- Line 59 – 60: While discussing previous studies on wintertime photochemical air pollution in the manuscript, please cite the source for the statement “Surprisingly high concentrations of OH radical, particularly under hazy conditions, have been observed and are largely attributed to HONO photolysis”.

Response: Added.

- Line 188 – 189: The statement “the precursor concentration of PAN is significantly lower than in the northern region” is not quite clear. Is it meant to be “the precursor concentration of PAN is significantly lower than that in the northern region” or “the precursor concentration of PAN is significantly lower in the northern region”? It would be helpful if the statement can be clarified.

Response: Thank you for your feedback. We have clarified the statement to specify that "the precursor concentrations of PAN, including NO_2 and VOCs, are significantly lower in the studied area compared to those in the northern region."

- Line 190 – 191: Could the authors provide a figure to support the statement “The correlation between the daily maximum values of PAN and BC is the strongest (R=0.85), followed by O3 (R=0.75)”?

Response: Thank you for your suggestion. We have included Fig. R1 in the revised manuscript to visually support the statement “The correlation between the maximum daily values of PAN and BC is the strongest (R=0.85), followed by O_3 (R=0.75)”.

- Line 274 – 276: Both the R2 and K values are discussed in the manuscript, but only the R2 values are defined (Line 152) and shown (Figure 3 (c)). The K values seem to be not defined in the manuscript. It might take the reader some time and efforts to notice that the K values potentially mean the slopes in Figure 3 (c). I would suggest the authors to clearly define the K values in the manuscript.

Response: Thank you for your suggestion. The K has been defined in the revised manuscript: Furthermore, the simulated values are closer to the observed values during clean period, reflected in a higher R^2 value (R^2=0.6782) and a slope value (K) closer to 1 (K=0.9097) (Fig. 3(c)).

- Line 279 – 280 (Figure 3): The legend of the figure shows “obs” and “sim” without their definitions. It would be great if the legend of the figure could be defined in and be consistent with the caption of the figure. For example, the caption could be modified as “Comparison of observed (obs) PAN and simulated (sim) PAN”.

Response: Thank you for your suggestion. We agree that clarifying the definitions of "obs" and "sim" in the figure legend would improve consistency and understanding. We have modified the caption to read: Comparison of observed (obs) PAN and simulated (sim) PAN, ensuring that the definitions are clearly stated.

- Line 283 and 285: Since there are 2 models, a box model and a machine learning model, being used in this study, it would be appreciated if the “target” and the “features” mentioned here are of which model can be clearly stated.

Response: Thank you for your suggestion. We aim to use a machine learning model to evaluate the reasons behind the biases in the box model simulations. To avoid ambiguity, we have added 'OBM' in the sentence: To identify the key factors influencing the performance of the OBM model simulation.

- Line 294 – 295: It is stated that “NO3- is the second most significant parameter influencing the bias between the two, contributing 11.33 %”. However, it seems that the number 11.33 % is not shown in Figure 4 (a). Could the authors clarify whether the contribution of 11.33 % is on average or obtained in some other ways? Would it be possible to show such a value in the corresponding figure as described in the manuscript?

Response: Thank you for your suggestion. This proportion is calculated by taking the absolute values of the SHAP values for all features, summing them up, and then dividing the absolute SHAP value of a particular feature by the total sum. In other words, it represents the average proportion of the absolute SHAP value for each feature during the whole observation period.

