Environmental Science & Technology

pubs.acs.org/est

Article

Mechanism-Constrained Learning for Regime-Aware Urban Ozone Prediction and Interpretation

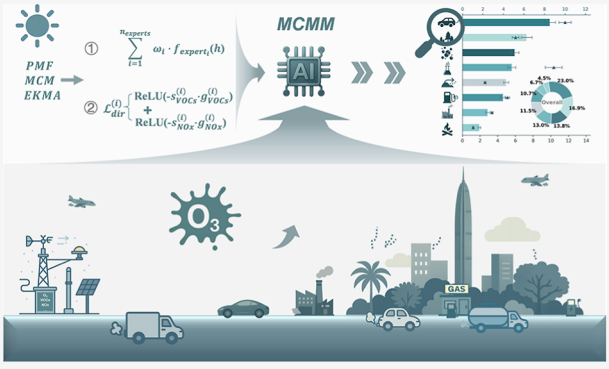
Yong Cheng, Yan Peng, Xiao-Feng Huang, Hui Zeng, and Ling-Yan He*

Cite This: https://doi.org/10.1021/acs.est.6c12789

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Accurately characterizing source-related and precursor factors associated with surface ozone (O_3) pollution remains a central challenge for air quality management. Data-driven models can achieve strong predictive performance, but mechanistically consistent interpretation remains challenging. Here we develop a mechanism-constrained diagnostic framework (MCDF) integrating observations, source-related factors of volatile organic compounds (VOCs) resolved by positive matrix factorization, and scenario-specific constraints from the empirical kinetic modeling approach (EKMA). Its core mechanism-constrained multitask model (MCMM) predicts O_3 while jointly learning precursor-control regime and signed ridge distance, with a direction-consistency loss guiding precursor responses toward the corresponding photochemical state. MCMM achieved stable repeated-split O_3 prediction performance and produced finite precursor responses more consistent with EKMA than an unconstrained model. Across 10−30% increases in nitrogen oxides (NOx), MCMM predicted negative O_3 responses in 90.1−92.3% of initially VOCs-limited cases. Applied to observations from a representative O_3-polluted megacity case in China, MCDF revealed regime-dependent changes in model-attributed precursor and source-related associations. Overall, MCDF integrates scenario-specific photochemical constraints with data-driven O_3 prediction, bridges predictive performance and mechanistic interpretation, and provides an interpretable, regime-aware framework for diagnosing precursor sensitivities and source-related variability in complex urban O_3 pollution.



KEYWORDS: *ozone pollution, machine learning, multitask learning, atmospheric chemistry, photochemical sensitivity, model interpretation*

1. INTRODUCTION

Humans increasingly rely on their capacity to both predict and understand air pollution in order to sustain livable cities.[1,2] Among the major pollutants, surface ozone (O_3) has evolved from a secondary concern to a central constraint on urban air quality management, especially in regions where particulate matter has declined even as photochemical pollution continues to intensify.[3−5] Elevated O_3 adversely affects human health, crop productivity, and climate,[6−9] yet its formation is governed by strongly nonlinear chemistry involving volatile organic compounds (VOCs) and nitrogen oxides (NOx) under solar radiation.[10] The resulting O_3 response depends not only on precursor abundance but also on meteorology, boundary-layer dynamics, atmospheric processing, and transport.[11−14] Moreover, the sensitivity of O_3 to NOx and VOCs can shift across photochemical regimes, so similar precursor changes may produce different O_3 responses under different atmospheric states.[15−17] These coupled chemical and environmental controls make it difficult to achieve both accurate prediction and physically meaningful interpretation of observed precursor−O_3 relationships, particularly as urban emission patterns continue to evolve.[18−23]

Although understanding of O_3 formation mechanisms has advanced substantially,[24−27] current approaches for diagnosing O_3 formation and source-related variability still exhibit notable limitations.[28,29] Mechanism-based models provide strong chemical interpretability and can diagnose VOC- or NOx-sensitive regimes, but their results can be sensitive to emission inventories, boundary conditions, transport representation, and the spatial and temporal resolution of input data.[30] In parallel, rapidly developing data-driven machine-learning methods have demonstrated strong performance in pollution prediction and model adaptability,[31−38] with recent studies combining positive matrix factorization (PMF) and interpretable models to relate

Received: August 18, 2026
Revised: September 3, 2026
Accepted: September 4, 2026

receptor-resolved VOC source-related factors to observed O_3 variability.[39−41] However, without explicit physical or chemical constraints, purely data-driven interpretation may remain dominated by statistical associations and provide limited information on chemically differentiated precursor-response behavior. Integrating mechanistic information into data-driven models while retaining their predictive capability has therefore become a key challenge for contemporary O_3 diagnostics and, more broadly, for Earth system science.[1,35]

To address this challenge, we develop a mechanism-constrained diagnostic framework (MCDF) for regime-aware O_3 prediction and interpretation. MCDF integrates time-resolved observational data, PMF-resolved VOC source-related factors, and scenario-specific photochemical sensitivity indicators derived from observation-constrained box-model and empirical kinetic modeling approach (EKMA) analyses within a multitask neural-network architecture. At the loss-function level, a direction-consistency term guides the learned precursor-response directions toward the corresponding EKMA-diagnosed photochemical regime. In this design, the mechanistic constraints serve as externally derived process-level priors that guide the data-driven model. To assess the applicability of this framework, we selected Shenzhen, a representative O_3-polluted megacity in China, as a case study and employed an advanced multicomponent photochemical observation platform to evaluate O_3 prediction, mechanistic consistency, and regime-dependent source-related patterns under real urban atmospheric conditions. Building on these analyses, we further examined how precursor sensitivities and model-attributed source-related patterns vary across photochemical regimes, providing an interpretable basis for regime-aware O_3 diagnosis and precursor-control assessment.

2. MATERIALS AND METHODS

2.1. Study Region and Data Sources

This study focuses on Shenzhen, a Chinese megacity in the coastal Pearl River Delta, characterized by a subtropical monsoon climate with pronounced seasonal variability. Local topography transitions from low hills in the southeast to coastal plains in the northwest, creating heterogeneous surface conditions that make Shenzhen a representative setting for investigating O_3 pollution. As a core city in the Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen has committed to green and sustainable development and has implemented some of the most stringent NOx and VOCs control measures in China. Nevertheless, the city continues to experience recurrent rebounds in ambient O_3 concentrations, pointing to important gaps in current control strategies regarding the underlying formation mechanisms.[26,42,43]

The primary observational site in this study is the Peking University Atmospheric Observation Supersite (PKU-AOSS; 22.60° N, 113.98° E), located in a mixed-use urban area comprising commercial districts, high-density residential neighborhoods and green spaces, with no major industrial point sources in the immediate vicinity (Figure S1). The station provides a representative urban background for O_3 pollution and overall air quality in Shenzhen. We focus on high-O_3 seasons during 2022 and 2023. During the observation period, VOC concentrations were measured with high precision using gas chromatography−mass spectrometry (GC−MS) and proton-transfer-reaction time-of-flight mass spectrometry (PTR-ToF-MS),[44] while O_3, CO, NOx, SO_2 and $PM_{2.5}$ were continuously monitored using instruments from Thermo Fisher Scientific (USA). To characterize photochemical processes and meteorological influences on O_3 formation, we deployed a photolysis rate measurement system (PFS-100) and a formaldehyde monitoring system (FMS-100). Meteorological variables, including temperature, relative humidity, wind speed and wind direction, were measured with a WXT520 sensor (Vaisala, Finland). All datasets were subjected to rigorous preprocessing to ensure quality suitable for modeling; details of instrumentation and data products are summarized in Table S1.

To further enhance data quality and accuracy, we implemented a multi-tier outlier management protocol. Biweekly data reviews were performed to identify and correct abnormal values and to diagnose and address instrument malfunctions or environmental interferences, thereby ensuring the completeness and reliability of the dataset. Additional details on quality assurance and quality control procedures are provided in our previous work.[28,32,45−47] The construction of the dataset for the Shenzhen case study, including temporal resolution and sample definition, is described in Section 3.

2.2. Overview of the Mechanism-Constrained Multitask Model

The mechanism-constrained multitask model (MCMM) developed in this study integrates multitask learning with EKMA-derived photochemical constraints to jointly support O_3 prediction and regime-aware interpretation of precursor and source-related variability. The model architecture was selected from predefined hyperparameter ranges, including shared-encoder hidden dimensions of 32−96, 3−5 expert subnetworks, and hidden dimensions of 64 for the expert and task-specific tower layers; the final settings are summarized in the Supporting Information . The shared encoder, expert subnetworks, and task-specific towers are implemented as fully connected feed-forward neural networks with rectified linear unit (ReLU) activation. This configuration was chosen to balance computational efficiency, model stability, and predictive performance.

MCMM is trained with three coupled tasks: O_3 concentration prediction as the primary task, together with two mechanistic auxiliary tasks, namely control-regime classification and signed ridge distance prediction. The signed ridge distance is defined as the signed shortest distance from each sample to the matched scenario-specific EKMA ridge, where negative values indicate samples on the VOCs-limited side of the ridge, positive values indicate samples on the NOx-limited side, and the magnitude reflects the degree of deviation from the ridge. The shared encoder first extracts common feature representations from the input data, which are then passed to the expert layers, where each expert applies a task-tailored transformation. A gating network dynamically combines the expert outputs by assigning input-dependent weights to each expert, thereby adapting the representation to the prevailing chemical regime. The gating network learns input-dependent expert weights from the shared representation, while EKMA-derived photochemical-state information is incorporated within the mechanism-constrained multitask formulation. The auxiliary tasks and direction-consistency loss further reinforce this mechanistic information during training. Formally, MCMM can be expressed as

$$h = f_{\text{shared}}(X) \tag{1}$$

where X is the input feature vector for each data sample and f_{shared} denotes the shared encoder. The encoded representation h is then passed to a set of expert networks. For the O_3 prediction task, the model output is given by a gated mixture of expert outputs

$$y_{O_3} = \sum_{i=1}^{n_{\text{exp erts}}} \omega_i \cdot f_{\text{exp ert}_i}(h) \tag{2}$$

where ω_i is the gating weight for the i-th expert and f_{experti} denotes the corresponding expert network. The gating weights are determined by a gating network learned from the shared representation.

For the control-regime classification task, the model first produces classification logits through a task-specific classification tower

$$z_{\text{class}} = f_{\text{class}}(h) \tag{3}$$

The predicted probability distribution over chemical regimes is then obtained as

$$\hat{p}_{\text{class}} = \text{Softmax}(z_{\text{class}}) \tag{4}$$

where f_{class} denotes the classification head, z_{class} is the output logit vector, and $\hat{p}_{\text{class}}$ is the predicted probability distribution over chemical

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

regimes. For the signed ridge distance prediction task, the model output is generated by a task-specific regression tower

$$y_{\text{Distance}} = f_{\text{Distance}}(h) \tag{5}$$

where f_{Distance} denotes the regression head that maps the shared representation h to the predicted signed ridge distance.

To align model behavior with expected chemical dynamics, we introduce a gradient-based direction consistency loss that penalizes violations of the mechanistic response directions implied by EKMA. The penalty for NOx- and VOC-related gradients is written as

$$\text{loss}_{\text{dir}} = \lambda_{\text{dir}} \sum_i \max(0, -g_i \cdot d_i) \tag{6}$$

where g_i is the gradient of the model output with respect to feature i, and d_i is the expected sign of the response direction for that feature, as prescribed by the EKMA based chemical regime.

For regularization, we adopt a combination of L1 and group Lasso penalties to promote sparsity in the model weights and enhance interpretability

$$\text{loss}_{\text{reg}} = \beta_1 \sum_j \|w_j\|_1 + \beta_2 \sum_g \|w_g\|_2 \tag{7}$$

where w_j and w_g denote the weights associated with individual features and feature groups respectively, and β_1 and β_2 control the strengths of the L1 and group Lasso penalties.

The overall loss function is a weighted combination of the task specific losses, the direction consistency term and the regularization penalties

$$\mathcal{L} = \lambda_1 \cdot \mathcal{L}_{O_3} + \lambda_2 \cdot \mathcal{L}_{\text{Distance}} + \lambda_3 \cdot \mathcal{L}_{\text{class}} + \lambda_4 \cdot \mathcal{L}_{\text{dir}} + \lambda_5 \cdot \mathcal{L}_{L1} + \lambda_6 \cdot \mathcal{L}_{\text{GroupLasso}} \tag{8}$$

here, $\mathcal{L}_{O_3}$ is the mean squared error loss for the O_3 prediction task, $\mathcal{L}_{\text{Distance}}$ is the loss for signed ridge distance prediction, $\mathcal{L}_{\text{class}}$ is the cross entropy loss for control regime classification, $\mathcal{L}_{\text{dir}}$ is the gradient based direction consistency loss and $\mathcal{L}_{L1}$ and $\mathcal{L}_{\text{GroupLasso}}$ are the L1 and group Lasso penalties that promote sparse and interpretable solutions. The weighting coefficients λ_i are tuned by grid search based on validation performance.

