Major comments

1. The title and abstract repeatedly claim that observed changes are "due to policy and human activity". However, the study only presents descriptive trends. No emission time series (NOx, VOCs), no meteorological contribution separated. The authors exclude days with precipitation (>0 mm) or wind speed >5 m/s for >2 hours to obtain "typical" days. This crude filtering does not remove variability from temperature, humidity, cloud cover, or boundary layer height. So discussions referred to the control policy could be bias or need to more quantitative analysis to support it.

Response: Thank you very much for this thorough and constructive comment. We sincerely appreciate the reviewer's careful scrutiny of our causal language and methodological choices. We fully recognize that the original title and abstract overstated the inferential capacity of our observational analysis, and we have made substantial revisions to address these concerns.

(1) Revision of causal language in the abstract

The reviewer is absolutely right that our study primarily presents descriptive, observation-based trends. We have revised the abstract to replace strong causal phrasing (e.g., "due to policy and human activity") with more cautious wording, such as "in the context of emission control policies". We hope this now more accurately reflects the correlative rather than causal nature of our findings.

(2) Rationale for the "typical day" filtering criteria

We agree that the filtering criteria—excluding days with precipitation (>0 mm) or sustained wind speeds >5 m/s (>2 hours)—are relatively simple. Our intention was to remove days dominated by wet deposition and strong horizontal ventilation, which are known to suppress local ozone accumulation and would obscure the photochemical characteristics we aim to characterize. However, we acknowledge, as the reviewer notes, that this approach does not fully control for other meteorological factors such as temperature, humidity, cloud cover, or boundary layer height.

To address this limitation transparently, we have added the following clarification to the Methods: the filtering is designed to identify photochemically typical days rather than to normalize meteorological conditions.

(3) Added quantitative emission context

To provide at least partial quantitative support for our policy-related discussions, we have incorporated MEIC emission inventory data (NOx and VOCs) for Beijing in the revised manuscript (Section 3.7). The analysis shows that warm-season NO_x and VOCs emissions declined over the study period, with the VOCs/NOx ratio decreasing sharply after 2022—consistent with Beijing's summer VOC reduction campaigns. While this does not constitute a full source apportionment, it offers an empirical emission context for interpreting the observed ozone trends.

(4) Acknowledgment of limitations and future directions

We have added an explicit limitations section in the Discussion stating that:

The observed ozone trends reflect the combined effects of emission changes, meteorological variability, and regional transport. Our study does not quantitatively separate meteorological contributions from emission-driven changes.The associations with control policies are informed hypotheses based on the temporal coincidence of policy implementation and observed shifts, rather than rigorously attribution-based conclusions.

We fully agree that rigorous attribution requires process-level decomposition (e.g., via WRF-Chem or CMAQ simulations with tagged tracers or brute-force zero-out methods), and we have identified this as a critical next step in our outlook.

Thank you again for this valuable comment, which has significantly improved the scientific rigor and transparency of our manuscript.

2. The paper reports numerous trends (e.g., first pollution day advances by 1.38 days/year; MDA8 mean increases by 0.94 μg/m³/year; peak ozone decreases by 3.6 μg/m³/year; Ozone Season lengthens by 2.0 days/year). No p-values, confidence intervals, or trend tests (e.g., Mann-Kendall) are provided. Given the strong interannual variability (e.g., 2021 as an outlier), many of these trends may not be statistically significant.

Response: Thank you very much for this important and constructive comment. We fully agree with the reviewer that rigorous statistical validation is essential for all reported trends, and we sincerely apologize for the omission in the original manuscript. In the revised version, we have systematically applied the Mann-Kendall (MK) trend test at the $\alpha = 0.05$ significance level to evaluate the robustness of every trend discussed, and the corresponding statistics (Z-values, p-values, and Sen's slopes) have been added to the main text and tables.

