Earth System Model Aerosol-Cloud Diagnostics Package (ESMAC Diags) Version 1: Assessing E3SM Aerosol Predictions Using Aircraft, Ship, and Surface Measurements

Shuaiqi Tang[1], Jerome D. Fast[1], Kai Zhang[1], Joseph C. Hardin[2], Adam C. Varble[1], John E. Shilling[1], Fan Mei[1], Maria A. Zawadowicz[3] and Po-Lun Ma[1]

[1]Pacific Northwest National Laboratory, Richland, WA, USA
[2]ClimateAI Inc, San Francisco, CA, USA
[3]Brookhaven National Laboratory, Upton, NY, USA

Correspondence to: Shuaiqi Tang (shuaiqi.tang@pnnl.gov)

Abstract. An Earth System Model (ESM) aerosol-cloud diagnostics package is developed to facilitate the routine evaluation of aerosols, clouds and aerosol-cloud interactions simulated by the Department of Energy's (DOE) Energy Exascale Earth System Model (E3SM). The first version focuses on comparing simulated aerosol properties with aircraft, ship, and surface measurements, most of them are measured in-situ. The diagnostics currently covers six field campaigns in four geographical regions: Eastern North Atlantic (ENA), Central U.S. (CUS), Northeastern Pacific (NEP) and Southern Ocean (SO). These regions produce frequent liquid or mixed-phase clouds with extensive measurements available from the Atmospheric Radiation Measurement (ARM) program and other agencies. Various types of diagnostics and metrics are performed for aerosol number, size distribution, chemical composition, CCN concentration and various meteorological quantities to assess how well E3SM represents observed aerosol properties across spatial scales. Overall, E3SM qualitatively reproduces the observed aerosol number concentration, size distribution and chemical composition reasonably well, but overestimates Aitken mode and underestimates accumulation mode aerosols over the CUS and ENA regions, suggesting that processes related to particle growth or coagulation might be too weak in the model. The current version of E3SM struggles to reproduce new particle formation events frequently observed over both the CUS and ENA regions, indicating missing processes in current parameterizations. The diagnostics package is coded and organized in a way that can be extended to other field campaign datasets and adapted to higher-resolution model simulations.

1. Introduction

Aerosol number, mass, size, composition, and mixing state affect how aerosol populations scatter and absorb solar radiation and influence cloud albedo, amount, lifetime, and precipitation (Twomey, 1977; Albrecht, 1989) by acting as cloud condensation nuclei (CCN) (e.g., Petters and Kreidenweis, 2007). However, there are still knowledge and measurement gaps on the physical and chemical mechanisms regulating the sources, sinks, gas-to-particle partitioning (e.g., secondary formation processes), and spatiotemporal distribution of aerosol populations. Consequently, the representation of the aerosol lifecycle and the interaction of aerosol populations with clouds and radiation in Earth system models (ESMs) still suffer from large uncertainties (Seinfeld et al., 2016; Carslaw et al., 2018), which impacts the ability of ESMs to predict the evolution of the climate system (IPCC, 2013).

To facilitate model evaluation and document the performance of parameterizations in ESMs, many modeling centers have developed standardized diagnostics packages. Some examples focus on meteorological metrics include the U.S. National Center of Atmospheric Research (NCAR) Atmospheric Model Working Group (AMWG) diagnostics package (AMWG, 2021), the U.S. Department of Energy (DOE) Energy Exascale Earth System Model (E3SM, Golaz et al., 2019) diagnostics (E3SM, 2021), the European Union (EU) Earth System Model Evaluation Tool (ESMValTool, Eyring et al., 2016), and the Program for Climate Model Diagnosis and Intercomparison (PCMDI) Metric Package (PMP, Gleckler et al., 2016). Some recent efforts focus on process-oriented diagnostics (POD) that are designed to provide insights into parameterization developments to address long-standing model biases. Maloney et al. (2019) summarizes the activities by the U.S. National Oceanic and Atmospheric Administration (NOAA) Modeling, Analysis, Prediction, and Projections program (MAPP) Model Diagnostics Task Force (MDTF) to apply community-developed PODs to climate and weather prediction models. Zhang et al. (2020) developed a diagnostics package that utilizes statistics derived from long-term ground-based measurements from the DOE Atmospheric Radiation Measurement (ARM) User Facility for climate model evaluation. Aerosol properties, however, are not included in these diagnostics packages.

The international collaborative AeroCom project (Myhre et al., 2013; Schulz et al., 2006) focuses on evaluation of aerosol predictions using available measurements and includes intercomparisons among global models to assess uncertainties in seasonal and regional variations in aerosol properties and their potential impact on climate. Their diagnostics heavily rely on satellite remote sensing products (e.g., aerosol optical depth) which have global coverage but poor spatial and temporal resolution that hinders a process-level understanding of the sources of model uncertainty. More recently, the Global Aerosol Synthesis and Science Project (GASSP, Reddington et al., 2017; Watson-Parris et al., 2019) has developed a global database of aerosol observations from fixed surface sites as well as ship and aircraft

63   platforms from 86 field campaigns between 1990 and 2015 that can be used for model evaluation. Recent
64   field campaigns after year 2015 are not included in this effort.

65   Many aerosol properties are difficult to measure directly. Remote sensing instruments (e.g., ground
66   and satellite radiometers) that only measure radiative properties of column-integrated aerosols, such as
67   optical depth, are frequently used to evaluate model predictions. Instruments such as ground lidars (e.g.,
68   Campbell et al., 2002) or lidars onboard aircraft (e.g., Müller et al., 2014) and satellite (e.g., CALIPSO,
69   Winker et al., 2009) platforms can provide vertical profiles of aerosol extinction, backscatter, and/or
70   depolarization, but they do not directly measure aerosol number, size and composition. Therefore, the
71   quantities measured by remote sensing instruments cannot be used alone to assess model predictions of
72   aerosol-radiation-cloud-precipitation interactions. Surface monitoring sites provide long-term in situ
73   aerosol property measurements but are limited to land locations with far fewer operational sites compared
74   to those dedicated to routine meteorological sampling. Ship and aircraft platforms are commonly
75   deployed during field campaigns to obtain in situ and remote sensing aerosol property measurements in
76   remote or poorly sampled locations such as over the ocean and within the free troposphere, which are
77   highly valuable when studying spatial variations of aerosols. Aircraft platforms also provide a means to
78   obtain coincident measurements of aerosol and cloud properties needed to understand their interactions.
79   Although in-situ ship and airborne aerosol measurements are usually limited to specific locations for short
80   time periods, the increasing number of completed field campaigns conducted over a range of atmospheric
81   conditions provides an opportunity to use them for model evaluation.

