Responses to Reviewer #1 Comments

This study investigates the seasonal and size-dependent variability of the aerosol mixing state index (χ) based on hygroscopicity parameter probability distributions from multiple sites, and establishes a climatological dataset of aerosol mixing state. The authors find that ambient aerosol populations generally tend toward internal mixing rather than the external mixing limit. Overall, the study is built upon a promising methodological framework and reports interesting observational findings. However, the depth of the analysis needs to be strengthened before the manuscript can be accepted for publication.

> We thank the reviewer for the thoughtful assessment. We have revised the manuscript to clarify the scope of the study, strengthen the discussion of methodological limitations, and align the interpretations with the evidence provided by the observations. In particular, we now define χ explicitly as a hygroscopicity-based metric under the binary LH–MH framework, distinguish it from composition-resolved chemical mixing state, and present source and process explanations as hypotheses unless supported by independent observations. Researchers who contributed the MHD and HACPL observations and their interpretation have joined the revised manuscript as co-authors. The page and line numbers referenced below correspond to the marked-up version of the revised manuscript.

Several specific aspects warrant further improvement:

1. The observed variations in mixing state are frequently attributed to diverse aerosol source populations (e.g., marine vs. continental, anthropogenic vs. biogenic). However, the current analysis relies solely on κ-PDFs and χ, without direct support from chemical composition measurements or backward trajectory analysis.

> We thank the reviewer for this constructive suggestion. We agree that the original source and process interpretations were not sufficiently supported by independent observations. We therefore added available chemical-composition measurements (Figs. S11–S13) and 72 h back-trajectory analyses (Figs. 5–6 and S14–S18). The composition measurements provide context for temporal changes in the aerosol mixture, while the trajectory analysis identifies the main transport pathways and their corresponding size-resolved χ distributions. We also removed unsupported source-specific assignments and limited the discussion to interpretations supported by the available composition and transport evidence. The corresponding additions and revisions are provided in Section 4.2.

2. The study focuses primarily on seasonal and size-dependent patterns. However, the diurnal cycle of aerosol mixing state can provide critical insights into the competing effects of primary emissions (e.g., traffic rush hours) and atmospheric aging. A comparative analysis of diurnal variability across the multiple sites should be add.

> We thank the reviewer for this suggestion. We have added a diurnal analysis for seven sites. Representative results for CRG, EPC, HOU, SGP, MHD, and BJIAP are shown in Figure 4, the filtered and unfiltered ENA results are provided in Figure S10. HACPL is not included because its diameter-specific hourly κ distributions were sparsely sampled.
>
> We added the diurnal analysis in Section 4.1 (Figures 4 and S10). The calculation is described as follows:
>
> - P14L331–338: "Figure 4 shows the diurnal cycles of χ at six sites. For each date and measured diameter, observations within each available hour were combined into a κ-PDF, from which an hourly χ value was calculated. The plotted median and interquartile range at each local clock hour summarize those hourly χ values across the available dates. HACPL was excluded because hourly measurements at each diameter were too sparse."

3. The authors may consider to use machine learning approaches (e.g., Random Forest or XGBoost combined with SHAP analysis), which could be employed to identify and rank the relative importance of meteorological parameters (e.g., temperature, RH, boundary layer height), chemical precursors, and source proxies.

We thank the reviewer for this suggestion. We conducted an exploratory machine-learning analysis and present its design and results here in the response (Figures R1–R3). We do not include it in the manuscript or Supplement because the available predictors differ among sites and the results do not support the requested common ranking. Figure R1 evaluates target and predictor availability at one representative accumulation-mode diameter per site (150 nm, except 165 nm at MHD). Eligibility required at least 200 valid daily χ observations spanning at least 330 d, at least three seasons with 15 or more observations each, and one non-temporal predictor with at least 75% coverage. EPC, HOU, SGP, ENA, MHD, and HACPL met these criteria. CRG and BJIAP were excluded because their available records did not provide sufficient duration and seasonal representation for the blocked validation design. Predictor-category coverage also differs substantially among the eligible sites.

