Response to referee #2

Comments: Tropospheric ozone is a globally important air pollutant and a short-lived climate forcer, with substantial impacts on human health, climate change, and terrestrial ecosystems. Understanding the relationship between ozone concentration changes and their driving factors is essential for developing effective control strategies. This study utilizes ground-based observational data from the Chinese monitoring network during 2013-2023 and develops a machine-learning-based method to quantitatively disentangle the contributions of meteorological conditions and anthropogenic emissions. The analysis is further extended to evaluate the sensitivity of ozone to climate change. In addition, the authors employ satellite retrievals to explore the changes in precursor ratios and to diagnose the shifts in chemical regimes. The proposed analytical framework provides valuable insights and will be highly informative for future studies. Overall, the manuscript is clearly structured, well-designed, and well-written, and it fits well within the scope of ACP. I would recommend publication after the following issues are addressed:

Response: We sincerely thank the reviewer for the encouraging and insightful evaluation of our study. We are very grateful for the recognition of our analytical framework and its relevance to understanding ozone evolution and control strategies. Following the reviewer's valuable suggestions, we have carefully revised the manuscript to improve its clarity and completeness. All corresponding modifications have been incorporated into the revised version, with detailed point-by-point responses provided below.

Title suggestion: Consider revising the title to "Tracking surface ozone responses to clean air actions under a warming climate in China" for clarity and stronger alignment with the scope.

Response: We greatly appreciate the reviewer's constructive suggestion regarding the title. Following this advice, we have revised the title to:

"Tracking surface ozone responses to clean air actions under a warming climate in China using machine learning."

This revised title provides clearer expression and more accurately captures the methodological framework and research scope of our study.

Lines 52-54: This sentence requires additional references. In particular, the 2021 IPCC report should be cited and carefully verified.

Response: We appreciate the reviewer's helpful suggestion. In response, we have thoroughly reviewed the relevant literature and added the IPCC Sixth Assessment Report (AR6, 2021) as an authoritative reference to strengthen and substantiate the statement in lines 57–58. The corresponding text has also been carefully checked and revised to ensure both scientific accuracy and contextual consistency.

Line 81: Provide the full name of XGBoost when first introduced.

Response: Thank you for your suggestion. We have now provided the full name of XGBoost ("eXtreme Gradient Boosting") upon its first mention in line 85-86 to improve clarity and readability.

Line 89: Please remove the word "monsoon".

Response: Thank you for your suggestion. This word "monsoon" has been removed in the revised manuscript.

Line 109: The study develops an innovative machine-learning framework for attribution analysis, including an extension to climate change. This is a key contribution, but I suggest adding more technical details, such as a conceptual diagram of the methodology, to improve clarity and accessibility for readers.

Response: We sincerely appreciate the reviewer's positive recognition of our machine-learning-based analytical framework. To improve the clarity and accessibility of the methodology, we have added a conceptual diagram (Figure 1) that outlines the overall workflow of the framework and its application to trend attribution and climate change impact assessment. This addition visually summarizes the key analytical steps and enhances the reader's understanding of the underlying processes and logical structure of the study.

