

We provide detailed responses to all comments below; the responses are shown in blue, and all revisions in the manuscript are highlighted in red, with specific changes bolded.

The manuscript investigates the triggering mechanisms of new particle formation (NPF) under various O₃ pollution scenarios, focusing specifically on the dynamic balance between precursor sources and the background condensation sink. The study concludes that NPF is highly dependent on the synchronicity between source enhancement and sink attenuation—a process regulated by the evolution of the planetary boundary layer. Overall, this is a highly interesting and well-structured study that provides valuable quantitative insights through the application of interpretable machine learning methods. The methodology is scientifically sound, and the data robustly support the final conclusions. I recommend publication following minor revisions. The authors are requested to consider and respond to the following specific comments:

Response: We sincerely thank the reviewer for the careful reading of our manuscript and for the very encouraging overall assessment of our work. We also greatly appreciate the constructive and detailed comments, which have helped us further improve the clarity, rigor, and overall quality of the manuscript. All comments have been carefully considered, and the manuscript has been revised accordingly. Our point-by-point responses are provided below.

Major comments

1. Section 3.2 emphasizes the breakup of the morning inversion layer and the subsequent rapid vertical mixing as the key process leading to the synchronization of source enhancement and sink weakening. Considering that the temporal resolution of the UAV vertical observations is 3 hours, whereas this process may occur on a much shorter timescale, is this resolution sufficient to capture such rapid evolution?

Response: We are deeply grateful to the reviewers for raising such detailed and critical methodological questions. We agree that processes such as the breakdown of the morning inversion layer and the rapid development of the boundary layer may indeed occur on time scales shorter than three hours. However, we maintain that this resolution remains sufficient to support the conclusions of the present study, for the following reasons:

First, the objective of utilizing UAV vertical observation data in this paper is not to capture minute-scale turbulent processes, but rather to focus on the structural differences and evolutionary trends of the boundary layer across various stages—such as pre-sunrise, post-sunrise, and mid-morning—under different scenarios; a three-hour resolution is sufficient to characterize these stage-specific differences.

Second, under the NPF scenario, phenomena such as the synchronous decline of CS and N_{Acc} around 07:00 (Figs. 2c and S5c), the accumulation of SO₂ between 07:00 and 09:00 (Fig. S4h), and the downward transport of O₃ (Fig. S4d) are all finely characterized by high-resolution ground-based data, and they align highly consistently—both temporally and causally—with the evolution of the boundary layer.

2. Section 3.3: In the Non-O₃ scenario, RH shows a strong positive SHAP contribution, whereas T, SO₂, and O₃ show negative contributions. Could the authors provide more scientific interpretation for this?

Response: We are very grateful to the reviewer for raising this insightful question. SHAP values reflect the marginal contribution of each variable to the model's prediction of the Nuc mode particle number concentration, rather than strictly representing physical causal effects. The differences in the signs of the SHAP contributions under the Non-O₃ scenario essentially reflect the statistical covariation between the individual variables and the Nuc-mode number concentration within a low-ozone background, rather than the NPF nucleation mechanisms themselves.

From a physical perspective, the Non-O₃ scenario is characterized by a combination of low O₃, low T, and low SO₂ levels (Fig. S4)—conditions that fall far short of those required to drive new particle formation. However, high values of these variables tend to occur during the afternoon hours when photochemical activity is relatively high; this coincides with the dilution process of near-surface particulate matter, thereby resulting in a statistically negative correlation with the Nuc mode.

High RH may promote the hygroscopic growth of background particles and influence the distribution of the particle size spectrum in the sub-micron range, leading the model to statistically associate higher RH with higher NNuc values. However, it must be emphasized that this positive contribution does not signify an active NPF process. No distinct nucleation signatures were observed in the Non-O₃ scenario (Fig. S3b), indicating that—even though RH shows a positive contribution to the model's predictions—source limitation remains the dominant factor inhibiting the occurrence of NPF.

Modified content (Section 3.3, Lines 305-314): Under the Non-O₃ scenario, the mean SHAP values of T, SO₂ and O₃ are −94.4, −34.5, and −26.1, respectively, and high-value samples are predominantly distributed in the negative SHAP value region (Fig. 6c). **This may be attributed to the fact that higher values of these photochemically related variables typically occur in the afternoon, when intensified boundary layer mixing dilutes near-surface particles. Such co-variation explains their generally negative SHAP contributions to the predicted Nuc values.** Notably, RH exhibits a high mean SHAP value of 153.5 (Fig. 6a), and its dependence curve increases monotonically from 80% and remained positive thereafter (Fig. 7b), indicating that high humidity conditions are more likely to correspond to higher predicted Nuc values. **This is likely because high RH promotes hygroscopic growth of background particles and modulates the small-size end of the particle size distribution, which the model statistically associates with elevated N_{Nuc} .**

Minor comments

1. Line 29-33: The first paragraph of the Introduction is rather difficult to read and it is recommended to be reorganized.