For each eligible site, we compared three XGBoost models using temporal predictors, process-related predictors, or both predictor groups against a baseline that predicts the mean training-set χ for each season. Process predictors were grouped as meteorology, chemical composition, trace gases and source proxies, particle number, and physical 72 h trajectory metrics. Comparison with the seasonal baseline tests whether the XGBoost models capture variability beyond the seasonal cycle.

Model performance varied substantially among sites (Figure R2). Only EPC showed clear predictive skill beyond the seasonal baseline, with a positive R^2 and lower prediction error. At HOU, ENA, MHD, and HACPL, some models reduced the error, but their R^2 confidence intervals included zero, while SGP showed no improvement. The analysis therefore provides evidence for predictive relationships at EPC but does not support consistent prediction or attribution across the sites.

We calculated SHAP values and grouped them by predictor category (Figure R3). Temporal variables have the highest SHAP-based importance at EPC, meteorology has the highest importance at HOU, and transport has the highest importance at ENA and HACPL. At MHD, temporal variables and meteorology have similar contributions. Given the uncertain predictive skill outside EPC, these results do not support a robust ranking of the processes controlling χ across sites.

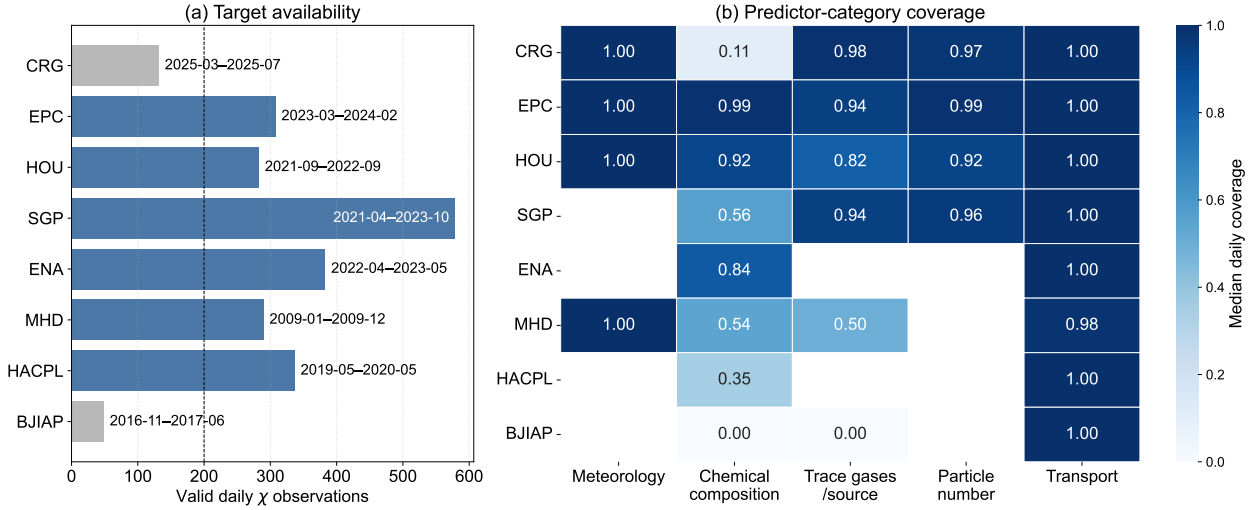


Figure R1: Availability of daily χ at the representative accumulation-mode diameter (150 nm, except 165 nm at MHD) and of the predictors used in the exploratory machine-learning analysis. (a) Number and measurement period of valid daily χ observations. Blue and gray bars denote eligible and excluded sites, respectively, and the dashed line marks the 200-day criterion. (b) Median daily coverage of variables in each predictor category. Blank cells denote unavailable categories. ENA observations were screened using the receptor-local contamination filter.

