Identifying the geospatial relationship of surface ozone pollution in China: Implications for key pollution control regions
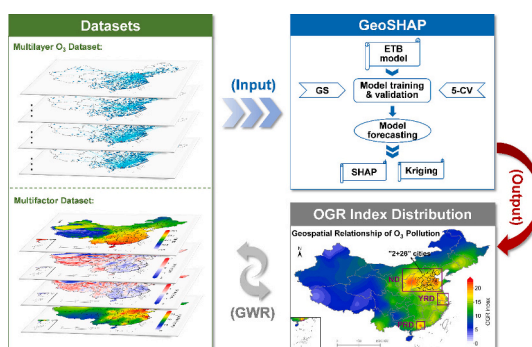
Yong Cheng, Yan Peng, Li-Ming Cao, Xiao-Feng Huang [*], Ling-Yan He

Laboratory of Atmospheric Observation Supersite, School of Environment and Energy, Peking University Shenzhen Graduate School, Shenzhen 518055, China

HIGHLIGHTS

- Introduced a novel tool to explore the geospatial relationship of ozone pollution
- High ozone geospatial relationship stems from anthropogenic and natural synergy
- Revealed key regions needing joint control to reduce ozone pollution in China

GRAPHICAL ABSTRACT

ARTICLE INFO

ABSTRACT

Surface ozone pollution, as a pressing environmental concern, has garnered widespread attention across China. Due to air mass transport, effective control of ozone pollution is highly dependent on collaborative efforts across neighboring regions. However, specific regions with strong internal interactions of ozone pollution are not yet well identified. Here, we introduced the Geospatial SHapley Additive exPlanation (GeoSHAP) approach, which primarily involves machine learning and geostatistical algorithms. Based on extensive atmospheric environmental monitoring data from 2017 to 2021, machine learning models were employed to train and predict ozone concentrations at the target location. The R^2 values on the test sets of different scale regions all reached 0.98 in the overall condition, indicating that the core model has good accuracy and generalization ability. The results highlight key regions with high ozone geospatial relationship (OGR) index, predominantly located in the Northern District (ND), spanning the Fen-Wei Plain, the Loess Plateau, and the North China Plain, as well as within portions of the Yangtze River Delta (YRD) and the Pearl River Delta (PRD). Further investigation indicated that high geospatial relationships stem from a synergy between anthropogenic and natural factors, with anthropogenic factors serving as a pivotal element. This study revealed key regions with the most urgent need for joint control of anthropogenic sources to mitigate ozone pollution.

1. Introduction

Surface ozone, a typical secondary pollutant that poses risks to human health and vegetation ecology (Monks et al., 2015; Agathokleous et al., 2020; Unger et al., 2020; Feng et al., 2022), forms through photochemical reactions of volatile organic compounds (VOCs) and nitrogen oxides (NOx) (Jacob, 2000). This process involves a complex atmospheric formation mechanism influenced by various factors, including natural processes and anthropogenic activities (Shindell et al., 1998; Wang et al., 2017; Li et al., 2019; Lu et al., 2019). Currently, China is confronted with substantial challenges regarding surface ozone pollution, which frequently extends beyond local boundaries, exhibiting distinctive regional and interregional characteristics (Li et al., 2014; Xue et al., 2020; Li et al., 2022; Wang et al., 2023c). Previous studies have indicated that such regional characteristics of ozone pollution are influenced by the proximity of source and receptor regions, East Asian monsoon circulations, typhoons, and topography (Shen et al., 2022; Wang et al., 2022). Furthermore, atmospheric transport has been found to significantly modulate regional interactions of ozone pollution in China (Shen et al., 2022). Consequently, effective control of ozone pollution heavily depends on collaborative prevention and control measures involving multiple areas (Guo et al., 2023; Lyu et al., 2023; Yao et al., 2023). A scientific understanding of the geospatial relationship of ozone pollution, and an exploration of the primary influencing factors, hold paramount significance for the mitigation policy of air pollution.

Currently, chemical transport models (CTMs) based on emission inventory data are one of the primary models for exploring regional air pollution concerns (Huang et al., 2019; Weng et al., 2023). Nevertheless, limitations in the timeliness and uncertainty of emission inventory data further constrain the performance of CTMs (Xing et al., 2020; Chen et al., 2023; Wu et al., 2023). Such constraints pose challenges in accurately simulating atmospheric regional pollution under current conditions, particularly for long-term and large-scale. In recent years, China's ground-based atmospheric pollution monitoring system has witnessed continuous refinement and enhancement. Due to air pollution data from ground-based monitoring more directly and accurately representing regional pollution conditions, an increasing number of researchers resort to traditional statistical methods for the analysis and exploration of data from ground-based monitoring stations (Lu et al., 2018; Li et al., 2020; Fan et al., 2021; Wang et al., 2023a). Nevertheless, these traditional statistical methods encounter challenges in revealing the inherent relationships of atmospheric pollution between regions. Given the rapid advancements in computer technology and big data capabilities, we contend that artificial intelligence (AI) represents a promising avenue to transcend these current limitations (Bi et al., 2023). Machine learning (ML) models, as integral components of artificial intelligence, are exceptionally suited to elucidating intricate, nonlinear relationships between different variables (Cheng et al., 2019; Zahrt et al., 2019; Shrock et al., 2020; Ogata et al., 2021; Rybarczyk and Zalakeviciute, 2021; Zhong et al., 2021; Xu et al., 2023). Moreover, they can complement CTMs effectively (Huang et al., 2021a; Chen et al., 2023; Wang et al., 2023c). However, to the best of our knowledge, it is still difficult to define the geospatial relationship of air pollution between areas at the mechanistic level, and there are no studies using ML-related methods to systematically explore the geospatial relationship of air pollution.

Here, we introduce the GeoSHAP method for the first time, employing the concept of the ozone geospatial relationship (OGR) index to quantify the geospatial relationship of ozone pollution. The GeoSHAP method comprises a combination of the ensemble tree-based (ETB) model, SHAP (Shapley, 1953; Lundberg et al., 2020), and the kriging algorithm (Bayraktar and Turalioglu, 2005). Notably, the ETB model, an algorithm we previously introduced, primarily relies on tree-based benchmark models, specifically including XGBoost, LightGBM, CatBoost, and Random forest. In practical tasks, these models can conduct optimal model training and prediction based on the characteristics of the target data (Cheng et al., 2023). Compared to other ML models, such as deep learning models, it maintains a commendable balance between predictive performance and interpretability (Cheng et al., 2021; Peng et al., 2023). Furthermore, because the GeoSHAP method primarily relies on statistical theory and does not involve intricate chemical reaction mechanisms, it offers significant advantages over traditional mechanism-based models in terms of computational efficiency and model generalization. Our investigation also delved into the influence of various factors on the geospatial relationship of ozone pollution, leading to a fresh understanding of the interplay of ozone pollution among different regions. Moreover, based on the OGR index, we revealed the key regions with the most urgent need for joint control of anthropogenic sources to mitigate ozone pollution in China.

2. Materials and methods

2.1. Data source

We have collected three distinct types of data in China: meteorological reanalysis data, source emission inventory data, and ground station monitoring data. In terms of meteorological data, we collected surface wind speed, surface temperature, and surface pressure data spanning from 2017 to 2021 from the ERA5 reanalysis data (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form, last access: Aug. 2023). Concerning source inventory data, we collected NOx and VOCs emission inventory data for the years 2017 to 2020 from the MEIC (http://meicmodel.org.cn/?page_id=541&lang=en, last access: Aug. 2023) (Li et al., 2017; Zheng et al., 2018). Regarding ground station monitoring data, the concentrations of O_3 and NO_2 used in our study were based on hourly ground-based data, and have all been uniformly converted to micrograms per cubic meter under reference conditions (298.15 K, 1013.25 hPa). This data was collected by the China National Environmental Monitoring Center network from 1597 monitoring stations across mainland China (the actual number of monitoring stations is increasing year by year, and the 1597 stations refer to the number of stations involved in our study. For the specific distribution of these stations, please refer to Fig. S1).

It's essential to highlight that ground station data may encounter certain challenges, including the presence of invalid values or outliers, stemming from instrument calibration or other circumstances. Therefore, we initiated thorough data preprocessing to generate the improved version of the fundamental dataset necessary for our research. Further, for better analysis, we employed an ordinary kriging algorithm with a spherical variogram model (Supporting Information S1.1) to interpolate the improved version data to obtain the final dataset (Yu et al., 2019; Pulliainen et al., 2020; Chossière et al., 2021). Furthermore, it is important to mention that among all the collected datasets, O_3 data were used for model training and prediction, while other data, such as meteorological variables and NO_2, were primarily used to investigate their impact on the geospatial relationship of ozone pollution.

2.2. Overview of the OGR index

In this study, we have introduced a novel concept of OGR index for quantifying the geospatial relationship of ozone pollution based on environmental big data and the GeoSHAP method from a statistical perspective. While traditional methods such as correlation analysis and pollutant analysis can identify ozone pollution hotspots and seasonal changes, they often neglect the complex interactions and nonlinear characteristics within the geospatial data of ozone pollution. The use of the GeoSHAP method enables us to account for the effects of geographical location factors on pollution distribution and the mutual influences of ozone pollution across different areas while processing large-scale environmental data. This approach allows us to quantify and visualize the relationships of ozone pollution across different

geographical locations, thereby revealing the intricate geospatial relationships and patterns within regional pollution interactions. Additionally, this provides new tools and insights for understanding and addressing complex environmental challenges in the field of atmospheric pollution research.

A detailed introduction to the concept of OGR is as follows. At the individual point level, the OGR index for a specific point signifies the intensity of its impact on the ozone concentrations in other points. A higher OGR index indicates a stronger impact on the ozone concentration of surrounding points, while values closer to 0 suggest a smaller impact on the nearby points. On the regional level, this involves numerous points where points interact with and influence each other, forming distinct feature regions. When adjacent regional points have the same or similar OGR index, it indicates a strong geospatial relationship. Conversely, significant differences in the OGR index between regional points suggest a weaker geospatial relationship. If the average OGR index of adjacent regional points is high and exhibits a strong geospatial relationship, we classify this area as a high OGR area or a high geospatial relationship of ozone pollution area. Such regions are also the areas of primary focus in our study. From a regional statistical perspective, the concept of OGR emphasizes uncovering commonalities in mutual influences between regions and offers a fresh quantification of the geospatial relationships of ozone pollution among different areas.

