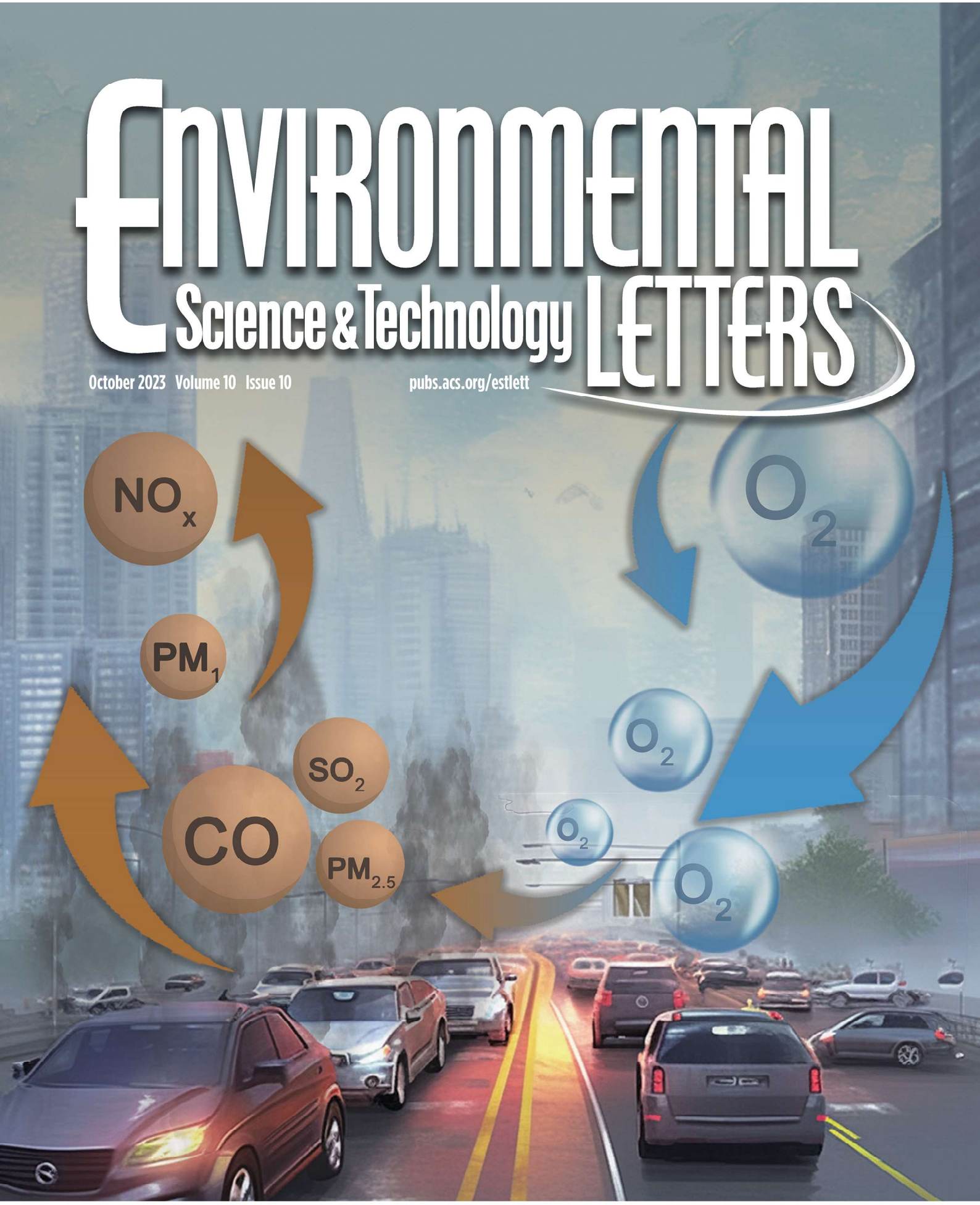
Environmental Science & Technology Letters

October 2023 Volume 10 Issue 10 pubs.acs.org/estlett



ACS Publications
Most Trusted. Most Cited. Most Read.

www.acs.org

Environmental Science & Technology Letters

pubs.acs.org/journal/estlcu

Letter

Revealing the Covariation of Atmospheric O_2 and Pollutants in an Industrial Metropolis by Explainable Machine Learning

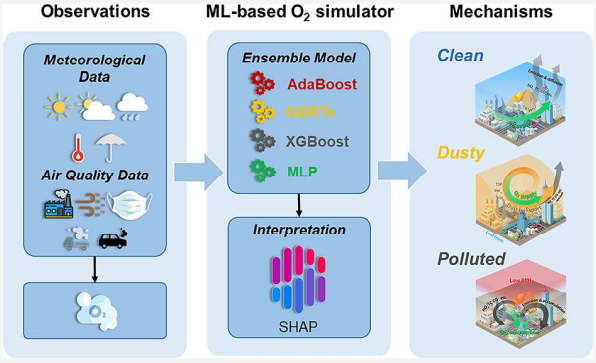
Xiaoyue Liu, Li Wang, Jianping Huang,* Yongqi Wang, Changyu Li, Lei Ding, Xinbo Lian, and Jinsen Shi

Cite This: *Environ. Sci. Technol. Lett.* 2023, 10, 851−858

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In urban areas, atmospheric O_2 actively participates in the process of anthropogenic emissions and energy consumption. However, the covariation between atmospheric O_2 and the emitted pollutants has yet to be thoroughly explored. This study examines the covariations between atmospheric O_2 and pollutants in Lanzhou, a semi-arid industrial metropolis. A machine learning (ML)-based O_2 simulator coupled with a SHapley Additive exPlanation (SHAP) algorithm is established to explore and interpret their covariations under diverse conditions. Our findings indicate an increase of 16.3 ppm in the O_2 concentration associated with the atmospheric transport of natural dust particles during dusty weather events. This suggests that natural dust transport can mitigate the depletion of atmospheric O_2 caused by primary emissions and secondary formation of anthropogenic particulate matter. Furthermore, we identify a nonlinear relationship between the concentrations of O_2 and pollutant concentrations, which likely arises from their distinct diffusive abilities. The study highlights the unique pollution characteristics in a semi-arid urban downtown and demonstrates the ability of ML-based methodology to reproduce and interpret the environmental and anthropogenic impacts on the local carbon−oxygen balance.



