Reviewer #2

General comments:

The manuscript presents an explainable deep transfer learning framework (PINN-ResMLP) to predict urban isoprene concentrations and attribute their variability across Chinese and international cities. It further explores long-term drivers in Hong Kong and London (1990–2023) and projects future isoprene and ozone responses under CMIP6/SSP scenarios, including NO_x-control sensitivity. The study fills an important gap: robust isoprene prediction without detailed local emission inventories or explicit chemistry, and interpretable attribution that links meteorology, greenspace, and traffic to observed and modeled trends. The approach is timely and impactful for urban air quality management in a warming climate. The integration of physics-informed constraints with transfer learning is a notable strength, as is the explicit discuss ability (SHAP-based) of model predictions. The Hong Kong–London contrast is compelling and policy-relevant.

Response: We sincerely thank Reviewer #2 for the thorough and positive evaluation of our work. We greatly appreciate the recognition of the novelty and impact of our study, particularly the explainable deep transfer learning framework (PINN-ResMLP), its ability to predict isoprene concentrations without detailed local emission inventories, and the interpretability provided by SHAP analysis. We also thank the reviewer for emphasizing the relevance of the Hong Kong–London comparison and the importance of integrating physics-informed constraints with transfer learning. The constructive feedback and supportive comments are highly encouraging and have helped us further clarify and refine the manuscript. Below, we provide our point-by-point responses to each comment, with our replies highlighted in blue and the corresponding revisions marked in red.

Specific comments:

1. Equations (5–6) use sign functions over partial derivatives. Please clarify how the gradients with respect to inputs are computed for monotonicity (e.g., via automatic *differentiation), and whether local monotonicity is enforced pointwise or globally. Also specify α and β values and sensitivity.*

Response: We thank the reviewer for this comment. In our PINN-ResMLP model, the gradients with respect to input features (VI and $BC_{traffic}$) are computed using PyTorch's automatic differentiation (autograd). Monotonicity is enforced pointwise, ensuring that the derivative of the model output with respect to each input at each data point satisfies the expected monotonicity constraint. During training, we experimented with different values for the loss weight coefficients α and β in the multi-loss function. Specifically, α was tested in [0.01, 0.1, 1] and β in [0.0001, 0.001, 0.01, 0.1]. We found that α = 1 and β = 0.0001 yielded the best performance in terms of capturing monotonicity without degrading predictive accuracy. Additionally, we corrected the implementation of the sign function in Equations (5) and (6). The corrected form is:

$$\mathcal{L}_{monotonicity} = \frac{1}{N}\sum_{i=1}^{N}\left[-\frac{sign\left(\frac{\partial ISOP}{\partial VI}\right) + sign\left(\frac{\partial ISOP}{\partial BC_{traffic}}\right)}{2}\right] \tag{5}$$

$$sign(\theta) = \begin{cases} -1, & \theta < 0 \\ 0, & \theta \geq 0 \end{cases} \tag{6}$$

2. $\mathcal{L}_{structure}$ is defined as sum of squared weights per layer ($W_i^2 + b_i^2$). Are there any architectural constraints (e.g., skip connections in ResMLP, layer widths) chosen to improve stability? Include a small ablation (ResMLP vs. ResMLP+PINN vs. PINN alone) if possible.

Response: To improve training stability, our ResMLP architecture incorporates residual (skip) connections and carefully chosen layer widths. These design choices help mitigate vanishing/exploding gradient issues and ensure robust convergence. The model is trained using the Adam optimizer with a weight decay term to further regularize the network and stabilize training. While we did not perform a formal ablation study, our results show that incorporating the PINN framework improves predictive accuracy and enforces physical constraints (Figure 2), highlighting the effectiveness of combining ResMLP with physics-informed constraints.

3. For overseas sites, you fine-tune on 70% and validate on 30%. Clarify whether the split preserves temporal ordering (to reduce leakage) and whether performance is robust to different splits (report variance across splits).

Response: For the overseas validation experiments, due to the limited data size (~1,000 samples), the data split did not preserve temporal ordering. Instead, we performed cross-validation to assess model robustness. The results consistently show that the PINN-ResMLP$_T$ framework improves predictive accuracy across different sites, demonstrating the stable performance of the model.

4. The authors showed that WRF-Chem performed poorly in isoprene simulations. Provide configuration details (chemistry mechanism, emissions, resolution, boundary conditions) and whether the MEGAN parameterization and land-use were tuned to urban greenspace. This contextualizes the performance gap and its causes (e.g., grid dilution, canopy-scale processes).

Response: In this study, we used the Weather Research and Forecasting model with Chemistry (WRF-Chem, version 3.7) to simulate urban isoprene concentrations. Meteorological initial and lateral boundary conditions were provided by the NCEP FNL dataset at 1° × 1° resolution, and Four-Dimensional Data Assimilation (FDDA) was applied to improve the meteorological fields. The Noah land surface model and the MM5 Monin–Obukhov surface layer scheme were used to represent land–atmosphere exchange processes, while the planetary boundary layer was simulated using the YSU scheme. Gas-phase chemistry and aerosol processes were represented using the CBMZ mechanism and the MOSAIC module, respectively. Biogenic VOC emissions were calculated using the Model of Emissions of Gases and Aerosols from Nature (MEGAN v2.1; Guenther et al., 2012). The vegetation-related static inputs were updated using MODIS PFT (MCD12Q1) and LAI (MCD15A2H) products. Anthropogenic emissions were obtained from the updated 2020-based MEIC inventory for regions within China and the MIX inventory (Li et al., 2017) for regions outside China, both at 0.25° × 0.25° resolution and including major emission sectors (transportation, industry, power plants, residential, and agriculture). These configurations of WRF-Chem have been successfully applied in our previous studies (Huang et al., 2020; Huang et al., 2021; Wang et al., 2021), indicating a reliable performance across China.

