

Review #1

This study develops a machine-learning framework to retrieve gap-free, 1-km² daily concentrations of six major pollutants across China and uses these datasets to investigate atmospheric responses to rapid emission perturbations. The results highlight strong pollutant-specific responses and nonlinear interactions among emissions, meteorology, and atmospheric chemistry during emission reduction and recovery periods. Overall, this work is well-conducted and suitable for publication after addressing the comments listed below.

Major comments:

1. The datasets used in this study span 2019–2022. However, emissions, meteorological conditions, and other influencing factors may vary substantially across different years. Therefore, it is important to assess whether the model performance remains stable over time, and year-specific validation for each pollutant is recommended.

Response: Thanks for the suggestion. We have conducted year-specific validation for each pollutant and added the results to the revised manuscript to assess the temporal stability of model performance, as described in the main text:

“Furthermore, year-specific validation demonstrates stable model performance over time, showing consistent sample-based CV-R² (RMSE) values during 2019–2022, ranging from 0.89–0.93 (12.51–20.21 $\mu\text{g m}^{-3}$) for PM₁₀, 0.93–0.94 (6.40–8.79 $\mu\text{g m}^{-3}$) for PM_{2.5}, 0.93–0.95 (10.11–13.59 $\mu\text{g m}^{-3}$) for O₃, 0.86–0.87 (5.80–7.19 $\mu\text{g m}^{-3}$) for NO₂, 0.76–0.81 (3.13–4.87 $\mu\text{g m}^{-3}$) for SO₂, and 0.75–0.82 (0.16–0.20 mg m^{-3}) for CO, indicating consistent reliability despite interannual variations in emissions and meteorology.”

Table S5. Year-specific overall estimation accuracy of daily air pollutants in China

Variable	Year	Slope	R ²	RMSE	MAE
PM ₁₀	2019	0.86	0.91	17.11	9.26
	2020	0.86	0.89	16.64	7.93
	2021	0.89	0.93	20.21	8.06
	2022	0.90	0.93	12.51	6.97
PM _{2.5}	2019	0.90	0.93	8.79	5.15
	2020	0.90	0.93	7.97	4.46
	2021	0.90	0.93	7.25	4.22
	2022	0.91	0.94	6.40	4.00
O ₃	2019	0.90	0.93	13.59	9.32
	2020	0.91	0.94	11.08	7.68
	2021	0.91	0.94	10.64	7.43
	2022	0.93	0.95	10.11	6.92
NO ₂	2019	0.82	0.86	7.19	5.17
	2020	0.83	0.87	6.35	4.54
	2021	0.83	0.87	6.10	4.42

SO ₂	2022	0.81	0.86	5.8	4.12
	2019	0.77	0.81	4.87	2.54
	2020	0.74	0.79	4.08	2.14
	2021	0.71	0.77	3.64	1.91
	2022	0.71	0.76	3.13	1.74
CO	2019	0.78	0.82	0.20	0.13
	2020	0.77	0.81	0.18	0.11
	2021	0.72	0.75	0.18	0.11
	2022	0.74	0.78	0.16	0.10

Moreover, the number and spatial distribution of monitoring stations for each year should be clearly described, as changes in station coverage may affect model performance and comparability.

Response: The number of monitoring stations shows a moderate increase over time, ranging from 1474 to 1896 (Fig. S4), while the spatial distribution remains broadly stable. This ensures that the interannual stability of model performance is not confounded by changes in sampling density or spatial coverage of the monitoring network. This has been clarified in the revised manuscript.

