Dear Editor,

We would like to express our gratitude for the comments provided by the reviewer. We have revised our manuscript to address these comments and improve the quality of our work. Your patience and assistance throughout the review process are greatly appreciated.

Your sincerely,

Xiaohong

Prof. Xiaohong Yao (Ph.D)

Ocean University of China

Response to Reviewer 1

This is a good manuscript that should be published. It is well written and well organized. I have a few suggestions.

1. Are the times given standard time or daylight time? During summer most countries advance the clock 1 hour. According to the longitudes noted standard time should be 8 hours in advance of GMT for 120 degrees east. This is important for sun time and for anthropogenic influences that are generally a function of local time. In the time plots are the hash marks with date at the beginning of each day (00:00) or are they at the middle of each day (12:00)?

Response: Thanks for your suggestions. Local time is used through this study. In each time plot, the harsh marks with the data represent the beginning of each day at 00:00. This has been clarified in the revision.

2. Reveal how many data points are in each of the correlation plots such as Fig. 2b, c and d.

Response: The numbers have been added in the revised Fig 2b (N=117), c (N=176) and d (N=186).

3. Fig. 3 the abscissa should have equal intervals. E is not designated.

Response: Corrected.

4. Fig. 5 is not cited in the text.

Response: In the original version, Fig. 5 is cited in the P6. L14-17.

5. P1. L36. Delete (cloud condensation nuclei). Already defined it in L31.

Response: Deleted in the revision.

6. P2. L10. Insert these after of. L11. Delete events. L21-22. Delete to some extent.

Response: Done.

7. P3. L32. Remove ing. disappear.

Response: Done

8. P4. L4. Add ing to play. L14. Delete was prepared for. Change ments to d. change between to from. Change and to to. Change comma to period. Change with to Thus it had. L20. Delete used for. Change tion to ted. Insert here after evaluated.

Response: Done

9. P5. L1. J should be the same J as in L2. It is different. L16. Delete last the. L17. Delete period from. Change to -. Insert good agreement after not. Change and to or. L18-19. Define mean fractional bias and mean fractional error. L20-21. Move US EPA in front of benchmarks. Explain what such a benchmark is and why it is relevant to this research. L24. Insert not after are. Delete to be. L25. Change try to tried. Change and to to. Delete poor. L26. Change The to those. Insert However, before those. Change for to to. Change thereby to therefore.

Response: Corrected. In the revision, we clarified the benchmark and it reads "No benchmark is available regarding the statistical metrics for simulated atmospheric particle number concentrations but we adopt the benchmarks and the goal values widely used in air quality studies for $PM_{2.5}$ mass concentrations (US EPA, 2007, the benchmark with absolute MFB: ≤50 %; MFE: ≤75 %; the goal values with absolute MFB: ≤30 %; MFE: ≤50 %). During the NPF events period, the MFB of 24 % and the MFE of 66 % on June 29–July 6 met the benchmarks."

10. P6. L4. Delete 1st and 2nd the. L25. Delete with the events being most frequently. L26. Delete observed on. Delete 2nd and 3rd July. Delete occurrence. L28. Delete meteorological. L30. This needs further explanation.

Response: Done. The details on the events were presented in the companion paper, which has been accepted for publishing in ACP. The reference has been cited here in the revision.

11. P7. L2. Change as no to because. Insert not after were. L5. Delete respectively. Nighttime should include two dates. L6. But there was no data for most of July 5. L9. Delete The. L23. Delete the. L25. Add s to present. L35. What date?

Response: Done.

12. P8. L13. Delete values. L26. Explain respectively. L38. Add s to show.

Response: Done. The revision for L29 reads as "Fig S6a-c present the comparison of CCN simulation with the observation under 0.2 % SS during the NPF events on July 1, 3, and 6, respectively. The comparison shows that the simulated Nccn at 0.2 % SS were clearly underestimated by several folds." We also revise the title of Fig S6 accordingly.

13. P9. L20. Insert greater after even.

Response: Done.

14. P10. L11. Change ing to ed. L29. Probably because of the lack of data on that

date.

Response: Done. The sentence at L29 (original version) has been revised as "However, the occurrence frequency of preexisting particle growth was much less than that of NPF events on basis of the observations in this study alone".

15. P11. L21. Delete 1st the. L22. Delete 1st the.

Response: Done.