KEYWORDS: *urban O_2 observations, explainable machine learning, air pollution, anthropogenic impact, emission inventory*

INTRODUCTION

Cities are highly concentrated areas of human activities, contributing to >70% of global CO_2 emissions.[1] The emission of CO_2 from the combustion of fossil fuel-related sources in cities involves the removal of a corresponding amount of O_2 from the urban atmosphere, making these areas a significant contributor to anthropogenic O_2 consumption, especially in rapidly urbanizing regions with high levels of industrialization.[2] The current rate of decline in the level of atmospheric O_2 is approximately 3000 times faster than that documented in ice cores over the past 1 million years.[3] However, only a few stations around the world have conducted long-term atmospheric O_2 measurements at the parts per million level in urban regions.[2,4]

O_2 is the only reactant in fossil fuel combustion processes $[C_xH_y + (x + y/4)O_2 \rightarrow xCO_2 + (y/2)H_2O]$[5] that can be directly measured in the atmosphere. In addition, O_2 is essential for vital bodily functions such as nerve and brain function, muscle movement, digestion, and hormone synthesis.[6,7] Elucidating covariations of O_2 and pollutants reveals drivers of pollution episodes and informs mitigation strategies for air pollution impacts.[8−10]

Machine learning (ML) models have been demonstrated to be a powerful tool for reconstructing, simulating, and predicting atmospheric pollution, including $PM_{2.5}$,[11−13] O_3,[9,14] NO_x,[12,15] etc., outperforming finely designed chemical transport models.[16] The use of ML models provides greater flexibility and efficiency when utilizing real-world data and is especially adept at revealing complex and hidden nonlinear correlations[17,18] that might not be easily identified using traditional physical models, providing new insights into the underlying mechanisms of the studied phenomena.[19] Recently, ensemble learning models that integrate multiple machine learning algorithms are widely applied to obtain better predictive performance,[20,21] because they can combine the strengths of multiple models, each of which may be good at different aspects of the problem. However, the opacity of the majority of ML models, or the "black box" effect, makes it challenging to physically interpret the output. As feature attribution methods progress, tools for analyzing feature importance are becoming increasingly accessible. The SHapley

Received: July 15, 2023
Revised: August 31, 2023
Accepted: September 1, 2023
Published: September 6, 2023

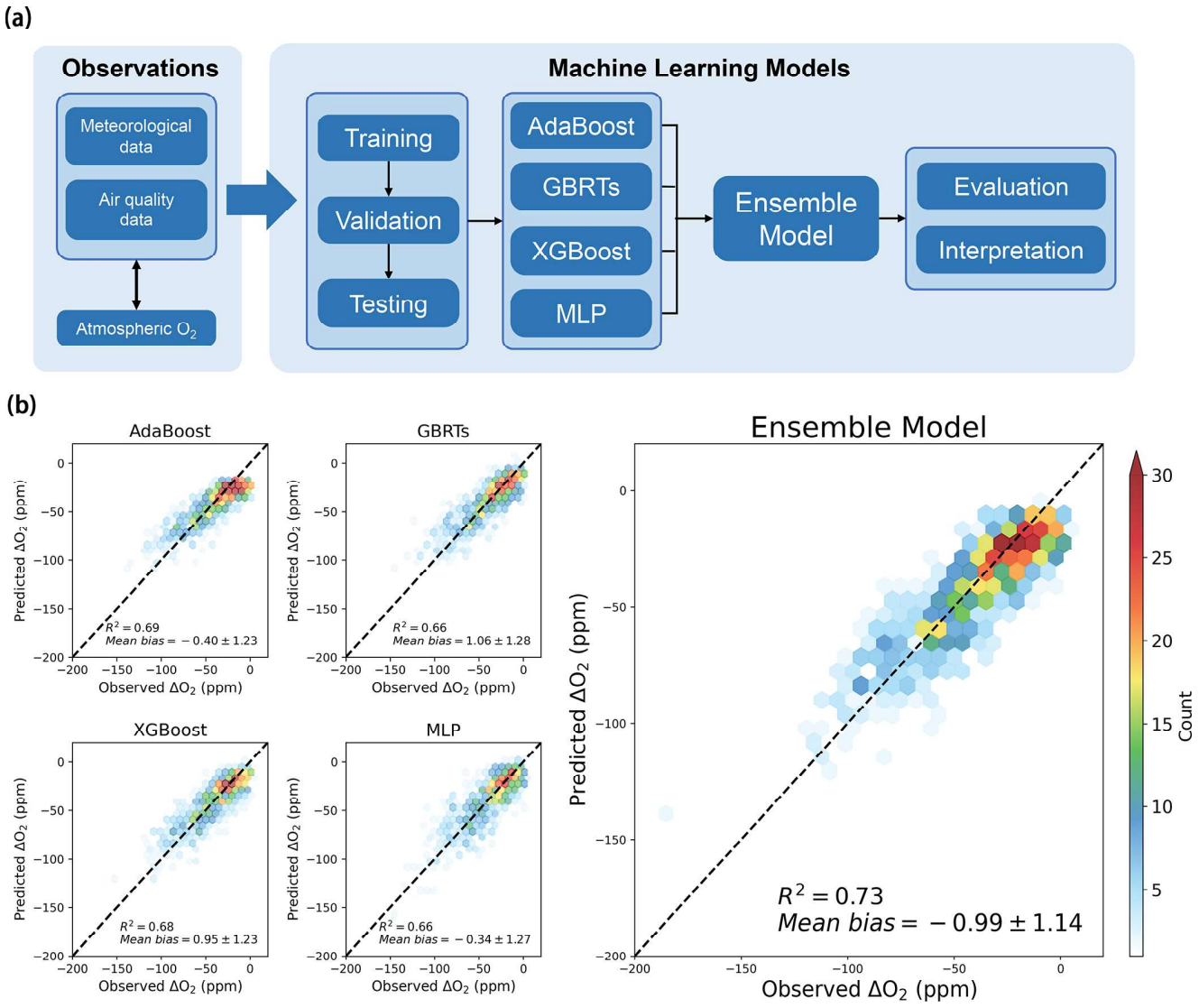


Figure 1. Model structure and performance evaluation. (a) Schematic diagram of the ML-based O_2 simulator. (b) Scatter plot of hourly observed vs simulated ΔO_2 in testing sets from AdaBoost, GBRTs, XGBoost, MLP, and the ensemble model (the final predictor), with black dashed lines denoting the 1/1 relationship.

Additive exPlanations (SHAP) framework, for instance, offers insights into the impact of a feature on model outcomes.[22] ML models combined with explainable tools have been extensively used in various aspects of air pollution modeling,[9,12,23,24] yet their application in reproducing variations in O_2 concentrations,[25] particularly in urban settings,[2] has been rather scant.

This study investigates the covariations between atmospheric O_2 and pollutants in Lanzhou, China. Lanzhou's distinct topography, with a narrow river valley surrounded by mountains, hinders atmospheric dispersion and prolongs pollutant retention in the valley.[26,27] The city is also affected by frequent dust storms originating from the nearby Gobi Desert.[28] With air pollutants from natural and anthropogenic sources, Lanzhou has been classified as one the most polluted cities in China, according to a World Health Organization (WHO) database.[29] We developed an ensemble ML-based O_2 simulator using concurrent observations of critical air pollutants and meteorological data. Employing the SHAP approach, we quantify the influence of natural and anthropogenic pollutants on atmospheric O_2 levels under various conditions. This study is one of the first to leverage an ML-based method to disentangle the complex contributions of natural and anthropogenic factors to atmospheric O_2 variations. It can potentially guide policy makers to effectively improve local air quality management, minimize potential health risk, and provide useful insights for the sustainable development of global urban regions.