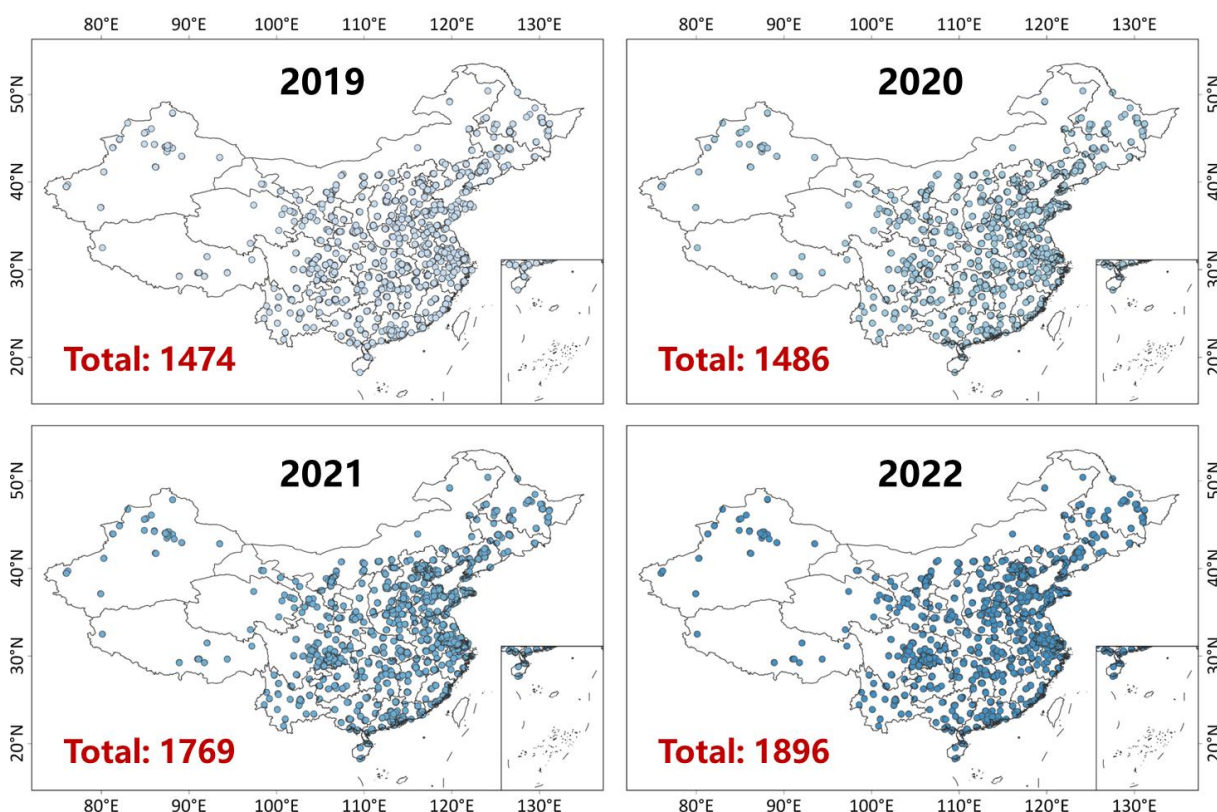


Figure S4. Spatial distribution of ground-based monitoring stations across China from 2019 to 2022. Red numbers shown in each panel indicate the total number of stations for the corresponding year.

2. On Page 7, Lines 208–210, the manuscript describes the conversion process for IAQI values.

It is unclear how values falling between category thresholds, i.e., 50–51, 100–101... are handled during the conversion. For example, are these values rounded, truncated, or excluded? Because such handling may influence the calculated IAQI values and consequently affect the subsequent analysis, the authors should clarify the exact conversion procedure. Please verify the original implementation method to ensure the calculation is performed correctly; otherwise, the analysis should be recomputed accordingly.

Response: We utilize raw floating-point pollutant concentrations (C_p) throughout the conversion process without any rounding, truncation, or exclusion. $IAQI_p$ is computed using standard piecewise linear interpolation between breakpoint concentrations, ensuring that values between category thresholds (e.g., 50–51, 100–101) are handled continuously. All intermediate values are retained in floating-point precision to avoid discretization bias, and IAQI values are capped at 500. We have clarified this in Section 2.2 by explicitly presenting category definitions using continuous inequality notation based on Table S2.

Minor comments:

1. Page 3, Line 66, If no statistical significance test was conducted, the term “significant” should not be used. Please carefully check the manuscript and revise such expressions accordingly.

Response: We agree and have replaced the terms “significant” with more appropriate descriptors (e.g., “substantial”, “pronounced”, or “notable”) in contexts where no statistical significance tests were performed.

