Environmental Science and Ecotechnology 24 (2025) 100524

Contents lists available at ScienceDirect

Environmental Science and Ecotechnology

journal homepage: www.journals.elsevier.com/environmental-science-and-ecotechnology/

Original Research

Causal-inference machine learning reveals the drivers of China's 2022 ozone rebound

Lin Wang [a, 1], Baihua Chen [a, 1], Jingyi Ouyang [a, b], Yanshu Mu [c], Ling Zhen [a, b], Lin Yang [a], Wei Xu [a, **], Lina Tang [a, *]

[a] Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, 361021, China
[b] University of Chinese Academy of Sciences, Beijing, 100049, China
[c] China School of Mathematics, Jilin University, Changchun, 130012, China

ARTICLE INFO

Article history:
Received 16 May 2024
Received in revised form
4 January 2025
Accepted 5 January 2025

Keywords:
Ground-level ozone
Interpretable machine learning
Causal forest
Ozone rebound

ABSTRACT

Ground-level ozone concentrations rebounded significantly across China in 2022, challenging air quality management and public health. Identifying the drivers of this rebound is crucial for designing effective mitigation strategies. Commonly used methods, such as chemical transport models and machine learning, provide valuable insights but face limitations—chemical transport models are computationally intensive, while machine learning often fails to address confounding factors or establish causality. Here we show that elevated temperatures and increased solar radiation, as primary meteorological drivers, collectively account for 57 % of the total ozone increase, based on an integrated analysis of ground-based monitoring data, satellite observations, and meteorological reanalysis information using explainable machine learning and causal inference techniques. Compared to the year 2021, 90 % of the stations reported an increase in the Formaldehyde to Nitrogen ratio, implying a growing sensitivity of ozone formation to nitrogen oxide levels. These findings highlight the significant causal role of meteorological changes in the ozone rebound, urging the adoption of targeted ozone mitigation strategies under climate warming, particularly through varied regional strategies that consider existing anthropogenic emission levels and the prospective increase in biogenic volatile organic compounds. This identification of causal relationships in air pollution dynamics can support data-driven and accurate decision-making.

1. Introduction

Ground-level ozone is among the most concerning air pollutants because it poses significant hazards to public health [1] and ecosystems [2]. Short-term ozone exposure can contribute to adverse respiratory effects [3]. By contrast, long-term exposure to ozone has a broader and more severe impact and is associated with increased cardiovascular and respiratory mortality and a decline in lung function [4,5]. Comprehensive ecological investigations have found that elevated ozone levels can impede photosynthetic activity, resulting in significant reductions in crop productivity and disruptions in ecological processes [6]. In addition, ozone may exacerbate particulate matter (PM) pollution by oxidising volatile compounds, ultimately leading to secondary pollutant formation [7].

Having implemented stringent regulatory measures, such as enhanced emissions standards, vehicle exhaust controls, and industrial upgrades, China has documented substantial improvements in air pollution control in recent years [8]. Consequently, the annual average $PM_{2.5}$ concentration in China fell by 57 % from 2013 to 2022 [9]. However, the ozone levels escalated from 2013 to 2019, particularly during the warm seasons, with significant pollution surges in the Beijing–Tianjin–Hebei (BTH), Yangtze River Delta (YRD), and Fenwei Plain (FWP) regions [10]. This rising trend in ozone concentrations temporarily moderated from 2019 to 2021, but a widespread rebound was observed in 2022 when over 70 % of Chinese cities recorded increases in annual ozone levels [9]. Yang et al. [11] identified the unprecedented heat as the primary driver of the 2022 summertime ozone rebound. Qiao et al. [12] further

* Corresponding author.
** Corresponding author.
E-mail addresses: wxu@iue.ac.cn (W. Xu), lntang@iue.ac.cn (L. Tang).
[1] Lin Wang and Baihua Chen contributed equally.

https://doi.org/10.1016/j.ese.2025.100524

elucidated that the heat primarily exacerbated the ozone levels through increased net chemical production. However, ground-based measurements indicated that this ozone rebound was not confined to the summer but persisted into the cold seasons. A comprehensive investigation of the causal factors of the annual ozone surge across China remains incomplete.

Ground-level ozone is formed through photochemical reactions involving nitrogen oxides (NO_x) and volatile organic compounds (VOCs) in the presence of sunlight [13]. This process is influenced by a complex interplay of meteorological conditions, including air temperature, humidity, wind speed, boundary layer height, surface pressure, precipitation, and surface downward radiation [14]. These factors critically affect the physical-chemical properties of ozone and determine its spatial-temporal dynamics [15]. High temperatures and strong sunlight significantly accelerate the formation of ozone, whereas increased wind speeds and precipitation mitigate ozone levels through dispersion and deposition. Atmospheric stagnation, characterised by low boundary layer heights and reduced wind speeds, often results in an increase in ozone due to the accumulation of precursors. Pollutants other than NO_x and VOCs, such as sulphur dioxide (SO_2), carbon monoxide (CO), and $PM_{2.5}$, also play vital roles in ozone chemistry [16–18]. SO_2 and CO serve as reactants in photochemical reactions, modifying the ozone formation rate, while $PM_{2.5}$ influences ozone through the absorption or scattering of sunlight and provides a medium for heterogeneous reactions that may either generate or degrade ozone [19]. This intricate, nonlinear system of chemistry and meteorology interactions complicates the identification of the primary causes of ozone rebound.

Numerical Chemical Transport Models (CTMs) are essential in advancing our understanding of ozone dynamics. However, CTMs often use coarse spatial resolutions and require intensive computing resources, and their reliance on up-to-date emission inventories and established chemical mechanisms introduces considerable uncertainties in ozone simulations [20]. Traditional statistical models, typically applied to analyse observational ozone data, often fail to capture the non-linear and high-dimensional relationships within the atmosphere due to their inherent methodological simplicity [21]. Recently developed statistical models, such as machine learning (ML) techniques, are specifically designed to handle these complexities and have proven effective in uncovering the nonlinear and high-dimensional relationships crucial for ozone formation [22]. ML models, including random forest (RF), light gradient boosting (LightGBM), convolution neural networks (CNN), deep neural networks (DNN), and long short-term memory models (LSTM), have been utilised to explore the complex sensitivities involved in ozone generation and to quantify the importance of multiple atmospheric drivers [18,23]. For instance, Mai et al. [24] deployed CNNs and Shapley additive explanations (SHAP) to investigate the influence of three-dimensional spatial meteorological variations on Shenzhen's ozone dynamics, attributing the 2022 ozone rebound primarily to raised temperatures and dominant northerly winds rather than the changes in emissions. Tan et al. [25] used RF models and the SHAP approach to isolate the effects of pollutants and meteorological factors on the ozone uptick in lockdown cities, finding that increases in temperature and VOC emissions were the main drivers of this uptick, particularly during heatwaves and droughts.

However, predictive ML models predominantly capture associations rather than causality, which could compromise the accuracy and applicability of the decision-making process [26,27]. Improving predictive accuracy and causal understanding in ozone formation studies requires the integration of causal inference insights to specifically address the complex interplay of meteorological and environmental factors influencing ozone levels [28]. Causal ML effectively merges causal interpretability algorithms with ML's proficiency in detecting statistical patterns within high-dimensional, unstructured data [29,30]. This synthesis can clarify the cause-and-effect relationship between variables while maintaining robustness against confounders, untangling the causal mechanisms that drive observable phenomena over mere associations. Causal forest (CF) is a tree-based causal ML method that exhibited high performance in analysing the causal effect of various interventions on outcomes [31–33].

In this study, we employed ML-aided causal inference techniques to investigate the drivers and discuss the causality of the widespread ozone rebound in 2022. RF models were applied to quantify the effects of both anthropogenic emissions and meteorological factors on ozone variation. Subsequent SHAP and partial dependence analyses illuminated how specific predictors variably influenced spatiotemporal ozone levels across China. These analyses were supported by CF techniques to pinpoint critical causes affecting the ozone escalation in 2022. This study attempts to provide causal evidence of the environmental factors driving the 2022 ozone rebound, thereby enhancing decision-making for air pollution control.

2. Materials and methods

2.1. Data sources and processing

This study integrated multi-source heterogeneous data, including ground-based measurements, satellite-derived ozone precursors, and meteorological reanalysis datasets.

Hourly ground-level concentrations of ozone, $PM_{2.5}$, NO_2, CO, and SO_2 collected from 1558 air quality monitors between 2019 and 2022 were obtained from the China National Environmental Monitoring Center (CNEMC; http://www.cnemc.cn/). We utilised the maximum daily 8-h average (MDA8) ozone concentration (containing at least six valid hourly values) as the dependent variable in the development of our model, adhering to the World Health Organisation's Air Quality Guidelines (2005), a standard commonly adopted in air quality regulations and health-related studies. Overall, there are 953,343 valid observations in our model.

The Sentinel-5 Precursor (S5P) satellite, deployed by the European Space Agency to monitor air pollution, orbits in a sun-synchronous path near the poles and crosses the equator daily at 13:30 local time [34]. The TROPOspheric Monitoring Instrument (TROPOMI) equipped with the S5P satellite offers detailed global mappings of atmospheric trace gases at a high spatial resolution. Utilising Google Earth Engine, we accessed TROPOMI's daily offline Level 3 (OFFL L3) measurements of formaldehyde (THCND) and NO_2 (THNO2) from 2019 to 2022, all at a spatial resolution of 1113.2 m (source: Google Earth Engine Dataset Catalog). The harpconvert tool using the bin_spatial operation is used to convert the original TROPOMI data to OFFL L3 in Google Earth Engine. The harpconvert tool, part of the HARP (Harmonized Atmospheric Research Product) package, aggregates the original TROPOMI measurements onto a regular grid, handling missing values and producing seamless Level 3 data. The fifth-generation reanalysis data, ERA5, from the European Centre for Medium-Range Weather Forecast (ECMWF), embodies a comprehensive record of the global atmosphere (C3S, 2017). The dataset of ERA5-Land hourly data from 1950 to the present (0.1° resolution) and ERA5 hourly data on single levels from 1940 to the present (0.25° resolution) were integrated for the MDA8 ozone concentration prediction (https://cds.climate.copernicus.eu). We incorporated essential meteorological variables including surface solar radiation downwards (SSRD), 2 m temperature (T_{2m}), relative humidity (RH), boundary layer height (BLH), surface pressure (SP), total precipitation (TP), 10 m u-

component of wind (U_{10m}), and 10 m v-component of wind (V_{10m}). Wind speed (WS_{10}) was calculated by U_{10m} and V_{10m}. The averages of ERA5 meteorology data from 9:00 a.m. to 6:00 p.m. were processed to match with the MDA8 ozone concentration.