Mechanistic information is incorporated into MCMM through complementary model-level and training-level constraints. EKMA-derived precursor-control regime and signed ridge distance provide scenario-specific photochemical-state information within the multitask formulation. The model treats O_3 concentration prediction as the primary task and includes two mechanistic auxiliary tasks: precursor-control regime classification (NOx-limited, VOCs-limited, or transitional) and signed ridge-distance prediction. These auxiliary tasks encourage the learned representation to retain information on the prevailing photochemical state. At the loss-function level, the conventional objective is further augmented with a direction-consistency term. Specifically, backpropagation is used to compute the partial derivatives of predicted O_3 with respect to NOx and VOCs, and the signs of these gradients are guided according to the corresponding EKMA-diagnosed control regime. For samples in the NOx-limited regime, the NOx gradient is required to be positive; for samples in the VOCs-limited regime, the VOC gradient is required to be positive; transitional-regime samples are subject to a weaker directional guidance. When the learned gradient directions deviate from the expected mechanistic directions, a directional penalty is added to the loss and weighted by the magnitude of the signed ridge distance. This design guides the learned local precursor-response directions toward the EKMA-derived photochemical sensitivity pattern, while retaining the data-driven prediction objective.

For the repeated-split evaluation, valid hourly samples were randomly divided into training, validation, and test subsets using a 64%/16%/20% split. This procedure was repeated 10 times with independent random seeds. For each split, the validation subset was used for early stopping and hyperparameter selection through grid search over predefined parameter ranges, whereas the test subset was held out for final evaluation. Together with the composite multitask loss, this repeated-split protocol was used to assess whether O_3 prediction performance and EKMA-consistent mechanistic diagnostics remained stable across different data partitions.

2.3. Positive Matrix Factorization Model

In this study, positive matrix factorization (PMF) was employed to resolve VOC source-related factors in the urban atmosphere of Shenzhen. PMF is an unsupervised, matrix factorization-based receptor model that, under non-negativity constraints, extracts latent source factors and quantifies their contributions to the observed concentration data.[48,49] For VOC source apportionment, PMF assumes that the observed VOC concentration matrix X can be expressed as the product of a source contribution matrix G and a source profile matrix F, plus a residual matrix E

$$X_{ij} = \sum_{k=1}^{p} G_{ik} \cdot F_{kj} + E_{ij} \tag{9}$$

where X_{ij} denotes the concentration of the j-th VOC species in the i-th sample, G_{ik} is the contribution of factor k to sample i, F_{kj} is the profile of factor k for species j, and E_{ij} is the residual representing the deviation between modeled and observed concentrations. The PMF solution is obtained by minimizing the weighted sum of squared residuals subject to non-negativity constraints on G and F, yielding optimal estimates of source contributions and source profiles.

Uncertainty treatment was incorporated into the PMF setup by assigning each observation an uncertainty based on measurement precision, detection limits, and data-screening procedures. These uncertainties were used to weight the residuals in the PMF optimization, thereby reducing the influence of low-confidence observations on the resolved factor profiles and contributions. To evaluate solution robustness and factor-number selection, 5-, 6-, 7-, and 8-factor solutions were compared using Q diagnostics, marginal Q reduction, convergence behavior, solution stability, residual structure, source-profile interpretability, and diurnal patterns. Although Q values generally decreased as the number of factors increased, Q reduction alone was not used as the sole criterion. The seven-factor solution provided the best balance between model fit, numerical stability, and physical interpretability, whereas the eight-factor solution showed reduced stability and a greater risk of over-splitting.

The selected seven-factor PMF solution resolved vehicle emissions, gas evaporation, solvent usage, biomass burning, industrial emissions, a background-influenced factor, and natural sources. The background-influenced factor was not interpreted as a distinct local emission source, but as a background- and mixing-influenced VOC factor with comparatively smooth temporal variability and weaker source-specific tracer characteristics. The detailed factor-number evaluation, apportionment procedure, and factor interpretation are provided in Section S1.1 of the Supporting Information and Figures S2 and S3. These PMF-resolved factors were used as observation-derived source-related predictors in MCMM, rather than as direct O_3 production contributions.

2.4. Observation-Constrained Box Model

In this study, an observation-constrained box model was used to diagnose the chemical processes governing O_3 formation in Shenzhen. Three-dimensional chemical transport models, such as the Weather Research and Forecasting−Community Multiscale Air Quality (WRF-CMAQ) and Weather Research and Forecasting model coupled with Chemistry (WRF-Chem), are essential for representing regional transport and emission-driven pollutant evolution over large spatial scales.[50,51] In contrast, the box model used here serves as a complementary process-level tool for examining local photochemical sensitivity under observation-constrained conditions. It is constrained primarily by in situ observations and coupled with chemical kinetic equations, and its outputs are interpreted as diagnostics of photochemical response rather than as a full local-versus-regional O_3 budget.[47] Instrument deployment and associated parameter settings for the observation period are summarized in Table S1.

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

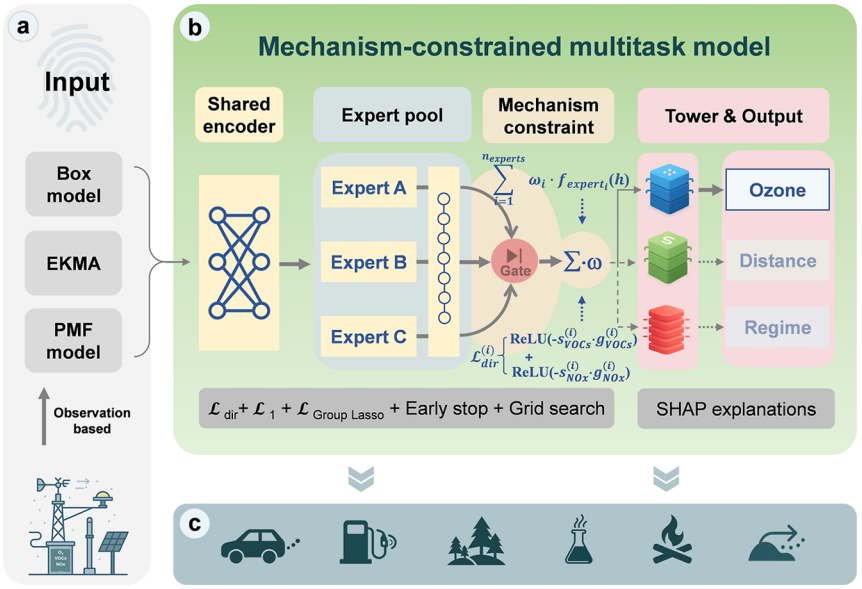


Figure 1. Mechanism-constrained diagnostic framework for regime-aware urban ozone prediction and interpretation. (a) Information integration module. (b) Mechanism-constrained learning module. (c) Regime-aware interpretation module.

We used Master Chemical Mechanism version 3.3.1 (MCM v3.3.1) as the chemical mechanism because it provides a detailed representation of VOC and NOx oxidation chemistry and is widely used for process-level atmospheric chemistry analysis. This mechanism enables a more explicit simulation of radical chemistry, secondary pollutant formation, and O_3 production pathways than simplified mechanisms (Table S2). The VOC species constrained in the box model and their parameter configurations are listed in Table S3. Because the box model is a zero-dimensional (0-D) framework,[52] background and transport influences are not explicitly resolved in three dimensions, but are represented in an effective sense through prescribed background concentrations and a first-order dilution term

$$\left.\frac{d[X]}{dt}\right|_{dil} = -k_{dil}(t)([X] - [X]_b) \tag{10}$$

where $[X]$ is the concentration of a given species, $[X]_b$ is its prescribed background concentration, and $k_{dil}(t)$ is an effective first-order dilution rate.

To reduce the dependence of the mechanistic constraints on a single representative dilution condition, the box-model and EKMA calculations were implemented using a meteorology- and ventilation-stratified approach. Specifically, daytime temperature, relative humidity, and a ventilation index (VI) based on boundary-layer height (BLH) and wind speed (WS) were used to construct temperature−relative humidity−ventilation index (T−RH−VI) scenarios. For each retained scenario, a representative 24 h box-model input profile was generated from the observed mean diurnal cycle of the corresponding days, including meteorological variables, photolysis rates, precursor concentrations, and VOC composition. Details of the ventilation index, scenario thresholds, and scenario screening are provided in Section S1.2 of the Supporting Information.

Sensitivity analysis of O_3 formation was then performed using the empirical kinetic modeling approach (EKMA). Within each retained T−RH−VI scenario, VOC and NOx levels were perturbed to construct a scenario-specific EKMA surface.[53] The EKMA ridge, signed ridge distance, and precursor control-regime classification were diagnosed separately for each scenario, and each hourly observation was assigned the mechanistic indicators from its matched scenario rather than from a single period-mean EKMA surface. Scenario-specific EKMA surfaces are shown in Figures S4 and S5. Within MCDF, the EKMA-derived indicators provide scenario-specific photochemical-state information to MCMM. Their mechanistic structure is further reinforced through the auxiliary tasks and direction-consistency loss. This design embeds scenario-specific photochemical sensitivity information into the data-driven learning framework, while recognizing that the box-model and EKMA analyses remain observation-constrained 0-D diagnostics rather than explicit simulations of three-dimensional transport.

3. RESULTS

3.1. Overview of the Mechanism-Constrained Diagnostic Framework for Regime-Aware Ozone Prediction and Interpretation

To improve the chemical interpretability of data-driven O_3 prediction while retaining predictive capability, we developed MCDF. As shown in Figure 1, MCDF comprises three coupled modules: an information integration module, a mechanism-constrained learning module, and a regime-aware interpretation module. These correspond to three stages of the diagnostic chain, namely input construction, mechanism-constrained modeling, and model interpretation, and together define a workflow in which data-driven learning is guided by externally derived photochemical constraints.

The information integration module (Figure 1a) serves as the front end of the framework. It combines time-resolved observations of VOCs, O_3, and other atmospheric constituents with photochemical process information generated by a box model to construct input features and auxiliary indicators for mechanism-constrained learning. First, PMF is applied to resolve VOC source-related factors, such as vehicle emissions, gas evaporation, and natural sources. Second, a T−RH−VI-stratified EKMA scenario library is constructed using a MCM-based box model to simulate O_3 response surfaces under perturbed VOC and NOx levels. From the EKMA responses, we derive two mechanistic diagnostic indicators that provide external mechanistic priors for the mechanism-constrained learning module: (1) precursor control-regime classification and (2) precursor signed ridge distance.

The mechanism-constrained learning module (Figure 1b) forms the modeling core of the diagnostic framework. To incorporate EKMA-derived photochemical sensitivity into data-driven learning, we develop a mechanism-constrained multitask model (MCMM). The role of MCMM is not to replace EKMA