MATERIALS AND METHODS

Data Sources. We trained the ML-based O_2 simulator using hourly air quality and meteorological parameters from January 2021 to December 2021. Hourly observations of air quality and meteorological parameters were obtained from the Lanzhou Atmospheric Components Monitoring Superstation (LACMS; 36.05°N, 103.87°E).[9,26,30] The air quality data include concentrations of particulate matter (TSP, PM_{10}, $PM_{2.5}$, and PM_1, reported in micrograms per cubic meter), particulate carbon (TC = EC + OC, reported in micrograms per cubic meter), and trace gases (NO_x, SO_2, and CO, reported in micrograms per cubic meter). Instead of directly

https://doi.org/10.1021/acs.estlett.3c00505
Environ. Sci. Technol. Lett. 2023, 10, 851−858

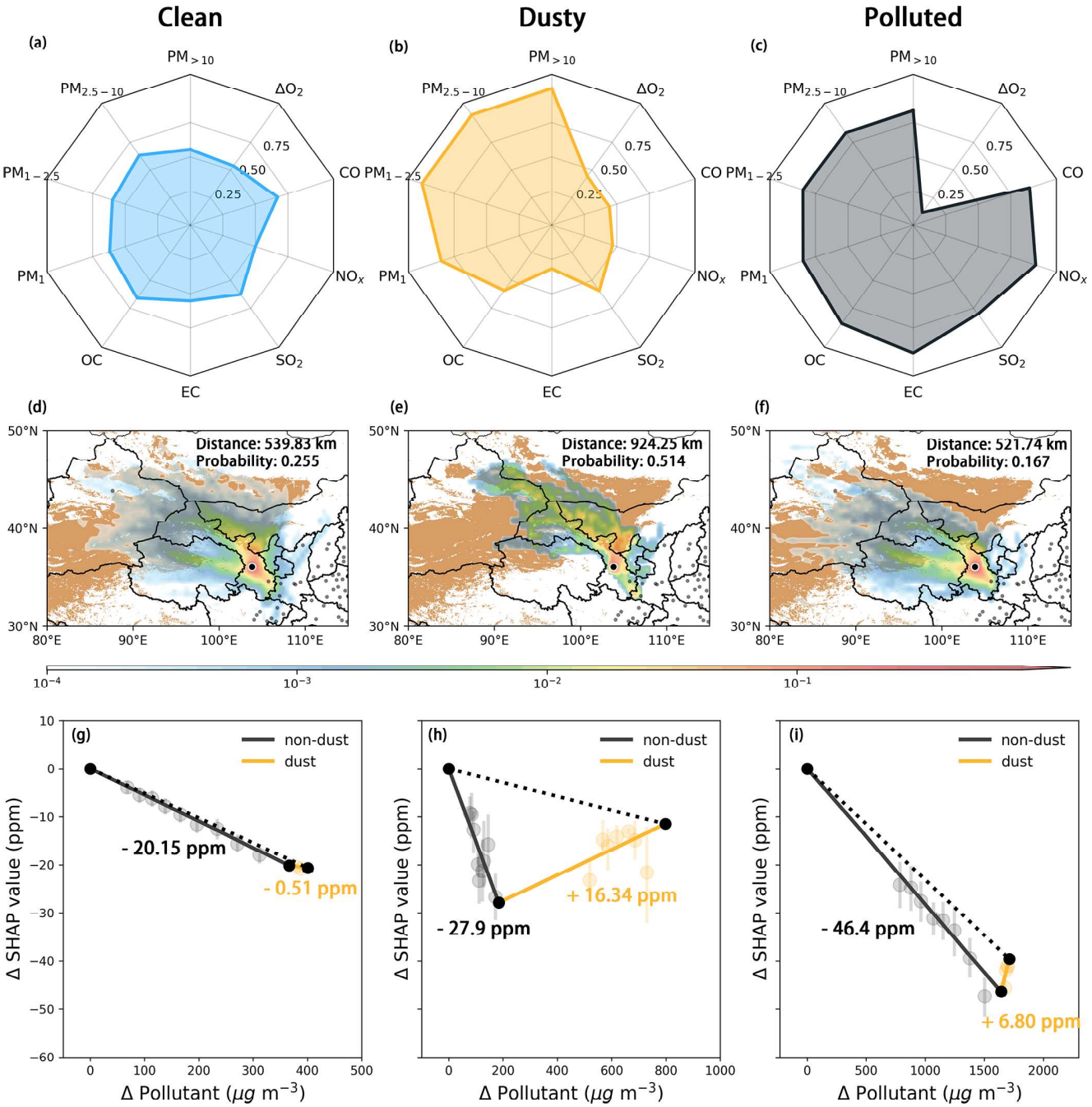


Figure 2. Covariation of atmospheric O_2 and pollutants under (a, d, and g) clean, (b, e, and h) dusty, and (c, f, and i) polluted conditions. (a–c) Radar plots of normalized ΔO_2 and pollutant levels under clean, dusty, and polluted conditions, respectively. (d–f) Probability density distributions of 48 h backward trajectories for Lanzhou under clean, dusty, and polluted conditions, respectively, at 500 m heights based on the HYSPLIT model simulation from January 1, 2021, to December 31, 2021. The color scale indicates the probability of air masses passing through a 0.5 × 0.5 grid cell. The orange shade indicates the potential dust source area (barren or sparsely vegetated), derived from the MODIS land cover data (MCD12C1 version 6.1).[39] The gray dots denote cities with populations exceeding 3 million. (g–i) Urban ΔO_2 budgets revealed by SHAP values, with solid gold lines denoting the impact from dust pollutants (i.e., $PM_{>10}$, $PM_{2.5-10}$, and $PM_{1-2.5}$) and solid black lines denoting the impact from nondust pollutants (i.e., NO_x, CO, SO_2, EC, OC, and PM_1). See Text S5.1 for details of panels g–i.

including TSP, PM_{10}, and $PM_{2.5}$, we introduced $PM_{1-2.5}$, $PM_{2.5-10}$, and $PM_{>10}$ into our model. These variables capture the distinctions between $PM_{2.5}$ and PM_1, PM_{10} and $PM_{2.5}$, and TSP and PM_{10}, respectively (see Figure S1). Observed meteorological variables include air temperature (degrees Celsius), atmospheric pressure (kilopascals), relative humidity (percent), and wind speed and direction (WS and WD in meters per seconds and degrees, respectively). The model training also utilizes the leaf area index (LAI) in square meters per square meter, total precipitation in meters, and boundary layer height (BLH) in meters from ERA5 reanalysis data.[31]

Atmospheric O_2 concentrations were measured on the rooftop of a building on the campus of Lanzhou University (~300 m from LACMS), using a gas chromograph (7890B,

https://doi.org/10.1021/acs.estlett.3c00505
Environ. Sci. Technol. Lett. 2023, 10, 851–858