Response to Reviewer 2

The manuscript entitled "Investigating the contribution of grown new particles to cloud condensation nuclei with largely varying preexisting particles - Part 2: Modeling chemical drivers and 3-D NPF occurrence" is a modeling study of new particle formation events in the North China Plane. Modeling studies such as these can be helpful for understanding new particle formation, as they provide possible insights into difficult-to-measure aspects of new particle formation such as the horizontal and vertical extent of events. The often-repeated quote "All models are wrong, but some are useful." (attributed to George Box) is a useful mantra to consider when assessing the scientific value of a study such as this. Certainly, models such as WRF-Chem are quite sophisticated in their ability to handle some chemical and meteorological conditions, but I feel this paper also displays some of the pitfalls of modeling studies, such as their reliance on accurate emissions data for compounds that are actually quite difficult to measure. In my opinion, the most significant weakness of this manuscript is the lack of objective standards applied to the assessment of model performance: specifically the use of statements such as "the model does a good job …" without objective standards to base such statements on.

In my view, this manuscript may be publishable in Atmos. Chem. Phys., but only after addressing numerous concerns outlined below. My comments are preceded by a numbering system that states the page and line number of the passage in question.

Response: We thank the constructive comments and revised the manuscript accordingly.

comments

1. 5-16: The use of the term "good" is not really useful. What means "good" to the authors does not mean the same to the reader. Indeed, I would object that the agreement is good here, even in the time period selected by the authors (and I agree that the fit is horrible outside the period). I strongly suggest removing "good" here and letting the statistical tests tell the story so that the reader can make a judgement on whether this is acceptable.

Response: Agree. We revised the issue through the manuscript.

2. 5-21: If there is a benchmark provided by the USEPA then I would say that going back to my prior comment "meets the benchmarks" should be replaced with "good". But I cannot find this in the references. If this is an oversight. the authors need to provide a reference to this benchmark.

Response: We are sorry that the cited report was missed in the references. The report has been added in the revision. In the Appendix B of Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone,

PM2.5, and Regional Haze, Vol EPA-454/B-07e002, USEPA, 2007 (Page 243-253), it summarized model performance evaluations for PM$_{2.5}$, and recommend MFE $\leq$75%, absolute MFB $\leq$50% as the benchmarks. In the revision, it reads "No benchmark is available regarding the statistical metrics for simulated atmospheric particle number concentrations but we adopt the benchmarks and goal values widely used in air quality studies for PM$_{2.5}$ mass concentrations (US EPA, 2007, the benchmark with absolute MFB: $\leqslant$50 %; MFE: $\leqslant$75 %; the goal values with absolute MFB: $\leqslant$30 %; MFE: $\leqslant$50 %). During the NPF events period, the MFB of 24 % and the MFE of 66 % on June 29–July 6 met the benchmarks."

3.5-28: Again, I strongly suggest that the authors refrain from making this sweeping statement because, as they state in subsequent sentences, the fit appears to be good for sulfate but quite bad for nitrate and ammonium.

Response: Agree. Revised through the whole manuscript.

4. 5-34: It's easy to blame measurements here, but I think it's even more likely that the model is not getting source emissions correct. For their analysis to be respected I suggest that the authors reflect on how their model could be getting things wrong as well.

Response: Agree. In the revision, it reads "Although the differences of simulated concentrations from observed values can be due to the model weaknesses such as poorly predicting meteorological parameters and estimating air pollutant emissions, and lacking of key mechanisms (US EPA, 2007; Matsui et al., 2011; Liu et al., 2021; Shen et al., 2022), it should be noted that the poor performance in overestimating the observed organics, NO$_3^-$, and NH$_4^+$ could also be partially attributed to sampling artifacts, given their higher volatility compared to ammoniated sulfate acid (Yao et al., 2002; Chow et al., 2010)."

5. 5-35: I am not sure why we are talking about Beijing measurements here. It doesn't really seem to add a lot to this study, since it's a different locale.

Response: We are sorry that we cannot agree with the comments. NPF events occurred for several hours and should be considered as so-called regional NPF events. Thus, it is important to examine the modeling results in a large region, if possible.

6. 6-14: there are a lot of plots in the main manuscript and some, like Figs 4 and 5, I think can easily go into the supplement. That is because they are referring not to the site that is the focus of this study, but to an urban Beijing site.

Response: Agree. The two Figures have been moved to SI in the revision.

7. 6-20: I would maintain that the correlations between measured and modeled ammonium are just about as bad as those for nitrate, even during the "NPF" period. Please comment on this.