82   As noted by Reddington et al. (2017), the considerable investment in collecting field campaign
83   measurements of aerosol properties is underexploited by the climate modeling community. This can be
84   largely attributed to datasets located in disparate repositories and the lack of a standardized file format
85   that requires excessive time and effort spent on manipulating the datasets to facilitate comparisons
86   between observed and simulated values, especially for those unfamiliar with measurement techniques,
87   assumptions, and uncertainties. With many field campaigns conducted since 2015 being available but
88   rarely used for model evaluation, this study describes the first version of the ESM Aerosol-Cloud
89   Diagnostics (ESMAC Diags) package to facilitate the evaluation of ESM-predicted aerosols, utilizing
90   recent measurements from aircraft, ship and surface platforms collected by the U.S. DOE ARM and
91   National Science Foundation (NSF) NCAR user facilities, most of which are in-situ measurements. The
92   overall structure of ESMAC Diags is designed similar to the Aerosol Modeling Testbed for the Weather
93   Research and Forecasting (WRF) model described in Fast et al. (2011), except that ESMAC Diags uses
94   Python to interface the measurements with ESM output and does not preprocess the observational dataset
95   into a common format. The diagnostics package is firstly designed with and applied to E3SM Atmosphere

96 Model version 1 (EAMv1, Rasch et al., 2019). EAMv1 uses an improved modal aerosol treatment

97 implemented based on the 4-mode version of the modal aerosol module (MAM4, Liu et al., 2016), such

98 as improved treatment of H_2SO_4 vapor for new particle formation (NPF), improved secondary organic

99 aerosol (SOA) treatment, new marine organic aerosol (MOA) species, improvements to aerosol

100 convective transport, wet removal, resuspension from evaporation and aerosol-affected cloud

101 microphysical processes (Wang et al., 2020).

102

2. Introduction of ESMAC Diags

104 The workflow of ESMAC Diags v1 is illustrated in Figure 1. Most field campaign datasets are directly

105 read by the diagnostics package. In some field campaigns, more than one instrument is used to measure

106 aerosol size distribution over different size ranges. We therefore merge these datasets to create a more

107 complete description of the size distribution. These data are introduced in Section 2.1. Model outputs are

108 extracted at the ground sites and along the flight tracks or ship tracks. The simulation and preprocessing

109 details are provided in Section 2.2. ESMAC Diags reads in these field campaign and model data with

110 quality controls and generates a set of diagnostics and metrics listed in Section 2.3. The diagnostics

111 package is designed to be flexible so that additional measurements and functionality can be included in

112 the future. Figure 2 depicts the directory structure to illustrate the organization of the datasets and code.

113 Most of the datasets used in ESMAC Diags are in a standardized netCDF format (NETCDF, 2021);

114 however, some ARM aircraft measurements use different ASCII formats. Currently, the diagnostic

115 package reads observational data directly from their original format. In the long term, we may standardize

116 the observational data format in a similar manner as was done in GASSP project (Reddington et al.,

117 2017).

118

119

120

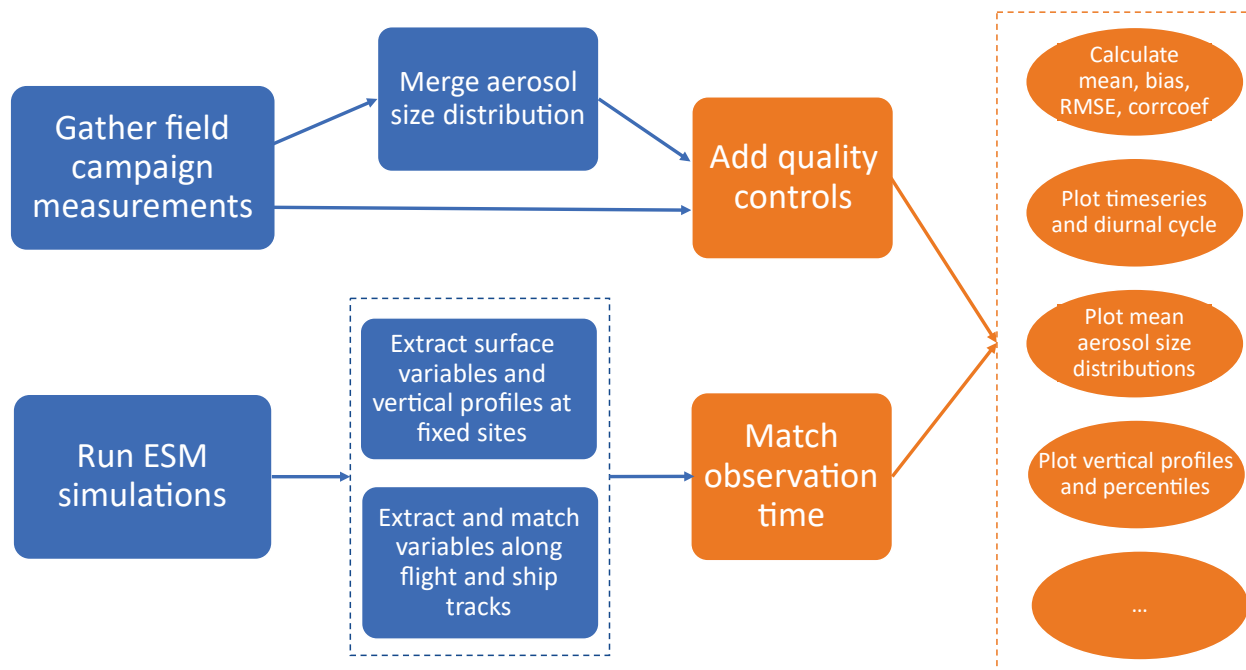



Figure 1: Workflow of ESMAC Diags. Data preprocessing and input are indicated by blue;

diagnostics and plotting are indicated by orange.

