Pol. J. Environ. Stud. Vol. XX, No. X (XXXX), 1-14
DOI: *10.15244/pjoes/215216* ONLINE PUBLICATION DATE: 2026-03-04

Original Research

Spatial and Temporal Characteristics of Ozone Pollution and Sensitivity Analysis of Its Influencing Factors in Hangzhou

Huansang Chen[1], Kexin Yan[2], Jingyi Qiu[1], Mengjing Zhu[1], Yuting He[2], Ru Yi[3], Mingjuan Shi[2], Tao Ding[2]°*

[1]Zhejiang Lin'an Atmospheric Background National Observation and Research Station, Zhejiang Institute of Meteorological Sciences, Hangzhou 311300, China
[2]Department of Environmental Engineering, China Jiliang University, Hangzhou 310018, China
[3]College of Quality and Standardization, China Jiliang University, Hangzhou 310018, China

Received: 03 July 2025
Accepted: 05 December 2025

Abstract

In recent years, ozone (O_3) pollution in Hangzhou has become increasingly prevalent, posing a significant threat to public health and the ecological environment. This study investigated the spatiotemporal characteristics and sensitivity of ozone pollution in Hangzhou by analyzing monitoring data collected from 2018 to 2022 at 12 monitoring stations. The dataset included ozone concentrations, meteorological parameters, and pollutant information. The results show that: (1) The spatiotemporal distribution of ozone concentrations is characterized by a spatial pattern where concentrations are lower in the west and higher in the east. The monthly variation displays a bimodal pattern with peaks occurring in late spring and summer, while the diurnal variation demonstrates a unimodal pattern peaking in the afternoon. Higher ozone concentrations were consistently recorded during weekends in contrast to weekdays. (2) Sensitivity analysis demonstrated that higher temperatures and reduced NO_2 concentrations significantly facilitated ozone formation, whereas high relative humidity (>55%) suppressed this process; $PM_{2.5}$ concentrations demonstrated ozone-promoting effects within the range of 11-40 μg/m^3, while its impact plateaued when concentrations exceeded 77 μg/m^3. This study provides scientific evidence to support precise prevention and control strategies for ozone pollution in Hangzhou, Zhejiang Province.

Keywords: ozone pollution, machine learning, Hangzhou city, spatiotemporal characteristics, sensitivity analysis

*e-mail: dingtao@cjlu.edu.cn
Tel.: +86-571-13757133532
°ORCID iD: 0000-0003-4323-7625

Introduction

Ground-level ozone pollution has become a critical global environmental challenge. Accelerated urbanization and population growth have exacerbated the adverse impacts of atmospheric pollutants on human health and ecosystems [1-3]. As a typical secondary pollutant, ozone is formed through complex photochemical reactions driven by precursor compounds (NO_x and VOCs) under conditions of high temperature and intense solar radiation. Research confirms that ozone pollution not only increases the risks of respiratory and cardiovascular diseases but also poses a significant threat to crop productivity and ecosystem stability [4-6]. Therefore, a comprehensive analysis of ozone's spatiotemporal distribution patterns and driving factors is essential for developing targeted and effective mitigation strategies.

Ozone pollution demonstrates substantial spatiotemporal heterogeneity [7-9]. Spatially, the area exceeding O_3 concentration thresholds in China has been expanding annually in recent years, revealing pronounced regional disparities between urban, suburban, and rural areas. For instance, Yan et al. [10] analyzed data from a variety of monitoring sites across Europe (1995-2012) and discovered that ozone concentrations were generally higher in suburban and rural areas than in urban centers. Ren et al. [11] conducted a systematic investigation into the spatial evolution of ozone concentrations in Beijing from 2014 to 2020, demonstrating that high-ozone zones exhibited distinct migratory patterns. Initially concentrated in the northern suburbs, these zones subsequently expanded northeastward, southwestward, and into eastern urban core areas. Temporally, ozone concentrations exhibit pronounced seasonal fluctuations and diurnal variation patterns, and in certain regions, the "weekend effect" is observed. Notably, in recent years, the peaks of ozone pollution have shifted earlier, occurring more frequently in spring rather than in summer. For example, Chengdu exhibits a bimodal monthly pattern with peaks occurring in May and August, while Lishui City in the Yangtze River Delta shows an M-shaped trend with dual peaks in April-May and September. These variations are likely attributable to regional climatic differences and the evolution of precursor emissions.

Ensemble algorithms based on bagging strategies, such as Random Forest, utilize parallelized training to construct multiple base learners. By aggregating predictions, these algorithms enhance model accuracy and have demonstrated superior performance in studies of atmospheric pollution [12-14]. For instance, Ma et al. conducted a systematic investigation into the spatiotemporal evolution of ozone pollution in the Beijing-Tianjin-Hebei region from 2010 to 2017 using the Random Forest framework. Boosting-class algorithms, such as XGBoost, utilize iterative optimization to progressively minimize the prediction errors of base learners, thus constructing models with improved stability and accuracy [15]. Notably, Li et al. demonstrated substantial enhancements in the accuracy of air pollutant concentration predictions by integrating the WRF-Chem chemical transport model with the XGBoost machine learning algorithm [16].

Hangzhou, the second-largest city in the Yangtze River Delta, has witnessed its urban population surpass 9 million [17]. Between 2018 and 2022, ozone concentrations (measured as the 90th percentile of daily maximum 8-hour averages) increased by 4.9%, reaching 170 μg/m^3 in 2022, with a notable rise in exceedance days. Ozone pollution poses a significant threat to public health and acts as a barrier to achieving sustainable urban development. However, existing research has inadequately addressed the spatiotemporal heterogeneity of ozone pollution in Hangzhou and lacks sensitivity analyses of key driving factors. Notably, there is a scarcity of studies that integrate machine learning methods to explore multi-element synergistic effects. Therefore, in-depth investigation into the spatiotemporal characteristics of ozone pollution in Hangzhou and sensitivity analyses of its influencing factors have significant implications for formulating scientific pollution control policies [18, 19], enhancing public health awareness [20-22], advancing relevant academic research [23], strengthening regional collaborative governance and promoting sustainable development [24, 25].

Materials and Methods

Monitoring Data and Preprocessing

The data for this study were obtained from the China National Environmental Monitoring Centre (https://www.cnemc.cn/). The environmental monitoring network in Hangzhou City includes 12 stations: Municipal Government Building (S1), Town No.2 Middle School (S2), Yunqi (S3), Wolong Bridge (S4), Xixi (S5), Hemu Primary School (S6), Chengxiang Town (S7), Binjiang (S8), Fire Brigade (S9), Linping Town (S10), Zhejiang Agricultural University (S11), and Xiasha (S12). The geographical distribution of the aforementioned monitoring stations is depicted in detail in Fig. 1. Additionally, to comprehensively analyze the formation mechanisms of ozone pollution, this study integrates meteorological monitoring data as supplementary information. The specific locations of the meteorological monitoring stations are also shown in Fig. 1. However, since the actual locations of the meteorological monitoring stations do not perfectly align with those of the environmental monitoring stations, during data collection, the meteorological data from the station closest to each environmental monitoring station was selected. This ensures that the meteorological data accurately represent the climatic conditions at the corresponding site.