The framework of OGR index is depicted in Fig. 1. Initially, we collected and preprocessed ozone data from the period 2017–2021. Subsequently, we dynamically transformed the dataset into an n-dimensional format, with 'n' representing the number of stations. This transformation is based on the relationships between independent variables and the target variable, where one station's data serves as the target variable, and others act as independent variables. Based on the first law of geography proposed by Waldo Toble (Tobler, 1970), "everything is related to everything else, but near things are more related than distant things", and considering that it is difficult to properly and scientifically define the optimal boundary of the target variable, we ultimately input all the ground stations data into the GeoSHAP method for calculation, resulting in the distribution map of OGR index for China at a resolution of $0.1°×0.1°$. It is noteworthy that, given surface ozone is a secondary product of daytime photochemical reactions, our primary focus for analysis and discussion is on the daytime period (8:00–19:00).

2.3. GeoSHAP method

The GeoSHAP method primarily involves machine learning and geostatistical algorithms. The core of the model-related information can be found in Supporting Information S1.2. The following is the detailed process of the GeoSHAP method:

(1) Input the preprocessed ozone data (which constitutes 80 % of the total dataset) into the ETB model in sequence for model training and validation. The hyperparameter settings of the tree-based benchmark models in the ETB model are mainly based on the grid search (GS) method (Supporting Information S1.3). In addition, 5-fold cross-validation is used for model training and validation. Finally, according to the performance of the model, the optimal model (R^2 as the main evaluation metric) is selected for prediction.

(2) Use the SHAP algorithm (Supporting Information S1.4) to analyze the prediction behavior of the optimal model. The analytical results can be used to measure the impact of different independent variable stations on the target station (measured by SHAP value).

(3) The above two steps need to be repeated n (number of stations) times, which means that n*4 models (excluding parameter adjustment and cross-validation) need to be established for training, and finally the mutual influence matrix between different stations is obtained. Through matrix calculation, we can finally get the raw OGR value of each station.

(4) For enhanced visualization, we simultaneously scale up the values of all results by n times, and then use an ordinary kriging algorithm with a spherical variogram model, a classical geostatistical algorithm, to interpolate the values and obtain the final OGR dataset.

2.4. GWR model

The geographically weighted regression (GWR) model is a local geospatial regression method that differs from traditional global regression by accounting for geographical space heterogeneity. While global regression treats all observations as belonging to a single model, GWR introduces geospatial variability in model parameters. In essence, the GWR model calculates a distinct set of regression coefficients for each geographical location, whether it's a point or a region. This approach enables the model to capture variations in the relationship between variables across different locations. In the following introduction, we will use the point as this research object to describe. The specific
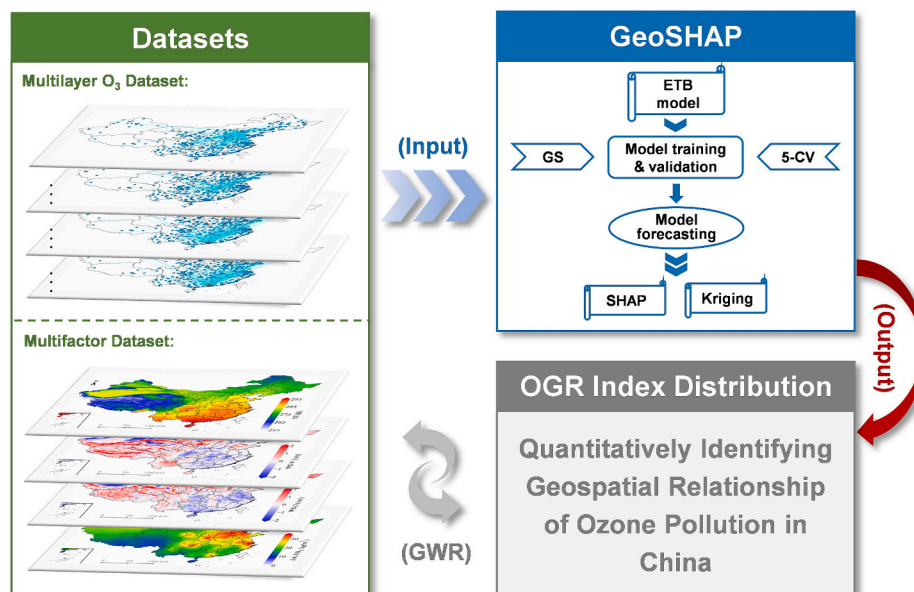

Fig. 1. Overview of the OGR index framework. This demonstrates the comprehensive workflow spanning from data preprocessing to the computation of the OGR index, involving primarily the use of machine learning and geostatistical algorithms.

formula of GWR is as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \qquad (1)$$

where y_i denotes the dependent variable, and i corresponds to a point within the study area; anthropogenic activities (AA), surface temperature (ST), east-west surface wind speed (WS-U) and north-south surface wind speed (WS-V) are the four major driving factors denoted as $x_{ik}(k = 1, 2, 3, 4)$. Within the model, β_{ik} represents the estimated coefficient, and (u_i, v_i) indicates the location of i, while ε_i accounts for random error.

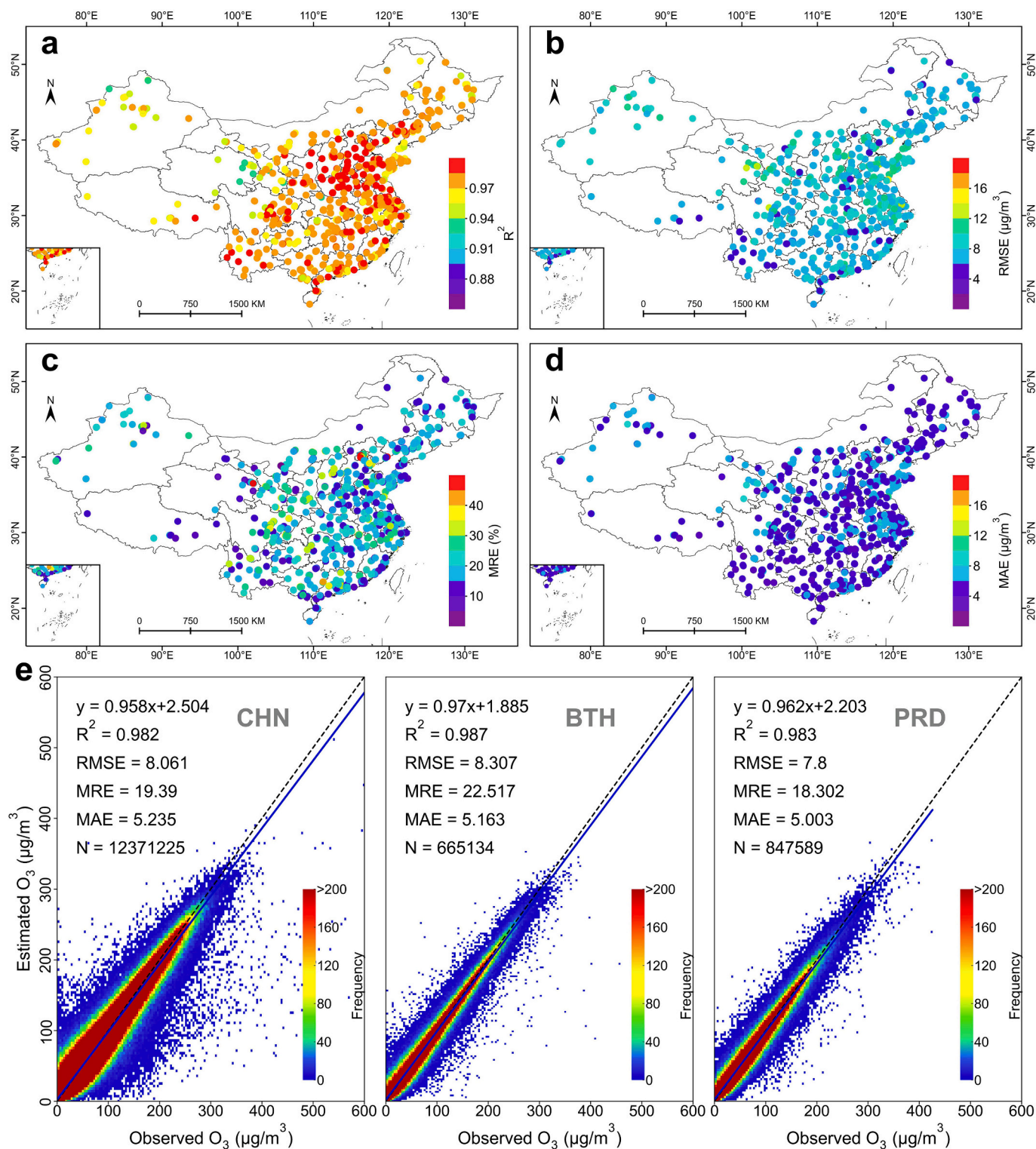In the model, points surrounding a particular study area hold more

Fig. 2. The performance of the core model at various stations across China in the overall condition. **(a)** Distribution of coefficient of determination (R^2). **(b)** Distribution of root mean squared error (*RMSE*). **(c)** Distribution of mean relative error (*MRE*). **(d)** Distribution of mean absolute error (*MAE*). **(e)** Model performance in three representative areas, including mainland China (CHN), BTH, and PRD.

significance. The influence of this effect diminishes gradually with increasing distance between the two points. Consequently, the weight values adhere to a bell-shaped curve based on distance. The weight matrix signifies the relative importance among points. The estimated coefficients can be computed using the subsequent matrix formulation, with T denoting the matrix transpose operation.

$$\widehat{\beta} = \left(X^T W_i X\right)^{-1} X^T W_i Y \tag{2}$$

where $\widehat{\beta}$ represents the estimated coefficient for the target point i; X denotes the independent variable matrix; and X^T signifies the matrix transpose operation applied to the matrix. W_i is the weight matrix of target raster i; Y is the dependent variable.

