





Understanding Regional Background Ozone by Multiple Methods: A Case Study in the Shandong Region, China, 2018–2020

F. T. Wang^{1,2}, K. Zhang^{1,2}, J. Xue^{1,2}, L. Huang^{1,2} , Y. J. Wang^{1,2}, H. Chen^{1,2} , S. Y. Wang^{1,2}, J. S. Fu³ , and L. Li^{1,2} 

¹School of Environmental and Chemical Engineering, Shanghai University, Shanghai, China, ²Key Laboratory of Organic Compound Pollution Control Engineering, Shanghai University, Shanghai, China, ³Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, USA

Key Points:

- Regional background ozone is quantified for SD using multiple methods for the years 2018–2020
- There are clear seasonal variations in the regional background O₃
- Local photochemical formation is much more significant in in-land cities

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

L. Li,
lily@shu.edu.cn

Citation:

Wang, F. T., Zhang, K., Xue, J., Huang, L., Wang, Y. J., Chen, H., et al. (2022). Understanding regional background ozone by multiple methods: A case study in the Shandong region, China, 2018–2020. *Journal of Geophysical Research: Atmospheres*, 127, e2022JD036809. <https://doi.org/10.1029/2022JD036809>

Received 21 MAR 2022
Accepted 28 OCT 2022

Author Contributions:

Conceptualization: L. Li
Data curation: F. T. Wang, J. Xue
Formal analysis: F. T. Wang, K. Zhang, J. Xue, L. Huang, Y. J. Wang, H. Chen, S. Y. Wang
Funding acquisition: L. Li
Investigation: L. Huang, L. Li
Methodology: F. T. Wang, K. Zhang, J. Xue, L. Li
Project Administration: L. Li
Resources: L. Li
Supervision: L. Li
Validation: L. Huang, Y. J. Wang, S. Y. Wang
Visualization: F. T. Wang, Y. J. Wang
Writing – original draft: F. T. Wang
Writing – review & editing: L. Huang, J. S. Fu, L. Li

Abstract Uprising ground-level ozone (O₃) and its regional pollution in Northern China are attracting more attention. Besides local precursor emissions and photochemistry, background ozone and long-range transport also contribute significantly to O₃ concentrations. To quantify the regional background O₃ concentrations and their temporal and spatial variations, multiple methods, including the principal component analysis (PCA) and the Texas Commission on Environmental Quality (TCEQ) method, were applied as a case study in Shandong province in Northern China, where serious O₃ pollution occurred frequently yet the background contributions have not been well quantified. We used four methods to quantify the regional background O₃: Method 1 is PCA analysis with only ambient O₃ data as input; Method 2 is PCA analysis considering O₃ and meteorological parameters; Method 3 combines multiple linear regression and the traditional PCA method; and Method 4 is based on TCEQ and consists of the lowest MAD8 O₃ measured to represent regional background O₃ concentrations. Results derived from multiple methods show an overall consistent trend with 2018–2020 averaged regional background O₃ (MDA8) of 41.5 ppb, accounting for 79.4% of the total O₃ in the region. From 2018 to 2020, the changes in regional MDA8 O₃ estimated by Methods 1–4 are –1.8, 0.7, –2.4, and 0.4 ppb, respectively. Clear seasonal variations in the regional background O₃ are observed, showing a pattern of summer > spring > autumn > winter. In addition, the regional ozone contribution at coastal cities was larger than that for inland cities with local O₃ contribution gradually increasing from coastal areas to inland areas. The 3-year average sea-land breeze contribution to summertime O₃ in the eastern coastal cities was estimated to be around 2.1%, while the local photochemistry to O₃ in the inland cities was about 29.7% during ozone pollution episodes, with maximum contribution estimated up to 55.8%. Overall, our study provides insights into the regional background ozone and local photochemistry in Northern China.

Plain Language Summary Surface Ozone is harmful to human health and plants, which is a result of both local emissions and regional contributions. In this study, multiple methods were applied to estimate the regional background ozone in a typical region, Shandong province in North China for the years 2018–2020. Temporal and spatial changes in the regional background ozone were estimated. In addition to the regional background ozone, contributions from local emissions and local meteorology (such as sea-land breeze circulation) to ozone are also analyzed.

1. Introduction

Tropospheric ozone is a typical secondary pollutant, which adversely affects public health, crop yields, and air quality (Chen et al., 2007; Schaubberger et al., 2019; Suciú et al., 2017; Tai & Martin, 2017). Additionally, O₃ has a significant impact on global climate change (IPCC, 2021; Morgenstern et al., 2014). While a small amount of tropospheric O₃ is transported from the stratosphere; it is mainly produced via photochemical reactions among reactive precursors (NO_x, VOCs, and CO). In general, at any location, the measured surface O₃ is the sum of the regional background O₃ and locally produced O₃ (Berlin et al., 2013; Nielsen-Gammon, Tobin, & McNeel, 2005; Nielsen-Gammon, Tobin, McNeel, et al., 2005). Regional background O₃ refers to the amount of O₃ transported into the area by large-scale winds (Langford et al., 2009; Nielsen-Gammon, Tobin, & McNeel, 2005; Nielsen-Gammon, Tobin, McNeel, et al., 2005), that is, the global background and the regional transport of O₃, which mainly includes the photochemical effects of natural emissions of VOCs, NO_x, and CO;

long-range transport of O₃ from distant emission sources; and O₃ from stratosphere-troposphere gas exchange (Langford et al., 2009; Vingarzan, 2004). In most studies, regional background O₃ is defined as the O₃ that would be present in the absence of anthropogenic emissions (Skipper et al., 2021).

O₃ pollution has become increasingly prominent with obvious regional pollution characteristics (Dai et al., 2021; Dang & Liao, 2019). To prevent and control O₃ pollution, it is essential to quantify the background and local O₃ contributions so that the O₃ reduction efficiency by controlling anthropogenic precursors can be investigated (Vingarzan, 2004). Regarding the concentration of regional background O₃, existing research has not fully addressed this problem. The most commonly used methods for calculating regional background O₃ concentrations include the background in-situ measurement, the principal component analysis (PCA) method, the Texas Commission on Environmental Quality (TCEQ) regional background O₃ estimation method (Wu et al., 2017), and modeling simulations (Skipper et al., 2021). The PCA method has been used in estimating background O₃ in the past decade. Langford et al. (2009) were the first to use PCA to analyze the regional background O₃ concentration for Texas in 2006 and identified the first principal component (explained variance of 84%) as the regional background O₃ concentration, which was demonstrated by the spatial distribution of load and meteorological conditions. Using the same method, Liang et al. (2018) analyzed the regional background of O₃ in the Yangtze River Delta region in May 2016, demonstrating that local production had a significant contribution to the high concentrations of O₃. Based on the aforementioned method, Suciú et al. (2017) innovatively inserted meteorological parameters into the PCA to restrict the regional background O₃ and obtained a more reasonable result. The TCEQ regional background O₃ estimation method is simpler than PCA but has higher requirements for the number of monitoring stations, representativeness of the regional distribution, and the integrity of the monitoring data (Wu et al., 2017). The TCEQ method defines the minimum maximum daily 8h average (MDA8) O₃ for all monitoring sites in the study area as the regional background O₃, and the difference between the maximum and minimum values as locally generated O₃ (Nielsen-Gammon, Tobin, & McNeel, 2005; Nielsen-Gammon, Tobin, McNeel, et al., 2005). Xue et al. used the TCEQ method to study the relative contributions from regional background O₃ and local formation to the O₃ level in Hong Kong and further investigated the long-term trend in regional background O₃ from 2002 to 2013. They found that the regional background contribution accounted for approximately 70% of the total O₃, and the increase in regional background O₃ concentration was the major factor for the increase in urban O₃ concentration (Xue et al., 2014). However, estimations of regional background O₃ derived by different methods have significant differences and uncertainties. It is of scientific significance to get an overview of the regional background of O₃ levels estimated by different methods and understand their variations.

China has experienced significant O₃ pollution in recent years, particularly in the North China Plain (NCP) region, one of the most economically developed and polluted regions (Ma et al., 2016; Sun et al., 2021). Shandong (SD) is one of the provinces with fast economic development, active anthropogenic activities, high emission intensities, and severe air pollution, where the 90th percentile of the annual average MDA8 O₃ climbed from 71.9 ppb in 2015 to 86.8 ppb in 2019 (Zhang et al., 2021). According to the annual air quality report issued by the Ministry of Ecology and Environment of China, 2–5 cities were listed among the 20 cities with the worst air quality among the 168 key cities during 2018–2020, showing that the air pollution situation in the SD region is very serious. SD is a large province, with an area of 155,800 km² covering inland and coastal cities. To continuously improve air quality, China issued the Blue Sky Protection Strategy (BSPS) for the years 2018–2020, setting targets for reducing emissions and controlling air pollution. Therefore, we choose the years 2018–2020 to investigate changes in the regional background O₃, so that changes in the local generation of O₃ can be further estimated, which can be used to get an insight into the effects of BSPS implementation. We used multiple methods, including the PCA method, PCL-MLR method, TCEQ method, and background in-situ measurement method to quantify the regional background O₃ concentrations in the SD region. There are two ways to run PCA, with a single variable (only MDA8 O₃) and multivariable (MDA8 O₃, NO₂, wind direction [WD], wind speed [WS], and temperature [T]) as inputs, respectively. The results of PCA, PCA-MLR, TCEQ methods, and background in-situ measurement were compared. On the basis of the aforementioned analysis, we estimated annual changes in the regional background O₃ concentrations, their seasonal variations, and their spatial distributions in the SD region to evaluate the regional contributions of O₃ and provide effective scientific and technological support for the mitigation of O₃ pollution in the SD province and even other regions.