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

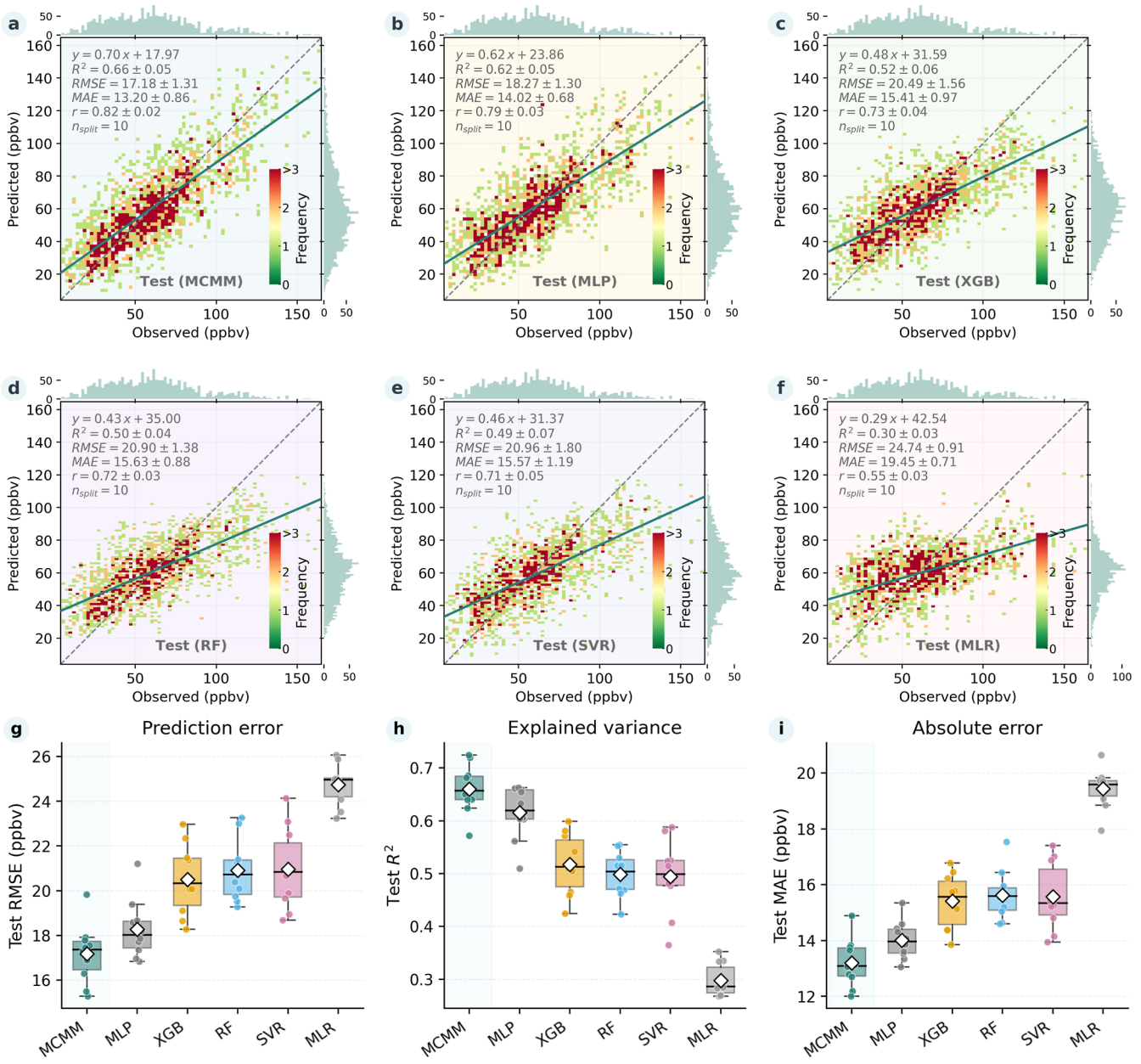


Figure 2. Repeated-split evaluation of ozone prediction performance. (a–f) Pooled test-set predictions versus observations for MCMM and baseline models across 10 repeated train/validation/test splits. Solid lines indicate fitted linear regressions, dashed lines indicate the 1:1 reference, and color shading indicates point frequency. Performance metrics are reported as mean ± SD across the 10 split-level evaluations, whereas the scatter-density distributions are based on pooled test predictions from all splits. (g–i) Distributions of test RMSE, R^2, and MAE across the 10 repeated splits. Diamonds denote mean values.

in diagnosing control regimes, nor to independently recover regime structure from observations alone. Rather, it uses EKMA-derived information as mechanistic priors to guide the learning of precursor-response relationships and model-attributed predictor patterns while encouraging photochemically consistent response behavior. By integrating observational inputs, PMF-resolved source-related predictors, and mechanistic constraints within a unified learning framework, MCMM supports regime-aware O_3 prediction and interpretation under varying pollution conditions. Detailed descriptions of the MCMM architecture, multitask learning design, and constraint implementation are provided in the Methods section.

The regime-aware interpretation module (Figure 1c) forms the downstream component of the framework and focuses on translating model outputs into chemically interpretable and control-relevant diagnostics. By interrogating the model-attributed patterns produced by MCMM, this module identifies key source-related predictors associated with O_3 variability under different pollution and chemical-regime conditions and examines how precursor-response behavior varies across photochemical regimes, including nonlinear transitions between NOx- and VOC-sensitive conditions. These diagnostics provide model-based evidence for how predictor attribution and precursor-response behavior vary across different pollution and chemical-regime conditions.

Taken together, MCDF links mechanistic information extraction, source-related predictor construction, and regime-aware model interpretation within an observation-constrained

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

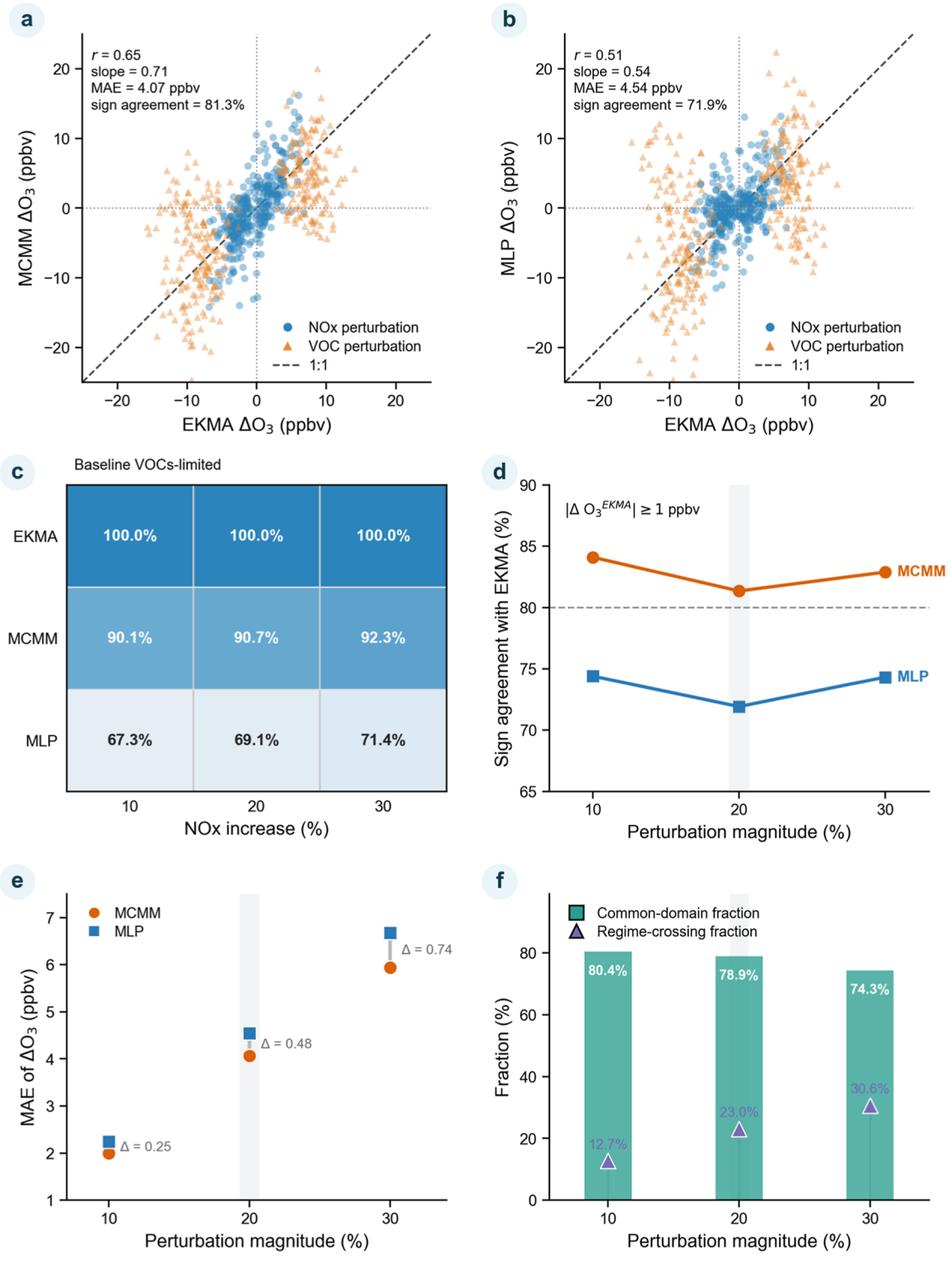


Figure 3. EKMA-referenced evaluation of state-consistent finite precursor responses across perturbation magnitudes. (a,b) Comparison of O_3 responses predicted by MCMM and MLP, respectively, with the corresponding EKMA responses under ±20% NOx and VOC perturbations. Dashed lines denote the 1:1 relationship. Directional agreement is the fraction of cases for which model- and EKMA-derived ΔO_3 have the same sign; cases with $|\Delta O_3^{EKMA}|<1$ ppbv are excluded from this metric to avoid near-zero reference responses with weakly defined directions. (c) Fraction of initially VOCs-limited cases exhibiting a negative O_3 response following 10%, 20%, and 30% increases in NOx for EKMA, MCMM, and MLP. (d) Directional agreement with EKMA across precursor perturbation magnitudes, using the same $|\Delta O_3^{EKMA}|\geq 1$ ppbv criterion. (e) *MAE* of model-predicted ΔO_3 relative to EKMA, with Δ denoting $MAE_{MLP}-MAE_{MCMM}$. (f) Fraction of counterfactual cases remaining within the common EKMA−ML supported domain and fraction crossing the EKMA-defined photochemical-regime boundary. Methodological details are provided in Section S1.4 of the Supporting Information.

diagnostic workflow. By integrating multi-source observations, PMF-resolved source-related factors, EKMA-derived photochemical indicators, and machine-learning algorithms, it reduces reliance on purely statistical interpretation and supports regime-

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

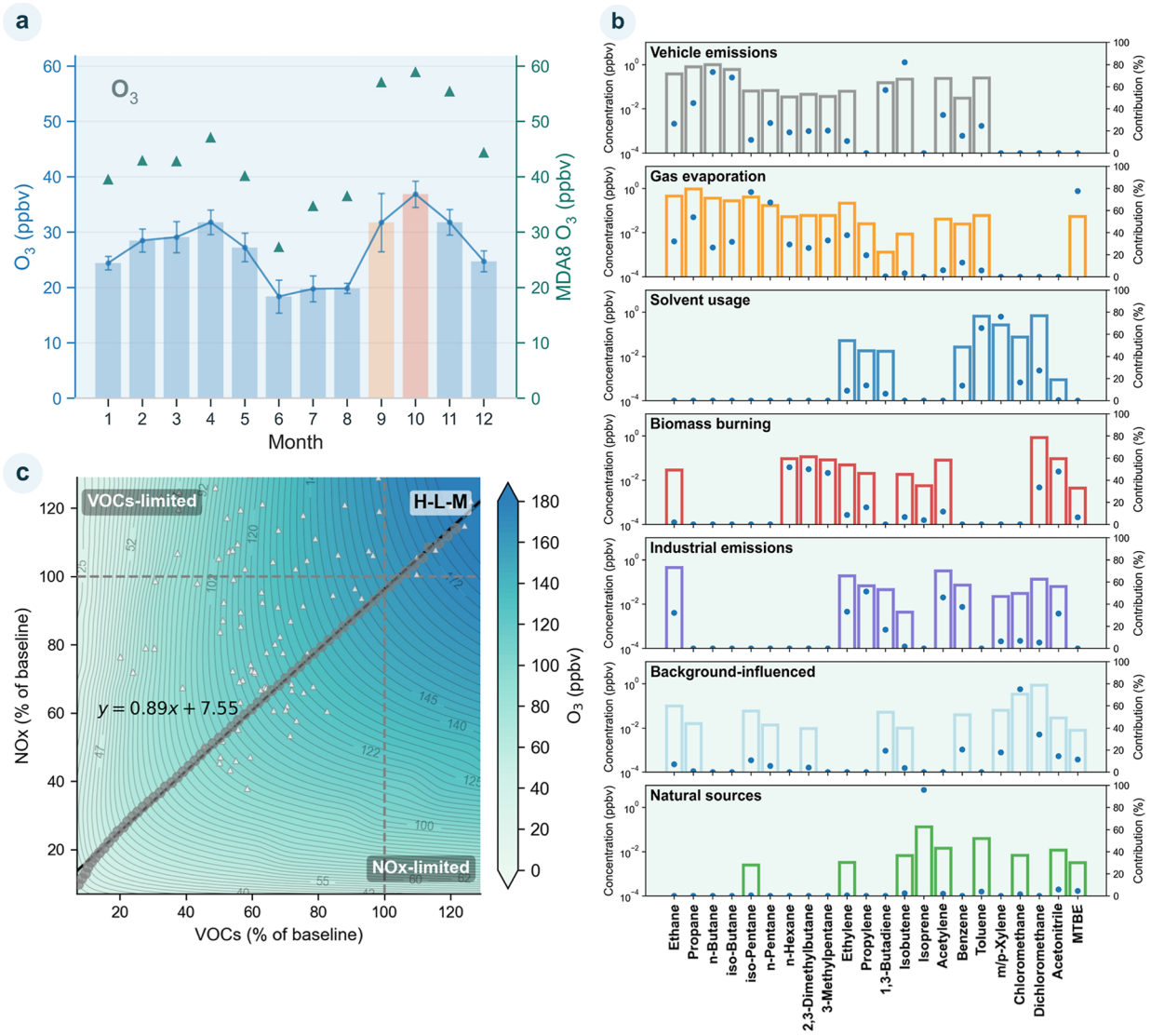


Figure 4. Observed ozone characteristics, VOC source-related factor profiles, and representative EKMA-based regime diagnosis. (a) Monthly mean O_3 and MDA8 O_3 during the study period, with the September−October photochemically active period highlighted. (b) Source-related factor profiles from the seven-factor PMF solution. Open bars represent factor-specific species concentrations, and blue dots represent the relative contribution of each factor to individual VOC species. (c) Representative scenario-specific EKMA surface for the H−L−M T−RH−VI scenario, where H−L−M denotes high temperature, low relative humidity, and middle ventilation. This panel is shown as an example of the scenario-specific EKMA diagnostics; the complete set of scenario-specific EKMA surfaces is provided in the Supporting Information.

aware interpretation of precursor and source-related variability associated with O_3. In doing so, MCDF provides a methodological basis for mechanistically informed O_3 prediction and precursor-control diagnosis.