In this study, we employ an adaptive kernel to determine the optimal number of neighboring points using the Akaike information criterion (AIC), recognized as the most suitable metric for selecting the right number of neighbors for the target point. To calculate the weights between points, we utilize a bi-square weighting function, computed as follows:

$$w_{ij} = \begin{cases} \left[1 - \left(\dfrac{d_{ij}}{d_{max}}\right)^2\right]^2, & d_{ij} \leq d_{max} \\ 0, & d_{ij} > d_{max} \end{cases} \tag{3}$$

w_{ij} represents the influence between position i and position j; d_{ij} denotes the geographical distance between target grid i and surrounding grid j; d_{max} is the farthest point from target point i (m represents the selected

neighboring raster with the optimal number of points). The optimal bandwidth distance can be automatically optimized in the model.

3. Results and discussions

3.1. The performance of the core model at different stations

The pivotal step in the OGR index calculation involves the GeoSHAP method, which significantly depends on the predictive performance of the ETB model. To assess the performance of the ETB model, we utilized test sets comprising 20 % of the total dataset. Comprehensive statistical evaluations were carried out across three representative regions at different scales (Fig. S1): CHN, BTH, and PRD—under different conditions (overall and during ozone pollution events), as depicted in Figs. 2 and S2. The evaluation employed several metrics for model performance, such as R^2, RMSE, MRE, and MAE, with their respective formulas outlined in the Supporting Information S1.5.

As shown in Fig. 2, in the overall scenario, the model performed exceptionally well for the CHN region, with over 98 % of the sites having R^2 values exceeding 0.95. >86 % of the monitoring stations reported RMSE and MRE values below 10 μg/m³ and 27, respectively. Moreover, over 99 % of the stations maintained MAE values under 10 μg/m³. For the BTH and PRD regions, the R^2 values and other metrics generally surpassed the CHN benchmarks. During ozone pollution events (defined as hourly concentration values exceeding 160 μg/m³), as illustrated in Fig. S2, the BTH region demonstrated the best performance with an R^2 of 0.935, followed by the CHN ($R^2 = 0.906$) and PRD regions ($R^2 = 0.897$). Although the MRE values improved compared to the overall condition, a
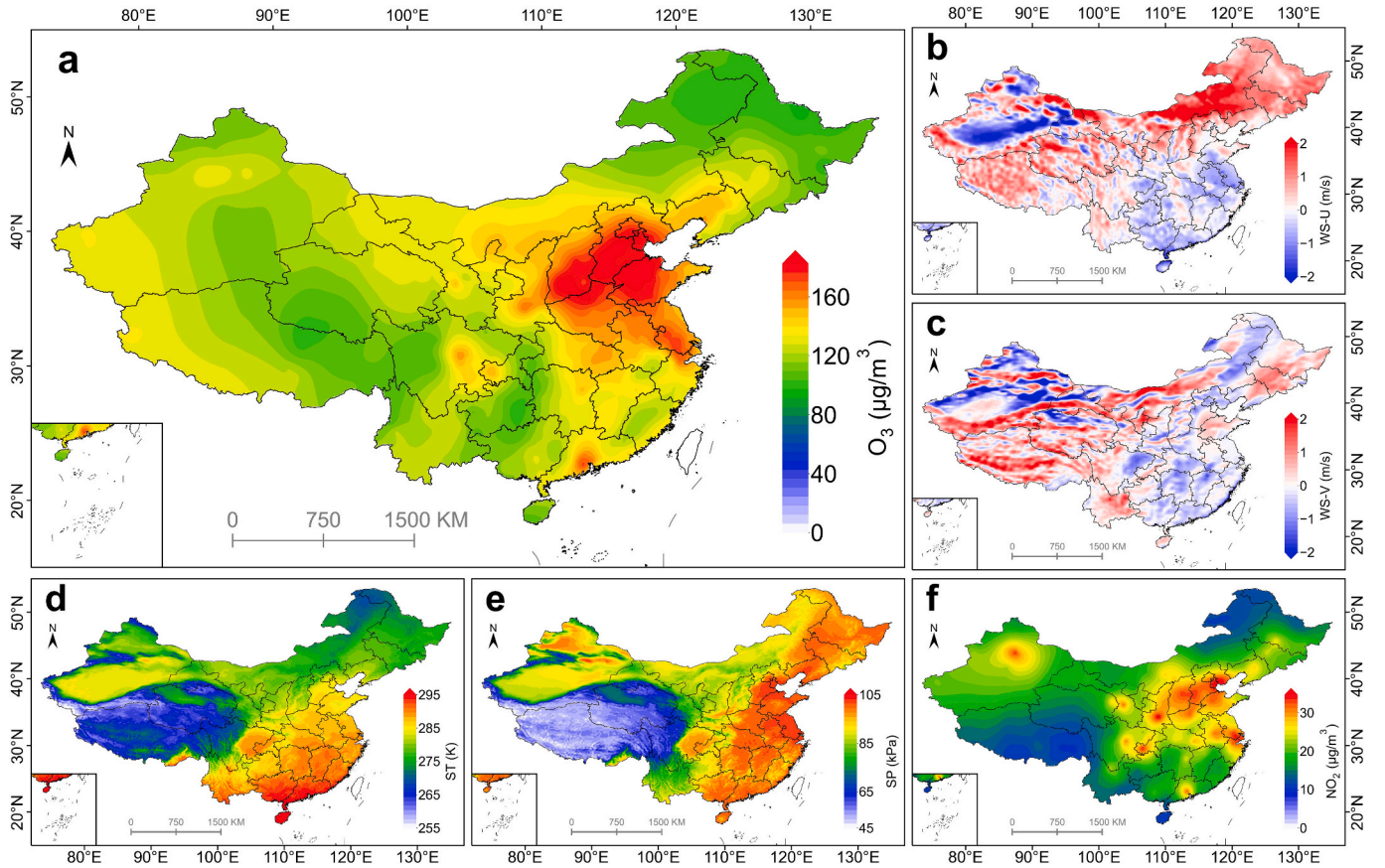

Fig. 3. Geospatial distribution of ozone concentrations and various factors (2017–2021). **(a)** The 90th percentiles of the geospatial distribution of maximum daily 8-hour average (MDA8) ozone concentrations. **(b)** Geospatial distribution of mean daytime east-west surface wind speeds. **(c)** Geospatial distribution of mean daytime north-south surface wind speeds. **(d)** Geospatial distribution of mean daytime surface temperature. **(e)** Geospatial distribution of mean daytime surface pressure. **(f)** Geospatial distribution of mean daytime NO_2 concentrations. In the present study, the daytime period refers to 8:00–19:00, and the characteristics of MDA8 ozone concentrations and various factors are presented in Table S1.

slight decrease in other metrics was observed; nevertheless, the performance remained at a high standard across all metrics.

Overall, the model established in this study demonstrated exceptional accuracy while maintaining a satisfactory level of generalization.

3.2. Geospatial distribution characteristics of ozone pollution

Fig. 3a illustrates the geospatial distribution of the 90th percentile of MDA8 ozone concentrations in China from 2017 to 2021. It is evident that there is significant spatial heterogeneity in the distribution of ozone pollution. High-concentration areas are primarily located in the central and northern regions of China, spanning multiple areas, including the Beijing-Tianjin-Hebei (BTH) region, Shandong, Shanxi, Shaanxi, Henan, as well as the YRD and PRD regions. In contrast, low-concentration areas are predominantly located in western and northern China. Given the potential significant adverse impacts of high ozone concentrations on public health and ecosystems, the high-concentration regions in the figure indicate the presence of potential environmental and health risks.

In the study of ground-level ozone pollution, VOCs, NOx, along with a range of natural factors such as surface temperature, wind speed and direction, are considered key factors influencing the formation and distribution of ozone (Mishra et al., 2021; Zhou et al., 2023). Among these, VOCs and NOx, as crucial precursors to ozone, play an essential role in photochemical reactions (Kumar and Sinha, 2021; Zhu et al., 2021; Xiong et al., 2023). However, obtaining large-scale regional VOC species concentration data faces numerous challenges, such as the complexity or stability of the relevant measuring instruments (gas chromatography–mass spectrometry and proton transfer reaction mass spectrometry), making it difficult to acquire a broad range of usable VOC species concentration data. Furthermore, the dataset published by the China National Environmental Monitoring Center network currently does not include NOx concentrations data. Therefore, this study does not delve into detailed exploration of these two precursors. Nonetheless, considering that most VOC species and NOx pollutant are generated by anthropogenic activities, and the distribution of ground-level NO_2 concentrations can as one index to describe the characterization and intensity of anthropogenic activities to some extent (Laughner and Cohen, 2019; Li et al., 2021; Cooper et al., 2022), we collected and processed daytime average NO_2 data spanning from 2017 to 2021, and the distribution results are shown in Fig. 3f. It is apparent that areas with high concentration hotspots are often characterized by intense economic activities and higher levels of pollution emissions.

By comparing Fig. 3a and f, it becomes evident that there is some overlap between areas with high ozone concentrations and regions with elevated anthropogenic activities, particularly in eastern China. This highlights that in most regions and situations, areas with higher anthropogenic activities tend to experience more ozone pollution, suggesting the pivotal role of anthropogenic activities factors in influencing ozone pollution. Regarding natural factors, considering the potential strong collinearity among different natural factors, we initially selected basic variables that may directly affect the distribution of ozone, such as various wind directions and speeds, surface temperature, and surface pressure, while some factors that have undergone secondary processing, such as ventilation coefficient, have not been included yet. The geographical distribution of different natural factors is shown in Fig. 3b–e. It is difficult to directly observe their relationship with ozone pollution from the figures, but from a mechanistic perspective, natural factors are one of the important aspects that cannot be overlooked. They can not only affect the rate of ozone formation, like surface temperature factors, but also influence the transport of ozone between different regions, such as factors like wind direction, wind speed, and surface pressure.

