General Overview:

This manuscript reports the deep connection between the Peroxyacetyl nitrate (PAN) formation with summertime haze and photochemical air pollution in Xiamen. The authors have used multiple observations of trace gases, aerosol chemical composition, and meteorological parameters, combined with a model simulation with the Master Chemical Mechanism (MCMv3.3.1) and a machine learning model. The manuscript is well written and I recommend its publication after addressing the major and minor comments given below in order for it to be published.

Response: Thank you for your thorough review and positive feedback on our manuscript. We appreciate your insightful comments and have addressed both the major and minor points you raised to enhance the clarity and quality of our work. In general replies, we use blue font; red font indicates parts added in the revised manuscript, and blue italic font denotes references.

Major comments

Line 198. The main text describes two periods: "haze" and "clean" conditions based on the maximum diurnal hourly concentration (> 35 μg/m³). I believe another condition should be considered, as the period from July 26 to 31 shows distinct characteristics compared to other periods (e.g., 11 and 13 of July). This behavior remains consistent over time, indicating increased pollution and greater atmospheric reactivity (more precursors, radicals, etc.), which is reflected in the concentrations of O_3, NO_2, and $PM_{2.5}$. I think this period should be treated as a special event, as something significant occurred during that time.

Response: Yes, the "haze" and "clean" conditions are based on the maximum daily concentration (> 35 $\mu g \cdot m^{-3}$). The period from July 26 to 31 was continuously influenced by the West Pacific Subtropical High (WPSH), providing relatively stable (ws = 1.04 $m \cdot s^{-1}$), high-temperature (maximum daily average of 37.82 °C), high-humidity (maximum daily average of 81.65 %), and strong UV(maximum daily average of 46.71$W \cdot m^{-2}$) atmospheric conditions, which are conducive to the accumulation and secondary transformation of pollutants. As a result, we observed an increase in the concentrations of various pollutants (including O_3, NO_2, and $PM_{2.5}$) during this period. On July 13, the Xiamen also affected by WPSH. The days of July 11 and 21-23 were influenced by the periphery of typhoons. This indicates that the rise in $PM_{2.5}$ concentrations was not solely due to the WPSH. Although July 12-15 was also influenced by the WPSH, there was no significant increase in pollutant concentrations, likely because the impact of Typhoon No. 9—"Shanshen" began on the 16th. If we categorize based solely on the WPSH, it would diverge from the theme of this paper.

One factor that could be related to the "clean" and "haze" conditions is wind direction, which differs between the period of July 14 to 19, considered as "clean conditions," and July 26 to 31, considered as "haze conditions." Is there any influence from air mass intrusions? Could there be an impact from pollution transport? I would suggest running back trajectories and exploring this aspect further, as it could be an important contributing factor.

Response: The 72-hour backward trajectories from July 14-19 originated from the East China Sea and passed through the coastal areas of Fujian Province (Fig. R1(a)). In contrast, the 72-hour backward trajectories from July 26-31 originated from the South China Sea and passed through the eastern part of Guangdong Province (Fig. R1(b)). This phenomenon indicates a significant difference in air masses between the two periods. As shown in Fig. R1(c), we can observe that the 72-hour backward trajectories from July 14-19 are associated with clean air masses, suggesting that this airflow primarily plays a role in removing pollutants. Additionally, when considering wind speed and direction, we found that during the clean period, the wind predominantly came from the northeast (Fig. R2(a)). During the haze period, the wind direction was dominated by other directions except for the northeast, particularly the southeast wind (Fig. R2(b)). As shown in Fig. R2(c), when the wind blows from the northeast, the pollutant concentrations are very low, indicating that the northeast wind primarily plays a role in clearing pollutants. However, during other wind directions, the overall $PM_{2.5}$ concentrations were relatively high, with high values mainly occurring at lower wind speeds, suggesting that pollution during the haze period primarily comes from local generation.

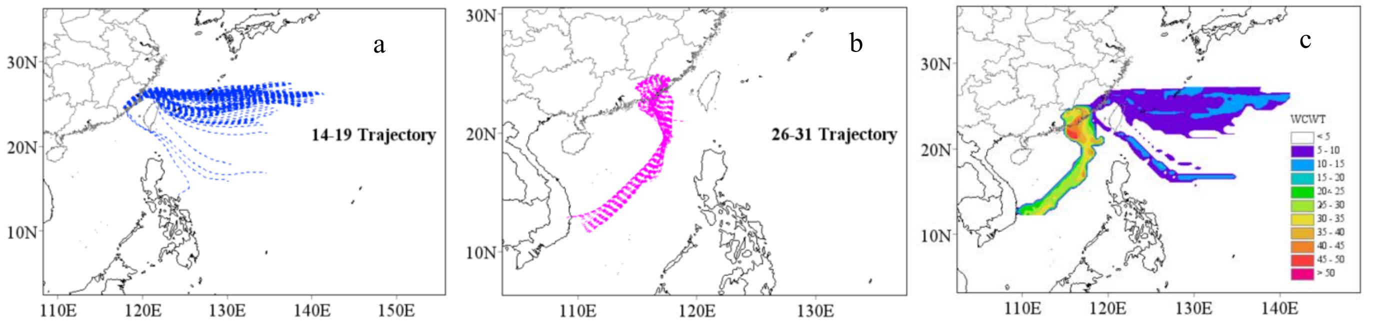

Fig. R1 72 h backward trajectories of July 14-19 (a), and July 26-31 (b), and WCWT plots for $PM_{2.5}$ concentrations including July 14-19 and July 26-31 at observation site (c).

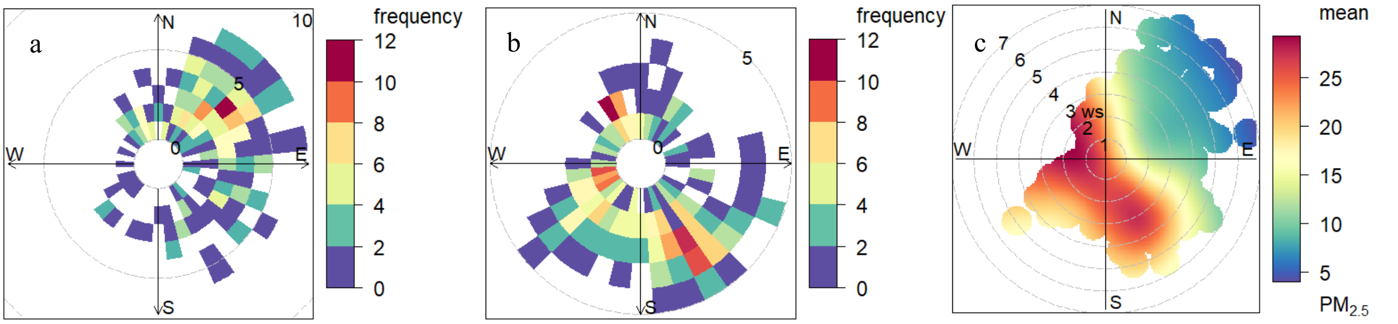

Fig. R2 Polar frequency of wind speed and wind direction for clean (a) and haze period (b), polar plot of $PM_{2.5}$ concentrations for whole observation period (c)

I'm not entirely sure that the periods of July 11 and 13 are relevant, as there are peaks exceeding 35 µg m^{-3}. I believe a "haze" event should have a consistent prevalence over time (at least for a couple of hours), and not be defined by peak. Please consider this option.