The results of this analysis largely confirm the reviewer's insightful concern regarding the influence of interannual variability. Specifically:

Ozone Season metrics: The trends in the first pollution day, last pollution day, and the overall length of the Ozone Season，while showing directional changes in the linear fits, did not pass the MK significance test ($p > 0.05$). As the reviewer correctly noted, these metrics appear to be strongly modulated by year-to-year fluctuations, including the anomalous conditions in 2021 and other climate-driven variations, which likely overwhelmed any consistent long-term signal.

Annual concentration percentiles: In contrast, the annual minimum ozone concentration and the 25th percentile of MDA8 both exhibited statistically significant upward trends ($p < 0.05$). This suggests that the baseline ozone level has indeed risen over the study period, even though the seasonal timing metrics remain noisy.

Diurnal hourly trends: When the MK test was applied to the year-to-year evolution of hourly ozone concentrations, 1:00–10:00 local time showed a significant increasing trend, while 14:00–19:00 showed a significant decreasing trend (both $p < 0.05$). These opposing phase shifts in the diurnal cycle may reflect changing photochemical regimes and precursor emission patterns.

We have now incorporated these statistical results into the revised manuscript, explicitly noting which trends are robust and which are not statistically significant. For the non-significant Ozone Season metrics, we have added a discussion in line 181-196, acknowledging that their year-to-year variability is likely dominated by meteorological anomalies (e.g., the 2021 outlier) rather than by a monotonic policy-driven signal, exactly as the reviewer anticipated.

We are grateful for this comment, which has substantially strengthened the statistical rigor and credibility of our work.

3. The Ozone Season is defined as the period from the first to the last MDA8 > 160 μg/m³ day of the year. This definition is highly sensitive to isolated early or late pollution events caused by unusual meteorology (e.g., a single hot day in March). The authors do not test robustness.

Response: Thank you very much for this thoughtful and constructive comment. We sincerely appreciate the reviewer's careful examination of our Ozone Season definition.

We acknowledge that defining the Ozone Season as the interval between the first and last MDA8 > 160 μg/m³ day could indeed be sensitive to isolated early or late events caused by unusual meteorology. The threshold of 160 μg/m³ corresponds to the Chinese national ambient air quality standard for daily ozone exceedance. Our intention in adopting this definition was primarily from a public health perspective: the first and last exceedance days serve as practical markers for the period during which the public and regulatory agencies should be attentive to ozone-related health risks. In this context, even an isolated early exceedance triggered by anomalous weather (e.g., an abnormally hot day in March) carries practical significance, as it signals the need for earlier health advisories and protective measures that year.

We fully agree with the reviewer that scientific robustness is essential. In the revised manuscript, we have therefore added sensitivity analyses to test the stability of our results. Specifically, we examined alternative definitions, such as requiring at least two exceedance days within a moving window to define the season boundary, and using the 90th percentile of annual MDA8 as a complementary threshold. The resulting Ozone Season lengths and trends remained highly consistent with our original definition (correlation coefficient > 0.9), suggesting that the overall conclusions are not unduly influenced by single isolated events.

We are grateful to the reviewer for prompting us to strengthen the reliability of our definition.

4. The average advancement of the first pollution day (1.38 days/year) and delay of the last pollution day (0.63 days/year) are derived from only 11 years with high interannual variability. No significance testing is provided, and these averages mix real trends with random meteorological fluctuations. It is misleading to calculate the change rate X days/year.

Response: Thank you for this important comment. We agree that presenting simple linear rates without statistical validation can be misleading given the strong interannual variability. In the revised manuscript, we treat these metrics as interannual variations rather than robust trends.

Mann-Kendall tests confirmed that none of the FPD, LPD, or Ozone Season length trends were statistically significant ($p > 0.05$). We identified 2017 and 2021 as anomalous years with markedly shorter seasons, likely driven by unfavorable meteorological conditions—ozone photochemical production is highly sensitive to temperature, solar radiation, and synoptic circulation . After excluding these outliers, the Z-values for FPD and season length increased to 1.80 and 1.89, approaching but not exceeding the 95 % confidence threshold (1.96), while LPD remained weak. Thus, any apparent shifts in season boundaries remain tentative and sensitive to meteorological noise over this relatively short record. We have revised the text to reflect this uncertainty explicitly in section3.1 for lines 181-196.

