

Review 1 comment on esd 2024-97

This paper describes a combined dataset of observations from several ARM campaigns. The data are from aircraft observations and include thermodynamic, dynamic, cloud, aerosol, and gas-phase measurements. The instrumentation and the dataset are described clearly, and several example figures are shown. A combined dataset like this one is useful to achieve continued, high-quality, and repeatable science.

Minor comments: Line 167 should reference Figure 3, not 4. The text on Figures 4 and 5 should be larger.

Response: We sincerely appreciate the reviewer's comments and suggestions. We have revised the manuscript and corrected the Figure number and also updated Figures 4 and 5.

Review 2 comment on essd 2024-97

The manuscript gives an overview of the detailed data collected during those seven field campaigns by the ARM program and how airborne measurements catch the atmosphere's detailed atmospheric processes. A data set integrated from these campaigns is an important resource for the study of various atmospheric phenomena but more so for aerosols, clouds, and trace gases. The authors have indeed succeeded in compiling, standardizing, and making these datasets available, which will support ongoing and future research in our community. Overall, this manuscript is well-organized and of clear scientific significance. I want to give him some advice:

We sincerely appreciate the reviewer's comments and suggestions. We have provided responses below and revisions to the manuscript to address each comment.

1. The manuscript may benefit from more discussions of the limitations and uncertainties associated with the data, particularly regarding the comparison between airborne and ground-based measurements. While the paper touches on the challenges of comparing these two types of measurements, a more thorough exploration of the potential biases introduced by differences in spatial and temporal resolution would help us understand more about the dataset.

Response: Thank you for this constructive advice. We have added Table S3 concerning uncertainties associated with the data. Regarding the comparison between airborne and ground-based measurements, we revised section 3 to the below paragraphs.

" The AAF's airborne measurements are considered accurate and reliable because they are obtained directly from the atmosphere using well-calibrated instruments. Numerous prior studies have systematically assessed ARM data quality, employing methods such as laboratory evaluation based on community-accepted standards, comparing similar properties across different instruments, and conducting intercomparisons across diverse platforms (from ground to airborne or airborne to airborne) (Bond et al., 1999; Lance et al., 2010; Kassianov et al., 2015; Kassianov et al., 2018; Mei et al., 2020; Zawadowicz et al., 2021; Kulkarni et al., 2023; Tang et al., 2023; Zhang et al., 2023b). Table S3 lists AAF measurements and uncertainties of atmospheric properties, including temperature, humidity, aerosol concentrations, cloud particle sizes, and radiation levels. (Mei et al., 2020) "

" This study further demonstrates the comparison of AAF data with ground-based remote sensing retrieval. The ARM G-1 aircraft was deployed above or near the ENA and SGP sites during field campaigns like ACE-ENA and HI-SCALE. Various instruments employing diverse observational techniques have measured atmospheric parameters, aerosol, and cloud properties from ground and airborne perspectives. These coordinated deployments enable a thorough assessment of robustness and statistical representativeness across collocated measurements.

Comparing airborne and ground-based measurements involves evaluating data from two platforms that differ in spatial and temporal resolutions, and measurement techniques. Thus, three potential biases exist in the measurements – spatial, temporal, and instrumental. The instrumental bias is typically due to the differences in sensors, calibration, and data processing techniques between the two platforms. Airborne measurements usually provide in situ spatiotemporal data over leveled flight legs at different altitudes and can capture data over various, even difficult-to-access, terrains. Meanwhile, ground-based remote sensing data usually provide continuous monitoring at a fixed location with limited spatial coverage or less vertical resolution. Our efforts focus on minimizing the temporal and spatial biases to ensure accurate and meaningful comparisons. We selected the comparison period by aligning the data acquisition times for both airborne and ground-based measurements as closely as possible. To ensure that both airborne and ground-based measurements are georeferenced accurately. For instance, ground-based remote sensing uses height or the altitude above the ground level (AGL) as the vertical geographic coordinate. In contrast, airborne data usually uses the mean sea level (MSL) altitude, which can be converted to the AGL. We then use interpolation techniques to match the spatial resolutions of airborne and ground-based data. "

2. The discussion of data quality could also include the differences and specific challenges between different field campaigns. For instance, variations in environmental conditions across different campaigns may have unique limitations in data collection and processing.

Response: Thank you very much for this suggestion. We have added section 3.3 "Data collection – challenges and future potential" to share lessons learned from each campaign in Table 3.

3. The manuscript adds more case studies or examples of how the dataset has been or could be used in specific research applications, which would provide practical context for its utility in past research.

Response: Thank you so much for this suggestion. We have provided more case studies and examples based on past research in section 3.