121

122

123

124

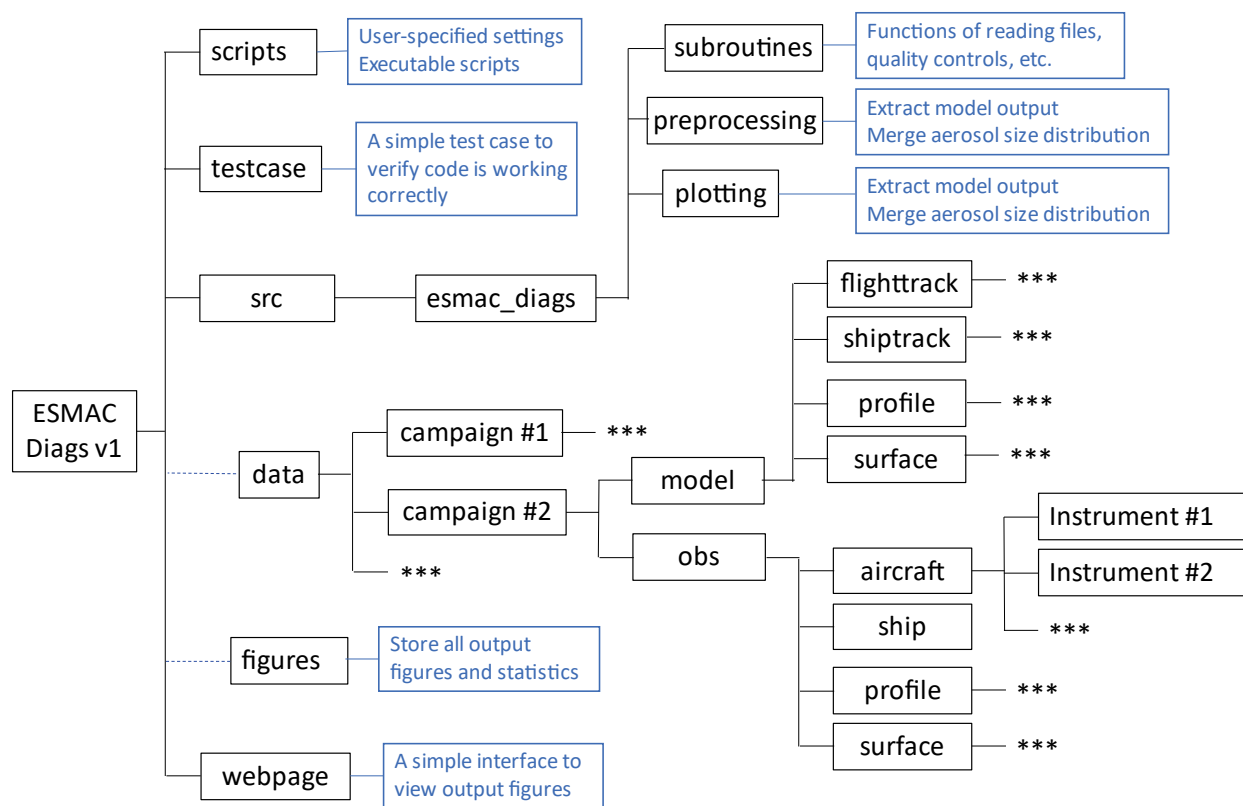



Figure 2: Structure of ESMAC Diags. The "scripts" directory contains executable scripts and user-specified settings. The "src" directory contains all source code including code used to preprocess model output, read files, merge measurements from different instruments, compute observed versus simulated statistical relationships, and plot results. All observational and model data in the "data" directory are organized by field campaign. The diagnostic plots and statistics are put in the "figures" directory, also organized by field campaign. The "testcase" directory includes a small amount of input and verify data to test if the package is installed properly. The "webpage" directory provides an interface to view diagnostics figures. Boxes in blue describe the functions of the directory. Asterisks represent boxes that follow the same format as those shown in parallel.

2.1 Field observations and merged aerosol size distribution

We initially focus on four geographical regions where liquid clouds occur frequently and extensive measurements are available from ARM and other agencies: Eastern North Atlantic (ENA), Northeastern Pacific (NEP), Central U.S. (CUS, where the ARM Southern Great Plains, SGP, site is located), and Southern Ocean (SO). Aerosol properties also vary among these regions. Six field campaigns from these four testbeds are selected in the version 1 of ESMAC Diags (Table 1). HI-SCALE and ACE-ENA are

140   based on long-term ARM ground sites with aircraft field campaigns sampling below, within, and above

141   convective and marine boundary layer clouds, respectively, within a few hundred kilometers around the

142   sites. CSET and MAGIC are field campaigns with aircraft and ship platforms, respectively, sampling

143   transects between California and Hawaii characterized by a transition between stratocumulus and trade

144   cumulus dominated regions. SOCRATES and MARCUS are field campaigns with aircraft and ship

145   platforms, respectively, based out of Hobart, Australia. Aircraft transects during SOCRATES extended

146   south to around 60°S, while ship transects during MARCUS extended southwest from Hobart to

147   Antarctica. The aircraft (black) and ship (red) tracks for these field campaigns are shown in Figure 3.

148   **Table 1. Descriptions of the field campaigns used in this study. Numbers after aircraft or ship**

149   **represent number of flights or ship trips in each field campaign or IOP.**

Campaign*	Period	Platform	Typical Conditions	Reference
HI-SCALE	IOP1: 24 Apr – 21 May 2016 IOP2: 28 Aug – 24 Sep 2016	Ground, aircraft (IOP1: 17, IOP2: 21)	Continental cumulus with high aerosol loading	(Fast et al., 2019)
ACE-ENA	IOP1: 21 Jun – 20 Jul 2017 IOP2: 15 Jan – 18 Feb 2018	Ground, aircraft (IOP1: 20, IOP2: 19)	Marine stratocumulus with low aerosol loading	(Wang et al., 2021)
MAGIC	Oct 2012 – Sep 2013	Ship (18)	Marine stratocumulus to cumulus transition with low aerosol loading	(Lewis and Teixeira, 2015; Zhou et al., 2015)
CSET	1 Jul – 15 Aug 2015	Aircraft (16)	Same as above	(Albrecht et al., 2019)
MARCUS	Oct 2017 – Apr 2018	Ship (4)	Marine liquid and mixed phase clouds with low aerosol loading	(McFarquhar et al., 2021)
SOCRATES	15 Jan – 24 Feb, 2018	Aircraft (14)	Same as above	(McFarquhar et al., 2021)

150   * Full names of the listed field campaigns:
151   HI-SCALE: Holistic Interactions of Shallow Clouds, Aerosols and Land Ecosystems
152   ACE-ENA: Aerosol and Cloud Experiments in the Eastern North Atlantic
153   MAGIC: Marine ARM GCSS Pacific Cross-section Intercomparison (GPCI) Investigation of Clouds
154   CSET: Cloud System Evolution in the Trades
155   MARCUS: Measurements of Aerosols, Radiation and Clouds over the Southern Ocean
156   SOCRATES: Southern Ocean Cloud Radiation and Aerosol Transport Experimental Study
157