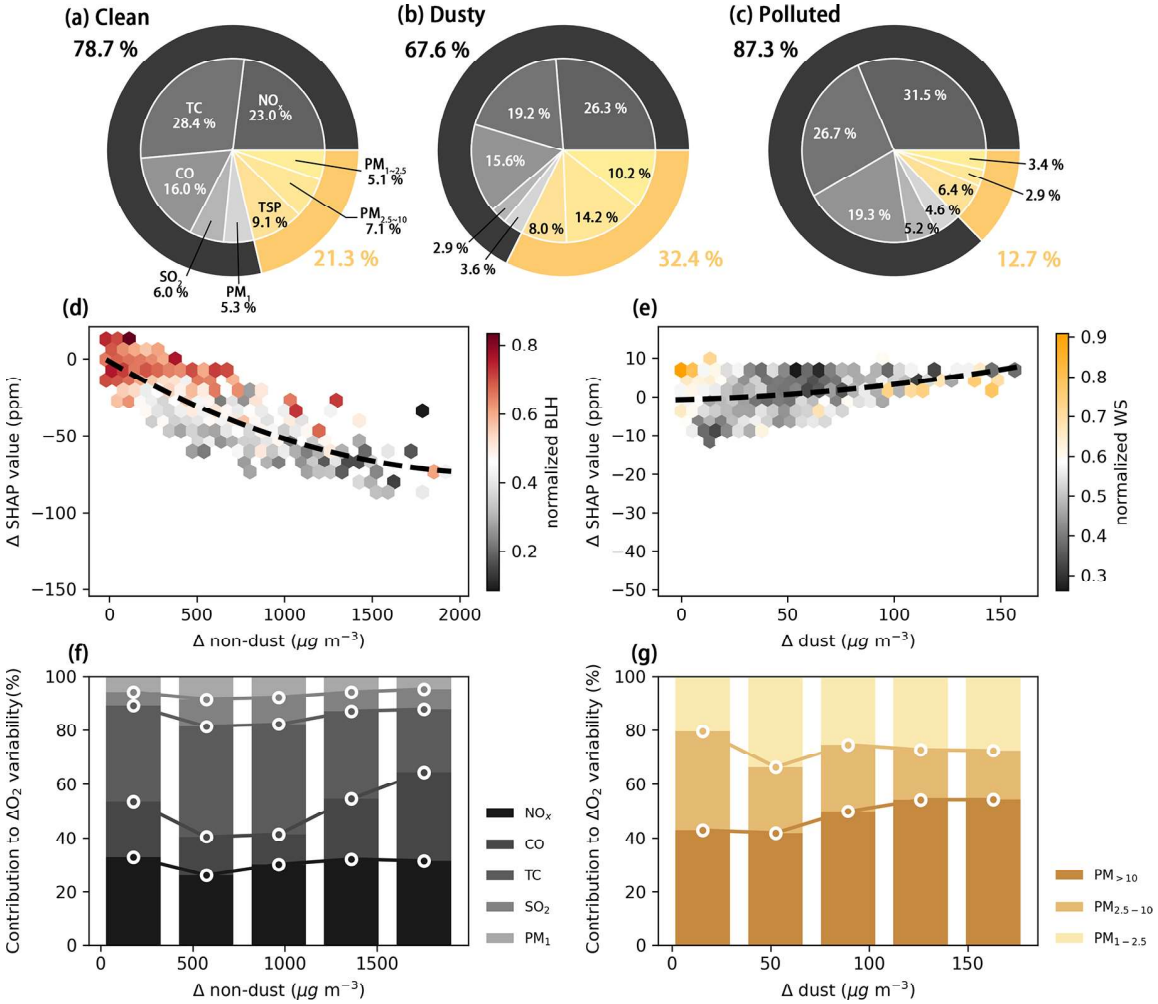


Figure 3. Influence of dust and nondust pollutants on ΔO_2. (a–c) Contributions of nondust and dust pollutants to pollutant-related ΔO_2 variability under clean, dusty, and polluted conditions, respectively. (d) SHAP dependence plot of nondust pollutants vs its SHAP value, with shadings denoting the normalized boundary layer height. (e) Same as panel d, but for dust pollutant, with shadings denoting the normalized wind speed. (f) Changes in the relative contributions of nondust pollutants to ΔO_2 variability with varying levels of pollution. (g) Same as panel f but for dust pollutants.

Agilent).[32] To weaken the dilution effect, it is customary to report O_2 fluctuations in terms of the relative deviation in the O_2/N_2 ratio from a reference value:

$$\Delta(O_2/N_2) = \left[\frac{(O_2/N_2)_{sample}}{(O_2/N_2)_{reference}} - 1\right] \times 10^6$$

where the subscripts sample and reference denote the sample air and reference gas, respectively. $\Delta(O_2/N_2)$ is expressed in per meg, where 1 per meg = 1×10^{-6}. To compare O_2 with pollutant concentrations, the ratio of 4.8 per meg/ppm is used to convert the observed changes in the O_2/N_2 ratio [$\Delta(O_2/N_2)$] to the ΔO_2 anomaly (ΔO_2) relative to a reference concentration that has been arbitrarily chosen. For more information about atmospheric O_2 measurements, see ref 2.

The observation sites are located in the Chengguan District, which is predominantly composed of commercial and residential buildings. The sites are located adjacent to Tianshui Road, one of the busiest thoroughfares in Lanzhou, and are thus strongly affected by traffic-related O_2 consumption and pollutant emission processes (see Figure S2). We also calculated the $\Delta O_2/\Delta CO_2$ ratios (1.33 ± 0.017), which identifies coal and oil as the major fossil fuel types in urban Lanzhou (see Text S1 and Figure S3).

ML-Based O_2 Simulator. We developed an ensemble ML-based O_2 simulator that integrates Adaptive Boosting (AdaBoost), Gradient Boosting Regression Trees (GBRTs), eXtreme Gradient Boosting (XGBoost), and Multilayer Perceptron (MLP) using a ridge regression model[33,34] (see Text S2 and Figure 1a). A detailed list of the predictors is given in Table S1. We validated and evaluated the model performance using a 10-fold cross-validation based on coefficients of determination (R^2) and mean bias, as shown in Figure 1b. When the four models were combined, the performance of the ultimate estimator was improved, leading to an increase in R^2 to 0.73. Using the SHAP approach (see Text S2.3 for details), we can evaluate the influence of each variable on each predicted O_2 sample.[22] This allows us to infer the exact process responsible for the variation in the O_2 concentration for each sample.

RESULTS AND DISCUSSION

Covariation of O_2 and Critical Pollutants under Distinctive Conditions. To extract the typical covariation

https://doi.org/10.1021/acs.estlett.3c00505
Environ. Sci. Technol. Lett. 2023, 10, 851−858

pattern between O_2 and pollutants, we applied K-means clustering on the observed ΔO_2 and pollutants (see Text S3 for details). Panels a–c of Figure 2 illustrate the clustering results in radar charts, with concentrations normalized between 0 and 1. Samples in the Clean cluster (Figure 2a) exhibited the highest O_2 content and the lowest concentration for all pollutants. The Dusty cluster (Figure 2b) is characterized by an extraordinarily high concentration of particulate pollutants ($PM_{>10}$, $PM_{2.5-10}$, and $PM_{1-2.5}$), indicating dust pollution of a natural origin. The Polluted clusters (Figure 2c) had the lowest levels of O_2 and the highest concentrations of anthropogenic pollutants. We find $PM_{>10}$, $PM_{2.5-10}$, and $PM_{1-2.5}$ have the highest weights in determining the Dusty cluster (see Text S3.3 and Figure S5 for details), and ΔO_2 responds differently to the two groups of pollutants [dust group: $PM_{>10}$, $PM_{2.5-10}$, and $PM_{1-2.5}$; nondust group: NO_x, CO, PM_1, TC, and SO_2 (see Figure S7)]. Consequently, we divide the pollutants into dust and nondust pollutants for subsequent analysis.