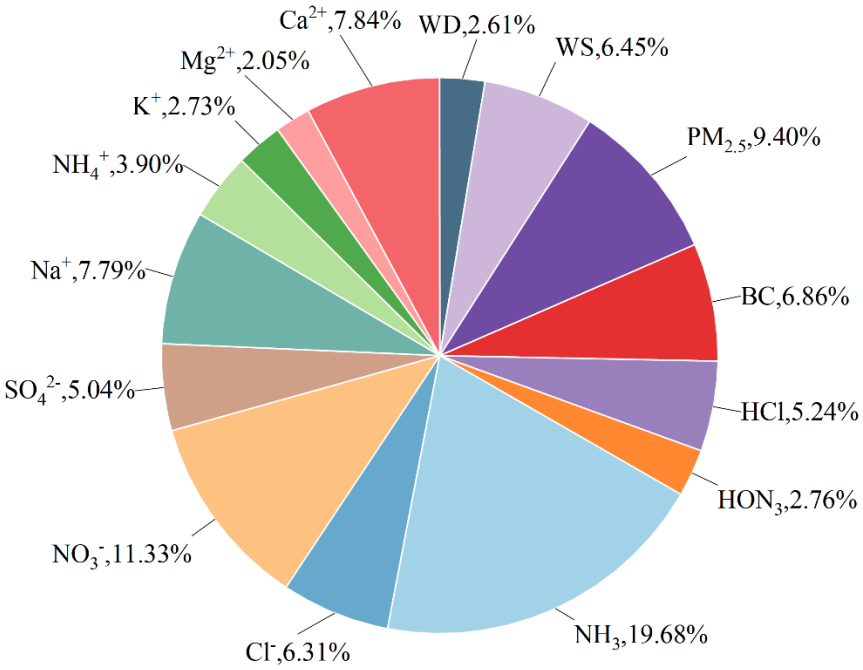

Fig. R1 The average proportion of the absolute SHAP value for each feature during the whole observation period.

- Line 312 – 313: I would appreciate it if it could be clearly stated whether the net production rate of PAN is simulated net production rate or observed net production rate.

Response: Thank you for your question. The net production rate of PAN mentioned in Lines 312 – 313 is based on the model simulation results. We have ensured this is clearly stated in the revised manuscript: From 6:00 to 12:00 during the haze period, the simulated net production rate of PAN is positive, with an average value of 0.19 ppb·h^{-1}. During the clean period, from 6:00 to 12:00, the simulated net production rate of PAN is 0.12 ppb·h^{-1}.

- Line 313 – 314: Is the diurnal variation of PAN based on observation or simulations? I would suggest the authors to clearly state it here.

Response: The result in Lines 313 – 314 is based on observations. We have included this clarification in the revised manuscript: The observed diurnal variation of PAN shows that from 6:00 to 12:00, the average net production rates during the haze and clean periods are 0.20 ppb·h^{-1} (Fig. 2(a)) and 0.09 ppb·h^{-1} (Fig. 2(b)), respectively.

- Line 331, 333, 335, 337, and 347 (Figure 6): Could the authors clearly state in the text and caption of the figure whether the net production rate of PAN is simulated net production rate of PAN or not?

Response: Thank you for your suggestion. We have clarified in the text and the figure caption that the net production rate of PAN refers to the simulated net production rate.

- Line 381 (Figure 7): Are the PA radical production and destruction rates simulated or not?

Response: Thank you for your question. Yes, both the PA radical production and destruction rates were simulated in our study. We have clarified this point in the revised manuscript to avoid any confusion: Figure 7. PA radical production and destruction pathways simulated by OBM on (a) clean days and (b) haze days.

- Line 429 – 430: It would be helpful if the definition of "The difference of HO2, OH, RO2, NO2, and NO between base scenario with PAN mechanism and scenario without PAN mechanism" could be expressed mathematically as what values minus what values for the reader to be clear about the mathematic definition.

Response: Thank you for your comment. We have clarified the mathematical definition by explicitly stating that it represents the values from the base scenario with the PAN mechanism minus the values from the scenario without the PAN mechanism: **Figure 9.** The time series of ΔHO_2, ΔOH, ΔRO_2, ΔNO_2, and ΔNO. The ΔHO_2, ΔOH, ΔRO_2, ΔNO_2, and ΔNO is calculated as the base scenario with the PAN mechanism minus the scenario without the PAN mechanism.

- Figure S6: The term “their product” in the caption should be clearly described along with the mathematic expression $O_3 \times JO_1D$ to be consistent with the axis labels of the figure.

Response: Revised.