Response: The ambient air quality standards (GB 3095-2012) define that if the 24-hour average value of $PM_{2.5}$ exceeds 35 µg·m^{-3}, it surpasses the first-level standard; if it exceeds 75 µg·m^{-3}, it surpasses the second-level standard. Xiamen, recognized for its excellent air quality, ranked among China's top 10 cities from 2018 to 2023, achieving 7th in 2018, 4th in both 2019 and 2020, 6th in 2021, 9th in 2022, and returning to 7th in 2023 (mee.gov.cn, last assessed October 30, 2014). In this site, the overall pollutant concentrations are relatively low, with 24-hour average values not exceeding 75 µg·m^{-3}. The maximum value was observed on July 30, at only 50.38 µg·m^{-3}, indicating that, according to strict standards, there is no haze present. The 24-hour average exceeded 35 µg·m^{-3} on only three days: July 28, 29, and 30. If we use the criterion of having at least two consecutive hours exceeding 35 µg·m^{-3} to make a determination, then only the days from July 26 to 31 meet this requirement. We conducted a correlation analysis between the simulated and observed PAN values and found that the original classification method performed better. The R^2 value during the clean period decreased from 0.68 to 0.54, and during the haze period, it decreased from 0.47 to 0.36 (Fig. R3), indicating that the new classification method does not effectively differentiate their patterns. Additionally, when examining the time series of the net generation rate of PAN with and without PAN constraints, we see that this classification method precisely selects the scenario where the net generation rate is negative and exceeds -1 when constrained by PAN (Fig. R4), also suggesting that this classification method better distinguishes between the two different generation mechanisms.

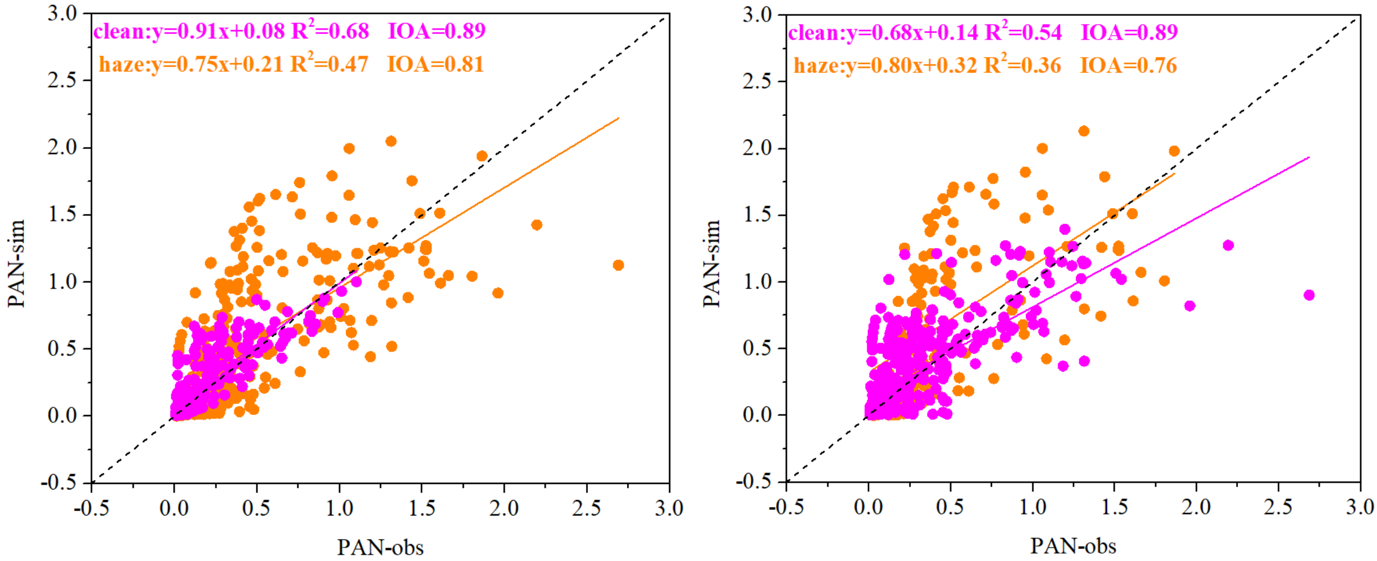

Fig. R3 Correlation between PAN observations and simulated values, where magenta represents the clean period and orange represents the haze period. The figure shows the original classification (a); The figure presents the new classification, with July 26-31 designated as the haze stage and the rest as the clean period (b).

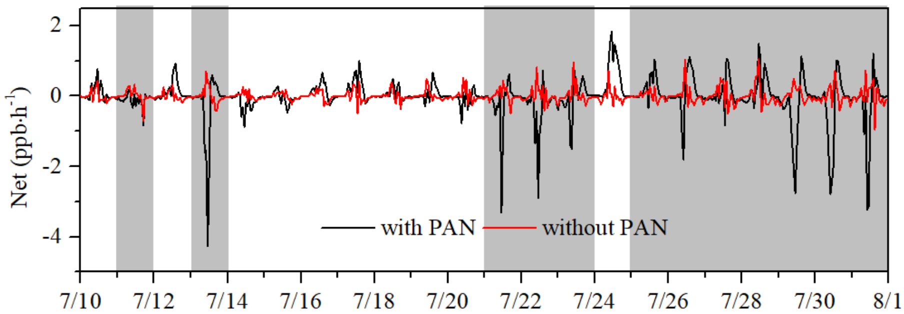

Fig. R4 The time series of PAN net production rate with PAN constrained and without PAN constrained

There is no detailed description of the VOCs measured in the analysis or how they were characterized, and there is no information about the instrument employed. Additionally, there is a lack of information regarding oxidation products and PAN precursors. I believe it is essential to include this information, along with diurnal variability and time series of individual species crucial for PAN formation.