Lanzhou is situated near the Gobi Desert and is therefore vulnerable to dust storms that arise when cold fronts pass through. This usually facilitates the long-range transport of air masses and dust of natural origin. We performed 48 h backward trajectory simulations for Lanzhou to evaluate the impact of atmospheric transport on O_2 and dust concentrations in Lanzhou (Figure 2d–f; see Text S4 for details). We found that air masses arriving within the atmospheric boundary layer in Lanzhou, particularly those in dusty clusters, are more likely to experience long-range atmospheric transport from northern and northwestern desert sources (probability of 0.514) than to originate locally (mean trajectory distance of 924.25 km). Panels g–i of Figure 2 demonstrate how dust pollutants and anthropogenic pollutants can alter the ΔO_2 values under different pollution patterns. The anthropogenic pollutants can result in decreases in ΔO_2 of 20.2, 27.9, and 46.4 ppm under clean, dusty, and polluted conditions, respectively, while dust pollutants can lead to increases in ΔO_2 of 16.3 and 6.8 ppm under dusty and polluted conditions, respectively. This stands in sharp contrast to what is commonly observed and anticipated in other cities almost untouched by dust storms, where an increase in particulate matter (PM) concentration typically signals anthropogenic pollution that leads to increased CO_2 levels[8,35−37] (and decreased O_2 levels as expected). Lanzhou experiences PM pollution from both natural and anthropogenic sources, via long-range transport, and anthropogenic sources, through primary emissions and secondary formations.[27,38] The positive contribution from dust pollutants to the O_2 levels in Lanzhou indicates that atmospheric transport may introduce a volume of O_2 greater than that consumed by primary emissions and secondary formations. This finding underscores the potential counterbalancing effect of atmospheric transport that offsets the negative effects of primary emissions and secondary formations, which directly and indirectly decrease the level of O_2 in the atmosphere. It is worth mentioning that discussions concerning the variability of O_2 resulting from reactions between O_2 and pollutants are beyond the scope of this work, because we focus on the variation of O_2 on the parts per million scale (see Text S2.3).

Mechanisms Driving O_2 Variability in an Urban Atmosphere. We further quantified the contribution of nondust and dust pollutants to pollutant-related ΔO_2 variability, as shown in panels a–c of Figure 3 (see Figures S11–S13 and Texts S5.2 and S5.3 for SHAP dependency plots of other variables). We find that the emission of nondust pollutants is the primary determinant of ΔO_2 variability, with contributions of 78.7%, 67.6%, and 87.3% under clean, dusty, and polluted conditions, respectively. Among the pollutants, NO_x, TC, and CO are the leading contributors, and their contributions increase significantly under polluted conditions, indicating the emission of these pollutants as the primary O_2 sinks in urban Lanzhou. The contribution of nondust pollutants also exhibits a diurnal cycle, particularly under polluted conditions. O_2 consumption is significantly higher during working hours, as illustrated in Figure S14.

Panels d and e of Figure 3 demonstrate the response of ΔO_2 to increases in the concentrations of nondust and dust pollutants, as well as the interactions between BLH and WS. Higher concentrations of nondust pollutants are associated with decreases in ΔO_2, which is typically accompanied by a decrease in BLH. Moreover, a decrease in BLH can cause a decrease in ΔO_2, even at the same pollutant levels. In contrast, increased levels of dust pollutants can result in an increased ΔO_2. In addition to the monotonic relationships between ΔO_2 and pollutants, we also found that the slope of the response curve is sensitive to varying pollutant levels (see Figure S16 for details). For nondusty pollutants, the reduction of ΔO_2 tends to decouple from the accumulation of pollutants, suggesting that the same amount of O_2 consumed can generate more pollutants in a more polluted atmosphere. A previous observational study[40] during the COVID-19 lockdown periods reported a similar decoupling effect; the authors found that the declines in classic air pollutants (particularly NO_x) were significantly larger than the associated decrease in co-emitted CO_2. They attributed this decoupling effect to the diverging impact of lockdown policies on urban combustion sources, i.e., more significant reductions in traffic-related fossil fuel combustions than residential, commercial, and public sectors. In addition to changes in emission patterns, the distinct atmospheric lifetimes of CO_2 and pollutants like NO_x indicate differences in their ability to react and travel in the atmosphere, leading to varying degrees of concentration changes in response to emission alterations. Therefore, in this study, the observed decoupling should not be solely attributable to emissions, as chemical processes and atmospheric transport also play significant roles[8,41] (see Texts S6.1 and S6.2).

We further investigated the major emission sources that drive O_2 consumption under different pollution levels and found increased contributions from CO and NO_x and decreased contributions from EC, OC, SO_2, and PM_1 (Figure 3f). Though CO and NO_x are co-emitted from broadly similar sources, the tropospheric lifetime of NO_x is shorter.[10] The chemical loss of atmospheric NO_x and SO_2, which do not cause ΔO_2 variability at the parts per million level, resulted in weakening contributions.[8] This explains why the growth in contribution from CO to ΔO_2 is significantly larger as the pollutant accumulates. Compared with fossil fuel combustion, which directly removes O_2 and releases pollutants, chemical loss processes could consume pollutants in the atmosphere without significant O_2 depletion. This would result in a steeper response curve under heavily polluted conditions, which is inconsistent with the results depicted in Figure 3d.

We have discovered that emission and chemical processes alone are insufficient to fully account for the decoupling between ΔO_2 and nondust pollutants (see Text S6.1 for details). This decoupling can be sensibly explained only by the distinct diffusion and transport of O_2 and pollutants in the

https://doi.org/10.1021/acs.estlett.3c00505
Environ. Sci. Technol. Lett. 2023, 10, 851−858

