospheric environmental management, bringing about a general improvement in global air conditions. Located as a core hub in the Beijing-Tianjin-Hebei urban agglomeration, Beijing has achieved notable outcomes in the prevention and treatment

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.DL35114

of air pollution. Nevertheless, heavy pollution events still occur from time to time. The compound pollution problem caused by PM2.5 and O_3 has become increasingly serious, turning into a major obstacle restricting further pollution reduction [2]. After effective control over pollutant emission sources, it is urgent and necessary to clarify the independent and interactive impacts of meteorological conditions on air pollutants, so as to provide support for the long-term steady improvement of regional atmospheric quality.

Existing studies have shown that meteorological factors, including temperature, pressure, precipitation, wind speed, and humidity, can obviously affect the concentration of PM2.5 and O_3 by changing the deposition, diffusion, and dilution of pollutants and the intensity of photochemical reaction [3]. Relevant studies have shown that Jin et al. used multiple linear regression to analyze the correlation between meteorological conditions and AQI in the Beijing-Tianjin-Hebei region, and found that average temperature hurt AQI, while air pressure and humidity had a positive impact [4]. In addition, Yu et al. analyzed three heavy pollution events in Shijing Shan District of Beijing in 2024 by using the HYSPLIT model, clarifying that weather conditions have a great impact on the development of pollution events [5]. However, most of the current studies only focus on the variation of a single pollutant, and the synergistic effect of meteorological factors under PM2.5 and O_3 compound pollution is insufficient. The internal mechanism of the combined effect still needs to be further explored and improved.

In the study, daily meteorological and air quality monitoring data of Beijing in 2025 were collected. Combined with Pearson correlation analysis and the multiple linear regression method, the actual effects of meteorological conditions such as temperature, air pressure, wind speed, precipitation, and humidity on PM2.5 and O_3 concentrations were analyzed. The study further clarifies the interaction between meteorological factors in different weather scenarios by quantifying the intensity and contribution of various meteorological indicators. The relevant conclusions drawn from this analysis can also provide a realistic basis and feasible reference for the local optimization of air quality forecasting and the formulation of joint pollution prevention and control programs.

2. Research methods

2.1. Data sources and descriptions

This study focuses on Beijing and uses complete monitoring data for the whole year of 2025 for analysis [6]. Daily records cover local air quality and meteorological indicators. The air quality part includes the daily concentration records of major pollutants in Beijing, mainly PM2.5 and O_3. Meteorological records include conventional monitoring indicators such as temperature, precipitation, and wind speed. The data set is comprehensive and complete, which is very suitable for exploring how various weather conditions affect these two pollutants and their comprehensive pollution characteristics.

To build stable and credible statistical models, all raw data were cleaned and organized via SPSS. Invalid records with missing dates or missing meteorological and pollutant indicators were excluded upfront to ensure the overall integrity of research data. The IQR approach was applied to identify abnormal PM2.5 values, with the lower quartile Q_1 and upper quartile Q_3 calculated first to determine the interquartile range. Values lower than $Q_1 - 1.5 \times \mathrm{IQR}$ or higher than $Q_3 + 1.5 \times IQR$ were judged as outliers, and any records containing such abnormal data were removed to maintain overall data quality. Strict screening yielded 346 valid daily samples from 2025, capable of effectively supporting subsequent statistical analysis.

DOI: 10.54254/2753-8818/2026.DL35114

2.2. Indicator selection and explanation

This study focuses on Beijing and analyzes the internal correlation between meteorological elements and major air pollutants. In this study, the daily mean values of PM2.5 and O_3 were used as dependent variables, and five meteorological factors, such as temperature, air pressure, wind speed, precipitation, and humidity, were used as independent variables to explore their driving effects on pollutant diffusion, transformation, and synergistic change.

2.3. Method introduction

This study integrated two analytical methods to elucidate the independent effects of meteorological elements on PM2.5 and O_3 levels. It also further analyzes the comprehensive interactive performance affected by the common weather background.

In this study, Pearson correlation analysis was used to evaluate the linear relationship between various meteorological indicators and the changes in PM2.5 and O_3 concentrations. The analysis tool can directly reflect the intensity and trend of their internal correlation, and objectively reflect the positive and negative effects of different weather conditions. The detailed discussion of each meteorological element is helpful to distinguish the key influencing factors and clarify their actual characteristics, so as to better support the follow-up research content. The specific calculation formula of the Pearson correlation coefficient is as follows:

$$r_{xy} = \frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}} \tag{1}$$

Where r_{xy} is the Pearson correlation coefficient, x_i and y_i are the i-th observed values of the two variables, $\bar{x}$ and $\bar{y}$ are the mean values of the observed values of the two variables, and n is the sample size.