Response: Agree. In the revision, it reads "For the modeled NH_4^+, the MFB (-1%) and MFE (57%) met the benchmarks during the frequent-NPF period (Fig. 4g). Excluding the data on July 5 only slightly increased R from 0.41 to 0.45."

8. 8-30: I would say that there are poor correlations for all plots, especially given that the data are presented on a log-scale. If the authors think that a correlation is "reasonable," then please let the reader know by what basis you judge this.

Response: The senescence has been revised as "However, the model reproduced N_{ccn} at 0.4 % SS during the NPF events on July 1 and 3 with the performance meeting the benchmark (Fig 10d-e)."

9. 11-2: Again, if there is a metric for assessing whether the model did a good job it's important to state that rather than just say it did a good job.

Response: The sentence has been revised as "Specifically, the model reproduced CN_{10-40}, N_{ccn} at 0.4 % SS, mass concentrations of $PM_{2.5}$, mass concentrations of SO_4^{2-} in $PM_{1.0}$ and TSP, ORG and NH_4^+ in $PM_{1.0}$, and other variables, with the performance meeting the benchmark."

Minor comments:

1. 1-30: In the abstract especially, it is important to avoid ambiguous statements. Thus I would avoid the term "CCN size," since the ability of a particle to become CCN depends on size and composition as well as SS.

Response: Agree and correct.

2. 2-12: These reviews are quite old. Please supplement with more recent reviews on the topic.

Response: More recent review papers have been added in the revision, i.e., NPF studies in China by Chu et al., 2019; NPF studies at mountain sites by Sellegri et al., 2019, and NPF mechanisms by Lee et al., 2019.

3. 2-30: replace with "becoming CCN" since size AND composition is important.

Response: Correct.

4. 2-31: There have been many recent studies of NPF in the NCP, so what do the authors mean about "updated study"? What about those studies have not addressed this question?

Response: The part reads as "After considering the situation outlined above, it is imperative and essential to conduct an updated study that quantifies the diverse contributions of chemical drivers to grow newly formed particles becoming CCN and to modify pre-existing particles simultaneously over NCP (Wei et al., 2023)." in the revision.

5. 3-1: This line of reasoning suggests that 3D modeling is on par with measuring approaches, but each has unique aspects that they contribute and each has issues. The discussion needs to be more nuanced.

Response: To be more nuanced, it reads as "Although observations are important to characterize NPF events and explore related mechanisms, the use of one fixed-site to observe condensable vapors is not sufficient to explain NPF occurring downwind, and it is difficult to perform Lagrangian observations with moving air masses. Thus, three-dimensional (3-D) modeling studies are needed to determine where NPF events initially occurred."

6. 3-9: remove "In the literature"

Response: Done.

7. 3-25: "in terms of" is not the correct phrase. replace with something like "focusing on"

Response: Done.

8. 3-32: I am not sure what this means for a modeling study to discuss observations and what a "particle signal" refers to. Certainly "disappearing" should be replaced with "disappear" in order to be grammatically correct ... but what is a "signal"?

Response: Thanks. We agree that "particle signal" is not a right scientific term. In the revision, it reads "Section 3.5 will address what happened for grown new particles after the particles disappeared from observations." The issue has been corrected through the whole manuscript accordingly.

9. 4-3: Usually when a sentence starts with "on one hand" it is because later there will be a different perspective provided (the "other hand"). Why is this phrase necessary for this sentence?

Response: The authors trust the language-editing by ChatGPT. The phrase has been deleted in the revision.

10. 4-13: what does this mean "a combination was used"? For what? I also note that an SMPS system always includes a CPC. Is this an additional one used for some other

purpose? I think the authors mean to say that an SMPS was used for the size distribution during this time.

Response: The sentence has been revised as "For measurements taken between June 14 and 30, 2019, both the SMPS and the additional CPC were used."

11. 4-17: it would be helpful for the reader to know about the inlet dimensions (length and diameter) in order to assess the degree to which losses might be important.

Response: The information has been added "The instruments were located on the third floor of the main station building and connected ambient using the conductive silicone tubing (TSI Inc., product number 3001788, inner diameter 0.19 inch, outer diameter 0.375 inch) in ~ 2 m length."

12. 5-10: units need to be added on each statement of the rate constant, even though you state the units above.

Response: Added.

13. 5-32: Shouldn't Fig. 4 be referenced here? If not, how is this statement justified?

Response: It is Fig 2b-d. Added.

14. 6-33: This is likely not due to the shrinkage of grown new particles, but rather an air mass change that brings in particles of differing age. Please rephrase this. Perhaps adding wind direction to the plot would also provide insights into the cause of this change in the size distribution.