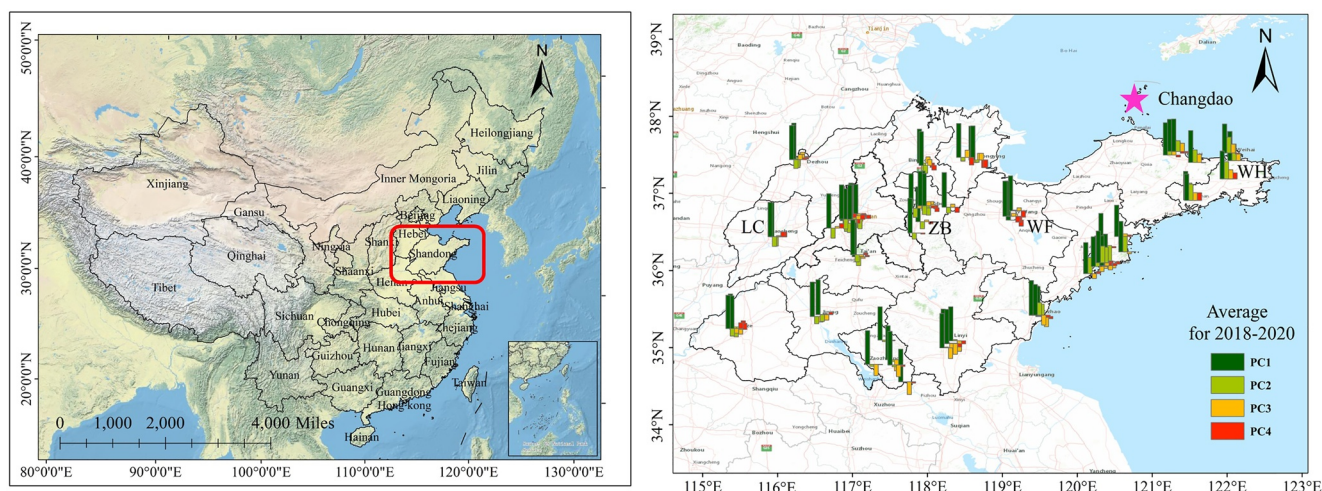


Figure 1. Location of Shandong province; Spatial distribution of the average values of component coefficients (loadings) for PC1, PC2, PC3, and PC4 during 2018–2020 derived by Method 1. Column length represents the size of the loading. The purple star marks the location of the national background site (Changdao, 120.7414°E, 38.1872°N).

2. Methodology

2.1. Data Collection and Preprocessing

Hourly concentrations of O_3 and NO_2 were collected from 96 Air Quality Monitoring Stations (AQMS) in SD Province from 2018 to 2020. These data were measured and released by the China National Environmental Monitoring Center (<http://www.cnemc.cn>). The data processing method used in this study is similar to the process reported by Chu et al. (2020). First, we deleted the missing values and zero values of the site data. Second, we calculated the efficiency of the data (Shamsipour et al., 2014), and sites with an efficiency lower than 90% were excluded. Thus, 66 AQMS were selected after data screening, which covered all cities. Third, the missing values and zero values were filled with linear interpolation to calculate MDA8 O_3 (Ottosen & Kumar, 2019). For data missing for more than 3 consecutive days, linear interpolation was not used, instead, the average of the continuous observational data at the remaining sites was used as a replacement for such data. The spatial distribution of these sites is shown in Figure 1.

Meteorological data were extracted from the National Centers for Environmental Prediction (NCEP) final operational global analysis data files with temporal and spatial resolutions of 6 hr and $2.5^\circ \times 2.5^\circ$, respectively (<https://www.psl.noaa.gov/data/gridded/data.ncep.reanalysis.html>). A large subset of these data is available from the Physical Sciences Laboratory in its original four-times-daily format and as daily averages. Seven grids cover the SD region, and the corresponding grid meteorological data are matched with the AQMS. The meteorological data contained three elements: temperature, u -wind, and v -wind; the temperature was derived from the data for 2 m above ground and both types of wind were set at the 995 sigma level.

2.2. PCA Method

The O_3 concentration varies significantly over time and is influenced by the emissions of O_3 precursors as well as meteorological conditions (J. Z. Wang et al., 2019). When the meteorological conditions are relatively stable, O_3 concentrations are more likely to be affected by the local emissions and photochemistry (Shan et al., 2009) whereas the influence of regional transmission increases as the atmospheric diffusion conditions improve.

Background O_3 (defined as the O_3 that would be present in the absence of anthropogenic emissions) is typically quantified by using chemical transport models (CTMs). The most common approach is the zero-out method or tagged method like CAMx-OSAT or CMAQ-ISAM. However, results are uncertain due to potential errors in model process description and inputs (Skipper et al., 2021). Many different models have been used to quantify background O_3 , often providing estimates that differ significantly (Fiore et al., 2014; Jaffe et al., 2018; McDonald-Buller et al., 2011). In addition, if we want to look at the long-term trends of the regional background

O₃, it takes huge amounts of time to conduct modeling simulations. Statistical methods based on observations provide a way to quantify regional background O₃. The basic principle is: The ozone measured at a specific site is the sum of the background O₃, and the local photochemically generated O₃ (Berlin et al., 2013; Langford et al., 2009; Suciuc et al., 2017). These proportions can be well quantified by using PCA analysis. The regional background O₃ transported into the area by the larger-scale winds, and over a large region tend to show similar patterns at different monitoring sites (Berlin et al., 2013; Liang et al., 2018), which are often shown to be the most dominant component (PC1). The PCA method can avoid the interference of complicated meteorological conditions and local circulation to a certain extent, but the arithmetic process is relatively complex and has strict requirements on the quality of the monitoring network. To strip out the regional background O₃, we used the PCA method to analyze the multi-site MDA8 O₃ and the single-site MDA8 O₃ with NO₂, WD, WS, and T data at various sites in the SD region, and interpreted the results of PCA in combination with meteorological data.

PCA is effective for dimensionality reduction and simplifying the system structure by converting multiple indicators into several uncorrelated comprehensive indicators (principal components) under the premise of less information loss through the correlation coefficient (or variance-covariance) matrix (Murtagh & Heck, 1987). In general, the first few principal components can explain most of the variance in the original variables, and the results of these principal components are used to explain the original observations (Abdul-Wahab et al., 2005). PCA can be combined with multiple linear regression (MLR) methods, where the resolved principal components are considered as ozone sources, factor scores are considered as independent variables, and pollutant concentrations after normalization are considered as dependent variables, to predict and further determine their contributions to O₃ (Jolliffe, 2005; Statheropoulos et al., 1998).

In this study, PCA was used to calculate the regional background O₃ concentration. First, we assumed that all stations in the study area were affected by regional transport air mass; therefore, the principal component representing the regional background could be extracted (Wu et al., 2017). Next, using the prior methods as a reference (Langford et al., 2009; Suciuc et al., 2017), we used the results of loadings and factor scores to explain which principal component represent the regional background and then inversely calculated the regional background O₃ according to Equation 1. This method has been widely applied in O₃ regional background research (Huang et al., 2021; Liang et al., 2018; Yao et al., 2021).

$$O_3 = \bar{O}_3 + \sigma(O_3) \sum_I^{N=66} f_i \alpha_i(t) \quad (1)$$

$$O_3^{PC1} = \bar{O}_3 + \sigma(O_3) f_1 \alpha_1(t) \quad (2)$$

where \bar{O}_3 is the mean of the 3-year MDA8 O₃ at 66 sites, $\sigma(O_3)$ is the standard deviation of the data set, f_i is the PC_{*i*} variance contribution of the results of the PCA, and α_i is the daily PC_{*i*} amplitudes. When PC1 represents the regional background, Equation 2 was used to calculate 8-hr regional background O₃.