Response: Thank you for pointing this out. The VOC measurements were conducted using a gas chromatography mass spectrometer (GC-FID/MS, TH-300B, Wuhan, China) at an hourly time resolution. This included key VOC species such as alkanes, alkenes, aromatics, and oxygenated VOCs. Detailed information regarding the VOC detection system and calibration procedures is available in our previous study (Liu et al., 2022). We added time series plots of VOCs by category and their daily variations, separately listing the species C_5H_8 due to its unique source, which aligns with the subsequent RIR analysis. The concentration of alkanes is the highest, followed by alkenes, OVOCs and aromatics, while halogenated hydrocarbons and C_5H_8 exhibit lower concentrations (Fig. R5). Furthermore, VOC concentrations for various species are elevated during haze periods compared to clean periods (Fig. R5). As shown in Fig. R6, the diurnal variation of VOC concentrations for various species are not significant during clean periods, likely due to higher wind speeds that facilitate the dispersion of pollutants. In contrast, during haze periods, the daily variations are evident, with peaks occurring before sunrise, followed by a decline, and then an increase after sunset. This is because the haze period is relatively stable at nighttime, which allows for the accumulation of pollutants, while during the daytime, sunlight converts VOCs into photochemical products like O_3 and PAN.

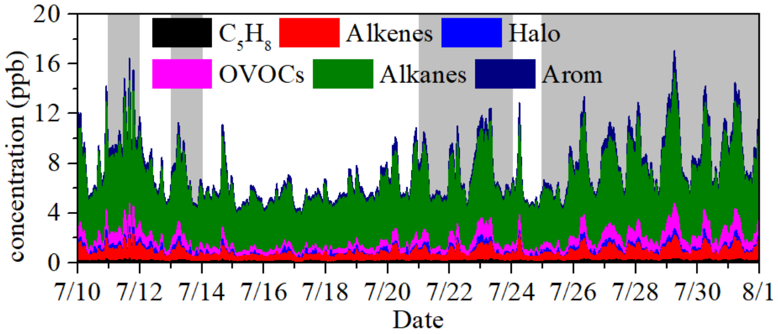

Fig. R5 Time series of VOCs observed at IUE during 10-31 July 2018. The gray shading represents days when the $PM_{2.5}$ hourly daily maximum value exceeded 35 μg·m^{-3}.

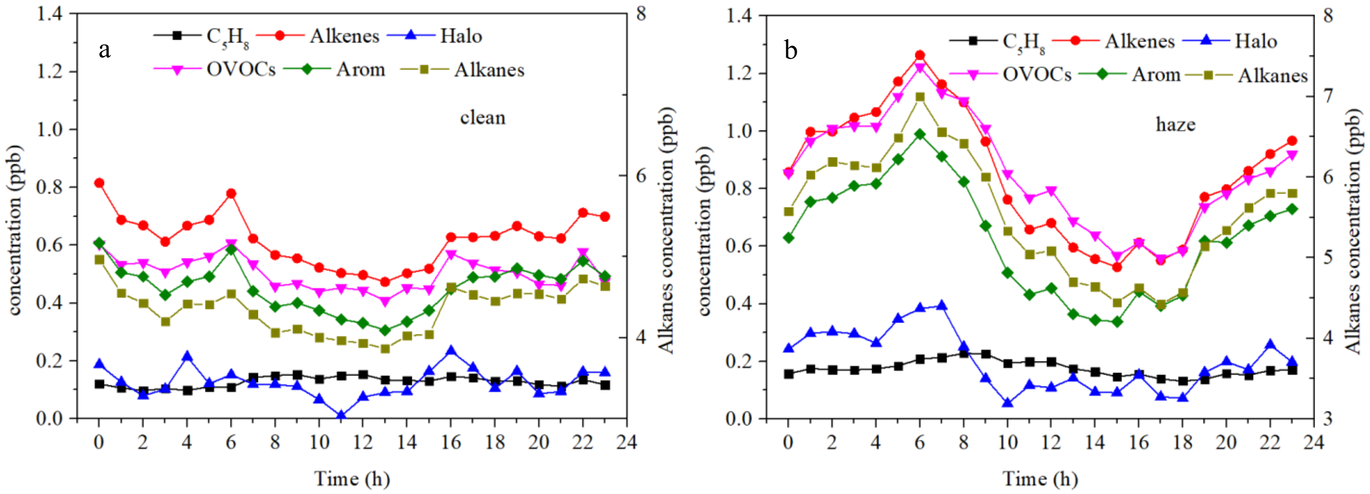

Fig. R6 The diurnal variations of VOCs during clean (a) and hazy (b) periods. Due to the significantly higher concentrations of alkanes compared to other species, the right vertical axis is used for alkanes, while the remaining VOCs are represented on the left vertical axis.

Minor comments:

Abstract: It is not clear at this place of the manuscript where the measurements were conducted.

Response: Thank you for the comment. We have revised the abstract to explicitly state that the measurements were conducted in Xiamen, a coastal city in southeastern China: We addressed the puzzle of summertime PAN formation and its association with aerosol pollution under high O_3 conditions in Xiamen, a coastal city in southeastern China, by analyzing continuous high temporal resolution data utilizing box modeling in conjunction with the master chemical mechanism (MCM).

Line 31-32: The abbreviations IOA, RIR, and EKMA are not defined.

Response: Thank you for your valuable feedback. We have updated the manuscript to include definitions for the abbreviations as follows: With an index of agreement (IOA) value of 0.75, the MCM model proves to be an ideal tool for investigating PAN photochemical formation. The model performed better during the clean period (R^2: 0.68, slope K: 0.91) than during the haze period (R^2: 0.47, slope K: 0.75). Through the machine learning method of XGBoost, we found that the top three factors leading to simulation bias were NH_3, NO_3, and $PM_{2.5}$. Moreover, the net production rate of PAN becomes negative with PAN constrained, suggesting the existence of an unknown compensatory mechanism. Both relative incremental reactivity (RIR) and empirical kinetic modeling approach (EKMA) analyses indicate that PAN formation in this region is VOC-controlled.

Line 38: PAN was also defined in line 20. Maybe just define the terms once.

Response: Thank you for pointing this out. We have revised the manuscript to define PAN (peroxyacetyl nitrate) only once, at its first mention in line 20, and removed the redundant definition in line 38.

Line 43: VOCs already define in line 21

Response: Thank you for your observation. We have revised the manuscript to remove the redundant definition of VOCs (volatile organic compounds) in line 43, as it is already defined in line 21. This adjustment helps maintain conciseness in the text.

