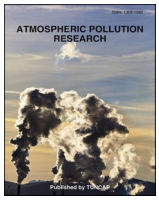
Atmospheric Pollution Research 16 (2025) 102694

Contents lists available at ScienceDirect

Atmospheric Pollution Research

journal homepage: www.elsevier.com/locate/apr



Evaluating the impact of anthropogenic drivers and meteorological factors on air pollutants by explainable machine learning in Shandong Province, China

Yue Yuan [a,b,c,1], Fuzhen Shen [d,*,1], Chunyan Sheng [a,e,**], Zeming Zhang [a,b], Weihua Guo [a,b], Wengang Zhu [a,e], Hui Zhu [a,f]

[a] Key Laboratory of Meteorological Disaster Prevention and Mitigation of Shandong Province, Jinan, 250031, China
[b] Jining Meteorological Bureau, Shandong, 272000, China
[c] Key Laboratory of Ecological Environment and Meteorology of Qinling and Loess Plateau, Shaanxi Meteorological Bureau, Shaanxi, 710016, China
[d] Institute of Climate and Energy Systems: Stratosphere (ICE-4), Forschungszentrum Jülich GmbH, Jülich, 52428, Germany
[e] Shandong Institute of Meteorological Sciences, Jinan, 250031, China
[f] Shandong Meteorological Data Center, Jinan, 250031, China

ARTICLE INFO

Keywords:
Machine learning
SHAP model
Anthropogenic drivers
Meteorological factor

ABSTRACT

Unexpected haze in North China Plain during the COVID-19 lockdown has been regarded as a natural window to explore the meteorological impact on formatting $PM_{2.5}$ pollution but with limitations in quantifying weather elements' contributions. In this study, daily data of six air pollutants (including $PM_{2.5}$, PM_{10}, SO_2, NO_2, O_3, and CO) and six meteorological factors (including temperature, pressure, relative humidity (RH), wind speed (WS), wind direction (WD), and precipitation) from 2015 to 2020 across 16 capital cities in Shandong province, China, was used to drive the Machine Learning and the SHapley Additive exPlanation (SHAP) models. By applying these models, contributions from anthropogenic drivers to pollutant reductions and contributions from meteorological factors to the haze event were investigated. Results show that the COVID-19 lockdown measures reduced concentrations of NO_2, $PM_{2.5}$, PM_{10}, CO and SO_2 by −52.1 %, −40.0 %, −45.5 %, −29.4 % and −38.7 % respectively. On average, an 18.9 % increase in O_3 was observed. $PM_{2.5}$ pollution was mainly driven by temperature with a SHAP value of 19.7 μg/m^3, followed by RH (5.8 μg/m^3), precipitation (0.9 μg/m^3), WD (0.3 μg/m^3), pressure (0.1 μg/m^3) and WS (0.1 μg/m^3) during the haze period. Relative to the post-haze period, high-pressure systems coupled with lower temperatures and weakened surface winds hindered the dispersion of $PM_{2.5}$ whilst higher RH was in favour of $PM_{2.5}$ production during the haze period. This study underscores the intricate interplay between emissions, meteorological conditions, and regulatory measures in air pollution, offering critical insights into future air quality management strategies by air pollution prediction.

1. Introduction

Air pollution remains a critical global environmental challenge with substantial implications for public health, ecosystem integrity, and the global economic sustainability (Nel, 2005; Shen et al., 2020). According to the World Health Organization (WHO), approximately 7 million premature deaths occur annually as a result of exposure to air pollutants (Kuehn, 2014), with fine particulate matter ($PM_{2.5}$) and ground-level ozone (O_3) identified as major contributors (Lelieveld et al., 2015; Shen et al., 2024). These pollutants are strongly associated with increased incident of respiratory and cardiovascular diseases, highlighting the urgent necessity for effective air quality management strategies at both national and global levels (Burnett et al., 2014; Lelieveld et al., 2015).

The outbreak of the COVID-19 pandemic in late 2019 and the subsequent implementation of widespread lockdown restrictions in early 2020 resulted in drastic reductions in industrial operations, vehicular traffic, and other anthropogenic activities. Those abrupt changes led to

* Corresponding author.
** Corresponding author. Key Laboratory of Meteorological Disaster Prevention and Mitigation of Shandong Province, Jinan, 250031, China.
E-mail addresses: f.shen@fz-juelich.de (F. Shen), sdqxscy@126.com (C. Sheng).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.apr.2025.102694
Received 9 March 2025; Received in revised form 6 August 2025; Accepted 12 August 2025
Available online 14 August 2025

notable declines in nitrogen dioxide (NO_2) and particulate matter (PM) concentrations (Shen et al., 2022). Numerous studies have confirmed that both variations in emissions and meteorological conditions affect regional air quality significantly (He et al., 2017; Shen et al., 2022). Despite the significant reduction in anthropogenic emissions, several studies also reported that the North China Plain (NCP) still experienced an unexpected haze event during the early 2020 lockdown. This abnormal haze has been attributed to complex atmospheric chemistry processes and special weather conditions (Le et al., 2020; Shen et al., 2024). As a result, the combination of COVID-19 lockdown measures induced emission reductions and the concurrent severe haze event could be regarded as a 'natural window' to better disentangle the respective roles of anthropogenic emission reduction and meteorological drivers in air pollution, thereby informing more effective mitigation strategy. To evaluate the contributions of anthropogenic emission reduction and meteorology to the change in air pollutants during the COVID-19 lockdown, Machine Learning (ML) models (Rybarczyk and Zalakeviciute, 2021; Shi et al., 2021; Wang et al., 2020b; Yang et al., 2021) and chemistry transport model (CTM) simulations (Li et al., 2020; Lv et al., 2020; Marlier et al., 2020; Xing et al., 2020) are two effective approaches by designing comparison between different scenarios, such as a 'business as usual' (BAU) scenario to the factual scenario. In contrast, the ML model outperforms the CTM model due to the latter's reliance on outdated emission inventories and limited spatial resolution. Moreover, ML algorithms can also capture the complex nonlinear relationships between predictors and local features' characteristics (Shen et al., 2022; Xiao et al., 2018). At last, with the exception of the above advantages, ML models can be integrated with explainable AI tools, such as SHapley Additive exPlanations (SHAP) module, which allow for the quantification of feature importance and enhance the physical interpretability of key drivers or major contributors to air pollution (Hou et al., 2022; Men et al., 2023).

Shandong Province, located on the eastern coast of China, is among the most industrialized regions in the country and has experienced persistent air pollution challenges due to its diverse industrial sectors and coal-dominated energy structure (Weng et al., 2022; Zhou et al., 2021). A national assessment of air quality in 168 major cities in China reported that five cities in Shandong, including Zaozhuang, Linyi, Liaocheng, Jinan and Zibo, ranked among the bottom 20 for air quality in 2019 (MEEP, 2019). Moreover, several cities within the province, including Jinan, Zibo, Jining, and Dezhou, are located along a key air pollution transport channel of the NCP region, further exacerbating regional air quality issues in Shandong Province (Zhou et al., 2021). Given these characteristics, Shandong Province could serve as an ideal study domain for investigating the individual roles of anthropogenic emission reduction and meteorology in the local air quality change during the COVID-19 lockdown.

To evaluate the contribution of anthropogenic emission reduction to the air quality change in Shandong Province, a machine learning framework based on the Gradient Boosting Machine (LightGBM) was employed to construct a BAU scenario during the COVID-19 lockdown. Subsequently, the SHAP interpretability module was coupled with the ML model to identify the major meteorological factors contributing to the haze event in this region. Additionally, to better understand how various meteorological factors synergistically to promote haze formation, the impact of atmospheric physical fields on $PM_{2.5}$ changes was analyzed as well.

2. Materials and methods

2.1. Station data of air pollutants and meteorological factors

Shandong province (114°47.5′ - 122°42.3′E, 34°22.9′ −38°24.01′ N), located along the eastern coast of China (Fig. S1), was selected as the research domain. Hourly mass concentrations of six air pollutants, including $PM_{2.5}$, PM_{10}, SO_2, NO_2, O_3, and CO, were collected across 16 cities from China National Environmental Monitoring Center (NEMC) (http://www.cnemc.cn/). Those hourly measurements were aggregated into daily values from January 1, 2015, to April 30, 2020. Data quality control procedures were implemented in accordance with the methodology described in Shen et al. (2020). The spatial distribution of air quality sampling sites across 16 cities in Shandong Province is shown in Fig. 1. Daily surface meteorological factors (e.g., Temperature: T, °C; Relative Humidity: RH, %; Wind Speed: WS, m/s; Wind Direction: WD, °; Precipitation: Precip, mm; Pressure: Pres, hPa) were also collected for the same period. For the ML model development, the dataset was divided into the training part (January 1, 2015–December 30, 2019) and the test part (January 1, 2020–April 30, 2020). Note that among 123 National Meteorological Stations in Shandong Province, only 16 stations were selected because they are situated closest to the 16 air pollutant monitoring sites. Meteorological data can be accessed from the National Meteorological Science Data Center (https://data.cma.cn/).