To ensure consistency across our dataset, we pre-processed and fused data from multiple sources. We employed the L1 norm methodology to identify the nearest grid point to each station. The L1 norm methodology, also known as the Manhattan distance or least absolute deviations, involves calculating the sum of the absolute differences between coordinates. This method is effective for determining proximity in a grid-based system. By using this approach, we accurately assigned data from ERA5 and TROPOMI at these grid points to the respective stations.

2.2. Methods

In this study, we adopted multiple methods to understand the rebound of ground-level ozone in China during 2022, utilising RF modelling, SHAP additive explanations, and CF modelling. The RF models were developed for each air pollution monitoring station, integrating comprehensive datasets from ground-based measurements, satellite-derived precursors, and meteorological reanalysis data. Feature importance scores, SHAP, and partial dependence analyses were systematically applied to delineate the statistical association of each predictor on the ozone rebound across regions. Ultimately, we performed the CF models at each station to assess the causal effects of predictors on ozone levels, comparing these effects with the RF's partial dependence results, thus validating the actual causal factors behind the 2022 ozone rebound (Fig. 1).

2.2.1. RF modelling

Utilising RF algorithms, we constructed models to analyse the relationships between MDA8 ozone data and an array of meteorological and air pollution predictors at all monitoring sites across China. RF, an ensemble method detailed by Liaw and Wiener [35], leverages bootstrap sampling and random variable selection to reduce the 'out-of-bag' error by optimising data splits and variable selection. It is not a complete 'black-box' approach, as the importance of the predictors during the RF's construction can be used to infer some underlying mechanisms influencing ozone. The detailed predictors and formula are shown in Supplementary Material Text S1.1. We developed RF models for each monitoring station, with the number of valid entries ranging from 301 to 1166 per station. We selected the data from 2019 to 2021 for training and reserved the data in 2022 for testing. After all 1588 models underwent an initial training phase, their hyperparameters were adjusted based on performance evaluation to ensure optimal predictive efficacy.

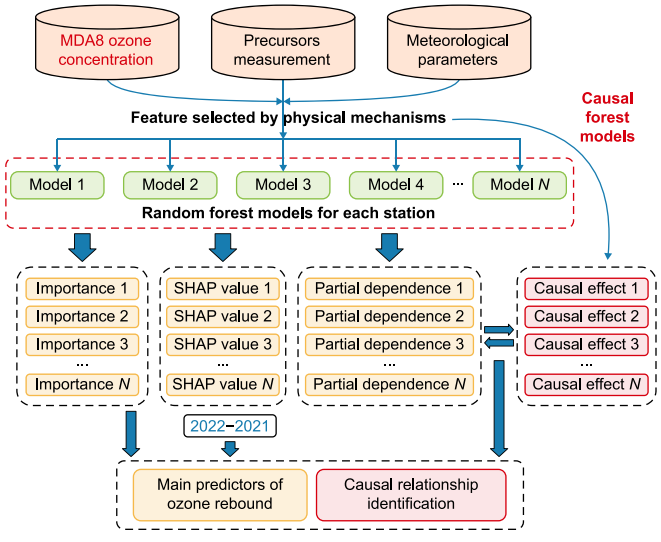

Fig. 1. The schematic diagram for identifying the cause factors of ozone rebound in 2022.

2.2.2. SHAP analysis

SHAP values were applied to elucidate the contributions of individual features to the outcomes of the RF models we deployed at each monitoring site, as detailed by Lundberg et al. [36,37]. The local SHAP value reveals both the magnitude and direction of how each feature in an individual sample influences the target outcome (Supplementary Material Text S1.2). We further assessed the SHAP value difference (SVD) between the year 2022 and the average of the years 2020–2021 to quantify the contribution of different parameters to the ozone rebound (Supplementary Material Text S1.2).

2.2.3. CF modelling

The CF is derived from generalised RFs [32]. Unlike traditional RFs, which are predominantly used for classification and regression tasks, CFs aim to estimate the treatment effect rather than merely predict outcomes. This operation involves estimating the average treatment effect (ATE) as well as the conditional average treatment effect (CATE) across different subgroups identified within the dataset. The CF achieves unbiased and consistent estimation of treatment effects using statistical methods such as propensity score estimation, data partitioning focused on treatment effect heterogeneity, honesty in tree construction, variance reduction through ensembling, local centring, and constructing confidence intervals based on asymptotic normality. This approach is adept at managing the complexities of high-dimensional data and mitigating the influence of irrelevant covariates. The ability of CFs to delve beyond mere associations into causal mechanisms offers invaluable insights for decision-making in areas such as public policy and healthcare, thereby ensuring that interventions are based on robust, statistically inferred causal relationships. Moreover, CFs can offer asymptotically Gaussian and centred sampling distributions to identify the true treatment effects, which assist in constructing asymptotic confidence intervals around estimates, facilitating robust statistical inference in complex scenarios.

For each air quality monitoring station, we applied default construction parameters to develop a CF model composed of over 200 causal trees. This model was instrumental in validating the causal impacts on ozone levels previously identified by using the RF model to conduct a comparative analysis of the ATEs. The CF was conducted using the R package 'grf'. The detailed formulas are shown in Supplementary Material Text S1.3.

3. Results and discussion

3.1. Spatial and temporal rebound pattern of MDA8 ozone and potential drivers

In 2022, a significant rebound in ozone levels was recorded at over 70 % of China's air pollution monitoring sites in contrast to the decline observed from 2019 to 2021. The annual national mean MDA8 ozone concentration for all monitoring sites dropped from 111.14 ± 0.48 to 101.79 ± 0.32 μg m^{-3} from 2019 to 2021 (Fig. 2a). Ni et al. [10] reported that the decline was more pronounced in eastern China (26°–42° N, 110°–122.5° E), where there was a decrease of 14.9 μg m^{-3}, while the rest of China decreased by 4.8 μg m^{-3}, indicating distinct regional variations in ozone reductions. However, the annual mean MDA8 ozone concentration escalated to 111.62 ± 0.45 μg m^{-3} in 2022.

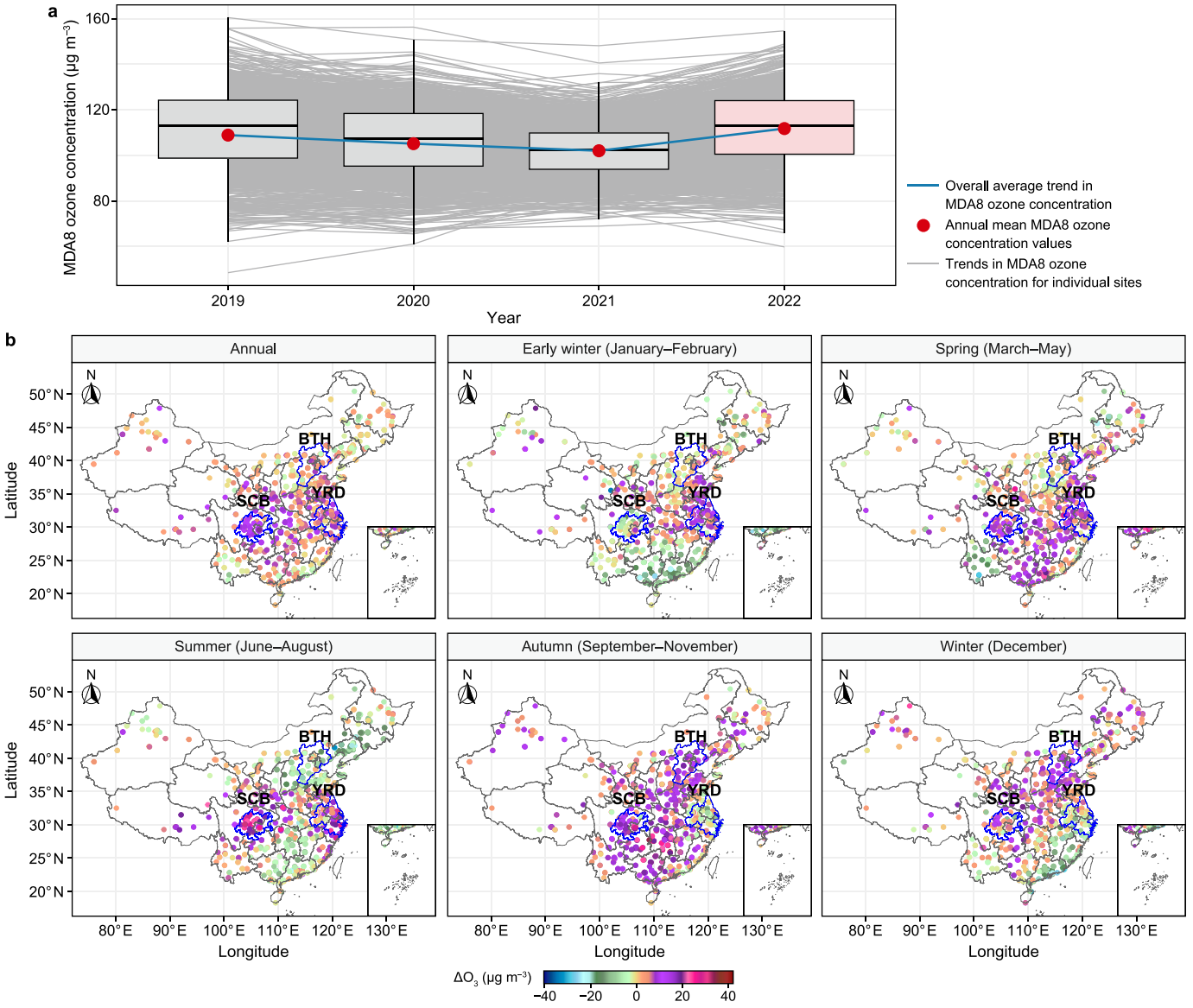

Fig. 2. The maximum daily 8-h average (MDA8) ozone concentration variation between 2022 and 2019–2021. **a**, Annual MDA8 ozone concentration from 2019 to 2022. **b**, Annual and seasonal spatial MDA8 ozone concentration variation between 2022 and 2021. Key regions include Beijing–Tianjin–Hebei (BTH), Sichuan Basin (SCB), and Yangtze River Delta (YRD), marked with blue boundaries.