2.3. TCEQ Method

The TCEQ method was proposed by the TCEQ. A rural site in the upwind direction was chosen among all the monitoring sites in the study area and the O₃ concentration at the site was utilized as the regional background (Langford et al., 2009; Wu et al., 2017). Nielsen-Gammon, Tobin, and McNeel (2005) and Nielsen-Gammon, Tobin, McNeel, et al. (2005) presented a TCEQ method based on a larger air quality monitoring network, which is simpler, more reliable, and more widely adopted, while its disadvantages are manifested in its susceptibility to local meteorological conditions and the higher requirements of the monitoring network than PCA. This approach calculates the highest 8-hr O₃ concentration at each site and uses the lowest 8-hr O₃ concentration measured across all sites as the regional background value. The improved TCEQ method considers data from a well-established monitoring network with good coverage in all directions in the study area, ensuring that at least one site is not affected by local emissions regardless of wind direction changes. Additionally, the difference between the highest and lowest 8-hr O₃ concentrations at each site is defined as the O₃ generated by local photochemical reactions. The daily 8-hr regional background O₃ and locally generated O₃ can be calculated by Equations 3 and 4.

$$O_{3(R)} = O_{3_MIN} \quad (3)$$

$$O_{3(L)} = (O_{3_MAX}) - (O_{3_MIN}) \quad (4)$$

Table 1
Summary of Parameters for Methods of Calculating Regional Background O_3

Approach	Observational stations	Input parameters
Method 1 (PCA)	66 AQMS in the SD region	MDA8 O_3
Method 2 (pPCA ^a)	5 AQMS sites in the SD region	MDA8 O_3 , NO_2 , WD, WS, T
Method 3 (PCL/MLR)	66 AQMS in the SD region	MDA8 O_3
Method 4 (TCEQ)	65 AQMS in the SD region	MDA8 O_3
Background measurement	Changdao site, Tuoji Island	MDA8 O_3

^aTo distinguish PCA in Method 1, pPCA (precise PCA) is used to describe the PCA used in Method 2.

where $O_{3(R)}$ represents the regional background O_3 , and $O_{3(L)}$ represents the locally generated O_3 .

2.4. Experimental Design

We conducted three distinct PCA calculations to analyze single and multiple variables determining regional background O_3 concentrations at various stations in the SD region. Method 1 was the most conventional approach. We used only MDA8 O_3 to run PCA for the selected 66 AQMS in the SD region from 2018 to 2020. Method 2 considered more information such as meteorological parameters (WD, WS, and T) and precursors (NO_2) with fewer sites than in Method 1, and these sites were distributed in different regions of the SD region to better represent the regional characteristics. We ran five independent PCAs on the selected five sites to extract the regional background O_3 concentrations (Suciu et al., 2017). Unlike Method 1 (single variable, multiple sites) and 2 (multiple variables, multiple sites), Method 3 is a relatively innovative method that combines PCA with MLR, usually used for pollutant source analysis (Bian et al., 2013; Feng et al., 2020). Method 3 includes three steps. First, assuming that regional contribution, local contribution, and other contributions such as sea-land breeze and local small air masses are several sources of ozone. Second, using PCA-MLR to analyze MDA8 O_3 from 66 AQMS. The PC factor score is derived from PCA as the independent variable and the standardization result of the mean value of all sites is used as a dependent variable to predict the contribution rate of different sources. Finally, regional background O_3 was estimated based on the regional contribution rates and factor scores.

In addition, regional background O_3 was calculated using the TCEQ method, which is named Method 4 in this study. However, considering the influence of the MDA8 O_3 minimum data and the location of specific sites on the results, we first screened the frequency distribution of the sites with the smallest MDA8 O_3 values and found that one site (SBQZZ) had the smallest MDA8 value among all the sites on 377 days over 3 years, which could not adequately capture regional transport air masses due to the strong titration by intense NO emissions and the favorable meteorological conditions. Thus, this site was removed. Second, the remaining data were cleaned using a phase-line approach, deleting outliers higher than $Q3 + 1.5(Q3 - Q1)$ or less than $Q1 - 1.5(Q3 - Q1)$ from the MDA8 O_3 sub-data set ($Q1$ and $Q3$ represent the first and third quartiles, respectively) (Mousavinezhad et al., 2021; Yin et al., 2019). Moreover, to evaluate the reliability of the results of the four distinct methods, the regional background O_3 was calculated by the different methods and compared with the observations at Changdao station, which was defined as a national background site (http://www.cnemc.cn/zzjj/jcwl/dqjcw/201711/t20171108_645109.shtml). Specific information for each method is presented in Table 1.

3. Results and Discussion

3.1. Regional and Local Contributions to MDA8 O_3 (Method 1—PCA)

After cleaning the data of all AQMS in the SD region from 2018 to 2020, 66 sites fulfilled the data-quality requirements. We ran three independent PCAs on the MDA8 O_3 at these sites per year, and only the components with eigenvalues greater than 1 were judged as the main components. Results are summarized in Table 2. The PCA resulted in four components for MDA8 O_3 over 3 years: the first principal component could explain the highest (nearly 80%) variance of O_3 , and the cumulative variance of the four principal components is higher than 90%.

Table 2
Results of PCA Analysis (Method 1)

PC	2018			2019			2020		
	Eigenvalue	Variance contribution	Cumulative variance	Eigenvalue	Variance contribution	Cumulative variance	Eigenvalue	Variance contribution	Cumulative variance
PC1	50.99	72.26	72.26	54.09	81.95	81.95	52.26	79.18	79.18
PC2	5.68	8.60	85.87	4.22	6.40	88.35	4.67	7.07	86.25
PC3	1.98	3.00	88.87	1.66	2.52	90.87	2.39	3.62	89.87
PC4	1.10	1.67	90.54	1.005	1.52	92.39	1.16	1.76	91.63

A notable clustering phenomenon was observed when we mapped the principal component loadings for each site to reveal its spatial distribution characteristics (Figure 1). Different colors represent different principal components, and the coefficients ranging from -1 to $+1$ represent the mean contribution of each component to each site during 2018–2020. The length of the column represents the size of the load, with the upward direction corresponding to positive values and the downward to negative values. The amplitude (scores) and loads jointly determine the daily increase or decrease in the O_3 concentration at a specific site. The loading coefficient and amplitude with positive and negative values indicate that the O_3 concentration increases or decreases at the sites, respectively. By comparing the spatial and temporal information provided by the scores and loadings with meteorological data such as wind and temperature, the potential physical and chemical processes could be inferred.

The spatial distribution of loadings is shown in Figure 1. Loadings associated with each principal component using Method 1 for the specific years 2018, 2019, and 2020 are presented in Table S1 in Supporting Information S1. The loadings range from $+0.63$ to $+0.97$ and the PC1 averagely accounts for 77.8% of the variance at each of the 66 sites. The widespread cluster of PC1 suggests that the O_3 and PC1 values at the sites were mostly controlled by the regional background O_3 . This interpretation is supported by Figure 2, which compares the PC1 amplitudes against the NCEP winds. For PC1, the spatial load coefficients of all sites were positive; according to the vector scatter plot of PC1 amplitude and wind speed, the principal components on O_3 exceedance days were positive as well, indicating that PC1 contributed positively to the O_3 concentration at all sites. PC1 represents the regional background, and the southerly wind prevails on the day when the O_3 exceeds the ambient air quality standards of China (~ 75 ppb).

The positive loadings of PC2 are distributed in the coastal area, which shows that PC2 contributes significantly to the stations in the coastal area and is largely influenced by the sea-land breeze. Based on the relationship between PC2 scores and meteorological variables, the influences of WS and WD on PC2 were analyzed. On the monthly scale, PC2 scores were low in the high O_3 season, which was related to air mass transportation in the eastern coastal region. Therefore, we interpreted that PC2 represented mainly the sea-land breeze circulation. The spatial distributions of PC3 and PC4 with positive loadings also showed an obvious feature: PC3 was mainly distributed in the northern region of the SD region, and PC4 was low in the central region and high on the east and west sides. Thus, PC3 and PC4 might be affected mostly by local photochemistry. Based on the temporal variations in PC3 and PC4 scores and their relationship with the meteorological variables (Figures S2 and S3 in Supporting Information S1), there is no clear evidence showing potential influences from either specific meteorology or regional transport, which were therefore named as contributions from local generation.

Based on the spatial distribution of the sites in Figure 1, four sites (Weihai: SDFX; Weifang: HTJZ; Zibo: DFHGC; and Liaocheng: QZF) with relatively complete data were selected and marked on the map. The O_3 season (April–September) was used to illustrate the changes from inland to coastal areas. The difference between the measurements and the regional background O_3 represents the local contribution, which includes not only locally produced O_3 but also the O_3 formed via local circulation. As shown in Figure 3, the local contribution increases as the distance to the coast (from Weihai to Liaocheng), which was expected due to the dense anthropogenic emissions in the inland SD region. In summer, the PC2 amplitude was mostly negative (Figure S1 in Supporting Information S1), and the local contribution to Weihai becomes negative when O_3 from the ocean is transported to this region. In April, May, and September, PC2 was generally positive, and the local contribution increased in Weihai.