Current ozone pollution in China exhibits distinct regional distribution characteristics, and regions with similar characteristics are more susceptible to pollutant dispersion and transmission (Li et al., 2014; Xue et al., 2020; Li et al., 2022; Shen et al., 2022; Wang et al., 2022; Wang

et al., 2023c). This is a primary reason for the emphasis by the Chinese government on regional collaborative prevention and control efforts. From Fig. 3a, it can be observed that, aside from several typical high-pollution regions, the central and northern regions of China, as well as the YRD region, seem to be gradually forming a larger high-concentration ozone region. This suggests that not only is ozone pollution likely to occur within individual subregions, but it may also extend between different high-concentration larger areas. However, the precise quantification of the geospatial relationship of ozone pollution across China's diverse subregions and between different mesoscale areas remains a significant challenge.

3.3. Characteristics of ozone geospatial relationship

We calculated the geospatial distribution of daytime average OGR index in various areas of China for the period 2017–2021, as depicted in Fig. 4a. In this figure, areas with similar colors, signifying similar OGR index, indicate a stronger geospatial relationship of ozone pollution in those specific areas. Areas with high OGR index, are primarily depicted in yellow or red in Fig. 4a, mainly in the eastern part of China, and most of these regions located within the HH areas in the map of local indicators spatial autocorrelation cluster analysis (as shown in Fig. 4c), indicating a clear positive spatial autocorrelation in geography (global spatial autocorrelation test has been conducted before, and Moran's I = 0.99 with a p-value of <0.001). Conversely, areas with low OGR index, represented in blue or light white on Fig. 4a, are primarily situated in the western and northernmost regions of China. These areas tend to belong to the LL areas in Fig. 4c, which also demonstrates a positive spatial autocorrelation in geography and showcases distinctive geospatial clustering patterns. These discernible high and low value clustering patterns across various regions suggest that the distribution of OGR index is non-random and likely significantly influenced by specific factors.

By directly comparing the distribution of the OGR index with other related factors (as shown in Fig. 3b–f), it was observed that areas of anthropogenic activity hotspots largely overlap with regions of high OGR index, suggesting a significant correlation between the two. This implies that anthropogenic activities may have a substantial impact on the geospatial relationship of ozone pollution. The outbreak of the COVID-19 pandemic in 2020 and the subsequent urban lockdowns led to a rapid decrease in the intensity of anthropogenic activities in the eastern regions of China (Huang et al., 2020), thereby providing an opportunity to assess the relationship between the intensity of anthropogenic activities and the OGR index. We calculated the distribution of the OGR index during the lockdown period (defined in this study as January 24 to February 18, 2020) in the eastern regions of China, as well as for the corresponding periods in 2019 and 2021. To avoid interference from the "Spring Festival effect", the calculations excluded the Spring Festival period in the lockdown periods of 2019 and 2021 (Huang et al., 2021b), with the final results shown in Fig. S3a. It is evident from the figure that, during the lockdown period, the OGR index in the eastern regions of China significantly decreased compared to the same periods in 2019 and 2021, with most areas in the east nearing an OGR index close to zero. While it's undeniable that the planetary boundary layer (PBL) height in the eastern regions of China decreased to varying degrees during the lockdown, which is not conducive to regional pollution transmission (Le et al., 2020), it is noticeable from the figure that typical areas with relatively high OGR index values, such as in Central and Southern China, almost completely disappeared. This indicates that changes in the intensity of anthropogenic activities can have a significant impact on the geospatial relationship of regional ozone pollution.

From this basic analysis, it is evident that the formation of high OGR index areas requires suitable natural conditions and a higher intensity of anthropogenic activities. In other words, areas with a high OGR index are also those that most urgently requiring joint control to address ozone
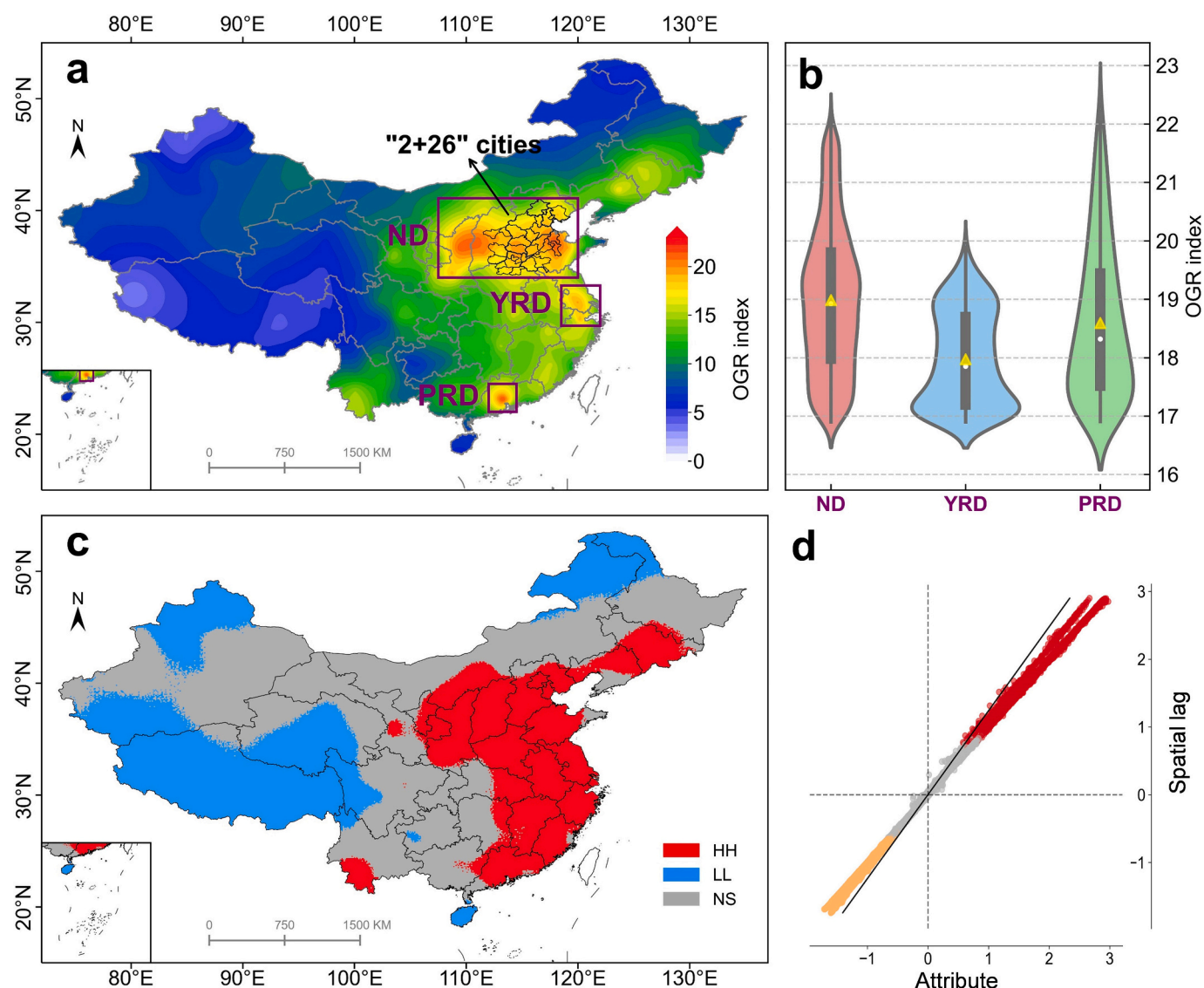
Fig. 4. Geospatial distribution of the OGR index in China and related statistical results. **(a)** Geospatial distribution of mean daytime OGR index. "2 + 26" cities include Beijing and Tianjin municipalities, 8 cities in Hebei Province, 4 cities in Shanxi Province, 7 cities in Shandong Province, and 7 cities in Henan Province. **(b)** The statistical distribution of the OGR index in different focal regions. **(c)** Local indicators spatial autocorrelation (LISA) cluster patterns in different regions of China. High-high (HH) areas, depicted in red on the map, indicate high values surrounded by similar high values, suggesting positive spatial autocorrelation and a tendency to cluster. Low-low (LL) areas, shown in blue on the map, display low values surrounded by similar low values, reflecting positive spatial autocorrelation and a tendency for clustering among low values. Not significant (NS) areas, represented in gray, denote areas where clustering is not statistically significant. **(d)** Moran local scatterplot.

pollution. In this study, Our analysis identified the top three key regions in China with high OGR index, as indicated by the purple box in Fig. 4a. These regions, ranked by regional mean OGR index (Fig. 4b), include the ND, as well as within portions of the YRD and the PRD. Since 2017, the Ministry of Environmental Protection of the People's Republic of China has designated a group of cities known as the "2 + 26" cities as key areas of air pollution concern (Wang et al., 2023b). Fig. 4a illustrates that the ND, one of the high OGR areas, not only encompasses all "2 + 26" cities but also includes adjacent areas such as the Fen-Wei plain and northern Shaanxi Province. This suggests the potential need for expanding the key concern areas in northern China. Conversely, the YRD area could consider reducing the control scope within the original key urban areas (encompassing Shanghai, most cities in Jiangsu, Zhejiang, and Anhui provinces), focusing instead on areas with high OGR index. This strategic shift aims to optimize regional resource allocation further, thereby enhancing the effectiveness of regional air pollution prevention and control measures. Additionally, the results indicate that the current

setup of joint prevention and control areas in the PRD region of southern China is relatively reasonable.

Based on the geospatial distribution of the OGR index across different seasons as shown in Fig. S3b, it is evident that the aforementioned three key areas require intensified prevention and control measures in different seasons. The ND and the YRD need enhanced control efforts primarily in the summer, while the PRD region requires focused attention in the autumn. However, in reality, ozone pollution can occur in any region during any season—spring, summer, autumn, or winter. Therefore, we believe that implementing a normalized regional joint prevention and control strategy, as suggested by Fig. 4a, is a more effective solution for tackling ozone pollution. Additionally, calculating the mean distribution of the OGR index over an extended period (2017–2021) as shown in Fig. 4a also yields more accurate results.