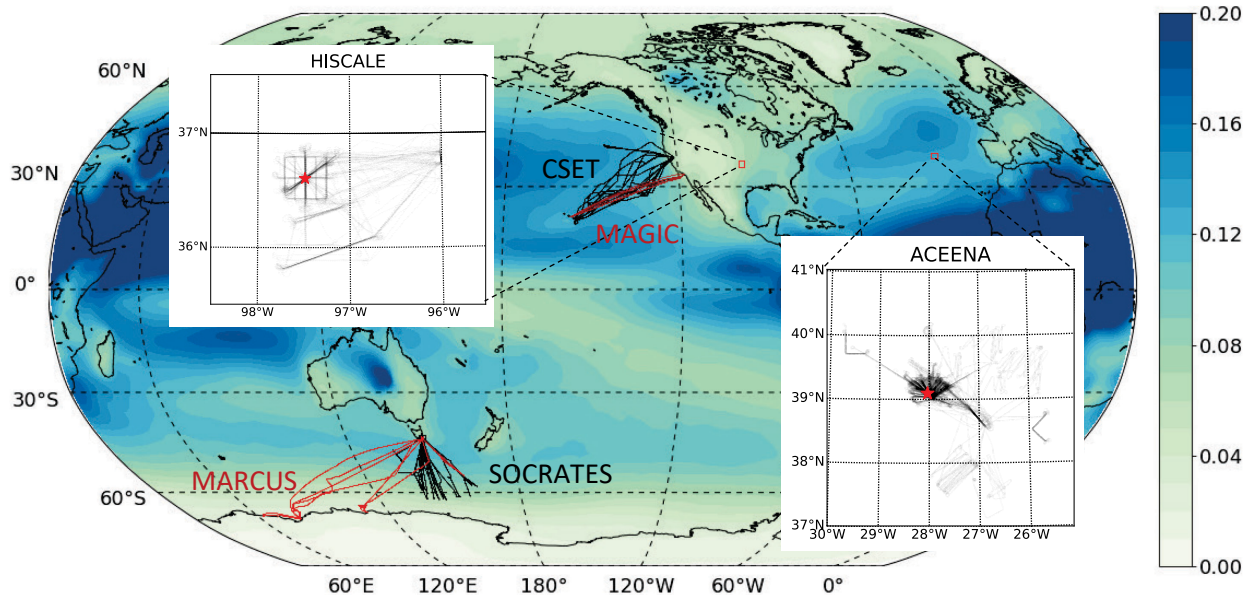



Figure 3. Aircraft (black) and ship (red) tracks for the six field campaigns. Overlaid is aerosol optical depth at 550nm averaged from 2014 to 2018 simulated in EAMv1.

The instruments and measurements used in ESMAC Diags version 1 are listed in Table 2. All in-situ measurements are converted to under ambient temperature and pressure. Note that some instruments are only available for certain field campaigns or failed operationally during certain periods, so that model evaluation is limited by the availability of data collected in each field campaign. ARM data usually include quality flags indicating bad or indeterminate data. These flagged data are filtered out, except surface condensation particle counter (CPC) measurements for HI-SCALE, that data flagged as greater than maximum value (8000 cm^{-3}) are retained since aerosol loading can be higher than that during NPF events. This exception ensures a reasonable diurnal cycle shown in Section 3.3. For some data that do not have a quality flag, a simple minimum and maximum threshold is applied (e.g., 500 cm^{-3} maximum threshold is used for each UHSAS bin from the NCAR research flight measurements).

Table 2. List of instruments and measurements used in ESMAC Diags v1.

Instrument	Platform	Measurements	Available campaigns	DOIs or References
Surface meteorological station (MET)	Ground, ship	Temperature, relative humidity, wind, pressure	HI-SCALE, ACE-ENA, MAGIC, MARCUS	HI-SCALE, ACE-ENA: (Kyrouac and Shi, 2018) MAGIC: (ARM, 2014) MARCUS: 10.5439/1593144
Scanning mobility particle sizer (SMPS)	Ground	Aerosol size distribution (20-700 nm)	HI-SCALE	(Howie and Kuang, 2016)

Nano scanning mobility particle sizer (nanoSMPS)	Ground	Aerosol size distribution (2-150 nm)	HI-SCALE	(Koontz and Kuang, 2016)
Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Ground, aircraft, ship	Aerosol size distribution (60 – 1000 nm), number concentration	HI-SCALE, ACE-ENA, MAGIC, MARCUS, CSET, SOCRATES	HI-SCALE, MAGIC, MARCUS: (Koontz and Uin, 2018) ACE-ENA: (Uin et al., 2018) CSET: 10.5065/D65Q4T96 SOCRATES: 10.5065/D6M32TM9
Condensation particle counter (CPC)	Ground, aircraft, ship	Aerosol number concentration (> 10 nm)	HI-SCALE, ACE-ENA, MAGIC, MARCUS	HI-SCALE (ground): (Kuang et al., 2016) ACE-ENA (ground), MAGIC: (Kuang et al., 2018a) MARCUS: (Kuang et al., 2018b) HI-SCALE (aircraft): (ARM, 2016b) ACE-ENA (aircraft): (Mei, 2018)
Condensation particle counter – ultrafine (CPCU)	Ground, aircraft	Aerosol number concentration (> 3 nm)	HI-SCALE, ACE-ENA	HI-SCALE (ground): 10.5439/1046186 HI-SCALE (aircraft): (ARM, 2016b) ACE-ENA (aircraft): 10.5439/1440985
Condensation nuclei counter (CNC)	Aircraft	Aerosol number concentration (11-3000 nm)	CSET, SOCRATES	CSET: 10.5065/D65Q4T96 SOCRATES: 10.5065/D6M32TM9
Cloud condensation nuclei (CCN) counter	Ground, aircraft, ship	CCN number concentration (0.1% to 0.5% supersaturation* depending on the platform)	HI-SCALE, ACE-ENA, MAGIC, MARCUS, SOCRATES	HI-SCALE (ground), ACE-ENA (ground), MARCUS: 10.5439/1342133 MAGIC: 10.5439/1227964 SOCRATES: 10.5065/D6Z036XB HI-SCALE (aircraft): (ARM, 2016a)
Aerosol chemical speciation monitor (ACSM)	Ground	Aerosol composition	HI-SCALE, ACE-ENA	10.5439/1762267
Microwave radiometer (MWR)	Ground, ship	Liquid water path, precipitable water vapor	MAGIC, MARCUS	10.5439/1027369
Counterflow virtual impactor (CVI)	Aircraft	Separates large droplets or ice crystals	HI-SCALE, ACE-ENA, SOCRATES	HI-SCALE: (ARM, 2016a) ACE-ENA: 10.5439/1406248 SOCRATES: 10.5065/D6M32TM9
Fast integrated mobility spectrometer (FIMS)	Aircraft	Aerosol size distribution (10 – 425 nm)	HI-SCALE, ACE-ENA	HI-SCALE: (ARM, 2017) ACE-ENA: (ARM, 2020)
Passive cavity aerosol spectrometer (PCASP)	Aircraft	Aerosol size distribution (120 – 3000 nm)	HI-SCALE, ACE-ENA, CSET	HI-SCALE: (ARM, 2016a) ACE-ENA: (ARM, 2018) CSET: 10.5065/D65Q4T96

Optical particle counter (OPC)	Aircraft	Aerosol size distribution (390 – 15960 nm)	ACE-ENA	(ARM, 2018)
Interagency working group for airborne data and telemetry systems (IWG)	Aircraft	Navigation information and atmospheric state parameters	HI-SCALE, ACE-ENA	HI-SCALE: (ARM, 2017) ACE-ENA: (ARM, 2018)
High-resolution time-of-flight aerosol mass spectrometer (AMS)	Aircraft	Aerosol composition	HI-SCALE, ACE-ENA	HI-SCALE: (ARM, 2017) ACE-ENA: 10.5439/1468474
Water content measuring system (WCM)	Aircraft	Cloud liquid and total water content	HI-SCALE, ACE-ENA	HI-SCALE: (ARM, 2016a) ACE-ENA: 10.5439/1465759
Doppler lidar (DL)	Ground	Boundary layer height	HI-SCALE	10.5439/1726254
Reprocessed CN and CCN data to remove ship exhaust influence	Ship	CN, CCN number concentration	MARCUS	10.25919/ezp0-em87

173 * For measured supersaturations (SS) that vary over time, a $\pm$ 0.05% window is applied (e.g., 0.5% SS

174 includes samples with SS between 0.45% and 0.55%).