Line 76: There is no detailed description of the city of Xiamen, and there is no basic meteorological information to provide an idea of the island. I would suggest including general information to get an idea of the location of the study.
Response: Thank you for your valuable suggestion. In the revised manuscript, we have included a detailed description of the city of Xiamen, along with essential meteorological information: Xiamen is one of the fastest urbanizing regions in the southeast China and is also one of the cities with the best air quality in China, where the air quality could represent the future of other Chinese urban regions. Between 2018 and 2023, Xiamen ranked among the top 10 cities in China, achieving positions of 7th in 2018, 4th in both 2019 and 2020, 6th in 2021, 9th in 2022, and returning to 7th in 2023 (mee.gov.cn, last assessed October 30, 2014). Xiamen is located in a low-latitude coastal area, with abundant sunlight and long daylight hours during the summer, resulting in strong solar radiation and rapid photochemical conversion rates. The city is typically influenced by the East Asian monsoon and serves as a transport channel for atmospheric pollutants from both the Yangtze River Delta and Pearl River Delta regions. Additionally, during the summer, Xiamen is often affected by complex meteorological conditions such as typhoons and the West Pacific Subtropical High (WPSH). The WPSH creates weather conditions that promote the formation and accumulation of photochemical pollutants and particulate matter (Wu et al., 2019). This setting provides an ideal "laboratory" for investigating the complexities of summertime PAN formation and its relationship with aerosol pollution under high O_3 concentrations. In summer, especially in July, high temperatures, high humidity, and intense radiation are likely to accelerate both the formation and consumption rates of PAN.

Line 78: There are 2 definitions of West Pacific Subtropical High (WPSH) and Western Pacific Subtropical high (WPSH) in line 78 and line 164, respectively. Please define once.
Response: Thank you for noticing this duplication. We have revised the manuscript to define the West Pacific Subtropical High (WPSH) only once, at its first mention in line 78, and removed the redundant definition in line 164.

Line 129. Relative incremental reactivity (RIR) is defined in line 129 but it should be defined in line 31.
Response: Thank you for your feedback. We have moved the definition of relative incremental reactivity (RIR) to line 31, ensuring that it is introduced earlier in the manuscript.

Line 175: Consider homogenizing the way you describe Xiamen, in some parts it is as Xiamen City, Xiamen Island or just Xiamen.
Response: Thank you for your feedback. We have standardized the term "Xiamen City" to "Xiamen." However, we retain the designation "Xiamen Island" to specifically refer to the older urban area, which is more densely populated, has heavier traffic, and is economically more developed. As shown in Fig. R7 (a), Xiamen is located in the southeastern coastal area of China, highlighted in light red, while Xiamen Island, which is part of Xiamen, is situated to the south of the observation site (IUE).

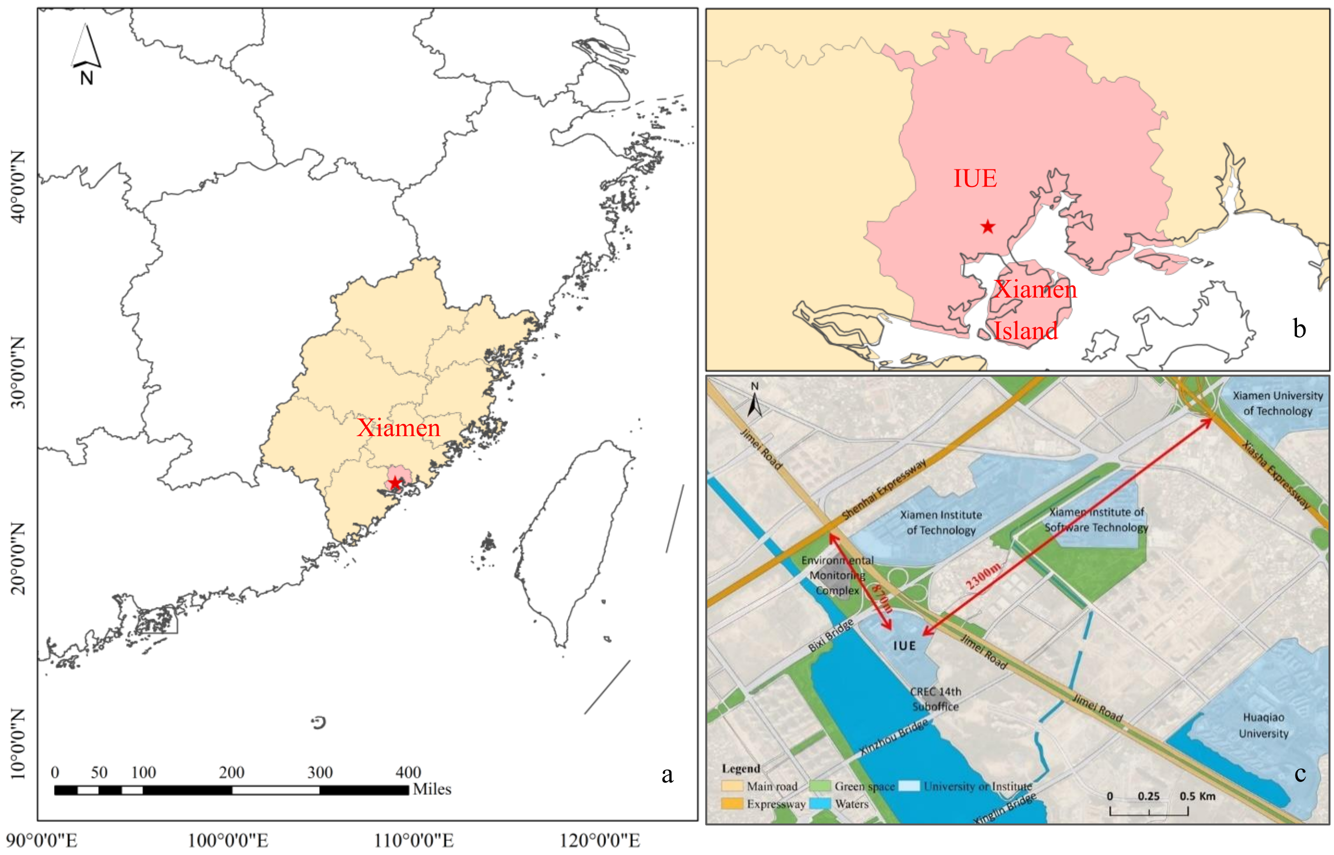

Fig. R7 Location of Xiamen (a), position of IUE in Xiamen (b) and surrounding of IUE (c).

Line 193 Figure 1: Please consider changing the colors of each subplot. The same color for each variable is a bit confusing. Consider including $PM_{2.5}$, in one of the first panels of Figure 1. I consider it important to describe the figure in the main text in order according to the figure itself. It would be easier for the reader.

It is mentioned that the gray area represents the daytime average equal to or greater than 35 ugm^{-3}, however, on July 25, the concentration was lower. I would suggest correcting the shaded period. The last period should be from July 26 to 31.