Next, a multiple linear regression model is constructed to objectively evaluate the independent influence and real contribution of each meteorological factor to pollutant concentration in the comprehensive environment. To prevent bias due to arbitrary selection, all five relevant meteorological factors were included as independent variables. The validity of the model is confirmed by R and F tests, so that it can compare and analyze in detail how these factors affect the two pollutants differently. The form of the regression model is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \tag{2}$$

Where Y represents the concentration of pollutants (PM2.5 or O_3), β_0 is the intercept term, $\beta_1, \beta_2, \cdots, \beta_k$ are the regression coefficients of the respective independent variables, $X_1, X_2, \cdots, X_k$ are the significantly correlated meteorological factors, and ε is the random error term.

DOI: 10.54254/2753-8818/2026.DL35114

3. Results

3.1. Pearson correlation analysis

The results of Pearson correlation analysis are shown in Table 1, showing significant differences in the effects of various meteorological factors on the two pollutants. In terms of the correlation between PM2.5 concentration and meteorological elements, wind speed was significantly negatively correlated with PM2.5 concentration. The increase in wind speed accelerates the diffusion and dilution of particulate matter and effectively reduces the concentration of PM2.5.

The correlation analysis between O_3 concentration and meteorological indicators showed that there was a highly significant positive correlation between temperature, wind speed, and O_3 content. The increase in ambient temperature can accelerate the progress of atmospheric photochemical reactions and effectively promote the production of O_3. Appropriate wind speed supports the mixing and accumulation of pollutant precursors in the air, which further promotes the continuous increase of O_3 concentration. In contrast, atmospheric pressure is significantly negatively correlated with O_3. Low-pressure systems are usually accompanied by strong solar radiation, which is a favorable condition for the formation and stable accumulation of O_3 in the atmosphere.

Table 1. Pearson correlation analysis results between PM2.5, O_3 concentrations, and meteorological elements in Beijing

	PM2.5	O_3
Temperature	-0.046	0.310**
Wind speed	-0.209**	0.282**
Atmospheric pressure	-0.001	-0.404**
Precipitation	-0.063	-0.065
Humidity	0.114*	-0.057
* p<0.05 ** p<0.01		

3.2. Multiple linear regression analysis

Table 2 shows the results of multiple linear regression analysis. The PM2.5 concentration model is significant. Temperature, wind speed, precipitation, and humidity are the main negative factors. Higher temperatures and wind speeds promote air diffusion, and precipitation contributes to the deposition of particles in the air. High-pressure weather conditions promote the diffusion of pollutants, all of which contribute to the reduction of PM2.5 concentration. The humidity showed a significant positive driving effect. Higher humidity promotes the hygroscopic expansion of particulate matter, resulting in an increase in concentration. From the perspective of the overall model indicators, although the explanatory power of meteorological factors on PM2.5 concentration is not high, the overall model has passed a very important test. The VIF values of all variables are less than 5, indicating that there are no multicollinearity problems and the model results are reliable.

Table 2. Results of the Multiple Linear Regression Model for PM2.5 Concentration in Beijing

	Unstandardized Coefficient		Standardized Coefficient	t	p Value	Collinearity Diagnostic	
	B	Std. Error	Beta			VIF	Tolerance
Constant	1167.224	617.521	-	1.89	0.06	-	-
Temperature	-1.221	0.541	-0.216	-2.257	0.025*	3.358	0.298

DOI: 10.54254/2753-8818/2026.DL35114

Table 2. (continued)

Wind speed	-1.917	0.55	-0.191	-3.489	0.001**	1.107	0.903
Atmospheric pressure	-0.994	0.599	-0.16	-1.661	0.098	3.426	0.292
Precipitation	-33.795	13.701	-0.144	-2.467	0.014*	1.249	0.8
Humidity	0.397	0.177	0.143	2.248	0.025*	1.493	0.67
R-squared			0.076				
Adjust R-squared			0.062				
F Test			F (5,340)=5.560,p=0.000				
RMSE			63.108				
D-W			1.389				
Note: Dependent Variable = PM2.5							
* p<0.05 ** p<0.01							

As shown in Table 3, the regression model of O_3 concentration reached a significant level as a whole. In terms of driving effect, wind speed has a significant positive driving effect on O_3. Appropriate wind speed contributes to the mixing reaction of O_3 precursors and promotes the formation of O_3. At the same time, atmospheric pressure and humidity show a significant negative driving effect. High-pressure environments easily lead to the accumulation of pollutants, while high humidity inhibits photochemical reactions, both of which lead to a decrease in O_3 concentration. The effects of temperature and precipitation on O_3 concentration do not meet the significance test, so they are not the main factors affecting their changes. The collinearity test of the model shows that the VIF value is less than 5, which excludes multicollinearity and ensures reliable results.