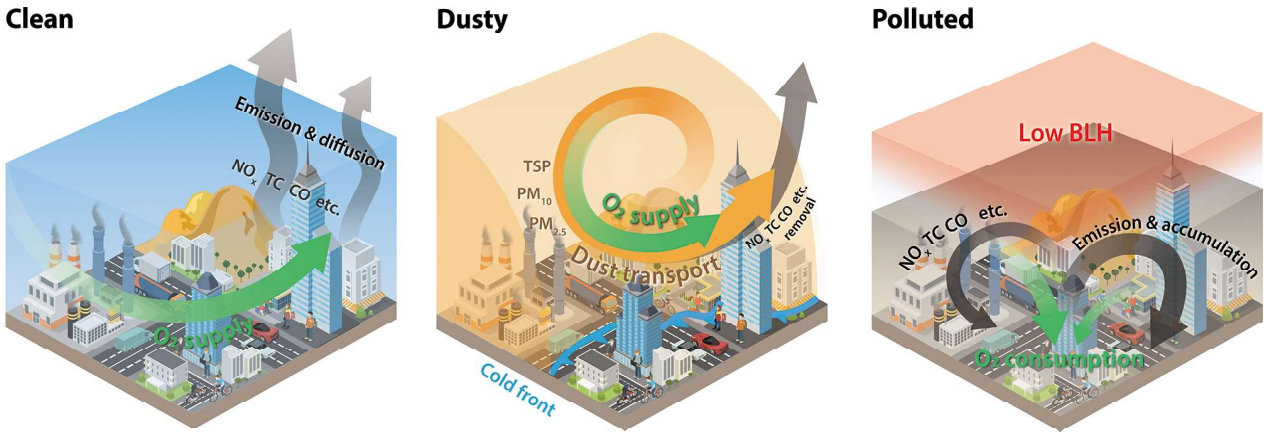


Figure 4. Schematic diagram. Three typical types of pollution patterns (clean days, dusty days, and polluted days) with different levels of O_2. The black arrows show the air flows that transport the anthropogenic pollutants (NO_x, TC, CO, etc.). The green arrows represent the air flow of atmospheric O_2, and the orange arrows denote the transport of dust aerosols from the dust source region (Gobi Desert). During clean days, good atmospheric ventilation supplies enough O_2 to the urban region and diffuses urban pollutants. Dusty days feature cold front activities that bring both fresh air with higher O_2 content and dust from the natural backgrounds and blow the urban pollutants away. During polluted days, near-surface temperature inversion and downdrafts trap pollutants within the urban boundary layer, which leads to poor air quality and a decrease in the level of O_2.

atmosphere. In the near surface, pollutant emission is characterized by upward fluxes, while the emission of O_2 is consumed via downward fluxes. High pollutant concentrations often coincide with stable atmospheric conditions, marked by a reduced lapse rate or temperature inversion. Such conditions could suppress the upward dispersion of pollutants while promoting the downward transport of O_2. This divergence in vertical transport likely causes slower O_2 declines alongside enhanced pollutant accumulation in the lower atmosphere. These processes resemble gravitational fractionation, which has been documented in the upper troposphere and stratosphere[42] but rarely studied in the boundary layer under stagnant weather (see Text S6.3). Further research using field measurements and model simulations will be important to validate the proposed mechanisms driving divergent vertical transport of O_2 and pollutants within the near surface.

For dust pollutants, ΔO_2 tends to increase faster as concentrations of dust pollutants increase, particularly under high-wind speed conditions, as evidenced by the steeper response curve shown in Figure 3e. We discovered that particles of larger aerodynamic diameters had an increasingly greater contribution to ΔO_2 as the concentrations of the dust pollutants increased. Unlike the nondust pollutants discussed above, atmospheric transportation processes are the primary drivers of the variability in dust pollutants of natural origin. A higher concentration of dust pollutants, accompanied by an increase in wind speed, suggests enhanced atmospheric transport, leading to a concurrent increase in O_2 and dust concentrations (see Figure S11). In this situation, an air mass can exhibit greater atmospheric diffusivity than dust, thus resulting in a quicker ΔO_2 increase.

The nonlinear responses of ΔO_2 to varying levels of pollutants illustrate the influence of atmospheric dispersion on elucidating and interpreting the emission and/or consumption and atmospheric transport processes derived from observations. Atmospheric components with varying diffusion characteristics and directions of fluxes can result in discrepancies between the observation-derived and actual emission ratios, oxidative (consumption) ratios, etc. Accounting for non-uniform diffusion enhances the validation and calibration of urban emission inventories, while also constraining biological processes.[43]

The mechanisms that drive urban variation of the O_2 for Lanzhou during clean, dusty, and polluted episodes are illustrated in Figure 4. O_2 levels of urban areas respond differently to dust and nondust pollutants, and the roles of atmospheric transport and diffusion may further complicate the relationship between O_2 and pollutants. Our study also indicates that dust pollutants could be an indicator for an increase in the level of O_2 in Lanzhou. This study is among the first to use a machine learning-based approach to unravel the intricate impacts of both natural and anthropogenic elements on atmospheric O_2 levels and emphasizes the pollution characteristic of urban areas adjacent to dust sources. The ML-based O_2 simulator can potentially serve as a useful tool for exploring the complex emission, chemical, and transport processes in the urban atmosphere, but it is necessary to exercise caution when interpreting the results of our research due to certain limitations (see Text S7 for details).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.estlett.3c00505.

Texts S1−S7, Figures S1−S19, and Tables S1 and S2 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jianping Huang − *Collaborative Innovation Center for Western Ecological Safety, Lanzhou University, Lanzhou 730000, China;* orcid.org/0000-0003-2845-797X; Email: hjp@lzu.edu.cn

Authors

Xiaoyue Liu − *Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China;* orcid.org/0000-0001-8074-3362

https://doi.org/10.1021/acs.estlett.3c00505
Environ. Sci. Technol. Lett. 2023, 10, 851−858

Li Wang − *Collaborative Innovation Center for Western Ecological Safety, Lanzhou University, Lanzhou 730000, China*
Yongqi Wang − *Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China*
Changyu Li − *Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China*
Lei Ding − *Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China*
Xinbo Lian − *Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China*
Jinsen Shi − *Collaborative Innovation Center for Western Ecological Safety, Lanzhou University, Lanzhou 730000, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.estlett.3c00505

Author Contributions

X. Liu, J.H., and L.W. designed the study and contributed to the ideas, data analysis, and manuscript writing. L.W. contributed to the data processing, interpretation, and manuscript writing. All of the authors contributed to the discussion and interpretation of the manuscript. All of the authors reviewed the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was jointly supported by the National Science Foundation of China (41888101, 41991231, and 91937302), the China 111 project (B13045), the Natural Science Foundation of Gansu Province of China (21JR7RA528), and the Fundamental Research Funds for the Central Universities (lzujbky-2022-kb10, lzujbky-2021-64).