Response: Thank you for your suggestion. We have rewritten and split the first paragraph of the introduction to make the logic clearer and the sentences more fluent.

Modified content (Section 1, Lines 29-33): Atmospheric new particle formation (NPF) is one of the major sources of atmospheric aerosols (Golden et al., 2017; Jiang et al., 2021) and serves as a critical pathway for the formation of cloud condensation nuclei (CCN) (Matsui et al., 2011; Zhu et al., 2021). **As a result**, it not only significantly impacts air quality (Kulmala et al., 2021; Tang et al., 2021), but also further modulates cloud microphysical processes, radiative balance, and climate effect (Makkonen et al., 2012; Sebastian et al., 2021; Sullivan et al., 2018).

2. Lines 107-109: Please delete the space after the hyphen in “Low- O₃” and “High- O₃”.

Response: We appreciate you pointing this out and have already made the correction.

Modified content (Section 2.2, Lines 108-109): Finally, by combining the NPF event identification with the ozone pollution classification, the observation period was divided into four scenarios:

- (1) NPF: 30 May, 3 June, and 11–13 June;
- (2) Non-NPF (Non-O₃): 31 May, 5 June, and 9–10 June (hereafter Non-O₃);
- (3) Non-NPF (**Low-O₃**): 1–2 June, 6–7 June, and 14–15 June (hereafter Low-O₃);
- (4) Non-NPF (**High-O₃**): 4 June and 8 June (hereafter High-O₃).

3. Line 120: “coagulation sink (Coags, S⁻¹)” is recommended to be revised to s⁻¹.

Response: Thank you for pointing out this unit formatting issue.

Modified content (Section 2.3, Line 120): And the coagulation sink (Coags, s⁻¹) describes the loss rate of target particles due to coagulation with background particles and is defined as (Kulmala et al., 2001a):

4. Line 122: In the equation, should “Cogas” be corrected to Coags? It seems to be a spelling error.

Response: We thank the reviewer’s meticulous review. This was indeed a spelling error. We have consistently corrected "Cogas" to "Coags" throughout the formulas.

Modified content (Section 2.3, Line 123):

$$Coags(D_p) = \int K(D_p, D'_p)n(D'_p)dD'_p$$

5. Section 2.4: Why was GAM chosen instead of random forest or XGBoost?

Response: We thank the reviewer for the careful reading of the manuscript. We selected GAM primarily because the objective of this study is not merely to achieve high predictive accuracy, but rather to identify the relative contributions of various influencing factors to the number concentration of Nuc mode particles under different O₃ pollution scenarios, as well as their nonlinear characteristics.

RF and XGBoost tend to function more as "black-box" models, which can often result in a loss of physical interpretability. Although they typically demonstrate strong fitting capabilities in predictive tasks, for the purposes of this study, our primary focus lies on the marginal effects of different factors and their underlying physical implications, rather than solely on the performance of the model itself.

In contrast, GAM is more easily associated with actual atmospheric physical and chemical processes. While breaking through the limitations of traditional linear regression, it also avoids the shortcomings of deep machine learning models that lack physical meaning, making it more suitable for our mechanism-oriented requirements.

6. Line 178: "As show in Table 1" should be corrected to "As shown".

Response: Thank you for pointing this out. We have made the revisions in the manuscript.

Modified content (Section 3.1, Lines 179-180): As shown in Table 1, the intensity and duration of the NPF process are not controlled by a single factor, instead, they are governed by the dynamic balance between precursor strength (source) and background particle removal capacity (sink).

7. Line 181 & Table1: "Cogas" is also a spelling error here.

Response: We thank the reviewer for pointing out this error. All "Cogas" in the manuscript have been uniformly corrected to "Coags".

Modified content (Section 3.1):

Line 182: Taking June 13th as an example, although both CS ($6.20 \times 10^{-2} \text{ s}^{-1}$) and Coags ($18.74 \times 10^{-5} \text{ s}^{-1}$) reached the highest levels among the five events, meaning that newly formed particles and precursors would be strongly removed, which is theoretically unfavorable for nucleation.

Table1:

Table 1. Summary of key kinetic parameters for the identified NPF events.

Date	FR ($\text{cm}^{-3} \cdot \text{s}^{-1}$)	GR ($\text{nm} \cdot \text{h}^{-1}$)	C ($\times 10^7 \text{ cm}^{-3}$)	Q ($\times 10^6 \text{ cm}^{-3} \cdot \text{s}^{-1}$)	CS ($\times 10^{-2} \text{ s}^{-1}$)	Coags ($\times 10^{-5} \text{ s}^{-1}$)
------	--	---	--	--	---	--

2018.05.30	0.10	1.73	2.37	1.88	5.61	8.90
2018.06.03	0.03	5.99	8.21	2.73	3.30	6.56
2018.06.11	0.05	1.11	1.53	0.63	3.43	6.07
2018.06.12	0.25	0.38	0.52	0.51	4.77	9.80
2018.06.13	0.71	5.80	7.95	3.81	6.20	18.74

8. Lines 219-220: “In contrast...” is repeated from Lines 165–166, please delete it.

Response: Thank you for the correction. We have removed the duplicate statement.