2.2. ERA5 dataset

In addition, hourly surface meteorological reanalysis data from the ERA5 dataset, provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), were used. Those datasets have a spatial resolution of 0.25° × 0.25° and include variables such as 2m surface temperature, dew-point temperature (T_d, °C), 10m wind, mean sea level pressure (MSLP) and geopotential height, RH, wind, and air temperature at pressure levels from 850 to 1000 hPa. The data were accessed from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/).

2.3. Machine learning (ML) model

The LightGBM, an ensemble ML method based on decision trees, was employed to predict the mass concentrations of six air pollutants in BAU in 2020. The LightGBM has demonstrated strong capability in handling complex nonlinear relationships, performing internal feature selection, and efficiently providing feature importance (Ke et al., 2017). Moreover, its ability to capture site-specific features makes it particularly suitable for predicting pollutants in diverse cities, including those lacking emission inventories or located in the desert and plateau regions (Shen et al., 2022). The input dataset of ML model not only comprises six meteorological factors and six air pollutants but also includes four time variables (e.g., Julian Day (JD), Day of Week (DOW), Holidays, and Chinese New Year (CNY)) from 2015 to 2020. To select an optimal ML model, cross-validation by using the time-series split rolling method executed before the implementation of the final predictions. After the step of cross-validation, the performance of the ML model was evaluated based on root-mean-squared error (RMSE) and Pearson correlation coefficient (PCC). Typically, a ML model's optimal performance is defined by the lowest RMSE and the highest PCC values.

2.4. SHapley additive exPlanation (SHAP) method

Measuring the importance of input features for ML model is as vital as attaining high predictive accuracy. To achieve this goal, the SHAP method, based on game theory, is used to quantify the importance of each feature input (Lundberg et al., 2020). SHAP measures the impact of each feature by comparing prediction outcomes with and without the feature (Hou et al., 2022). This approach provides a detailed understanding of each variable's contribution to the model's predictions. For a model with n variables applied to a sample (x_i) and its predicted output $f(x_i)$, the prediction function can be expressed as follows:

$$f(x_i) = E_0(f, x) + \sum_{m=1}^{n} E_m(f, x_i) \quad (1)$$

where x_i represents the input with variable m in the prediction model f, producing the SHAP value $E_m(f, x_i)$. $E_0(f, x)$ is the expected value for the prediction model over the entire dataset. In general, the SHAP value

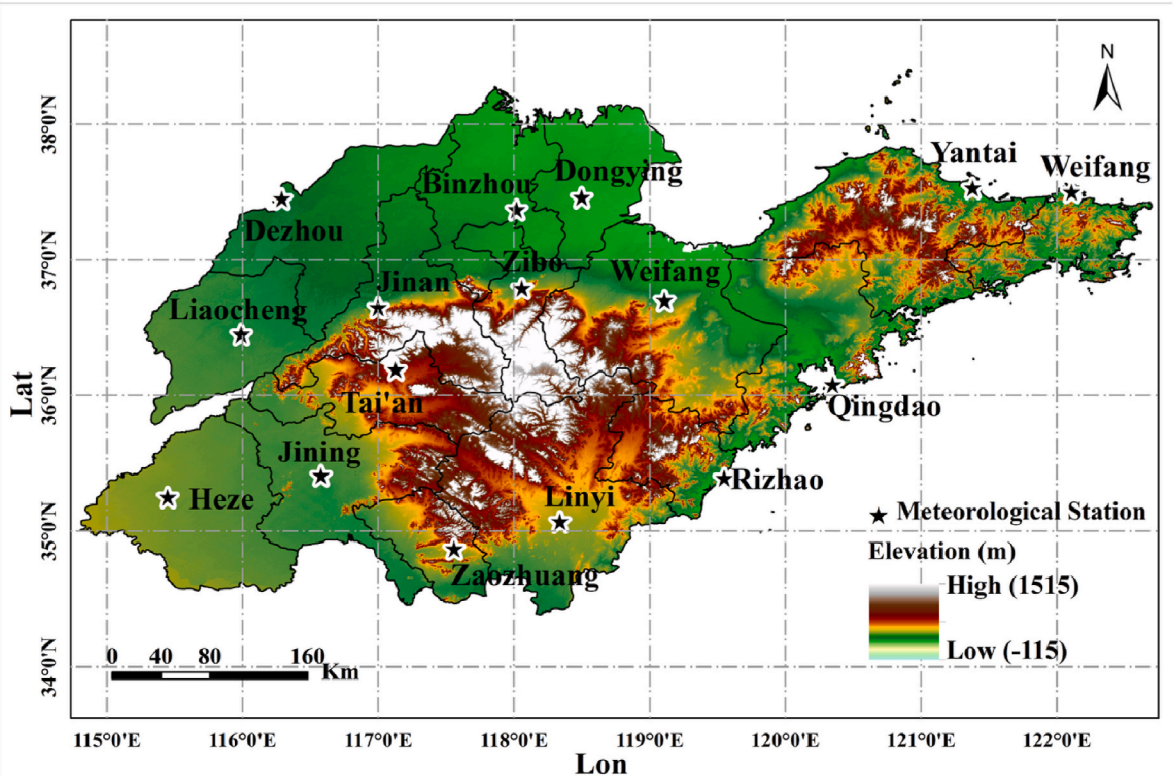


Fig. 1. The locations of 16 surface meteorological observation sites in Shandong Province, China. The color bar indicates the altitude in Shandong.

explains the contribution of each variable to the prediction. This contribution is measured from the base value (the predicted mean of the expected value) to the final model output. Variables that increase the prediction are shown as positive, while those that lower it are displayed as negative.

2.5. Specification of the study period

In early 2020, the COVID-19 outbreak prompted local authorities nationwide to implement diverse restrictive measures. These measures resulted in a significant reduction in population movement, industrial production and economic activities (Nie et al., 2021). Notably, on January 25, 2020, Shandong Province initiated a Level I response to address a major public health emergency. Subsequently, on 8 March, the emergency response level for preventing and controlling the Candidatus Pneumonia outbreak in Shandong Province was elevated to Level II. Considering the above two different levels, we defined the entire Level I period as the 'Total COVID-lockdown' for analysis in section 3.1. Detailed classifications of the different lockdown phases are shown in Table S1. The CNY period, defined by the lunar calendar, typically starts about a week before and lasts approximately 10 days after the CNY's day. The periods labeled 'Before CNY', 'Extended COVID-lockdown', and 'Total COVID-lockdown' are classified within the first three months of 2020.

During the COVID-19 lockdown period, the NCP region experienced an abnormal haze pollution. To better understand and quantify the contribution of meteorological factors to the haze event, the interval from January 21 to February 9, 2020 was defined as the 'During-haze' period; the time before this interval is categorized as 'Before-haze' and the time after identified as 'Post-haze'. Detailed definitions of these period are provided in Table 1.

Table 1
Classification and characteristics of haze-related period in early 2020.

Label	Event Name	Time Span	Characteristics
P1	Before-haze	January 1-January 20, 2020	High emissions; high $PM_{2.5}$ level; moderate meteorology
P2	During-haze	January 21-February 09, 2020	Low emissions; high $PM_{2.5}$ levels; high RH; moderate T
P3	Post-haze	February 10-February 29, 2020	Low emissions; low $PM_{2.5}$ level; Precip; high T; high WS

3. Results

3.1. Quantify contributions of different drivers for six air pollutants

Fig. 2 demonstrates not only the observed but also the predicted mass concentrations of all six air pollutants during the COVID-19 lockdown. Except for O_3, the time series of the other air pollutant concentrations all exhibited a fluctuating downward trend in early 2020. This decline is attributed to the related epidemic closure and/or control measures and associated anthropogenic emission reduction of the primary pollutants and the precursors of secondary pollutants (Wang et al., 2021a). In contrast, the continued increase in O_3 could be attributed to the reduction of NOx in the VOC-limited O_3 chemistry regime, which results in the O_3 enhancement due to decreased titration by NO (Shen et al., 2022; Wang et al., 2020b). Among six air pollutants, their predictions and the corresponding observations agree with each other well, with the highest PCC of 0.86 for SO_2 and CO, followed by PCC of 0.84 for O_3, 0.78 for PM_{10}, 0.77 for $PM_{2.5}$ and 0.76 for NO_2, the average PCC of which is above the previous studies in China (PCC = 0.71) (Shen et al., 2022) and Spain (PCC = 0.72) (Petetin et al., 2020), indicating robust predictive performance by the ML model. Note that before selecting the optimal ML model, the time series cross validation was applied in four experiments, with the training datasets covering 2015, 2015–2016, 2015–2017,