3.2. Predictive Performance and Mechanistic Consistency of the Mechanism-Constrained Multitask Model

As the predictive core of the framework, MCMM was evaluated for both O_3 prediction and consistency of its precursor-response behavior with EKMA-diagnosed photochemical states. Valid hourly samples were partitioned into training, validation, and test subsets using 10 independent random splits with fixed proportions. For each split, model fitting, early stopping, and hyperparameter selection were performed using only the training and validation subsets, while the test subset was held out for final evaluation. MCMM was compared with five commonly used data-driven models, including a multilayer perceptron (MLP), extreme gradient boosting (XGB), random forest (RF), support vector regression (SVR), and multiple linear regression (MLR). All models were trained using the same hourly observational dataset and observational predictor set, whereas MCMM additionally incorporated EKMA-derived photochemical-state information within its mechanism-constrained multitask formulation. Hyperparameter ranges and final model settings are provided in Tables S4−S9.

MCMM showed the best overall O_3 prediction performance and remained stable across repeated data partitions (Figure 2). Across the 10 test splits, MCMM achieved R^2 = 0.66 ± 0.05, *RMSE* = 17.18 ± 1.31 ppbv, *MAE* = 13.20 ± 0.86 ppbv, and r = 0.82 ± 0.02. Among the unconstrained models, MLP showed the strongest overall performance, with R^2 = 0.62 ± 0.05, RMSE = 18.27 ± 1.30 ppbv, MAE = 14.02 ± 0.68 ppbv, and r = 0.79 ± 0.03. XGB, RF, SVR, and MLR showed lower R^2 values or larger prediction errors, with R^2 ranging from 0.30 ± 0.03 to 0.52 ± 0.06 and RMSE from 20.49 ± 1.56 to 24.74 ± 0.91 ppbv. The repeated-split distributions of RMSE, R^2, and MAE further

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

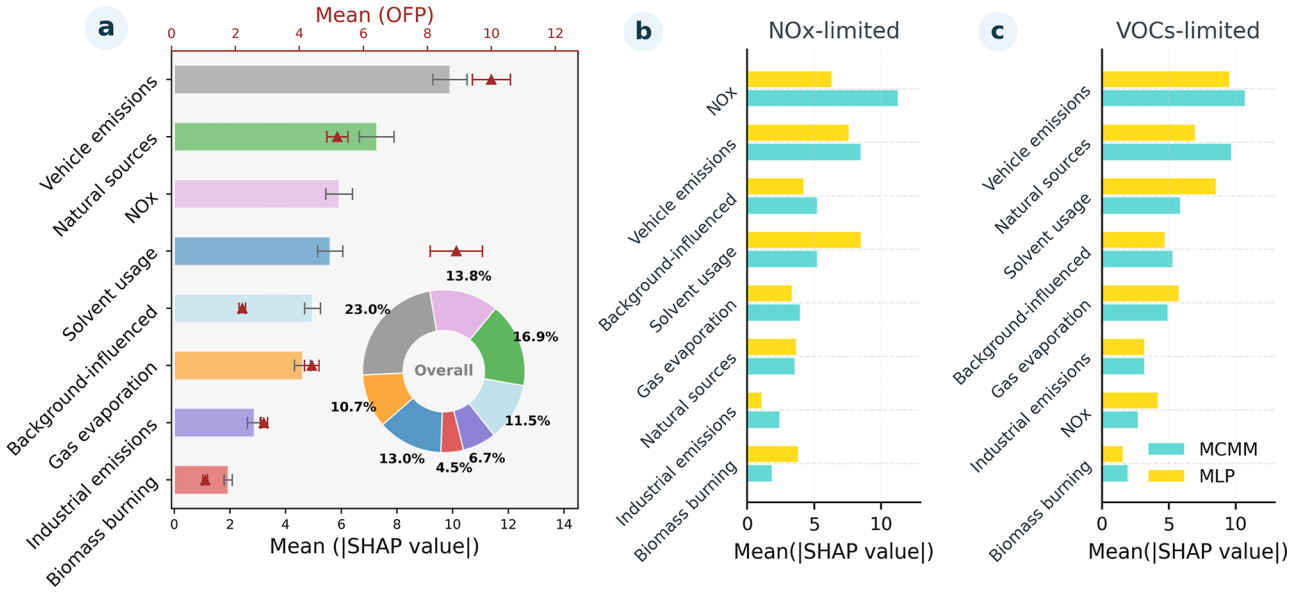


Figure 5. Model-attributed importance of ozone predictors across photochemical regimes in Shenzhen. (a) Comparison of MCMM mean absolute SHAP importance with OFP-based VOC reactivity. Horizontal bars show mean absolute SHAP values, and red ticks along the upper axis indicate OFP-based activity for VOC source-related factors. The inset pie chart summarizes the fractional contribution of each VOC source-related factor to total OFP. (b) Comparison of mean absolute SHAP importance between MCMM and the unconstrained MLP under NOx-limited conditions. (c) Same as (b), but under VOCs-limited conditions. Mean absolute SHAP values in all panels represent attribution magnitude and do not indicate attribution direction.

showed that the performance advantage of MCMM persisted across alternative data partitions rather than arising from a single favorable split (Figure 2g–i). Because MLP was the strongest unconstrained baseline and was architecturally closest to MCMM, it was used for the subsequent mechanistic-response comparisons. The single-task MLP contained no EKMA-derived signed ridge distance or regime information and was trained without auxiliary mechanistic tasks or directional regularization. The auxiliary-task outputs further provided internal diagnostics of how the EKMA-derived photochemical-state information was represented within the multitask formulation (Figures S6 and S7; for detailed metric definitions and calculations, see Section S1.3 of the Supporting Information). For signed ridge-distance regression, the predicted values tracked variations in distance from the EKMA ridge, with a split-level R^2 of 0.65 ± 0.05. For precursor-control regime classification, MCMM achieved an accuracy of 0.82 ± 0.02 and a Macro-F1 score of 0.66 ± 0.05 across repeated splits. The row-normalized confusion matrix showed a clear diagonal structure, with the greatest uncertainty for transitional samples near the EKMA ridge. The two auxiliary tasks captured complementary information on the continuous ridge position and the corresponding precursor-control regime.

We next examined whether this photochemical-state information was reflected in the modeled response to precursor variability. Fixed-state partial-dependence analysis showed clearer regime-dependent precursor responses in MCMM than in the unconstrained MLP (Figure S8). The positive NOx response was strongest under NOx-limited conditions and substantially attenuated under VOCs-limited conditions, whereas responses to VOC-related predictors became more pronounced under VOC-sensitive conditions. We further examined these regime-dependent responses using a state-consistent finite-perturbation analysis (Section S1.4 of the Supporting Information). Starting from held-out observational states, NOx or VOC-related predictors were perturbed by prescribed finite amounts, while the corresponding EKMA-derived photochemical state was updated consistently with each perturbed precursor condition. The resulting O_3 changes predicted by MCMM and MLP were compared with the corresponding responses from the matched scenario-specific EKMA surface (Figure 3). For the ±20% precursor perturbations shown in Figure 3a,b, MCMM showed closer correspondence with the EKMA response than MLP, with r = 0.65, a regression slope of 0.71, an *MAE* of 4.07 ppbv, and directional agreement of 81.3% for cases with $|\Delta O_3^{EKMA}| \geq 1$ ppbv (Figure 3a). The corresponding MLP values were r = 0.51, slope = 0.54, *MAE* = 4.54 ppbv, and directional agreement = 71.9% (Figure 3b). Slopes below unity indicated that both models attenuated the EKMA response amplitude, although MCMM more closely followed both its direction and magnitude.

The clearest regime-specific contrast occurred for samples initially classified as VOCs-limited (Figure 3c). Increasing NOx by 10%, 20%, and 30% produced negative EKMA O_3 responses in all evaluated cases. MCMM predicted a negative response in 90.1%, 90.7%, and 92.3% of cases, respectively, compared with 67.3%, 69.1%, and 71.4% for MLP. Across the 10–30% perturbation range, directional agreement with EKMA remained approximately 81–84% for MCMM and 72–74% for MLP (Figure 3d). MCMM also maintained a lower MAE in ΔO_3 relative to EKMA at each perturbation magnitude, with MLP–MCMM MAE differences of 0.25, 0.48, and 0.74 ppbv at 10%, 20%, and 30%, respectively (Figure 3e). Larger perturbations increasingly moved samples across the EKMA-defined photochemical-regime boundary (Figure 3f). The regime-crossing fraction increased from 12.7% at 10% perturbation to 23.0% at 20% and 30.6% at 30%, while 80.4%, 78.9%, and 74.3% of the corresponding counterfactual cases remained within the common EKMA–ML supported domain. MCMM therefore combined improved O_3 prediction with precursor-response behavior that was more consistent with the EKMA-informed photochemical constraints than that of the unconstrained MLP.

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX–XXX

3.3. Regime-Aware Interpretation of Source-Related Ozone Variability in Shenzhen

To evaluate the case-study applicability of MCDF under real urban atmospheric conditions, we applied it to Shenzhen, a representative Chinese megacity with pronounced O_3 pollution. Shenzhen is a national pilot city for O_3 control and is characterized by a complex mixture of emission sources and meteorological conditions. In autumn, the city frequently experiences high temperatures, strong solar radiation, and low wind speeds, which together favor active photochemistry (Figures S9 and S10). We therefore selected September to October in recent years as a representative analysis window, during which the citywide mean maximum daily 8 h average O_3 (MDA8 O_3) remained elevated (Figures 4a, S11 and S12), exhibiting typical characteristics of autumn O_3 pollution. This period provided a suitable setting for evaluating the framework under active photochemical conditions. For this case study, all variables were harmonized to an hourly resolution, resulting in 2568 valid samples after preprocessing. Each sample included the target O_3 concentration together with VOC measurements, conventional air pollutants, photolysis-related parameters, and meteorological variables (Table S1). We used PMF to resolve seven VOC source-related factors in Shenzhen, including vehicle emissions, gas evaporation, solvent usage, biomass burning, industrial emissions, a background-influenced factor, and natural sources, and derived hourly time series of factor contributions (Figures 4b, S2 and S3). In parallel, T−RH−VI-stratified EKMA surfaces were constructed using a MCM-based box model. For each retained T−RH−VI scenario, the EKMA ridge, signed ridge distance, and precursor control-regime classification were diagnosed separately, and each hourly sample was assigned the mechanistic indicators from its matched scenario. Figure 4c shows one representative scenario-specific EKMA surface, and the complete scenario-specific diagnostics are provided in Section S1.5 of the Supporting Information.

We then used MCMM to calculate mean absolute SHAP values for each predictor and compared the MCMM-derived importance of the PMF-resolved VOC source-related factors with O_3 formation potential (OFP)-based activity (Section S1.6 of the SI).[39,40,54] Mean absolute SHAP values quantify the magnitude of model-attributed influence on predicted O_3 variability under observed conditions but do not distinguish whether a predictor shifts the prediction upward or downward. OFP, in contrast, represents a VOC-only reactivity-based metric calculated under standardized assumptions.[55,56] The two diagnostics therefore characterize different aspects of the observed system and should not be interpreted as equivalent source-contribution measures. Figure 5a shows that vehicle emissions, natural sources, and solvent usage had the highest mean absolute SHAP values among the PMF-resolved source-related factors, while NOx also ranked among the leading predictors when precursor variables were considered together. Among the anthropogenic source-related factors, vehicle emissions showed the highest model-attributed importance,[57,58] while natural sources and solvent usage also showed substantial importance under the observed high-O_3-season conditions. In comparison, the OFP-based ranking emphasizes VOC reactivity without accounting for the regime-dependent role of NOx or covariation among source-related factors, meteorology, and photochemical state. The two diagnostics therefore provide complementary perspectives on VOC reactivity and observation-conditioned model attribution.

We further compared the magnitude of model-attributed importance between MCMM and the unconstrained MLP under different EKMA-diagnosed control regimes (Figures 5b,c and S13). Under NOx-limited conditions, MCMM assigned the highest mean absolute SHAP importance to NOx, followed by vehicle emissions and several VOC source-related factors, including the background-influenced factor, solvent usage, gas evaporation, and natural sources. Relative to MLP, the MCMM attribution pattern shifted more strongly toward NOx under NOx-limited conditions. Under VOCs-limited conditions, both models identified vehicle emissions as an important predictor, but MCMM assigned substantially lower relative importance to NOx and greater importance to several VOC-related factors, including vehicle emissions, natural sources, solvent usage, and the background-influenced factor. These differences indicate a clearer regime-dependent redistribution of attribution magnitude under the mechanism-constrained representation. Because mean absolute SHAP values do not contain directional information, we examined the direction of these model-attributed relationships separately using signed SHAP values.