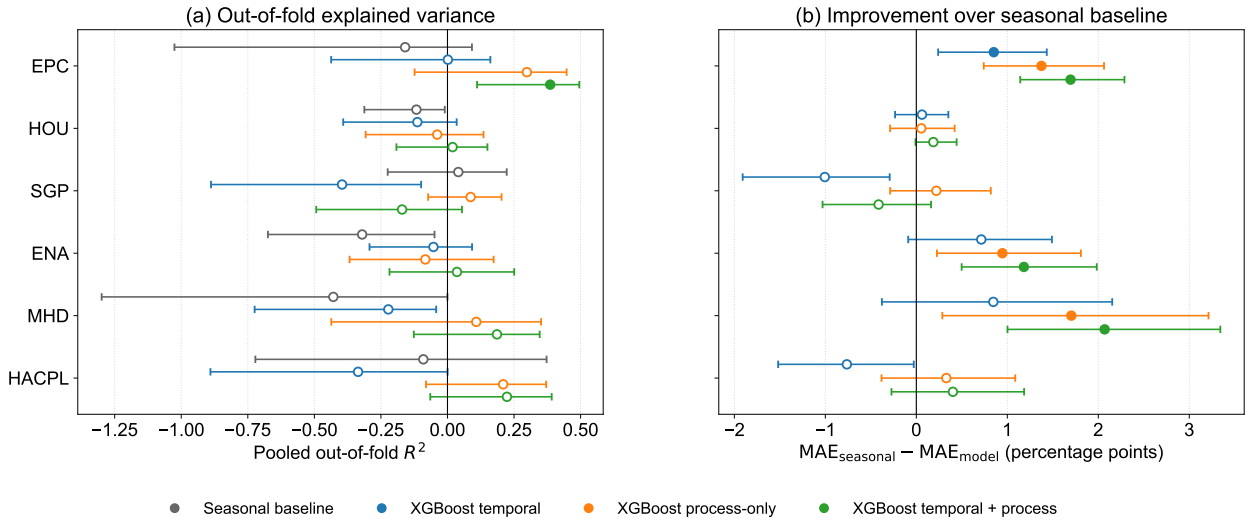


Figure R2: Predictive performance of the seasonal baseline and XGBoost models using temporal predictors, process-related predictors, or both. Each site record was divided into time blocks. Each block was predicted using a model trained on the other blocks, and the combined predictions were used to calculate the out-of-fold (OOF) metrics. (a) OOF R^2. (b) Reduction in OOF mean absolute error relative to the seasonal baseline, with positive values indicating lower error. Error bars show 95% confidence intervals estimated by resampling observations in calendar-month blocks. Filled markers indicate intervals entirely above zero, whereas open markers indicate intervals that include zero.