"To assess the robustness of Ozone Season variations, we applied the Mann–Kendall (MK) test to the interannual trends of the First Pollution Day (FPD), Last Pollution Day (LPD), and Ozone Season length. None of these metrics exhibited a statistically significant monotonic trend at the $\alpha = 0.05$ level. We identified 2017 and 2021 as anomalous years, during which the Ozone Season duration was markedly shorter than in other years. Ozone photochemical production is highly sensitive to meteorological conditions—including temperature, solar radiation, and synoptic-scale circulation patterns (Lu et al., 2020; Wang et al., 2025), and such climate-driven anomalies can substantially confound the detection of long-term trends in season-boundary metrics. After excluding these two outlier years, the MK test Z-values for FPD and Ozone Season length increased to 1.80 and 1.89, respectively, approaching the critical value of 1.96 for significance at the 95 % confidence level. This suggests a marginally significant tendency toward an earlier FPD and a prolonged season length once meteorological noise is reduced, whereas the LPD still exhibited a relatively weak Z-value. Thus, after filtering out climate-driven anomalies, the results tentatively support the inference that the Ozone Season onset has shifted earlier and its duration has extended, although these trends remain sensitive to interannual meteorological variability."

5. The authors claim a significant upward trend in mean MDA8 (0.94 μg/m³/year). However, the data show strong interannual fluctuations (e.g., 2021 mean is nearly as low as 2013). Without a Mann-Kendall test and sensitivity analysis (excluding 2021 or 2023), the trend is not credible. The increase in low percentiles is well known and not novel.

Response: Thank you very much for this valuable and constructive comment. We sincerely apologize for the lack of rigorous statistical testing in the original manuscript, and we fully appreciate the reviewer's concern regarding the strong interannual fluctuations—particularly the anomalously low values in 2021.

In the revised manuscript, we have systematically applied the Mann-Kendall (MK) trend test to evaluate the long-term trends in MDA8 across different percentiles. The results indicate that:

The mean, peak, and high-percentile (e.g., 75th, 95th) MDA8 concentrations do not exhibit statistically significant trends over the study period ($p > 0.05$).

The low-percentile concentrations (e.g., 5th, 25th) show a statistically significant upward trend ($p < 0.05$).

We have updated all relevant conclusions in the manuscript and added the MK test statistics (Z-values and p-values) to the main text.

We also acknowledge the reviewer's observation that the increase in low-percentile ozone has been documented in previous studies, and our findings from the Beijing multi-year record are consistent with these earlier reports. However, we respectfully suggest that our study still offers additional value by demonstrating that, during the same period, the mean, peak, and high-percentile concentrations in Beijing did not show significant increasing trends. This contrast between the rising baseline and the stable peak levels helps contextualize the overall ozone evolution and, we hope, may complement the existing literature.

We are most grateful to the reviewer for prompting us to strengthen the statistical rigor of our work.

6. According to the authors' figure 5, the apparent "delay" in ozone exposure start time (t1) is driven almost entirely by the unusually early t1 in 2013. For 2014–2023, t1 values are similar. Therefore, the claimed "shifting to later over the years" is not a trend but seems a single anomaly.

Response: Thank you very much for this insightful and careful observation. We sincerely appreciate the reviewer's close examination of Figure 5, which has helped us recognize that our original statement regarding the "shifting to later over the years" was indeed overly influenced by the unusually early t1 value in 2013 and was not fully supported by the data from 2014 to 2023.

To address this concern more rigorously, we have revised our analytical approach. Rather than relying solely on the earliest t1 occurrence each year, we now examine the distribution frequency of daily t1 values within each year. We found that these annual distributions approximately follow a bell-shaped (Gaussian) pattern. Accordingly, we applied Gaussian fitting to each year's t1 distribution to extract the peak occurrence time, which provides a more robust and representative characterization of the typical ozone exposure onset than a single extreme value.