3.4. The impact of different factors on the ozone geospatial relationship

Based on a preliminary understanding of the distribution of the OGR index and anthropogenic/natural factors across regions, this study further explored the impact of different factors on the ozone geospatial relationship. We conducted Pearson correlation analysis and variance inflation factor (VIF) analysis (Fig. S4 and Table S2). Four key factors were identified: anthropogenic activities (AA), surface temperature (ST), east-west surface wind speeds (WS-U), and north-south surface wind speeds (WS-V). By comparing the scatterplots of these four key factors with the OGR index (Fig. S5), it becomes evident that AA and ST have a relatively significant impact on the OGR index and exhibit a positive correlation trend in their distribution. Conversely, WS-U and

WS-V have a relatively minor impact on the OGR index and exhibit a negative correlation trend in their distribution. However, the scatterplot distribution of the global relationship does not reveal the specific geospatial impact of different factors, especially the interplay between anthropogenic activities and natural factors (ST, WS-U, and WS-V) on the OGR index in specific geographical locations. Therefore, we utilized a GWR model for in-depth analysis, as it is well-suited for addressing and identifying geographical relationships among various factors in spatial analysis, making it a classic model in the field of Earth sciences (Martinez-Morata et al., 2022).

Before conducting the GWR modeling, we performed a VIF test on these four independent variables. The VIF test results for all four variables consistently remained below 2 (Table S2), indicating a low level of
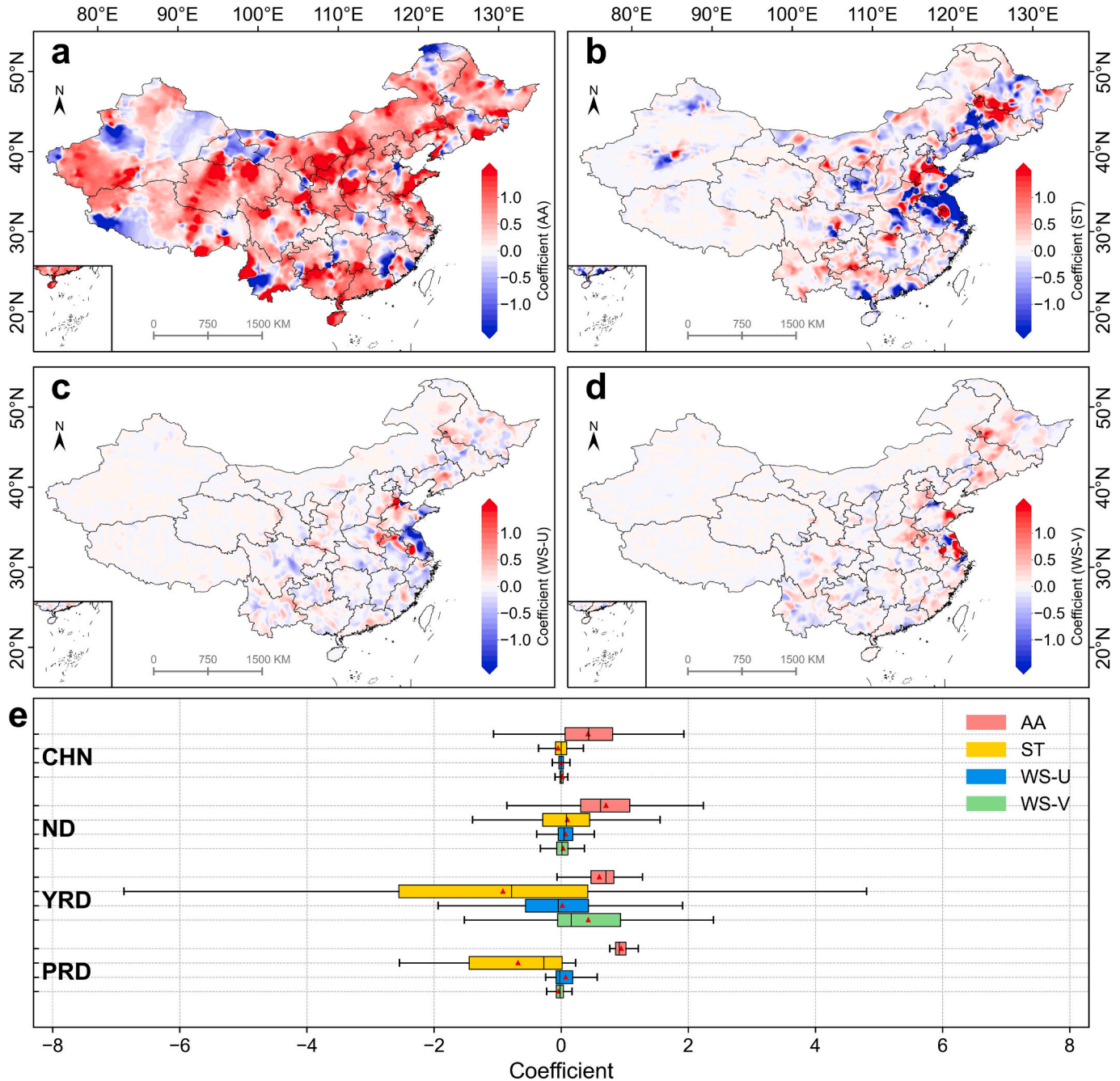

Fig. 5. Geospatial distribution of coefficients in the GWR model. **(a)** Geospatial distribution of AA coefficient. **(b)** Geospatial distribution of ST coefficient. **(c)** Geospatial distribution of WS-U coefficient. **(d)** Geospatial distribution of WS-V coefficient. **(e)** Boxplots of coefficients for different areas, including CHN, ND, as well as within portions of the YRD and the PRD.

collinearity among these variables and satisfying the prerequisite conditions for model construction. Analysis of the model results revealed a mean local R^2 value of 0.9289 (Fig. S6), signifying excellent fitting performance. The distributions of coefficients for the different factors are visually represented in Fig. 5. Fig. 5a depicts the impact of anthropogenic activities factors on OGR index, whereas Fig. 5b–d illustrate the influence of natural factors on OGR index.

Based on the results of the GWR model (Figs. 5 and S7), in most regions of China, the impact of anthropogenic activities on the geographical distribution of the OGR index is more significant than that of natural factors, with this impact demonstrating a positive promotive effect. Among the natural factors, surface temperature, in particular, shows significant spatial heterogeneity in its impact on the distribution of the OGR index (Fig. 5b). Given that surface temperature is a highly comprehensive factor, which can also be affected by atmospheric pollutants or other natural factors, it is somewhat challenging to discern the pattern of its impact on OGR values directly through the distribution of surface temperature coefficient. To delve deeper into its influence on the distribution of the OGR index, we explored the relationship between surface temperature, surface temperature coefficient, and the OGR index, with results shown in Fig. S8. From the figure, it is apparent that when surface temperature is below 273 K, its impact on the OGR index is minimal, and the value of the OGR index is also low. In the range of 273–293 K, it can be seen that as surface temperature increases, higher OGR index values are formed. However, the distribution of the coefficient indicates that surface temperature exerts both positive and negative effects on the OGR index throughout this process, with the negative impact potentially predominating at higher temperatures. Above 293 K, the impact of temperature on the OGR value tends towards being overall negative.

We speculate that the reasons for these observations may be as follows: On one hand, as surface temperatures rise, the rate of photochemical reactions in the atmosphere usually increases, which means that the rate of ozone formation may increase. However, this also accelerates reactions that lead to ozone depletion, especially in the lower atmosphere near the ground, potentially causing a decrease in ozone levels close to the source. Secondly, increased temperatures may lead to the formation of local high-pressure systems, under which air descends, suppressing vertical mixing and possibly leading to reduced wind speeds, diminishing the horizontal dispersion of ozone and its precursors, and causing ozone to remain more confined to its area of origin. On the other hand, this suggests that surface temperature may not be the core natural factor determining the distribution of OGR index at a regional level. Often, it is the synergistic effect with other natural factors (such as wind speed and direction) that influences the distribution of OGR index, leading to the significant spatial heterogeneity in the distribution of temperature coefficients without clear patterns or modes.

Moreover, in the three identified key regions, various factors exhibit diverse impacts on the distribution of OGR index. In the ND, the influence of these factors on the OGR index distribution closely aligns with the CHN level, with anthropogenic activities maintaining its dominant role. Yet, in certain areas within portions of the YRD and the PRD, the impact of WS-U and WS-V factors is noteworthy and should not be overlooked. Our research reveals that in the part of the YRD region, WS-U hinder the formation of high OGR areas, whereas WS-V promote their formation. Our analysis indicates that the prevalence of sea-land winds in the east-west direction in the eastern coastal area of China, where the part of the YRD region is situated, disrupts the local pollution distribution pattern. This leads to a detrimental effect of WS-U on the OGR index distribution in the majority of the part of the YRD region. Conversely, the WS-V in the part of the YRD region is predominantly affected by inland winds, significantly increasing the potential for regional transport of ozone and related pollutants from neighboring areas (Yao et al., 2023). Consequently, this reinforces the local pollution, resulting in a favorable impact of WS-V on the high OGR index distribution in the part of the YRD region. In the part of the PRD region, the scenario is reversed

due to its location in the southern coastal region of China. Dominant WS-U promote the development of high OGR areas, while WS-V hinder this process (Huang et al., 2018). Nevertheless, due to the relatively low anthropogenic activities in the northern adjacent areas of the PRD and its eastern proximity to the ocean (Zheng et al., 2018), both WS-U and WS-V have a smaller impact on the local distribution of the OGR index compared to the part of the YRD region.