9. Lines 226-227: The statement “According to the classification criteria in Sect. 2.2, both the NPF and Low-O₃ scenarios had similarly low O₃ pollution levels (160-215 µg/m³)” does not seem to be mentioned in Section 2.2. Please clarify.

Response: Thank you for this insightful comment. The observed O₃ concentrations during the NPF days also belong to the Low-O₃ range (160 < MDA8-O₃ < 215 µg/m³). This detail was originally omitted in the text for brevity. We apologize for any confusion this omission may have caused and have now revised the manuscript to clarify this point.

Modified content (Section 2.2, Lines 106-107): Finally, by combining the NPF event identification with the ozone pollution classification, the observation period was divided into four scenarios:

- (1) NPF (Low-O₃): 30 May, 3 June, and 11–13 June (hereafter NPF);
- (2) Non-NPF (Non-O₃): 31 May, 5 June, and 9–10 June (hereafter Non-O₃);
- (3) Non-NPF (Low-O₃): 1–2 June, 6–7 June, and 14–15 June (hereafter Low-O₃);
- (4) Non-NPF (High-O₃): 4 June and 8 June (hereafter High-O₃).

10. Line 254: The citation “Ali et al., 2025” is inconsistent with Ali in the References (Line 392); please revise it.

Response: We thank the reviewer for this comment. We have made corrections.

Modified content (Section 3.2, Line 255): This enhancement was likely associated with persistently high near-surface RH (Fig. S4b), which promoted hygroscopic growth of background particles and thereby increased CS (Ali et al., 2025).

11. Lines 229-300: The phrase “turning negative beyond approximately 0.085 m·s⁻¹” is discussing CS, so the authors are advised to check the unit. Should it be s⁻¹ instead?

Response: We thank the reviewer for the careful examination, and we note that this comment likely refers to Lines 299–300. The unit designation here was indeed incorrect, and we have

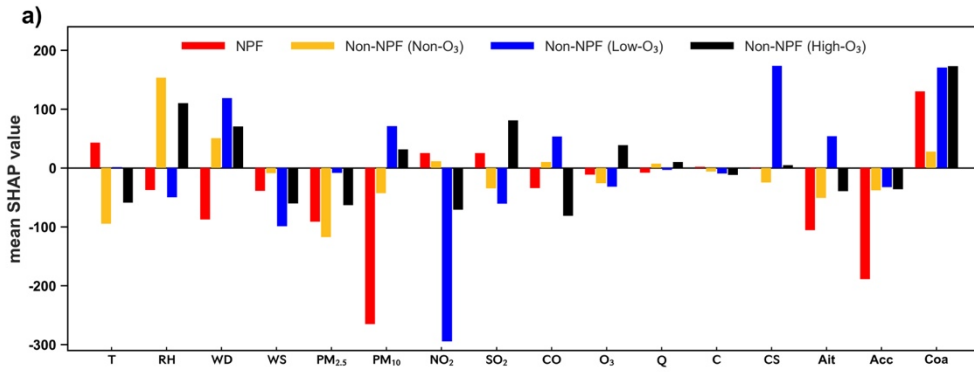
since corrected it.

Modified content (Section 3.3, Line 301): Its dependence plot shows that the fitted curve initially increases and then decreases with increasing CS values, turning negative beyond approximately 0.085 s^{-1} (Fig. 7m), consistent with a mean SHAP value of -0.9 (Fig. 6a).

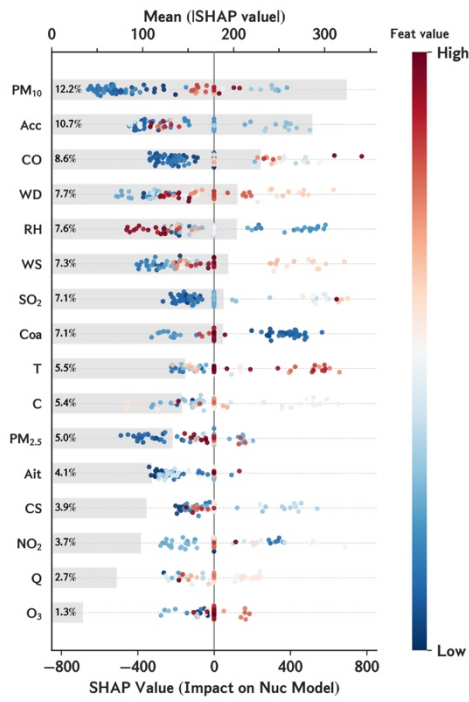
12. Lines 316–317: Please revise the caption of Figure 6, as “(a) Mean SHAP values across for scenario” is not grammatically smooth.