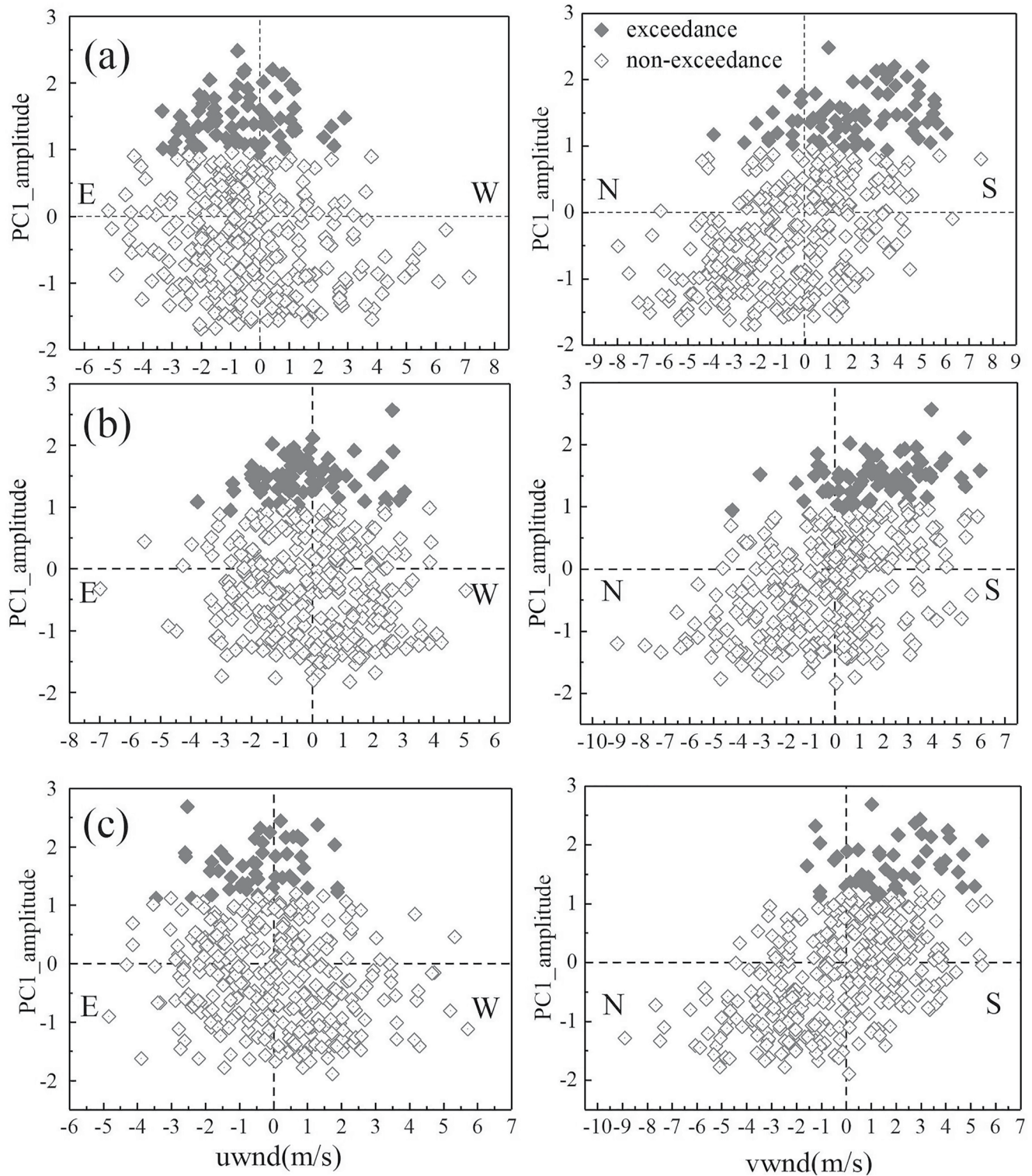


Figure 2. Scatterplots of relations between PC1 amplitudes and mean NCEP reanalysis winds. Solid diamonds represent O₃ exceedance days (MDA8 O₃ > 75 ppb); open diamonds represent O₃ non-exceedance days (MDA8 O₃ < 75 ppb). (a)–(c) represent 2018–2020, respectively.

In addition to the differences in local contributions from inland to coastal, the contribution of sea-land winds to coastal cities and locally generated O₃ in inland cities was further explored. SD region is usually affected by the south-easterly summer monsoon in summer, while coastal areas are also influenced by the sea-land breeze.

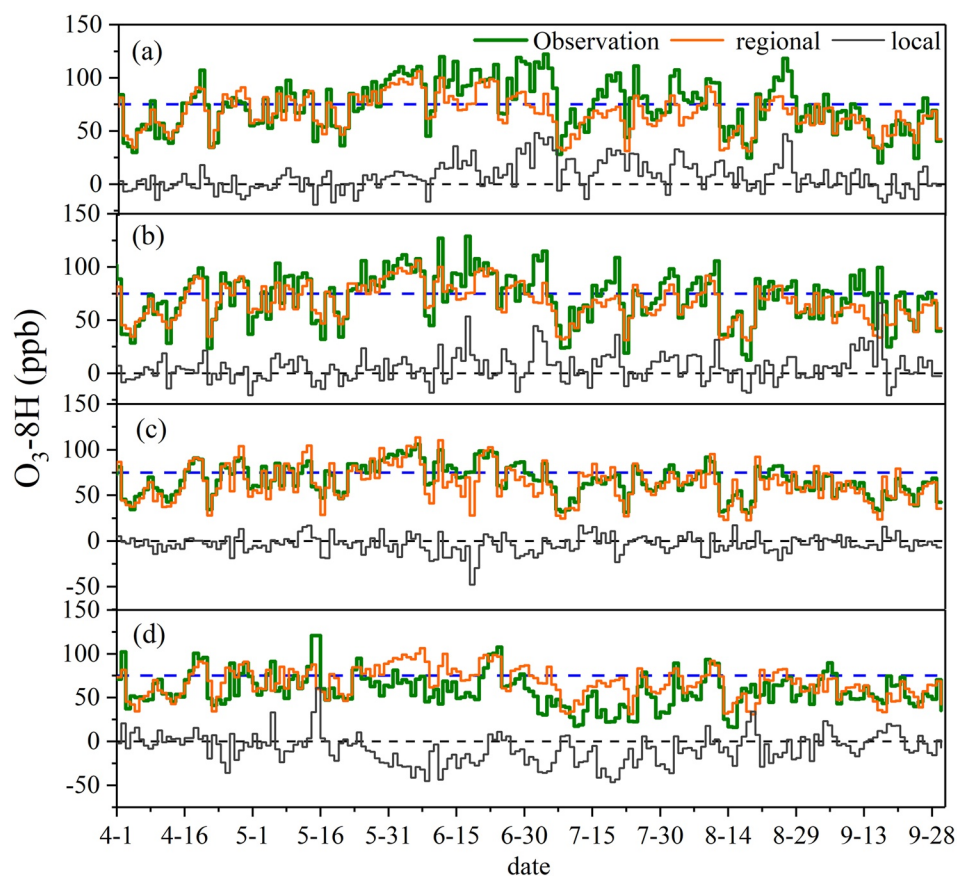


Figure 3. Daily 8-hr maximum O_3 measured at Liaocheng (a), Zibo (b), Weifang (c), and Weihai (d), compared with the regional background O_3 from inland to coast derived from PCA. The green line represents the observed O_3 concentration in the city; the orange line represents the regional background O_3 ; the blue dashed line represents the ozone exceedance limit. The dotted line represents the local contribution defined as the difference between the measurements and the regional background O_3 .

Therefore, we considered the impact of sea-land wind on O_3 in coastal cities from June to August 2018–2020, and the O_3 concentration affected by sea-land wind is calculated using Equations 1 and 2. Its contribution is calculated using the average value of ozone concentration in coastal cities, and the results show that the contribution of sea-land wind to coastal cities in the past 3 years was 4.1%, 2.4%, and 1.8%, respectively. In terms of O_3 exceedance days (with MDA8 O_3 larger than 75 ppb), the contribution of sea-land breeze to coastal cities in the past 3 years was 0.4%, 0.4%, 1.8%, respectively, which is clearly lower than that without pollution, indicating that when ozone pollution occurs, the contribution of sea-land breeze is lower, and is more likely to be affected by inland air mass and photochemistry. O_3 exceedance days during the ozone season (April–September) were chosen to illustrate the contribution of local generation to O_3 in inland cities. Results indicate that the local generation of O_3 during ozone season in 2018, 2019, and 2020 was 35.5%, 29.0%, and 24.7% in 2018–2020, while during O_3 exceedance days, the max contribution of local formation was up to 50.3%, 42.9%, 55.8%, in the year 2018, 2019, and 2020, respectively. Regarding the contribution of regional background O_3 , it can be seen from Figure 3 that higher background O_3 makes it susceptible for O_3 concentrations to exceed the limit when local ozone generation is significant. During the O_3 season, the average contribution for 2018–2020 is 77.5%, 80.5%, and 83.0%, respectively, while for ozone exceedance days, the increased contribution from local generation leads to a slightly lower contribution from the regional background, 76.1%, 79.1%, and 80.3% for 2018–2020, respectively.

