Reviewer 2:

This manuscript describes the ESMAC Diags version 1.0 package and provides useful examples of its application. I have only minor concerns and some hopefully useful suggestions below, but otherwise I believe the manuscript is ready for prompt publication. Kudos to the authors for this nice service to the community, and I hope the package gets good use. -MD

We would like to thank Michael Diamond for taking the time to review this paper and provide helpful comments to improve the paper. The comments are repeated below in black with our reply in blue.

General comments:

Addition of future campaigns:

Would it be possible to discuss more about **which other campaigns are being considered for inclusion in future versions of the diagnostics package**? The southeast Atlantic smoke-cloud campaigns (NASA ORACLES, DOE LASIC, plus CLARIFY and AEROCLO-SA internationally) in particular could be great testbeds for aerosol representation and have good ground- and air-based sampling. ATom could also be really interesting for its global reach.

Although the ongoing version 2 of ESMAC Diags is focusing on clouds and aerosol-cloud interactions for current field campaigns, this package can also be extended to other campaigns or other ESMs in the future. We added the following statement in the summary:

"In the future, this diagnostics package may also be extended to include other field campaigns that provide valuable data on aerosol properties and cloud-aerosol interactions, such as the ARM Layered Atlantic Smoke Interactions with Clouds (LASIC, Zuidema et al., 2018), NASA ObseRvations of Aerosols above CLouds and their intEractionS (ORACLES, Redemann et al., 2021), or NASA Atmospheric Tomography Mission (ATom, Brock et al., 2019) campaigns. As an open-source package, ESMAC Diags can also be applied by any user to other ESMs with small modifications on model preprocessing."

Treatment of observations as "truth":

At some locations in the text (e.g., "underestimation" in Line 303) the language sounds like the observations are being treated as base "truth." Other locations more thoroughly discuss limitations in the observed data as well. It might be helpful to address the issue of how observations are treated (not truth, but useful baseline given limitations are known) more in the introduction or methods sections.

We agree that observations have their own limitations and uncertainties although they are usually treated as “truth” when evaluating models. As suggested, we added the following discussion in Section 2.1:

“Although these measurements are considered as “truth” when evaluating ESMs, we note that they are subject to limitations and uncertainties due to theoretical/methodological formulations, sampling representativeness, instrumental accuracy and precision, imperfect calibration, random errors, etc. In addition, sampling volumes differ between observations and model output and are not reconcilable. It is difficult to quantify every aspect of observational uncertainty within the context of interpreting comparisons with model output, but we try to discuss some of them in this study to the best of our knowledge. Percentiles (either 25% - 75% or 5% - 95%) are used in some analyses of this study to approximate data variability that is likely to be much higher than measurement uncertainty.”

Specific comments:

Line 164: Specify that size is referring to **aerodynamic dry diameter** (or whatever it is you are using) for all uses thereafter. I'm assuming diameter but I don't remember seeing it in the text (but it is in some figure labels).

We added the following sentence to specify what “size” means:

“In ESMAC Diags v1, aerosol “size” refers to mobility and optical dry diameter of particles.”

Line 208: Why choose only latitude? Are there any longitudinal variation issues that should be addressed?

The purpose of this bullet of diagnostics is examining variations due to climate regime transitions along aircraft or ship tracks across the Northeastern Pacific or the Southern Ocean. We chose latitude in this diagnostics package because the latitudinal gradient dominates over the Southern Ocean, while over the Northeastern Pacific variations exist both longitudinally and latitudinally. We can add similar diagnostics along longitude if additional field campaigns are incorporated where longitudinal variation is relevant.

Figures 4-5: It might be helpful to also place markers binned for >10 nm and >100 nm for easy comparison to Table 3 and the discussion in the text.

The figures are updated with grid lines to easily check 10 nm and 100 nm bins, and this modification has been made to ESMAC Diags output.

Line 256 (Table 3): Are there any issues worth discussing between the PCASP and UHSAS data, e.g., different size cutoffs and bins?

It is a data availability issue that for aircraft measurements during HI-SCALE and ACE-ENA, only PCASP is available. UHSAS is available on surface measurements during HI-SCALE and ACE-ENA, and in other field campaigns. Size cuts of the respective size distribution measurements are given in Table 2. We revised the sentence as below to avoid confusion:

“PCASP is available only on aircraft for HI-SCALE and ACE-ENA. UHSAS is available only in surface measurements for HI-SCALE and ACE-ENA, and in other field campaigns.”

Lines 297-299: A citation from a relevant HI-SCALE paper would be useful here.

We revised this sentence and added a citation as below:

"All observed aerosol properties decrease with height since the major sources of aerosols (anthropogenic, biogenic, and biomass burning) (Liu et al., 2021) are from precursors emitted near the surface and chemical formation within the PBL."