" The AAF airborne data are also often used as a benchmark or standard for other measurements, especially those from remote sensing technologies such as satellites, ground-based radars, and lidars. Junghenn Noyes et al. validated remote sensing retrievals with help from ground-based and airborne measurements. Their study enhanced the understanding of smoke particle behavior and its implications for remote sensing. (Junghenn Noyes et al., 2020) Mech et al. showcased how integrating airborne data into the PAMTRA (Passive and Active Microwave TRAnsfer) validation process enhances the model's skill in accurately simulating microwave measurements. The detailed comparison between simulated and observed data helps understand the model's performance in real-world conditions, leading to a more robust and reliable tool for atmospheric research. (Mech et al., 2020) Yang et al. developed a new method to estimate supersaturation fluctuations in stratocumulus clouds using ground-based remote-sensing retrievals. Then, they

used the airborne data to validate these estimations. (Yang et al., 2019) Wu et al. retrieved profiles of marine boundary layer (MBL) cloud and drizzle microphysical properties from ground-based observations, validated by aircraft measurements over the Azores. (Wu et al., 2020) Zhang et al. (2023) evaluated cloud droplet number concentrations using multiple ground-based methods validated through aircraft in situ measurements. (Zhang et al., 2023a)

Furthermore, research based on collected aircraft data led to advancements, characterization, and understanding of atmospheric processes. (Martin et al., 2017; Fast et al., 2011; Fast et al., 2019a; Fast et al., 2019b; Varble et al., 2021; Wang et al., 2022) For example, various studies have utilized aircraft measurements to characterize aerosol and cloud properties while advancing the understanding of aerosol chemistry and cloud microphysical properties and processes, including investigations over the North Atlantic, Amazon basin, and Southern Great Plains. (Shrivastava et al., 2019; Shilling et al., 2018; Zawadowicz et al., 2020; Wang et al., 2023b; Fast et al., 2024) The airborne data has also been used to examine the vertical variability of aerosol properties over the Southern Great Plains, contributing to a better understanding of the distribution and impact of aerosols at different atmospheric levels. (Wang et al., 2016; Fast et al., 2022) "

4. The manuscript may discuss the potential for future updates or expansions of the dataset, including the incorporation of additional variables or the development of new analytical tools to enhance its utility.

Response: We appreciate this advice and added section 3.3 "Data collection – challenges and future potential" to discuss the potential for future updates or expansions based on the science community needs.

" We summarize the challenges and data collection limitations encountered during the 7 field campaigns in Table 3. The lessons learned from these campaigns suggest that ensuring data quality and enhancing data collection variability require the following key strategies:

- Regular sensor calibration to maintain accuracy with scheduled, impromptu, and well-documented validation methods ensuring reliable measurements.
- Cross-validating and monitoring sensor performance using redundancy, dataset fusion, and statistical techniques to identify inconsistencies and malfunctions.
- Diversified sampling strategies to ensure comprehensive data representation across varying conditions.
- Leveraging model simulations and statistical studies based on previous measurements to refine data collection methods and anticipate potential issues, leading to more effective and targeted data acquisition.

These combined approaches ensure robust data collection, improve measurement accuracy, and provide reliable data products for the community of users.

Future work can expand data collection and merge and facilitate further investigations into atmospheric chemistry, aerosol properties, aerosol-cloud interactions and their representation in Earth System Models. To support future research needs, the ARM data center plans to work with the AAF instrument mentors and community experts to standardize this merged data product. Future deployments will use a new airborne platform (Challenger 850) and include more baseline airborne measurements ([https://arm.gov/capabilities/instruments?type\[0\]=armobs&category\[0\]=Airborne%20Observations](https://arm.gov/capabilities/instruments?type[0]=armobs&category[0]=Airborne%20Observations)). For example, we plan to add the liquid water content measurements from the Multi-Element Water Content System (<https://arm.gov/capabilities/instruments/wcm-air>) and solar radiation measurements from multifilter radiometers (<https://arm.gov/capabilities/instruments/mfr-air>) into the future merged dataset.

In addition to the baseline measurements, we plan to offer data integration options through the ARM Data Integrator tool (details in section 4), allowing campaign principal investigators or community users to flexibly incorporate additional data into the merged data product. This approach provides users with the ability to target specific science themes. For example, one proposed data product is to include additional aerosol optical properties with this AAF merged dataset. The atmospheric community can expand research on the evolution of aerosol particles from wildfires, particularly on how different combustion phases (flaming vs. smoldering) result in varying chemical compositions and quantities of emitted aerosols with additional chemical composition and gaseous phase concentration data. This custom-built merged dataset is crucial for improving air quality models and understanding the climate impacts of biomass burning. Continuing to investigate the influence of urban pollution on natural aerosol formation, similar to the studies conducted in the Amazon during GoAmazon2014/5, will provide new insights into aerosol composition. Potential data for such a study could be non-airborne remote-sensing data. Combining airborne data with ground-based remote sensing data allows the exploration of interactions between different aerosol types and evolving cloud and precipitation patterns. Quantifying these interactions can improve models and understanding, aiding the development of strategies for mitigating anthropogenic impacts on natural environments.

By strategically combining long-term ground-based remote sensing measurements with high-resolution airborne data, researchers can achieve more robust analyses of atmospheric processes, leading to more accurate scientific findings and better-constrained models. For instance, ground-based sensors can continuously monitor a specific location, while targeted airborne missions can capture critical in situ measurements during specific events that are not retrievable by remote sensing to better study cloud evolution, pollutant transport, or extreme weather."