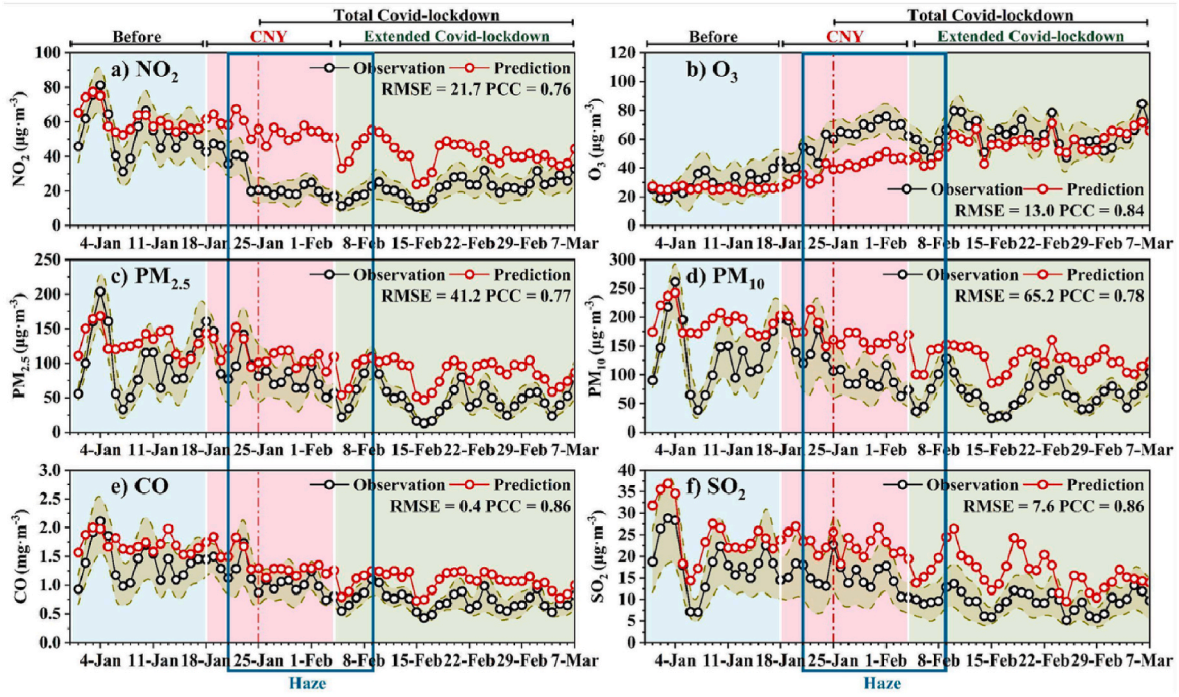


Fig. 2. Time series distribution of observed and predicted daily mean mass concentrations of six air pollutants in Shandong Province. The black circle-line indicates observed concentrations, the red circle-line indicates predicted concentrations and the blue rectangle indicates the haze event in 2020.

2015–2018 and 2015–2019, and the test datasets covering the first three months of 2016, 2017, 2018, 2019 and 2020, respectively. As shown in Fig. S2-S7, six pollutants' mass concentration predictions generally agree with the observations well except for some extreme high or low values generally due to anthropogenic emission changes suddenly.

To evaluate the influence of various anthropogenic drivers, fractional changes in pollutant concentrations between the BAU and factual scenarios were calculated for different periods (Table 2). Before the CNY period, the ML model systematically overestimated the concentrations of pollutants (excluding O_3) relative to the observed six air pollutants' mass concentrations (Fig. 2). These overestimations likely reflect the effect of local or national air quality improvement regulations, such as the national Clean Air.

Plan in 2018–2020 ($CAP_{2018\text{-}2020}$) (Shen et al., 2022). On average, these measures have led to a reduction of −31.1 % for PM_{10}, followed by reductions of −26.4 %, −23.5 %, −18.8 % and −11.7 % for SO_2, $PM_{2.5}$, CO and NO_2, respectively. For O_3, the reduction of O_3 chemistry precursors, especially NOx, attributable to CAP measures could trigger increased O_3 production (12.6 %) due to its nonlinear response characteristic of VOC-limited photochemical regimes (Zhang et al., 2022). At the beginning of the CNY period, a larger gap between observed and predicted concentrations of NO_2, SO_2, PM_{10} and O_3 started to appear due to the CNY effect. The above gaps further to widen until the end of the CNY period attributable to concurrent CNY effect and COVID-19 lockdown effect. However, this widening gap was not observed for $PM_{2.5}$ in the same period, indicating the presence of additional production of $PM_{2.5}$ driven by specific meteorological conditions (Fig. 2c). During the whole CNY period, the reduction portion of NO_2 (−51.3 %) was higher than that (−47.4 %) in the extended COVID-lockdown period due to the compound effect of the CNY related and the specific COVID-19 lockdown related reductions. In response to the biggest NO_2 reduction during this special period, O_3 production became more intensive and led to a much higher increase (51.2 %) in O_3 concentration than that (10.0 %) in the extended COVID-lockdown period (Sun et al., 2020). By contrast, reductions in $PM_{2.5}$ (−21.7 %), PM_{10} (−31.8 %), SO_2 (−32.7 %) and CO (−19.7 %) during the CNY period were relative modest and less than those during the extended COVID-lockdown period ($PM_{2.5}$: 43.8 %; PM_{10}: 46.3 %; CO: 32.2 %; SO_2: 42.1 %). Different magnitudes of reduction in PM, CO and SO_2 in the above three periods could not only be attributed to the primary emission reductions but also to secondary formation process such as the gas-to-particle conversions (Chen et al., 2020). These processes are known to be enhanced under stagnant meteorological conditions (Sun et al., 2020).

To better investigate the air quality situation during the haze event, the spatial distribution of each pollutant was investigated in Shandong Province. As shown in Fig. 3, except for O_3, the mass concentrations of air pollutants were higher in the western regions compared to the coastal areas. The average $PM_{2.5}$ concentration across Shandong was 76.2 μg/m^3, with particularly elevated levels observed in Dezhou (101.8 μg/m^3), Liaocheng (101.5 μg/m^3), and Heze (106.3 μg/m^3), where values exceeded 100 μg/m^3. Similarly, the average PM_{10} concentration in Shandong was 96.7 μg/m^3, with the highest record in Heze at 142.0 μg/m^3. In contrast, elevated O_3 concentrations were found mainly concentrated in southern Shandong, with the highest value (92.1 μg/m^3) in Jining. Notably, during the period of the haze event in Shandong, only the average mass concentration of $PM_{2.5}$ exceeded the China Ambient Air Quality Standards Grade II (CAAQS-II) threshold of 75 μg/m^3 on average, while the levels of other pollutants remained below their corresponding CAAQS-II thresholds. These results underscore the severity of $PM_{2.5}$ pollution in Shandong Province during this period, particularly in the inland areas.

Table 2
The COVID-lockdown effects in different time periods.

Pollutants	Before CNY	CNY	Extented COVID-lockdown	Total COVID-lockdown
NO_2	−11.7 %	−51.3 %	−47.4 %	−52.1 %
O_3	12.6 %	51.2 %	10.0 %	18.9 %
$PM_{2.5}$	−23.5 %	−21.7 %	−43.8 %	−40.0 %
PM_{10}	−31.1 %	−31.8 %	−46.3 %	−45.5 %
CO	−18.8 %	−19.7 %	−32.2 %	−29.4 %
SO_2	−26.4 %	−32.7 %	−42.1 %	−38.7 %

Effect (%) = ((Obs - Pre)/Pre) × 100.

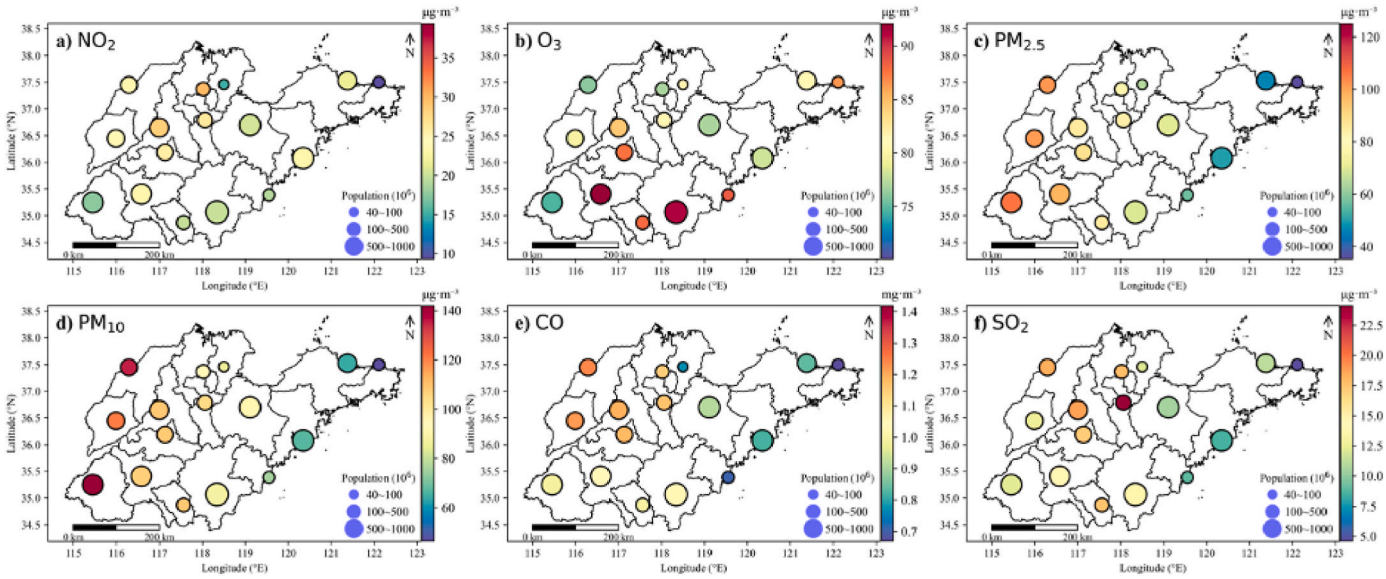


Fig. 3. The spatial distributions of observed mean values of six air pollutants during the haze event in 2020. The size of blue solid circle indicates the population size in Shandong Province.