2. For datasets used in the study, corresponding references should be provided if available, e., Page 5, Line 147-154, ABaCAS-EI; Page 15-16, Oxford Coronavirus Government Response Tracker (OxCGRT) stringency index...

Response: We have added the corresponding references for all datasets used in this study, including ABaCAS-EI and the Oxford Coronavirus Government Response Tracker (OxCGRT) stringency index, in the revised manuscript.

3. Page 4, Line 133-135, although the authors state that these functions and method can be useful for capturing daily variations and seasonal cycles in air pollution, however, the meaning of each function is not clearly explained. Please clarify.

Response: We have clarified the meaning and role of these functions in the revised manuscript as follows:

“A key improvement over the original deep forest model is the incorporation of spatiotemporal information. Spatial variability is represented using three spherical coordinate features derived from longitude (Lon) and latitude (Lat), i.e., $P_s \sim [S_1, S_2, S_3] = [\sin\left(2\pi \frac{Lon}{360}\right), \cos\left(2\pi \frac{Lon}{360}\right) \sin\left(2\pi \frac{Lat}{180}\right), \cos\left(2\pi \frac{Lon}{360}\right) \cos\left(2\pi \frac{Lat}{180}\right)]$, which capture spatial autocorrelation patterns. Temporal variability is represented using three helix-shaped trigonometric features based on the normalized day of the year ($\frac{DOY}{N}$), i.e., $P_t \sim [T_1, T_2, T_3] = [\frac{DOY}{N}, \cos\left(2\pi \frac{DOY}{N}\right),$

$\sin\left(2\pi\frac{DOY}{N}\right)$], which capture both linear temporal progression and periodic (seasonal) variations. Together, these features enable the model to better represent daily variations and seasonal cycles in air pollution.”

4. Page 6, Line 168, The manuscript refers to “variables for PM”? Please clarify whether this refers to PM₁₀, PM_{2.5}, or both, and provide a clear description.

Response: The term “PM” refers to both PM_{2.5} and PM₁₀, and we have clarified this in the revised manuscript

5. Page 7, Line 144-195, Could this section be made more concise? Some parts appear repetitive and may benefit from further streamlining.

Response: We have carefully revised this section to improve conciseness by removing repetitive descriptions and streamlining the text.

6. Page 8, Line 215, in equation (7), what is $IAQI_{p,l}$? Provide a clear definition.

Response: We apologize for the error in the formula. The term should be $IAQI_{p,j-l}$ and we have corrected this in the revised manuscript.

7. Page 9, Line 258, the word “changes” appears to contain a grammatical issue. Please carefully proofread the manuscript and consider professional language editing

Response: We have revised “air pollution conditions changes” to “changes in air pollution conditions”. In addition, the entire manuscript has been carefully proofread to correct grammatical issues and improve language clarity.

8. Page 10, Line 295, for the VIF calculation of PM₂₅ and O₃ predictors, some variables may exhibit collinearity. Were any variables screened or removed based on the VIF results?

Response: We assessed multicollinearity among all PM_{2.5} and O₃ predictors using the Variance Inflation Factor (VIF) across national, regional, and urban scales. A threshold of VIF < 10 was adopted. Although some variables showed relatively high VIF values (e.g., 9.86 for air temperature and 8.79 for surface pressure in the national O₃ model), all remained below this threshold. Therefore, no variables were excluded, and multicollinearity is unlikely to substantially affect the results. This has been clarified in Section 2.4.

9. Page 19, Line 493, the legend shows “2019 vs. 2019”, which appears to be an error. Please verify whether similar inconsistencies exist elsewhere in the manuscript.

Response: We have corrected the labeling error (it should be “2020 vs. 2019”) in the revised manuscript. In addition, we have carefully checked all figures and captions to ensure that no similar inconsistencies remain.

10. Several spelling errors were noted in the manuscript, Page 20, Line 513, “conditionss”? The author should carefully check the manuscript for spelling mistakes.

Response: We have corrected the spelling error and also carefully proofread the entire manuscript to eliminate similar mistakes.