As shown in Figure 4, the seasonal variations in the regional background O_3 showed the characteristic pattern of summer > spring > autumn > winter from 2018 to 2020. The regional background O_3 ranged from 28 to 71 ppb. The seasonal variation of background O_3 is the highest in summer, during which background O_3 arises from biogenic emissions of NO_x and VOCs peaks (Chang et al., 2014; Y. Wang et al., 2011). In terms of interannual

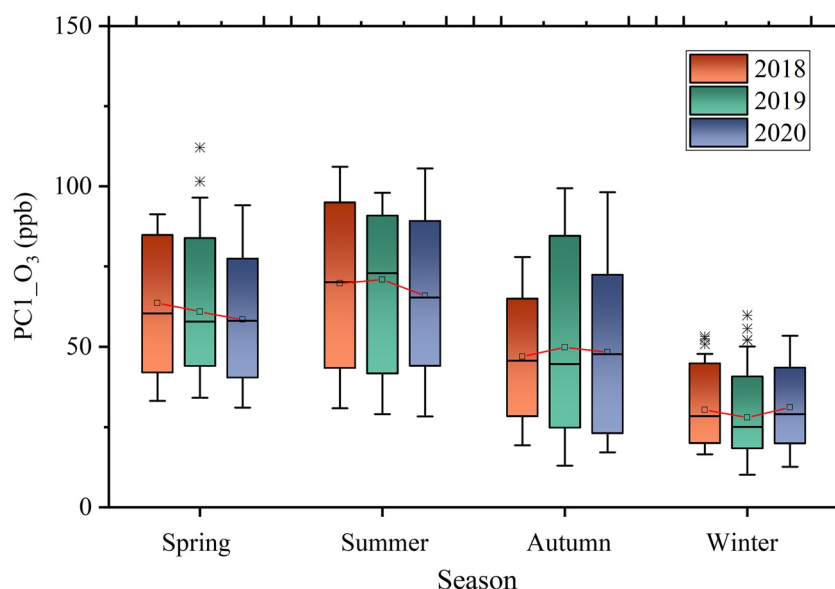


Figure 4. Regional background O₃ in different seasons of 2018–2020 (Method 1—PCA).

variation, the regional background O₃ decreased slightly in spring and summer and increased slightly in autumn and winter during the investigated years. Based on the principle of Method 1, the trend of the regional background ozone basically depends on α_i , which is output by PCA, indicating that the relative contribution of each station during the investigation could be greater or less than the average contribution, so that the regional background O₃ calculated by this method could have a consistent trend with the input observational data. During the study period (2018–2020), the regional background O₃ across four seasons was found to change by -5.1 , -3.8 , 1.4 , and 0.8 ppb, respectively.

3.2. Regional and Local Contributions to MDA8 O₃ (Method 2—PCA)

Method 2 differs from Method 1 because it uses not only MDA8 O₃ but also considers both O₃ precursors (NO₂) and meteorological variables (WS, WD, and T), and selects fewer sites (five sites) with required site distribution. Additionally, site distribution is required, and data must be complete. Data from five sites were used for the analysis: Zibo, Qingdao, Taian, Weihai, and Binzhou. PCA was performed on the five parameters: MDA8 O₃, daily mean NO₂, WD, WS, and T for the five sites from 2018 to 2020; the meteorological data sources were the daily mean data from the NCEP reanalysis data. Results are shown in Table 3, where two components with eigenvalues greater than 1 were extracted for each site, and the eigenvalues of PCs from each site were similar; the mean value was approximately 1.6. The first component explained approximately 40% of the variance in the original variables, and the second component explained approximately 25% of the variance, indicating that both PCs were important in explaining the original variables.

We infer the meaning of the components by considering the relationship between each principal component loading (absolute values greater than or equal to 0.5) and the variables. From the loadings of the two principal components at each site (Table 4), a clear pattern emerges: for each site, PC1 has high loadings on the factors O₃, NO₂, and T, reflecting the chemical photochemistry; PC2 at all sites had larger values on the factors NO₂ and WS, reflecting the physical transport process under the impact of local meteorology. Taking Weihai as an example, we further demonstrated the influence of local meteorological circulation. PC1 increases with O₃ and T but reduces with the decrease of NO₂ (Figure S4 in Supporting Information S1), which reflects the large-scale O₃ photochemical production. PC2 scores did

Table 3
Results of PCA Analysis (Method 2)

City	PC	Eigenvalue	Variance contribution	Cumulative variance
Zibo	PC1	2.324	46.475	46.475
	PC2	1.106	22.122	68.597
Qingdao	PC1	2.071	41.420	41.420
	PC2	1.170	23.405	64.825
Taian	PC1	2.390	47.796	47.794
	PC2	1.140	22.810	70.604
Weihai	PC1	1.835	36.692	36.679
	PC2	1.275	25.508	62.201
Binzhou	PC1	2.155	43.092	43.092
	PC2	1.224	24.486	67.579

Table 4
Loading or Correlations of Components With Variables at Each Site From Method 2

City	PC1					PC2				
	O ₃	NO ₂	T	WD	WS	O ₃	NO ₂	T	WD	WS
Zibo	0.916	-0.674	0.881	0.498	0.081	0.091	0.403	0.014	0.490	-0.834
Qingdao	0.775	-0.582	0.900	0.561	-0.087	0.213	0.596	-0.005	0.199	-0.854
Taian	0.892	-0.726	0.922	0.443	0.139	0.142	0.426	0.036	0.581	-0.775
Weihai	0.731	-0.210	0.807	0.662	-0.409	0.031	-0.854	0.112	0.009	0.729
Binzhou	0.922	-0.488	0.909	0.484	0.075	0.087	0.695	-0.050	0.704	-0.486

not have a significant relationship with T and they increased with WS (Figure S5 in Supporting Information S1), which reflects regional transport effects. Thus, further evidence suggests that PC1 and PC2 were primarily associated with chemical and physical processes, respectively.

Finally, based on the PCA results, referring to the method of Suciu et al. (2017), the PC scores for regional background O₃ were substituted as the mean of PC2 scores at each site, and the PC scores for local contributions were replaced by the mean of PC1 scores at the site, and the cumulative contribution of the PCs was replaced by the results of the standardization of each component. Based on this method, the regional background O₃ was back-calculated, and the results are shown in Figure 5. Compared with other methods, there were no significant seasonal changes. This is mainly because this method considers five sites, and the deviation is relatively small. That is, the estimated regional background O₃ fluctuates around the mean observed ozone value at all stations during the study period, and the regional background O₃ was approximately 51 ppb in each season. The regional background O₃ almost kept constant during 2018–2020.

3.3. Regional and Local Contributions to MDA8 O₃ (Method 3—PCA/MLR)

PCA/MLR (Method 3), as a relatively novel method, uses the idea of source resolution and continues to use MLR to estimate the O₃ regional background based on the results of Method 1 (PC1 represents the regional background). Using the factor score of Method 1 as the independent variable, and the standardized results of the mean MDA8 O₃ of 66 AQMS sites in the SD region as the dependent variable, after MLR processing, the contribution proportions of the 2018–2020 regional background O₃ were obtained as follows: 60.2%, 57.3%, and 57.3%, with

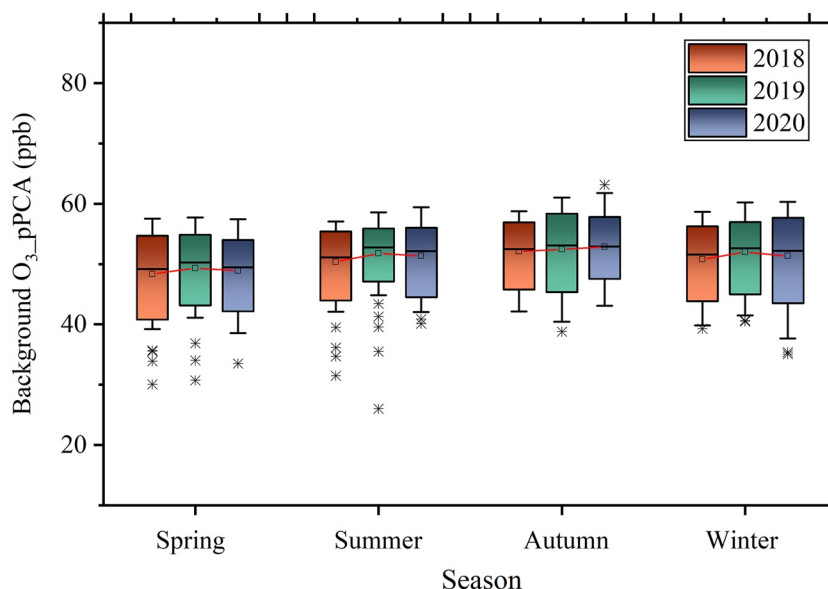


Figure 5. Regional background O₃ during different seasons of 2018–2020 (Method 2—PCA).