Table 3. Results of the multiple linear regression model for O_3 concentration in Beijing

	Unstandardized Coefficient		Standardized Coefficient	t	p Value	Collinearity Diagnostic	
	B	Std. Error	Beta			VIF	B
Constant	1212.346	224.678	-	5.396	0.000**	-	-
Temperature	-0.115	0.197	-0.051	-0.587	0.558	3.358	0.298
Wind speed	0.795	0.2	0.198	3.977	0.000**	1.107	0.903
Atmospheric pressure	-1.167	0.218	-0.469	-5.36	0.000**	3.426	0.292
Precipitation	-0.34	4.985	-0.004	-0.068	0.946	1.249	0.8
Humidity	-0.16	0.064	-0.144	-2.49	0.013*	1.493	0.67
R-squared			0.238				
Adjust R-squared			0.227				
F Test			F (5,340)=21.294,p=0.000				
RMSE			22.961				
D-W			0.948				
Note: Dependent Variable = O_3							
* p<0.05 ** p<0.01							

The comparative analysis of the model shows that there are obvious differences in the effects of five meteorological factors on PM2.5 and O_3. Temperature, wind speed, precipitation, and humidity work together on PM2.5, and wind speed is the core factor to inhibit its concentration increase. O_3 is

DOI: 10.54254/2753-8818/2026.DL35114

mainly controlled by wind speed, air pressure, and humidity. Wind speed can effectively diffuse and reduce PM2.5, but it will increase O_3 concentration. The two show obvious reverse changes, which fully reflect the complexity of the interaction between the two types of pollutants.

3.3. Analysis of the synergistic effect of meteorological factors on PM2.5 and O_3 concentrations

Table 4 shows the results of the correlation analysis. There is a significant negative correlation between PM2.5 and O_3 levels in Beijing, showing a significant reverse relationship. This is consistent with the differential driving effect of meteorological factors.

Table 4. Pearson correlation analysis results of PM2.5 and O_3 concentrations in Beijing

		O_3
PM2.5	Coefficient	-0.448**
	*P*Value	0
	n	346
	* p<0.05 ** p<0.01	

Meteorological factors have different effects on these two pollutants. In particular, wind speed and humidity, which are opposite to each other, are the main drivers of their interrelated fluctuations. When explaining the change of O_3 concentration, meteorological factors are dominant, highlighting the sensitivity of O_3 and classifying it as a meteorological-driven pollutant. On the contrary, PM2.5 levels are more profoundly affected by non-meteorological factors, especially emission sources.

4. Discussion

Through rigorous multiple linear regression analysis, this study revealed a significant negative correlation between PM2.5 and O_3 concentration in Beijing. Temperature, wind speed and relative humidity are the key meteorological driving factors affecting the distinction of these pollutants, thus clarifying their negative correlation. The research conducted by Li et al. emphasized the important role of temperature in promoting O_3 pollution in the Beijing-Tianjin-Hebei region, which is consistent with the research results [7]. Jia et al. clarified the potential mechanism of PM2.5 reduction promoting O_3 production [8]. This shows that the combined pollution observed in Beijing is a typical representative of the national regional model, making a single control measure redundant. It is necessary to develop a sound prevention and control strategy, considering the synergistic relationship and the meteorological model.

Nevertheless, this study is not without shortcomings. The PM2.5 multiple linear regression model used here shows a relatively low goodness of fit. The level of PM2.5 is mainly affected by human emissions and regional transfer. However, the study only considered meteorological factors, resulting in insufficient explanation of concentration changes. In addition, the data span is short and not differentiated based on season or warming period, which limits the wide applicability of the conclusions [9, 10].

The follow-up study can extend the span of observation time series, incorporate pollution source-related data to optimize multi-factor analysis, compare the driving differences of meteorological conditions in different periods, further improve the interpretation ability of the model, and strengthen the practical reference and application value of the research conclusions.