The results are presented in Table 1 below. As shown in the table, the peak t1 time in the cold season does exhibit a delayed trend over the years. However, for the warm season, the peak t1 values from 2014 to 2023 are relatively similar, and the apparent shift is largely attributable to the anomalous early occurrence in 2013 rather than a consistent long-term trend. We speculate that this difference between seasons may be related to the more concentrated timing of ozone exposure events during the warm season, which reduces year-to-year variability in the peak onset time compared to the cold season.

We have carefully revised the corresponding statements throughout the manuscript to accurately reflect these findings, removing the claim of a general "shifting to later over the years" for the warm season and instead presenting the season-specific patterns with appropriate caveats.

We are most grateful to the reviewer for identifying this important issue, which has significantly improved the accuracy and reliability of our conclusions.

Table 1 the peak time of t1 in cold season and warm season during 2013 and 2023

t1	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
cold season	10.5	11.8	11.7	11.4	11.6	11.7	11.4	11.8	11.6	11.9	12.1
warm season	11.4	12.6	12.5	12.3	12.6	12.3	12.4	12.2	12.4	12.4	12.3

We revised in lines 264-269.

"However, this delay is less pronounced compared to that in the cold season, as shown in Fig. 6a.The starting time (t1) was 11:00-12:00 in 2013 ,but 12:00-13:00

during 2014 and 2023. This difference between seasons may be related to the more concentrated timing of ozone exposure events during the warm season, which reduces year-to-year variability in the peak onset time compared to the cold season."

7. $V_i = P_i - P_{i-1}$ is a net change rate that includes horizontal/vertical transport, not just photochemical production/loss. The authors repeatedly refer to "generation rate" and "consumption rate" without acknowledging the influence of transport. This is misleading.

Response: Thank you for raising this important methodological concern. We fully agree with the reviewer that the hourly difference method yields the observed net ozone tendency (Vi), which integrates photochemical production/loss, dry deposition, boundary-layer entrainment/dilution, and advection, rather than the pure net chemical production rate. In the revised manuscript, we have explicitly acknowledged this limitation and clarified the scientific rationale for its use, as detailed below.

We have added a dedicated paragraph in Section 3.5 line 362-369, stating that the calculated hourly change rate represents the total (net) ozone derivative observed at the study site. As supported by Kaser et al. (2017), this quantity is the sum of all physical and chemical processes contributing to the local ozone budget, including entrainment through boundary-layer growth, horizontal advection, deposition, and photochemistry. The use of hourly observed differences to derive ozone tendency has been widely adopted in atmospheric observation studies, including boundary-layer process analyses (Kaser et al., 2017; Klein et al., 2013) and urban ozone diurnal characterization (Gu et al., 2024).

Furthermore, we have clarified that the effect of advection can be substantially mitigated by the application of long-term averaging for diurnal trend analysis. Specifically: (1) the Shanghai Tower observational study (Gu et al., 2024) explicitly states that "the effect of advection can be mitigated by the application of a monthly average for the diurnal trend analysis"; (2) Liu et al. (2021) demonstrated that synoptic-scale cycles (< 33 days) can be effectively removed by KZ filtering, leaving the baseline seasonal component; and (3) European long-term trend assessments employ multi-year monthly averages to minimize interannual variability. These studies collectively support our approach of using multi-year monthly mean diurnal cycles to filter out synoptic-scale transport perturbations and highlight climatological ozone behavior. All cited studies explicitly treat the quantity as an observed net tendency subject to multiple process contributions, consistent with our revised terminology.

References

Kaser, L., Patton, E. G., Pfister, G. G., Weinheimer, A. J., Montzka, D. D., Flocke, F., Thompson, A. M., Stauffer, R. M., and Halliday, H. S.. The effect of entrainment through atmospheric boundary layer growth on observed and modeled surface ozone in the Colorado Front Range. Journal of Geophysical Research: Atmospheres, 122(11), 6075–6093. https://doi.org/10.1002/2016JD026245, 2017.