175

176 For some field campaigns (HI-SCALE and ACE-ENA), there are several instruments (e.g., FIMS,

177 PCASP, OPC for aircraft; SMPS and nanoSMPS for ground) measuring aerosol size distribution over

178 different size ranges. These datasets are merged to create a more complete size distribution. In ESMAC

179 Diags v1, aerosol "size" refers to mobility and optical dry diameter of particles. The aerosol

180 concentrations in the "overlapping" bins measured by multiple instruments are weighted by the

181 uncertainty of each instrument based on the knowledge of the ARM instrument mentors. An example of

182 the merged aerosol size distribution and individual measurements for one flight in ACE-ENA is shown in

183 Figure 4. Ranging from 10^1 to 10^4 nm, the merged aerosol size distribution data account for ultrafine,

184 Aitken, and accumulation modes.

185 Although these measurements are considered as "truth" when evaluating ESMs, we note that they are

186 subject to limitations and uncertainties due to theoretical/methodological formulations, sampling

187 representativeness, instrumental accuracy and precision, imperfect calibration, random errors, etc. In

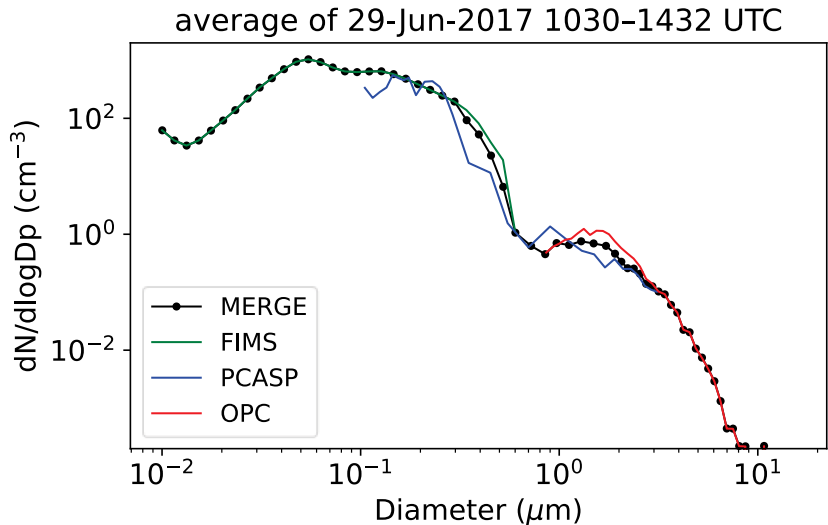
188 addition, sampling volumes differ between observations and model output and are not reconcilable. It is

189 difficult to quantify every aspect of observational uncertainty within the context of interpreting

190 comparisons with model output, but we try to discuss some of them in this study to the best of our

191 knowledge. Percentiles (either 25% - 75% or 5% - 95%) are used in some analyses of this study to

192 approximate data variability that is likely to be much higher than measurement uncertainty.

193

194 **Figure 4. An example of a mean aerosol number distribution merged from FIMS, PCASP and OPC**

195 **instruments for ACE-ENA aircraft measurements on 29 June 2017.**

196

197 **2.2 Preprocessing of model output**

198 In this study, we run EAMv1 from 2012 to 2018, covering all six field campaign periods introduced

199 previously, with enough time for model spin-up. The model is configured to follow the Atmospheric

200 Model Intercomparison Project (AMIP) protocol (Gates et al., 1999) with real-world forcings (e.g.,

201 greenhouse gases, sea surface temperature, aerosol emissions, etc.). For each simulation year, we use the

202 year 2014 emission data from CMIP6, since the emission data does not cover years after 2014. The

203 simulated horizontal winds are nudged towards the Modern-Era Retrospective analysis for Research and

204 Applications, Version 2 (MERRA-2, Gelaro et al., 2017) with a relaxation time scale of 6 hours. Previous

205 studies (Sun et al., 2019; Zhang et al., 2014) showed that with such nudging configuration the large-scale

206 circulation is well constrained in the nudged simulation, especially for the mid- and high- latitude regions.

207 The simulation uses a horizontal grid spacing of ~1° (NE30, the number of elements along a cube face of

208 the E3SM High-Order Methods Modeling Environment, HOMME, dynamics core) with a 30-minute

209 timestep. We saved hourly output to compare with field campaign measurements. The diagnostics

210 package post-processes 3-D model variables associated with aerosol concentration, size, composition,

211 optical properties, precursor concentration, CCN concentration, and atmospheric state variables. The size

212 of output data is reduced by saving 3-D variables only over the field campaign regions. The model

213 configuration and execution scripts are uploaded as electronic supplement to this paper. Users can apply it
214 in their own E3SM simulations (or output similar variables if running other models) to use this package.

215 We extracted model output along the aircraft (ship) tracks using an "aircraft simulator" (Fast et al.,
216 2011) strategy to facilitate comparisons of observations and model predictions. At each aircraft (ship)
217 measurement time, we find the nearest model grid cell, output time slice, and vertical level of the aircraft
218 altitude (or the lowest level for ship) to obtain the appropriate model values. Since there are both spatial
219 and temporal mismatch existing between model output and field measurements, the evaluation focuses on
220 overall statistics. We also calculate the aerosol size distribution from 1 nm to 3000 nm at 1 nm increments
221 from the individual size distribution modes in MAM4 to facilitate comparisons with observed aerosol
222 number distribution that has different size ranges for different instruments. All these variables are saved
223 in separate directories according to the specific aircraft (ship) tracks, as indicated in Figure 2.

224 **2.3 List of diagnostics and metrics**

225 Currently, ESMAC Diags produces the following diagnostics and metrics:

226 - Mean value, bias, root mean square error (RMSE) and correlation of aerosol number
227   concentration.
228 - Timeseries of aerosol variables (aerosol number concentration, aerosol number size distribution,
229   chemical composition, CCN number concentration) for each field campaign or intensive
230   observational period (IOP) at the surface or along each flight (ship) track.
231 - Diurnal cycle of aerosol variables at the surface.
232 - Mean aerosol number size distribution for each field campaign or IOP.
233 - Percentiles of aerosol variables by height for each field campaign or IOP.
234 - Percentiles of aerosol variables by latitude for each field campaign or IOP.
235 - Pie/bar charts of observed and predicted aerosol composition averaged over each field campaign
236   or IOP.
237 - Vertical profile of cloud fraction and LWC composite of aircraft measurements for each field
238   campaign or IOP.
239 - Timeseries of atmospheric state variables.
240 - Aircraft and ship track maps.

241 In the next section we will demonstrate these diagnostics and metrics by providing several examples.

242

243 **3. Examples**

244 Aerosol number concentration, size distribution, and chemical composition (that controls hygroscopicity)

245 are key quantities that impact aerosol-cloud interactions, such as the activation of cloud droplets. Errors in

246 model predictions of these aerosol properties contribute to uncertainties in aerosol direct and indirect

247 radiative forcing. These aerosol properties vary dramatically depending on location, altitude, season, and

248 meteorological conditions due to variability in emissions, formation mechanisms, and removal processes

249 in the atmosphere. This section shows some examples to illustrate the usage of this diagnostics package

250 on evaluating global models.