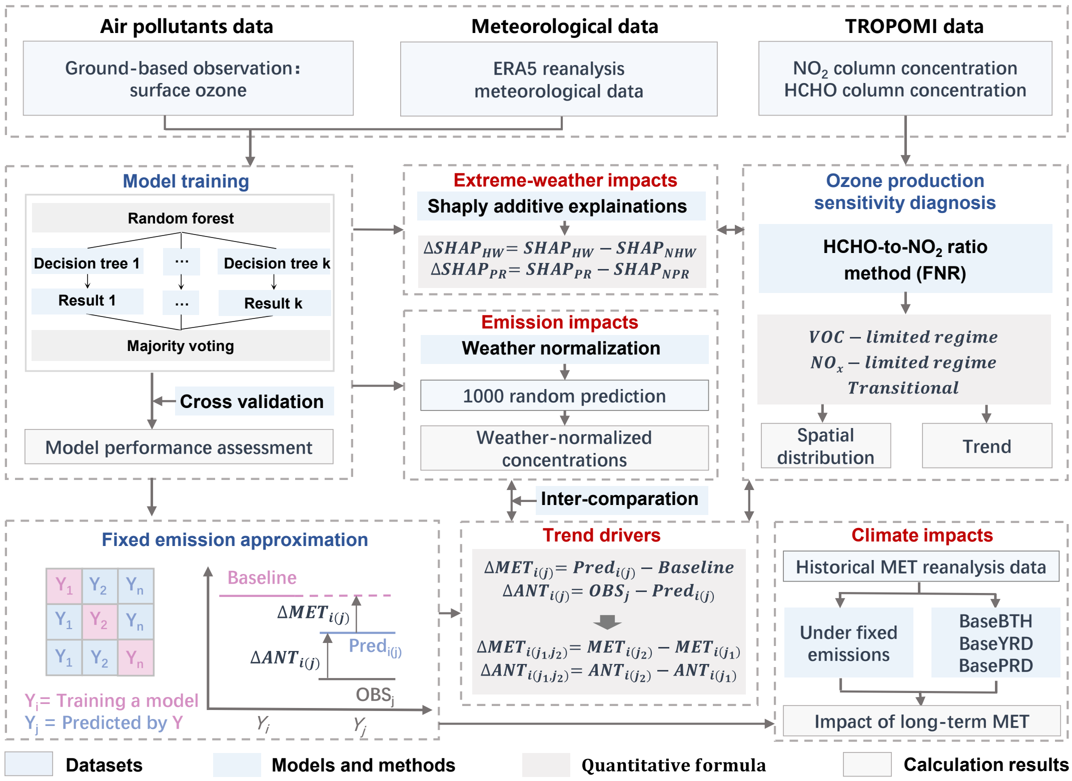

Figure 1. Schematic framework of data analysis and methodology. *This study integrates multi-dimensional datasets, including ground-based observations, meteorological reanalysis, and satellite remote sensing. A fixed emission approximation (FEA) approach, developed based on the random forest (RF) model, is employed to quantitatively disentangle the contributions of meteorological conditions (MET) and anthropogenic emissions (ANT) to ozone trend variations, and its performance is compared with the conventional meteorological normalization method. The SHAP technique is further applied to assess the influence of extreme weather events, such as heatwaves (HW) and extreme precipitation (PR). The satellite-derived formaldehyde-to-nitrogen dioxide ratio (FNR) is used to diagnose ozone production sensitivity, to explain and verify the impact of extreme weather and anthropogenic emissions on ozone. Finally, the FEA framework is extended to evaluate the long-term impacts of climate change on ozone trends since 1970.*

Line 114: The role of time variables requires clarification. Were the diurnal and seasonal/monthly variables included to remove short-term and seasonal variability, leaving the long-term trend for quantitative attribution? Please explain explicitly.

Response: We thank the reviewer for this insightful comment. Yes, the diurnal and seasonal/monthly variables were incorporated as proxies for short-term, periodic variations in emissions and meteorological conditions. Their inclusion allows the model to effectively separate these regular temporal patterns from the long-term interannual trends that are the main focus of our quantitative attribution analysis. The revised manuscript now clarifies this point, and the relevant text has been updated to read as follows:

The time variables – hour (hour of day) and month (month of year) – are used as emission surrogates to capture regular diurnal and seasonal variations in anthropogenic activity. A similar strategy is widely applied in previous studies about long-term trends in air pollutants (e.g., Grange et al., 2018; Vu et al., 2019) to separate short-term cyclical emission variability from long-term trends.

These temporal emission surrogates, including month and hour, represent short-term regular emission patterns (e.g., diurnal cycles), thereby enabling the model to isolate the long-term emission-driven component of ozone changes (Grange et al., 2018; Meng et al., 2025; Shi et al., 2021; Vu et al., 2019).