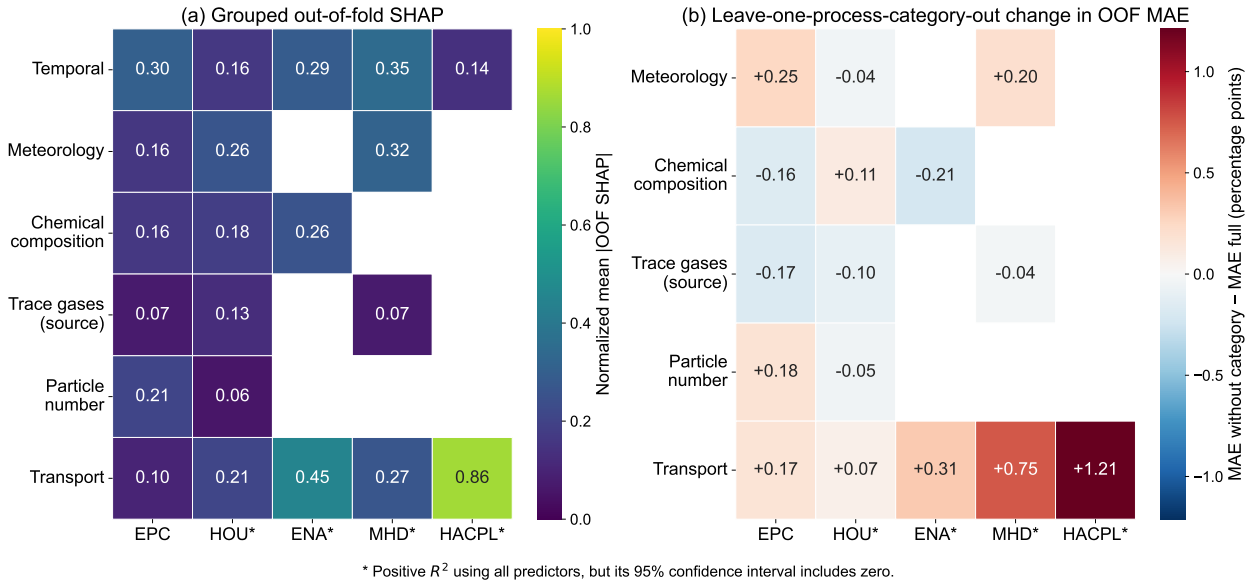


Figure R3: Importance of predictor categories in the XGBoost model using temporal and process-related predictors. (a) Normalized mean absolute SHAP values. (b) Change in prediction error after removing each predictor category. Positive values in (b) indicate increased error. Blank cells denote unavailable predictor categories. An asterisk indicates that R^2 is positive using all predictors, but its 95% confidence interval crosses zero.

Specific technical and interpretative concerns:

1. Line 70: At the marine-influenced sites, κ_{MH} is set to 1.5 to represent highly hygroscopic species such as sea salt. A sensitivity test using alternative κ_{MH} values would help quantify the robustness of the retrieved χ values and should be included.

We thank the reviewer for raising this point. The upper end members are operational bounds used to calculate χ, rather than fixed hygroscopicity values assigned to particular chemical species. We selected the baseline values to encompass the resolved structures of the measured κ-PDFs without truncating their upper modes, while remaining consistent with reported hygroscopicity ranges. We have now quantified the dependence of the retrieval on this choice. The results are presented separately for the marine and non-marine sites in Figures S3–S4. The alternative end members produce systematic shifts in retrieved χ, with ENA showing the greatest sensitivity.

We made the following changes:

- P4L88–102: "For the continental, urban, coastal-urban, and mountainous sites, we follow Liu et al. (2026) and prescribe $\kappa_{LH} = 0$ as the non-hygroscopic limit and $\kappa_{MH} = 0.65$ as representative of highly hygroscopic inorganic salts, including sulfate- and nitrate-containing salts. For the marine-influenced sites ENA (Eastern North Atlantic) and MHD (Mace Head, Ireland), we use $\kappa_{MH} = 1.5$. The baseline end members were selected to encompass the resolved structures of the measured κ-PDFs without truncating their upper modes, while remaining consistent with reported hygroscopicity ranges for common aerosol components (Petters and Kreidenweis, 2007; Carrico et al., 2008; Zieger et al., 2017; Lewis, 2019; Chen et al., 2022). They are effective retrieval bounds rather than species-specific κ values. The broader interval used at ENA and MHD makes the binary representation more approximate because common inorganic particles can be represented as LH–MH mixtures rather than pure MH material. We therefore use the ENA and MHD retrievals to examine temporal and size-dependent variations at these sites and do not compare their χ values quantitatively with values retrieved using $\kappa_{MH} = 0.65$."
- P7L184–P8L194: "We evaluated the sensitivity of the χ retrievals to analysis choices that differ among the datasets, including the prescribed upper end member, the RH used for GF-to-κ conversion, and the κ-PDF bin width (Figures S3–S6). Each alternative retrieval was paired with the corresponding baseline retrieval for the same day, site, and particle diameter, and the sensitivity was calculated as

$$\Delta\chi = \chi_{\text{alternative}} - \chi_{\text{baseline}}.$$

Changing the upper end member produced systematic shifts in χ, with the largest response at ENA. For MHD and BJIAP, changing the RH used in the GF-to-κ conversion from 90% to 89% or 91% changed median χ by approximately 1 percentage point at MHD and 2–3 percentage points at BJIAP. Using coarser κ-PDF bins produced smaller changes in χ across most sites and particle diameters. The resulting $\Delta\chi$ distributions represent sensitivity to these prescribed perturbations rather than complete retrieval uncertainty because comparable measurement and inversion uncertainty estimates are unavailable across the datasets."