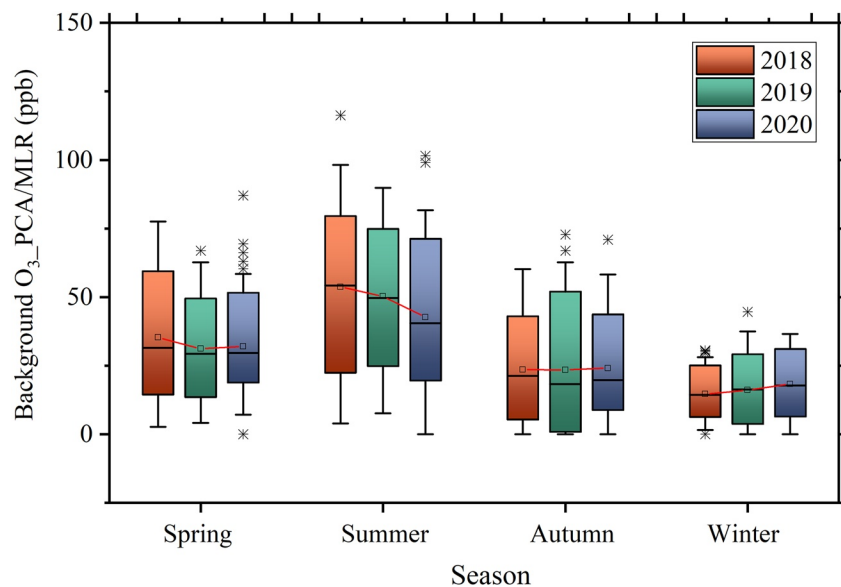


Figure 6. Regional background O₃ in the seasons of 2018–2020 (Method 3—PCA/MLR).

a decrease in the latter 2 years, compared to 2018; the regional background O₃ was then calculated using eqn. 10 of Bian et al. (2013), results are shown in Figure 6. The seasonal pattern of regional background O₃ remained consistent with that of Method 1, but the annual variation varied slightly by season, especially in summer, for which there was a decreasing trend from year to year, mainly because of the interannual decreasing regional background contribution and the decreasing O₃ concentration. In contrast, the calculated regional background contribution of ozone derived by Method 1 was found to be higher in the summer of 2019 than in 2018 such that the estimated regional background O₃ exhibited an increase.

3.4. Regional and Local Contributions to MDA8 O₃ (Method 4—TCEQ)

The TCEQ method was also used to estimate the regional background O₃ in the SD region. Since the lowest MDA8 O₃ at the AQMS selected by the TCEQ method represents the regional background concentration, to reduce the inaccuracy caused by the specific site, the distribution of the minimum MDA8 O₃ at all sites was calculated to ensure that it would not be affected by a single low-value site. Figure 7 shows the regional background O₃ derived by the TCEQ method from 2018 to 2020, ranging from 12.6 to 49.8 ppb. It can be observed that the O₃ pattern of summer > spring > autumn > winter exhibited trends. Summer and autumn show an increasing trend followed by a decrease while the other two seasons present the opposite pattern. Overall, they all show a slightly increasing characteristic across four seasons over the 3 years, rising by −0.3, 1.5, 1.7, and 1.8 ppb, respectively. The results are slightly different from that was calculated by PCA from Method 1, the results in summer also showed a rising characteristic followed by decreasing, but exhibited an overall decline trend from 2018 to 2020.

To further illustrate the contribution of regional background O₃ to coastal and inland cities in different years and seasons, the mean MDA8 O₃ was calculated for all AQMS in coastal and inland cities in the SD region, and the ratio of the regional background O₃ to the mean MDA8 O₃ was defined to reflect the magnitude of the contribution of the regional background O₃. As shown in Table 5, the contribution of the regional background to the coastal cities is higher than inland, which is consistent with the conclusion in PCA Method 1 that the local contribution from the coast to the inland is increasing. Regarding interannual variability, the regional contribution of O₃ to both showed an increasing pattern, and for seasonal variability, it decreased sequentially from spring to winter.

3.5. Comparisons Among Multiple Methods

Due to the differences in the principles used to estimate the regional background O₃ concentrations, there were differences in the calculated results. In general, the results of Methods 1 and 2 were approximately 20 ppb larger than those of Method 3 and the TCEQ, but they are closer to the background site. Additionally, Method 2 has a

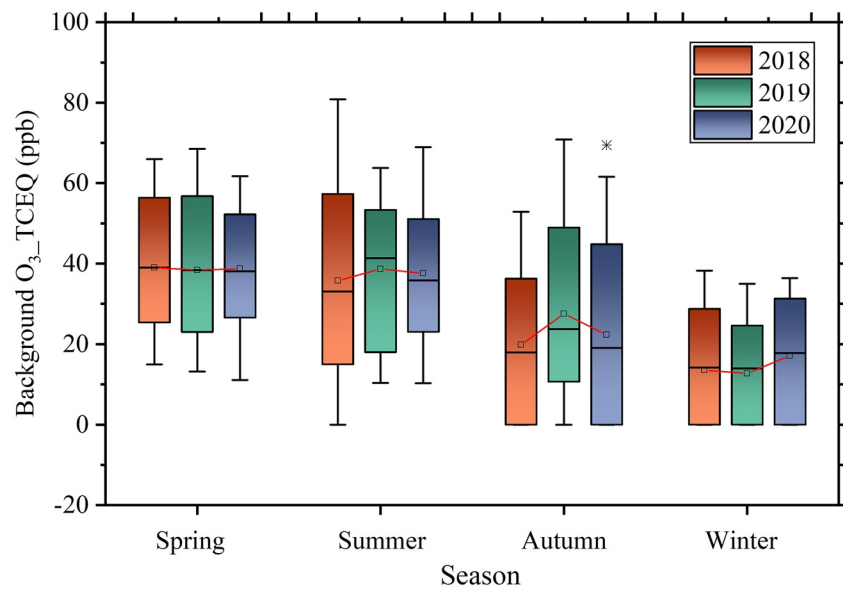


Figure 7. Regional background O₃ in the seasons of 2018–2020 (Method 4—TCEQ).

smooth seasonal trend, with results that are marginally higher than the background site results in the autumn and winter. Therefore, the results of Method 1 were chosen as being more accurate for estimating the regional background O₃. For Method 3, this difference was mainly because after the MLR processing, the resulting regional contribution decreased, to only about 60%. Thus, the lower regional background O₃ levels were estimated. In the TCEQ method, the lowest MDA8 O₃ was selected to represent the regional background, and the selected site may be influenced by urban sites that do not capture the regional background well and are therefore lower in magnitude compared with Methods 1 and 2.

As shown by the prior analysis of the results, the seasonal changes of the regional background O₃ were generally consistent across the three methods, with a clear monthly variation except for Method 2, which adds meteorological parameters as constraints and has a smooth trend. This phenomenon may be because Method 2 considers the meteorological factors of the station, indicating that the main component of the regional background value has almost no relationship with temperature. Therefore, there are no obvious monthly variations. The results of each method for interannual variability are presented in Table 6. Methods 1 and 3 use different analysis methods for the same data set, the annual changes for both are consistent, showing a slight decrease of 1.8 and 2.4 ppb for each of the 3 years. The results of Method 2, the TCEQ method, and the background sites show a consistent pattern of increasing and then decreasing, but overall, the values increase by 0.7, 0.4, and −0.8 ppb, respectively, over the 3 years. Although the annual changes vary among different methods, an overall increase in the estimated regional background O₃ during the past 3 years can be consistently observed. Additionally, to reduce the error of a single method, the average value of the aforementioned results is expressed as the regional background O₃ in the SD region in the past 3 years, which were 41.6, 41.9, and 40.9 ppb, respectively, and the 3-year average value is 41.5 ppb.

Table 5
Contribution of Regional Background O₃ to Coastal and Inland Cities in the Seasons of 2018–2020

Season	Coastal cities			Inland cities		
	2018	2019	2020	2018	2019	2020
Spring	65.1%	66.8%	72.0%	62.2%	63.6%	67.1%
Summer	57.6%	61.9%	63.7%	47.0%	52.2%	55.6%
Autumn	43.1%	54.7%	45.4%	42.8%	54.2%	44.0%
Winter	41.8%	43.2%	52.7%	41.3%	41.9%	51.1%

3.6. Comparisons With Previous Studies

Figure 8 summarizes the regional background O₃ concentrations reported in previous studies estimated by different methods around the world. We compared the results of this study with other studies (Berlin et al., 2013; Huang et al., 2021; Liang et al., 2018; Sahu et al., 2021; Souri et al., 2016; Suciú et al., 2017; Y. Wang et al., 2011; Xue et al., 2014). Method 2 is referenced in Suciú's Approach B. In our results, PC2 represents the regional background, which differs from the results of Suciú et al. (2017), and the seasonal variation in regional background O₃ is not significant but remains similar in terms of magnitude. Compared with the results of Berlin et al. (2013), the

Table 6
Comparison of All Approaches in This Study and the Literature

Method	2018		2019		2020		Average	
	AVE	SD	AVE	SD	AVE	SD	AVE	SD
Method 1 (PCA)	52.8	21.5	52.5	23.2	51.0	21.7	52.1	21.7
Method 2 (pPCA)	50.4	4.9	51.4	4.9	51.1	4.8	51.0	4.9
Method 3 (PCL/MLR)	32.5	21.4	31.6	19.6	30.1	18.4	31.4	19.5
Method 4 (TCEQ)	30.9	14.6	32.1	14.9	31.3	14.3	31.4	14.4
Average	41.6	19.7	41.9	19.8	40.9	18.2	41.5	4.9
Background site	50.6	22.0	52.9	21.9	49.8	19.3	51.1	21.1

Note. AVE represents the average value of a method for a given year. SD represents the standard deviation of a method for a given year.

region background O₃ is calculated by the same method, and both exhibit larger PCA results than those of TCEQ, by roughly 8.2 ppb. Our results are lower than those of Sourì et al. (2016), focused on the regional background O₃ under different wind directions and showed that the regional background O₃ was greatest with east-northeast winds. The results of Liang et al. and Huang et al. do not provide specific regional background O₃ concentrations but report ranges of 32.1–72.2 and 31.0–102.6 ppb, respectively. The maximum values are higher than the overall average value because both study periods are in the O₃ season. In addition, we also compared with results from modeling studies. Sahu et al. (2021) simulated the regional background O₃ across China using the CMAQv5.2 model and found that in the SD region, the regional background O₃ ranged from 21 to 33.6 ppb in 2015, accounting for 71%–94% of the observed O₃ concentration, and the seasonal characteristics are consistent with this study. Wang et al. used the nested-grid GEOS-Chem model and found that the annual mean regional background O₃ over China was 44.1 ppbv in 2006, while in SD it ranged between 30 and 50 ppbv. Slightly different from this study, the results of Wang et al. showed higher background O₃ in spring than in summer, mainly because spring can be attributed to enhanced stratosphere-troposphere exchange (Y. Wang et al., 2011). Overall, our results are comparable with previous results but provide more insights into the O₃ pollution in the SD region.