Line 125: Why was the modeling performed separately for each city, rather than by grouping cities into regions? Please explain the rationale.

Response: Thank you for this valuable comment. We conducted the modeling separately for each city to minimize uncertainties arising from surface and emission heterogeneity within broader regions. Cities across China exhibit distinct characteristics in terms of land use patterns, emission structures, local meteorology, and boundary-layer dynamics, all of which can strongly influence ozone formation and variability. Modeling at the city level allows the framework to better capture these localized processes and maintain higher fidelity in attribution analysis. We have revised the manuscript to clarify this rationale, as follows:

Our modeling strategy involves building and predicting models for individual cities and for each year from 2015 to 2023, which helps in minimizing the uncertainty caused by surface heterogeneity.

Nonetheless, we acknowledge that this approach may also introduce certain limitations. Specifically, the current implementation does not explicitly resolve grid-scale spatial heterogeneity, vegetation activity, or land-use dynamics, which may influence local ozone formation. To address this, we have included additional discussion in the conclusion section as follows:

Nonetheless, some limitations remain. The current implementation did not explicitly resolve grid-scale spatial heterogeneity, vegetation, or land-use dynamics, which may influence ozone formation. Moreover, potential sensitivities to spatial resolution warrant further investigation through coupled applications of machine learning and chemical transport models.

Lines 152-161: The uncertainty analysis, particularly for the Fixed Emission Approximation (FEA) method, is highly valuable. I strongly recommend moving these results (currently Figure S3) into the main text.

Response: We fully agree with the reviewer that the uncertainty analysis of the Fixed Emission Approximation (FEA) method represents a key component of the study. To improve its visibility and enhance the transparency of our methodological evaluation, we have moved these results from the Supplementary Information (previously Figure S3) to the main text as the new Figure 2. This adjustment allows readers to more directly assess the robustness and reliability of the FEA framework, thereby strengthening the methodological clarity and scientific rigor of the paper.

Lines 202-205: The manuscript highlights several regions in China. Please explain why these regions were emphasized and include a map showing their geographic distribution for better context.

Response: We thank the reviewer for this valuable comment. The selected regions – BTH, YRD, FWP, SCB, and PRD – were chosen because they are representative urban clusters of different parts of China and capture the diversity in emission characteristics and atmospheric conditions. For example, the BTH region represents northern inland cities dominated by anthropogenic emissions, while the PRD region represents southern coastal cities with substantial biogenic emissions. The YRD exhibits strong anthropogenic emissions influenced by southern biogenic sources, and SCB reflects the pollution characteristics of central and southwestern China. These regional divisions are consistent with prior studies and provide a meaningful framework for analyzing ozone variability across China. To enhance clarity, we have added a new map illustrating the geographic distribution of these regions, now included as Figure 3 in the revised manuscript. This addition allows readers to easily contextualize the regional analyses. The relevant text has been updated to read as follows:

Figure 3 presents the spatial distribution of the average summertime (2018-2023) maximum daily 8-hour average (MDA8) ozone, surface NO_2, and TROPOMI NO_2, HCHO column concentrations across China, along with the locations of the country's five major city clusters: Beijing-Tianjin-Hebei (BTH), Fenwei Plain (FWP), Yangtze River Delta (YRD), Sichuan Basin (SCB), and Pearl River Delta (PRD). Across these five major city clusters, the average summer ozone concentrations ranged from 88.9 to 161.3 $\mu g\ m^{-3}$ – substantially exceeding the 43.0 $\mu g\ m^{-3}$ threshold associated with ecosystem productivity loss (Gong et al., 2021) and the World Health Organization (WHO, 2021)-recommended peak seasonal average of 60 $\mu g\ m^{-3}$. TROPOMI satellite observations of NO_2 column concentration show notably elevated concentrations over the five major city clusters, particularly in the BTH, YRD, and FWP, which align with surface NO_2 distribution patterns and *confirm the scale of anthropogenic NO_x emissions in these regions (Zheng et al.,2021). TROPOMI satellite observations of HCHO column concentrations similarly reveal these city clusters as hotspots for VOC emissions (Fig. 3d). These concurrent high levels of NO_2 and HCHO suggest a strong photochemical ozone pollution potential, as the abundant precursors in these urban clusters could drive substantial ozone production during the summer months. This highlights the significant risks posed by summertime ozone in China's most urbanized and industrialized regions, with implications for both human and ecosystem health.*