251 **3.1 Aerosol size distributions and number concentrations**

252 Aerosol properties are highly dependent on location and season. Figure 5 shows the mean aerosol size

253 distribution for each of the four testbed regions. For HI-SCALE and ACE-ENA, the two IOPs operated in

254 different seasons are shown separately. Table 3 shows the mean aerosol number concentration from these

255 field campaigns, for two particle size ranges: >10 nm and >100 nm. The inter-quartile range (25% and

256 75% percentiles) are also shown to illustrate the variability in space and time. Among the four testbed

257 regions, the CUS region has the largest aerosol number concentrations since the other field campaigns are

258 primarily over open ocean. Overall, EAMv1 overestimates Aitken mode (10 – 70 nm) aerosols and

259 underestimates accumulation mode (70 – 400 nm) aerosols for the CUS and ENA regions, suggesting that

260 processes related to particle growth or coagulation might be too weak in the model. Over the NEP region,

261 EAMv1 overestimates aerosol number for particle sizes >100 nm and >10 nm (Table 3), both at the

262 surface and aloft. Over the SO region, which is considered a pristine region with low aerosol

263 concentration, observations show a significant number of particles <200 nm in both aircraft and ship

264 measurements (Figure 5). The mean aerosol number concentration over SO region is comparable or even

265 greater than the other ocean testbeds (Table 3). In contrast, EAMv1 simulates a clean environment with

266 the lowest aerosol number concentrations among the four regions. These types of comparisons

267 demonstrate the need for additional analyses to understand why SO has similar aerosol number with other

268 ocean regions and why EAMv1 cannot simulate this feature. The observed 75% percentiles are sometimes

269 smaller than the mean value (Table 3), indicating skewed aerosol size distribution with long tail in large

270 aerosol size. EAMv1 usually produces smaller inter-quartile range than the observations, likely because

271 the current model resolution is too coarse to capture the observed spatial variability in aerosol properties.

272

273

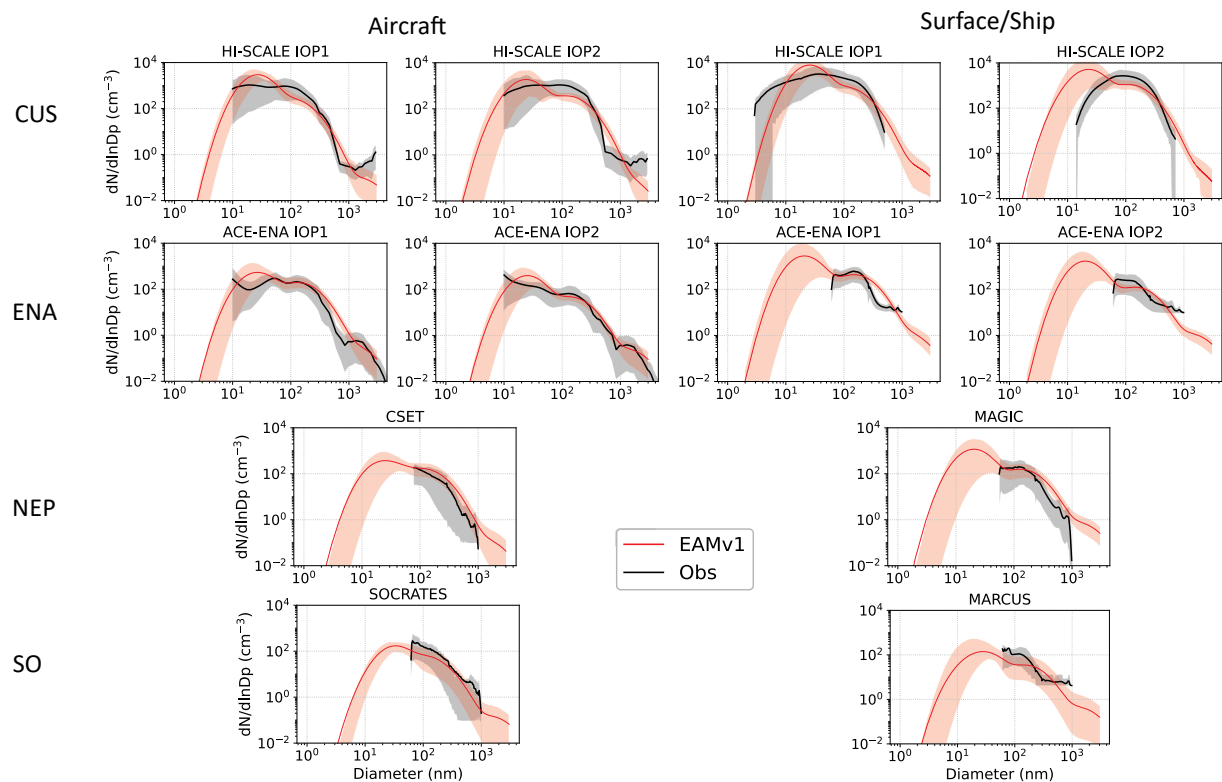



Figure 5: Mean aerosol number distribution averaged for each field campaign or IOP. Shadings denote the range between 10% and 90% percentiles.

276

Table 3: Mean aerosol number concentration and inter-quartile range (25% and 75% percentiles, small numbers in parenthesis) for two size ranges averaged for each field campaign (or each IOP for HI-SCALE and ACE-ENA). Aircraft measurements 30 minutes after take-off and before landing are excluded to remove possible contamination from the airport.

Unit: #/cm^3			>10 nm		>100 nm	
			CPC	E3SMv1	UHSAS/PCASP*	E3SMv1
CUS	Surface (HI-SCALE)	IOP1	4095 (2198, 4943)	4566 (2865, 5984)	675.1 (393.2, 929.5)	321.3 (229.7, 400.8)
		IOP2	N/A	N/A	N/A	N/A
	Aircraft (HI-SCALE)	IOP1	4206 (1132, 5013)	3872 (2803, 4946)	465.7 (112.6, 616.1)	159.6 (112.2, 200.5)
		IOP2	4121 (1610, 3829)	2514 (1332, 3584)	789.1 (444.4, 1088.0)	383.6 (280.7, 483.8)
ENA	Surface	IOP1	610 (343, 711)	1723 (600, 1650)	206.1 (134.5, 267.1)	209.8 (155.3, 255.5)

	(ACE-ENA)	IOP2	458 (239, 505)	843 (320, 1152)	59.6 (25.0, 71.9)	61.9 (53.6, 71.9)
	Aircraft (ACE-ENA)	IOP1	576 (264, 677)	919 (562, 917)	135.6 (65.3, 185.1)	199.9 (146.6, 266.3)
		IOP2	356 (132, 383)	521 (279, 627)	72.8 (22.2, 72.8)	50.3 (41.6, 62.3)
NEP	Ship (MAGIC)		417 (117, 285)	1271 (356, 1652)	113.6 (47.0, 139.9)	143.0 (93.7, 148.5)
	Aircraft (CSET)		408 (155, 386)	607 (353, 675)	81.5 (17.0, 73.4)	134.5 (81.2, 151.3)
SO	Ship (MARCUS)		354 (244, 415)	303 (164, 326)	68.5 (36.8, 94.0)	54.5 (32.5, 71.7)
	Aircraft (SOCRATES)		988 (327, 991)	237 (169, 270)	56.2 (14.1, 50.4)	32.3 (13.2, 42.2)

281 * PCASP is available only on aircraft for HI-SCALE and ACE-ENA. UHSAS is available only in surface

282 measurements for HI-SCALE and ACE-ENA, and in other field campaigns.