4. Conclusions

Three PCA methods with different parameters and the TCEQ method were used to estimate the regional background O₃ concentration in the SD region, where O₃ pollution is severe in recent years. The regional background O₃ calculated using different PCA and TCEQ methods did not differ significantly and showed an overall consistent trend. Method 1 is the most commonly used method for resolving regional background O₃ using PCA and produces the highest regional background O₃ concentration. Method 2 incorporates O₃ precursor (NO₂) and meteorological

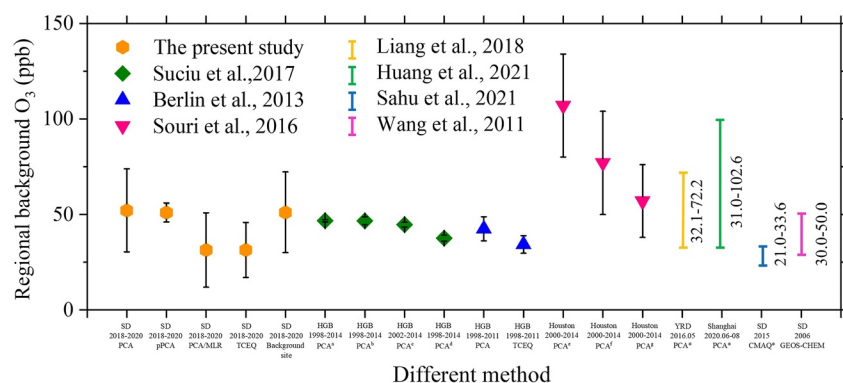


Figure 8. Comparisons between this study and previous results, ^aO₃ data using MDA8 (multi-site). ^bO₃ data using MDA8 (single-site). ^cSame as b, but less time period is used. ^dO₃ data using hourly measurements. ^eConstrained by wind direction from east-northeast. ^fConstrained by wind direction from east-southeast. ^gConstrained by wind direction from south-southeast. ^hModel-based results.

parameters as constraints and yields a flat monthly trend. Method 3 combines PCA with MLR and resolves relatively lower O₃ background concentrations, which makes it close to the result of the TCEQ method. The results of the four methods are similar and slightly different. Based on the results of the four methods and ozone concentration from national background sites, the 3-year regional background O₃ showed an overall slightly increasing pattern, and the 3-year average values for Methods 1, 2, 3, 4, and in-situ background measurement were 52.1 ± 21.7 , 51.00 ± 4.9 , 31.4 ± 19.5 , 31.4 ± 14.4 , and 51.1 ± 21.1 ppb, respectively. There was a clear seasonal pattern of regional background O₃, with high values in spring and summer and low concentrations in autumn and winter. Overall, we recommend Method 1 among all the four methods because results derived from it are most close to the national background site observations, in terms of seasonal trends and the scale of regional background O₃. Furthermore, the regional background O₃ differs spatially from the eastern coastal area seeing more influences from the marine environment. The concentration of locally generated O₃ gradually increased from coastal to inland cities while the opposite was observed for regional ozone contribution. Uncertainties exist in terms of estimating the regional background O₃ concentrations. Additional factors can also be considered for multivariate analysis, such as adding constraints to the precursor (e.g., VOCs) and additional relevant meteorological variables (e.g., solar radiation and relative humidity). Further research is necessary to reduce these uncertainties in the future.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

All data used in this paper have been deposited in an open research repository, which is available at <https://doi.org/10.17632/tp9k8cs6f.1> (F. Wang et al., 2022) or <https://pan.baidu.com/s/1LbWiMnjgADJE0tF89jJTHw?pwd=5pe9>, the file name is: "Summary of data for estimating regional background ozone in SD.xlsx." Hourly ambient air quality data are publicly available on <https://www.cnemc.cn>. Meteorological data were extracted from the National Centers for Environmental Prediction (NCEP) final operational global analysis data files with temporal and spatial resolutions of 6 hr and $2.5^\circ \times 2.5^\circ$, respectively (<https://www.psl.noaa.gov/data/gridded/data.ncep.reanalysis.html>). The IBM SPSS Statistics 18 was used to perform principal component analysis (PCA) and multiple linear regression analysis (MLR), accessible at <https://www.ibm.com/cn-zh/analytics/spss-trials>.

Acknowledgments

This study was financially supported by the National research program for key issues in air pollution control (DQGG202119) and the National Natural Science Foundation of China (42075144).

References

- Abdul-Wahab, S. A., Bakheit, C. S., & Al-Alawi, S. M. (2005). Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Environmental Modelling & Software*, *20*(10), 1263–1271. <https://doi.org/10.1016/j.envsoft.2004.09.001>
- Berlin, S. R., Langford, A. O., Estes, M., Dong, M., & Parrish, D. D. (2013). Magnitude, decadal changes, and impact of regional background ozone transported into the greater Houston, Texas, Area. *Environmental Science & Technology*, *47*(24), 13985–13992. <https://doi.org/10.1021/es4037644>
- Bian, L., Li, T., & Hou, J. (2013). Source apportionment of polycyclic aromatic hydrocarbons using two mathematical models for runoff of the Shanghai elevated inner highway, China. *Environmental Science*, *34*(10), 3840–3846. <https://doi.org/10.13227/j.hjxx.2013.10.027>
- Chang, C. C., Wang, J. L., Lung, S. C. C., Chang, C. Y., Lee, P. J., Chew, C., et al. (2014). Seasonal characteristics of biogenic and anthropogenic isoprene in tropical–subtropical urban environments. *Atmospheric Environment*, *99*, 298–308. <https://doi.org/10.1016/j.atmosenv.2014.09.019>
- Chen, T. M., Kuschner, W. G., Gokhale, J., & Shofer, S. (2007). Outdoor air pollution: Ozone health effects. *The American Journal of the Medical Sciences*, *333*(4), 244–248. <https://doi.org/10.1097/MAJ.0b013e31803b8e8c>
- Chu, B., Ma, Q., Liu, J., Ma, J., Zhang, P., Chen, T., et al. (2020). Air pollutant correlations in China: Secondary air pollutant responses to NO_x and SO₂ control. *Environmental Science & Technology Letters*, *7*(10), 695–700. <https://doi.org/10.1021/acs.estlett.0c00403>
- Dai, H., Zhu, J., Liao, H., Li, J., Liang, M., Yang, Y., & Yue, X. (2021). Co-occurrence of ozone and PM_{2.5} pollution in the Yangtze River Delta over 2013–2019: Spatiotemporal distribution and meteorological conditions. *Atmospheric Research*, *249*, 105363. <https://doi.org/10.1016/j.atmosres.2020.105363>
- Dang, R., & Liao, H. (2019). Radiative forcing and health impact of aerosols and ozone in China as the consequence of clean air actions over 2012–2017. *Geophysical Research Letters*, *46*(21), 12511–12519. <https://doi.org/10.1029/2019GL084605>
- Feng, J., Song, N., Yu, Y., & Li, Y. (2020). Differential analysis of FA-NNC, PCA-MLR, and PMF methods applied in source apportionment of PAHs in street dust. *Environmental Monitoring and Assessment*, *192*(11), 1–11. <https://doi.org/10.1007/s10661-020-08679-3>
- Fiore, A. M., Oberman, J. T., Lin, M. Y., Zhang, L., Clifton, O. E., Jacob, D. J., et al. (2014). Estimating North American background ozone in U.S. surface air with two independent global models: Variability, uncertainties, and recommendations. *Atmospheric Environment*, *96*, 284–300. <https://doi.org/10.1016/j.atmosenv.2014.07.045>
- Huang, Q., Huang, Y., Zhang, S., Jin, D., Gao, S., & Xiu, G. (2021). O₃ source characteristics of industrial area in Yangtze River Delta based on boundary observation. *Environmental Science*, *42*(10), 4621–4631. <https://doi.org/10.13227/j.hjxx.202101199>
- IPCC. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.). Cambridge University Press. Retrieved from <https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/>