3.2. *Quantify contributions of meteorological factors for air pollutants*

Following the evaluation of the contributions of various drivers for six air pollutants, a SHAP analysis was employed to quantify the importance of the model's input features in the BAU scenario. Fig. 4 shows the time series of SHAP values for all ML model's input features during the COVID-19 lockdown. Notably, time variables (JD, DOW, Holiday, CNY) were used as independent predictors. The time predictors were primarily used as proxies for emissions: the year reflects changes in emission factors, while JD, DOW, Holiday and CNY capture emission trends a day, a week, a holiday, and CNY days respectively (Wang et al., 2020b). Overall, the CNY variable showed a negative mean SHAP value for NO_2 (-3.0 μg/m^3), indicating the driver of the CNY holiday plays a dominant role in NO_2 reduction during the haze period (Fig. 4a). During the Total COVID-lockdown, in terms of negative SHAP, JD was the most significant contributor to $PM_{2.5}$ (-9.7 μg/m^3), PM_{10} (-12.9 μg/m^3), CO (-0.3 μg/m^3) and SO_2 (-14.5 μg/m^3) among all the features, reflecting the ongoing daily emission reductions of those four pollutants in Shandong Province (Fig. 4c–f). For O_3, the SHAP value of temperature (-13.6 μg/m^3) was the highest, followed by pressure (-5.1 μg/m^3), RH (-0.4 μg/m^3), WS (-0.1 μg/m^3) and WD (-0.1 μg/m^3) in Shandong (Fig. 4b). It reveals that those meteorological factors push O_3 prediction

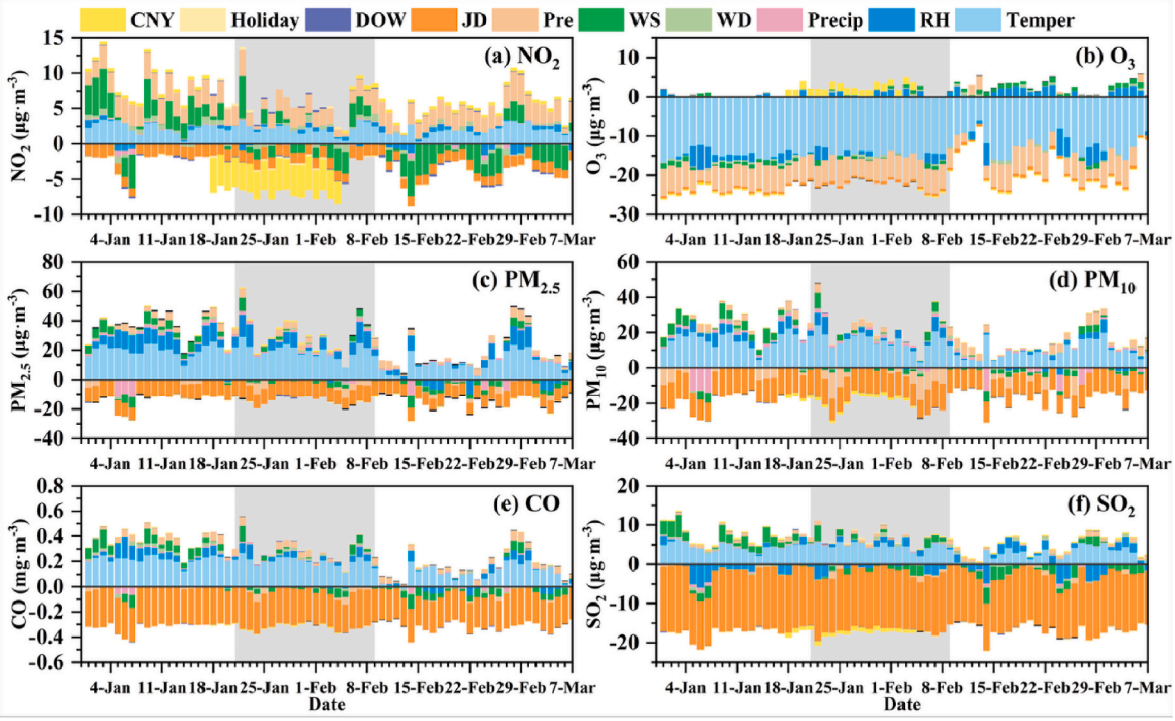


Fig. 4. Time series distribution of daily SHAP values of input variables (colored bar) for six air pollutants during the COVID-19 lockdown in Shandong, respectively (Note that the shaded area indicates the haze event period).

to a lower value, suggesting that these variables suppressed O_3 toward the baseline levels in Shandong (Fig. 2b). In contrast, SHAP values of temperature for $PM_{2.5}$ (16.5 μg/m^3), PM_{10} (13.2 μg/m^3), SO_2 (3.9 μg/m^3) and CO (0.2 μg/m^3) were still ranked the first among the six meteorological factors but turned to positive, indicating its role in promoting pollutant accumulation (Fig. 2c–f).

To better understand meteorology's impact rather than emission reductions on haze formation, we further compared the mean SHAP values of each meteorological factor for $PM_{2.5}$ between during haze (P2) and after haze periods (P3) (Fig. 5). It should be noted that P2 and P3 coincided with first-level emergency response in Shandong Province. On average, Shandong Province experienced an obvious $PM_{2.5}$ pollution phase with a value of 76.3 μg/m^3 in P2 and a clean phase with a value of 45.6 μg/m^3 in P3, respectively (Fig. 6). During P2, the temperature had a positive SHAP value for $PM_{2.5}$ (+19.7 μg/m^3), followed by RH (+5.8 μg/m^3), Precip (+0.9 μg/m^3), WD (+0.3 μg/m^3), Pres (+0.1 μg/m^3) and WS (+0.1 μg/m^3), respectively (Fig. 5c). In comparison, the SHAP values for meteorological factors in P3 showed a significant difference, with lower values for temperature (+11.3 μg/m^3) and RH (+1.6 μg/m^3), a higher value for Pres (+1.8 μg/m^3), and opposite values for Precip (−0.6 μg/m^3), WD (−0.1 μg/m^3) and WS (−1.7 μg/m^3), respectively (Fig. 5c). For temperature, the mean value in P2 (1.7 °C) is lower than that in P3 (5.1 °C), contributing to a lower planetary boundary layer height (PBLH) in P2, which facilitates the accumulation of $PM_{2.5}$ (Fig. 6f). For precipitation and wind, higher precipitation (P3:1.6 mm vs P2: 0.1 mm) and stronger winds (P3: 2.8 m/s vs P2: 2.4 m/s) in P3 promoted wet deposition and transportation, thereby leading to lower concentration of $PM_{2.5}$ (Fig. 6b–d).