2. Line 165: The interpretation of χ variations at sites such as SGP, CRG, HOU, and EPC requires more detailed explanation. In particular, for urban and coastal urban sites that are persistently influenced by anthropogenic emissions (including fresh primary organic aerosol and black carbon), it is not immediately intuitive why the aerosol population remains relatively internally mixed throughout the year. The authors should elaborate on the underlying mechanisms—e.g., rapid coagulation, condensation of secondary species, or regional aging and add appropriate references.

We thank the reviewer for this suggestion. We revised the interpretation of relatively high χ to distinguish similarity in particle hygroscopicity from uniform chemical composition. Persistent anthropogenic influence does not necessarily imply that fresh local emissions dominate the sampled aerosol.

We confirmed that the CRG measurements were collected at the rural site near Mount Airy rather than the urban Baltimore site and corrected its classification in Section 3.1 and Table 1. For HOU and EPC,

Section 4.1 now discusses condensation and coagulation as possible contributors to reduced hygroscopicity differences among particles, together with boundary-layer mixing and regional transport that may alter the contributions of fresh and processed aerosol.

For SGP, the added composition and trajectory analyses in Section 4.2 provide supporting observations for the revised interpretation in Section 4.3. This discussion relates wintertime nitrate enhancement to changes in hygroscopicity diversity and considers secondary-material condensation and the mixing of transported aerosol populations as possible contributors, without treating these mechanisms as established causes

References

Carrico, C. M., Petters, M. D., Kreidenweis, S. M., Collett Jr., J. L., Engling, G., and Malm, W. C.: Aerosol Hygroscopicity and Cloud Droplet Activation of Extracts of Filters from Biomass Burning Experiments, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD009274, 2008.

Chen, L., Zhang, F., Zhang, D., Wang, X., Song, W., Liu, J., Ren, J., Jiang, S., Li, X., and Li, Z.: Measurement Report: Hygroscopic Growth of Ambient Fine Particles Measured at Five Sites in China, Atmospheric Chemistry and Physics, 22, 6773–6786, https://doi.org/10.5194/acp-22-6773-2022, 2022.

Lewis, E. R.: The Dependence of Radius on Relative Humidity and Solute Mass at High Relative Humidities Up to and Including 100%, Journal of Geophysical Research: Atmospheres, 124, 2105–2126, https://doi.org/10.1029/2018JD030008, 2019.

Liu, Y., Wang, J., and Riemer, N.: From κ to χ: Evaluating Hygroscopicity-Based Mixing State Estimates with a Particle-Resolved Model, Atmospheric Chemistry and Physics, 26, 10 965–10 978, https://doi.org/10.5194/acp-26-10965-2026, 2026.

Petters, M. D. and Kreidenweis, S. M.: A Single Parameter Representation of Hygroscopic Growth and Cloud Condensation Nucleus Activity, Atmospheric Chemistry and Physics, 7, 1961–1971, https://doi.org/10.5194/acp-7-1961-2007, 2007.

Zieger, P., Väisänen, O., Corbin, J. C., Partridge, D. G., Bastelberger, S., Mousavi-Fard, M., Rosati, B., Gysel, M., Krieger, U. K., Leck, C., Nenes, A., Riipinen, I., Virtanen, A., and Salter, M. E.: Revising the Hygroscopicity of Inorganic Sea Salt Particles, Nature Communications, 8, 15 883, https://doi.org/10.1038/ncomms15883, 2017.