Reviewer Report on Manuscript: "Optimizing CCN predictions through inferred modal aerosol composition – a boreal forest case study"

The manuscript provides a thorough and insightful exploration of aerosol-cloud interactions, focusing specifically on cloud condensation nuclei (CCN) closure studies in boreal forest environments. Using a robust, multi-year dataset from SMEAR II (2016–2020), the authors employ forward and inverse modelling approaches to evaluate the impact of size-dependent chemical composition and hygroscopicity on CCN predictions. The study addresses key uncertainties in climate modelling and attempts to constrain modal aerosol composition using CCN observations — an approach with significant scientific merit. The manuscript presents an extensive dataset of ~6,200 concurrent ACSM, CCN and size distribution observations from Hyytiälä, Finland. The manuscript evaluates three methods for predicting CCN concentrations and explores seasonal variability in aerosol hygroscopicity and composition, with a particular focus on the differences between the Aitken and accumulation modes. This work addresses a significant challenge in the field of aerosol-cloud-climate research, particularly in the context of future scenarios involving declining anthropogenic emissions and increased contributions from natural aerosol sources. The authors implement a novel inverse closure technique using optimization to estimate mode-specific κ values. This comprehensive study is highly relevant to the atmospheric sciences community.

The study is well structured, with a clear methodology and appropriate referencing to prior work. However, there are some areas where the clarity could be improved, potential ambiguities could be resolved and minor corrections could be made to enhance the quality of the manuscript. The dataset is extensive and the topic is timely. Nevertheless, clarification or expansion of several methodological and interpretational aspects is required before the manuscript can be considered for publication.

Major Comments

1. The optimization of size-resolved composition is central to this study, yet the method remains a bit opaque. It is not very clear: What parameters are varied during optimization? Are any physical constraints or priors imposed (e.g., known hygroscopicity bounds for organics/inorganics)? Is the optimization performed independently per time point, season, or SS level?

2. While the optimized CCN predictions agree better with observations, the inferred composition remains unvalidated. Have the authors compared the mode-specific organic/inorganic fractions with any available independent chemical data (e.g., AMS, offline filters, or PTR-MS)? Without this, there could be a risk of overfitting the CCN closure.

Furthermore, the large seasonal variability in composition (e.g., +156% inorganic in winter) should be discussed in the context of known aerosol processes—such as wintertime transport, boundary layer dynamics, or nucleation suppression.

3. The improvement is mainly observed above 0.5% supersaturation. The authors should also discuss the implications at low SS and whether the optimization technique could be adapted or constrained to better handle this regime.

4. The optimization suggests that accumulation mode particles are more enriched in inorganics while the Aitken mode is more organic-richer. This is plausible, but a more mechanistic

explanation is needed. For example: Is this pattern consistent with condensation of oxidized VOCs on smaller particles and aqueous-phase processing on larger particles? How does this compare with seasonal biogenic activity or anthropogenic influence?
Including a more detailed interpretation supported by prior literature would strengthen the conclusions.

5. While the research questions are outlined at the end of the introduction, the manuscript would benefit from explicitly stating the working hypothesis earlier (perhaps around line 55 or 70). Suggest rephrasing and condensing the goals for better readability and alignment with subsequent methodology.

6. The inverse modeling framework is a major novelty in this work but is not adequately introduced in terms of assumptions, mathematical implementation, or validation strategies. Clarify what "inverse aerosol-CCN closure" means in practical terms—e.g., optimization method, objective function, constraints used.

7. The manuscript discusses organic aerosol extensively but does not explain how the complex properties of organics (e.g., surface tension depression, limited solubility) are accounted for in κ parameterization or closure attempts.

8. There's an implicit assumption that Hyytiälä data can be generalized to other forest regions or clean continental environments. This assumption should be stated explicitly and discussed in the limitations.

9. The study notes persistent overprediction errors not resolved by optimized κ values. It would strengthen the work to more directly explore model structural assumptions such as: constant surface tension, neglecting semi-volatile partitioning, mixing state (internal mixing assumption for size modes).

10. The assumption of stable size distributions during the CCN cycle is critical. Was this stability verified using size-resolved time series? Otherwise, this assumption should be treated more cautiously.

11. The text mentions calibration frequency for CCNc and invalidation criteria for aethalometer data (e.g., RH > 40%). Please clarify how data gaps or invalid data points were handled in the analysis. Were interpolation or gap-filling methods used? What fraction of data was excluded due to quality control?

Minor Comments

Line 43: "Aerosol particles are important in the formation…" consider rephrasing as "Aerosol particles play a critical role in the formation…"

Line 44: Check the phrasing. Suggest: "...by lowering the energy barrier for the heterogeneous nucleation of water, thus promoting cloud droplet activation…"

Line 46-47: Rephrase: "thereby changes in the CCN concentration" to "thus, changes in CCN concentration"

Line 57: "drivers of SS_{max} fluctuations". Define "SS_{max}" explicitly on the first use for clarity.

Line 68: Suggest moving the sentence “These inverse approaches…” earlier to clarify the inverse model’s novelty and importance.

Line 93: “Still, organic aerosol plays a significant role…”. Consider beginning with “Nevertheless,” to better connect to prior sentence.

Line 123: Confusing sentence. Suggest: “Specifically, median κ was 0.41 at 0.1% SS (corresponding to larger activation diameters), and 0.14 at 1.0% SS (smaller activation diameters)...”

Line 129: Add a clarifying phrase on what “systematic overprediction” means quantitatively.

Lines 196-214: The site description is thorough, but additional discussion on how representative the SMEAR II site is for boreal forest aerosols under varying seasonal anthropogenic influence would be valuable.

Line 310: “with by a Nafion dryer” to “with a Nafion dryer” (remove “by”).

Lines 314-315: The ACSM and eBC data are averaged over 1-hour intervals but converted to a 2-hour median to match CCN measurements. Please discuss potential impacts of this temporal averaging on capturing short-term variability in aerosol composition and CCN. Were any tests performed to ensure this does not bias the results?

Line 659–661: Consider rephrasing for clarity: “The relative difference in the median Aitken and accumulation κ…” to Perhaps “The seasonal variability in median κ between Aitken and accumulation modes is most pronounced in winter (~162%)...”

Line 676–678: Repetition – consider merging: “observed CCN concentrations are a valuable tool… Our study uses this approach…” to avoid redundancy.

Line 687–688: Suggest citing more recent or diverse κ-related parameterization studies for broader context.