3.3. Analyses disturbance of meteorological conditions during haze event

Weather systems played a critical factor in driving the haze event in Shandong Province during the COVID-19 lockdown. To elucidate their influence, we further compare the weather systems between the P2 and P3 periods, focusing on how temporal and spatial variations in atmospheric circulation and physical fields affect haze formation and dissipation (Figs. 7 and 8). At 850 hPa level (Fig. 7a–b), RH exhibited significant difference in southwestern Shandong, where RH decreased from 60 to 70 % in P2 to 50–60 % in P3. Similarly, across the 925–1000 hPa layers (Fig. 7c–f), RH declined by approximately 10 % in P3 relative to P2, indicating a province-wide reduction in atmospheric moisture. This decrease in RH highlights a pronounced decline in humidity conditions from P2 to P3, which likely limited secondary aerosol formation, contributing to the observed decrease in $PM_{2.5}$. During P2, the MLSP

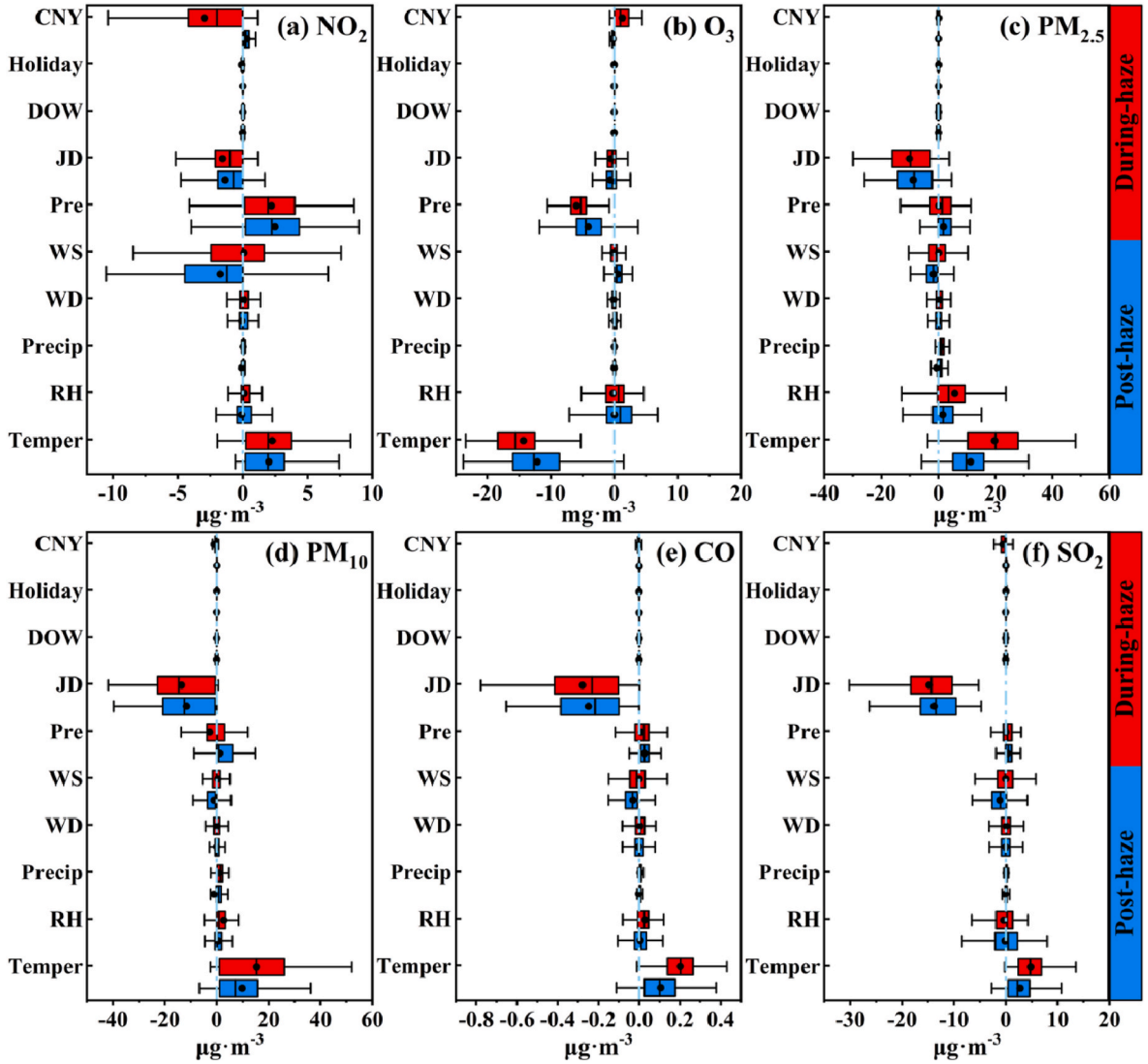


Fig. 5. The box-and-whisker plots represent the mean SHAP value of the input variables for six air pollutants during the haze (red color) and post haze period (blue color) in Shandong Province.

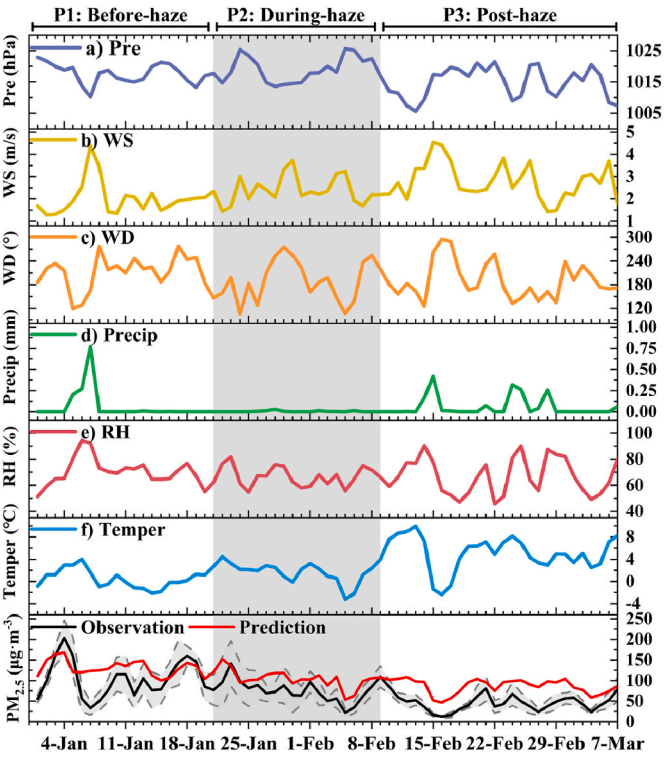


Fig. 6. Time series distribution of observed mean values of six meteorological factors and $PM_{2.5}$ during the COVID-19 lockdown period (Note that the shaded area indicates the haze event period).

field revealed a dominant high-pressure system over the NCP region, northeastern China, and Shandong Province. Shandong Province, located at the base of this high-pressure system, experienced strong horizontal divergence, which suppressed vertical atmospheric motion and trapped pollutants near the surface (Fig. 7g). In contrast, P3 saw a transition to a broader uniform-pressure environment (Fig. 7h), weakening atmospheric stability and providing more favorable conditions for pollutant dispersion. Notably, the temperature-dew point difference ($T - T_d$) approached zero in P3 (Fig. 7h), primarily due to precipitation-induced saturation, which facilitated the removal of $PM_{2.5}$ from the atmosphere. For surface wind field, inland Shandong in P2 experienced persistently low wind speeds, with near-calm conditions dominating in the southwestern region (Fig. 7g). These surface-stagnant meteorological conditions in P2, relative to P3, significantly hindered horizontal pollutant dispersion, leading to the accumulation of particulate matter. In terms of temperature from P2 to P3, as shown in Fig. 8, it increased by 2 °C–4 °C at the 850 hPa level and at the surface (2 m above ground), enhancing PBLH and improving vertical dispersion conditions, which contributed to better surface air quality. In addition to the influence of local meteorological factors on $PM_{2.5}$ pollution in P2, regional pollution transport also played a crucial role. This external pollution influx in P2 was driven by a pronounced warm ridge over the NCP region (Fig. 8a) coupled with prevailing northwesterly winds at 850–925 hPa (Fig. 7a–c), resulting in the transport of pollutants from the NCP region to Shandong.

Overall, specific weather patterns under the impact of different weather elements concurrently were attributed to the $PM_{2.5}$ accumulation and dispersion in Shandong Province during P2 and P3 periods. These results consistently highlight the critical role of weather systems in haze formation and provide helpful insights into meteorological modulation in air pollution in Shandong Province.