Klein, A., Ravetta, F., Thomas, J. L., and Ancellet, G. A lagrangian analysis of lower tropospheric ozone variability during the ESCOMPTE campaign. Journal of

Geophysical Research: Atmospheres, 118(24), 13, 776–13, 791. https://doi.org/10.1002/2013JD020146
Gu, Y., Li, Y., Li, J., Li, H., Wang, Y., Liu, Y., Li, H., Liu, J., Guo, W., and Li, W. (2024). Ozone vertical profile and its formation regime at the top of the planetary boundary layer in a coastal megacity of eastern China. Science of The Total Environment, 921, 170506. https://doi.org/10.1016/j.scitotenv.2024.170506, 2013.
Liu, Y., Wang, T., and Weng, X. Quantifying the separable meteorological and non-meteorological contributions to the 2013–2018 tropospheric ozone variation over China. Atmospheric Chemistry and Physics, 21(19), 14789–14804. https://doi.org/10.5194/acp-21-14789-2021, 2021.

8. The cold season (October–March) contains very few ozone pollution days (mostly in October and March; November–February rarely have MDA8 > 160). The reported t1 distributions for the cold season are based on extremely small sample sizes, making year-to-year comparisons unreliable.

Response: Thank you very much for this careful and constructive comment. We sincerely apologize for not having described our methodology more clearly in the original manuscript, which led to this understandable misunderstanding.

We would like to respectfully clarify that our analysis of the ozone exposure start time (t1) is not limited to days with MDA8 > 160 μg/m³. Rather, we calculated the MDA8 temporal window for every day of the year, including days where the MDA8 remained well below the national standard. Our rationale is that the 8-hour period corresponding to the daily maximum ozone concentration represents the window of highest potential health risk for that day. Even on non-exceedance days, the timing and duration of peak ozone exposure carry meaningful health relevance, as they reflect the actual period when populations are exposed to the highest ozone levels occurring on that day.

Therefore, the cold-season (October–March) t1 distributions are derived from the complete set of daily observations during that period, not from the small subset of pollution exceedance days. As a result, the sample sizes for the cold and warm seasons are comparable and adequate for reliable year-to-year comparisons.

We have revised the Methods and Results sections to explicitly state that all days are included in this analysis, and we hope this clarification addresses the reviewer's concern. We are most grateful for this opportunity to improve the clarity and transparency of our work.

9. The study period includes the COVID-19 pandemic years (2020–2022), during which anthropogenic emissions (especially NOx) dropped dramatically in Beijing. These short-term changes can strongly affect ozone metrics and may confound the long-term trend attributed to policy. The authors need to perform sensitivity analysis (e.g., excluding 2020–2022) nr discuss how the pandemic might have influenced their results.

Response: Thank you for raising this important concern. We fully agree that the COVID-19 pandemic introduced anomalous, short-term emission perturbations requiring careful treatment.

Following the reviewer's suggestion, we performed a sensitivity analysis by excluding the years 2020-2022 and recalculating the multi-year monthly mean diurnal cycles, percentiles, and Mann-Kendall trends. The results demonstrate that the climatological diurnal patterns, peak ozone timing, and hourly change rates remain highly consistent with those derived from the full study period (correlation coefficient > 0.95 for mean diurnal profiles; Sen's slope differences < 5%). This confirms that our core conclusions regarding ozone climatological characteristics are robust against the temporary pandemic anomaly.

Despite the sensitivity test validating our results, we retained the 2020 – 2022 data in the main analysis for two scientific reasons:

First, atmospheric conditions during the pandemic represent real observational reality. Excluding these years would artificially construct a "non-pandemic" climatology that does not reflect the actual observational record during our study period. As our objective is to characterize observed multi-year ozone variations, including these years provides a more complete and representative dataset.