Reference: Liu, J., Alexander, L., Fast, J. D., Lindenmaier, R., and Shilling, J. E.: Aerosol characteristics at the Southern Great Plains site during the HI-SCALE campaign, Atmos. Chem. Phys., 21, 5101-5116, https://doi.org/10.5194/acp-21-5101-2021, 2021.

Lines 400-401: I'm not sure this concern is warranted, as above-cloud CCN concentration has only limited relevance to cloud properties because the timescale for entraining above-cloud air into the cloudy boundary layer is on the order of days (Diamond et al., 2018; Mardi et al., 2019). The below-cloud CCN concentrations seems better-represented, and these should be the relevant metric for ACI considerations.

We agree with the reviewer and delete this statement.

Lines 440-441: Although this reads the "right" way based on the x-axis in Figures 13-14(a), it's backwards from the Lagrangian/cloud perspective. I'd recommend flipping it ("SSTs increase from CA to HI...").

We revised the corresponding statements to flip all description from CA to HI.

Lines 452-458: The commentary here is really sparse as compared to the other sections/campaigns. I'm not as familiar with this region, but I know of a few papers from Daniel and Isabel McCoy and colleagues that seem potentially relevant (listed below), and am sure there are many others that could be usefully discussed here.

Thank you for the suggestion. We have now expanded the discussions in Southern Ocean below:

"Similar latitudinal gradients of aerosol and CCN number concentrations along ship tracks from MARCUS and aircraft tracks from SOCRATES are shown in Figures 16 and 17, respectively. Over the SO region, NPF frequently occurs during austral summer when ample biogenic precursor gases (e.g., DMS) are released and rise into the free troposphere (McFarquhar et al., 2021; McCoy et al., 2021). Large values of ship-measured aerosol and CCN number concentration are observed near Antarctica corresponding to the coastal biological emissions of aerosol precursors, and also occur to the north of 45°S, indicating impacts from continental and anthropogenic sources. This is consistent with other studies (Sanchez et al., 2021; Humphries et al., 2021). EAMv1 underestimates aerosol and CCN number concentration near Antarctica. This bias, which may be related to too strong wet scavenging or insufficient NPF and growth, is commonly seen in many other ESMs (e.g., McCoy et al., 2020; McCoy et al., 2021). Aircraft flight paths during SOCRATES (Figure 17) do not extent as far south as the ship measurements (Figure 16). The observed aerosol properties have little latitudinal variation in general. EAMv1 underestimates aerosol number concentration for size > 10nm and CCN number concentration with SS=0.5%, but the predictions are closer to observed for aerosol size > 100 nm and CCN with SS=0.1% (Figure 17), consistent with the

mean aerosol size distribution in Figure 5. This indicates that the model performs better in simulating accumulation mode than Aitken mode particles. These model aerosol biases are highly relevant when considering their interaction with clouds and radiations, which will be included in version 2 of ESMAC Diags."