3.5. The implications from the OGR index

This study emphasizes that the OGR index can effectively quantify the geospatial relationship of ozone pollution in China, which can provide robust support for the formulation of China's current policies to mitigate ozone pollution, especially in regional collaborative control. Meantime, the study has unveiled and highlighted the substantial impact of anthropogenic activities on the geospatial distribution of the OGR index. Thus, we also can shift our perspective to examine the pattern of ozone-related pollution emissions by analyzing the OGR index distribution. High OGR index areas frequently indicate a significant presence of local ozone-related anthropogenic activities. For this investigation, we collected and processed the regional VOCs and NOx emission inventory data from the Multi-resolution Emission Inventory for China for the years 2017–2020 (Li et al., 2017; Zheng et al., 2018) (MEIC), and mapped out the patterns of VOCs and NOx emissions across different regions of China (as shown in Fig. 6). The results indicate a higher prevalence of pollution source emissions in the eastern regions of China, especially in the relatively developed areas, where the emissions and range of VOCs sources are notably larger and more extensive.

High OGR index areas typically indicate significant anthropogenic activities locally, with natural factors collectively favoring regional air pollution transport. Upon comparing Figs. 4a and 6, we noticed a strong correlation between areas of high VOCs and NOx emissions and most high OGR index areas, particularly in the PRD and YRD regions. Nevertheless, distinctive differences exist in the ND region. Remarkably, within the western part of the ND region, particularly in the northern areas of Shaanxi Province and the Fen-Wei Plain area (referring to the part within Shanxi Province), high OGR areas are predominant. Yet, the emissions of VOCs and NOx in this area are considerably low. Given this area's location within the Loess Plateau, which is at a higher elevation compared to the surrounding plains (such as the North China Plain) known for higher pollutant emissions at lower elevations, the transport of air pollution from these plains to the plateau is likely more obstructed and challenging. Therefore, we suggest the possibility of underreporting in the emission statistics for this region. It should be emphasized that some uncertainties remain with this speculation and further investigation is needed to fully explore it.

4. Conclusions

This study introduced the GeoSHAP method and the concept of the OGR index, facilitating an extensive investigation into the geospatial relationship of surface ozone pollution in China. In addition, based on the OGR index, we identified the most urgent key regions necessitating collaborative ozone control, encompassing the ND, as well as within portions of the YRD and the PRD. This underscores the spatial heterogeneity and regional characteristics of ozone pollution, emphasizing the urgent need for targeted pollution control measures within these distinctive patterns. Moreover, a thorough analysis using the GWR model explores the influence of various factors on regional ozone relationships. The findings suggest that the significant geospatial relationships stem from the synergistic interplay between anthropogenic and natural factors, with anthropogenic factors playing a pivotal role. Notably, in specific regions, especially in coastal regions like the YRD and PRD, natural factors exhibit a substantial impact. These regional distinctions emphasize the necessity for customized and localized strategies to effectively control surface ozone pollution.
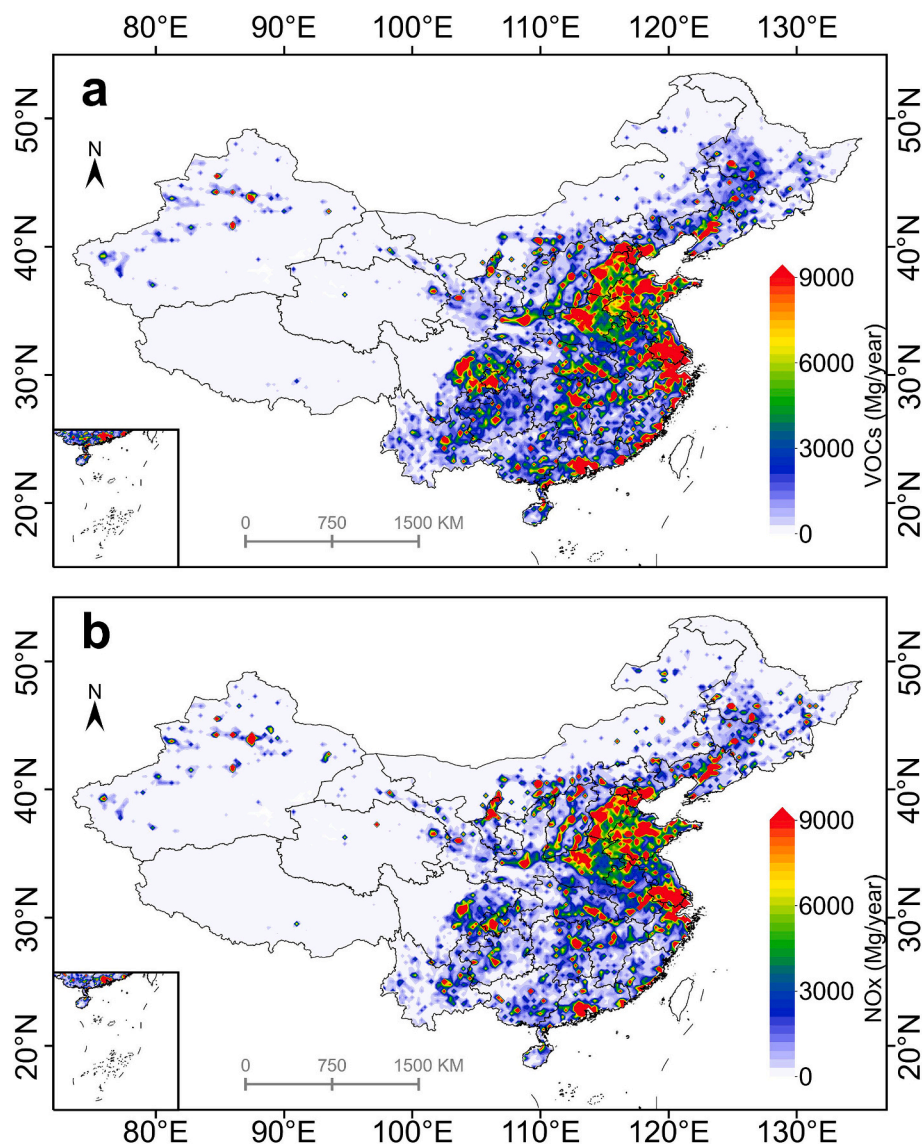
Fig. 6. The geospatial distribution of the mean VOCs and the NOx emission inventory (2017–2020). **(a)** VOCs source emission inventory data. **(b)** NOx source emission inventory data.

CRediT authorship contribution statement

Yong Cheng: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Yan Peng:** Visualization, Investigation, Data curation. **Li-Ming Cao:** Investigation, Data curation. **Xiao-Feng Huang:** Writing – review & editing, Formal analysis, Conceptualization. **Ling-Yan He:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Agathokleous, E., Feng, Z., Oksanen, E., Sicard, P., Wang, Q., Saitanis, C.J., Araminiene, V., Blande, J.D., Hayes, F., Calatayud, V., Domingos, M., Veresoglou, S. D., Peñuelas, J., Wardle, D.A., De Marco, A., Li, Z., Harmens, H., Yuan, X., Vitale, M., Paoletti, E., 2020. Ozone affects plant, insect, and soil microbial communities: a threat to terrestrial ecosystems and biodiversity. Sci. Adv. 6 (33), eabc1176 https://doi.org/10.1126/sciadv.abc1176.
Bayraktar, H., Turalioglu, F.S., 2005. A Kriging-based approach for locating a sampling site—in the assessment of air quality. Stoch Environ Res Risk Assess. 19 (4), 301–305. https://doi.org/10.1007/s00477-005-0234-8.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., Tian, Q., 2023. Accurate medium-range global weather forecasting with 3D neural networks. Nature 619 (7970), 533–538. https://doi.org/10.1038/s41586-023-06185-3.

Chen, X., Wang, M., He, T.-L., Jiang, Z., Zhang, Y., Zhou, L., Liu, J., Liao, H., Worden, H., Jones, D., Chen, D., Tan, Q., Shen, Y., 2023. Data- and model-based urban O3 responses to NOx changes in China and the United States. J. Geophys. Res. Atmos. 128 (20), e2022JD038228 https://doi.org/10.1029/2022JD038228.

Cheng, Y., Zhang, H., Liu, Z., Chen, L., Wang, P., 2019. Hybrid algorithm for short-term forecasting of PM2.5 in China. Atmos. Environ. 200, 264–279. https://doi.org/10.1016/j.atmosenv.2018.12.025.

Cheng, Y., Zhu, Q., Peng, Y., Huang, X.F., He, L.Y., 2021. Multiple strategies for a novel hybrid forecasting algorithm of ozone based on data-driven models. J. Clean. Prod. 326 https://doi.org/10.1016/j.jclepro.2021.129451.

Cheng, Y., Huang, X.F., Peng, Y., Tang, M.X., Zhu, B., Xia, S.Y., He, L.Y., 2023. A novel machine learning method for evaluating the impact of emission sources on ozone formation. Environ. Pollut. 316 (Pt 2), 120685 https://doi.org/10.1016/j.envpol.2022.120685.

Chossière, G.P., Xu, H., Dixit, Y., Isaacs, S., Eastham, S.D., Allroggen, F., Speth, R.L., Barrett, S.R.H., 2021. Air pollution impacts of COVID-19–related containment measures. Sci. Adv. 7 (21), eabe1178 https://doi.org/10.1126/sciadv.abe1178.

Cooper, M.J., Martin, R.V., Hammer, M.S., Levelt, P.F., Veefkind, P., Lamsal, L.N., Krotkov, N.A., Brook, J.R., McLinden, C.A., 2022. Global fine-scale changes in ambient NO2 during COVID-19 lockdowns. Nature 601 (7893), 380–387. https://doi.org/10.1038/s41586-021-04229-0.

Fan, H., Wang, Y., Zhao, C., Yang, Y., Yang, X., Sun, Y., Jiang, S., 2021. The role of primary emission and transboundary transport in the air quality changes during and after the COVID-19 lockdown in China. Geophys. Res. Lett. 48 (7), e2020GL091065 https://doi.org/10.1029/2020GL091065.

Feng, Z., Xu, Y., Kobayashi, K., Dai, L., Zhang, T., Agathokleous, E., Calatayud, V., Paoletti, E., Mukherjee, A., Agrawal, M., Park, R.J., Oak, Y.J., Yue, X., 2022. Ozone pollution threatens the production of major staple crops in East Asia. Nat. Food 3 (1), 47–56. https://doi.org/10.1038/s43016-021-00422-6.