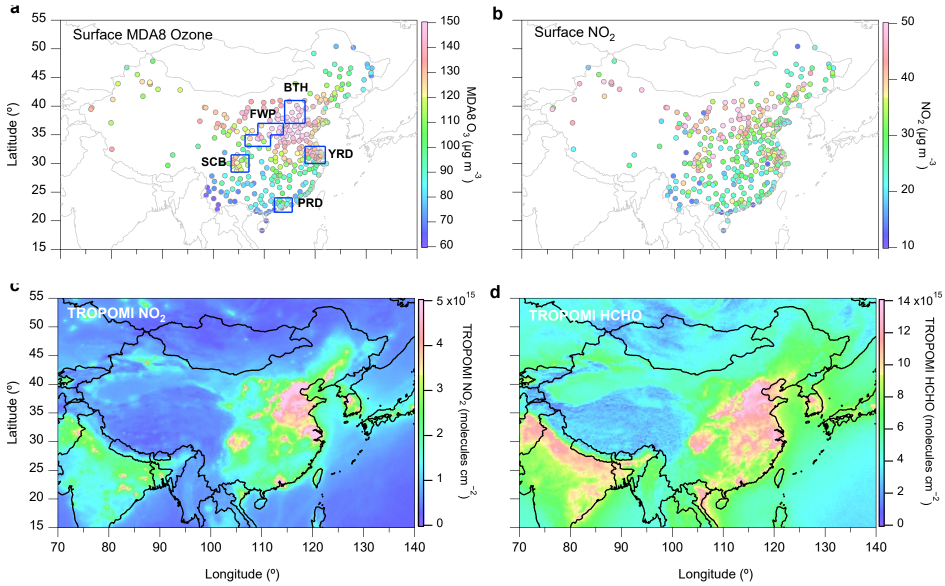

Figure 3. Spatial distribution of summertime MDA8 ozone, surface NO_2, and TROPOMI NO_2, HCHO across major city clusters in China. *The panels represent the average MDA8 ozone, surface NO_2, and TROPOMI NO_2, HCHO column concentrations for 354 cities in China during the summertime (June–August) from 2018 to 2023. The corresponding five regions includes BTH (37°–41°N, 114°–118°E); YRD (30°–33°N, 118.2°–122°E); SCB (28.5°–31.5°N, 103.5°–107°E); PRD (21.5°–24°N, 112°–115.5°E) and FWP (106.25–111.25°E, 33–35°N, and 108.75–113.75°E, 35–37°N).*

Line 227: The references here primarily address ecological impacts, yet the text mentions "both human and ecological health." Please provide more references specific to human health. Also, revise "ecological health" to "ecosystem health."

Response: We thank the reviewer for this helpful suggestion. In response, we have added references specifically addressing the impacts of tropospheric ozone on human health, including respiratory and cardiovascular outcomes. Additionally, we have revised the terminology from "ecological health" to "ecosystem health" throughout the manuscript to ensure precision. The modified text now reads:

Across these five major city clusters, the average summer ozone concentrations ranged from 88.9 to 161.3 μg m^{-3} – substantially exceeding the 43.0 μg m^{-3} threshold associated with ecosystem productivity loss (Gong et al., 2021) and the World Health Organization (WHO, 2021)-recommended peak seasonal average of 60 μg m^{-3}.

This highlights the significant risks posed by summertime ozone in China's most urbanized and industrialized regions, with implications for both human and ecosystem health.