Between 2021 and 2022, China experienced a decline in most of the pollutants that could lead to ozone reductions (Supplementary Materials Figs. S1–8). However, the VOC levels increased during this period, which also experienced meteorological conditions conducive to ozone formation (Supplementary Materials Figs. S1–8). $PM_{2.5}$, NO_2, SO_2, and CO emissions decreased by 9.22 %, 13.73 %, 7.13 %, and 8.01 %, respectively (Supplementary Material Fig. S2). The levels of HCHO, indicative of total VOC reactivity, rose by 6.32 %, particularly in the central, southern, and eastern regions. Meteorological data showed an average annual increase of 8.61 % in T_{2m} and 5.12 % in SSRD, with the most significant T_{2m} rise observed in central and eastern China. Notably, the proportion of the average of extremely hot days for all stations (defined as days with an average T_{2m} above the 95th percentile of summer T_{2m} from 2019 to 2022) increased by 5.43 % in the summer of 2022, and the maximum T_{2m} in 2022 was 1.64 °C higher than that in 2021. As regards the other meteorological parameters, TP decreased by 4.93 %, BLH increased by 6.09 %, RH decreased by 3.01 %, WS_{10m} increased by 1.80 %, and SP had a minimal decrease of 0.10 %.

Varied spatial patterns of ozone rebound can be observed across mainland China (Fig. 2b). The eastern and central regions, particularly the industrialised and densely populated YRD, BTH, and Sichuan Basin (SCB) regions and their surroundings, saw significant increases in ozone concentrations. As suggested by the observed increase in HCHO levels and changes in meteorological factors in Supplementary Materials Figs. S1–2, these surges were likely linked to extensive vegetation, substantial industrial activities, heavy traffic, and varying meteorological conditions [11]. For example, the elevated temperatures in these areas may have exacerbated this ozone pollution by intensifying VOC emissions from natural and certain temperature-sensitive anthropogenic sources [38]. Elevated temperatures can increase the emissions of biogenic VOCs, such as isoprene and monoterpenes, by altering metabolic activity and causing membrane damage in plants [39].

High temperatures significantly increase the evaporation and fugitive emissions of anthropogenic VOCs from sources such as petrochemical processing, vehicles, and solvent use, as power generation is also rising due to increased cooling and heating demands [40]. In contrast, the relatively minor changes observed in the western, southeastern, and some northern parts of China, as indicated by the modest shifts in HCHO levels and meteorological factors shown in Supplementary Materials Figs. S1–2, could suggest fewer fluctuations in VOC emissions and meteorological conditions that are less detrimental to ozone formation.

Distinct seasonal patterns in MDA8 ozone concentration were also found, with substantial increases during the spring, summer, and autumn. Spring witnessed significant ozone level rises, particularly in central, southern, and eastern China. During summer, intensified ozone levels were prominently observed in the SCB and YRD regions. Autumn saw an expansive increase in ozone levels across the mainland, excluding the YRD. In early winter and winter, while ozone levels rebounded in northern and eastern China, the increase was less pronounced than in other seasons (Fig. 2b).

3.2. Analysis of spatial distribution of variable importance

To investigate the cause of the widespread ozone rebound during 2022, we developed RF models for each monitoring station. To ensure generalisability and avoid over-fitting, the data from 2019 to 2021 were used as the training dataset, and those from 2022 were used as the test dataset. The test dataset consistently showed good performance with an average Pearson's *R* of 0.85 despite some spatial differences in model performance (Fig. 3a), indicating that the RF models, like other ozone models [41,42], could reproduce most of the ozone variation in 2022.

Permutation importance analyses were conducted for each station. The probability density of feature importance across the RF models is shown in Fig. 3c. Among the 12 predictors, T_{2m} and SSRD were the most important variables in the prediction models, aligning with the understanding that higher temperatures and solar radiation could increase ozone formation through enhanced photochemical activity [43–45]. $PM_{2.5}$ and NO_2 concentrations were highlighted as critical anthropogenic variables in ozone prediction models. Additional variables, including SP, BLH, RH, WS_{10m}, CO, and THCND, also contributed to predicting ozone levels, but their effects were relatively minor.

We further analysed the spatial distribution of the key variables in ozone prediction. The temperature was identified as a dominant factor in the BTH, YRD, and SCB regions of China (Fig. 3b). The widespread impact of solar radiation in the southwest and northeast, as well as the southern part of the YRD region, highlights its key role in the photochemical formation of ozone [46]. Moreover, $PM_{2.5}$ was revealed to be particularly important in the southern regions, underscoring the complex relationship between $PM_{2.5}$ and ozone. This relationship may be influenced by meteorological conditions, common precursors, atmospheric oxidation capacity, and $PM_{2.5}$–ozone interaction [16,47,48]. Specifically, studies have shown that heterogeneous uptake of NO_2 enhances ozone formation by converting NO_2 to nitrous acid (HONO), which can be a significant source of HO_x radicals ($OH + HO_2$) [49,50]. In Shenzhen, Tan et al. [25] found that HO_2 and NO_2 uptake altered MDA8 ozone by −4 % and 42 %, respectively, while $PM_{2.5}$ reduced ozone by 45 % through light extinction in October. Lyu et al. [16] combined multi-source data with an RF model to predict $PM_{2.5}$–ozone co-occurrence and identify common precursors in southern China, such as VOCs, as key contributing factors.

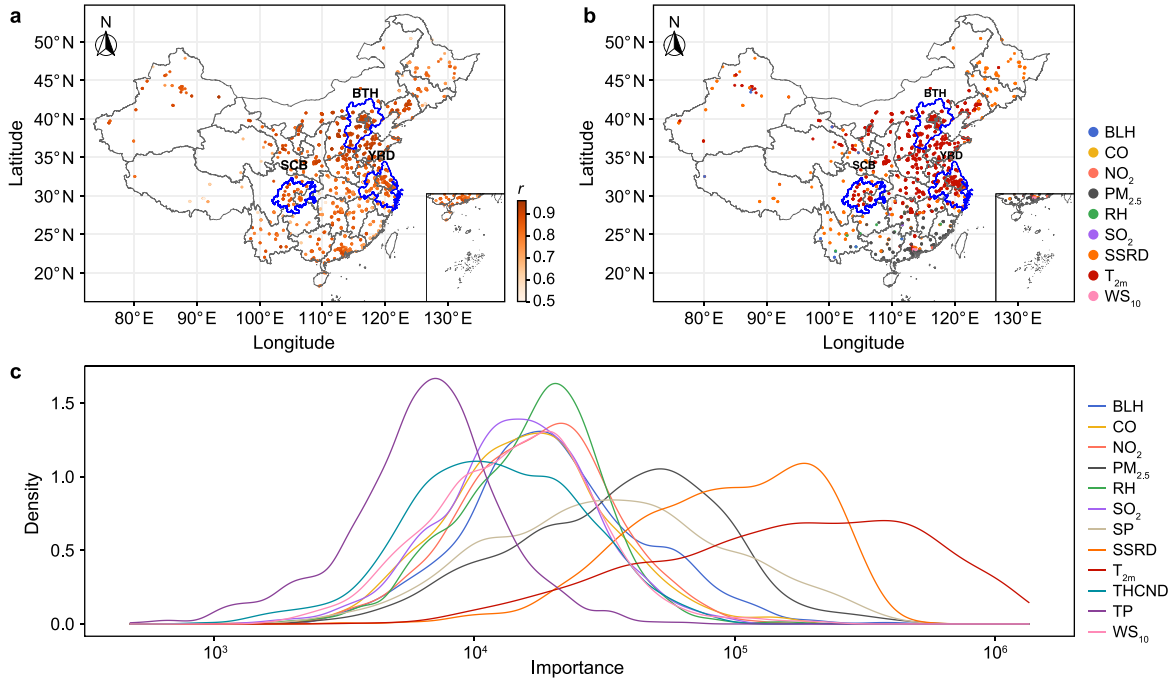

Fig. 3. Model performance and permutation importance analyses. **a**, Performance of the random forest prediction model measured by Pearson's correlation coefficient (*r*) at each monitoring site across China. The color scale represents the correlation values, with darker shades indicating higher correlations. **b**, The dominant variable in each prediction model. **c**, The probability density distribution of each variable's importance. The *x*-axis represents permutation importance, reflecting changes in model performance when feature values are shuffled. Larger values indicate higher predictor importance. Features include boundary layer height (BLH), CO, NO_2, $PM_{2.5}$, relative humidity (RH), SO_2, surface pressure (SP), surface solar radiation downwards (SSRD), 2 m temperature (T_{2m}), formaldehyde (THCHD), total precipitation (TP), and 10 m wind speed (WS_{10}). Key regions include Beijing–Tianjin–Hebei (BTH), Sichuan Basin (SCB), and Yangtze River Delta (YRD), marked with blue boundaries.

3.3. Quantifying the contribution of each variable to ozone rebound: the SHAP approach

A SHAP analysis was then performed to quantify the contribution of each variable to the rebound of MDA8 ozone in 2022. We computed the SVD value by subtracting the SHAP values of each variable in 2021 from their respective values in 2022 (Figs. 4 and 5). The SVD calculation differs from the permutation importance analysis by quantifying the contribution of each variable to the ozone rebound. For instance, if a variable was important in driving the daily or seasonal levels of ozone, its permutation importance ranking would be high, whereas if it showed little variation in 2022 compared to the previous three years, its SVD values would be minor.