283   Both observed and simulated aerosol size distribution and number concentration show large variability

284 during these field campaigns. Over the period of a few weeks or longer, aerosol number can vary by an

285 order of magnitude between the 10% and 90% percentiles, especially for small particles (Figure 5). Figure

286 6 shows mean aerosol size distributions for two flight days during HI-SCALE: one with a large number of

287 small (<70 nm) particles (14 May) and the other (3 September) with fewer small particles but more

288 accumulation mode (70 – 300 nm) particles. On both days, EAMv1 reproduced the observed planetary

289 boundary layer (PBL) height (PBLH) reasonably well with sufficient samples below and above PBL. On

290 14 May, EAMv1 reproduces the observed aerosol size distribution reasonably well both within the PBL

291 and in the lower free atmosphere. However, on 3 September EAMv1 produces too many aerosols in the

292 Aitken mode and too few accumulation mode aerosols in the PBL. In the free atmosphere, EAMv1

293 reproduces the lower concentration of Aitken mode aerosols but still underestimates the accumulation

294 mode. Such contrasting cases will be useful to help diagnose the specific processes contributing to model

295 uncertainties in future analyses. This large day-to-day variability also indicates that long-term

296 measurements are needed to avoid sampling bias in building robust statistics in aerosol properties. The

297 next version of ESMAC Diags will be extended to include the available long-term ARM measurements at

298 SGP, ENA and other sites outside of the field campaign time periods.

299

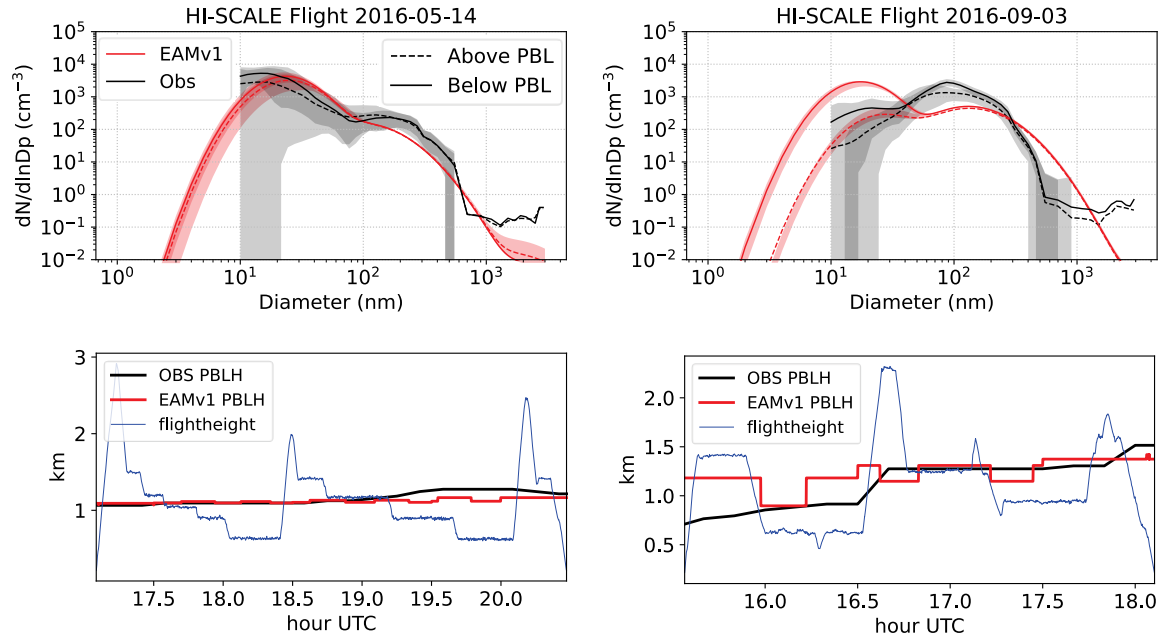



Figure 6: (Top) Mean aerosol number distribution for two flights during HI-SCALE: (left) 14 May 2016 and (right) 3 September 2016, for data above (dashed line) and below (solid line) observed PBLH. If there is cloud observed within a 1-hour window of the sample point, the above-PBL sample needs to be above cloud top and the below-PBL sample needs to be below cloud base for the sample point to be chosen. Shadings represent the data range between 10% and 90% percentiles. Relatively large particles with no shading indicate more than 90% of samples with zero values.

(bottom) Timeseries of observed (black) and simulated (red) PBLH overlaid with flight height (blue) during the two flight periods. The observed PBLH is derived from Doppler lidar measurements.

3.2 Vertical profiles of aerosol properties

A research aircraft is the primary platform to provide information on the vertical variations of key aerosol properties that cannot be obtained accurately by remote sensing instrumentation. In this section we show an example of evaluating vertical profiles of aerosol properties using aircraft measurements as well as illustrate the capability of evaluating multiple model simulations with ESMAC Diags. In addition to the standard EAMv1 simulation described in the previous section, we performed an EAMv1 simulation using the regionally refined mesh (RRM) (Tang et al., 2019). The model is configured to run with the horizontal grid spacing of ~0.25° over the continental U.S. and ~1° elsewhere. The two model configurations are identical, except for the higher spatial resolution (including primary aerosol emissions) in the RRM over

319   the continental U.S. All aircraft measurements with a cloud detected simultaneously (cloud flag = 1) were

320   excluded.

321   Figure 7 shows vertical percentiles of aerosol number concentration, composition and CCN number

322   concentration among all the HI-SCALE aircraft flights. Note that aircraft rarely flew above 3 km during

323   HI-SCALE so the sample size above that altitude is much smaller. The observed aerosol concentrations of

324   number and chemical composition decrease with height since the major sources of aerosols

325   (anthropogenic, biogenic, and biomass burning) (Liu et al., 2021) are from precursors emitted near the

326   surface and chemical formation within the PBL. EAMv1 generally simulates less variability than

327   observations except for sulfate. Overall, EAMv1 reproduces the observed mean aerosol number

328   concentration for aerosol size > 10 nm but underestimates the number of larger particles > 100 nm during

329   HI-SCALE (Table 3). The model also overestimates sulfate and underestimates organic matter

330   concentrations when compared to aircraft AMS measurements. Its underestimation of CCN number

331   concentration is consistent with underestimation of aerosol number concentration for diameter > 100 nm

332   but contrary with overestimation of sulfate. A similar relationship is seen for ACE-ENA, to be described

333   later in this section.