Guo, J., Zhang, X., Gao, Y., Wang, Z., Zhang, M., Xue, W., Herrmann, H., Brasseur, G.P., Wang, T., Wang, Z., 2023. Evolution of ozone pollution in China: what track will it follow? Environ. Sci. Technol. 57 (1), 109–117. https://doi.org/10.1021/acs.est.2c08205.

Huang, L., Liu, S., Yang, Z., Xing, J., Zhang, J., Bian, J., Li, S., Sahu, S.K., Wang, S., Liu, T.Y., 2021a. Exploring deep learning for air pollutant emission estimation. Geosci. Model Dev. 14 (7), 4641–4654. https://doi.org/10.5194/gmd-14-4641-2021.

Huang, X., Ding, A., Gao, J., Zheng, B., Zhou, D., Qi, X., Tang, R., Wang, J., Ren, C., Nie, W., Chi, X., Xu, Z., Chen, L., Li, Y., Che, F., Pang, N., Wang, H., Tong, D., Qin, W., Cheng, W., Liu, W., Fu, Q., Liu, B., Chai, F., Davis, S.J., Zhang, Q., He, K., 2020. Enhanced secondary pollution offset reduction of primary emissions during COVID-19 lockdown in China. Natl. Sci. Rev. 8 (2) https://doi.org/10.1093/nsr/nwaa137.

Huang, X.F., Zou, B.B., He, L.Y., Hu, M., Prévôt, A.S.H., Zhang, Y.H., 2018. Exploration of PM2.5 sources on the regional scale in the Pearl River Delta based on ME-2 modeling. Atmos. Chem. Phys. 18 (16), 11563–11580. https://doi.org/10.5194/acp-18-11563-2018.

Huang, X.F., Cao, L.M., Tian, X.D., Zhu, Q., Saikawa, E., Lin, L.L., Cheng, Y., He, L.Y., Hu, M., Zhang, Y.H., Lu, K.D., Liu, Y.H., Daellenbach, K., Slowik, J.G., Tang, Q., Zou, Q.L., Sun, X., Xu, B.Y., Jiang, L., Shen, Y.M., Ng, N.L., Prévôt, A.S.H., 2021b. Critical role of simultaneous reduction of atmospheric odd oxygen for winter haze mitigation. Environ. Sci. Technol. 55 (17), 11557–11567. https://doi.org/10.1021/acs.est.1c03421.

Huang, Z., Zheng, J., Ou, J., Zhong, Z., Wu, Y., Shao, M., 2019. A feasible methodological framework for uncertainty analysis and diagnosis of atmospheric chemical transport models. Environ. Sci. Technol. 53 (6), 3110–3118. https://doi.org/10.1021/acs.est.8b06326.

Jacob, D.J., 2000. Heterogeneous chemistry and tropospheric ozone. Atmos. Environ. 34 (12), 2131–2159. https://doi.org/10.1016/S1352-2310(99)00462-8.

Kumar, V., Sinha, V., 2021. Season-wise analyses of VOCs, hydroxyl radicals and ozone formation chemistry over north-west India reveal isoprene and acetaldehyde as the most potent ozone precursors throughout the year. Chemosphere 283, 131184. https://doi.org/10.1016/j.chemosphere.2021.131184.

Laughner, J.L., Cohen, R.C., 2019. Direct observation of changing NO<i>_x</i> lifetime in North American cities. Science 366 (6466), 723–727. https://doi.org/10.1126/science.aax6832.

Le, T., Wang, Y., Liu, L., Yang, J., Yung, Y.L., Li, G., Seinfeld, J.H., 2020. Unexpected air pollution with marked emission reductions during the COVID-19 outbreak in China. Science 369 (6504), 702–706. https://doi.org/10.1126/science.abb7431.

Li, D., Wu, Q., Wang, H., Xiao, H., Xu, Q., Wang, L., Feng, J., Yang, X., Cheng, H., Wang, L., Sun, Y., 2021. The Spring Festival Effect: the change in NO2 column concentration in China caused by the migration of human activities. Atmos. Pollut. Res. 12 (12), 101232 https://doi.org/10.1016/j.apr.2021.101232.

Li, K., Jacob, D.J., Liao, H., Shen, L., Zhang, Q., Bates, K.H., 2019. Anthropogenic drivers of 2013–2017 trends in summer surface ozone in China. Proc. Natl. Acad. Sci. U. S. A. 116 (2), 422–427. https://doi.org/10.1073/pnas.1812168116.

Li, K., Jacob, D.J., Shen, L., Lu, X., De Smedt, I., Liao, H., 2020. Increases in surface ozone pollution in China from 2013 to 2019: anthropogenic and meteorological influences. Atmos. Chem. Phys. 20 (19), 11423–11433. https://doi.org/10.5194/acp-20-11423-2020.

Li, M., Liu, H., Geng, G., Hong, C., Liu, F., Song, Y., Tong, D., Zheng, B., Cui, H., Man, H., Zhang, Q., He, K., 2017. Anthropogenic emission inventories in China: a review. Natl. Sci. Rev. 4 (6), 834–866. https://doi.org/10.1093/nsr/nwx150.

Li, X., Liu, J., Mauzerall, D.L., Emmons, L.K., Walters, S., Horowitz, L.W., Tao, S., 2014. Effects of trans-Eurasian transport of air pollutants on surface ozone concentrations over Western China. J. Geophys. Res. Atmos. 119 (21), 12,338–12,354. https://doi.org/10.1002/2014JD021936.

Li, X.-B., Yuan, B., Parrish, D.D., Chen, D., Song, Y., Yang, S., Liu, Z., Shao, M., 2022. Long-term trend of ozone in southern China reveals future mitigation strategy for air pollution. Atmos. Environ. 269, 118869 https://doi.org/10.1016/j.atmosenv.2021.118869.

Lu, X., Hong, J., Zhang, L., Cooper, O.R., Schultz, M.G., Xu, X., Wang, T., Gao, M., Zhao, Y., Zhang, Y., 2018. Severe surface ozone pollution in China: a global perspective. Environ. Sci. Technol. Lett. 5 (8), 487–494. https://doi.org/10.1021/acs.estlett.8b00366.

Lu, X., Zhang, L., Chen, Y., Zhou, M., Zheng, B., Li, K., Liu, Y., Lin, J., Fu, T.M., Zhang, Q., 2019. Exploring 2016–2017 surface ozone pollution over China: source contributions and meteorological influences. Atmos. Chem. Phys. 19 (12), 8339–8361. https://doi.org/10.5194/acp-19-8339-2019.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.I., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2 (1), 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Lyu, X., Li, K., Guo, H., Morawska, L., Zhou, B., Zeren, Y., Jiang, F., Chen, C., Goldstein, A.H., Xu, X., Wang, T., Lu, X., Zhu, T., Querol, X., Chatani, S., Latif, M.T., Schuch, D., Sinha, V., Kumar, P., Mullins, B., Seguel, R., Shao, M., Xue, L., Wang, N., Chen, J., Gao, J., Chai, F., Simpson, I., Sinha, B., Blake, D.R., 2023. A synergistic ozone-climate control to address emerging ozone pollution challenges. One Earth 6 (8), 964–977. https://doi.org/10.1016/j.oneear.2023.07.004.

Martinez-Morata, I., Bostick, B.C., Conroy-Ben, O., Duncan, D.T., Jones, M.R., Spaur, M., Patterson, K.P., Prins, S.J., Navas-Acien, A., Nigra, A.E., 2022. Nationwide geospatial analysis of county racial and ethnic composition and public drinking water arsenic and uranium. Nat. Commun. 13 (1), 7461. https://doi.org/10.1038/s41467-022-35185-6.

Mishra, A.K., Sinha, B., Kumar, R., Barth, M., Hakkim, H., Kumar, V., Kumar, A., Datta, S., Guenther, A., Sinha, V., 2021. Cropland trees need to be included for accurate model simulations of land-atmosphere heat fluxes, temperature, boundary layer height, and ozone. Sci. Total Environ. 751, 141728 https://doi.org/10.1016/j.scitotenv.2020.141728.

Monks, P.S., Archibald, A.T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K.S., Mills, G.E., Stevenson, D.S., Tarasova, O., Thouret, V., von Schneidemesser, E., Sommariva, R., Wild, O., Williams, M.L., 2015. Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer. Atmos. Chem. Phys. 15 (15), 8889–8973. https://doi.org/10.5194/acp-15-8889-2015.

Ogata, S., Takegami, M., Ozaki, T., Nakashima, T., Onozuka, D., Murata, S., Nakaoku, Y., Suzuki, K., Hagihara, A., Noguchi, T., Iihara, K., Kitazume, K., Morioka, T., Yamazaki, S., Yoshida, T., Yamagata, Y., Nishimura, K., 2021. Heatstroke predictions by machine learning, weather information, and an all-population registry for 12-hour heatstroke alerts. Nat. Commun. 12 (1), 4575. https://doi.org/10.1038/s41467-021-24823-0.

Peng, X., Xie, T.-T., Tang, M.-X., Cheng, Y., Peng, Y., Wei, F.-H., Cao, L.-M., Yu, K., Du, K., He, L.-Y., Huang, X.-F., 2023. Critical role of secondary organic aerosol in urban atmospheric visibility improvement identified by machine learning. Environ. Sci. Technol. Lett. https://doi.org/10.1021/acs.estlett.3c00084.

Pulliainen, J., Luojus, K., Derksen, C., Mudryk, L., Lemmetyinen, J., Salminen, M., Ikonen, J., Takala, M., Cohen, J., Smolander, T., Norberg, J., 2020. Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018. Nature 581 (7808), 294–298. https://doi.org/10.1038/s41586-020-2258-0.

Rybarczyk, Y., Zalakeviciute, R., 2021. Assessing the COVID-19 impact on air quality: a machine learning approach. Geophys. Res. Lett. 48 (4), e2020GL091202 https://doi.org/10.1029/2020GL091202.