The collected monitoring data covered a comprehensive range of air pollutants, including

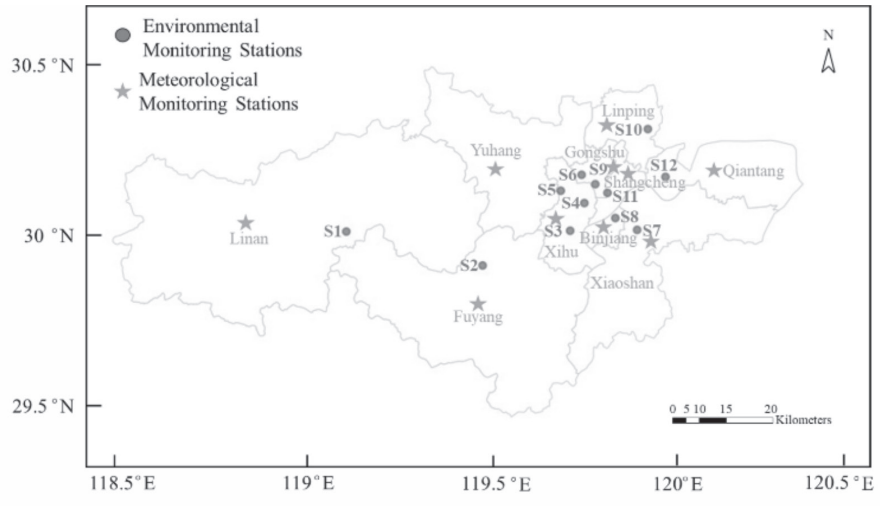

Fig. 1. Spatial distribution of environmental and meteorological monitoring stations in Hangzhou.

ozone (O_3), nitrogen dioxide (NO_2), fine particulate matter ($PM_{2.5}$), inhalable particulate matter (PM_{10}), carbon monoxide (CO), and sulfur dioxide (SO_2). Simultaneously, the meteorological parameters included temperature (T), relative humidity (RH), wind speed (WS), wind direction (WD), and precipitation (Precip).

The specific sample data were obtained from hourly monitoring records collected at each station. The data from Stations S3 to S12 cover hourly-averaged measurements over the period from 2018 to 2022, while the data for Stations S1 and S2 are limited to 2021-2022 owing to the relocation of these monitoring stations. The dataset collected across this temporal span provides a robust basis for analyzing the spatiotemporal characteristics of ozone pollution in Hangzhou and elucidating its formation mechanisms.

During the preprocessing of the collected data, this study utilized the Maximum Daily 8-Hour Average Ozone (MDA8-O_3) as the primary metric for assessing ozone concentrations. For the other monitored parameters, including NO_2, $PM_{2.5}$, PM_{10}, CO, SO_2, and meteorological factors (T, RH, WS, and WD), data were processed as daily averages. Precipitation data were derived by calculating daily precipitation totals. The aforementioned datasets collectively constitute the primary source of time-series data for the study.

Spatial Clustering Analysis Based on PAM

The PAM (Partitioning Around Medoids) clustering algorithm proposed by Kaufman and Rousseeuw in 2005 not only combines the advantages of the K-means clustering algorithm [26], such as simple principle, convenient operation, and wide application, but also performs well in handling outliers. The core of this algorithm lies in its main difference from the K-means algorithm, which is that PAM uses actual sample points as the center points of clusters (medoids), rather than relying solely on sample feature means. This method makes PAM more objective in extracting cluster features and helps improve the accuracy of clustering results.

Research Method

Machine Learning Model

Extreme Gradient Boosting (XGBoost), proposed by Chen et al. [27] in 2016, is a state-of-the-art boosting algorithm within the realm of ensemble learning techniques. Ensemble learning, a prominent area in modern machine learning, is based on the principle of aggregating multiple weak learners to achieve performance comparable to that of a strong learner. XGBoost has distinguished itself among numerous machine learning approaches, owing to its superior performance and robust generalization capabilities.

XGBoost is an additive model composed of k base models, and its iterative function can be mathematically expressed as:

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \tag{1}$$

Where $\hat{y}_i^{(t)}$ represents the prediction result of sample i after the t-th iteration; x_i represents the input sample; $f_k(x_i)$ represents the prediction result of the k-th tree; $\hat{y}_i^{(t-1)}$ represents the prediction result of the previous t-1 step; $f_t(x_i)$ represents the predicted value of the new model added for the t time.

Bayesian Optimization Method

Bayesian Optimization (BO) is a heuristic global optimization approach rooted in probability theory. It enables automated hyperparameter search for models, demonstrating exceptional efficacy in optimizing complex systems with high-dimensional parameter

spaces. This method has achieved widespread adoption within both the machine learning and deep learning communities.

BO is a method for estimating the extrema of a function when the objective function is unknown, utilizing information from existing sample points. Its core principle involves assuming a prior distribution, updating it to derive the posterior distribution, and refining the confidence level of the original model. Its mathematical expression is:

$$x^* = argmax_{x \in X} f(x) \tag{2}$$

Where x^* represents the optimal combination of hyperparameters; X represents the decision space; $f(x)$ represents the objective function.

Model Evaluation Indicators

To evaluate the accuracy of the model, this paper employs the coefficient of determination (R-Square, R^2), root mean square error (RMSE), and mean absolute error (MAE) as performance metrics to validate the constructed machine learning model. The calculation formula is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2} \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{5}$$

Where y_i represents the predicted values for the model; $\hat{y}_i$ represents the actual measured values; $\overline{y}_i$ represents the mean of the actual measured values; n represents the number of samples in the test set; $|y_i - \hat{y}_i|$represents the absolute error representing the i-th result.

Comparison with Other Machine Learning Studies

In addition to the studies conducted in Hangzhou, similar investigations have been carried out in other Chinese cities using machine learning techniques. While Wu et al. (2025) [28] utilized random forest algorithms to analyze ozone pollution across urban agglomerations in China, identifying key meteorological drivers like shortwave radiation and boundary layer temperature, the Hangzhou study employed an ensemble of techniques, including the BO-XGBoost model, to achieve higher predictive accuracy (R^2 = 0.8202) and lower error margins (RMSE = 0.1795). Similarly, Shen et al. (2025) [29] in Ningbo applied multiple machine learning models for ozone forecasting, with the gradient boosting regression tree (GBRT) model showing promising results (R^2 = 0.63503). However, the Hangzhou study's comprehensive integration of meteorological and pollutant data, coupled with the BO-XGBoost framework, offers a more nuanced understanding of ozone dynamics and superior performance in comparison.

Results and Discussion

Analysis of Spatial Characteristics of Ozone Pollution

This study utilizes ozone concentration data collected from 12 environmental monitoring stations in Hangzhou between 2018 and 2022. It calculates statistical metrics for each station, including MDA8-O_3, the number of heavily polluted days, and annual evaluation metrics to reveal disparities in pollution levels across sites and systematically assess interregional ozone pollution. The analysis aims to comprehensively characterize the spatial distribution patterns of ozone pollution in Hangzhou.

Spatial Clustering Results of Ozone Pollution

The PAM clustering method was used to perform clustering analysis on the daily time series of MDA8-O_3 at 12 monitoring stations in Hangzhou to reveal the spatial clustering pattern of ozone concentration. When analyzing different values of the clustering number k, the calculated contour coefficient shows that when k = 3, the contour coefficient reaches its maximum value, indicating that the number of cluster center points is the best choice at this time, and the clustering results obtained have high rationality and stability. Thus, Hangzhou can be objectively divided into three subregions, namely the western region (S1-S2), central region (S3-S7), and eastern region (S8-S12). The clustering results are shown in Fig. 2, which clearly illustrates the differences in ozone concentration distribution between different regions.