Line 229: The phrase "reflecting initial policy effectiveness" is unclear. Please rephrase for precision.

Response: We thank the reviewer for pointing this out. To avoid potential confusion, we have removed the phrase entirely, as Fig. S1 focuses solely on observed concentration changes rather than explicitly quantifying policy impacts. This modification enhances clarity and precision in the text.

Lines 230-232: The conclusion drawn here seems overstated, as the evidence provided is insufficient. This section mainly discusses temporal and spatial ozone concentration trends. A more cautious interpretation is recommended: instead of attributing trends directly to policy effectiveness, the authors could note that observed trends occurred under varying emission control backgrounds, while meteorology also played an important role.

Response: We appreciate the reviewer's suggestion. The text has been revised to adopt a more cautious interpretation, emphasizing that the observed temporal and spatial ozone trends occurred under varying emission control contexts and were influenced by meteorological variability. The revised sentence now reads:

Spatially, ozone hotspot regions expanded between 2013 and 2017 (Fig. S1 a-e), followed by contraction during 2018-2020 (Fig. S1 f-i). However, this progress stalled in 2021. A sharp reversal was observed in 2022, with widespread increases in MDA8 ozone (Fig. S1 k). These changes could be closely linked to emission control measures and meteorological conditions, which will be further discussed in Sections 3.2 and 3.3.

Line 243: Please define the parameters τ and p shown in Figure 1.

Response: We thank the reviewer for this suggestion. τ is a statistic used in the Mann-Kendall trend test to measure the correlation between data points in a sequence, but it is rarely used. The p-value is a statistic used to assess the statistical significance of the trend. We have now removed τ and explicitly defined the parameter p in the figure caption of Figure.

Line 276: Correct "Emission-driven" to "emission-driven."

Response: Thank you for pointing this out. We have corrected the capitalization, changing "Emission-driven" to "emission-driven" in the revised manuscript.

Lines 279-281: The logic here is confusing. The discussion first emphasizes the role of anthropogenic emissions, but then suggests that changes in emissions highlight the role of meteorology. Please clarify or restructure this argument.

Response: Thank you for your comment, and sorry for the confusion. We have restructured the argument for clarity. The revised text now reads:

These results indicate that while emission control policies initially produced substantial benefits, their effectiveness has gradually diminished, suggesting that ozone responses to further emission reductions may have reached a saturation point.

Lines 299-300: Please rephrase the sentence. The term "near-baseline" is ambiguous and requires clarification.

Response: Thank you for pointing this out. We have rephrased the sentence for clarity:

By 2023, HCHO concentrations had returned to pre-heatwave levels.

Line 342: There is an editorial error that needs correction.

Response: Thank you for your comment. We have identified the editorial error and corrected it. The revised text now reads:

Ozone decreases attributable to meteorology reached –14.4 ± 3.0 μg m^{-3} in the FWP, –15.9 ± 3.8 μg m^{-3} in the YRD, and –11.1 ± 2.4 μg m^{-3} in the SCB, explaining 58 ± 12%, 77 ± 18%, and 80 ± 17% of the total ozone decline, respectively.

Line 390: I recommend revising the y-axis labels for greater accuracy. For instance, in panel (a), the label currently suggests "extreme weather," but it actually represents only "extreme heatwave". In contrast, panel (b) provides a more specific description. The labeling should be made consistent and precise to avoid potential misinterpretation.

Response: Thank you for pointing this out. We have revised the y-axis labels to improve clarity and consistency. Specifically, panel (a) now explicitly indicates "Extreme Heatwave (HW)," while panel (b) retains the more specific description of pluvial events. This ensures accurate representation and avoids potential misinterpretation. The updated figure has been included in the revised manuscript.