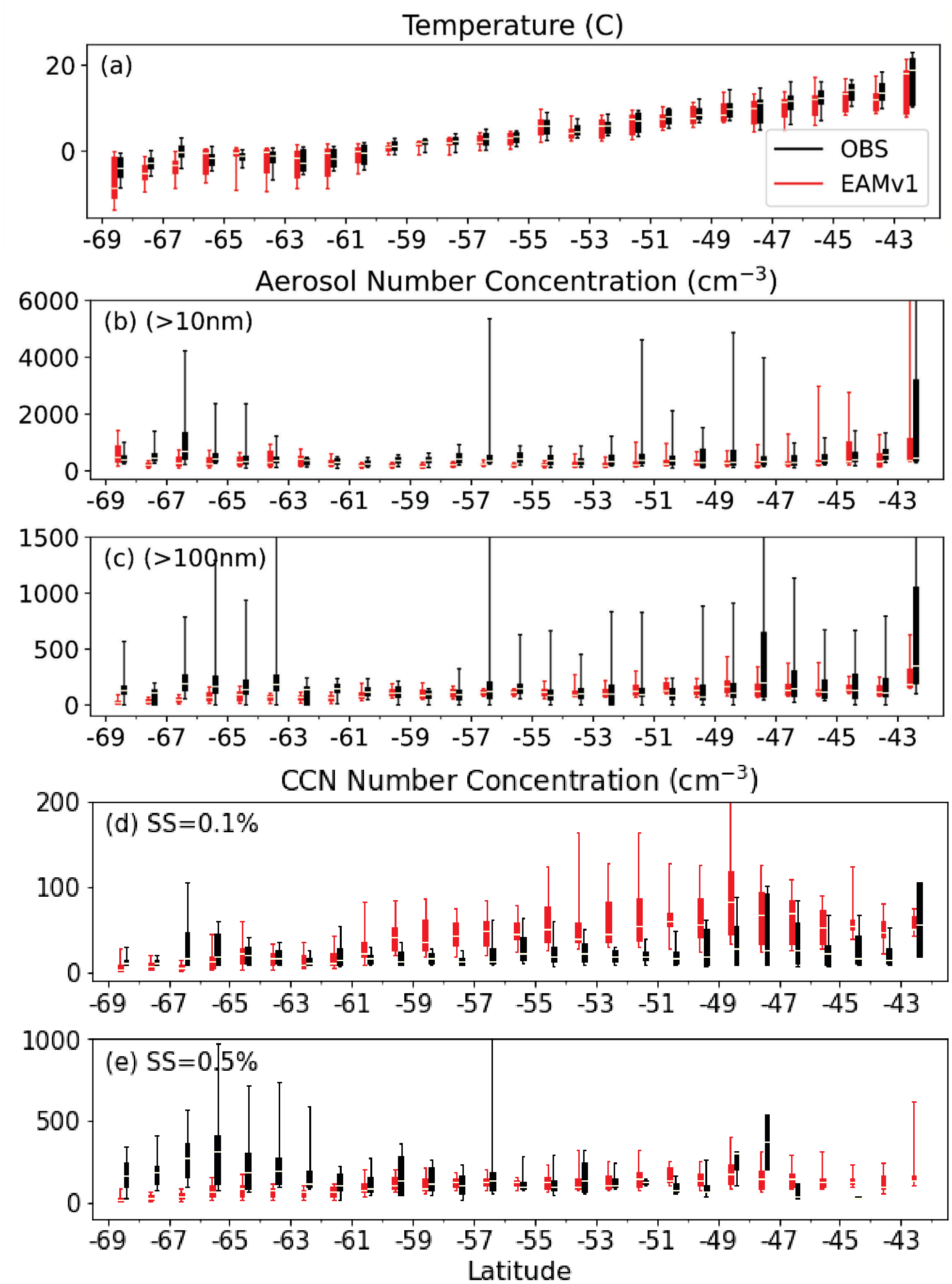

Figure 16: Percentiles of (a) air temperature, (b) aerosol number concentration for diameter >10 nm, (c) aerosol number concentration for diameter >100 nm, (d) CCN number concentration for supersaturation SS=0.1%, and (e) CCN number concentration for supersaturation SS=0.5% for all ship tracks in MARCUS binned by 1° latitude bins.

