

This study applies machine learning (random forest) to the spatiotemporal prediction of cloud condensation nuclei (CCN) concentrations in the heavily polluted North China Plain. It significantly reduces the CCN simulation bias from -59% to approximately -31% , and consequently lowers the simulation uncertainty of cloud radiative forcing from $1.89 \pm 0.78 \text{ W m}^{-2}$ to $0.81 \pm 0.63 \text{ W m}^{-2}$. This provides a new method for accurately assessing aerosol–cloud interactions and the climate benefits of pollution control. Moreover, by incorporating observational data such as PM_{2.5}, NO₂, and SO₂, the model captured the long-term decreasing trend in aerosol concentration from 2014 to 2018 and quantified the reduction in cloud radiative forcing uncertainty achieved by mitigating N_{CCN} simulation biases. This topic is highly relevant to the journal of Geoscientific Model Development. The model incorporates the CCN concentrations simulated by WRF-Chem as an input, with the observed CCN as the target. This represents a reasonable bias-correction approach. The methodology and results are interesting. Thus, I suggest several minor modifications to be made before publication.

Minor Concerns:

Line 160: "TAP" is used without definition; please provide the full name (Tracking Air Pollution in China) at first use.

Line 162: The OM and sulfate in the TAP dataset are "underestimated by approximately 50%" and therefore apply a "twofold correction factor." However, it is unclear whether this correction is applied during the training phase, the prediction phase, or both. If the correction is applied only during the prediction phase while the training phase uses uncorrected data, this would lead to a distribution mismatch between training and inference.

Line 199-201 and 276-277: The sentence "the campaign mean mass concentration of PM_{2.5} ranges from 35.6 to 160 $\mu\text{g m}^{-3}$, indicating that the observations can represent various atmospheric conditions" is repetitive with the introduction in Section 2.3.1.

Line 236: The authors interpret the SHAP values as evidence that OM has strong hygroscopicity driving CCN activation. However, SHAP values reflect association rather than causation. It is recommended to add a qualifying statement: "While SHAP indicates a strong association, the causal interpretation should be supported by the known hygroscopicity of OM in the NCP (Liu et al., 2021)."

Line 289: "During the GC2018_WIN campaign, the observed N_{CCN} is underestimated by as much as 71% by WRF-Chem (Fig. S6). Here, it might be referring to Fig. S5?"

Lin 289-296: It notes that the model's performance improves much more during severely polluted winter conditions than during cleaner summer conditions. It is recommended to emphasize more explicitly in the Conclusions that the improvement is particularly pronounced under highly polluted (cold-season) conditions, which has practical implications for CCN prediction in heavily polluted regions.

Line 297: "The improvements in RFRM model also demonstrate the effectiveness of the model trained on atmospheric variables to revise the simulation in model". The

phrase “in model” at the end of the sentence is unclear in meaning.

Line 298: The article has already used another observation at GC site to provide independent spatiotemporal validation. This should be discussed in detail to demonstrate that the model's generalizability has been effectively verified.

Line 307: “the underestimation of N_{CCN} by the WRF-Chem model is likely due to the overestimation of the organics and BC mass fraction induced by WRF-Chem (Fig. S8), **but** the underestimation of the hygroscopic parameter of organics, and the simplified prescriptions in particle size distribution”. “But the clauses before and after 'but' do not indicate a contrast.”

Line 329: Here they used long-term PNSD measurements and κ -Köhler theory to calculate the "observed" annual mean N_{CCN} (N_{CCN_obs}). However, the κ values themselves are derived from the TAP dataset (which is biased even after correction). Moreover, the authors note that " κ values are much less sensitive to changes in NCCN compared to the PNSD" (Lines 345–347). A sensitivity analysis is needed to quantify the impact of κ uncertainty on the final N_{CCN_obs} .

Line 476: In the “Limitations and outlook” section, the authors honestly acknowledge that the observational data come from only six campaigns at three sites. It is recommended to add a discussion on the spatiotemporal representativeness of these observations.

Line 732: Figure 2 shows an R^2 of 0.86–0.95 for the test set, but Figure 3c shows an R^2 of only 0.86 (RFRM vs. observations). These two R^2 values are calculated for different targets (the former may be RFRM vs. WRF-Chem? The latter is RFRM vs. observations). This should be clearly stated in the figure captions.

Line 732: In Figure 2b, the SHAP plot lacks units or dimensions. Please clarify in the caption: “SHAP values represent the contribution to N_{CCN} (in cm^{-3}).”

Line 737: Use N_{CCN} consistently throughout the manuscript. For example, in the title of Fig. 3 on page 34, it reads “Time series of the observed and predicted CCN number concentrations”, while the axis label reads “ N_{CCN} (cm^{-3})”

The manuscript uses both "RFRM" (Random Forest Regression Method) and "RF model" (line 138). It is recommended to use "RFRM model" consistently throughout.