4. Discussions

4.1. Anthropogenic drivers for the air pollution change during the COVID-19 lockdown period

Regarding the anthropogenic contributors, the COVID-19 lockdown measures, the $CAP_{2018\text{-}2020}$ policy, and the CNY holiday all made a certain contribution to the change in the major pollutants. Previous studies have demonstrated that both routine and special societal slowdowns, such as the CNY holiday, CAP implementation, and high-profile events like the Olympic Games and G20 Summits, are typically associated with decreased industrial and vehicular activity, subsequently leading to reductions in pollutant emissions (Pei et al., 2020). For instance, Wang et al. (2010) estimated that over 60 % of $PM_{2.5}$ reduction during the 2008 Olympics could be attributed to emissions control strategies. Similarly, Yan et al. (2012) reported lower near-surface SO_2 concentrations in the Pearl River Delta during the 2010 Asian Games. Huang et al. (2015) found a 47 % reduction in tropospheric NO_2 concentrations over the North China Plain during the 2014 APEC meeting

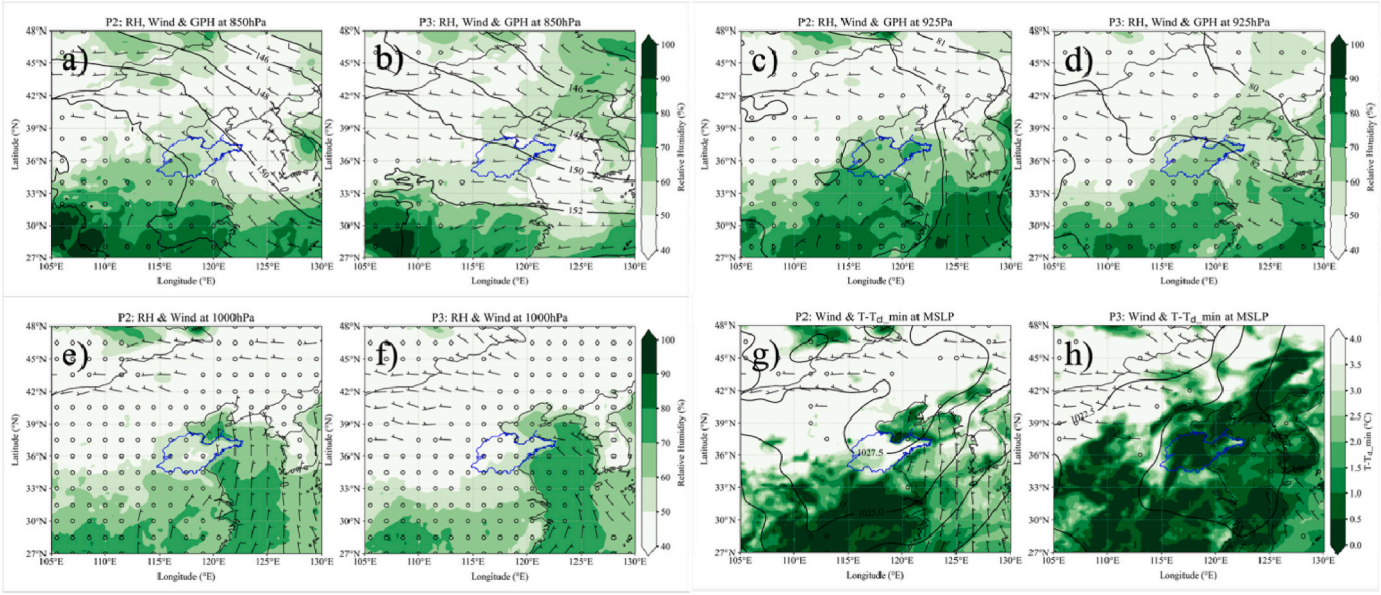


Fig. 7. Comparison of the RH and wind fields at 850 hPa, 925 hPa, 1000 hPa levels and T-Td field at MSLP between During-haze and Post-haze periods, respectively.

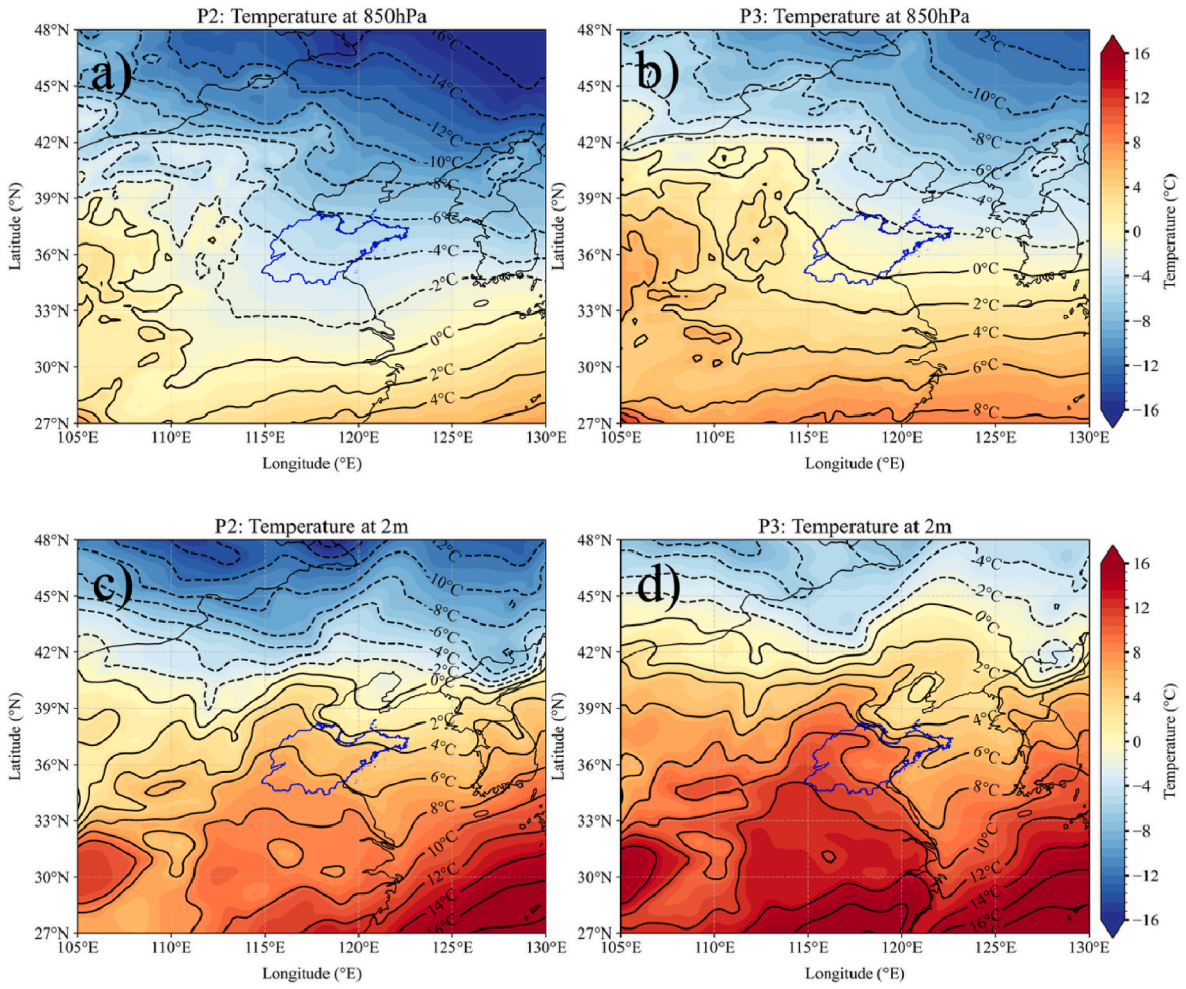


Fig. 8. Comparison of temperature fields at 850 hPa and 1000 hPa levels between During-haze and Post-haze periods, respectively.

relative to the same period in the preceding three years. Li et al. (2016) showed a 52–57 % reduction in PM_1 mass concentration during the 2015 Victory Day Parade after controlling for meteorological variability.

Over the past decade, a major routine anthropogenic intervention has been China's Clean Air policy initiatives. In 2013, the Chinese State Council launched a five-year Air Pollution Prevention and Control Action Plan (referred to as the "Ten Measures of Air"), targeting $PM_{2.5}$ reduction by the end of 2017 (Shen et al., 2020). Building on this, the government issued the "Three-Year Action Plan to Win the Blue Sky Defense War" in 2018 to address $PM_{2.5}$ and O_3 pollution jointly (Shen et al., 2022). Following implementation of these policies, concentrations of $PM_{2.5}$, PM_{10}, NO_2, SO_2, and CO decreased by 46.2 %, 43.1 %, 24.7 %, 55.5 %, and 37.2 %, respectively. In contrast, 8-h O_3 concentrations increased by 19.4 %, highlighting the complexity of secondary pollutant formation(Li et al., 2023).

The CNY holiday represents another routine social pause with moderate emission impacts. This social pause generally starts one week before the Chinese New Year Day and lasts two weeks or more (Miyazaki et al., 2020). Although industrial activity declines during this period, many manufacturing operations continue, and traffic flows shift—often moving from urban centers to rural areas(Guo et al., 2018). Unlike these routine events, the COVID-19 lockdown was an unprecedented, nationwide cessation of both industrial production and transportation for over a month(Tian et al., 2020). This unique event offered a rare opportunity to evaluate baseline atmospheric conditions in the absence of typical anthropogenic influences. Numerous studies have since reported heterogeneous air quality responses across regions during the lockdown (Nie et al., 2021; Shen et al., 2024; Wang et al., 2021a). Our findings presented here underscore the value of continued investigation into the differential impacts of anthropogenic activity on atmospheric composition.

4.2. *Influence of meteorological input features on haze event*

To better understand the influence of meteorological conditions on haze formation by SHAP in Shandong Province, the contribution of each input feature to local $PM_{2.5}$ concentration predictions was examined while accounting for intra-model variability. Here, intra-model variability of SHAP values refers to differences in SHAP attributions for the same feature value across different phases of a haze event, reflecting the complex mechanisms governing $PM_{2.5}$ formation and removal.