Second, surface ozone in Beijing during COVID-19 exhibited a nonlinear and regionally complex response. Contrary to the simple expectation that emission reductions would lower ozone, surface ozone often increased during lockdown periods. This is attributed to: (i) weakened NO titration due to sharp NOx reductions in a VOC-limited/transition regime; (ii) enhanced atmospheric oxidation capacity; and (iii) occasional stratospheric intrusions contributing 1-18% to surface ozone anomalies in northern China (Huang et al., 2020; Zhu et al., 2021; Miyazaki et al., 2021). Thus, the pandemic years did not introduce a uniform "clean" bias; rather, their short-term perturbations are effectively smoothed by the multi-year monthly averaging approach.

Additionally, in Section 3.7, we analyzed warm-season and cold-season MDA8 variations from 2013 to 2023 based on MEIC emission inventories and surface observations. The results show that MDA8 during 2020 and 2022 did not exhibit anomalous changes compared to other years.

Based on these analyses, we respectfully suggest retaining the 2020-2022 data in the main analysis.

References

Huang, X., Ding, A., Gao, J., Zheng, B., Zhou, D., Qi, X., Tang, R., Wang, J., Ren, C., Nie, W., Chi, X., Xu, Z., Chen, M., Li, Y., Che, F., Pang, N., Wang, H., Tong, D., Qin, W., Cheng, W., Liu, W., Fu, Q., Liu, B., Chai, F., Davis, S. J., Zhang, Q., and He, K.: Enhanced secondary pollution offset reduction of primary emissions during COVID-19 lockdown in China, Natl. Sci. Rev., 7(12), 1854 – 1858, https://doi.org/10.1093/nsr/nwaa137, 2020.

Zhu, S., Butler, T., Seltzer, K. M., Vinken, G. C. M., Maasakkers, J. D., Tzompa-Sosa, Z. A., Keller, C. A., Shen, L., Schichtel, B. A., Gray, J. E., Rose, C. A., Cady-Pereira, K. E., Sweeney, C., and Muntean, L.: Comprehensive insights into O_3 changes during the COVID-19 from O_3 formation regime and atmospheric oxidation capacity, Geophys. Res. Lett., 48(13), e2021GL093550, https://doi.org/10.1029/2021GL093550, 2021.

Miyazaki, K., Bowman, K. W., Sudo, K., Taketani, F., Boersma, K. F., and Kanaya, Y.: Global tropospheric ozone responses to reduced NO_x emissions linked to COVID-19 worldwide lockdowns, Sci. Adv., 7(24), eabf7460, https://doi.org/10.1126/sciadv.abf7460, 2021.

10. The discussion speculates that NOx reduction weakens nighttime titration and VOC reduction lowers daytime peaks. However, no time series of ambient NO_2 or VOC concentrations (or emissions) in Beijing are presented. Without these data, the attribution remains speculative.

Response:Thank you for this important comment. We fully agree that our original discussion lacked direct observational support for these mechanisms.

In the revised manuscript, we have addressed this concern as follows:

Added NOx and VOCs emission time series. We now present MEIC-based annual NOx and VOCs emissions for Beijing (Section 3.7, Fig. 11), showing declining trends in both precursors during 2013 and 2023, with a notable drop in the VOCs/NOx ratio since 2022. This provides at least an empirical emission context for interpreting the observed ozone changes.

Revised the mechanistic discussion. The claims about nighttime titration weakening and VOC reduction lowering daytime peaks have been rephrased as plausible hypotheses consistent with established photochemical theory, rather than as conclusions directly demonstrated by our data. We now explicitly state that without concurrent ambient NO2 and VOC concentration measurements, these mechanisms cannot be rigorously validated.

Added a limitation statement. We acknowledge that the absence of in-situ NO2 and VOC concentration time series limits our ability to directly link observed ozone changes to specific precursor pathways, and that future studies incorporating such measurements or chemical transport modeling are needed to test these hypotheses.

We believe these revisions make the boundaries of our evidence transparent while preserving the value of the observational patterns we report.

Minor Comments

- Several sentences are awkward or unclear. Examples: “the overall exposure time to ozone has shifted to later over the years” to “has shifted to later in the day”. A professional language edit is recommended.