Regime-specific signed SHAP distributions further resolved the direction and heterogeneity of model-attributed relationships (Figures S14 and S15). Under NOx-limited conditions, high NOx values were predominantly associated with positive SHAP contributions. Under VOCs-limited conditions, the signed SHAP distribution for NOx shifted toward more negative values, with high NOx values extending into a distinct negative tail, although the distribution was not uniformly negative. In contrast, higher values of the leading VOC source-related factors, including the vehicle-emission, natural-source, and solvent-usage factors, were predominantly associated with positive SHAP contributions under VOCs-limited conditions. The background-influenced factor showed a different pattern, with higher values tending to be associated with negative SHAP contributions in both regimes. This pattern may reflect broader background or air-mass variability covarying with this factor rather than a direct negative chemical effect on O_3. Median values and interquartile ranges summarize these regime-dependent patterns in Figure S14, while the full sample-level distributions and relative feature values are shown in Figure S15. Signed SHAP values are used here to describe the direction and heterogeneity of model attribution, complementing the magnitude information provided by mean absolute SHAP values; they are not interpreted as local chemical derivatives or causal source contributions.

Overall, the Shenzhen case study identified distinct observation-conditioned patterns in both the magnitude and direction of model attribution. Vehicle emissions, natural sources, and solvent usage showed the highest mean absolute SHAP importance among the PMF-resolved source-related factors, while NOx also showed high importance as a precursor predictor. Both the magnitude and directional patterns of these model-attributed relationships varied across EKMA-diagnosed precursor-control regimes. These results illustrate the value of combining observation-derived source-related predictors with photochemical-regime information when interpreting urban O_3 variability.

4. DISCUSSION

The MCDF developed in this study combines multi-source observations, PMF-resolved VOC source-related predictors, observation-constrained box-model and EKMA diagnostics, and data-driven learning. Its central contribution is to embed

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

externally derived photochemical information into O_3 prediction and interpretation rather than to replace mechanistic models or infer chemical regimes from observations alone. EKMA-derived control-regime information and signed ridge distance provide scenario-specific photochemical-state information within MCMM, while the auxiliary tasks and direction-consistency loss reinforce this mechanistic structure during learning. The state-consistent finite-perturbation analysis further evaluates how this EKMA-derived photochemical information propagates through MCMM into finite precursor-response behavior. MCDF therefore links predictive capability with photochemically constrained response behavior and provides a framework for regime-aware interpretation of data-driven O_3 predictions.

The Shenzhen case study further indicates that model-attributed precursor and source-related patterns depend on the prevailing photochemical regime rather than following a fixed ranking. Changes in both mean absolute and signed SHAP patterns between NOx- and VOCs-limited conditions show that the magnitude and direction of model attribution vary with photochemical state. This regime dependence helps explain why fixed source-related rankings or VOC-reactivity metrics alone may provide an incomplete view of observed urban O_3 variability. The value of the framework therefore lies in combining receptor-derived source-related information with scenario-specific photochemical constraints, allowing predictor relationships to be interpreted in the photochemical context in which they occur rather than as regime-invariant effects.

Despite the strengths of the proposed framework as an observation-constrained diagnostic approach, several limitations must be acknowledged. First, MCDF does not decompose observed O_3 into regional transport, background, and local photochemical production, nor does it quantify causal contributions of individual original emission sources to observed O_3. Regional background variability and transport-related influences are not explicitly resolved in the current framework. In the box model, these processes are represented in an effective sense through the background and dilution terms, while in the machine-learning framework, they may be partially embedded in observed meteorological states, precursor levels, and PMF-derived source-related factors. The PMF factors themselves represent receptor-side VOC variability after chemical processing, dilution, transport, and mixing and should therefore be interpreted as source-related predictor representations rather than direct O_3 source contributions. This limitation is particularly relevant on shorter time scales, when transient transport events or residual-layer effects may strongly influence observed O_3 variability. Second, both PMF-based source apportionment and SHAP-based interpretation are constrained by the representativeness of the observational inputs and the model-based nature of the analysis. Although mean absolute SHAP values and signed SHAP distributions provide useful diagnostics of model-attributed importance and attribution direction, respectively, they do not establish causal relationships or local chemical sensitivities. Future work could improve source-related interpretation and mechanistic resolution by incorporating complementary source tracers and transport diagnostics together with causal-inference designs, independent intervention-based analyses, or source-resolved chemical transport modeling. In addition, uncertainties associated with upstream models, including PMF for source apportionment and the box model for EKMA-based mechanistic constraints, may propagate into the resulting model-attributed patterns. A detailed discussion of these uncertainties is provided in Section S1.7 of the Supporting Information.

Finally, although the framework is conceptually extensible to other cities and chemical regimes, its current implementation still depends on relatively information-rich observations to support source-related predictor construction, mechanistic constraint construction, and model training. This dependence may limit its immediate applicability in data-sparse regions and may contribute to sampling variability in the reported performance metrics, despite the repeated-split evaluation. Accordingly, the present framework is better viewed as an adaptable mechanism-constrained diagnostic framework rather than a universally deployable fixed model. A practical pathway for broader application is to first evaluate the key mechanism-constrained relationships in representative cities and then reduce the observational burden for cross-city deployment through simplified feature sets, transfer learning, and other data-efficient extensions. Further evaluation across diverse cities, seasons, and chemical regimes will be important for assessing its robustness, scalability, and practical utility for broader air-quality applications.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.6c12789.

Additional details on PMF analysis, scenario-specific box-model and EKMA calculations, model evaluation, state-consistent finite-precursor perturbation analysis, SHAP interpretation, and uncertainty analysis; supplementary figures and tables supporting the main text (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Ling-Yan He − *Key Laboratory for Urban Habitat Environmental Science and Technology, Peking University Shenzhen Graduate School, Shenzhen 518055, China;* orcid.org/0000-0003-3955-3473; Email: hely@pku.edu.cn

Authors

Yong Cheng − *Key Laboratory for Urban Habitat Environmental Science and Technology, Peking University Shenzhen Graduate School, Shenzhen 518055, China;* orcid.org/0000-0001-9590-9548

Yan Peng − *Key Laboratory for Urban Habitat Environmental Science and Technology, Peking University Shenzhen Graduate School, Shenzhen 518055, China*

Xiao-Feng Huang − *Key Laboratory for Urban Habitat Environmental Science and Technology, Peking University Shenzhen Graduate School, Shenzhen 518055, China;* orcid.org/0000-0003-0488-3333

Hui Zeng − *Key Laboratory for Urban Habitat Environmental Science and Technology, Peking University Shenzhen Graduate School, Shenzhen 518055, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.6c12789

Notes

The authors declare no competing financial interest.

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

■ ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (42375101), the Science and Technology Plan of Shenzhen Municipality (SYSPG20241211173440002), and the China Postdoctoral Science Foundation (2026M791294).

■ REFERENCES

(1) Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature.* **2019**, *566* (7743), 195−204.

(2) Bi, K.; Xie, L.; Zhang, H.; Chen, X.; Gu, X.; Tian, Q. Accurate Medium-Range Global Weather Forecasting with 3d Neural Networks. *Nature.* **2023**, *619* (7970), 533−538.

(3) Dedoussi, I. C.; Eastham, S. D.; Monier, E.; Barrett, S. R. H. Premature Mortality Related to United States Cross-State Air Pollution. *Nature.* **2020**, *578* (7794), 261−265.

(4) Hu, J.; Zhang, Y. Reinforce the Coordinated Control of O3 and PM2.5 to Continuously Improve China's Ambient Air Quality. *Chin. Sci. Bull.* **2022**, *67* (18), 1975−1977.

(5) Wang, Y.; Yang, Y.; Yuan, Q.; Li, T.; Zhou, Y.; Zong, L.; Wang, M.; Xie, Z.; Ho, H. C.; Gao, M.; Tong, S.; Lolli, S.; Zhang, L. Substantially Underestimated Global Health Risks of Current Ozone Pollution. *Nat. Commun.* **2025**, *16* (1), 102.

(6) Yao, M.; Niu, Y.; Liu, S.; Liu, Y.; Kan, H.; Wang, S.; Ji, J. S.; Zhao, B. Mortality Burden of Cardiovascular Disease Attributable to Ozone in China: 2019 Vs 2050. *Environ. Sci. Technol.* **2023**, *57* (30), 10985−10997.