During phase P2, the highest average SHAP value for ambient temperature indicated a strong positive association with $PM_{2.5}$ concentrations, followed by RH (Fig. 5c). The statistical model successfully reproduced established patterns linking $PM_{2.5}$ species variability to temperature. For instance, sulfate formation is known to increase at higher temperatures due to enhanced oxidation rates of SO_2 (Li et al., 2017). Chen et al. (2025), using SHAP analysis, suggested that temperature is the dominated meteorological factors influencing $PM_{2.5}$ levels through changing the atmospheric oxidation capacity in Beijing-Tianjin-Hebei (BTH), China. In this study, the higher $PM_{2.5}$ levels observed in P2 at lower average temperatures (1.7 °C) compared to P3 (5.1 °C) are inconsistent with this phenomenon. This pattern is likely related to enhanced aqueous-phase sulfate formation under colder

and foggier conditions in Beijing (Wang et al., 2020a). Similarly, Stirnberg et al. (2021) reported other potential mechanisms at lower temperatures aligned with the SHAP results, such as the significant nitrate formation and the enhanced organic matter, which are the two well-documented processes (Petit et al., 2015; Wang et al., 2021b).

The relatively higher RH SHAP value for $PM_{2.5}$ also played a positive role in the high $PM_{2.5}$ levels during P2. RH could influence the formation mechanisms of nitrate and sulfate by promoting heterogeneous reactions and serves as an indicator of those occurrence (Ge et al., 2012; Sun et al., 2013). Liu et al. (2025), in a study of exploring the drivers of $PM_{2.5}$ pollution by the SHAP method, have suggested that SO_2 oxidation could be accelerated in a condition of RH exceeds 50 % but suppressed with RH above 90 %, likely due to the rainfall-induced scavenging of $PM_{2.5}$. The observed difference in SHAP value for $PM_{2.5}$ between P2 and P3 is consistent with this mechanism. In P2, the lower average RH (66.6 %) with respect to P3 likely lead to more $PM_{2.5}$ formation, as several rainfall events occurred in P3 (Fig. 6d).

4.3. Limitations

This study quantified the contributions of anthropogenic drivers and meteorological factors to pollutant reductions using machine learning and SHAP analysis. While, there are several limitations in disentangling anthropogenic drivers.

First, the treatment of anthropogenic drivers in this study was indirect and primarily represented by time-related variables (e.g., holidays, Chinese New Year, Julian day). Although these proxies can capture variations in human activity, they do not provide direct information about changes in specific emission sources or anthropogenic behaviors. As a result, the quantification of anthropogenic drivers may be less precise than the assessment of meteorological factors. Moreover, the study did not explicitly categorize anthropogenic drivers by emission sector (e.g., industry, transportation, residential combustion). The reliance on a residual-based approach derived from machine learning predictions makes the interpretation of anthropogenic contributions more abstract. Finally, the lack of direct emission data or more detailed quantitative indicators of social and economic activities limits the ability to fully separate and attribute anthropogenic effects. Future studies could address this limitation by designing a framework that integrates multi-data fusion analysis (e.g., detailed emission inventories, sectoral activity data, or real-time indicators of anthropogenic activity) using an ensemble machine learning method to improve attribution accuracy.

5. Conclusion

This study evaluates the contributions of different drivers and meteorological factors to the changes in six air pollutants at 16 surface sampling sites in Shandong Province during the COVID-19 lockdown. From ML results, it was found that the combined effect of the Total COVID-lockdown measures and regular CNY celebration contributed to an overall 52.1 % reduction in NO_2, followed by reductions of 45.5 %, 40.0 %, 38.7 %, and 29.4 % for PM_{10}, $PM_{2.5}$, SO_2 and CO, respectively. In contrast, an increase of 18.9 % in O_3 was observed during the Total COVID-lockdown in Shandong. Based on SHAP values and the analysis of meteorological disturbances, we found that temperature, with the SHAP values of 19.7 $\mu g/m^3$ in P2, was the dominant meteorological factor in the weather system, facilitating the formation of the haze event during the COVID-19 lockdown period in Shandong.

Overall, this study provides robust insights into the impact of meteorological factors on air pollution by quantifying their contributions during the COVID-19 lockdown in Shandong Province. By analyzing the weather patterns during the same period, it further proves the accuracy and effectiveness of the ML model and SHAP module, which aid in the development of more effective air quality management strategies. This approach highlights the potential for using data-driven methods to design more effective policies for improving air quality and safeguarding public health. However, in order to improve the performance of the ML model, our work needs to be extended to include other related meteorological factors.

CRediT authorship contribution statement

Yue Yuan: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Fuzhen Shen:** Writing – review & editing, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. **Chunyan Sheng:** Software, Resources, Investigation, Formal analysis. **Zeming Zhang:** Software, Resources, Formal analysis, Data curation. **Weihua Guo:** Visualization, Software, Resources, Data curation. **Wengang Zhu:** Visualization, Software, Resources, Formal analysis. **Hui Zhu:** Visualization, Software, Resources, Formal analysis.

Code availability

The python code and data used to run the LightGBM model can be obtained upon request from the corresponding authors.

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.

Acknowledgments

This work was supported by the Scientific Research Project of Shandong Meteorological Bureau (SDYBY2020-08, 2023SDQN16, 2023sdqxz05, 2023sdqxz08, 2024sdqxz11), the Meteorological Science and Technology Research Project of the Jining Meteorological Bureau (2023JNZL03, 2023JNZL05), and the Key Laboratory of Eco-Environment and Meteorology for the Qinling Mountains and Loess Plateau, Shaanxi Meteorological Bureau (2023Y-22). The project also received support from the Artificial Intelligence Meteorological Application Technology Innovation Team (SDCXTD2023-3).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apr.2025.102694.

References

Burnett, R.T., Pope III, C.A., Ezzati, M., Olives, C., Lim, S.S., Mehta, S., Shin, H.H., Singh, G., Hubbell, B., Brauer, M., 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ. Health Perspect. 122, 397–403.

Chen, H., Huo, J., Fu, Q., Duan, Y., Xiao, H., Chen, J., 2020. Impact of quarantine measures on chemical compositions of PM2.5 during the COVID-19 epidemic in Shanghai. China. Sci. Total Environ. 743, 140758.

Chen, M., Liu, J., Chu, B., Zhao, D., Li, R., Chen, T., Ma, Q., Wang, Y., Zhang, P., Li, H., 2025. Research on the influencing factors of PM2. 5 in China at different spatial scales based on machine learning algorithm. Environ. Int., 109536

Ge, X., Zhang, Q., Sun, Y., Ruehl, C.R., Setyan, A., 2012. Effect of aqueous-phase processing on aerosol chemistry and size distributions in Fresno, California, during wintertime. Environ. Chem. 9, 221–235.

Guo, Z., Zhou, K., Zhang, X., Yang, S., Shao, Z., 2018. Data mining based framework for exploring household electricity consumption patterns: a case study in China context. J. Clean. Prod. 195, 773–785.

He, J., Gong, S., Yu, Y., Yu, L., Wu, L., Mao, H., Song, C., Zhao, S., Liu, H., Li, X., 2017. Air pollution characteristics and their relation to meteorological conditions during 2014–2015 in major Chinese cities. Environ. Pollut. 223, 484–496.

Hou, L., Dai, Q., Song, C., Liu, B., Guo, F., Dai, T., Li, L., Liu, B., Bi, X., Zhang, Y., Feng, Y., 2022. Revealing drivers of haze pollution by explainable machine learning. Environ. Sci. Technol. Lett. 9, 112–119.