■ REFERENCES

(1) Gurney, K. R.; Liang, J.; Roest, G.; Song, Y.; Mueller, K.; Lauvaux, T. Under-Reporting of Greenhouse Gas Emissions in U.S. Cities. *Nat. Commun.* **2021**, *12* (1), 553.
(2) Liu, X.; Huang, J.; Wang, L.; Lian, X.; Li, C.; Ding, L.; Wei, Y.; Chen, S.; Wang, Y.; Li, S.; Shi, J. Urban Respiration" Revealed by Atmospheric O $_2$ Measurements in an Industrial Metropolis. *Environ. Sci. Technol.* **2023**, *57* (6), 2286−2296.
(3) Yan, Y.; Brook, E. J.; Kurbatov, A. V.; Severinghaus, J. P.; Higgins, J. A. Ice Core Evidence for Atmospheric Oxygen Decline since the Mid-Pleistocene Transition. *Sci. Adv.* **2021**, *7* (51), eabj9341 DOI: 10.1126/sciadv.abj9341.
(4) Ishidoya, S.; Tsuboi, K.; Kondo, H.; Ishijima, K.; Aoki, N.; Matsueda, H.; Saito, K. Measurement Report: Method for Evaluating CO_2 Emission from a Cement Plant by Atmosphere O_2/N_2 and CO_2 Measurements and Its Applicability to the Detection of CO_2 Capture Signals. *Atmos. Chem. Phys.* **2022**, DOI: 10.5194/acp-2022-570.
(5) Steinbach, J.; Gerbig, C.; Rödenbeck, C.; Karstens, U.; Minejima, C.; Mukai, H. The CO2 Release and Oxygen Uptake from Fossil Fuel Emission Estimate (COFFEE) Dataset: Effects from Varying Oxidative Ratios. *Atmos. Chem. Phys.* **2011**, *11* (14), 6855−6870.
(6) Keeling, R. F.; Powell, F. L.; Shaffer, G.; Robbins, P. A.; Simonson, T. S. Impacts of Changes in Atmospheric O_2 on Human Physiology. Is There a Basis for Concern? *Front. Physiol.* **2021**, *12* (March), 571137 DOI: 10.3389/fphys.2021.571137.
(7) Wei, Y.; Wu, J.; Huang, J.; Liu, X.; Han, D.; An, L.; Yu, H.; Huang, J. Declining Oxygen Level as an Emerging Concern to Global Cities. *Environ. Sci. Technol.* **2021**, *55* (12), 7808−7817.
(8) Bares, R.; Lin, J. C.; Hoch, S. W.; Baasandorj, M.; Mendoza, D. L.; Fasoli, B.; Mitchell, L.; Catharine, D.; Stephens, B. B. The Wintertime Covariation of CO 2 and Criteria Pollutants in an Urban Valley of the Western United States. *Journal of Geophysical Research: Atmospheres* **2018**, *123* (5), 2684−2703.
(9) Wang, L.; Zhao, Y.; Shi, J.; Ma, J.; Liu, X.; Han, D.; Gao, H.; Huang, T. Predicting Ozone Formation in Petrochemical Industrialized Lanzhou City by Interpretable Ensemble Machine Learning. *Environ. Pollut.* **2023**, *318*, 120798.
(10) Wu, D.; Liu, J.; Wennberg, P. O.; Palmer, P. I.; Nelson, R. R.; Kiel, M.; Eldering, A. Towards Sector-Based Attribution Using Intra-City Variations in Satellite-Based Emission Ratios between CO $_2$ and CO. *Atmos. Chem. Phys.* **2022**, *22* (22), 14547−14570.
(11) Song, C.; Liu, B.; Cheng, K.; Cole, M. A.; Dai, Q.; Elliott, R. J. R.; Shi, Z. Attribution of Air Quality Benefits to Clean Winter Heating Polices in China: Combining Machine Learning with Causal Inference. *Environ. Sci. Technol.* **2023**, DOI: 10.1021/acs.est.2c06800.
(12) Wen, Y.; Zhou, Z.; Zhang, S.; Wallington, T. J.; Shen, W.; Tan, Q.; Deng, Y.; Wu, Y. Urban−Rural Disparities in Air Quality Responses to Traffic Changes in a Megacity of China Revealed Using Machine Learning. *Environ. Sci. Technol. Lett.* **2022**, *9* (7), 592−598.
(13) Chen, B.; Song, Z.; Pan, F.; Huang, Y. Obtaining Vertical Distribution of PM2.5 from CALIOP Data and Machine Learning Algorithms. *Sci. Total Environ.* **2022**, *805*, 150338.
(14) Cheng, M.; Fang, F.; Navon, I. M.; Zheng, J.; Tang, X.; Zhu, J.; Pain, C. Spatio-Temporal Hourly and Daily Ozone Forecasting in China Using a Hybrid Machine Learning Model: Autoencoder and Generative Adversarial Networks. *Journal of Advances in Modeling Earth Systems* **2022**, *14* (3), No. e2021MS002806.
(15) Han, B.; Yao, T.; Li, G.; Song, Y.; Zhang, Y.; Dai, Q.; Yu, J. Marginal Reduction in Surface NO2 Attributable to Airport Shutdown: A Machine Learning Regression-Based Approach. *Environmental Research* **2022**, *214*, 114117.
(16) 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. From COVID-19 to Future Electrification: Assessing Traffic Impacts on Air Quality by a Machine-Learning Model. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (26), No. e2102705118.
(17) Mansfield, L. A.; Nowack, P. J.; Kasoar, M.; Everitt, R. G.; Collins, W. J.; Voulgarakis, A. Predicting Global Patterns of Long-Term Climate Change from Short-Term Simulations Using Machine Learning. *npj Climate and Atmospheric Science* **2020**, *3* (1), 1−9.
(18) Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat, N. Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature* **2019**, *566* (7743), 195−204.
(19) Zhong, S.; Zhang, K.; Bagheri, M.; Burken, J. G.; Gu, A.; Li, B.; Ma, X.; Marrone, B. L.; Ren, Z. J.; Schrier, J.; Shi, W.; Tan, H.; Wang, T.; Wang, X.; Wong, B. M.; Xiao, X.; Yu, X.; Zhu, J.-J.; Zhang, H. Machine Learning: New Ideas and Tools in Environmental Science and Engineering. *Environ. Sci. Technol.* **2021**, *55* (19), 12741−12754.
(20) Lyu, B.; Hu, Y.; Zhang, W.; Du, Y.; Luo, B.; Sun, X.; Sun, Z.; Deng, Z.; Wang, X.; Liu, J.; Wang, X.; Russell, A. G. Fusion Method Combining Ground-Level Observations with Chemical Transport Model Predictions Using an Ensemble Deep Learning Framework: Application in China to Estimate Spatiotemporally-Resolved PM2.5 Exposure Fields in 2014−2017. *Environ. Sci. Technol.* **2019**, *53* (13), 7306−7315.
(21) Requia, W. J.; Di, Q.; Silvern, R.; Kelly, J. T.; Koutrakis, P.; Mickley, L. J.; Sulprizio, M. P.; Amini, H.; Shi, L.; Schwartz, J. An Ensemble Learning Approach for Estimating High Spatiotemporal Resolution of Ground-Level Ozone in the Contiguous United States. *Environ. Sci. Technol.* **2020**, *54* (18), 11037−11047.
(22) Lundberg, S. M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing*