The SVD values were calculated for each variable at all sites (Figs. 4 and 5). The T_{2m} was the most significant factor associated with ozone rebound with a mean SVD value of 3.96 ± 0.09 μg m^{-3}, corresponding to an annual average temperature increase of 1.55 ± 0.05 °C. This finding can explain up to 38.90 % of the MDA8 ozone rebound. Notably, the SVD temperature values were most pronounced in the SCB and BTH regions and eastern China, with the highest recorded being 17.46 μg m^{-3}. A recent investigation into ozone levels during the summer of 2022 reveals that the BTH, YRD, and SCB regions experienced the highest ozone levels since 2015, largely attributed to a strong increase in net chemical production amid extreme heat [12]. Furthermore, Yang et al. [11] reported that the strong ozone–temperature correlation from 2019 to 2021 weakened at very high temperatures, probably due to the plant stomata closure that reduced biogenic VOCs (BVOCs) and caused significantly lower net solar radiation during extreme heat. High temperatures and increased humidity can also inhibit ozone photochemistry, especially in southern China. However, even under the elevated temperatures of 2022, this correlation persisted in the BTH and YRD regions, probably driven by a heat wave-related intensification of local net chemical production. However, the temperature rise not only promoted ozone formation in the summer (SVD value: 0.52 ± 0.08 μg m^{-3}) but also resulted in the most significant and widespread increase in ozone levels during the autumn, with a national average rise of 5.72 ± 0.13 μg m^{-3}, and a smaller increase of 1.54 ± 0.09 μg m^{-3} in spring (Supplementary Materials Figs. S9–14). This increase across multiple seasons highlights the robust relationship between temperature and ozone production, suggesting that temperature-driven chemical reactions

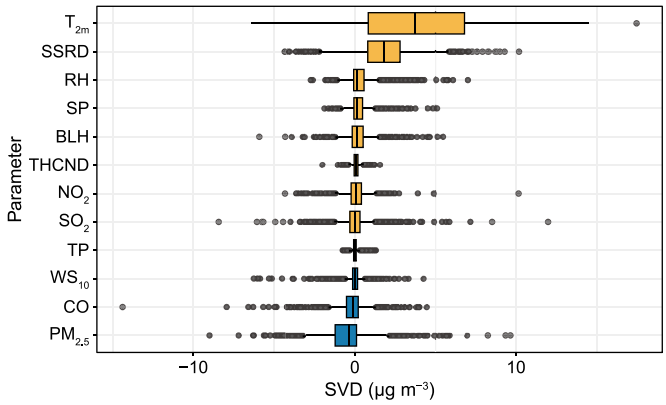

Fig. 4. The Shapley additive explanations (SHAP) value difference (SVD) value distribution for different parameters between the year 2022 and 2021. Features include boundary layer height (BLH), CO, NO_2, $PM_{2.5}$, relative humidity (RH), SO_2, surface pressure (SP), surface solar radiation downwards (SSRD), 2 m temperature (T_{2m}), formaldehyde (THCHD), total precipitation (TP), and 10 m wind speed (WS_{10}).

are crucial for ozone variability during summer and across transitional seasons.

A national increase in SSRD by $4.77 \times 10^5 \pm 9.57 \times 10^3$ J m^{-2} was associated with an average SVD value of 1.88 ± 0.05 μg m^{-3}, accounting for about 18.47 % of the MDA8 ozone rebound across China. However, this relationship has been little explored in the existing research on ozone rebound. Although the overall contribution of solar radiation to ozone rebound was less pronounced than that of temperature, it exhibited a broad and consistent influence across various geographical regions and throughout seasons, including the less sun-intensive periods of early winter and winter. This pattern underlines the crucial role of solar radiation in the photochemical dynamics influencing ozone rebound that is evident across different climatic conditions. In contrast, the influence of temperature, though more significant, tended to be more geographically and seasonally confined. Regarding the summer of 2022, the notable ozone rebound observed in the YRD and SCB regions can be largely attributed to the synergistic effects of temperature and solar radiation, predominantly governed by the westward expansion of the Western Pacific Subtropical High [12].

Additionally, other meteorological factors, such as RH, SP, and BLH, demonstrate positive SVD values of 0.37 ± 0.02, 0.27 ± 0.02, and 0.23 ± 0.02 μg m^{-3}, respectively, collectively accounting for approximately 8.55 % of the MDA8 ozone rebound. In 2022, the decreased RH in south-central China likely inhibited the hydroxyl radical formation necessary for ozone removal, while increased BLH and SP in this region may have facilitated the accumulation and chemical reaction of ozone precursors. These conditions, coupled with elevated temperatures and increased solar radiation, synergistically intensified the ozone rebound. The statistical associations derived from the RF models may suggest that these climate anomalies were the primary drivers of the increased observed ozone levels.

Regarding the emissions, the decrease in $PM_{2.5}$ concentrations by 3.12 ± 0.10 μg m^{-3} was associated with an SVD of −0.53 ± 0.04 μg m^{-3}, showing that the effects of $PM_{2.5}$ concentrations contributed less to ozone levels than those of temperature in 2022. Notably, the association between $PM_{2.5}$ and ozone varied significantly across regions. In northeast China and the SCB and YRD regions, reductions in $PM_{2.5}$ were consistently associated with comparable decreases in ozone levels. By contrast, in the northern YRD region, substantial $PM_{2.5}$ reductions were linked with only slight ozone decreases, while in the southeast YRD region, even minor $PM_{2.5}$ increases were associated with significantly larger ozone increases. This pattern in the southeast YRD region could be attributed to the shared precursors of $PM_{2.5}$ and ozone, such as NO_x and VOCs, which influence their interactions. The responses of $PM_{2.5}$ and ozone to precursor emissions are complex, exhibiting non-linear behaviour that varies across different locations and times [47]. Hence, identifying and targeting the common precursors in these areas could be particularly effective in simultaneously controlling $PM_{2.5}$ and ozone pollution. Additionally, decreases in other pollutants were associated with smaller SVD values; for example, CO and SO_2 decreased by 0.06 ± 0.01 and 0.07 ± 0.05 μg m^{-3}, respectively, and had SVD values of −0.21 ± 0.03 and −0.03 ± 0.03 μg m^{-3} respectively.

3.4. Distinct drivers of ozone rebound validated by causal inference

The permutation importance and SHAP analyses of the RF models identified the variables closely associated with ozone concentration at each monitoring site. Additional analyses are required to validate the causative effects of environmental variables on ozone levels. Thus, we framed this issue as a causal inference problem, with ozone concentration as the outcome and various

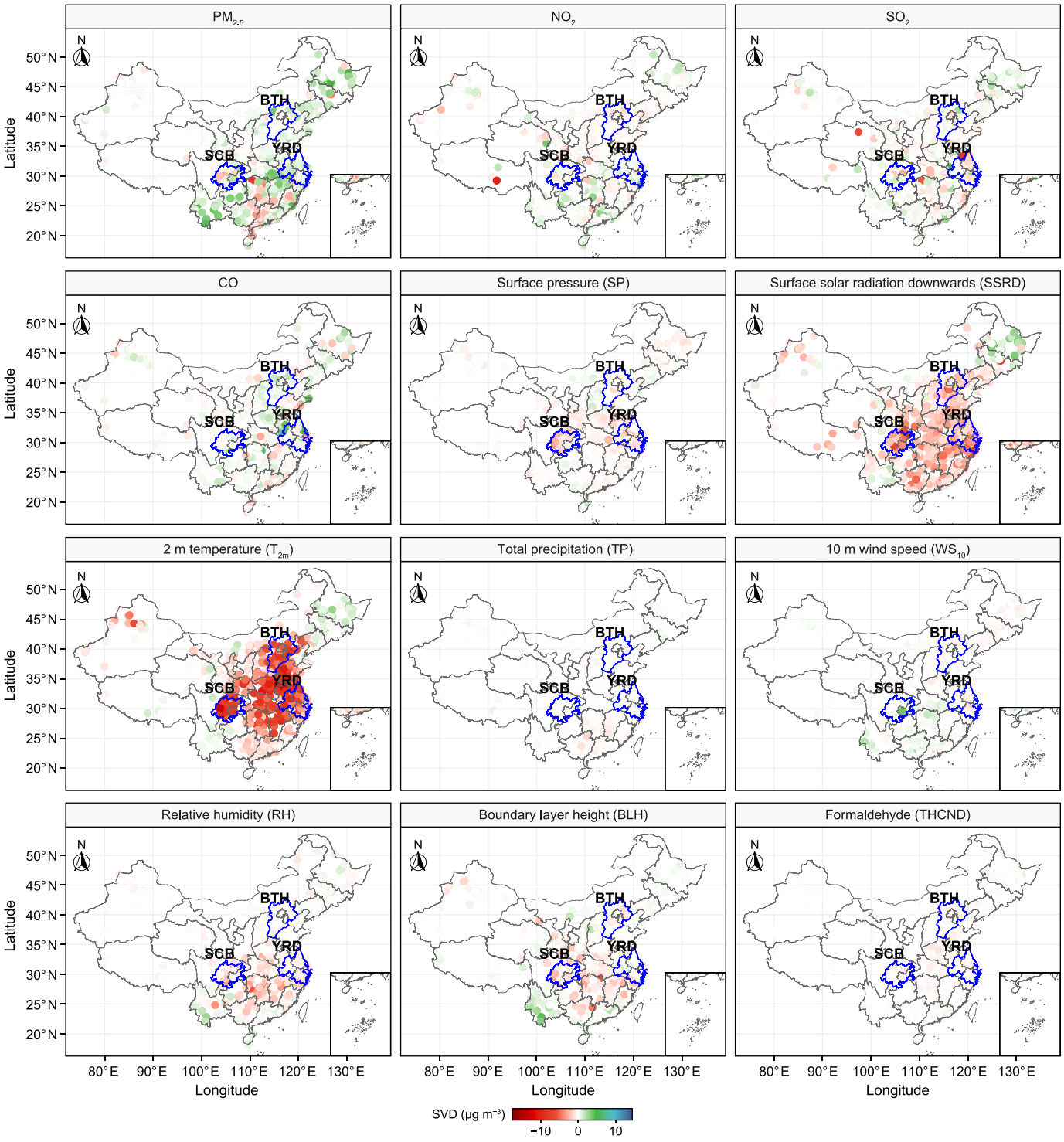

Fig. 5. Spatial distribution of the annual average SHAP value difference (SVD) for all features impacting ozone rebound, as determined by the random forest (RF) model. Shapley additive explanations (SHAP) represent the contribution of each feature to the model's prediction. Key regions include Beijing–Tianjin–Hebei (BTH), Sichuan Basin (SCB), and Yangtze River Delta (YRD), marked with blue boundaries.

meteorological and anthropogenic factors as the treatments. The aim is to calculate the ATE, which should be interpreted as the change in ozone concentration resulting from a unit increase in any given meteorological or emission variable. For example, when assessing the impact of temperature on ozone, temperature was considered a treatment, while emissions and other meteorological conditions were considered confounding covariates.