- Jaffe, D. A., Cooper, O. R., ArleneFiore, M., Henderson, B. H., Tonnesen, G. S., Russell, A. G., et al. (2018). Scientific assessment of background ozone over the U.S.: Implications for air quality management. *Elementa: Science of the Anthropocene*, 6, 56. <https://doi.org/10.1525/elementa.309>
- Jolliffe, I. (2005). Principal Component Analysis. *Encyclopedia of Statistics in Behavioral Science*. <https://doi.org/10.1002/0470013192.bsa501>
- Langford, A. O., Senff, C. J., Banta, R. M., Hardesty, R. M., Alvarez, R. J., Sandberg, S. P., & Darby, L. S. (2009). Regional and local background ozone in Houston during Texas Air Quality Study 2006. *Journal of Geophysical Research*, 114(D7). <https://doi.org/10.1029/2008JD011687>
- Liang, Y., Liu, Y., Wang, H., Li, L., Duan, Y., & Lu, K. (2018). Regional characteristics of ground-level ozone in Shanghai based on PCA analysis. *Acta Scientiae Circumstantiae*, 38(10), 3807–3815. <https://doi.org/10.13671/j.hjkxxb.2018.0209>
- Ma, Z., Xu, J., Quan, W., Zhang, Z., Lin, W., & Xu, X. (2016). Significant increase of surface ozone at a rural site, north of eastern China. *Atmospheric Chemistry and Physics*, 16(6), 3969–3977. <https://doi.org/10.5194/acp-16-3969-2016>
- McDonald-Buller, E. C., Allen, D. T., Brown, N., Jacob, D. J., Jaffe, D., Klob, C. E., et al. (2011). Establishing policy relevant background (PRB) ozone concentrations in the United States. *Environmental Science & Technology*, 45(22), 9484–9497. <https://doi.org/10.1021/es2022818>
- Morgenstern, O., Zeng, G., Dean, S. M., Joshi, M., Abraham, N. L., & Osprey, A. (2014). Direct and ozone-mediated forcing of the Southern Annular Mode by greenhouse gases. *Geophysical Research Letters*, 41(24), 9050–9057. <https://doi.org/10.1002/2014GL062140>
- Mousavinezhad, S., Choi, Y., Pouyaei, A., Ghahremanloo, M., & Nelson, D. L. (2021). A comprehensive investigation of surface ozone pollution in China, 2015–2019: Separating the contributions from meteorology and precursor emissions. *Atmospheric Research*, 257, 105599. <https://doi.org/10.1016/j.atmosres.2021.105599>
- Murtagh, F., & Heck, A. (1987). Principal components analysis. In *Multivariate data analysis* (Vol. 131, pp. 13–53). Springer. https://doi.org/10.1007/978-94-009-3789-5_2
- Nielsen-Gammon, J., Tobin, J., & McNeel, A. (2005). *A conceptual model for eight-hour ozone exceedances in Houston, Texas Part II: Eight-hour ozone exceedances in the Houston-Galveston Metropolitan area*, (p. 79). Houston Advanced Research Center. Retrieved from <https://hdl.handle.net/1969.1/158249>
- Nielsen-Gammon, J., Tobin, J., McNeel, A., & Li, G. (2005). *A conceptual model for eight-hour ozone exceedances in Houston, Texas Part I: Background ozone levels in eastern Texas* (p. 52). Houston Advanced Research Center. Retrieved from <https://hdl.handle.net/1969.1/158250>
- Ottosen, T. B., & Kumar, P. (2019). Outlier detection and gap filling methodologies for low-cost air quality measurements. *Environmental Science: Processes & Impacts*, 21(4), 701–713. <https://doi.org/10.1039/C8EM00593A>
- Sahu, S. K., Liu, S., Liu, S., Ding, D., & Xing, J. (2021). Ozone pollution in China: Background and transboundary contributions to ozone concentration & related health effects across the country. *Science of the Total Environment*, 761, 144131. <https://doi.org/10.1016/j.scitotenv.2020.144131>
- Schauberger, B., Rolinski, S., Schaphoff, S., & Müller, C. (2019). Global historical soybean and wheat yield loss estimates from ozone pollution considering water and temperature as modifying effects. *Agricultural and Forest Meteorology*, 265, 1–15. <https://doi.org/10.1016/j.agrformet.2018.11.004>
- Shamsipour, M., Farzadfar, F., Gohari, K., Parsaeian, M., Amini, H., Rabiei, K., et al. (2014). A framework for exploration and cleaning of environmental data—Tehran air quality data experience. *Archives of Iranian Medicine*, 17(12), 821–829. Retrieved from <http://aimjournal.ir/Article/752>
- Shan, W., Yin, Y., Lu, H., & Liang, S. (2009). A meteorological analysis of ozone episodes using HYSPLIT model and surface data. *Atmospheric Research*, 4(93), 767–776. <https://doi.org/10.1016/j.atmosres.2009.03.007>
- Skipper, T. N., Hu, Y., Odman, M. T., Henderson, B. H., Hogrefe, C., Mathur, R., & Russell, A. G. (2021). Estimating US Background Ozone Using Data Fusion. *Environmental Science & Technology*, 55(8), 4504–4512. <https://doi.org/10.1021/acs.est.0c08625>
- Souri, A. H., Choi, Y., Li, X., Kotsakis, A., & Jiang, X. (2016). A 15-year climatology of wind pattern impacts on surface ozone in Houston, Texas. *Atmospheric Research*, 174, 124–134. <https://doi.org/10.1016/j.atmosres.2016.02.007>
- Statheropoulos, M., Vassiliadis, N., & Pappa, A. (1998). Principal component and canonical correlation analysis for examining air pollution and meteorological data. *Atmospheric Environment*, 32(6), 1087–1095. [https://doi.org/10.1016/S1352-2310\(97\)00377-4](https://doi.org/10.1016/S1352-2310(97)00377-4)
- Suciu, L. G., Griffin, R. J., & Masiello, C. A. (2017). Regional background O₃ and NO_x in the Houston–Galveston–Brazoria (TX) region: A decadal-scale perspective. *Atmospheric Chemistry and Physics*, 17(11), 6565–6581. <https://doi.org/10.5194/acp-17-6565-2017>
- Sun, J., Shen, Z., Wang, R., Li, G., Zhang, Y., Zhang, B., et al. (2021). A comprehensive study on ozone pollution in a megacity in North China Plain during summertime: Observations, source attributions and ozone sensitivity. *Environment International*, 146, 106279. <https://doi.org/10.1016/j.envint.2020.106279>
- Tai, A. P., & Martin, M. V. (2017). Impacts of ozone air pollution and temperature extremes on crop yields: Spatial variability, adaptation and implications for future food security. *Atmospheric Environment*, 169, 11–21. <https://doi.org/10.1016/j.atmosenv.2017.09.002>
- Vingarzan, R. (2004). A review of surface ozone background levels and trends. *Atmospheric environment*, 38(21), 3431–3442. <https://doi.org/10.1016/j.atmosenv.2004.03.030>
- Wang, F. T., Zhang, K., Xue, J., Huang, L., Wang, Y. J., Chen, H. Chen H., et al. (2022). Understanding regional background ozone by multiple methods: A case study in the Shandong region, China, 2018–2020. [Dataset]. Mendeley Data. <https://doi.org/10.17632/ftp9k8cs6f.1>
- Wang, J. Z., Yang, Y., Zhang, Y., Niu, T., Jiang, X., Wang, Y., & Che, H. (2019). Influence of meteorological conditions on explosive increase in O₃ concentration in troposphere. *Science of the Total Environment*, 652, 1228–1241. <https://doi.org/10.1016/j.scitotenv.2018.10.228>
- Wang, Y., Zhang, Y., Hao, J., & Luo, M. (2011). Seasonal and spatial variability of surface ozone over China: Contributions from background and domestic pollution. *Atmospheric Chemistry and Physics*, 11(7), 3511–3525. <https://doi.org/10.5194/acp-11-3511-2011>
- Wu, L., Xue, L., & Wang, W. (2017). Review on the observation-based methods for ozone air pollution research. *Journal of Earth Environment*, 8(6), 479–491. <https://doi.org/10.7515/JEE201706001>
- Xue, L., Wang, T., Louie, P. K., Luk, C. W., Blake, D. R., & Xu, Z. (2014). Increasing external effects negate local efforts to control ozone air pollution: A case study of Hong Kong and implications for other Chinese cities. *Environmental science & technology*, 48(18), 10769–10775. <https://doi.org/10.1021/es503278g>
- Yao, Q., Ma, Z., Hao, T., Fan, W., Yang, X., Tang, Y., et al. (2021). Temporal and spatial distribution characteristics and background concentration estimation of ozone in Beijing-Tianjin-Hebei region. *China Environmental Science*, 41(11), 4999–5008. <https://doi.org/10.19674/j.cnki.issn1000-6923.20210706.007>
- Yin, C., Deng, X., Zou, Y., Solmon, F., Li, F., & Deng, T. (2019). Trend analysis of surface ozone at suburban Guangzhou, China. *Science of The Total Environment*, 695, 133880. <https://doi.org/10.1016/j.scitotenv.2019.133880>
- Zhang, M., Ding, C., Li, Y., Wang, G.-x., Lin, J.-j., Meng, H., & Xu, Y. (2021). Spatial and temporal distribution of ozone and influencing factors in Shandong Province. *Environmental Science*, 42(12), 5723–5735. <https://doi.org/10.13227/j.hjx.202102034>