It is mentioned that O_3 has a high correlation with PAN (R^2: 0.75) however it is noticeable that something happened from July 26 to 31. The O_3 concentration increased (around 110 ppb) from July 26 but remained constant from July 28 to 31. The $PM_{2.5}$ has the same behavior but the trend was different in comparison to the other days shaded by gray. Is there any explanation for this? please clarify this part.

Response: Thank you for your valuable feedback. As shown in Fig. R8, we have changed the colors of each subplot to ensure better differentiation between variables and enhance clarity in the presentation. We prioritize meteorological parameters in Figure 1 due to their significant impact on pollutant variations. Additionally, we describe the weather conditions before discussing pollutants, which involves meteorological data. Following this, we include the most important parameters representing haze ($PM_{2.5}$, BC) and photochemical pollution (O_3 and PAN). Subsequently, we present HONO, which greatly influences atmospheric oxidation, along with key precursors of O_3 and PAN (VOCs and NO_x). We believe this order provides a clearer understanding of the interactions.

We meant to indicate that the gray area represents the hourly maximum value exceeding 35 $\mu g \cdot m^{-3}$ rather than the daytime average. When using a daily maximum value greater than 35 $\mu g \cdot m^{-3}$ to define haze, there are 12 days that meet this criterion, as shown in the gray shading area of Fig. R8. When using a daily average value greater than 35 $\mu g \cdot m^{-3}$, only three days meet the requirement: July 28, 29, and 30. A detailed explanation for why July 26 to 31 was not chosen as the haze period is provided in the first question, so it will not be reiterated here.

In the original manuscript, we refer to the correlation coefficient R, not R^2. R refers to the correlation between the daily maximum values of O_3 and PAN. Considering daily maximum correlations, the correlation between BC and PAN is the strongest, with an R value reaching 0.85, while the correlation between O_3 and PAN has an R of only 0.75. Analyzing hourly values reveals that during haze periods, the correlation between PAN and O_3 is weaker, with an R^2 of 0.47, compared to 0.70

during clean periods. From the 25th to the 31st, a WPSH strengthened and controlled Xiamen, resulting in stable meteorological conditions with light winds (ws=1.04 $m \cdot s^{-1}$), persistently high temperatures (daily maximum average of 37.82 °C), and high relative humidity (maximum daily average of 81.65 %). These factors created an environment conducive to the accumulation of particulate matter and enhanced the photochemical formation of O_3 and PAN. The daily maximum averages of $PM_{2.5}$, O_3, and PAN were 49.26 $\mu g \cdot m^{-3}$, 93.62 ppb, and 1.37 ppb, respectively. The days of the 11th and 21st to 23rd were influenced by the typhoon's periphery, and the 13th also saw the WPSH. Overall, it is evident that when influenced by the WPSH, the trends of $PM_{2.5}$ and O_3 are relatively similar, while under the influence of the typhoon's periphery, their trends diverge. We further analyzed the correlation during haze periods based on whether they were influenced by the WPSH, finding a negative correlation between the daily maximum values of $PM_{2.5}$ and O_3 during typhoon influence, and a positive correlation when influenced by the WPSH, with the correlation significantly increasing from R^2 of 0.14 to 0.71. However, our focus is on the relationship between haze and photochemical pollution, and since the subtropical high is a weather pattern and an external factor, we still choose the daily maximum value being greater than 35 $\mu g \cdot m^{-3}$ as the criterion for determining haze.

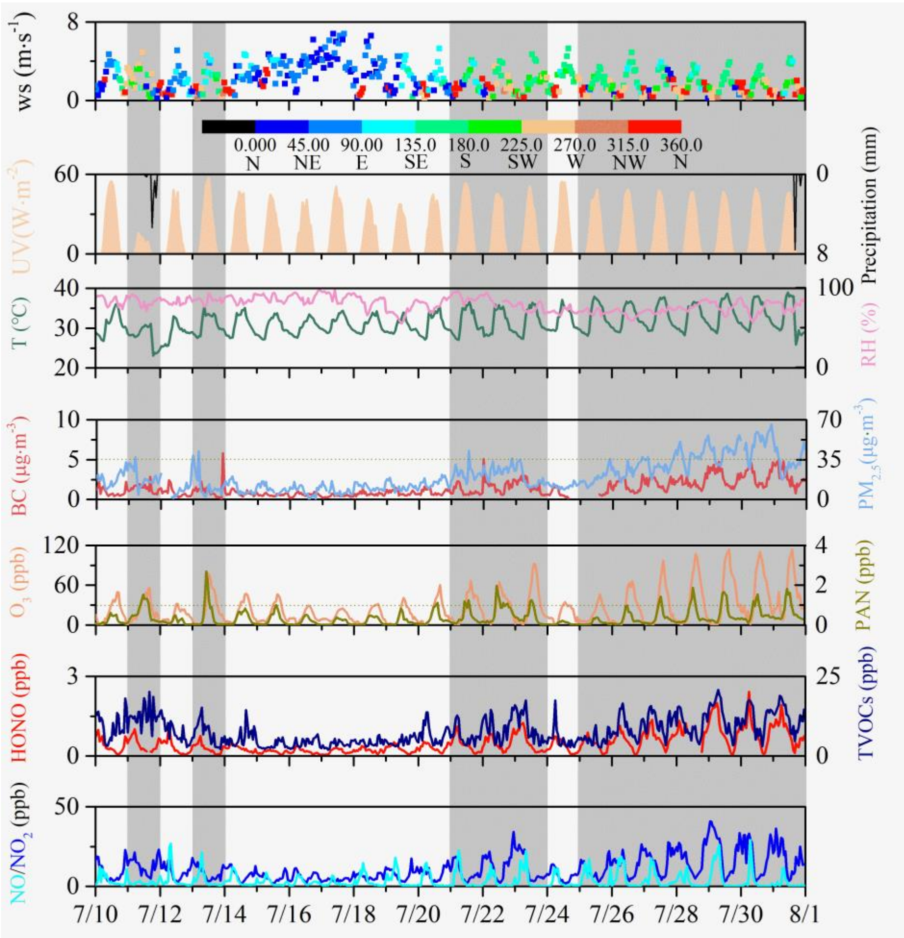

Fig. R8 Time series of trace gases and meteorological parameters observed at IUE during 10-31 July 2018. The gray shading represents days when the $PM_{2.5}$ hourly daily maximum value exceeded 35 $\mu g \cdot m^{-3}$.