Response: The part has been removed in the revision. However, the deletion does affect the science presented here. The observed new particles at nighttime had a younger age in comparison with those observed in the afternoon, is it possible? The issue is far complicated and beyond the scope of this study. The authors thereby delete the related description through the manuscript.

15. 7-9: Please comment on how and to what extent the model accounts for vertical mixing. In many studies, NPF is found to be initiated aloft and vertical transport is seen to be an essential part of the complete picture of the mechanisms of NPF. What does this model suggest? If this is discussed later in the paper, a reference to that here might be appropriate.

Response: We are sorry that we cannot capture "*how and to what extent the model accounts for vertical mixing.*" All 3-D air quality model can quantify vertical mixing, but the modeling results need to be verified by observations. However, vertical mixing observations were not available in this study. The same is generally true for most of modeling studies in the literature.

In the revision, we added "More discussion on PNCs in vertical direction can be found in Section 3.3. "

16. 7-11: Apologies if this was already stated earlier but what are the "goal values" and how were they assigned?

Response: It has been revised accordingly and reads as "When considering NPF and non-NPF days separately, the simulated N_{ccn} at SS = 0.2 % met the goal values on non-NPF days, with MFB = 19 % and MFE = 48 % (Fig. 7a)." The goal values have been defined in the revision, page 5, lines 24-29.

17. 7-27: How does the fact that the model dramatically overpredicts both ammonum and nitrate during these two days affect the interpretation of the modeling results? I would suggest a discussion of this at least. After reading the next paragraph, it seems that this topic comes up so perhaps including some reference to this here would be appropriate. Perhaps something like "an analysis of kappa confirms that this could be an overestimated effect"

Response: In the revision, we added "An analysis of kappa presented later in this section confirms that this could be an overestimated effect."

18. 8-19: I suggest putting Fig 9 into the supplement since there are so many plots shown in the main text.

Response: Done.

19. 8-36: I thought that the inclusion is that the model overestimated the organic composition during the NPF period? Again, the use of the word "reasonably" allows for a broad interpretation.

Response: The sentence has been revised as "Based on the fact that the model reproduced CN_{10-40} and organic drivers, that lead to the growth of newly formed particles at the ground level, to meet the benchmark during June 29–July 6,"

20. 9-4: use of "steer" in this context seems odd. It is not a synonym for control. This is probably a result of using ChatGPT for grammar (which I would NOT recommend!).

Response: It has been changed to "control" in the revision.

21. 9-18: Similar to my previous statements, I think it's OK to just show one of these plots (Fig 11 for instance) and put 12 in the SI.

Response: Done.

22. 9-22: "3D occurrence" is an odd phrase. I would suggest "3-D spatial distribution" or "3-D evolution"

Response: Correct.

23. 9-35: Could this observed spatial inhomogeneity contribute to the "shrinking" effect observed in the size distribution? It seems possible.

Response: We delete the description related to particle shrinkage through the manuscript since it is beyond the scope in this study.

24. 10-11: correct grammar "disappear". Also what does "signals" refer to? THis is a fundamental question for understanding this entire section.

Response: The sentence has been revised as "What happened for grown new particles after the particles disappear from observations?"

25. 10-14: See my question about about "signals no longer observable". In this entire section the word "signal" needs to be defined or replaced.

Response: Agree. It change to "the particles" in order to avoid the confusion through the whole manuscript.

26. 10-35: Lagrangian

Response: Done.

27. 11-4: I would disagree with Ncc(0.4%). I also don't think it did very well with some of the species mentioned later in this sentence.

Response: The sentence has been revised as "Specifically, the model reproduced CN_{10-40}, N_{ccn} at 0.4 % SS, mass concentrations of $PM_{2.5}$, mass concentrations of SO_4^{2-} in $PM_{1.0}$ and TSP, ORG and NH_4^+ in $PM_{1.0}$, and other variables, with the performance meeting the benchmark."

28. 11-13: Why, if growth is dominated by organics, is understanding inorganic species the key for modeling NPF?

Response: To be clear, the sentence has been revised as "This implies that the critical challenge in modeling contributions of NPF events to CCN budget may be accurately reproducing those inorganic species, accounting for a small but appreciable fraction, rather than SOA."

29. Fig 8 e-h - I personally feel that the 3D representation of these pie charts is distracting and unnecessary. In general I support the thought that the simpler way to represent data (without what Tufte called "Chartjunk") is best

Response: Revised accordingly.