Systems 30; Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc., 2017; pp 4765−4774.
(23) Hou, L.; Dai, Q.; Song, C.; Liu, B.; Guo, F.; Dai, T.; Li, L.; Liu, B.; Bi, X.; Zhang, Y.; Feng, Y. Revealing Drivers of Haze Pollution by Explainable Machine Learning. *Environ. Sci. Technol. Lett.* **2022**, *9* (2), 112−119.
(24) Song, C.; Becagli, S.; Beddows, D. C. S.; Brean, J.; Browse, J.; Dai, Q.; Dall'Osto, M.; Ferracci, V.; Harrison, R. M.; Harris, N.; Li, W.; Jones, A. E.; Kirchgäßner, A.; Kramawijaya, A. G.; Kurganskiy, A.; Lupi, A.; Mazzola, M.; Severi, M.; Traversi, R.; Shi, Z. Understanding Sources and Drivers of Size-Resolved Aerosol in the High Arctic Islands of Svalbard Using a Receptor Model Coupled with Machine Learning. *Environ. Sci. Technol.* **2022**, *56* (16), 11189−11198.
(25) Pickers, P. A.; Manning, A. C.; Le Quéré, C.; Forster, G. L.; Luijkx, I. T.; Gerbig, C.; Fleming, L. S.; Sturges, W. T. Novel Quantification of Regional Fossil Fuel CO2 Reductions during COVID-19 Lockdowns Using Atmospheric Oxygen Measurements. *Science Advances* **2022**, *8* (16), No. eabl9250.
(26) Du, T.; Wang, M.; Guan, X.; Zhang, M.; Zeng, H.; Chang, Y.; Zhang, L.; Tian, P.; Shi, J.; Tang, C. Characteristics and Formation Mechanisms of Winter Particulate Pollution in Lanzhou, Northwest China. *Journal of Geophysical Research: Atmospheres* **2020**, *125* (18), 1−17.
(27) Chen, S.; Zhang, X.; Lin, J.; Huang, J.; Zhao, D.; Yuan, T.; Huang, K.; Luo, Y.; Jia, Z.; Zang, Z.; Qiu, Y.; Xie, L. Fugitive Road Dust PM2.5 Emissions and Their Potential Health Impacts. *Environ. Sci. Technol.* **2019**, *53* (14), 8455−8465.
(28) Filonchyk, M. Characteristics of the Severe March 2021 Gobi Desert Dust Storm and Its Impact on Air Pollution in China. *Chemosphere* **2022**, *287*, 132219.
(29) World Health Organization (WHO). Air quality database: Update 2011. https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database/201web1 (accessed 2023-02-19).
(30) Wang, M.; Tian, P.; Wang, L.; Yu, Z.; Du, T.; Chen, Q.; Guan, X.; Guo, Y.; Zhang, M.; Tang, C.; Chang, Y.; Shi, J.; Liang, J.; Cao, X.; Zhang, L. High Contribution of Vehicle Emissions to Fine Particulate Pollutions in Lanzhou, Northwest China Based on High-Resolution Online Data Source Appointment. *Sci. Total Environ.* **2021**, *798*, 149310.
(31) Hersbach, H.; Bell, B.; Berrisford, P.; Biavati, G.; Horányi, A.; Muñoz Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Rozum, I.; Schepers, D.; Simmons, A.; Soci, C.; Dee, D.; Thépaut, J.-N. ERA5 hourly data on single levels from 1940 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. DOI: 10.24381/cds.adbb2d47 (accessed 2023-05-20).
(32) Tohjima, Y. Method for Measuring Changes in the Atmospheric O2/N2 Ratio by a Gas Chromatograph Equipped with a Thermal Conductivity Detector. *Journal of Geophysical Research Atmospheres* **2000**, *105* (D11), 14575−14584.
(33) Kidwell, J. S.; Brown, L. H. Ridge Regression as a Technique for Analyzing Models with Multicollinearity. *Journal of Marriage and Family* **1982**, *44* (2), 287−299.
(34) Mohammadi, S. A. Test of Harmful Multicollinearity: A Generalized Ridge Regression Approach. *Communications in Statistics - Theory and Methods* **2022**, *51* (3), 724−743.
(35) Wei, C.; Wang, M.; Fu, Q.; Dai, C.; Huang, R.; Bao, Q. Temporal Characteristics of Greenhouse Gases (CO2 and CH4) in the Megacity Shanghai, China: Association with Air Pollutants and Meteorological Conditions. *Atmos. Res.* **2020**, *235*, 104759.
(36) Elsunousi, A. A. M.; Sevik, H.; Cetin, M.; Ozel, H. B.; Ozel, H. U. Periodical and Regional Change of Particulate Matter and CO2 Concentration in Misurata. *Environ. Monit Assess* **2021**, *193* (11), 707.
(37) Jacobson, M. Z. Enhancement of Local Air Pollution by Urban CO2 Domes. *Environ. Sci. Technol.* **2010**, *44* (7), 2497−2502.
(38) Cheng, B.; Ma, Y.; Li, H.; Feng, F.; Zhang, Y.; Qin, P. Water-Soluble Ions and Source Apportionment of PM2.5 Depending on Synoptic Weather Patterns in an Urban Environment in Spring Dust Season. *Sci. Rep* **2022**, *12* (1), 21953.
(39) Friedl, M.; Sulla-Menashe, D. *MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG*, ver. 061; 2022.
(40) Lamprecht, C.; Graus, M.; Striednig, M.; Stichaner, M.; Karl, T. Decoupling of Urban CO_2 and Air Pollutant Emission Reductions during the European SARS-CoV-2 Lockdown. *Atmos. Chem. Phys.* **2021**, *21* (4), 3091−3102.
(41) Laughner, J. L.; Neu, J. L.; Schimel, D.; Wennberg, P. O.; Barsanti, K.; Bowman, K. W.; Chatterjee, A.; Croes, B. E.; Fitzmaurice, H. L.; Henze, D. K.; Kim, J.; Kort, E. A.; Liu, Z.; Miyazaki, K.; Turner, A. J.; Anenberg, S.; Avise, J.; Cao, H.; Crisp, D.; De Gouw, J.; Eldering, A.; Fyfe, J. C.; Goldberg, D. L.; Gurney, K. R.; Hasheminassab, S.; Hopkins, F.; Ivey, C. E.; Jones, D. B. A.; Liu, J.; Lovenduski, N. S.; Martin, R. V.; McKinley, G. A.; Ott, L.; Poulter, B.; Ru, M.; Sander, S. P.; Swart, N.; Yung, Y. L.; Zeng, Z.-C. the rest of the Keck Institute for Space Studies "COVID-19: Identifying Unique Opportunities for Earth System Science" study team. Societal Shifts Due to COVID-19 Reveal Large-Scale Complexities and Feedbacks between Atmospheric Chemistry and Climate Change. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (46), No. e2109481118.
(42) Ishidoya, S.; Sugawara, S.; Inai, Y.; Morimoto, S.; Honda, H.; Ikeda, C.; Hashida, G.; Machida, T.; Tomikawa, Y.; Toyoda, S.; Goto, D.; Aoki, S.; Nakazawa, T. Gravitational Separation of the Stratospheric Air over Syowa, Antarctica and Its Connection with Meteorological Fields. *Atmos. Sci. Lett.* **2018**, *19* (11), No. e857.
(43) Faassen, K. A. P.; Nguyen, L. N. T.; Broekema, E. R.; Kers, B. A. M.; Mammarella, I.; Vesala, T.; Pickers, P. A.; Manning, A. C.; Vilà-Guerau de Arellano, J.; Meijer, H. A. J.; Peters, W.; Luijkx, I. T. Diurnal Variability of Atmospheric O_2, CO_2, and Their Exchange Ratio above a Boreal Forest in Southern Finland. *Atmos. Chem. Phys.* **2023**, *23* (2), 851−876.

https://doi.org/10.1021/acs.estlett.3c00505
Environ. Sci. Technol. Lett. 2023, 10, 851−858