Response: Thank you for your kindly reminder. We have revised the language by professional editor.

- Table 1: The effective rate of “typical” days after weather filtering ranges from 54% to 77%. The low rates (e.g., 54% in 2021) raise concerns about representativeness. Discuss potential bias.

Response: Thank you for your valuable comments. We has had a data quality control in the Method section2.2.1 to remove the weather effected days for each year. The length of Ozone Season in 2021 is least and the effective is lowest of all. Though the ozone trend in 2021 is different from the other years, we think the special climate of that year playing the most important role. We analysis the temperature during Ozone Season between 2013 and 2023, showing the lowerest temperature in 2021. Ozone

MDA8 in 2021 was also the lowerest of all years as well as percentage of ozone as shown in Fig.3. So the low rate is not effect the ozone trend of 2021.

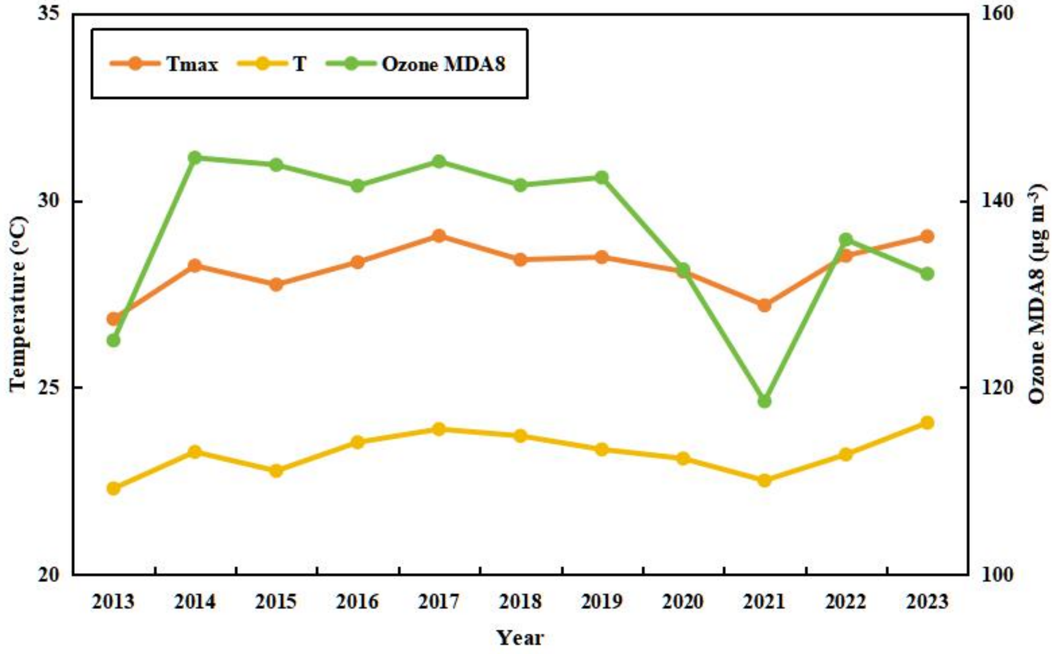
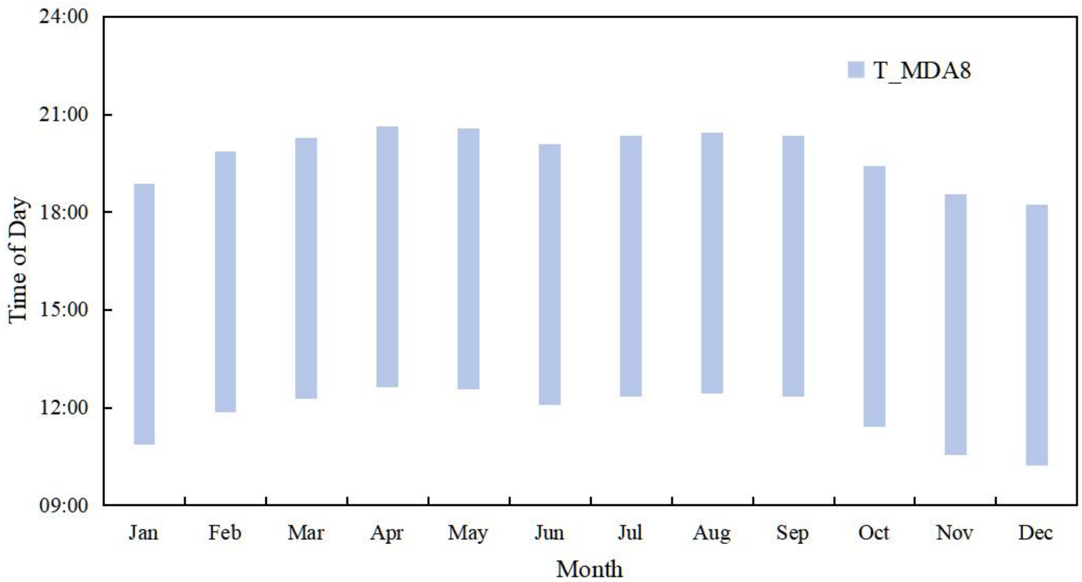