In the CF models, the ATE was determined as the weighted average of the ratio of the conditional covariance between treatment and outcome to the conditional variance of the treatment, with each conditioned on covariates. As indicated by the spatial variation of ATEs for different features (Fig. 6), temperature and

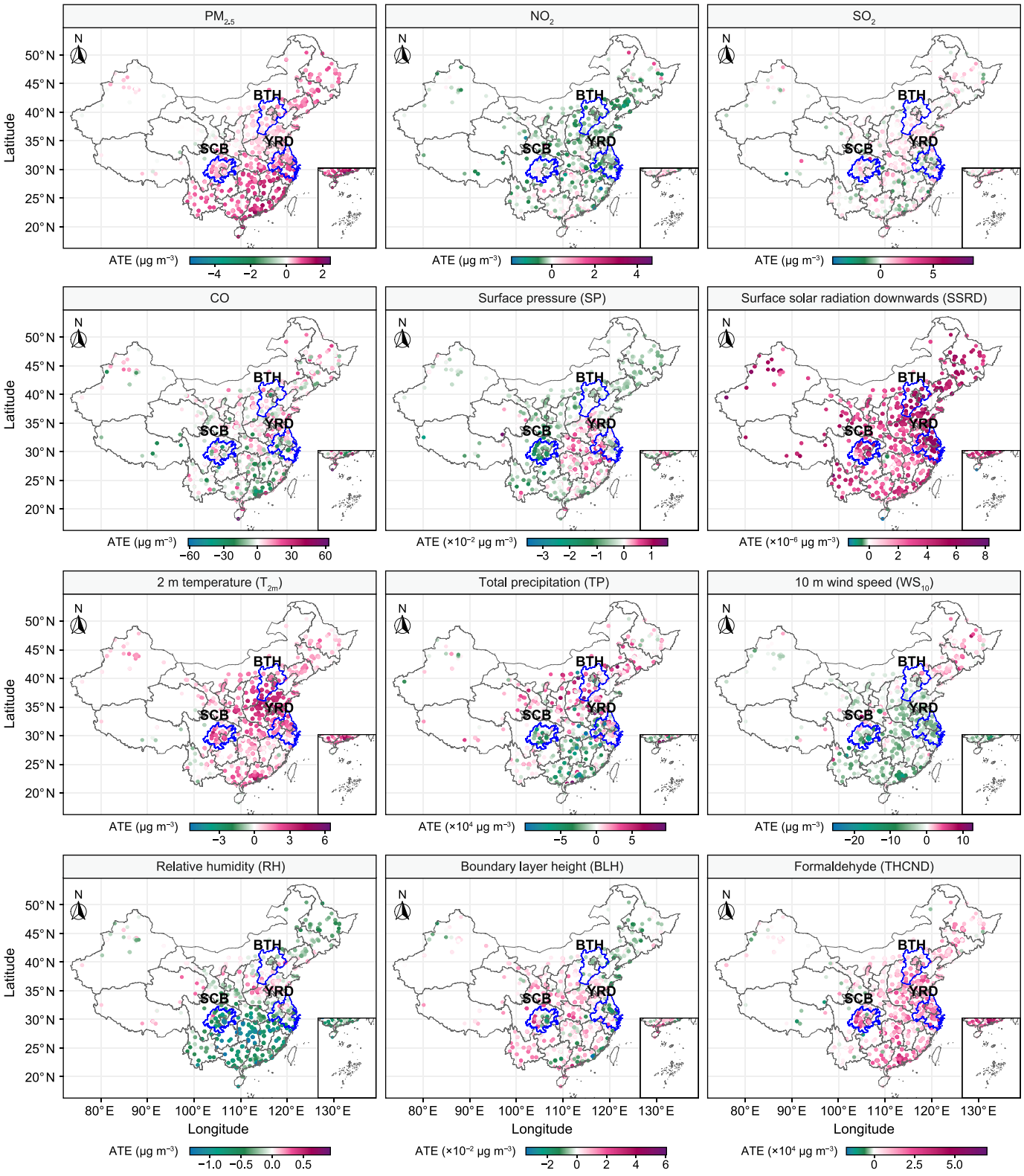

Fig. 6. Spatial variation of causal forest-based average treatment effect (ATE) across different features. Key regions include Beijing–Tianjin–Hebei (BTH), Sichuan Basin (SCB), and Yangtze River Delta (YRD), marked with blue boundaries.

solar radiation consistently exhibited positive ATE values across all monitoring sites, signifying that increases in these factors directly led to rising ozone levels. The ATE for temperature was particularly pronounced in central and eastern China, as well as in the SCB, while that for solar radiation extended beyond these regions, indicating that it had a broader impact. Their geographical

differences in ATE values align with their importance calculated by RF models, underscoring their influence on VOC emissions, potentially shaped by both geographical and socioeconomic factors [51–53]. Relative humidity generally reduces ozone concentrations across most regions of China, as it facilitates wet deposition and drives chemical reactions that diminish ozone levels [54,55]. Other meteorological parameters displayed varying ATEs across different regions. For instance, wind speed had a positive ATE in northeastern China, likely due to transport effects, as regional transport is reported to be a significant contributor to ozone pollution in this area [56]. In contrast, wind principally functioned in other regions to dilute and disperse pollutants, thereby reducing ozone levels [57]. Boundary layer height also shows inconsistent spatial effects on ozone concentrations [55]. As indicated by a study in Beijing [58], when the boundary layer height is below ~1000 m, an increase enhances chemical production due to greater light availability, while vertical mixing remains weak, resulting in a net gain in surface ozone concentration. However, as the boundary layer height rises, especially above ~2000 m, the effect shifts to a net ozone loss due to vigorous vertical mixing and reduced precursor concentrations at the surface. Surface pressure effects on ground ozone concentrations are inconsistent across China, mainly due to the influence of surface pressure on other meteorological factors, particularly wind speed and direction, which are further shaped by regional geographical conditions [59]. It should be noted that this does not necessarily imply that TP has a significant influence on ozone, as the ATE of TP refers to the change in ozone concentration resulting from a unit increase of TP.

Regarding emissions, $PM_{2.5}$ showed a positive ATE for ozone, particularly in southern regions, aligning with its importance as calculated by RF. NO_2 generally exhibited a negative ATE across most areas, while HCHO showed a positive ATE for ozone in many regions, indicating that ozone formation was more sensitive to VOCs in these areas. Both SO_2 and CO showed variable ATEs for ozone across different regions, reflecting their complex roles in photochemical reactions and their influence on the rate of ozone formation.

The ATE values were also calculated based on the partial dependence plot (PDP) obtained from the RF models. To determine the ATEs for different variables, we calculated the slopes between the points on the PDP that correspond to the maximum and minimum predicted values for each predictor [31] (Supplementary Material Fig. S15). As indicated in Fig. 7, the ATEs estimated by the CF models were generally comparable with those from the RF models for all variables (with Pearson's *R* values ranging from 0.21 to 0.73), supporting the argument that the identified associations for variables may reflect causal influences on ozone formation. This further confirms that temperature and solar radiation were the primary contributors to the ozone increase in 2022. Simultaneously, the spatial variation of ATEs for ozone across different variables, as derived from the CF and RF models, demonstrates notable consistency (Supplementary Material Fig. S16). Overall, these findings reveal that meteorological factors likely play a more important role than emissions in ozone rebound.

It should be noted that both the RF and CF have limitations in identifying causal relationships. PDP observes the changes induced by one feature by holding others constant, while SHAP values assess the marginal contributions of the features using a game-theoretic method. These two approaches make it possible to identify nonlinear ATE based on data associations. On the other hand, CF models, specifically designed for causal inference, directly evaluate the relationship between treatments and outcomes to estimate the ATE. CF models evaluate the causal impact of treatments by leveraging the conditional distributions of these treatments given the covariates, thereby isolating the impacts of treatments and minimising the biases from confounding variables. Nevertheless, the 'grf' package we used in this study, which is also the most widely employed CF tool for estimating ATE for continuous variables, may predominantly rely on linear relationships. However, temperature, the key driver of the ozone rebound in 2022, displayed substantial nonlinearity in its impacts on ozone concentrations across different temperature ranges. We applied a linear approximation to the PDP results to align the PDP-derived nonlinear ATE from the RF models with the linear ATE from the CF models. By assessing their magnitudes, we could use the causal effects estimated by CF to substantiate and interpret the nonlinear ATE estimated by RF as indicative of a causal relationship. The complementary use of CF and RF allowed us to validate that the relationships identified by RF are causally grounded, which further enabled us to confidently use non-linear SHAP values and SVD calculations to derive reliable insights into the drivers of the 2022 ozone rebound, effectively combining the non-linear predictive strengths of RF with the causal rigor of CF to accurately identify and quantify the factors influencing the rebound.

The causal inference supported the finding that the major driver of the ozone rebound in 2022 was the rising temperatures across the country. This effect can be exacerbated by strong surface solar radiation, which might have a broader geological and temporal coverage. Given our warming climate and the potential for stronger surface solar radiation [60,61], understanding the response of ozone to these meteorological conditions is essential for devising effective regional-specific strategies to control emission precursors.

3.5. Impact of shifting HCHO-NO_2 ratio on ozone rebound

Although our ML and causal inference analyses identified and quantified the drivers of the ozone rebound in 2022, it is important to recognise that NO_x and VOCs are critical ozone precursors. Therefore, controlling NO_x and VOC emissions remains a fundamental strategy for ozone mitigation. In 2022, satellite-derived pollution data from China revealed significant differences from 2021, with NO_2 column concentrations decreasing by an average of $(3.12 \pm 0.06) \times 10^{-5}$ mol m^{-2} at 87.01 % of stations and increases in HCHO column concentrations averaging $(9.45 \pm 0.05) \times 10^{-5}$ mol m^{-2} at 51.06 % of stations. Simultaneously, an enhanced positive correlation between ground-level ozone and NO_2 concentrations was observed, particularly in the southern and northeast regions, as evidenced in Supplementary Material Fig. S17.