DOI: 10.54254/2753-8818/2026.DL35114

5. Conclusion

Based on the daily meteorological and air quality monitoring data of Beijing in 2025, this study selected five core meteorological indicators of temperature, air pressure, wind speed, precipitation, and humidity to explore. Combined with Pearson correlation analysis and multiple regression model, the effects and interactions of various meteorological conditions on PM2.5 and O_3 concentrations were systematically explored.

From the point of view of the influence characteristics, the effect of meteorological elements on the two types of air pollutants is obviously different. Among them, wind speed and humidity are the core factors that regulate the growth and decline of the two pollutants: wind speed can accelerate the diffusion and dilution of near-surface particulate matter, and promote the diffusion and mixing of O_3 precursors, indirectly helping O_3 generation; the increase of air humidity will promote the hygroscopic expansion of fine particles, inhibit the process of atmospheric photochemical reaction, and then slow down the accumulation of O_3.

Further analysis found that the key meteorological driving conditions of the two pollutants were different. The concentration of PM2.5 was affected by four meteorological indicators, while the change of O_3 only significantly responded to wind speed, air pressure and humidity. The comparison results show that O_3 is more sensitive to the fluctuation of meteorological environment, and meteorological factors can explain 22.7 % of its concentration variation, which is much higher than 6.2 % of PM2.5.

The comprehensive results confirm that there is a significant negative correlation between PM2.5 and O_3 concentration in Beijing, forming a complex pollution pattern dominated by meteorological conditions, and also indicating that it is difficult to achieve efficient pollution control by a single treatment method. This study can provide meteorological data support and theoretical reference for the coordinated control of regional atmospheric composite pollution. In the future, the research sequence can be expanded, the pollution source emission data can be included, the analysis system can be further improved, and the reference value of the research conclusion can be improved.

References

[1] Hu, D.P. (2025) Simulation Study on the Impact of Climatic Conditions on Atmospheric Pollutant Diffusion and Concentration Changes. Environmental Science and Management, 50, 66-71.
[2] Li, C.L., Feng, C.L. and Sun, R. (2024) Analysis of PM2.5 and Ozone Complex Pollution Characteristics in Xuzhou City from 2019 to 2023. Sichuan Environment, 43, 7-14.
[3] Kong, T., Li, S., Huang, J., et al. (2023) Study on Meteorological Factors Affecting Atmospheric Pollutant Changes —A Case Study of Xinjiang Area. Environmental Science and Management, 48, 33-38.
[4] Jin, R.H., Zeng, G.J., Li, Y.X., et al. (2022) Analysis of Air Quality Distribution Characteristics and Meteorological Influencing Factors in Beijing-Tianjin-Hebei Region. Resources Conservation and Environmental Protection, 6, 41-44.
[5] Yu, F.R., Qu, L.Z., Yi, H.W., et al. (2026)Characteristics and Causes of Severe Air Pollution in Shijingshan District, Beijing: A Case Study of Three Severe Pollution Episodes in 2024. Journal of Capital Normal University (Natural Science Edition), 1-17.
[6] Kaggle. World Weather Repository (Daily Updating). (2023-08-29) [2026-02-14]. https: //www.kaggle.com/datasets/nelgiriyewithana/global-weather-repository.
[7] Li, H., Liu, X., Liu, Z., Li, M., Wu, T., Li, P. and Zhou, P. (2025) Tracking Long-Term Ozone Pollution Dynamics in Chinese Cities with Meteorological and Emission Attribution. Atmosphere, 16, 768.
[8] Jia, C., Yan, G., Yu, X., et al. (2025) Evidence for Coordinated Control of PM2.5 and O_3: Long-Term Observational Study in a Typical City of Central Plains Urban Agglomeration. Toxics, 13, 330.
[9] Li, B.W., Zhang, K., Gao, J., et al. (2024) Impact of Meteorological Conditions on Dual High PM2.5 and O_3 Concentrations in Beijing-Tianjin-Hebei and Surrounding Areas in Summer and Early Autumn. Environmental

DOI: 10.54254/2753-8818/2026.DL35114

Science, 45, 4375-4384.
[10] Zhang, M.M., Yan, S.Z., Chen, J.Y., et al. (2025) Characteristics of PM2.5 Pollution During Heating Period in Dalian City and Analysis of Potential Source Areas of Air Pollution. Environmental Monitoring and Early Warning, 17, 89-96.