MDA8-O_3

Ozone pollution in Hangzhou demonstrates significant spatial heterogeneity, as depicted in Fig. 3. Specifically, the MDA8-O_3 concentrations in the western region are significantly lower than those in the central and eastern regions, highlighting distinct spatial gradients. This disparity is further substantiated by statistical analyses: The upper quartile of ozone concentrations at three monitoring stations in the western region peaked at 214.4 μg/m^3, while the lowest

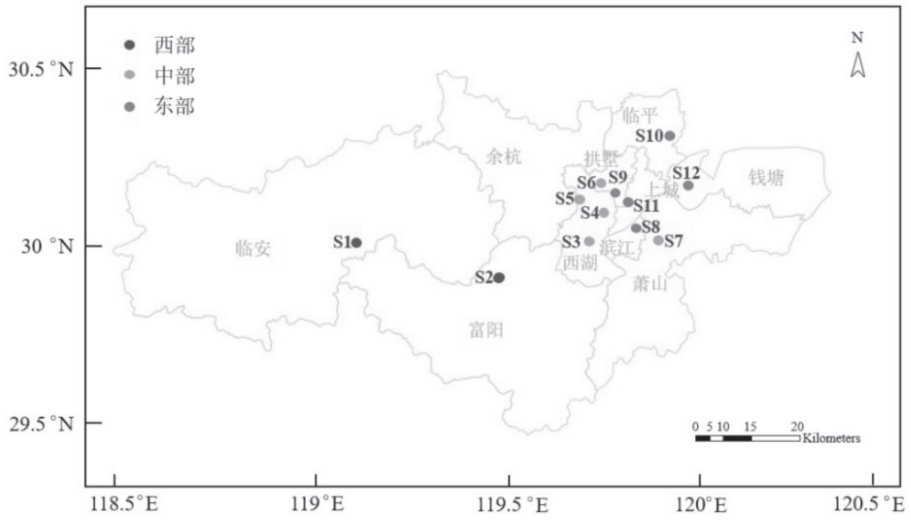

Fig. 2. Spatial clustering results of monitoring stations in Hangzhou city.

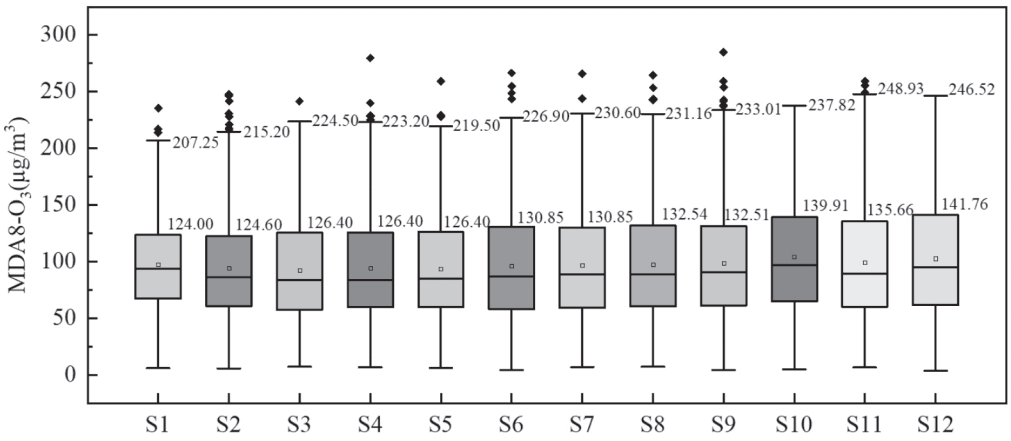

Fig. 3. MDA8-O_3 concentration box plot of monitoring stations in Hangzhou City.

upper quartile value in the central and eastern regions peaked at 231.5 μg/m^3, revealing a significant difference of 17.1 μg/m^3 between these areas. Concurrently, the maximum ozone concentration recorded in the western region was relatively low at 125.1 μg/m^3, which is below the lowest peak value observed in the central and eastern regions (135.9 μg/m^3). The data demonstrate that the central and eastern regions of Hangzhou exhibit elevated ozone pollution levels during peak concentration periods, which may be attributable to intensified economic activities, vehicular traffic, and industrial emissions characteristic of these regions.

The analysis of MDA8-O_3 indicators across different regions reveals the distinct spatial distribution patterns of ozone pollution in Hangzhou City. Specifically, there are significant disparities in pollution levels among the western, central, and eastern zones. The western region demonstrates relatively lower ozone concentrations, while the central and eastern regions demonstrate higher pollution intensity. This spatial heterogeneity reflects variations in regional pollution sources, meteorological conditions, and geographical characteristics.

Proportion of Days with Heavy Ozone Pollution

According to Ambient Air Quality Standards (GB 3095-2012), days with MDA8-O_3 concentrations equal to or exceeding 160 μg/m^3 were classified as severe ozone pollution events. A systematic analysis was performed to examine the proportion of days with severe ozone pollution across multiple monitoring stations in Hangzhou during the study period. The calculated proportions of severe ozone pollution days at each station are as follows: Monitoring station S1 accounts for 7.28%, S2 for 9.85%, S3 for 10.38%, S4 for 10.35%, S5 for 12.33%, S6 for 12.44%, S7 for 12.05%, S8 for 12.22%, S9 for 0.87%, S10 for 14.13%, S11 for 15.23%, and S12 for 13.61%. These data distinctly demonstrate spatial heterogeneity in the severity of ozone pollution among stations, as depicted in Fig. 4.

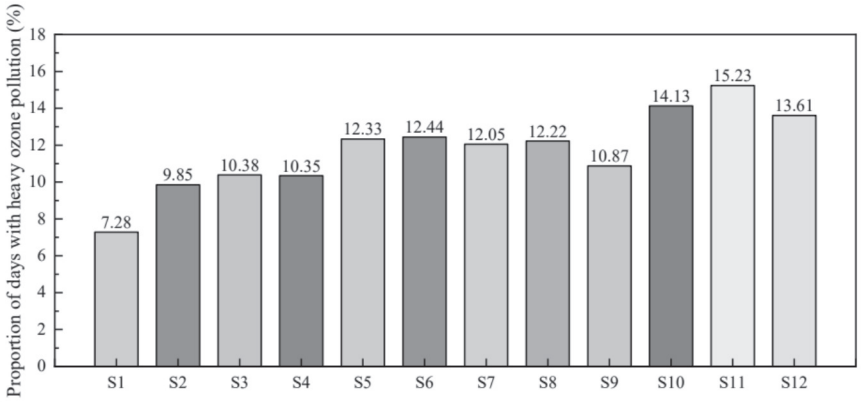

Fig. 4. Proportion of ozone heavily polluted days in Hangzhou.

Regional analysis demonstrates that, in the western districts, the proportion of days exceeding the ozone concentration threshold ranged from 7.28% to 9.85%, indicating relatively low pollution levels. Comparatively, the central sectors exhibited ozone exceedance frequencies ranging from 10.35% to 12.44%, while a significant increase was observed in the eastern sectors, with ozone exceedance frequencies varying between 10.87% and 15.23%. This spatiotemporal trend highlights significant spatial heterogeneity in ozone pollution gradients across Hangzhou's urban functional zones.

The spatial distribution of heavy pollution days across monitoring stations was quantified using regional averaging, resulting in proportions of 8.57% in the western region, 11.51% in the central region, and 13.21% in the eastern region. These results demonstrate a higher frequency of ozone pollution episodes in the eastern region than in other zones, highlighting greater challenges for environmental management and emission mitigation strategies in this area.

Annual Evaluation Index of Ozone

Ozone assessment under China's Ambient Air Quality Standards (GB 3095-2012) relies on the 90th percentile of maximum daily 8-hour average ozone concentrations (MDA8-O_3-90) as the primary evaluation metric, with a minimum pollution threshold set at 160 μg/m^3. Analysis of MDA8-O_3-90 measurements from twelve strategically located monitoring stations (S1-S12) in Hangzhou during the 2018-2022 period revealed a consistent upward trend with intermittent fluctuations in ozone pollution severity across all sampling locations, as depicted in Fig. 5.