Shapley, L.S., 1953. A value for n-person games. In: Harold William, K., Albert William, T. (Eds.), Contributions to the Theory of Games II. Princeton University Press, Princeton, pp. 307–317. https://doi.org/10.1515/9781400881970-018.

Shen, L., Liu, J., Zhao, T., Xu, X., Han, H., Wang, H., Shu, Z., 2022. Atmospheric transport drives regional interactions of ozone pollution in China. Sci. Total Environ. 830, 154634 https://doi.org/10.1016/j.scitotenv.2022.154634.

Shindell, D.T., Rind, D., Lonergan, P., 1998. Increased polar stratospheric ozone losses and delayed eventual recovery owing to increasing greenhouse-gas concentrations. Nature 392 (6676), 589–592. https://doi.org/10.1038/33385.

Shrock, E., Fujimura, E., Kula, T., Timms, R.T., Lee, I.-H., Leng, Y., Robinson, M.L., Sie, B.M., Li, M.Z., Chen, Y., Logue, J., Zuiani, A., McCulloch, D., Lelis, F.J.N., Henson, S., Monaco, D.R., Travers, M., Habibi, S., Clarke, W.A., Caturegli, P., Laeyendecker, O., Piechocka-Trocha, A., Li, J.Z., Khatri, A., Chu, H.Y., Villani, A.-C., Kays, K., Goldberg, M.B., Hacohen, N., Filbin, M.R., Yu, X.G., Walker, B.D., Wesemann, D.R., Larman, H.B., Lederer, J.A., Elledge, S.J., Lavin-Parsons, K., Parry, B., Lilley, B., Lodenstein, C., McKaig, B., Charland, N., Khanna, H., Margolin, J., Gonye, A., Gushterova, I., Lasalle, T., Sharma, N., Russo, B.C., Rojas-Lopez, M., Sade-Feldman, M., Manakongtreecheep, K., Tantivit, J., Thomas, M.F., Abayneh, B.A., Allen, P., Antille, D., Armstrong, K., Boyce, S., Braley, J., Branch, K., Broderick, K., Carney, J., Chan, A., Davidson, S., Dougan, M., Drew, D., Elliman, A., Flaherty, K., Flannery, J., Forde, P., Gettings, E., Griffin, A., Grimmel, S., Grinke, K., Hall, K., Healy, M., Henault, D., Holland, G., Kayitesi, C., LaValle, V., Lu, Y., Luthern, S., Marchewka, J., Martino, B., McNamara, R., Nambu, C., Nelson, S., Noone, M., Ommerborn, C., Pacheco, L.C., Phan, N., Porto, F.A., Ryan, E., Selleck, K., Slaughenhaupt, S., Sheppard, K.S., Suschana, E., Wilson, V., Alter, G., Balazs, A., Bals, J., Barbash, M., Bartsch, Y., Boucau, J., Chevalier, J., Chowdhury, F., Einkauf, K., Fallon, J., Fedirko, L., Finn, K., Garcia-Broncano, P., Hartana, C.,

Jiang, C., Kaplonek, P., Karpell, M., Lam, E.C., Lefteri, K., Lian, X., Lichterfeld, M., Lingwood, D., Liu, H., Liu, J., Ly, N., Michell, A., Millstrom, I., Miranda, N., O'Callaghan, C., Osborn, M., Pillai, S., Rassadkina, Y., Reissis, A., Ruzicka, F., Seiger, K., Sessa, L., Sharr, C., Shin, S., Singh, N., Sun, W., Sun, X., Ticheli, H., Trocha-Piechocka, A., Worrall, D., Zhu, A., Daley, G., Golan, D., Heller, H., Sharpe, A., Jilg, N., Rosenthal, A., Wong, C., 2020. Viral epitope profiling of COVID-19 patients reveals cross-reactivity and correlates of severity. Science 370 (6520), eabd4250. https://doi.org/10.1126/science.abd4250.

Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46, 234–240. https://doi.org/10.2307/143141.

Unger, N., Zheng, Y., Yue, X., Harper, K.L., 2020. Mitigation of ozone damage to the world's land ecosystems by source sector. Nat. Clim. Chang. 10 (2), 134–137. https://doi.org/10.1038/s41558-019-0678-3.

Wang, H., Wang, H., Lu, X., Lu, K., Zhang, L., Tham, Y.J., Shi, Z., Aikin, K., Fan, S., Brown, S.S., Zhang, Y., 2023a. Increased night-time oxidation over China despite widespread decrease across the globe. Nat. Geosci. 16 (3), 217–223. https://doi.org/10.1038/s41561-022-01122-x.

Wang, M., Sheng, H., Liu, Y., Wang, G., Huang, H., Fan, L., Ye, D., 2023b. 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. Sci. Total Environ. 888, 164121 https://doi.org/10.1016/j.scitotenv.2023.164121.

Wang, N., Huang, X., Xu, J., Wang, T., Tan, Z.-m., Ding, A., 2022. Typhoon-boosted biogenic emission aggravates cross-regional ozone pollution in China. Sci. Adv. 8 (2), eabl6166 https://doi.org/10.1126/sciadv.abl6166.

Wang, T., Xue, L., Brimblecombe, P., Lam, Y.F., Li, L., Zhang, L., 2017. Ozone pollution in China: a review of concentrations, meteorological influences, chemical precursors, and effects. Sci. Total Environ. 575, 1582–1596. https://doi.org/10.1016/j.scitotenv.2016.10.081.

Wang, Y., Zhao, Y., Liu, Y., Jiang, Y., Zheng, B., Xing, J., Liu, Y., Wang, S., Nielsen, C.P., 2023c. Sustained emission reductions have restrained the ozone pollution over China. Nat. Geosci. https://doi.org/10.1038/s41561-023-01284-2.

Weng, X., Li, J., Forster, G.L., Nowack, P., 2023. Large modeling uncertainty in projecting decadal surface ozone changes over city clusters of China. Geophys. Res. Lett. 50 (9), e2023GL103241 https://doi.org/10.1029/2023GL103241.

Wu, H., Kong, L., Tang, X., Zhu, L., Zhu, J., Wang, Z., 2023. Air quality forecasting with inversely updated emissions for China. Environ. Sci. Technol. Lett. 10 (8), 655–661. https://doi.org/10.1021/acs.estlett.3c00266.

Xing, J., Li, S., Jiang, Y., Wang, S., Ding, D., Dong, Z., Zhu, Y., Hao, J., 2020. Quantifying the emission changes and associated air quality impacts during the COVID-19 pandemic on the North China Plain: a response modeling study. Atmos. Chem. Phys. 20 (22), 14347–14359. https://doi.org/10.5194/acp-20-14347-2020.

Xiong, Y., Chai, J., Mao, H., Mariscal, N., Yacovitch, T., Lerner, B., Majluf, F., Canagaratna, M., Olaguer, E.P., Huang, Y., 2023. Examining the summertime ozone formation regime in Southeast Michigan using MOOSE ground-based HCHO/NO2 measurements and F0AM box model. J. Geophys. Res. Atmos. 128 (19), e2023JD038943 https://doi.org/10.1029/2023JD038943.

Xu, H., Yu, H., Xu, B., Wang, Z., Wang, F., Wei, Y., Liang, W., Liu, J., Liang, D., Feng, Y., Shi, G., 2023. Machine learning coupled structure mining method visualizes the impact of multiple drivers on ambient ozone. Commun. Earth Environ. 4 (1), 265. https://doi.org/10.1038/s43247-023-00932-0.

Xue, L., Ding, A., Cooper, O., Huang, X., Wang, W., Zhou, D., Wu, Z., McClure-Begley, A., Petropavlovskikh, I., Andreae, M.O., Fu, C., 2020. ENSO and Southeast Asian biomass burning modulate subtropical trans-Pacific ozone transport. Natl. Sci. Rev. 8 (6) https://doi.org/10.1093/nsr/nwaa132.

Yao, Y., Wang, W., Ma, K., Tan, H., Zhang, Y., Fang, F., He, C., 2023. Transmission paths and source areas of near-surface ozone pollution in the Yangtze River delta region, China from 2015 to 2021. J. Environ. Manag. 330, 117105 https://doi.org/10.1016/j.jenvman.2022.117105.

Yu, G., Jia, Y., He, N., Zhu, J., Chen, Z., Wang, Q., Piao, S., Liu, X., He, H., Guo, X., Wen, Z., Li, P., Ding, G., Goulding, K., 2019. Stabilization of atmospheric nitrogen deposition in China over the past decade. Nat. Geosci. 12 (6), 424–429. https://doi.org/10.1038/s41561-019-0352-4.

Zahrt, A.F., Henle, J.J., Rose, B.T., Wang, Y., Darrow, W.T., Denmark, S.E., 2019. Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. Science 363 (6424), eaau5631. https://doi.org/10.1126/science.aau5631.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., Zhang, Q., 2018. Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions. Atmos. Chem. Phys. 18 (19), 14095–14111. https://doi.org/10.5194/acp-18-14095-2018.

Zhong, S., Zhang, K., Bagheri, M., Burken, J.G., Gu, A., Li, B., Ma, X., Marrone, B.L., Ren, Z.J., Schrier, J., Shi, W., Tan, H., Wang, T., Wang, X., Wong, B.M., Xiao, X., Yu, X., Zhu, J.-J., Zhang, H., 2021. Machine learning: new ideas and tools in environmental science and engineering. Environ. Sci. Technol. 55 (19), 12741–12754. https://doi.org/10.1021/acs.est.1c01339.

Zhou, B., Guo, H., Zeren, Y., Wang, Y., Lyu, X., Wang, B., Wang, H., 2023. An observational constraint of VOC emissions for air quality modeling study in the Pearl River Delta region. J. Geophys. Res. Atmos. 128 (11), e2022JD038122 https://doi.org/10.1029/2022JD038122.

Zhu, B., Huang, X.-F., Xia, S.-Y., Lin, L.-L., Cheng, Y., He, L.-Y., 2021. Biomass-burning emissions could significantly enhance the atmospheric oxidizing capacity in continental air pollution. Environ. Pollut. 285, 117523 https://doi.org/10.1016/j.envpol.2021.117523.