(7) Murray, C. J. L.; Aravkin, A. Y.; Zheng, P.; Abbafati, C.; Abbas, K. M.; Abbasi-Kangevari, M.; Abd-Allah, F.; Abdelalim, A.; Abdollahi, M.; Abdollahpour, I.; Abegaz, K. H.; Abolhassani, H.; Aboyans, V.; Abreu, L. G.; Abrigo, M. R. M.; Abualhasan, A.; Abu-Raddad, L. J.; Abushouk, A. I.; Adabi, M.; Adekanmbi, V.; Adeoye, A. M.; Adetokunboh, O. O.; Adham, D.; Advani, S. M.; Agarwal, G.; Aghamir, S. M. K.; Agrawal, A.; Ahmad, T.; Ahmadi, K.; Ahmadi, M.; Ahmadieh, H.; Ahmed, M. B.; Akalu, T. Y.; Akinyemi, R. O.; Akinyemiju, T.; Akombi, B.; Akunna, C. J.; Alahdab, F.; Al-Aly, Z.; Alam, K.; Alam, S.; Alam, T.; Alanezi, F. M.; Alanzi, T. M.; Alemu, B. w.; Alhabib, K. F.; Ali, M.; Ali, S.; Alicandro, G.; Alinia, C.; Alipour, V.; Alizade, H.; Aljunid, S. M.; Alla, F.; Allebeck, P.; Almasi-Hashiani, A.; Al-Mekhlafi, H. M.; Alonso, J.; Altirkawi, K. A.; Amini-Rarani, M.; Amiri, F.; Amugsi, D. A.; Ancuceanu, R.; Anderlini, D.; Anderson, J. A.; Andrei, C. L.; Andrei, T.; Angus, C.; Anjomshoa, M.; Ansari, F.; Ansari-Moghaddam, A.; Antonazzo, I. C.; Antonio, C. A. T.; Antony, C. M.; Antriyandarti, E.; Anvari, D.; Anwer, R.; Appiah, S. C. Y.; Arabloo, J.; Arab-Zozani, M.; Ariani, F.; Armoon, B.; Ärnlöv, J.; Arzani, A.; Asadi-Aliabadi, M.; Asadi-Pooya, A. A.; Ashbaugh, C.; Assmus, M.; Atafar, Z.; Atnafu, D. D.; Atout, M. M. W.; Ausloos, F.; Ausloos, M.; Ayala Quintanilla, B. P.; Ayano, G.; Ayanore, M. A.; Azari, S.; Azarian, G.; Azene, Z. N.; Badawi, A.; Badiye, A. D.; Bahrami, M. A.; Bakhshaei, M. H.; Bakhtiari, A.; Bakkannavar, S. M.; Baldasseroni, A.; Ball, K.; Ballew, S. H.; Balzi, D.; Banach, M.; Banerjee, S. K.; Bante, A. B.; Baraki, A. G.; Barker-Collo, S. L.; Bärnighausen, T. W.; Barrero, L. H.; Barthelemy, C. M.; Barua, L.; Basu, S.; Baune, B. T.; Bayati, M.; Becker, J. S.; Bedi, N.; Beghi, E.; Béjot, Y.; Bell, M. L.; Bennitt, F. B.; Bensenor, I. M.; Berhe, K.; Berman, A. E.; Bhagavathula, A. S.; Bhageerathy, R.; Bhala, N.; Bhandari, D.; Bhattacharyya, K.; Bhutta, Z. A.; Bijani, A.; Bikbov, B.; Bin Sayeed, M. S.; Biondi, A.; Birihane, B. M.; Bisignano, C.; Biswas, R. K.; Bitew, H.; Bohlouli, S.; Bohluli, M.; Boon-Dooley, A. S.; Borges, G.; Borzì, A. M.; Borzouei, S.; Bosetti, C.; Boufous, S.; Braithwaite, D.; Breitborde, N. J. K.; Breitner, S.; Brenner, H.; Briant, P. S.; Briko, A. N.; Briko, N. I.; Britton, G. B.; Bryazka, D.; Bumgarner, B. R.; Burkart, K.; Burnett, R. T.; Burugina Nagaraja, S.; Butt, Z. A.; Caetano dos Santos, F. L.; Cahill, L. E.; Cámera, L. L. A.; Campos-Nonato, I. R.; Cárdenas, R.; Carreras, G.; Carrero, J. J.; Carvalho, F.; Castaldelli-Maia, J. M.; Castañeda-Orjuela, C. A.; Castelpietra, G.; Castro, F.; Causey, K.; Cederroth, C. R.; Cercy, K. M.; Cerin, E.; Chandan, J. S.; Chang, K. L.; Charlson, F. J.; Chattu, V. K.; Chaturvedi, S.; Cherbuin, N.; Chimed-Ochir, O.; Cho, D. Y.; Choi, J. Y. J.; Christensen, H.; Chu, D. T.; Chung, M. T.; Chung, S. C.; Cicuttini, F. M.; Ciobanu, L. G.; Cirillo, M.; Classen, T. K. D.; Cohen, A. J.; Compton, K.; Cooper, O. R.; Costa, V. M.; Cousin, E.; Cowden, R. G.; Cross, D. H.; Cruz, J. A.; Dahlawi, S. M. A.; Damasceno, A. A. M.; Damiani, G.; Dandona, L.; Dandona, R.; Dangel, W. J.; Danielsson, A. K.; Dargan, P. I.; Darwesh, A. M.; Daryani, A.; Das, J. K.; Das Gupta, R.; das Neves, J.; Dávila-Cervantes, C. A.; Davitoiu, D. V.; De Leo, D.; Degenhardt, L.; DeLang, M.; Dellavalle, R. P.; Demeke, F. M.; Demoz, G. T.; Demsie, D. G.; Denova-Gutiérrez, E.; Dervenis, N.; Dhungana, G. P.; Dianatinasab, M.; Dias da Silva, D.; Diaz, D.; Dibaji Forooshani, Z. S.; Djalalinia, S.; Do, H. T.; Dokova, K.; Dorostkar, F.; Doshmangir, L.; Driscoll, T. R.; Duncan, B. B.; Duraes, A. R.; Eagan, A. W.; Edvardsson, D.; El Nahas, N.; El Sayed, I.; El Tantawi, M.; Elbarazi, I.; Elgendy, I. Y.; El-Jaafary, S. I.; Elyazar, I. R.; Emmons-Bell, S.; Erskine, H. E.; Eskandarieh, S.; Esmaeilnejad, S.; Esteghamati, A.; Estep, K.; Etemadi, A.; Etisso, A. E.; Fanzo, J.; Farahmand, M.; Fareed, M.; Faridnia, R.; Farioli, A.; Faro, A.; Faruque, M.; Farzadfar, F.; Fattahi, N.; Fazlzadeh, M.; Feigin, V. L.; Feldman, R.; Fereshtehnejad, S. M.; Fernandes, E.; Ferrara, G.; Ferrari, A. J.; Ferreira, M. L.; Filip, I.; Fischer, F.; Fisher, J. L.; Flor, L. S.; Foigt, N. A.; Folayan, M. O.; Fomenkov, A. A.; Force, L. M.; Foroutan, M.; Franklin, R. C.; Freitas, M.; Fu, W.; Fukumoto, T.; Furtado, J. M.; Gad, M. M.; Gakidou, E.; Gallus, S.; Garcia-Basteiro, A. L.; Gardner, W. M.; Geberemariyam, B. S.; Gebreslassie, A. A. A. A.; Geremew, A.; Gershberg Hayoon, A.; Gething, P. W.; Ghadimi, M.; Ghadiri, K.; Ghaffarifar, F.; Ghafourifard, M.; Ghamari, F.; Ghashghaee, A.; Ghiasvand, H.; Ghith, N.; Gholamian, A.; Ghosh, R.; Gill, P. S.; Ginindza, T. G. G.; Giussani, G.; Gnedovskaya, E. V.; Goharinezhad, S.; Gopalani, S. V.; Gorini, G.; Goudarzi, H.; Goulart, A. C.; Greaves, F.; Grivna, M.; Grosso, G.; Gubari, M. I. M.; Gugnani, H. C.; Guimarães, R. A.; Guled, R. A.; Guo, G.; Guo, Y.; Gupta, R.; Gupta, T.; Haddock, B.; Hafezi-Nejad, N.; Hafiz, A.; Haj-Mirzaian, A.; Haj-Mirzaian, A.; Hall, B. J.; Halvaei, I.; Hamadeh, R. R.; Hamidi, S.; Hammer, M. S.; Hankey, G. J.; Haririan, H.; Haro, J. M.; Hasaballah, A. I.; Hasan, M. M.; Hasanpoor, E.; Hashi, A.; Hassanipour, S.; Hassankhani, H.; Havmoeller, R. J.; Hay, S. I.; Hayat, K.; Heidari, G.; Heidari-Soureshjani, R.; Henrikson, H. J.; Herbert, M. E.; Herteliu, C.; Heydarpour, F.; Hird, T. R.; Hoek, H. W.; Holla, R.; Hoogar, P.; Hosgood, H. D.; Hossain, N.; Hosseini, M.; Hosseinzadeh, M.; Hostiuc, M.; Hostiuc, S.; Househ, M.; Hsairi, M.; Hsieh, V. C. r.; Hu, G.; Hu, K.; Huda, T. M.; Humayun, A.; Huynh, C. K.; Hwang, B. F.; Iannucci, V. C.; Ibitoye, S. E.; Ikeda, N.; Ikuta, K. S.; Ilesanmi, O. S.; Ilic, I. M.; Ilic, M. D.; Inbaraj, L. R.; Ippolito, H.; Iqbal, U.; Irvani, S. S. N.; Irvine, C. M. S.; Islam, M. M.; Islam, S. M. S.; Iso, H.; Ivers, R. Q.; Iwu, C. C. D.; Iwu, C. J.; Iyamu, I. O.; Jaafari, J.; Jacobsen, K. H.; Jafari, H.; Jafarinia, M.; Jahani, M. A.; Jakovljevic, M.; Jalilian, F.; James, S. L.; Janjani, H.; Javaheri, T.; Javidnia, J.; Jeemon, P.; Jenabi, E.; Jha, R. P.; Jha, V.; Ji, J. S.; Johansson, L.; John, O.; John-Akinola, Y. O.; Johnson, C. O.; Jonas, J. B.; Joukar, F.; Jozwiak, J. J.; Jürisson, M.; Kabir, A.; Kabir, Z.; Kalani, H.; Kalani, R.; Kalankesh, L. R.; Kalhor, R.; Kanchan, T.; Kapoor, N.; Karami Matin, B.; Karch, A.; Karim, M. A.; Kassa, G. M.; Katikireddi, S. V.; Kayode, G. A.; Kazemi Karyani, A.; Keiyoro, P. N.; Keller, C.; Kemmer, L.; Kendrick, P. J.; Khalid, N.; Khammarnia, M.; Khan, E. A.; Khan, M; Khatab, K.; Khater, M. M.; Khatib, M. N.; Khayamzadeh, M.; Khazaei, S.; Kieling, C.; Kim, Y. J.; Kimokoti, R. W.; Kisa, A.; Kisa, S.; Kivimäki, M.; Knibbs, L. D.; Knudsen, A. K. S.; Kocarnik, J. M.; Kochhar, S.; Kopec, J. A.; Korshunov, V. A.; Koul, P. A.; Koyanagi, A.; Kraemer, M. U. G.; Krishan, K.; Krohn, K. J.; Kromhout, H.; Kuate Defo, B.; Kumar, G. A.; Kumar, V.; Kurmi, O. P.; Kusuma, D.; La Vecchia, C.; Lacey, B.; Lal, D. K.; Lalloo, R.; Lallukka, T.; Lami, F. H.; Landires, I.; Lang, J. J.; Langan, S. M.; Larsson, A. O.; Lasrado, S.; Lauriola, P.; Lazarus, J. V.; Lee, P. H.; Lee, S. W. H.; LeGrand, K. E.; Leigh, J.; Leonardi, M.; Lescinsky, H.; Leung, J.; Levi, M.; Li, S.; Lim, L. L.; Linn, S.; Liu, S.; Liu, S.; Liu, Y.; Lo, J.; Lopez, A. D.; Lopez, J. C. F.; Lopukhov, P. D.; Lorkowski, S.; Lotufo, P. A.; Lu, A.; Lugo, A.; Maddison, E. R.; Mahasha, P. W.; Mahdavi, M. M.; Mahmoudi, M.; Majeed, A.; Maleki, A.; Maleki, S.; Malekzadeh, R.; Malta, D. C.; Mamun, A. A.; Manda, A. L.; Manguerra, H.; Mansour-Ghanaei, F.; Mansouri, B.; Mansournia, M. A.; Mantilla Herrera, A. M.;