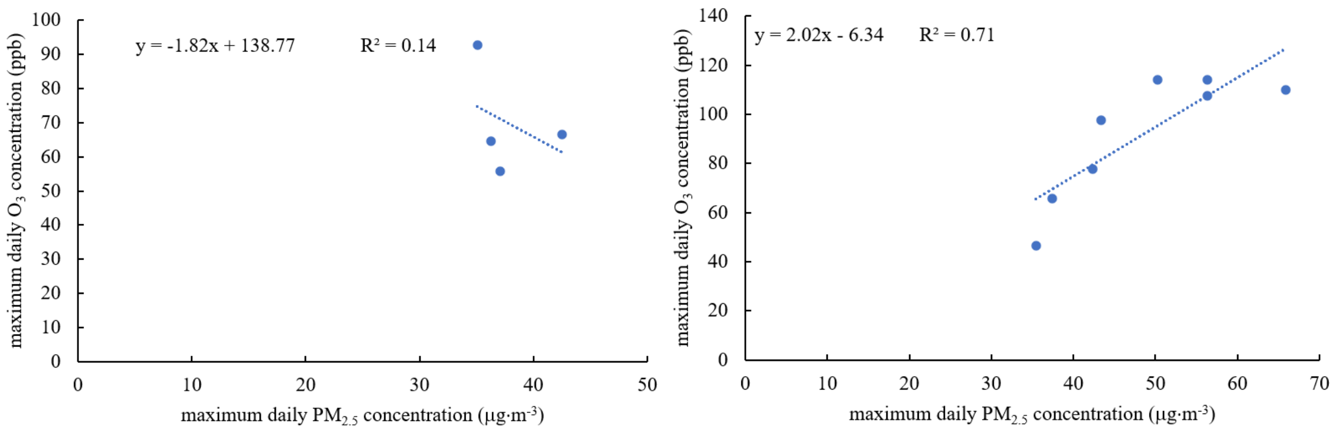

Fig. R9 Correlation between $PM_{2.5}$ and O_3 daily maximum concentrations during haze period affected by typhoon's periphery

(a) and by WPSH (b).

Line 228: it mentioned that PAN has unimodal pattern. Where can we observe that? There is no indication of any figure.
Response: Thank you for your observation. We have added a reference to Figure 2 in line 228 to clarify where the unimodal pattern of PAN can be observed.

Line 242: This information is from Mexico as a country or Mexico City? Please clarify it.
Response: Thank you for your question. We have clarified in the manuscript that the information refers specifically to Mexico city, rather than the country of Mexico as a whole.

Line 301: The font and quality of the figures are different compared to the other figures. Please consider increasing the font and resolution.
Response: Thank you for your feedback. We have increased the font size and resolution of Figure 4 to ensure consistency and improve clarity (Fig. R10).

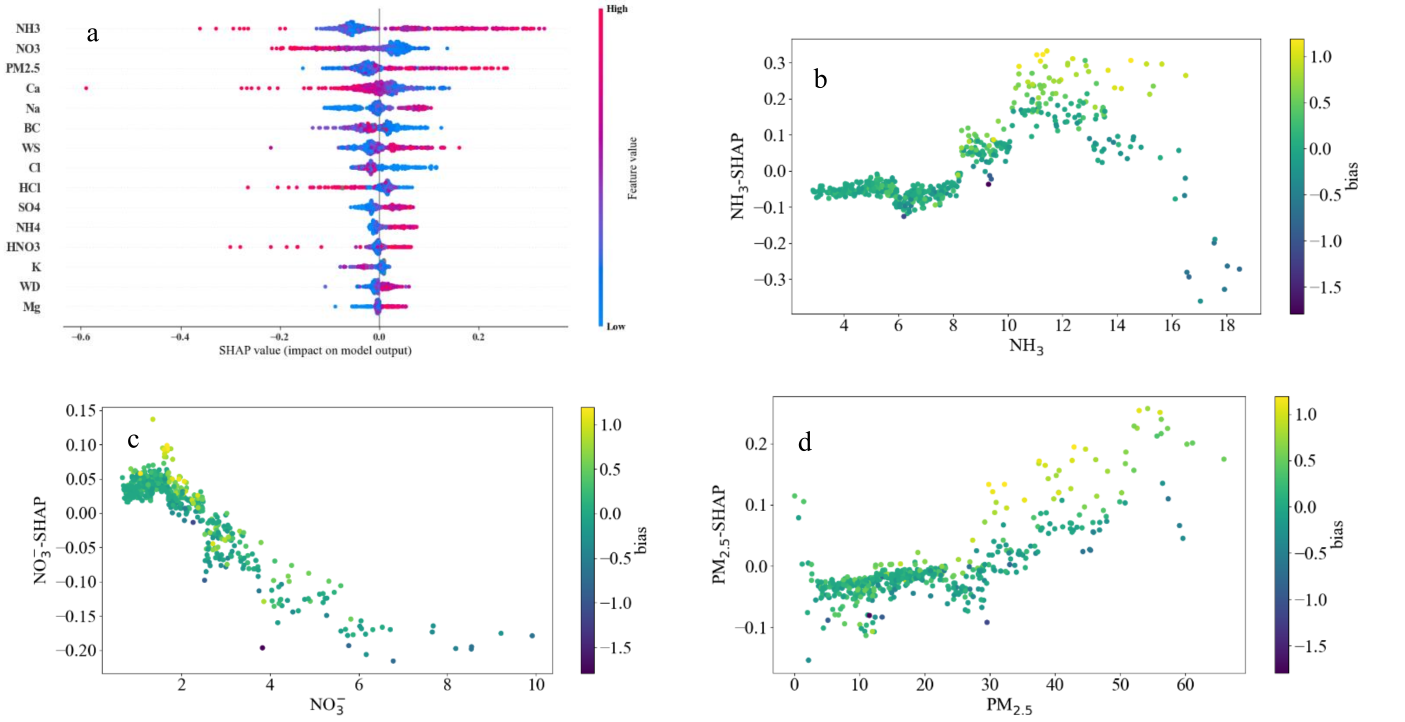

Fig. R10 Feature importance was obtained by XGBoost-SHAP method (a). The scatter plots between concentration of top three important features and their SHAP values (b, c and d), and colored with the bias (the model simulation minus the observed value).

Line 244. According to the main text, PAN undergoes thermal decomposition at high temperatures. However, the specific temperatures are not mentioned, as well as the temperature at 12:00 LT, when the decomposition commonly occurs. Please include this information.
Response: Thank you for your suggestion. The average temperature during the entire observation period was 31.39 °C, with an average temperature of 34.64 °C at 12:00 LT. As shown in Fig. R11, the thermal decomposition of PAN exhibits an exponential relationship with temperature. To highlight the significant impact of temperature, we included the thermal decomposition component in our analysis and found a substantial increase in the slope: from 0.021 during clean periods to 0.1504, and from 0.009 during haze periods to 0.1581 (Fig. R12).