A pronounced west-to-east spatial gradient in MDA8-O_3-90 variation was observed across Hangzhou, indicating a gradual increase in ozone pollution intensity from the western to the eastern sectors, as depicted in Fig. 6. Between 2018 and 2022, the western region exhibited fluctuating increases in ozone concentrations while still maintaining consistently lower levels compared to other regions. In contrast, the central zone demonstrated significant

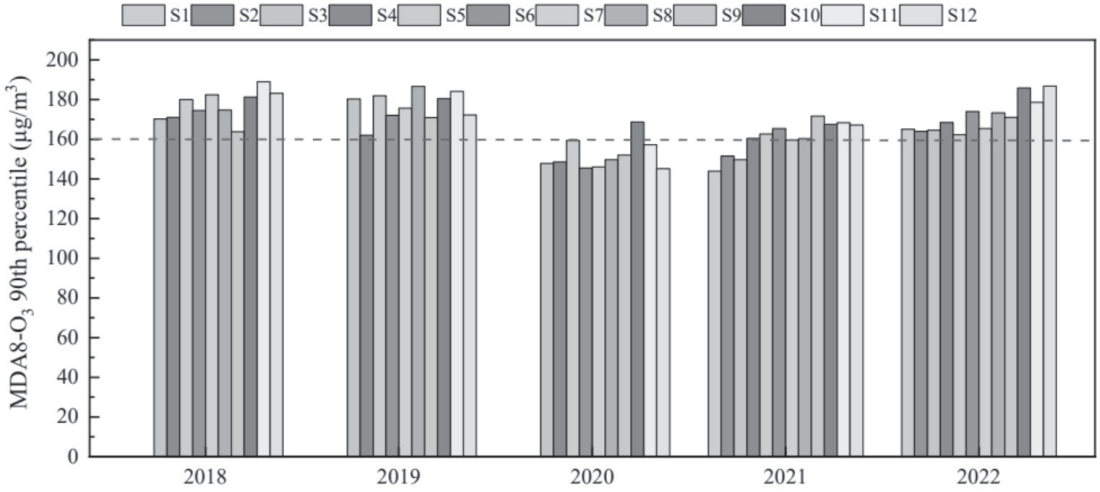

Fig. 5. MDA8-O_3-90 at stations in Hangzhou.

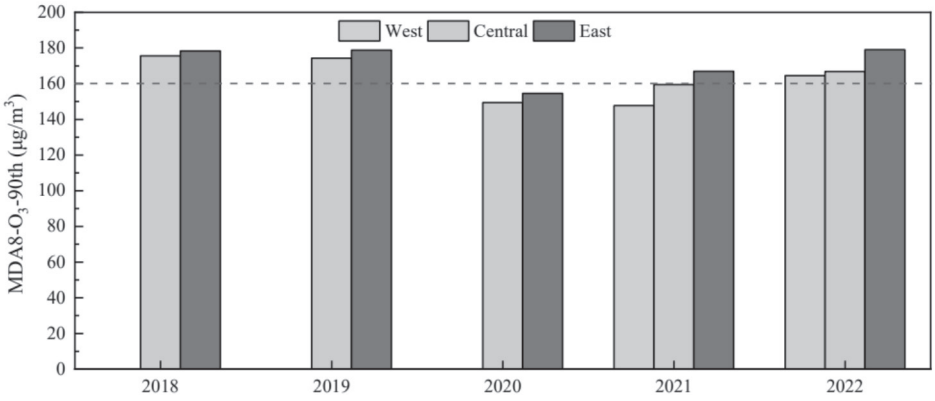

Fig. 6. MDA8-O_3-90 in the west, middle, and east of Hangzhou.

reductions in ozone pollution levels, indicating the potential effectiveness of the implemented control measures. The eastern sector exhibited relatively stable pollution patterns but consistently demonstrated comparatively higher pollution levels than those in other regions. Notably, the COVID-19 pandemic resulted in a reduction in human activities and industrial emissions in Hangzhou during 2020 and 2021. This decline coincided with a general downward trend in ozone pollution levels observed at 12 monitoring stations across the city during the same period, thereby highlighting the short-term environmental quality improvements attributable to pandemic containment measures.

Comprehensive spatial analysis revealed significant geospatial disparities in ozone pollution across Hangzhou, with the western sector consistently demonstrating higher ozone levels compared to the central and eastern sectors. This regional disparity not only reflects the distribution of pollution sources but also highlights the influence of varying geographical environments on air quality.

Analysis of Ozone Pollution Time Characteristics

Monthly Variation

Systematic analysis of the monitoring data revealed a bimodal monthly ozone variation pattern in Hangzhou, as depicted in Fig. 7. This characteristic distribution can be attributed to the combined effects of seasonal meteorological conditions and human activity patterns. Coordinated analysis of data from 12 ambient monitoring stations revealed two distinct annual ozone peaks occurring during the transition from spring to summer (May-June) and at the onset of autumn (September-October), with a suppression phase observed from June to July.

Hangzhou's ozone levels exhibit pronounced seasonal variations, with temperature and solar radiation increasing significantly from May to June. These factors drive enhanced photochemical reactions, leading to increased O_3 formation. From June to July, the plum rain season leads to reduced sunlight duration and high humidity, which suppresses O_3 formation, thereby

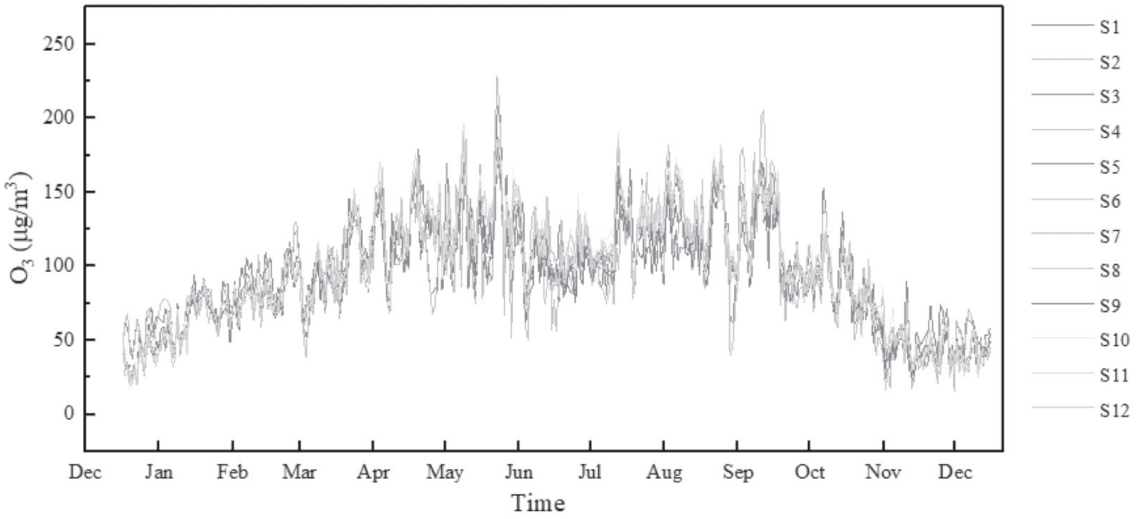

Fig. 7. Monthly variation characteristics of ozone pollution in Hangzhou.

resulting in lower concentrations. However, O_3 levels remain higher compared to winter values, possibly due to the persistence of higher temperatures and chemical processes triggered by sunlight. A secondary ozone peak occurs from September to October, driven by the sufficient sunlight typical of autumn and lower humidity levels that sustain photochemical activity. Following November, O_3 concentrations decrease significantly as meteorological conditions become less favorable. Research demonstrates that temperature and solar radiation serve as the primary factors influencing ozone concentrations, providing a scientific basis for the development of seasonal pollution control strategies.

Daily Variation

Analysis of the daily variation characteristics of ozone pollution in Hangzhou, according to collected monitoring data, as depicted in Fig. 8, reveals a distinct diurnal pattern characterized by a typical unimodal curve. Analysis of the monitoring data reveals that the lowest O_3 concentrations typically occur from 06:00 to 08:00, while peak concentrations are observed from 15:00 to 16:00.