Regarding the NOx titration, we further investigated the variation of the HCHO—NO_2 ratio (formaldehyde to nitrogen ratio, FNR) between 2021 and 2022. We applied the thresholds established by Wang et al. [62] to delineate ozone formation regimes in 2021 and 2022 based on the calculation of HCHO—NO_2 ratio. Thresholds of 2.3 and 4.2 differentiate VOC-limited from transition regimes and transition from NO_x-limited regimes, respectively (Supplementary Material Fig. S18). In 2021, 51.61 % of monitoring stations were identified as VOC-limited, predominantly in industrial and heavy traffic areas such as the SCB, BTH, and YRD regions and their adjacent areas, where the NO_x titration effect was notably significant [63,64]. Conversely, 15.96 % of the stations were under NO_x-limited conditions, mostly in the less developed regions of northeast and southwestern China. By 2022, 89.91 % of monitoring stations recorded increased FNR (Supplementary Material Fig. S18). There was a significant regime shift at 16.91 % of the stations from VOC-limited to transitional regimes, mainly in central and eastern China and a few stations in the BTH and SCB regions (Supplementary Material Fig. S19). Additionally, 10.22 % of the stations shifted to NO_x-limited regimes, mostly in southern and northeastern China. The shift observed in the SCB region and central, eastern, and southern China could be linked to elevated

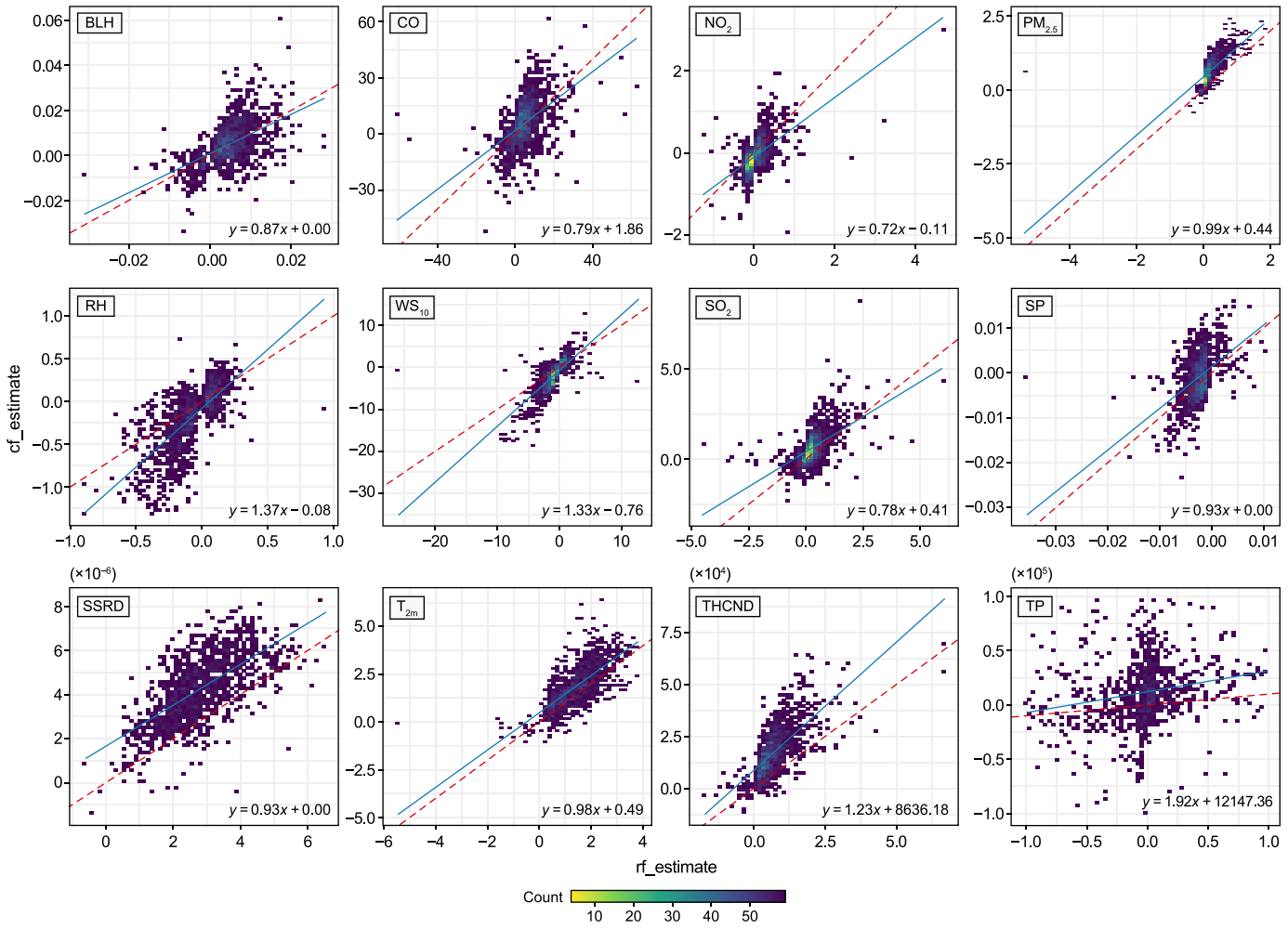

Fig. 7. Comparisons of the average treatment effect (ATE) of different variables from the random forest (RF) and causal forest (CF) models. Features include boundary layer height (BLH), CO, NO_2, $PM_{2.5}$, relative humidity (RH), SO_2, surface pressure (SP), surface solar radiation downwards (SSRD), 2 m temperature (T_{2m}), formaldehyde (THCHD), total precipitation (TP), and 10 m wind speed (WS_{10}).

temperatures that enhanced VOC emissions from biogenic and certain anthropogenic sources, as well as further reductions in NO_x, all of which increased ozone formation's sensitivity to NO_x (Supplementary Material Fig. S2).

Notably, projections from the Model of Emissions of Gases and Aerosols from Nature indicate that with ongoing climate warming, emissions of isoprene and monoterpene, which are key BVOCs, are expected to rise by 11.13 % and 25.20 %, respectively, under the Representative Concentration Pathway (RCP) 4.5 and 8.5 scenarios by the 2050s compared to 2015 [65]. These increases in BVOCs are most pronounced in the regions that experienced shifts towards transitional and NO_x-limited ozone regimes in 2022, including the SCB region and central, eastern, and southern China. Given the current emission inventories (http://meicmodel.org.cn), VOC and NO_x reductions are crucial in these regions, but the effectiveness of VOC mitigation on ozone will decline due to the rising BVOCs caused by climate change [66], emphasising the need to prioritise the reduction of NO_x. Referring to the decrements of NO_2 and HCHO shown in Supplementary Material Fig. S18, the observed shifts in the ozone regime in northeast China and parts of the BTH region could be predominantly attributed to a more significant reduction in NO_x than in VOCs, highlighting the critical need to manage both pollutants effectively. Furthermore, in the remaining regions classified as VOCs-limited, including parts of the BTH, YRD, and SCB regions, maintaining the priority to reduce VOC emissions is essential in effectively lowering ozone levels alongside careful management of NO_x emissions to prevent any potential reversal of ozone reduction through NO titration.

4. Conclusions

This research utilised RF modelling, SHAP additive explanations, and CF modelling to elucidate the primary drivers behind the widespread ozone rebound observed across China in 2022. The causal inference-aided analysis highlighted that meteorological changes, namely, the escalated temperature and enhanced solar radiation, contributed significantly to the surge in ozone levels. Specifically, temperature increases accounted for up to 38.90 % of the ozone rebound, with solar radiation contributing an additional 18.47 %. Temperature significantly raised ozone levels in the YRD, BTH, and SCB regions from the warmer months to autumn, while solar radiation consistently increased ozone across diverse regions and through all seasons. These climatic changes accelerated the photochemical reactions necessary for ozone formation, underscoring the sensitivity of ozone dynamics to meteorological variations. Specifically, the shifts in ozone formation regimes in 2022,

with increased sensitivity to NO_2, occurred primarily in the SCB region and central, eastern, and southern China. These regions coincide with projected increases in BVOCs due to climate change, underlining the need for sustained NO_x reductions to mitigate ozone levels in these areas effectively. The findings call for adaptive pollution control measures finely tuned to local environmental and climatic conditions. This research deepens our understanding of ozone dynamics and establishes a causal basis for crafting air-quality policies that are responsive to climate change.

CRediT authorship contribution statement

Lin Wang: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Baihua Chen:** Writing – review & editing, Formal analysis, Data curation. **Jingyi Ouyang:** Formal analysis. **Yanshu Mu:** Formal analysis. **Ling Zhen:** Validation. **Lin Yang:** Validation. **Wei Xu:** Writing – review & editing, Visualization, Methodology, Formal analysis, Conceptualization. **Lina Tang:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Data availability

The data and code availability are detailly described in the manuscript.

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.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (42307591), Fujian Provincial Natural Science Foundation project (2022J05096), National Key R&D Program of China (2022YFF1301302).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ese.2025.100524.