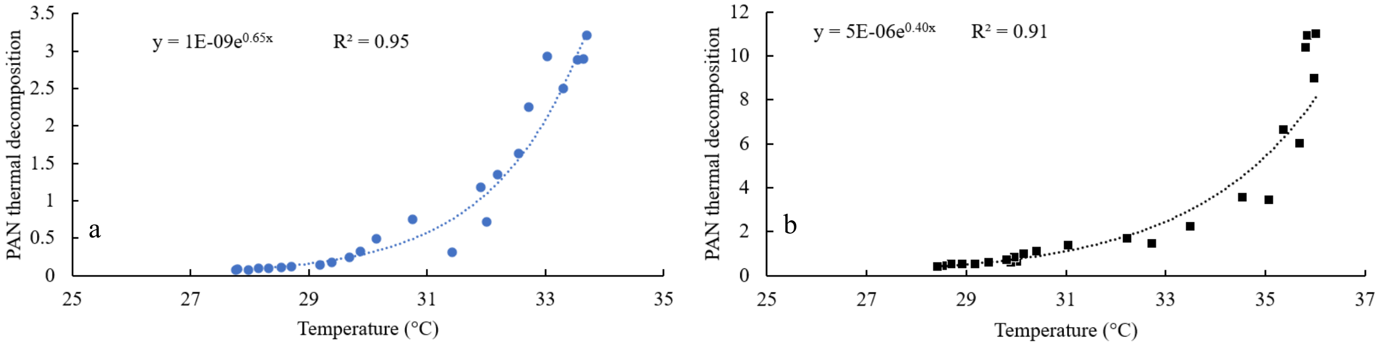

Fig. R11 Correlation between temperature and PAN thermal decomposition during clean (a) and haze (b) period.

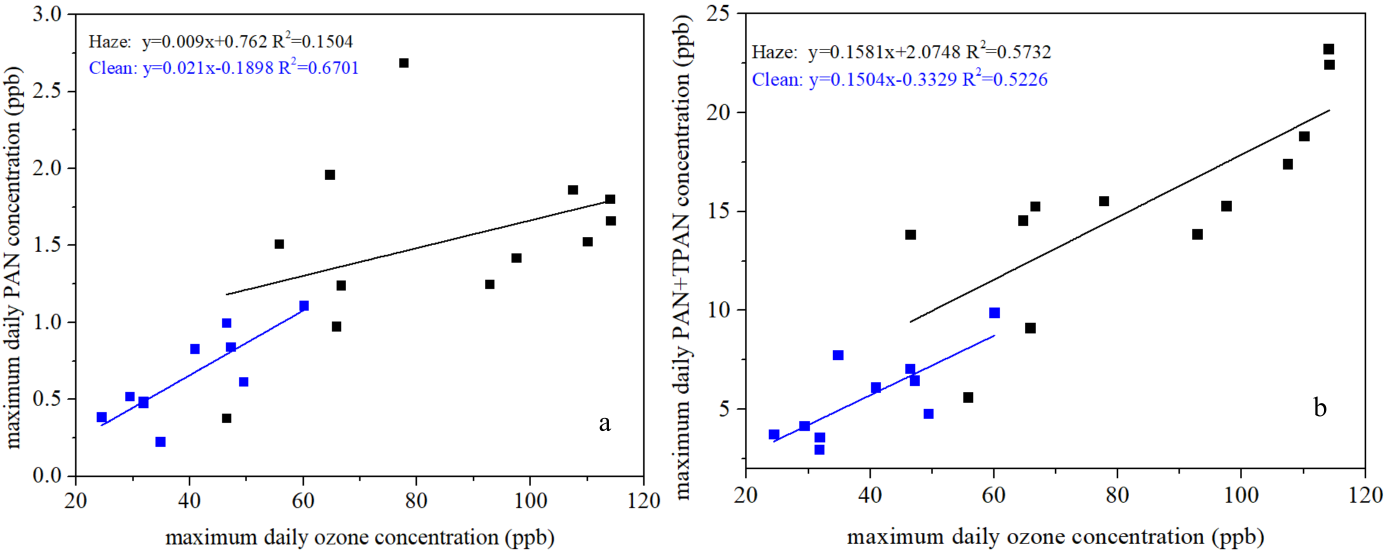

Fig. R12. Correlation between PAN and O_3 daily maximum concentrations during haze and clean (a), and correlation between PAN+TPAN and O_3 daily maximum concentrations during haze and clean (b)

Line 357: The R^2 is reported with a different number of decimal places. This R^2 has 2 number of decimal places but in line 275 R^2 has 4. Please be consistent.

Response: Thank you for bringing this to our attention. We have standardized the reporting of R^2 values throughout the manuscript to two decimal places for consistency.

Line 420: AOC is not defined in the main text.

Response: Thank you for pointing this out. We have added a definition for AOC (atmospheric oxidative capacity) in the main text to ensure that readers have a clear understanding of the term.

Line 452: It mentioned that the temperature is a key factor in determining PAN concentrations; however, it does not specify the temperature at which PAN decomposes or if there are chamber experiments that confirm this. It is also important to review laboratory experiments to accurately determine the decomposition conditions.

Response: Thank you for your insightful comments. In response to the previous question regarding temperature (Line 244), we have added detailed information about temperature and demonstrated the exponential relationship between PAN thermal decomposition and temperature (Fig. R11). Additionally, we noted that incorporating thermal decomposition significantly enhances the efficiency of PAN formation (Fig. R12). Regarding laboratory experiments, we reviewed several studies on the thermal decomposition of PAN. For instance, Tuazon et al. (1990) studied PAN in NO-NO_2 air mixtures over a temperature range of 283-313 K, yielding a rate constant $k = 2.52 \times 10^{16} \exp(-13573/T)\ s^{-1}$. Cox & Roffey (1977) found a rate constant of $k = 10^{14.90\pm0.60} \exp(-104.0 \pm 3.2\ kJ/RT)\ s^{-1}$ over 294-328 K. Senum et al. (1986) reported a rate constant of $k = 2.1\times10^{12} \exp(-24.800 \pm 1800\ cal/(Kmol/RT))\ s^{-1}$ over 298-338 K. Although there are differences in the coefficients from these studies, they all exhibit an exponential relationship, consistent with our findings. We added the following sentence in the revised manuscript:

Previous laboratory experiments also indicated that the thermal decomposition of PAN is exponentially related to temperature (Cox & Roffey 1977; Senum et al., 1986; Tuazon et al., 1990).

Reference

Tuazon, E. C., Carter, W. P., & Atkinson, R. (1991). Thermal decomposition of peroxyacetyl nitrate and reactions of acetyl peroxy radicals with nitric oxide and nitrogen dioxide over the temperature range 283-313 K. The Journal of Physical Chemistry, 95(6), 2434-2437.