The diurnal variation of O_3 concentrations exhibits a significant photochemical-driven pattern. During the early morning period (0:00-8:00), photochemical reactions were suppressed by the absence of sunlight and low temperatures, remaining relatively weak. Meanwhile, the continuous titration effect of NO_x led to persistent consumption of O_3, maintaining its concentration at a low level. During the daytime (8:00-16:00), photochemical reactions were significantly intensified by the increased solar radiation and rising temperatures. Emissions of NO_x and VOCs from transportation and industrial activities serve as significant precursors for ozone formation, leading to peak concentrations in the mid-afternoon (15:00-16:00) due to the synergistic effects of maximum solar irradiance and high temperatures. During nighttime (16:00-24:00), the restricted vertical dispersion of pollutants caused by a lowered boundary layer and stable meteorological conditions, combined with reduced solar radiation and temperature decreases, results in lower ozone formation rates and a subsequent gradual decline in concentrations. This diurnal variation pattern clearly reflects the coupling interactions among meteorological conditions, precursor emissions, and photochemical processes.

Weekend Effect

Analysis of monitoring data revealed a pronounced weekend ozone enhancement phenomenon in Hangzhou, as depicted in Fig. 9. Elevated ozone concentrations were observed during weekends (Saturday and Sunday) compared to weekdays (Monday through Friday), indicating a significant weekend effect. This observed anomaly contradicts conventional expectations, since weekends witness reduced industrial production and vehicular traffic, which should theoretically lead to a decrease in ozone precursor emissions. Paradoxically, elevated weekend ozone concentrations persist, potentially attributable to modified anthropogenic activity patterns and meteorological variability, among other factors.

The analysis of emission characteristics reveals significant weekend variations in pollutant discharge patterns. While stationary emission sources, such as industrial production and freight transportation, demonstrate reduced activity levels, mobile and residential sources exhibit distinct behavioral adaptations. For example, the residential emissions of VOCs have significantly increased, encompassing activities such as the use of household cleaning solvents, the volatilization of decorative coatings, and outdoor barbecuing. These activities release highly reactive VOC species, including olefins and aromatic hydrocarbons; the transformation of vehicular travel patterns, characterized by a decrease in commuter traffic and an

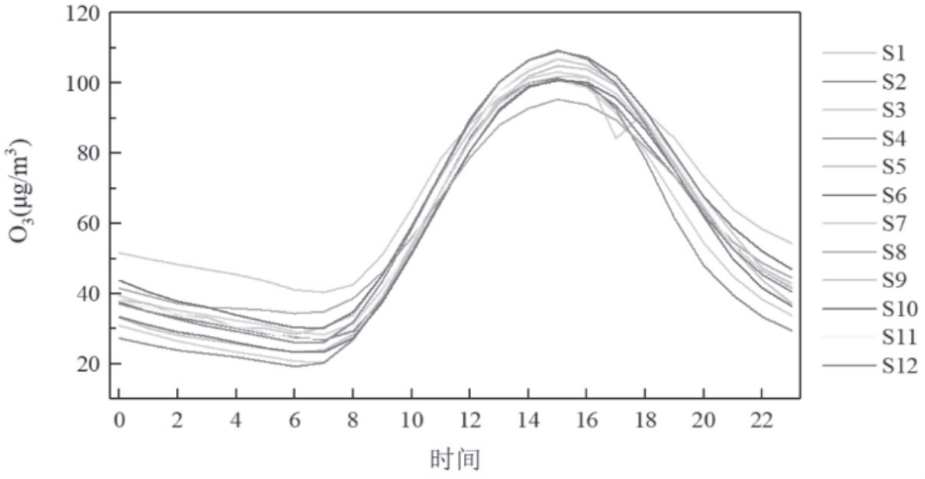

Fig. 8. Daily variation characteristics of ozone pollution in Hangzhou.

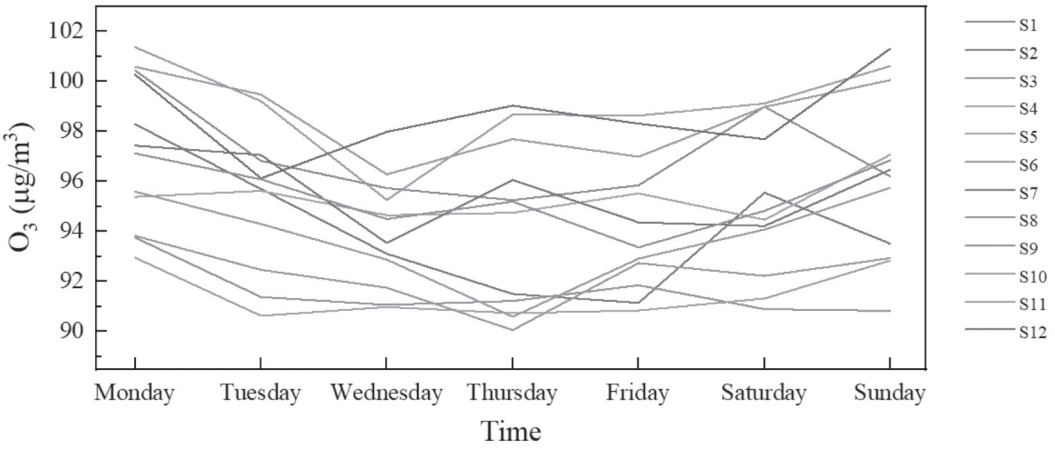

Fig. 9. Weekly variation characteristics of ozone pollution in Hangzhou.

increase in leisure trips, has resulted in the formation of localized emission hotspots for NO_x and VOCs around scenic and commercial areas. This shift in emission structure increases the VOCs/NO_x ratio, transitioning ozone formation mechanisms from being NO_x-limited to becoming VOCs-dominated. From a photochemical perspective, the altered emission characteristics have significantly influenced ozone formation: The increased VOCs/NO_x ratio enhances ozone production efficiency, while the reduction in NO_x emissions weakens the titration effect of NO on ozone. This enables the previously accumulated VOCs to continuously engage in photochemical reactions, thereby exacerbating ozone accumulation.

However, at specific monitoring stations (S1 and S2), while ozone concentrations increased on Saturday, a significant decrease was observed on Sunday. These stations are uniformly characterized by their geographical location farther from urban centers, which may subject them to stronger influences from atmospheric circulation patterns and pollutant dispersion mechanisms.

Sensitivity Analysis of Influencing Factors of Ozone Pollution

Averaged data (2018-2022) collected from monitoring stations S3-S12 were modeled using two distinct XGBoost configurations: a baseline implementation and its Bayesian-optimized counterpart, to fit the data. The complete dataset was systematically partitioned into training and testing subsets at a 7:3 allocation ratio, followed by standardized preprocessing. The BO framework was strategically employed for surrogate modeling to explore the hyperparameter space of the XGBoost architecture, achieving optimal configurations. The quantitative characterization of the comparative hyperparameter landscapes before and after optimization is presented in Table 1. The two models were evaluated, and the results of each evaluation metric are summarized in Table 2.

The optimized model demonstrated superior performance compared to the original model, with an increase in R^2 of 1.40%, a reduction in RMSE by 5.92%, and a decrease in MAE by 3.33%, indicating that BO effectively enhances the accuracy of the XGBoost model. In summary, the BO-XGBoost-based model demonstrates higher accuracy and lower error margins, indicating superior performance compared to the original model. Subsequent analysis, utilizing the optimized model's predictions, will explore the underlying patterns of ozone pollution.

To comprehensively investigate the impact patterns of meteorological factors and other pollutant indicators on ozone pollution in Hangzhou, this study focuses on analyzing the ozone pollution mechanisms associated with key influential factors – T, RH, $PM_{2.5}$, and NO_2 – based on the feature importance rankings derived from the BO-XGBoost model, as depicted in Fig. 10. This analysis aims to clarify the variation patterns of ozone concentrations under different driving factors.

Table 1. Initial and optimized values of hyperparameters in XGBoost model.

Regression model	Parameter	Initial value	Optimized value
XGBoost	learning_rate	0.01	0.04
	max_depth	5	4
	n_estimators	100	365

Table 2. Evaluation of XGBoost and BO-XGBoost models.