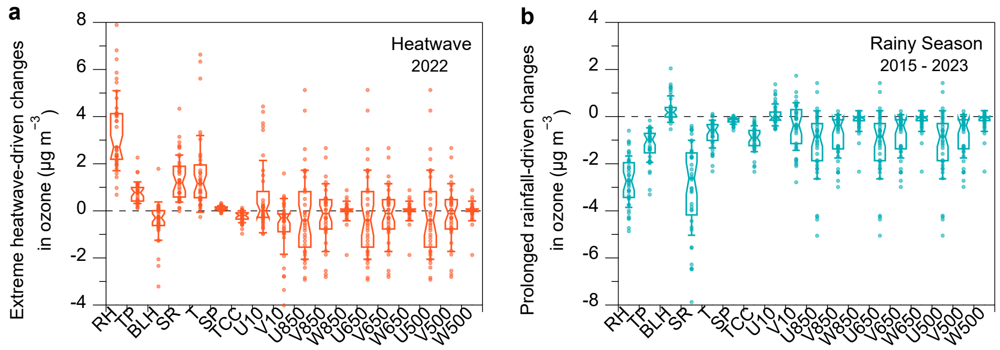

Figure 7. Meteorological impact on predicted ozone concentrations under heatwave and rainy weather conditions. *(a) Differences in SHAP values (ΔSHAP) between heatwave and non-heatwave periods in the Yangtze-Huaihe region during summer 2022. (b) Differences in SHAP values (ΔSHAP) between prolonged rainfall periods and non-prolonged rainfall periods in the same region from 2015 to 2023. Box plots show the distribution of ΔSHAP across cities; the center line indicates the median, boxes denote the interquartile range (25th-75th percentiles), and whisker line extends to one standard deviation.*

The title of this section should be revised, since the authors are not reconstructing the ozone trend per se. A more accurate option could be "Reshaping distributions of ozone controlled by a warming climate." This section is indeed interesting and methodologically innovative. However, the manuscript should elaborate more clearly on which specific factors are included in the climate-change-driven trend, especially considering the constraints posed by the limited length and coverage of historical observational records.

Response: Thank you for your insightful comments and suggestions. We have revised the title to "Reshaping distributions of ozone controlled by a warming climate" to more accurately reflect the content. Additionally, we have clarified the factors included in the climate-change-driven trend. Specifically, the trend incorporates temperature increases. We also note the constraints imposed by the limited length and spatial coverage of historical observational records, which are acknowledged in the discussion.

Line 331 (Figure 3): Ensure the map format is consistent with that in the Supplementary Figures.

Response: Thank you for your comment. In the revised manuscript, we have replaced the corresponding Figure with a map showing the trend of ozone sensitivity intervals across the five major city clusters of China from 2018 to 2023. Additionally, the map in the supplementary material has also been updated to reflect the latitude and longitude range corresponding to the five major city clusters. **The related modification is shown as Fig. 6 and Fig. S11.**

Lines 374-376: Since the SHAP interpreter is a key tool used to analyze predictor contributions, it should be briefly described in the Methods section.

Response: Thank you for your comment. We have moved the relevant description of the SHAP interpreter from the supplementary material to the main body of the manuscript. The updated content can now be found in Section 2.4.

Line 437: In Figure 5, the authors present both climate-change-driven and emission-driven trends. I am curious about how the results from the proposed FEA method compare with those from other widely used machine-learning approaches for trend analysis, such as de-weather. A comparison between different methods would not only be interesting but also serve as a useful validation of the robustness of the proposed framework.

Response: We thank the reviewer for this valuable suggestion. To address this, we have incorporated an analysis using the widely used "weather normalization" method and compared the results with those obtained from our FEA framework. The comparison shows that the trends derived from both approaches are highly consistent, demonstrating the robustness and reliability of our proposed FEA methodology. The inter-comparison results have been added to the revised manuscript for transparency and validation. The relevant text has been updated to read as follows:

2.3 Weather normalization analysis

To compare the FEA method with other commonly used statistical approaches, we also applied the widely adopted meteorological normalization technique based on the RF algorithm. This approach constructs a regression model that relates air pollutant concentrations to meteorological parameters and emission surrogate indicators (i.e., time variables such as unix time, day of year, day of month, and hour of day) (Grange et al., 2018; Vu et al., 2019). Once the model is trained, pollutant concentrations are predicted by randomly resampling meteorological variables from long-term historical meteorological datasets, thereby generating a new ensemble of predictions (Vu et al., 2019). These predictions are made under consistent meteorological conditions, enabling the isolation of meteorological influences from anthropogenic emission effects on air pollutant trends. The resulting weather-normalized pollutant concentrations (Fig. 1) represent the levels expected under average meteorological conditions, thus reflecting the impact of emission changes alone. This approach, first proposed by Grange et al. (2018), has been widely applied in the long-term attribution of air pollution trends and in assessing short-term emission reduction effects (Shi et al., 2021; Vu et al., 2019). In this study, the meteorological normalization follows this established framework, with meteorological *variables randomly sampled from the long-term dataset spanning 1970-2023. Each normalization process involves 1,000 iterations, and the arithmetic mean of these iterations' simulated values is adopted as the final normalized result. The alignment between FEA-based and weather-normalized trends (Fig. S4) affirms the robustness of the FEA framework.*

Supplement:

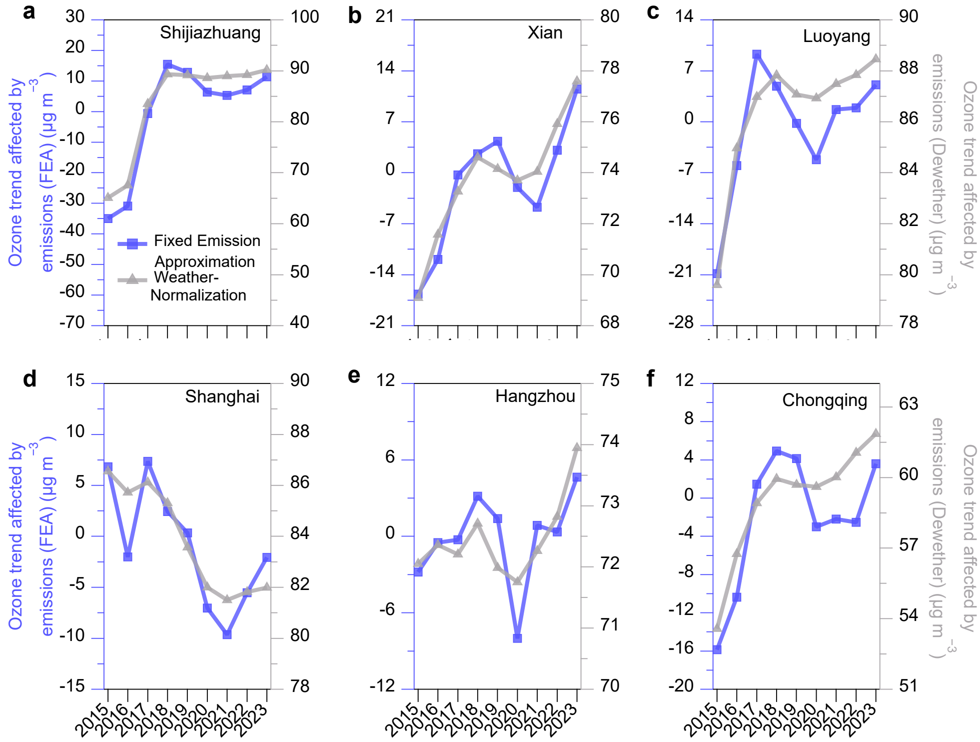

Figure S4. Trends in the average summertime ozone concentration changes from 2015 to 2023, driven by anthropogenic emission control. *The figure compares the ozone trend variations for six representative cities in key regions, based on both the FEA method and weather normalization method.*

Line 445 (Conclusion): The conclusion is overly lengthy. Please condense and refine this section for clarity and impact.

Response: Thank you for your suggestion. We have revised the conclusion to make it more concise and impactful, emphasizing the key findings and implications of our study.