Huang, K., Zhang, X., Lin, Y., 2015. The "APEC Blue" phenomenon: regional emission control effects observed from space. Atmos. Res. 164, 65–75.
Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. Lightgbm: a highly efficient gradient boosting decision tree. Adv. Neural Inf. Process. Syst. 30.
Kuehn, B.M., 2014. WHO: more than 7 million air pollution deaths each Year. JAMA 311, 1486, 1486.
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, 702–706.
Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525, 367–371.
Li, H., Zhang, Q., Duan, F., Zheng, B., He, K., 2016. The "Parade Blue": effects of short-term emission control on aerosol chemistry. Faraday Discuss 189, 317–335.
Li, L., Li, Q., Huang, L., Wang, Q., Zhu, A., Xu, J., Liu, Z., Li, H., Shi, L., Li, R., 2020. Air quality changes during the COVID-19 lockdown over the Yangtze River Delta Region: an insight into the impact of human activity pattern changes on air pollution variation. Sci. Total Environ. 732, 139282.
Li, R., Gao, Y., Xu, J., Cui, L., Wang, G., 2023. Impact of clean air policy on criteria air pollutants and health risks across China during 2013–2021. J. Geophys. Res. Atmos. 128, e2023JD038939.
Li, Z., Guo, J., Ding, A., Liao, H., Liu, J., Sun, Y., Wang, T., Xue, H., Zhang, H., Zhu, B., 2017. Aerosol and boundary-layer interactions and impact on air quality. Natl. Sci. Rev. 4, 810–833.
Liu, S., Wang, G., Kong, F., Zhao, N., Gao, W., Zhang, H., 2025. PM2. 5 pollution characteristics, drivers, and regional transport during different pollution levels in Linyi, China: an integrated PMF-ML-SHAP framework and transport models. J. Hazard. Mater., 138534
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, 56–67.
Lv, Z., Wang, X., Deng, F., Ying, Q., Archibald, A.T., Jones, R.L., Ding, Y., Cheng, Y., Fu, M., Liu, Y., 2020. Source–receptor relationship revealed by the halted traffic and aggravated haze in Beijing during the COVID-19 lockdown. Environ. Sci. Technol. 54, 15660–15670.
Marlier, M.E., Xing, J., Zhu, Y., Wang, S., 2020. Impacts of COVID-19 response actions on air quality in China. Environ. Res. Commun. 2, 075003.
MEEP, 2019. China Environmental Status Bulletin 2019. Ministry of Ecology and Environmental of the People's Republic of China(MEEP).
Men, Y., Li, Y., Luo, Z., Jiang, K., Yi, F., Liu, X., Xing, R., Cheng, H., Shen, G., Tao, S., 2023. Interpreting highly variable Indoor PM2.5 in rural North China using machine learning. Environ. Sci. Technol. 57, 18183–18192.
Miyazaki, K., Bowman, K., Sekiya, T., Jiang, Z., Chen, X., Eskes, H., Ru, M., Zhang, Y., Shindell, D., 2020. Air quality response in China linked to the 2019 novel coronavirus (COVID-19) lockdown. Geophys. Res. Lett. 47, e2020GL089252.
Nel, A., 2005. Air pollution-related Illness: effects of particles. Science 308, 804–806.
Nie, D., Shen, F., Wang, J., Ma, X., Li, Z., Ge, P., Ou, Y., Jiang, Y., Chen, M., Chen, M., Wang, T., Ge, X., 2021. Changes of air quality and its associated health and economic burden in 31 provincial capital cities in China during COVID-19 pandemic. Atmos. Res. 249, 105328.
Pei, Z., Han, G., Ma, X., Su, H., Gong, W., 2020. Response of major air pollutants to COVID-19 lockdowns in China. Sci. Total Environ. 743, 140879.
Petetin, H., Bowdalo, D., Soret, A., Guevara, M., Jorba, O., Serradell, K., Pérez García-Pando, C., 2020. Meteorology-normalized impact of the COVID-19 lockdown upon NO 2 pollution in Spain. Atmos. Chem. Phys. 20, 11119–11141.
Petit, J.-E., Favez, O., Sciare, J., Crenn, V., Sarda-Esteve, R., Bonnaire, N., Močnik, G., Dupont, J.-C., Haeffelin, M., Leoz-Garziandia, E., 2015. Two years of near real-time chemical composition of submicron aerosols in the region of Paris using an Aerosol Chemical Speciation Monitor (ACSM) and a multi-wavelength Aethalometer. Atmos. Chem. Phys. 15, 2985–3005.
Rybarczyk, Y., Zalakeviciute, R., 2021. Assessing the COVID-19 impact on air quality: a machine learning approach. Geophys. Res. Lett. 48, e2020GL091202.
Shen, F., Hegglin, M.I., Luo, Y., Yuan, Y., Wang, B., Flemming, J., Wang, J., Zhang, Y., Chen, M., Yang, Q., Ge, X., 2022. Disentangling drivers of air pollutant and health risk changes during the COVID-19 lockdown in China. npj Climate and Atmospheric Science 5, 54.
Shen, F., Hegglin, M.I., Yuan, Y., 2024. Impact of weather patterns and meteorological factors on PM2.5 and O3 responses to the COVID-19 lockdown in China. Atmos. Chem. Phys. 24, 6539–6553.
Shen, F., Zhang, L., Jiang, L., Tang, M., Gai, X., Chen, M., Ge, X., 2020. Temporal variations of six ambient criteria air pollutants from 2015 to 2018, their spatial distributions, health risks and relationships with socioeconomic factors during 2018 in China. Environ. Int. 137, 105556.
Shi, Z., Song, C., Liu, B., Lu, G., Xu, J., Van Vu, T., Elliott, R.J.R., Li, W., Bloss, W.J., Harrison, R.M., 2021. Abrupt but smaller than expected changes in surface air quality attributable to COVID-19 lockdowns. Sci. Adv. 7, eabd6696.
Stirnberg, R., Cermak, J., Kotthaus, S., Haeffelin, M., Andersen, H., Fuchs, J., Kim, M., Petit, J.-E., Favez, O., 2021. Meteorology-driven variability of air pollution (PM 1) revealed with explainable machine learning. Atmos. Chem. Phys. 21, 3919–3948.
Sun, Y., Lei, L., Zhou, W., Chen, C., He, Y., Sun, J., Li, Z., Xu, W., Wang, Q., Ji, D., Fu, P., Wang, Z., Worsnop, D.R., 2020. A chemical cocktail during the COVID-19 outbreak in Beijing, China: insights from six-year aerosol particle composition measurements during the Chinese New Year holiday. Sci. Total Environ. 742, 140739.
Sun, Y., Wang, Z., Fu, P., Jiang, Q., Yang, T., Li, J., Ge, X., 2013. The impact of relative humidity on aerosol composition and evolution processes during wintertime in Beijing, China. Atmos. Environ. 77, 927–934.
Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M.U., Li, B., Cai, J., Xu, B., Yang, Q., 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science 368, 638–642.
Wang, J., Lei, Y., Chen, Y., Wu, Y., Ge, X., Shen, F., Zhang, J., Ye, J., Nie, D., Zhao, X., Chen, M., 2021a. Comparison of air pollutants and their health effects in two developed regions in China during the COVID-19 pandemic. J. Environ. Manage. 287, 112296.
Wang, J., Li, J., Ye, J., Zhao, J., Wu, Y., Hu, J., Liu, D., Nie, D., Shen, F., Huang, X., 2020a. Fast sulfate formation from oxidation of SO2 by NO2 and HONO observed in Beijing haze. Nat. Commun. 11, 2844.
Wang, J., Ye, J., Zhang, Q., Zhao, J., Wu, Y., Li, J., Liu, D., Li, W., Zhang, Y., Wu, C., 2021b. Aqueous production of secondary organic aerosol from fossil-fuel emissions in winter Beijing haze. Proc. Natl. Acad. Sci. 118, e2022179118.
Wang, S., Zhao, M., Xing, J., Wu, Y., Zhou, Y., Lei, Y., He, K., Fu, L., Hao, J., 2010. Quantifying the air pollutants emission reduction during the 2008 Olympic Games in Beijing. Environ. Sci. Technol. 44, 2490–2496.
Wang, Y., Wen, Y., Wang, Y., Zhang, S., Zhang, K.M., Zheng, H., Xing, J., Wu, Y., Hao, J., 2020b. Four-month changes in air quality during and after the COVID-19 lockdown in six megacities in China. Environ. Sci. Technol. Lett. 7, 802–808.
Weng, Z., Wang, Y., Yang, X., Cheng, C., Tan, X., Shi, L., 2022. Effect of cleaner residential heating policy on air pollution: a case study in Shandong Province, China. J. Environ. Manage. 311, 114847.
Xiao, Q., Chang, H.H., Geng, G., Liu, Y., 2018. An ensemble machine-learning model to predict Historical PM2.5 concentrations in China from satellite data. Environ. Sci. Technol. 52, 13260–13269.
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, 14347–14359.
Yan, H.-h., Chen, L.-f., Tao, J.-h., Han, D., Su, L., Yu, C., 2012. SO2 long-term monitoring by satellite in the Pearl River Delta. Journal of Remote Sensing 16, 390–404.
Yang, J., Wen, Y., Wang, Y., Zhang, S., Pinto, J.P., Pennington, E.A., Wang, Z., Wu, Y., Sander, S.P., Jiang, J.H., Hao, J., Yung, Y.L., Seinfeld, J.H., 2021. From COVID-19 to future electrification: assessing traffic impacts on air quality by a machine-learning model. Proc. Natl. Acad. Sci. 118, e2102705118.
Zhang, K., Liu, Z., Zhang, X., Li, Q., Jensen, A., Tan, W., Huang, L., Wang, Y., de Gouw, J., Li, L., 2022. Insights into the significant increase in ozone during COVID-19 in a typical urban city of China. Atmos. Chem. Phys. 22, 4853–4866.
Zhou, M., Jiang, W., Gao, W., Gao, X., Ma, M., Ma, X., 2021. Anthropogenic emission inventory of multiple air pollutants and their spatiotemporal variations in 2017 for the Shandong Province, China. Environ. Pollut. 288, 117666.