References

[1] L. Conibear, C.L. Reddington, B.J. Silver, Y. Chen, C. Knote, S.R. Arnold, D.V. Spracklen, Sensitivity of air pollution exposure and Disease Burden to emission changes in China using machine learning Emulation, GeoHealth 6 (2022) e2021GH000570, https://doi.org/10.1029/2021gh000570.
[2] X. Long, Y. Han, Q.Y. Wang, X.K. Li, T. Feng, Y.C. Wang, Y. Wang, S.L. Zhang, Y.M. Han, G.H. Li, X.X. Tie, J.J. Cao, Y. Chen, Adverse effects of ozone pollution on net primary productivity in the North China plain, Geophys. Res. Lett. 51 (2024), https://doi.org/10.1029/2023gl105209.
[3] V. Barry, M. Klein, A. Winquist, H.H. Chang, J.A. Mulholland, E.O. Talbott, J.R. Rager, P.E. Tolbert, S.E. Sarnat, Characterization of the concentration-response curve for ambient ozone and acute respiratory morbidity in 5 US cities, J. Expo, Sci. Environ. Epidemiology 29 (2019) 267–277, https://doi.org/10.1038/s41370-018-0048-7.
[4] L.M. Paulin, A.J. Gassett, N.E. Alexis, K. Kirwa, R.E. Kanner, S. Peters, J.A. Krishnan, I.P. Robert, M. Dransfield, P.G. Woodruff, C.B. Cooper, R.G. Barr, A.P. Comellas, C.S. Pirozzi, M. Han, E.A. Hoffman, F.J. Martinez, H. Woo, R.D. Peng, A. Fawzy, N. Putcha, P.N. Breysse, J.D. Kaufman, N.N. Hansel, for S. investigators, Association of Long-Term Ambient Ozone Exposure with Respiratory Morbidity in Smokers, vol. 180, 2020, pp. 106–115, https://doi.org/10.1001/jamainternmed.2019.5498.
[5] S.-Y. Kim, E. Kim, W.J. Kim, Health effects of ozone on respiratory diseases, Tuberc. Respir. Dis. 83 (2020) S6–S11, https://doi.org/10.4046/trd.2020.0154.
[6] L. Emberson, Effects of ozone on agriculture, forests and grasslands, Philos. Trans. R. Soc. A 378 (2020) 20190327, https://doi.org/10.1098/rsta.2019.0327.
[7] K.S. Docherty, W. Wu, Y.B. Lim, P.J. Ziemann, Contributions of organic Peroxides to secondary aerosol formed from reactions of monoterpenes with O3, Environ. Sci. Technol. 39 (2005) 4049–4059, https://doi.org/10.1021/es050228s.
[8] Y. Wang, Y. Zhao, Y. Liu, Y. Jiang, B. Zheng, J. Xing, Y. Liu, S. Wang, C.P. Nielsen, Sustained emission reductions have restrained the ozone pollution over China, Nat. Geosci. 16 (2023) 967–974, https://doi.org/10.1038/s41561-023-01284-2.
[9] C.A. ASIAN, Executive summary of China air Report 2023. http://Www.Allaboutair.Cn/Uploads/231027/ChinaAir2023EN.Pdf, 2023.
[10] Y. Ni, Y. Yang, H. Wang, H. Li, M. Li, P. Wang, K. Li, H. Liao, Contrasting changes in ozone during 2019–2021 between eastern and the other regions of China attributed to anthropogenic emissions and meteorological conditions, Sci. Total Environ. 908 (2024) 168272, https://doi.org/10.1016/j.scitotenv.2023.168272.
[11] Li Yang, Chen Liao, Analysis of surface ozone pollution in China amid the record summertime extreme heat of 2022.pdf, Chin. J. Atmos. Sci. (2023), https://doi.org/10.3878/j.issn.1006-9895.2302.22211.
[12] W. Qiao, K. Li, Z. Yang, L. Chen, H. Liao, Implications of the extremely hot summer of 2022 on urban ozone control in China, Atmos. Ocean. Sci. Lett (2024) 100470, https://doi.org/10.1016/j.aosl.2024.100470.
[13] Z. Zhang, J. Jiang, B. Lu, X. Meng, H. Herrmann, J. Chen, X. Li, Attributing increases in ozone to accelerated oxidation of volatile organic compounds at reduced nitrogen oxides concentrations, PNAS Nexus 1 (2022) pgac266, https://doi.org/10.1093/pnasnexus/pgac266.
[14] H.J. Lee, T. Kuwayama, M. FitzGibbon, Trends of ambient O3 levels associated with O3 precursor gases and meteorology in California: Synergies from ground and satellite observations, Remote Sens. Environ. 284 (2023) 113358, https://doi.org/10.1016/j.rse.2022.113358.
[15] X. Yan, Y. Guo, Y. Zhang, J. Chen, Y. Jiang, C. Zuo, W. Zhao, W. Shi, Combining physical mechanisms and deep learning models for hourly surface ozone retrieval in China, J. Environ. Manag. 351 (2024) 119942, https://doi.org/10.1016/j.jenvman.2023.119942.
[16] Y. Lyu, H. Wu, X. Liu, F. Han, F. Lv, X. Pang, J. Chen, Co-Occurring extremes of fine particulate matter (PM2.5) and ground-level ozone in the summer of southern China, Geophys. Res. Lett. 51 (2024), https://doi.org/10.1029/2023gl106527.
[17] M. Ma, G. Yao, J. Guo, K. Bai, Distinct spatiotemporal variation patterns of surface ozone in China due to diverse influential factors, J. Environ. Manag. 288 (2021) 112368, https://doi.org/10.1016/j.jenvman.2021.112368.
[18] Y.-W. Chen, S. Medya, Y.-C. Chen, Investigating variable importance in ground-level ozone formation with supervised learning, Atmospheric Environ 282 (2022) 119148, https://doi.org/10.1016/j.atmosenv.2022.119148.
[19] K. Li, D.J. Jacob, H. Liao, J. Zhu, V. Shah, L. Shen, K.H. Bates, Q. Zhang, S. Zhai, A two-pollutant strategy for improving ozone and particulate air quality in China, Nat. Geosci. 12 (2019) 906–910, https://doi.org/10.1038/s41561-019-0464-x.
[20] Y. Sun, H. Yin, X. Lu, J. Notholt, M. Palm, C. Liu, Y. Tian, B. Zheng, The drivers and health risks of unexpected surface ozone enhancements over the Sichuan Basin, China, in 2020, Atmos. Chem. Phys. 21 (2021) 18589–18608, https://doi.org/10.5194/acp-21-18589-2021.
[21] H. Han, J. Liu, L. Shu, T. Wang, H. Yuan, Local and synoptic meteorological influences on daily variability in summertime surface ozone in eastern China, Atmos. Chem. Phys. 20 (2019) 203–222, https://doi.org/10.5194/acp-20-203-2020.
[22] H. Xu, H. Yu, B. Xu, Z. Wang, F. Wang, Y. Wei, W. Liang, J. Liu, D. Liang, Y. Feng, G. Shi, Machine learning coupled structure mining method visualizes the impact of multiple drivers on ambient ozone, Commun. Earth Environ. 4 (2023) 265, https://doi.org/10.1038/s43247-023-00932-0.
[23] J. Zhan, Y. Liu, W. Ma, X. Zhang, X. Wang, F. Bi, Y. Zhang, Z. Wu, H. Li, Ozone formation sensitivity study using machine learning coupled with the reactivity of volatile organic compound species, Atmos. Meas. Tech. 15 (2022) 1511–1520, https://doi.org/10.5194/amt-15-1511-2022.
[24] Z. Mai, H. Shen, A. Zhang, H.Z. Sun, L. Zheng, J. Guo, C. Liu, Y. Chen, C. Wang, J. Ye, L. Zhu, T.-M. Fu, X. Yang, S. Tao, Convolutional neural networks facilitate process understanding of Megacity ozone temporal variability, Environ. Sci. Technol. (2024), https://doi.org/10.1021/acs.est.3c07907.
[25] Y. Tan, Y. Zhang, T. Wang, T. Chen, J. Mu, L. Xue, Dissecting drivers of ozone pollution during the 2022 Multicity lockdowns in China Sheds light on Future control direction, Environ. Sci. Technol. (2024), https://doi.org/10.1021/acs.est.4c01197.
[26] J. Runge, A. Gerhardus, G. Varando, V. Eyring, G. Camps-Valls, Causal inference for time series, Nat. Rev. Earth Environ. 4 (2023) 487–505, https://doi.org/10.1038/s43017-023-00431-y.
[27] Q. Kang, X. Song, X. Xin, B. Chen, Y. Chen, X. Ye, B. Zhang, Machine learning-aided causal inference Framework for environmental data analysis: a COVID-19 Case study, Environ. Sci. Technol. 55 (2021) 13400–13410, https://doi.org/10.1021/acs.est.1c02204.
[28] T. Tesch, S. Kollet, J. Garcke, Causal deep learning models for studying the Earth system, Geosci. Model Dev. (GMD) 16 (2023) 2149–2166, https://doi.org/10.5194/gmd-16-2149-2023.
[29] A.K. Leist, M. Klee, J.H. Kim, D.H. Rehkopf, S.P.A. Bordas, G. Muniz-Terrera, S. Wade, Mapping of machine learning approaches for description, prediction,