Model	Evaluating indicator		
	R^2	RMSE	MAE
XGBoost	0.8089	0.1908	0.3304
BO-XGBoost	0.8202	0.1795	0.3194

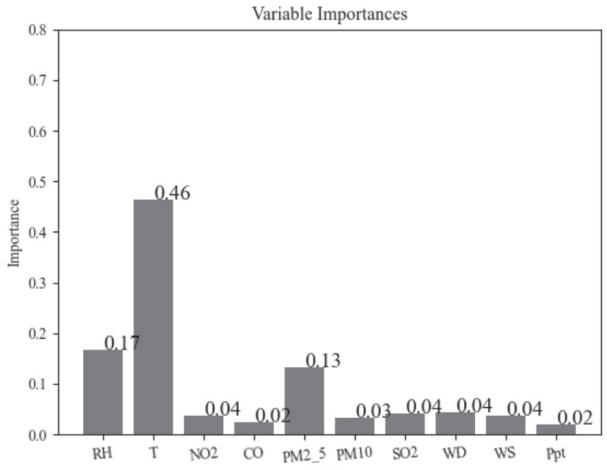

Fig. 10. Factor importance of BO-XGBoost model output.

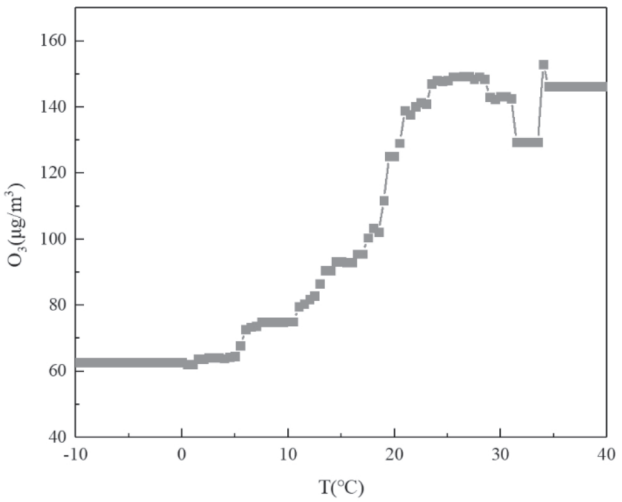

Fig. 11. Ozone versus temperature curve.

Temperature

The BO-XGBoost model was utilized to predict ozone concentration variations across temperature gradients. Model efficacy was ensured by maintaining fixed input parameters during training, with continuous variables set at their mean values and WD controlled using the most frequent categorical value. During the prediction phase, temperature was systematically perturbed in 0.5ºC increments to elucidate the dynamics between ozone and temperature, with the resultant patterns quantitatively depicted in Fig. 11. Analysis revealed a general positive correlation between temperature elevation and ozone accumulation. However, a significant decline in concentration was observed within the temperature range of 30-33ºC. This phenomenon demonstrates that the ozone-temperature relationship goes beyond simplistic linear assumptions, revealing a more complex and nuanced dynamic. A critical transition threshold was identified at 16ºC, beyond which ozone levels increased abruptly. These findings demonstrate complex nonlinear dynamics in the relationships between ozone and temperature, characterized by distinct dormant, transition, and active phases rather than monotonic progression. Within specific temperature ranges, the response of ozone concentration to temperature variations demonstrates nonlinear characteristics, necessitating more nuanced consideration of how temperature influences ozone formation.

When the ambient temperature increased from 16ºC to 24ºC, a significant rise in ozone concentration was observed. The sensitivity of ozone concentration variations to temperature was quantified by computing relative sensitivity coefficients. At this stage, the initial temperature was set to 16ºC with a variation of ±8ºC, while the ozone concentration had an initial value of 95.5 μg/m^3 and a variation of ±51.4 μg/m^3. According to these data, the relative rate of change in ozone concentration attributed to temperature was calculated to be 1.01.

The temperature-ozone relationship is elucidated through multiple critical mechanisms. Firstly, elevated temperatures are intrinsically linked to intensified solar radiation, which significantly promotes ozone formation under such meteorological conditions. Enhanced solar radiation supplies sufficient activation energy to drive photochemical reactions, thereby accelerating the formation of ozone. Secondly, the thermal decomposition of peroxyacetyl nitrate (PAN) under warmer conditions releases precursors such as NO_2, which significantly enhances the potential for ozone production. Furthermore, elevated temperatures can increase the emission of naturally occurring VOCs such as isoprene. These compounds can participate in photochemical reactions under sunlight, thereby promoting the formation of ozone. The synergistic interactions among these mechanisms collectively contribute to elevated ozone concentrations, thereby exacerbating the frequency and intensity of heavy pollution episodes.

Relative Humidity

The trained BO-XGBoost model was employed to predict ozone concentration variations at various RH levels. To investigate the humidity-dependent ozone formation dynamics, RH was systematically varied in increments of 0.5% while other influential parameters were maintained at their mean values during the pollution episode. For the WD feature, the most frequent WD was selected to ensure these variables remained unchanged throughout the forecasting period. As depicted in Fig. 12, the model outputs demonstrate a relationship between RH and ozone accumulation.

Analysis revealed a nonlinear relationship between ozone concentration and RH, characterized by threshold-dependent behavior that varies at different humidity levels. A significant inverse correlation emerged when RH exceeded 55%. Concentration levels exhibited

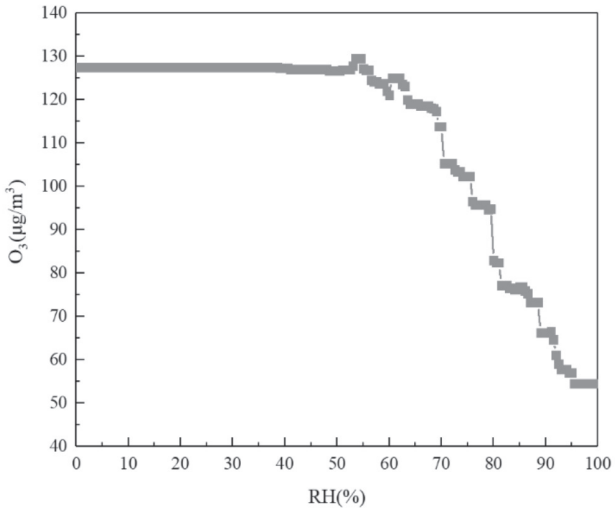

Fig. 12. Ozone versus relative humidity curve.

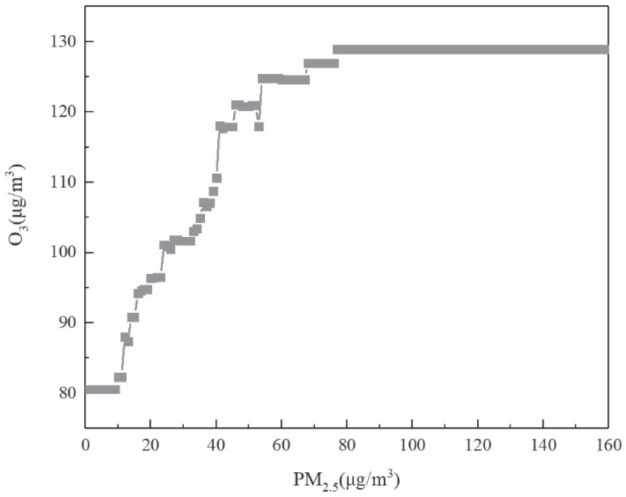

Fig. 13. Variation curve of ozone to $PM_{2.5}$

a sharp decline when the RH increased to approximately 69%. The sensitivity of ozone concentration to variations in RH was quantified by analyzing the relative change rate during the abrupt depletion phase. During this phase, the RH was initially set at 69% with a variation of ±12%, while the corresponding ozone concentration had an initial value of 113.8 μg/m^3 and a variation of -36.9 μg/m^3. Calculations revealed a relative change rate of -1.86 for ozone concentration with respect to RH. The impact of RH on ozone pollution can be attributed to the following mechanisms. First, elevated RH is frequently associated with an increased probability of precipitation. During precipitation events, wet deposition efficiently scavenges ozone molecules from the atmosphere, leading to a rapid decrease in ozone concentrations. Second, higher RH increases atmospheric water vapor content, which suppresses ozone formation by impeding key photochemical reactions. Third, the increase in RH leads to elevated hydroxyl radical concentrations, accelerating the process of ozone depletion. Finally, elevated RH promotes the conversion of NO_2 to HNO_3, a process that not only suppresses ozone formation but also amplifies the extent of its reduction.