Maravilla, J. C.; Marks, A.; Martin, R. V.; Martini, S.; Martins-Melo, F. R.; Masaka, A.; Masoumi, S. Z.; Mathur, M. R.; Matsushita, K.; Maulik, P. K.; McAlinden, C.; McGrath, J. J.; McKee, M.; Mehndiratta, M. M.; Mehri, F.; Mehta, K. M.; Memish, Z. A.; Mendoza, W.; Menezes, R. G.; Mengesha, E. W.; Mereke, A.; Mereta, S. T.; Meretoja, A.; Meretoja, T. J.; Mestrovic, T.; Miazgowski, B.; Miazgowski, T; Michalek, I. M.; Miller, T. R.; Mills, E. J.; Mini, G.; Miri, M.; Mirica, A.; Mirrakhimov, E. M.; Mirzaei, H.; Mirzaei, M.; Mirzaei, R.; Mirzaei-Alavijeh, M.; Misganaw, A. T.; Mithra, P.; Moazen, B.; Mohammad, D. K.; Mohammad, Y.; Mohammad Gholi Mezerji, N.; Mohammadian-Hafshejani, A.; Mohammadifard, N.; Mohammadpourhodki, R.; Mohammed, A. S.; Mohammed, H.; Mohammed, J. A.; Mohammed, S.; Mokdad, A. H.; Molokhia, M.; Monasta, L.; Mooney, M. D.; Moradi, G.; Moradi, M. D.; Moradi-Lakeh, M.; Moradzadeh, R.; Moraga, P.; Morawska, L.; Morgado-da-Costa, J.; Morrison, S. D.; Mosapour, A.; Mosser, J. F.; Mouodi, S.; Mousavi, S. M.; Mousavi Khaneghah, A.; Mueller, U. O.; Mukhopadhyay, S.; Mullany, E. C.; Musa, K. I.; Muthupandian, S.; Nabhan, A. F.; Naderi, M.; Nagarajan, A. J.; Nagel, G.; Naghavi, M.; Naghshtabrizi, B.; Naimzada, M. D.; Najafi, F.; Nangia, V.; Nansseu, J. R.; Naserbakht, M.; Nayak, V. C.; Negoi, I.; Ngunjiri, J. W.; Nguyen, C. T.; Nguyen, H. L. T.; Nguyen, M.; Nigatu, Y. T.; Nikbakhsh, R.; Nixon, M. R.; Nnaji, C. A.; Nomura, S.; Norrving, B.; Noubiap, J. J.; Nowak, C.; Nunez-Samudio, V.; Oţoiu, A.; Oancea, B.; Odell, C. M.; Ogbo, F. A.; Oh, I. H.; Okunga, E. W.; Oladnabi, M.; Olagunju, A. T.; Olusanya, B. O.; Olusanya, J. O.; Omer, M. O.; Ong, K. L.; Onwujekwe, O. E.; Orpana, H. M.; Ortiz, A.; Osarenotor, O.; Osei, F. B.; Ostroff, S. M.; Otstavnov, N.; Otstavnov, S. S.; Øverland, S.; Owolabi, M. O.; Owolabi, S.; Padubidri, J. R.; Palladino, R.; Panda-Jonas, S.; Pandey, A.; Parry, C. D. H.; Pasovic, M.; Pasupula, D. K.; Patel, S. K.; Pathak, M.; Patten, S. B.; Patton, G. C.; Pazoki Toroudi, H.; Peden, A. E.; Pennini, A.; Pepito, V. C. F.; Peprah, E. K.; Pereira, D. M.; Pesudovs, K.; Pham, H. Q.; Phillips, M. R.; Piccinelli, C.; Pilz, T. M.; Piradov, M. A.; Pirsaheb, M.; Plass, D.; Polinder, S.; Polkinghorne, K. R.; Pond, C. D.; Postma, M. J.; Pourjafar, H.; Pourmalek, F.; Poznańska, A.; Prada, S. I.; Prakash, V.; Pribadi, D. R. A.; Pupillo, E.; Quazi Syed, Z.; Rabiee, M.; Rabiee, N.; Radfar, A.; Rafiee, A.; Raggi, A.; Rahman, M. A.; Rajabpour-Sanati, A.; Rajati, F.; Rakovac, I.; Ram, P.; Ramezanzadeh, K.; Ranabhat, C. L.; Rao, P. C.; Rao, S. J.; Rashedi, V.; Rathi, P.; Rawaf, D. L.; Rawaf, S.; Rawal, L.; Rawassizadeh, R.; Rawat, R.; Razo, C.; Redford, S. B.; Reiner, R. C., Jr; Reitsma, M. B.; Remuzzi, G.; Renjith, V.; Renzaho, A. M. N.; Resnikoff, S.; Rezaei, N.; Rezaei, N.; Rezapour, A.; Rhinehart, P. A.; Riahi, S. M.; Ribeiro, D. C.; Ribeiro, D.; Rickard, J.; Rivera, J. A.; Roberts, N. L. S.; Rodríguez-Ramírez, S.; Roever, L.; Ronfani, L.; Room, R.; Roshandel, G.; Roth, G. A.; Rothenbacher, D.; Rubagotti, E.; Rwegerera, G. M.; Sabour, S.; Sachdev, P. S.; Saddik, B.; Sadeghi, E.; Sadeghi, M.; Saeedi, R.; Saeedi Moghaddam, S.; Safari, Y.; Safi, S.; Safiri, S.; Sagar, R.; Sahebkar, A.; Sajadi, S. M.; Salam, N.; Salamati, P.; Salem, H.; Salem, M. R. R.; Salimzadeh, H.; Salman, O. M.; Salomon, J. A.; Samad, Z.; Samadi Kafil, H.; Sambala, E. Z.; Samy, A. M.; Sanabria, J.; Sánchez-Pimienta, T. G.; Santomauro, D. F.; Santos, I. S.; Santos, J. V.; Santric-Milicevic, M. M.; Saraswathy, S. Y. I.; Sarmiento-Suárez, R.; Sarrafzadegan, N.; Sartorius, B.; Sarveazad, A.; Sathian, B.; Sathish, T.; Sattin, D.; Saxena, S.; Schaeffer, L. E.; Schiavolin, S.; Schlaich, M. P.; Schmidt, M. I.; Schutte, A. E.; Schwebel, D. C.; Schwendicke, F.; Senbeta, A. M.; Senthilkumaran, S.; Sepanlou, S. G.; Serdar, B.; Serre, M. L.; Shadid, J.; Shafaat, O.; Shahabi, S.; Shaheen, A. A.; Shaikh, M. A.; Shalash, A. S.; Shams-Beyranvand, M.; Shamsizadeh, M.; Sharafi, K.; Sheikh, A.; Sheikhtaheri, A.; Shibuya, K.; Shield, K. D.; Shigematsu, M.; Shin, J. I.; Shin, M. J.; Shiri, R.; Shirkoohi, R.; Shuval, K.; Siabani, S.; Sierpinski, R.; Sigfusdottir, I. D.; Sigurvinsdottir, R.; Silva, J. P.; Simpson, K. E.; Singh, J. A.; Singh, P.; Skiadaresi, E.; Skou, S. T.; Skryabin, V. Y.; Smith, E. U. R.; Soheili, A.; Soltani, S.; Soofi, M.; Sorensen, R. J. D.; Soriano, J. B.; Sorrie, M. B.; Soshnikov, S.; Soyiri, I. N.; Spencer, C. N.; Spotin, A.; Sreeramareddy, C. T.; Srinivasan, V.; Stanaway, J. D.; Stein, C.; Stein, D. J.; Steiner, C.; Stockfelt, L.; Stokes, M. A.; Straif, K.; Stubbs, J. L.; Sufiyan, M. B.; Suleria, H. A. R.; Suliankatchi Abdulkader, R.; Sulo, G.; Sultan, I.; Szumowski, Ł.; Tabarés-Seisdedos, R.; Tabb, K. M.; Tabuchi, T.; Taherkhani, A.; Tajdini, M.; Takahashi, K.; Takala, J. S.; Tamiru, A. T.; Taveira, N.; Tehrani-Banihashemi, A.; Temsah, M. H.; Tesema, G. A.; Tessema, Z. T.; Thurston, G. D.; Titova, M. V.; Tohidinik, H. R.; Tonelli, M.; Topor-Madry, R.; Topouzis, F.; Torre, A. E.; Touvier, M.; Tovani-Palone, M. R. R.; Tran, B. X.; Travillian, R.; Tsatsakis, A.; Tudor Car, L.; Tyrovolas, S.; Uddin, R.; Umeokonkwo, C. D.; Unnikrishnan, B.; Upadhyay, E.; Vacante, M.; Valdez, P. R.; van Donkelaar, A.; Vasankari, T. J.; Vasseghian, Y.; Veisani, Y.; Venketasubramanian, N.; Violante, F. S.; Vlassov, V.; Vollset, S. E.; Vos, T.; Vukovic, R.; Waheed, Y.; Wallin, M. T.; Wang, Y.; Wang, Y. P.; Watson, A.; Wei, J.; Wei, M. Y. W.; Weintraub, R. G.; Weiss, J.; Werdecker, A.; West, J. J.; Westerman, R.; Whisnant, J. L.; Whiteford, H. A.; Wiens, K. E.; Wolfe, C. D. A.; Wozniak, S. S.; Wu, A. M.; Wu, J.; Wulf Hanson, S.; Xu, G.; Xu, R.; Yadgir, S.; Yahyazadeh Jabbari, S. H.; Yamagishi, K.; Yaminfirooz, M.; Yano, Y.; Yaya, S.; Yazdi-Feyzabadi, V.; Yeheyis, T. Y.; Yilgwan, C. S.; Yilma, M. T.; Yip, P.; Yonemoto, N.; Younis, M. Z.; Younker, T. P.; Yousefi, B.; Yousefi, Z.; Yousefinezhadi, T.; Yousuf, A. Y.; Yu, C.; Yusefzadeh, H.; Zahirian Moghadam, T.; Zamani, M.; Zamanian, M.; Zandian, H.; Zastrozhin, M. S.; Zhang, Y.; Zhang, Z. J.; Zhao, J. T.; Zhao, X. J. G.; Zhao, Y.; Zhou, M.; Ziapour, A.; Zimsen, S. R. M.; Brauer, M.; Afshin, A.; Lim, S. S.; Lim, A. Global Burden of 87 Risk Factors in 204 Countries and Territories, 1990−2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *Lancet* **2020**, *396* (10258), 1223−1249.

(8) Feng, Z.; Xu, Y.; Kobayashi, K.; Dai, L.; Zhang, T.; Agathokleous, E.; Calatayud, V.; Paoletti, E.; Mukherjee, A.; Agrawal, M.; Park, R. J.; Oak, Y. J.; Yue, X. Ozone Pollution Threatens the Production of Major Staple Crops in East Asia. *Nat. Food* **2022**, *3* (1), 47−56.

(9) Agathokleous, E.; Feng, Z.; Oksanen, E.; Sicard, P.; Wang, Q.; Saitanis, C. J.; Araminiene, V.; Blande, J. D.; Hayes, F.; Calatayud, V.; Domingos, M.; Veresoglou, S. D.; Peñuelas, J.; Wardle, D. A.; De Marco, A.; Li, Z.; Harmens, H.; Yuan, X.; Vitale, M.; Paoletti, E. Ozone Affects Plant, Insect, and Soil Microbial Communities: A Threat to Terrestrial Ecosystems and Biodiversity. *Sci. Adv.* **2020**, *6* (33), No. eabc1176.

(10) Jacob, D. J. Heterogeneous Chemistry and Tropospheric Ozone. *Atmos. Environ.* **2000**, *34* (12−14), 2131−2159.

(11) Li, K.; Jacob, D. J.; Liao, H.; Shen, L.; Zhang, Q.; Bates, K. H. Anthropogenic Drivers of 2013−2017 Trends in Summer Surface Ozone in China. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116* (2), 422−427.

(12) Lu, X.; Zhang, L.; Chen, Y.; Zhou, M.; Zheng, B.; Li, K.; Liu, Y.; Lin, J.; Fu, T. M.; Zhang, Q. Exploring 2016−2017 Surface Ozone Pollution over China: Source Contributions and Meteorological Influences. *Atmos. Chem. Phys.* **2019**, *19* (12), 8339−8361.

(13) Shindell, D. T.; Rind, D.; Lonergan, P. Increased Polar Stratospheric Ozone Losses and Delayed Eventual Recovery Owing to Increasing Greenhouse-Gas Concentrations. *Nature.* **1998**, *392* (6676), 589−592.

(14) Wang, N.; Wang, H.; Huang, X.; Chen, X.; Zou, Y.; Deng, T.; Li, T.; Lyu, X.; Yang, F. Extreme Weather Exacerbates Ozone Pollution in the Pearl River Delta, China: Role of Natural Processes. *Atmos. Chem. Phys.* **2024**, *24* (2), 1559−1570.

(15) Lyu, X.; Li, K.; Guo, H.; Morawska, L.; Zhou, B.; Zeren, Y.; Jiang, F.; Chen, C.; Goldstein, A. H.; Xu, X.; Wang, T.; Lu, X.; Zhu, T.; Querol, X.; Chatani, S.; Latif, M. T.; Schuch, D.; Sinha, V.; Kumar, P.; Mullins, B.; Seguel, R.; Shao, M.; Xue, L.; Wang, N.; Chen, J.; Gao, J.; Chai, F.; Simpson, I.; Sinha, B.; Blake, D. R. A Synergistic Ozone-Climate Control to Address Emerging Ozone Pollution Challenges. *One Earth.* **2023**, *6* (8), 964−977.

(16) Xue, L. K.; Wang, T.; Gao, J.; Ding, A. J.; Zhou, X. H.; Blake, D. R.; Wang, X. F.; Saunders, S. M.; Fan, S. J.; Zuo, H. C.; Zhang, Q. Z.; Wang, W. X. Ground-Level Ozone in Four Chinese Cities: Precursors, Regional Transport and Heterogeneous Processes. *Atmos. Chem. Phys.* **2014**, *14* (23), 13175−13188.

(17) Wang, Y.; Zhao, Y.; Liu, Y.; Jiang, Y.; Zheng, B.; Xing, J.; Liu, Y.; Wang, S.; Nielsen, C. P. Sustained Emission Reductions Have Restrained the Ozone Pollution over China. *Nat. Geosci.* **2023**, *16* (11), 967−974.

(18) Cheng, Y.; Peng, Y.; Cao, L. M.; Huang, X. F.; He, L. Y. Identifying the Geospatial Relationship of Surface Ozone Pollution in