Cox, R. A., & Roffey, M. J. (1977). Thermal decomposition of peroxyacetylnitrate in the presence of nitric oxide. Environmental Science & Technology, 11(9), 900-906.

Senum, G. I., Fajer, R., & Gaffney, J. S. (1986). Fourier transform infrared spectroscopic study of the thermal stability of peroxyacetyl nitrate. The Journal of Physical Chemistry, 90(1), 152-156.

Technical corrections

Line 41: Use subscript in NOx like this: NO_x

Response: Thank you for your suggestion. We have updated "NOx" to use the correct subscript format.

Line 53: The space between "+" and "products" is missing.

Response: Thank you for your careful review. We have corrected the spacing issue between the "+" and "products" in line 53 to ensure proper formatting.

Line 56: The space between "+" and "NO_2" is missing.

Response: Thank you for pointing this out. We have corrected the spacing issue between the "+" and "NO_2" in line 56.

Line 80: It is already defined as "O_3" however you mention "ozone" (e.g., lines 68, 81, 204, 232, 432, 435, 436, 438, 439) and "O_3" (e.g., lines 21, 23, 34, 39, 43, 45, 57, 69, 73, 86) in the text. It should be consistent.

Response: Thank you for your valuable feedback. We have revised the manuscript to ensure consistency in the use of terms. We have used "O_3" uniformly throughout the text, aligning with the definitions provided earlier.

Line 93: x in NOx is in italics, but should be in upright font.

Response: Thank you for your observation. We have corrected the font for "x" in NOx to be upright, as suggested.

Figure 2: Please align all legends.

Response: Thank you for your feedback. We have aligned all legends in Figure 2 as requested.

Figure 3: The resolution of the figure needs to improve. The font of the figure does not match the others. Please improve it.

Response: Thank you for your feedback. We have improved the resolution of Figure 3 and ensured the font matches that of the other figures.

Line 306: The figure caption is incomplete. The description of Figure 6b is missing.

Response: Thank you for your observation. We have updated the figure caption to include a complete description of Figure 6b: Feature importance was obtained by XGBoost-SHAP method (a). The scatter plots between concentration of top three important features and their SHAP values (b, c and d), and colored with the bias (the model simulation minus the observed value).

Figure 8: Please change the colors of Figure 8b. It is very confusing with Figure 8a.

Response: Thank you for your feedback regarding Figure 8. As shown in Fig. R13, we have revised the colors in Figure 8b to ensure they are distinct from those in Figure 8a.

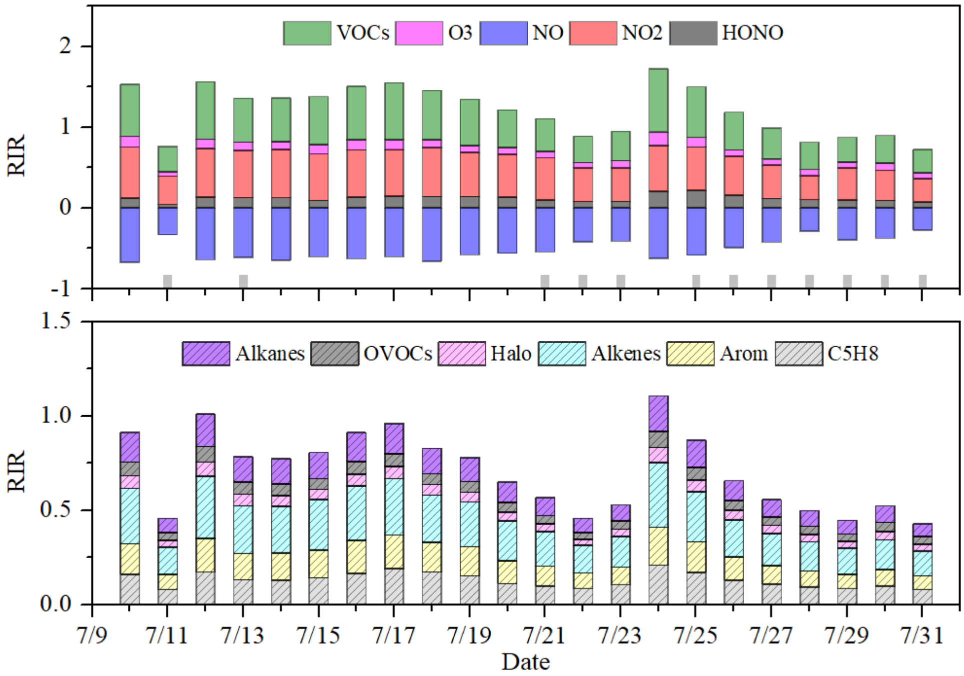

Fig. R13 RIR of PAN formation to major precursors (a), the impact of different VOCs species (b)

Figure 10. Please increase the font of the text and the resolution of the Figure.

Response: Thank you for your suggestion. As shown in Fig. R14, we have increased the font size of the text and enhanced the resolution of Figure 10.

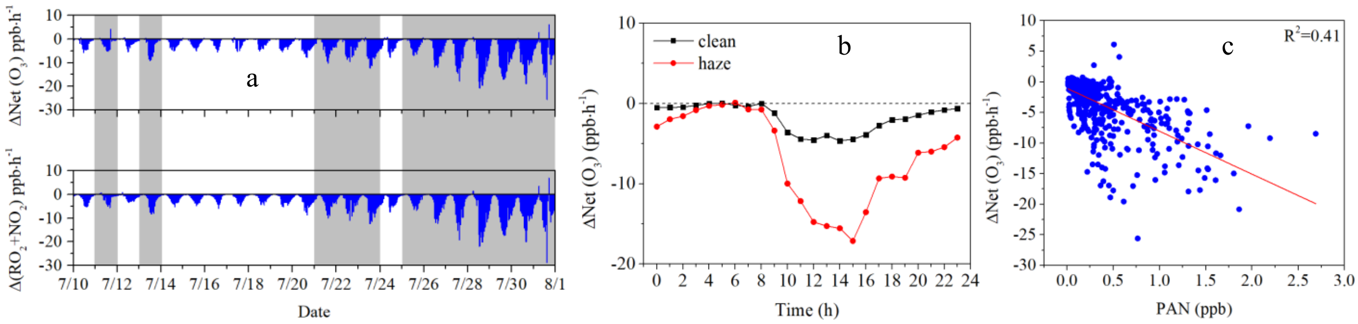

Fig. R14 (a) Time series plot of ΔNet (O_3) and the reaction of Δ(RO_2+NO_2), (b) Diurnal variation of ΔNet (O3) during clean and hazy conditions, (c) Correlation between ΔNet (O_3) and PAN. ΔNet (O_3) is calculated as the base scenario with the PAN mechanism minus the scenario without the PAN mechanism.