$PM_{2.5}$

The influence of $PM_{2.5}$ variations on ozone concentrations was assessed using a pre-trained BO-XGBoost model. To ensure predictive robustness, the covariates were standardized to their mean values during model training, and WD was constrained to its maximum frequency. These parameters were kept constant throughout the prediction phase. Systematic sensitivity analyses were performed by incrementally adjusting $PM_{2.5}$ concentrations in 1 μg/m^3 intervals, as depicted in Fig. 13. Results indicate that elevated $PM_{2.5}$ levels enhance ozone formation in Hangzhou's urban atmosphere. The most significant ozone enhancement occurred when $PM_{2.5}$ concentrations increased from 11 to 40 μg/m^3, leading to an increase in ozone levels from 82.3 μg/m^3 to 117.9 μg/m^3. A normalized sensitivity coefficient of 1.64 was calculated for this regime, indicating a significant role of $PM_{2.5}$ in promoting ozone formation. However, ozone production plateaued when $PM_{2.5}$ concentrations exceeded 77 μg/m^3, indicating a saturation threshold in aerosol-radical interactions.

The observed ozone-$PM_{2.5}$ relationship, characterized by an initial enhancement followed by stabilization, can be attributed to multiple mechanisms. First, the accumulation of $PM_{2.5}$ provides a substantial surface area, which promotes the occurrence of various chemical reactions. Heterogeneous reactions involving a variety of ozone precursors occur on $PM_{2.5}$ surfaces, effectively reducing ozone formation by depleting these precursors. Second, elevated $PM_{2.5}$ concentrations increase aerosol optical depth, leading to alterations in light conditions and thereby constraining the rates of photochemical reactions. Photochemical processes constitute the primary driver of ozone production, with the radiative attenuation of actinic flux directly regulating ozone formation rates and establishing a dynamic equilibrium in ambient ozone concentrations.

NO_2

The BO-XGBoost model was employed to predict ozone concentration variations caused by NO_2 fluctuations. Non-target covariates were standardized to their mean values during model training, and WD was constrained to its maximum mode as derived. These parameters were rigorously maintained at constant levels throughout the prediction phases to ensure covariate stability. A gradient analysis with 1 μg/m^3 NO_2 concentration intervals was implemented to examine the relationship between ozone pollution patterns and variations in NO_2 concentrations, as depicted in Fig. 14. The findings demonstrate that in the atmospheric

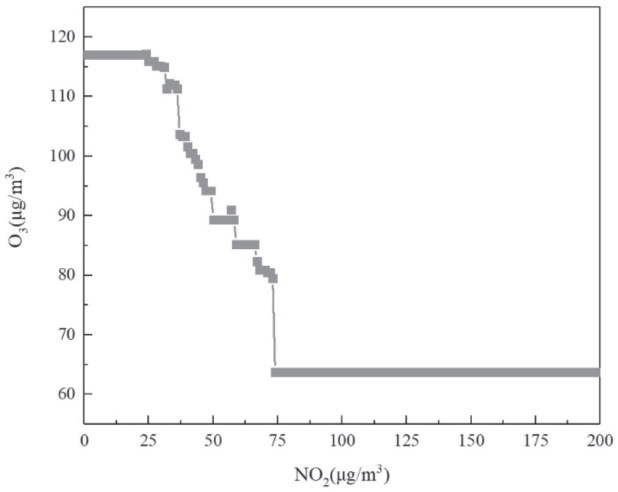

Fig. 14. Ozone versus NO_2 curve.

environment of Hangzhou, the ozone formation process is closely linked to the consumption of NO_2.

As illustrated, the maximum ozone formation rates occurred during the reduction of NO_2 concentrations from 74 to 36 μg/m³. This corresponded to an initial ozone concentration of 63.7 μg/m³ that increased by 47.4 μg/m³, resulting in a NO_2-ozone relative change ratio of -1.45. Beyond the 74 μg/m³ NO_2 threshold, ozone concentrations exhibited minimal variation despite increased surface emission levels. This plateau effect suggests incomplete NO_2 conversion to ozone. The cycling of NO_2 is regulated by multifaceted environmental factors, and its concentration dynamics do not inherently correlate with proportional variations in ozone levels. Elevated concentrations of NO_2 initiate chemical interactions with coexisting atmospheric components, leading to the formation of diverse chemical species and further transformation into other chemical forms, thereby influencing ozone formation. Furthermore, elevated NO_2 levels can potentially reduce the efficiency of photochemical cycling and decrease ozone production rates. Consequently, the relationship between NO_2 and O_3 concentrations manifests non-linear dynamics influenced by multifactorial drivers.

Conclusions

This study aimed to analyze the spatiotemporal characteristics and influencing factors of ozone pollution in Hangzhou, addressing the knowledge gap in understanding its regional heterogeneity and precursor sensitivity. By integrating machine learning and sensitivity analysis, we successfully elucidated the spatial gradients and temporal patterns of O₃ concentrations. Furthermore, our findings revealed critical thresholds for meteorological and pollutant parameters that drive O₃ formation, addressing the specific research objectives outlined in the Introduction. The BO-XGBoost model's high predictive accuracy validates its efficacy in deciphering complex pollution mechanisms, thereby filling the gap in precise O_3 prediction and control strategies for the region. The research results will help develop more targeted strategies for ozone pollution prevention and control, such as strengthening emission reduction measures in specific seasons and time periods and implementing differentiated pollution control policies based on the characteristics of different regions. The main conclusions are as follows:

(1) The ozone concentration in Hangzhou shows significant spatial distribution differences, with lower concentrations in the western region and higher concentrations in the eastern region. The MDA8-O_3-90 and the proportion of days with severe ozone exceedance both exhibit a fluctuating upward trend from S1 to S12 monitoring stations (western to eastern Hangzhou).

(2) Regarding the temporal characteristics of ozone concentration in Hangzhou city, in terms of temporal distribution: the annual variation in ozone concentration exhibits a distinct bimodal pattern with dual peaks occurring during May to July (summer season) and August to October (autumn season), while its diurnal variation demonstrates a typical unimodal pattern. The lowest values are observed between 06:00 and 08:00, while peak concentrations occur daily between 15:00 and 16:00, with a pronounced weekend effect noted.

(3) Based on predictions from the BO-XGBoost model, the key factors influencing ozone concentration variations are T, RH, $PM_{2.5}$, and NO_2. Temperature and $PM_{2.5}$ exhibit a positive correlation with ozone concentration, while RH and NO_2 demonstrate a negative correlation with ozone levels. A pronounced increase in ozone concentration is observed when temperatures rise to approximately 16ºC, with a temperature-ozone relative change rate of 1.01; a sharp decline in ozone concentration is observed when RH increases to approximately 69%, corresponding to a relative change rate of -1.86 between RH and ozone concentration; the ozone concentration demonstrates the most significant increase as $PM_{2.5}$ levels rise from 11 μg/m³ to 40 μg/m³, corresponding to a $PM_{2.5}$-ozone relative change rate of 1.64; the most significant increase in ozone concentration occurs when NO_2 levels decrease from 74 μg/m³ to 36 μg/m³, corresponding to a NO_2-ozone relative change rate of -1.45. Therefore, temperature is identified as the most influential factor based on model feature importance analysis, while RH is determined to be the most sensitive factor associated with ozone concentration variations according to the relative change rate analysis of influencing factors and ozone.