Response: We sincerely appreciate your careful review. The word “for” was indeed a typographical error and should have been “four”. We have corrected this spelling mistake and further polished the entire caption of Figure 6 for better clarity.

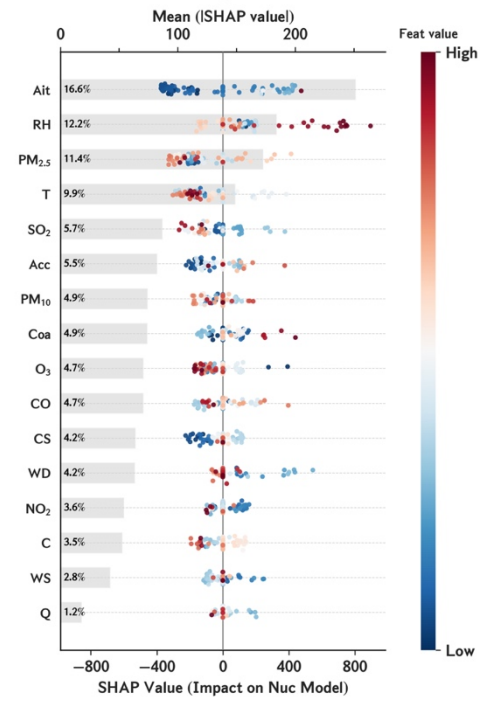
Modified content (Section 3.3, Lines 321-327):



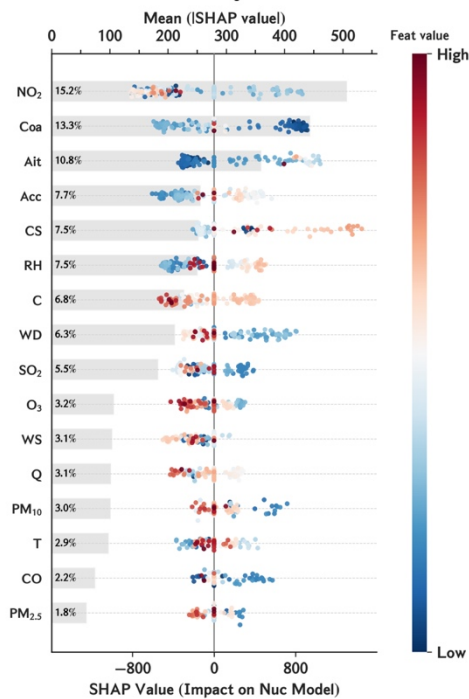
b) NPF



c) Non-NPF(Non-O₃)



d) Non-NPF(Low-O₃)



e) Non-NPF(High-O₃)

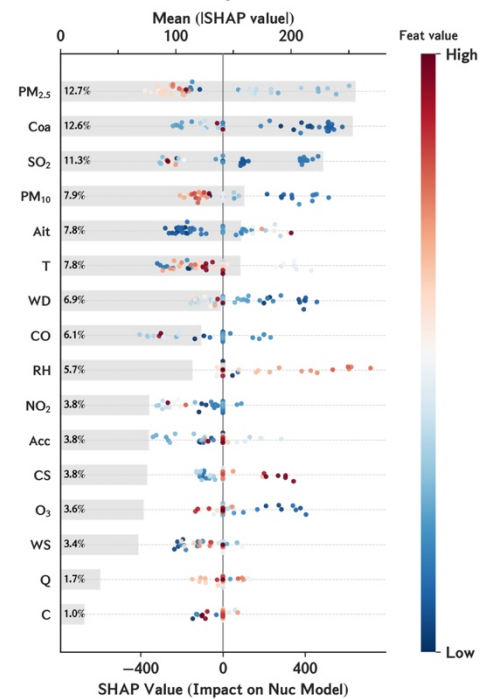


Figure 6. SHAP-based analysis of drivers for particle number concentrations in the Nuc mode. (a) Mean SHAP values across the **four** scenarios, with red, yellow, blue, and black bars represent the NPF, Non-NPF (Non-O₃), Non-NPF (Low-O₃) and Non-NPF (High-O₃) scenarios, respectively. (b-e) Feature importance and summary plots for (b) NPF, (c) Non-NPF (Non-O₃), (d) Non-NPF (Low-O₃), (e) Non-NPF (High-O₃) scenario. **The bar charts indicate the mean absolute SHAP values, while the scatter plots show the distribution of SHAP values, with colors transitioning from blue (low feature values) to red (high feature values).**