We note that the MEGAN and land-use datasets used in this study could not capture BVOC emissions from urban greenspace (which requires high-resolution data, e.g., 10 m × 10 m) or canopy-scale processes, as the MODIS satellite product is limited to 500 m × 500 m. Combined with the relatively coarse model resolution (WRF-Chem grid resolution: 25 km × 25 km), these limitations may lead to grid dilution of urban vegetation signals and underestimation of peak isoprene emissions, helping to explain the performance gap observed between WRF-Chem and our PINN-ResMLP framework.

Reference:

Huang, X., Ding, A., Wang, Z., et al.: Amplified transboundary transport of haze by aerosol－boundary layer interaction in China, Nature Geoscience, 13, 428-434, 10.1038/s41561-020-0583-4, 2020.

Huang, X., Ding, A. J., Gao, J., et al.: Enhanced secondary pollution offset reduction of primary emissions during COVID-19 lockdown in China, National Science Review, 8, 10.1093/nsr/nwaa137, 2021.

Wang, N., Xu, J., Pei, C., et al.: Air Quality During COVID-19 Lockdown in the Yangtze River Delta and the Pearl River Delta: Two Different Responsive Mechanisms to Emission Reductions in China, Environ. Sci. Technol., 55, 5721-5730, 10.1021/acs.est.0c08383, 2021.

5. State whether SHAP is computed on the fine-tuned model per site, the background dataset used, and whether interaction SHAP was explored (temperature × radiation) to reflect coupled sensitivities.

Response: Thank you for the valuable suggestion. In our study, SHAP values for the Chinese sites were computed based on the pre-trained model, while for Hong Kong and London, SHAP values were computed on the fine-tuned neural network model for each site, using the corresponding site’s training dataset as the background dataset. Regarding interaction SHAP values (e.g., temperature × radiation), we did not calculate them in the current manuscript. The widely used DeepSHAP implementation in the official "shap" package (v0.46.0, based on DeepExplainer) only computes marginal (main-effect) SHAP values and does not provide exact pairwise or higher-order interaction terms. Exact SHAP interaction values, as implemented in TreeSHAP for tree-based models, are currently not available for deep neural networks in any mature, computationally tractable form. The fundamental reason is that tree models have discrete decision paths that allow precise attribution of output changes to arbitrary feature coalitions, whereas neural networks exhibit highly nonlinear, continuous interactions across all features simultaneously (Janzing et al., 2020; Zern et al., 2023). Although some existing approximation approaches (e.g., Integrated Hessians) can estimate interactions between paired features in neural networks (Janizek et al., 2021), in this study we focused on the main SHAP effects to identify the key drivers of isoprene variability. We agree that exploring feature interactions is important and plan to investigate this in future work using alternative methods.

Reference:

Janizek, J. D., Sturmfels, P., and Lee, S.-I.: Explaining explanations: axiomatic feature interactions for deep networks, 22, Article 104, 2021.

Janzing, D., Minorics, L., and Bloebaum, P.: Feature relevance quantification in explainable AI: A causal problem, Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research 2020.

Zern, A., Broelemann, K., and Kasneci, G.: Interventional SHAP Values and Interaction Values for Piecewise Linear Regression Trees, Proceedings of the AAAI Conference on Artificial Intelligence, 37, 11164-11173, 10.1609/aaai.v37i9.26322, 2023.

6. For future projections, please explicitly acknowledge that future greenspace, urban form, and anthropogenic emissions will also evolve.

Response: We thank the reviewer for this comment. In this work, we adopt a pseudo-global-warming (PGW)–based approach to isolate the effect of climate warming on isoprene emissions and the consequent O_3 responses. PGW directly impose selected changes (e.g. temperature changes) in the climate system onto a historical regional climate simulation by modifying the initial and boundary conditions (Brogli et al., 2023). Following this framework, our future simulations were designed to vary only temperature while holding other precursors and environmental drivers fixed, thereby quantifying the chemical sensitivity of isoprene and O_3 to warming alone. Accordingly, the future projections do not account for potential changes in greenspace, urban form, or other anthropogenic emissions. We acknowledge that these factors may evolve over time and that incorporating them could further refine future predictions. We have added a note in the revised manuscript to explicitly clarify this limitation.

Please see our revisions in lines 219-227 of the manuscript: "It is worth noting that the diurnal profiles of other O_3 precursors, such as VOCs and carbon monoxide, were kept unchanged throughout all the simulations. Meanwhile, our future projections are designed to isolate the chemical response of O_3 to changes in temperature and isoprene and do not explicitly incorporate potential future changes in greenspace, urban morphology, or other anthropogenic emissions. Although these factors are expected to evolve under urban development and climate mitigation pathways, the present analysis focuses on quantifying the impacts of climate warming on isoprene emissions and the consequent O_3 responses.**"**

Reference:

Brogli, R., Heim, C., Mensch, J., et al.: The pseudo-global-warming (PGW) approach: methodology, software package PGW4ERA5 v1.1, validation, and sensitivity analyses, Geosci. Model Dev., 16, 907-926, 10.5194/gmd-16-907-2023, 2023.

7. There are two "the" in line 280.

Response: The redundant "the" in line 280 has been removed in the revised manuscript.