(4) Compared with the original XGBoost model, the BO-XGBoost model showed an increase in R2 of 1.40%, a decrease in RMSE of 5.92%, and a decrease in MAE of 3.33%, indicating that the BO-XGBoost model has higher accuracy and lower errors.

Regarding limitations, they generally include: 1) The manuscript only analyzed NO_2 and $PM_{2.5}$. Some precursors lack VOCs data to improve the study

of ozone generation mechanisms; 2) Based on the abnormal weekend ozone concentration discovered in the manuscript, in future research, changes in emission sources can be further analyzed to provide a basis for targeted control; 3) Hangzhou is located in the Yangtze River Delta, and the manuscript does not specify the impact of regional transmission. In the future, modeling and analysis of pollution inputs from surrounding cities are needed to support cross-regional policy formulation.

Acknowledgments

This work was supported by the Study on the Spatiotemporal Characteristics and Formation Patterns of Ozone Pollution in Hangzhou (LAZD202504).

Conflict of Interest

The authors declare no conflict of interest.

References

1. CAKMAK S., HEBBERN C., PINAULT L., LAVIGNE E., VANOS J., CROUSE D.L., TJEPKEMA M. Associations between long-term $PM_{2.5}$ and ozone exposure and mortality in the Canadian Census Health and Environment Cohort (CANCHEC), by spatial synoptic classification zone. Environment International. **111**, 200, **2018**.
2. WANG F., QIU X., CAO J., PENG L., ZHANG N., YAN Y., LI R. Policy-driven changes in the health risk of $PM_{2.5}$ and O_3 exposure in China during 2013-2018. Science of the Total Environment. **757** (7794), 143775, **2021**.
3. GUAN Y., XIAO Y., CHU C., ZHANG N., YU L. Trends and characteristics of ozone and nitrogen dioxide related health impacts in Chinese cities. Ecotoxicology and Environmental Safety. **241**, **2022**.
4. WHITE M.C., ETZEL R.A., WILCOX W.D., LLOYD C. Exacerbations of childhood asthma and ozone pollution in Atlanta. Environmental Research. **65** (1), 56, **1994**.
5. WESCHLER C.J. Roles of the human occupant in indoor chemistry. Indoor Air. **26** (1), 6, **2016**.
6. DU C., PEI J., FENG Z. Unraveling the complex interactions between ozone pollution and agricultural productivity in China's main winter wheat region using an interpretable machine learning framework. Science of the Total Environment. **954**, **2024**.
7. LIU B., LI Y., WANG L., ZHANG L., QIAO F., NAN P., JI D., HU B., XIA Z., LOU Z. Evaluating the effects of meteorology and emission changes on ozone in different regions over China based on machine learning. Atmospheric Pollution Research. **16** (3), **2025**.
8. LYU Y., JU Q., LV F., FENG J., PANG X., LI X. Spatiotemporal variations of air pollutants and ozone prediction using machine learning algorithms in the Beijing-Tianjin-Hebei region from 2014 to 2021. Environmental Pollution. **306**, **2022**.
9. DING S., WEI Z., LIU S., ZHAO R. Uncovering the evolution of ozone pollution in China: A spatiotemporal characteristics reconstruction from 1980 to 2021. Atmospheric Research. **307** (33), 107472, **2024**.
10. YAN Y., LIN J., POZZER A., KONG S., LELIEVELD J. Trend reversal from high-to-low and from rural-to-urban ozone concentrations over Europe. Atmospheric Environment. **213**, 25, **2019**.
11. REN J., HAO Y., SIMAYI M., SHI Y., XIE S. Spatiotemporal variation of surface ozone and its causes in Beijing, China since 2014. Atmospheric Environment. **260**,118556, **2021**.
12. LIU Z., QI Z., NI X., DONG M., MA M., XUE W., ZHANG Q., WANG J. How to apply O_3 and $PM_{2.5}$ collaborative control to practical management in China: A study based on meta-analysis and machine learning. Science of the Total Environment. **772**, **2021**.
13. YIN S., LIU H., DUAN Z. Hourly $PM_{2.5}$ concentration multi-step forecasting method based on extreme learning machine, boosting algorithm and error correction model. Digital Signal Processing. **118**, **2021**.
14. OUFDOU H., BELLANGER L., BERGAM A., KHOMSI K. Forecasting Daily of Surface Ozone Concentration in the Grand Casablanca Region Using Parametric and Nonparametric Statistical Models. Atmosphere. **12** (6), **2021**.
15. KUMAR A., GHOSH A.K., IEEE Ensemble Machine Learning Methods for Unsteady Aerodynamics Modeling from Flight Data. Singapore, SINGAPORE, **2024**.
16. LI J., AN X., LI Q., WANG C., YU H., ZHOU X., GENG Y.-A. Application of XGBoost algorithm in the optimization of pollutant concentration. Atmospheric Research. **276**, **2022**.
17. FENG R., LUO K., FAN J.-R. Decoding Tropospheric Ozone in Hangzhou, China: from Precursors to Sources. Asia-Pacific Journal of Atmospheric Sciences. **56** (3), 321, **2020**.
18. WANG M., SHENG H., LIU Y., WANG G., HUANG H., FAN L., YE D. Research on the diurnal variation characteristics of ozone formation sensitivity and the impact of ozone pollution control measures in "2+26" cities of Henan Province in summer. Science of the Total Environment. **888**, **2023**.
19. YU M., ZHU Y., LIN C.-J., WANG S., XING J., JANG C., HUANG J., HUANG J., JIN J., YU L. Effects of air pollution control measures on air quality improvement in Guangzhou, China. Journal of Environmental Management. **244**, 127, **2019**.
20. LIU Z., CHEN X., LU Q. Blowin' in the wind of an invisible killer: long-term exposure to ozone and respiratory mortality in the USA. Journal of Population Economics. **38** (3), **2025**.
21. YANG B.-Y., LIU Y., HU L.-W., ZENG X.-W., DONG G.-H. Urgency to Assess the Health Impact of Ambient Air Pollution in China. In: Ambient Air Pollution and Health Impact in China, G.H. Dong Ed. Volume **1017**, pp.1, **2017**.
22. REAMES T.G., BRAVO M.A. People, place and pollution: Investigating relationships between air quality perceptions, health concerns, exposure, and individual- and area-level characteristics. Environment International. **122**, 244, **2019**.
23. JIAN Z., CAI J., CHEN R., NIU Y., KAN H. A bibliometric analysis of research on the health impacts of ozone air pollution. Environmental Science and Pollution Research. **31** (11), 16746, **2024**.
24. DEWAN S., BAMOLA S., LAKHANI A. Addressing ozone pollution to promote United Nations sustainable development goal 2: Ensuring global food security. Chemosphere. **347**, 140693, **2024**.
25. LI P., YUAN X.-Y., DAI L.-L., FENG Z.-Z. Progress in the Interactions of Ozone Pollution and Other Environmental

Stress on Urban Forests in China. Huan jing ke xue= Huanjing kexue. **42** (11), 5075, **2021**.
26. KAUFMAN L., ROUSSEEUW P.J. Finding Groups in Data: An Introduction to Cluster Analysis. Finding Groups in Data: An Introduction to Cluster Analysis, **2009**.
27. CHEN T., GUESTRIN C., ASSOC COMP M. XGBoost: A Scalable Tree Boosting System. San Francisco, CA, **2016**.
28. WU D.-N., SHI L.-L., XING B. HAN Y., SONG J.P., LV Y., PANG X.-B. Machine learning reveals the driving factors of ozone meteorology in Chinese urban agglomerations. Ecology and Environmental Monitoring of Three Gorges. 1, **2025**.
29. SHEN Z.-B., PANG L., WANG W.D., ZHENG Y., ZHANG G.-C. Research on Ozone Pollution Forecasting Based on Machine Learning: A Case Study of Zhenhai District, Ningbo City. Sichuan Environment. **44** (3), 28, **2025**.