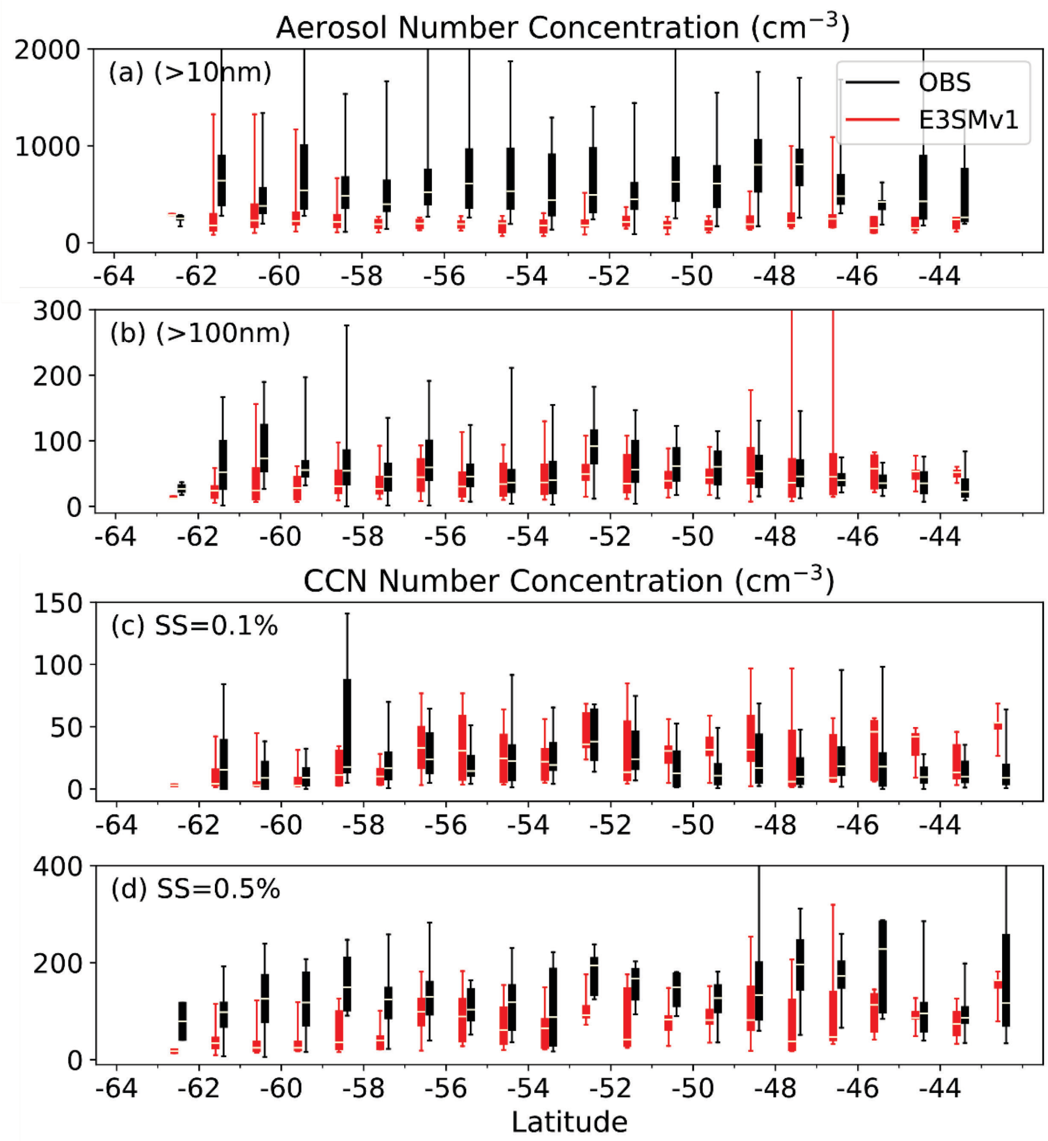

Figure 17: Percentiles of (a) aerosol number concentration for diameter >10 nm, (b) aerosol number concentration for diameter >100 nm, (c) CCN number concentration for supersaturation SS=0.1%, and (d) CCN number concentration for supersaturation SS=0.5% for all aircraft measurements between 0-3 km in SOCRATES binned by 1° latitude bins.

Line 453: Do we not need to worry about ice water path here as well, given the mixed-phase regime? It seems difficult to interpret LWP-only results here, unless this is being subset for confidently warm or supercooled clouds only?

We agree that ice is important over the Southern Ocean and the analysis of LWP-only may be difficult to interpret. Since this paper focuses on aerosols, we remove the plots and discussions on clouds (as another reviewer suggested) for the Southern Ocean, and only discuss aerosols instead. The diagnostics of cloud properties will be included in ESMAC Diags v2.

References:

Diamond, M. S., Dobracki, A., Freitag, S., Small Griswold, J. D., Heikkila, A., Howell, S. G., Kacarab, M. E., Podolske, J. R., Saide, P. E., and Wood, R.: Time-dependent entrainment of smoke presents an observational challenge for assessing aerosol–cloud interactions over the southeast Atlantic Ocean, Atmospheric Chemistry and Physics, 18, 14623-14636, 10.5194/acp-18-14623-2018, 2018.

Mardi, A. H., Dadashazar, H., MacDonald, A. B., Braun, R. A., Crosbie, E., Coggon, M. M., Aghdam, M. A., Woods, R. K., Jonsson, H. H., Flagan, R. C., Seinfeld, J. H., and Sorooshian, A.: Effects of Biomass Burning on Stratocumulus Droplet Characteristics, Drizzle Rate, and Composition, Journal of Geophysical Research: Atmospheres, 124, 12301– 12318, 10.1029/2018jd029134, 2019.

McCoy, D. T., Burrows, S. M., Wood, R., Grosvenor, D. P., Elliott, S. M., Ma, P.-L., Rasch, P. J., and Hartmann, D. L.: Natural aerosols explain seasonal and spatial patterns of Southern Ocean cloud albedo, Science Advances, 1, e1500157, 10.1126/sciadv.1500157, 2015.

McCoy, I. L., McCoy, D. T., Wood, R., Regayre, L., Watson-Parris, D., Grosvenor, D. P., Mulcahy, J. P., Hu, Y., Bender, F. A. M., Field, P. R., Carslaw, K. S., and Gordon, H.: The hemispheric contrast in cloud microphysical properties constrains aerosol forcing, Proceedings of the National Academy of Sciences of the United States of America, 117, 18998-19006, 10.1073/pnas.1922502117, 2020.

McCoy, I. L., Bretherton, C. S., Wood, R., Twohy, C. H., Gettelman, A., Bardeen, C. G., and Toohey, D. W.: Influences of Recent Particle Formation on Southern Ocean Aerosol Variability and Low Cloud Properties, Journal of Geophysical Research: Atmospheres, 126, 10.1029/2020jd033529, 2021.