China: Implications for Key Pollution Control Regions. *Sci. Total Environ.* **2024**, *930*, 172763.
(19) Wang, N.; Huang, X.; Xu, J.; Wang, T.; Tan, Z. m.; Ding, A. Typhoon-Boosted Biogenic Emission Aggravates Cross-Regional Ozone Pollution in China. *Sci. Adv.* **2022**, *8* (2), No. eabl6166.
(20) Lefohn, A. S.; Malley, C. S.; Simon, H.; Wells, B.; Xu, X.; Zhang, L.; Wang, T. Responses of Human Health and Vegetation Exposure Metrics to Changes in Ozone Concentration Distributions in the European Union, United States, and China. *Atmos. Environ.* **2017**, *152*, 123−145.
(21) Hogrefe, C.; Liu, P.; Pouliot, G.; Mathur, R.; Roselle, S.; Flemming, J.; Lin, M.; Park, R. J. Impacts of Different Characterizations of Large-Scale Background on Simulated Regional-Scale Ozone over the Continental United States. *Atmos. Chem. Phys.* **2018**, *18* (5), 3839−3864.
(22) Qin, M. M.; She, Y. L.; Wang, M.; Wang, H. L.; Chang, Y. H.; Tan, Z. F.; An, J. Y.; Huang, J.; Yuan, Z. B.; Lu, J.; Wang, Q.; Liu, C.; Liu, Z. X.; Xie, X. D.; Li, J. Y.; Liao, H.; Pye, H. O. T.; Huang, C.; Guo, S.; Hu, M.; Zhang, Y. H.; Jacob, D. J.; Hu, J. L. Increased Urban Ozone in Heatwaves Due to Temperature-Induced Emissions of Anthropogenic Volatile Organic Compounds. *Nat. Geosci.* **2025**, *18* (1), 50−56.
(23) Cheng, Y.; He, L. Y.; Huang, X. F. Development of a High-Performance Machine Learning Model to Predict Ground Ozone Pollution in Typical Cities of China. *J. Environ. Manage.* **2021**, *299*, 113670.
(24) Lyu, X.; Li, H.; Lee, S.-C.; Xiong, E.; Guo, H.; Wang, T.; de Gouw, J. Significant Biogenic Source of Oxygenated Volatile Organic Compounds and the Impacts on Photochemistry at a Regional Background Site in South China. *Environ. Sci. Technol.* **2024**, *58* (45), 20081−20090.
(25) Li, X.; Liu, J.; Mauzerall, D. L.; Emmons, L. K.; Walters, S.; Horowitz, L. W.; Tao, S. Effects of Trans-Eurasian Transport of Air Pollutants on Surface Ozone Concentrations over Western China. *J. Geophys. Res. Atmos.* **2014**, *119* (21), 12338−12354.
(26) Wang, W.; Li, X.; Cheng, Y.; Parrish, D. D.; Ni, R.; Tan, Z.; Liu, Y.; Lu, S.; Wu, Y.; Chen, S.; Lu, K.; Hu, M.; Zeng, L.; Shao, M.; Huang, C.; Tian, X.; Leung, K. M.; Chen, L.; Fan, M.; Zhang, Q.; Rohrer, F.; Wahner, A.; Pöschl, U.; Su, H.; Zhang, Y. Ozone Pollution Mitigation Strategy Informed by Long-Term Trends of Atmospheric Oxidation Capacity. *Nat. Geosci.* **2024**, *17* (1), 20−25.
(27) Zhou, X.; Li, M.; Huang, X.; Liu, T.; Zhang, H.; Qi, X.; Wang, Z.; Qin, Y.; Geng, G.; Wang, J.; Chi, X.; Ding, A. Urban Meteorology−Chemistry Coupling in Compound Heat−Ozone Extremes. *Nat. Cities.* **2025**, *2* (9), 847−856.
(28) Zhu, B.; Huang, X. F.; Xia, S. Y.; Lin, L. L.; Cheng, Y.; He, L. Y. Biomass-Burning Emissions Could Significantly Enhance the Atmospheric Oxidizing Capacity in Continental Air Pollution. *Environ. Pollut.* **2021**, *285*, 117523.
(29) Zhan, J.; Ma, W.; Song, B.; Wang, Z.; Bao, X.; Xie, H.-B.; Chu, B.; He, H.; Jiang, T.; Liu, Y. The Contribution of Industrial Emissions to Ozone Pollution: Identified Using Ozone Formation Path Tracing Approach. *npj Clim. Atmos. Sci.* **2023**, *6* (1), 37.
(30) Xiong, Y.; Chai, J.; Mao, H.; Mariscal, N.; Yacovitch, T.; Lerner, B.; Majluf, F.; Canagaratna, M.; Olaguer, E. P.; Huang, Y. Examining the Summertime Ozone Formation Regime in Southeast Michigan Using Moose Ground-Based Hcho/NO2 Measurements and F0AM Box Model. *J. Geophys. Res. Atmos.* **2023**, *128* (19), No. e2023JD038943.
(31) Cheng, Y.; Zhu, Q.; Peng, Y.; Huang, X. F.; He, L. Y. Multiple Strategies for a Novel Hybrid Forecasting Algorithm of Ozone Based on Data-Driven Models. *J. Clean. Prod.* **2021**, *326*, 129451.
(32) Jiang, Z.; Tang, M.-X.; He, L.; Li, J.-H.; Hong, Y.; He, L.-Y.; Huang, X.-F. Unveiling the HONO Offsetting Effect: Rethinking NOx Emission Controls During Urban Ozone Pollution Episodes. *Environ. Sci. Technol.* **2025**, *59* (41), 22050−22059.
(33) Callaghan, M.; Schleussner, C.-F.; Nath, S.; Lejeune, Q.; Knutson, T. R.; Reichstein, M.; Hansen, G.; Theokritoff, E.; Andrijevic, M.; Brecha, R. J.; Hegarty, M.; Jones, C.; Lee, K.; Lucas, A.; van Maanen, N.; Menke, I.; Pfleiderer, P.; Yesil, B.; Minx, J. C. Machine-Learning-Based Evidence and Attribution Mapping of 100,000 Climate Impact Studies. *Nat. Clim. Change* **2021**, *11* (11), 966−972.
(34) Mai, Z.; Shen, H.; Zhang, A.; Sun, H. Z.; Zheng, L.; Guo, J.; Liu, C.; Chen, Y.; Wang, C.; Ye, J.; Zhu, L.; Fu, T.-M.; Yang, X.; Tao, S. Convolutional Neural Networks Facilitate Process Understanding of Megacity Ozone Temporal Variability. *Environ. Sci. Technol.* **2024**, *58* (35), 15691−15701.
(35) Zhang, A.; Fu, T.-M.; Feng, X.; Guo, J.; Liu, C.; Chen, J.; Mo, J.; Zhang, X.; Wang, X.; Wu, W.; Hou, Y.; Yang, H.; Lu, C. Deep Learning-Based Ensemble Forecasts and Predictability Assessments for Surface Ozone Pollution. *Geophys. Res. Lett.* **2023**, *50* (8), No. e2022GL102611.
(36) Ghahremanloo, M.; Choi, Y.; Lops, Y. Deep Learning Mapping of Surface MDA8 Ozone: The Impact of Predictor Variables on Ozone Levels over the Contiguous United States. *Environ. Pollut.* **2023**, *326*, 121508.
(37) Dimri, R.; Choi, Y.; Salman, A. K.; Park, J.; Singh, D. Agatnet: An Adaptive Graph Attention Network for Bias Correction of CMAQ-Forecasted PM2.5 Concentrations over South Korea. *JGR: Machine Learning and Computation* **2024**, *1* (3), No. e2024JH000244.
(38) Hao, P.; Wang, Y.; Zheng, Z.; Mei, H.; DeMarsh, K. E.; Moo, Z.; Wang, H.; Pan, S.; Zhang, X. VOC Observations Improve Transferable Ozone Prediction Using Interpretable Machine Learning. *npj Clean Air.* **2026**, *2* (1), 53.
(39) Cheng, Y.; Huang, X. F.; Peng, Y.; Tang, M. X.; Zhu, B.; Xia, S. Y.; He, L. Y. A Novel Machine Learning Method for Evaluating the Impact of Emission Sources on Ozone Formation. *Environ. Pollut.* **2023**, *316*, 120685.
(40) Nelson, D.; Choi, Y.; Sadeghi, B.; Yeganeh, A. K.; Ghahremanloo, M.; Park, J. A Comprehensive Approach Combining Positive Matrix Factorization Modeling, Meteorology, and Machine Learning for Source Apportionment of Surface Ozone Precursors: Underlying Factors Contributing to Ozone Formation in Houston, Texas. *Environ. Pollut.* **2023**, *334*, 122223.
(41) Wang, C.; Qin, X.; Zhang, Y.; Liu, D.; Tao, W.; Wang, M.; Zhang, S.; Yang, J.; Du, J.; Cui, S.; Gu, D.; Sun, Y.; Lv, C. Machine Learning Integrated PMF Model Reveals Influencing Factors of Ozone Pollution in a Coal Chemical Industry City at the Jiangsu-Shandong-Henan-Anhui Boundary. *Atmos. Environ.* **2025**, *342*, 120916.
(42) Wang, W.; Parrish, D. D.; Wang, S.; Bao, F.; Ni, R.; Li, X.; Yang, S.; Wang, H.; Cheng, Y.; Su, H. Long-Term Trend of Ozone Pollution in China During 2014−2020: Distinct Seasonal and Spatial Characteristics and Ozone Sensitivity. *Atmos. Chem. Phys.* **2022**, *22* (13), 8935−8949.
(43) Xue, L.; Wang, T.; Louie, P. K. K.; Luk, C. W. Y.; Blake, D. R.; Xu, Z. Increasing External Effects Negate Local Efforts to Control Ozone Air Pollution: A Case Study of Hong Kong and Implications for Other Chinese Cities. *Environ. Sci. Technol.* **2014**, *48* (18), 10769−10775.
(44) Coggon, M. M.; Stockwell, C. E.; Claflin, M. S.; Pfannerstill, E. Y.; Xu, L.; Gilman, J. B.; Marcantonio, J.; Cao, C.; Bates, K.; Gkatzelis, G. I.; Lamplugh, A.; Katz, E. F.; Arata, C.; Apel, E. C.; Hornbrook, R. S.; Piel, F.; Majluf, F.; Blake, D. R.; Wisthaler, A.; Canagaratna, M.; Lerner, B. M.; Goldstein, A. H.; Mak, J. E.; Warneke, C. Identifying and Correcting Interferences to PTR-ToF-MS Measurements of Isoprene and Other Urban Volatile Organic Compounds. *Atmos. Meas. Technol.* **2024**, *17* (2), 801−825.
(45) Li, Z. J.; He, L. Y.; Ma, H. N.; Peng, X.; Tang, M. X.; Du, K.; Huang, X. F. Sources of Atmospheric Oxygenated Volatile Organic Compounds in Different Air Masses in Shenzhen, China. *Environ. Pollut.* **2024**, *340*, 122871.
(46) Huang, X. F.; Cao, L. M.; Tian, X. D.; Zhu, Q.; Saikawa, E.; Lin, L. L.; Cheng, Y.; He, L. Y.; Hu, M.; Zhang, Y. H.; Lu, K. D.; Liu, Y. H.; Daellenbach, K.; Slowik, J. G.; Tang, Q.; Zou, Q. L.; Sun, X.; Xu, B. Y.; Jiang, L.; Shen, Y. M.; Ng, N. L.; Prévôt, A. S. H. Critical Role of Simultaneous Reduction of Atmospheric Odd Oxygen for Winter Haze Mitigation. *Environ. Sci. Technol.* **2021**, *55* (17), 11557−11567.
(47) Cheng, Y.; Huang, X. F.; Peng, Y.; Cao, L. M.; Peng, X.; Wu, J.; He, L. Y. Enhancing Time Resolution of Ambient VOC Measurement Data by Machine Learning: From One-Hour to Five Minutes. *J. Geophys. Res. Atmos.* **2025**, *130* (5), No. e2024JD042477.

(48) Yang, Y.; Liu, B.; Hua, J.; Yang, T.; Dai, Q.; Wu, J.; Feng, Y.; Hopke, P. K. Global Review of Source Apportionment of Volatile Organic Compounds Based on Highly Time-Resolved Data from 2015 to 2021. *Environ. Int.* **2022**, *165*, 107330.
(49) Dai, Q.; Liu, B.; Bi, X.; Wu, J.; Liang, D.; Zhang, Y.; Feng, Y.; Hopke, P. K. Dispersion Normalized PMF Provides Insights into the Significant Changes in Source Contributions to PM2.5 after the COVID-19 Outbreak. *Environ. Sci. Technol.* **2020**, *54* (16), 9917−9927.
(50) Thomas, A.; Huff, A. K.; Hu, X.-M.; Zhang, F. Quantifying Uncertainties of Ground-Level Ozone within WRF-Chem Simulations in the Mid-Atlantic Region of the United States as a Response to Variability. *J. Adv. Model. Earth Syst.* **2019**, *11* (4), 1100−1116.
(51) Xu, D. N.; Yuan, Z. B.; Wang, M.; Zhao, K. H.; Liu, X. H.; Duan, Y. S.; Fu, Q. Y.; Wang, Q.; Jing, S. G.; Wang, H. L.; Zhao, X. Multi-Factor Reconciliation of Discrepancies in Ozone-Precursor Sensitivity Retrieved from Observation- and Emission-Based Models. *Environ. Int.* **2022**, *158*, 106952.
(52) Wolfe, G. M.; Marvin, M. R.; Roberts, S. J.; Travis, K. R.; Liao, J. The Framework for 0-D Atmospheric Modeling (F0AM) V3.1. *Geosci. Model Dev.* **2016**, *9* (9), 3309−3319.
(53) Hui, L.; Liu, X.; Tan, Q.; Feng, M.; An, J.; Qu, Y.; Zhang, Y.; Jiang, M. Characteristics, Source Apportionment and Contribution of VOCs to Ozone Formation in Wuhan, Central China. *Atmos. Environ.* **2018**, *192*, 55−71.
(54) Chen, X.; Ma, W.; Zheng, F.; Wang, Z.; Hua, C.; Li, Y.; Wu, J.; Li, B.; Jiang, J.; Yan, C.; Petäjä, T.; Bianchi, F.; Kerminen, V.-M.; Worsnop, D. R.; Liu, Y.; Xia, M.; Kulmala, M. Identifying Driving Factors of Atmospheric N2O5 with Machine Learning. *Environ. Sci. Technol.* **2024**, *58* (26), 11568−11577.
(55) Qiu, W.; Liu, Y.; Tan, Z.; Chen, X.; Lu, K.; Zhang, Y. Calculation of Maximum Incremental Reactivity Scales Based on Typical Megacities in China. *Chin. Sci. Bull.* **2020**, *65* (7), 610−621.
(56) Carter, W. P. L. Development of Ozone Reactivity Scales for Volatile Organic Compounds. *Air Waste* **1994**, *44* (7), 881−899.
(57) Atkinson, R.; Baulch, D. L.; Cox, R. A.; Crowley, J. N.; Hampson, R. F.; Hynes, R. G.; Jenkin, M. E.; Rossi, M. J.; Troe, J.; Subcommittee, I. Evaluated Kinetic and Photochemical Data for Atmospheric Chemistry: Volume II - Gas Phase Reactions of Organic Species. *Atmos. Chem. Phys.* **2006**, *6* (11), 3625−4055.
(58) Niu, Y. B.; Huang, X. F.; Wang, H. C.; Wang, S. X.; Lin, X. Y.; Chen, Y.; Zhu, B.; Zhu, Q.; He, L. Y. Effects of Nighttime Heterogeneous Reactions on the Formation of Secondary Aerosols and Ozone in the Pearl River Delta. *Chin. Sci. Bull.* **2022**, *67* (18), 2060−2068.

https://doi.org/10.1021/acs.est.6c12789
Environ. Sci. Technol. XXXX, XXX, XXX−XXX