and causal inference in the social and health sciences, Sci. Adv. 8 (2022), https://doi.org/10.1126/sciadv.abk1942 eabk1942.
[30] P. Sanchez, J.P. Voisey, T. Xia, H.I. Watson, A.Q. O'Neil, S.A. Tsaftaris, Causal machine learning for healthcare and precision medicine, R. Soc. Open Sci. 9 (2022) 220638, https://doi.org/10.1098/rsos.220638.
[31] Y. Ma, D. Woolf, M. Fan, L. Qiao, R. Li, J. Lehmann, Global crop production increase by soil organic carbon, Nat. Geosci. 16 (2023) 1159–1165, https://doi.org/10.1038/s41561-023-01302-3.
[32] S. Athey, J. Tibshirani, S. Wager, Generalized random forests, Ann. Stat. 47 (2019) 1148–1178, https://doi.org/10.1214/18-aos1709.
[33] H. Gulen, C. Jens, T.B. Page, An Application of causal forest in Corporate finance: how does financing affect Investment? SSRN Electron. J. (2020) https://doi.org/10.2139/ssrn.3583685.
[34] J.P. Veefkind, I. Aben, K. McMullan, H. Förster, J. de Vries, G. Otter, J. Claas, H.J. Eskes, J.F. de Haan, Q. Kleipool, M. van Weele, O. Hasekamp, R. Hoogeveen, J. Landgraf, R. Snel, P. Tol, P. Ingmann, R. Voors, B. Kruizinga, R. Vink, H. Visser, P.F. Levelt, TROPOMI on the ESA Sentinel-5 Precursor: a GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sens. Environ. 120 (2012) 70–83, https://doi.org/10.1016/j.rse.2011.09.027.
[35] A. Liaw, M. Wiener, Classification and Regression by Randomforest, R. News 2 (n.d.) 18–22. http://CRAN.R-project.org/doc/Rnews/.
[36] S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell. 2 (2020) 56–67, https://doi.org/10.1038/s42256-019-0138-9.
[37] S.M. Lundberg, B. Nair, M.S. Vavilala, M. Horibe, M.J. Eisses, T. Adams, D.E. Liston, D.K.-W. Low, S.-F. Newman, J. Kim, S.-I. Lee, Explainable machine-learning predictions for the prevention of hypoxaemia during surgery, Nat. Biomed. Eng. 2 (2018) 749–760, https://doi.org/10.1038/s41551-018-0304-0.
[38] E.Y. Pfannerstill, C. Arata, Q. Zhu, B.C. Schulze, R. Ward, R. Woods, C. Harkins, R.H. Schwantes, J.H. Seinfeld, A. Bucholtz, R.C. Cohen, A.H. Goldstein, Temperature-dependent emissions dominate aerosol and ozone formation in Los Angeles, Science 384 (2024) 1324–1329, https://doi.org/10.1126/science.adg8204.
[39] F. LORETO, C. BARTA, F. BRILLI, I. NOGUES, On the induction of volatile organic compound emissions by plants as consequence of wounding or fluctuations of light and temperature, Plant, Cell Environ 29 (2006) 1820–1828, https://doi.org/10.1111/j.1365-3040.2006.01561.x.
[40] W. Wu, T.-M. Fu, S.R. Arnold, D.V. Spracklen, A. Zhang, W. Tao, X. Wang, Y. Hou, J. Mo, J. Chen, Y. Li, X. Feng, H. Lin, Z. Huang, J. Zheng, H. Shen, L. Zhu, C. Wang, J. Ye, X. Yang, Temperature-dependent evaporative anthropogenic VOC emissions significantly exacerbate regional ozone pollution, Environ. Sci. Technol. 58 (2024) 5430–5441, https://doi.org/10.1021/acs.est.3c09122.
[41] W. Wang, X. Liu, J. Bi, Y. Liu, A machine learning model to estimate ground-level ozone concentrations in California using TROPOMI data and high-resolution meteorology, Environ. Int. 158 (2022) 106917, https://doi.org/10.1016/j.envint.2021.106917.
[42] Z. Li, H. Dong, Z. Zhang, L. Luo, S. He, Estimation of near-ground ozone with high Spatio-temporal resolution in the Yangtze River Delta region of China based on a temporally ensemble model, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 16 (2023) 7051–7061, https://doi.org/10.1109/jstars.2023.3298996.
[43] Y. Xia, Y. Hu, Y. Huang, J. Bian, C. Zhao, J. Wei, Y. Yan, F. Xie, J. Lin, Concurrent hot extremes and high ultraviolet radiation in summer over the Yangtze Plain and their possible impact on surface ozone, Environ. Res. Lett. 17 (2022) 064001, https://doi.org/10.1088/1748-9326/ac6c3c.
[44] Y. Cao, Y. Qu, J. Ma, Classification of ozone pollution and analysis of meteorological factors in the Yangtze River Delta, Big Earth Data 7 (2023) 318–337, https://doi.org/10.1080/20964471.2022.2157093.
[45] L. Zhang, L. Wang, B. Liu, G. Tang, B. Liu, X. Li, Y. Sun, M. Li, X. Chen, Y. Wang, B. Hu, Contrasting effects of clean air actions on surface ozone concentrations in different regions over Beijing from May to September 2013–2020, Sci. Total Environ. 903 (2023) 166182, https://doi.org/10.1016/j.scitotenv.2023.166182.
[46] J. Mao, L. Wang, C. Lu, J. Liu, M. Li, G. Tang, D. Ji, N. Zhang, Y. Wang, Meteorological mechanism for a large-scale persistent severe ozone pollution event over eastern China in 2017, J. Environ. Sci. 92 (2020) 187–199, https://doi.org/10.1016/j.jes.2020.02.019.
[47] L. Qi, J. Yin, J. Li, X. Duan, Detecting causal relationships between fine particles and ozone based on observations in four typical cities of China, Environ. Res. Lett. 19 (2024) 054006, https://doi.org/10.1088/1748-9326/ad376d.
[48] P. Qiu, L. Zhang, X. Wang, Y. Liu, S. Wang, S. Gong, Y. Zhang, A new approach of air pollution regionalization based on geographically weighted variations for multi-pollutants in China, Sci. Total Environ. 873 (2023) 162431, https://doi.org/10.1016/j.scitotenv.2023.162431.
[49] Z. Tan, K. Lu, X. Ma, S. Chen, L. He, X. Huang, X. Li, X. Lin, M. Tang, D. Yu, A. Wahner, Y. Zhang, Multiple impacts of aerosols on O3 production are largely Compensated: a Case study Shenzhen, China, Environ. Sci. Technol. 56 (2022) 17569–17580, https://doi.org/10.1021/acs.est.2c06217.
[50] C. Ye, H. Wang, X. Li, K. Lu, Y. Zhang, Atmospheric reactive nitrogen species weaken the air quality response to emission reductions in China, Environ. Sci. Technol. 58 (2024) 6066–6070, https://doi.org/10.1021/acs.est.3c10927.
[51] H. Wang, X. Liu, C. Wu, G. Lin, Regional to global distributions, trends, and drivers of biogenic volatile organic compound emission from 2001 to 2020, Atmos. Chem. Phys. 24 (2024) 3309–3328, https://doi.org/10.5194/acp-24-3309-2024.
[52] G. Wang, N. Zhao, H. Zhang, G. Li, G. Xin, Spatiotemporal distributions of ambient volatile organic compounds in China: Characteristics and sources, Aerosol Air Qual. Res. 22 (2022) 210379, https://doi.org/10.4209/aaqr.210379.
[53] M. Song, X. Li, S. Yang, X. Yu, S. Zhou, Y. Yang, S. Chen, H. Dong, K. Liao, Q. Chen, K. Lu, N. Zhang, J. Cao, L. Zeng, Y. Zhang, Spatiotemporal variation, sources, and secondary transformation potential of volatile organic compounds in Xi'an, China, Atmos. Chem. Phys. 21 (2021) 4939–4958, https://doi.org/10.5194/acp-21-4939-2021.
[54] L. Tong, H. Zhang, J. Yu, M. He, N. Xu, J. Zhang, F. Qian, J. Feng, H. Xiao, Characteristics of surface ozone and nitrogen oxides at urban, suburban and rural sites in Ningbo, China, Atmos. Res. 187 (2017) 57–68, https://doi.org/10.1016/j.atmosres.2016.12.006.
[55] C. Gong, H. Liao, A typical weather pattern for ozone pollution events in North China, Atmos. Chem. Phys. 19 (2019) 13725–13740, https://doi.org/10.5194/acp-19-13725-2019.
[56] H. Wang, K. Ding, X. Huang, W. Wang, A. Ding, Insight into ozone profile climatology over northeast China from aircraft measurement and numerical simulation, Sci. Total Environ. 785 (2021) 147308, https://doi.org/10.1016/j.scitotenv.2021.147308.
[57] Y. Liu, T. Wang, Worsening urban ozone pollution in China from 2013 to 2017 – Part 1: the complex and varying roles of meteorology, Atmos. Chem. Phys. 20 (2020) 6305–6321, https://doi.org/10.5194/acp-20-6305-2020.
[58] C. Zhang, Z. Jiang, M. Liu, Y. Dong, J. Li, Relationship between summer time near-surface ozone concentration and planetary boundary layer height in Beijing, Atmos. Res. 293 (2023) 106892, https://doi.org/10.1016/j.atmosres.2023.106892.
[59] Z. Chen, R. Li, D. Chen, Y. Zhuang, B. Gao, L. Yang, M. Li, Understanding the causal influence of major meteorological factors on ground ozone concentrations across China, J. Clean. Prod. 242 (2020) 118498, https://doi.org/10.1016/j.jclepro.2019.118498.
[60] C. Gutirrez, S. Somot, P. Nabat, M. Mallet, L. Corre, E. van Meijgaard, O. Perpin, M. Gaertner, Future evolution of surface solar radiation and photovoltaic potential in Europe: investigating the role of aerosols, Environ. Res. Lett. 15 (2020) 034035, https://doi.org/10.1088/1748-9326/ab6666.
[61] J. Wu, J. Niu, Q. Qi, C.A. Gueymard, L. Wang, W. Qin, Z. Zhou, Reconstructing 10-km-resolution direct normal irradiance dataset through a hybrid algorithm, Renew. Sustain. Energy Rev. 204 (2024) 114805, https://doi.org/10.1016/j.rser.2024.114805.
[62] W. Wang, R. van der A, J. Ding, M. van Weele, T. Cheng, Spatial and temporal changes of the ozone sensitivity in China based on satellite and ground-based observations, Atmos. Chem. Phys. 21 (2021) 7253–7269, https://doi.org/10.5194/acp-21-7253-2021.
[63] Y. Tan, Y. Zhang, T. Wang, T. Chen, J. Mu, L. Xue, Dissecting drivers of ozone pollution during the 2022 Multicity lockdowns in China Sheds light on Future control direction, Environ. Sci. Technol. (2024), https://doi.org/10.1021/acs.est.4c01197.
[64] Z. Zhang, J. Jiang, B. Lu, X. Meng, H. Herrmann, J. Chen, X. Li, Attributing increases in ozone to accelerated oxidation of volatile organic compounds at reduced nitrogen oxides concentrations, PNAS Nexus 1 (2022) pgac266, https://doi.org/10.1093/pnasnexus/pgac266.
[65] S. Liu, J. Xing, H. Zhang, D. Ding, F. Zhang, B. Zhao, S.K. Sahu, S. Wang, Climate-driven trends of biogenic volatile organic compound emissions and their impacts on summertime ozone and secondary organic aerosol in China in the 2050s, Atmos. Environ. 218 (2019) 117020, https://doi.org/10.1016/j.atmosenv.2019.117020.
[66] Y. Gao, F. Yan, M. Ma, A. Ding, H. Liao, S. Wang, X. Wang, B. Zhao, W. Cai, H. Su, X. Yao, H. Gao, Unveiling the dipole synergic effect of biogenic and anthropogenic emissions on ozone concentrations, Sci. Total Environ. 818 (2022) 151722, https://doi.org/10.1016/j.scitotenv.2021.151722.