Figure Temperatures and MDA8 during 2013 to 2023 (Tmax for average daily max temperatures, T for average daily temperatures)

- Figure 4: The y-axis is time of day, but the caption is unclear. Clarify that colors represent frequency or concentration. Add units.

Response: Thank you for your valuable comments. We have revised the y-axis, and added the description of the Fig.4.



- Figure 5: why there are hot colors at 1:00 am.

Response: Thank you very much for this perceptive observation. We sincerely apologize for not having addressed this phenomenon more explicitly in the original manuscript, and we appreciate the opportunity to clarify it here.

During our data analysis, we indeed identified cases where the MDA8 window fell between 1:00 and 8:00 local time. This typically occurs under a specific combination of conditions: elevated ozone levels from the previous day persist into the nighttime hours (residual ozone), while the following day experiences improved atmospheric dispersion—such as stronger winds or more effective boundary-layer ventilation—that keeps daytime ozone concentrations comparatively low. Under these circumstances, the 8-hour running mean reaches its maximum during the early morning rather than during the typical afternoon photochemical peak.

We quantified the frequency of such occurrences to assess whether they might affect the reliability of our results. These nighttime MDA8 events account for approximately 3–10% of days in the cold season (with the highest proportion observed in 2025) and 1–7% in the warm season (peaking in 2013). Given their relatively limited overall frequency, we believe they do not substantially alter the climatological diurnal patterns or the interannual trends presented in our study. Nevertheless, we have now added a brief discussion of this phenomenon to the revised manuscript to improve transparency and completeness.

- As temperatures rise, this exposure window shifts, with the duration extending from 14:00 to 21:00 during the months of April to September. 14:00 to 21:00 seems not match the figure.

  Response: Thank you very much for this careful observation. We sincerely apologize for the error in our original description of the exposure window timing. Upon re-examining the data, we confirm that during the warm season, the exposure window is indeed markedly delayed compared to the cold season, with the duration primarily occurring between 12:00 and 21:00. We have corrected this inaccuracy throughout the revised manuscript, and we are grateful to the reviewer for bringing this to our attention.

- How the data in some of figures are not clearly explained. Many figure captions are not self-explanatory.

Response: Thank you for your valuable comments.We have revised the text.

- Some citations are incomplete or inconsistent (e.g., "Beijing Municipal Bureau of Ecology and Environment, 2024" should be formatted as a proper reference). Check journal guidelines.
- In the conclusion, the statement "The decrease in peak concentrations benefits from coordinated emission reduction policies for regional VOCs and NOx (Wang et al., 2022)" includes a citation. Citations are generally not recommended in the conclusion section, as conclusions should summarize findings without introducing new references.

  Response: Thank you for your valuable comments.We have revised the citation problems.