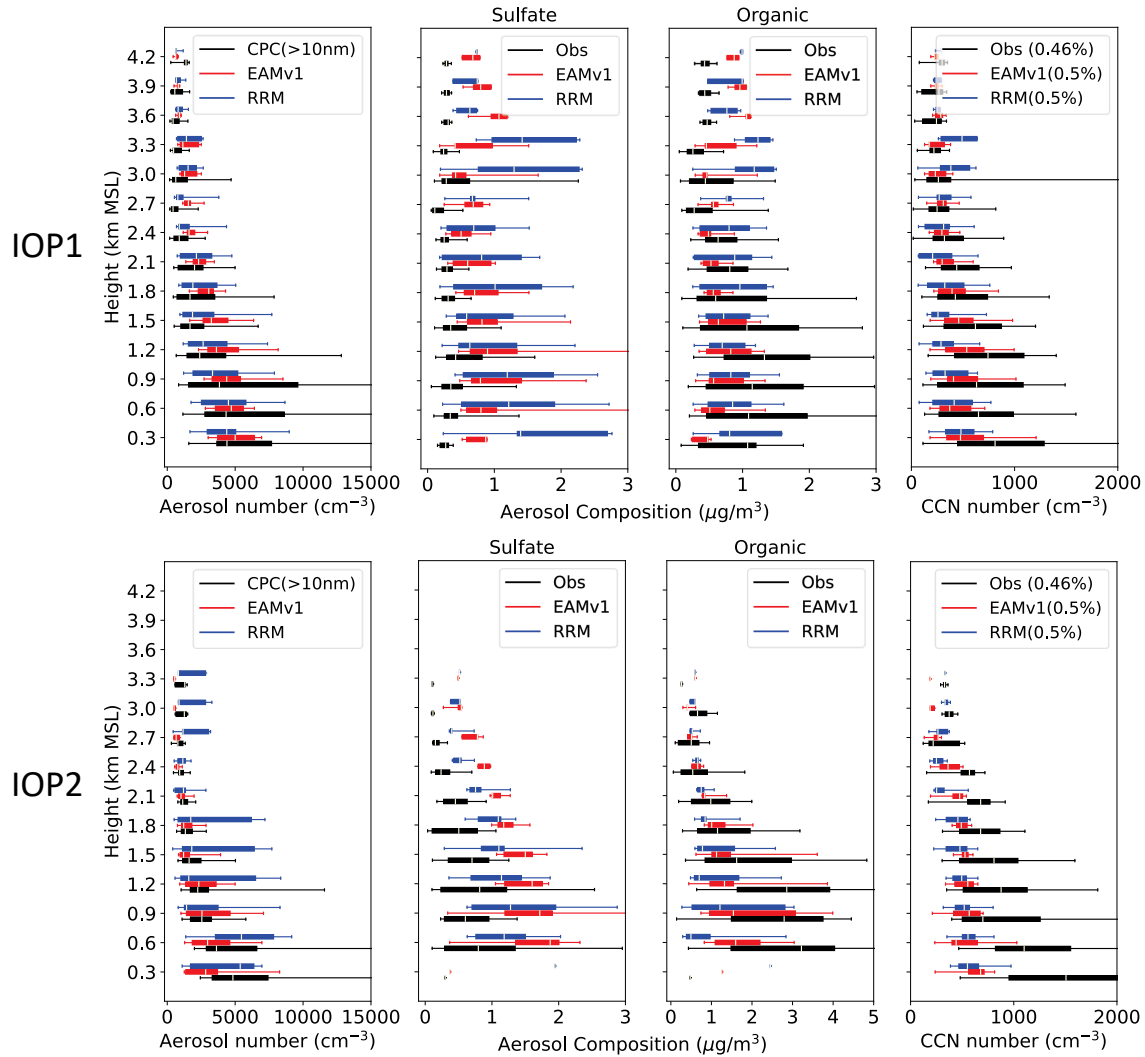
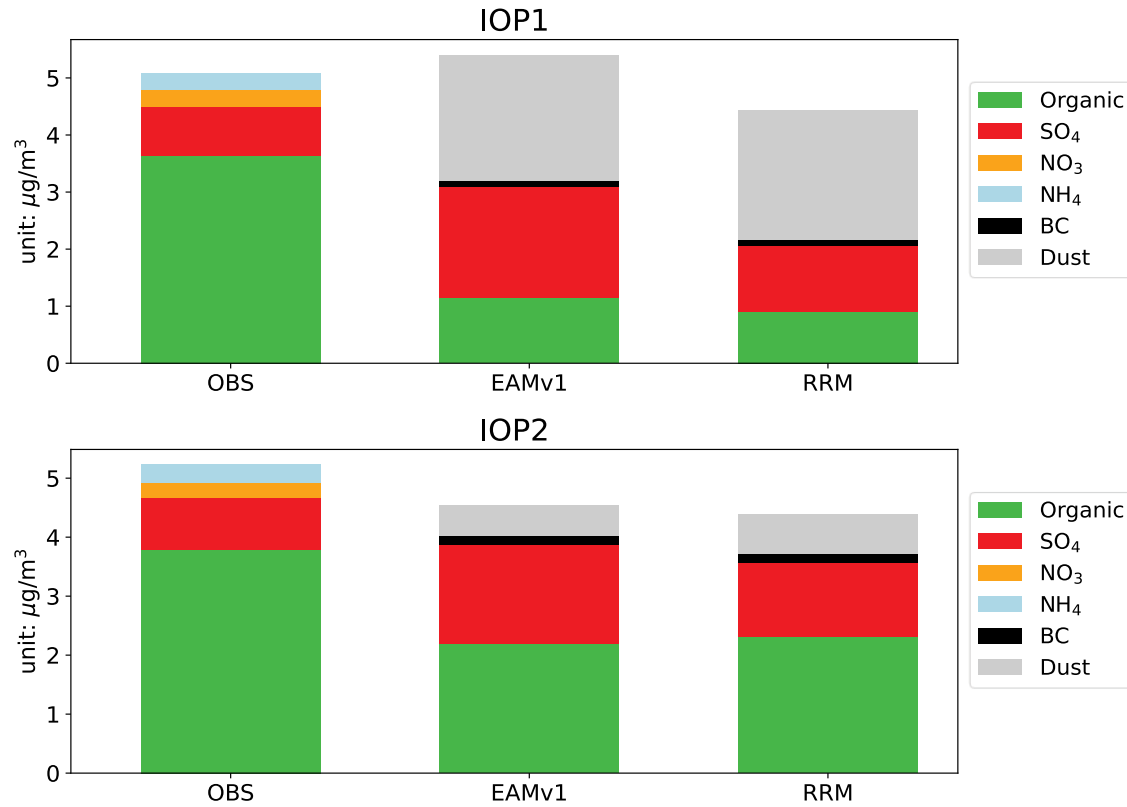



Figure 7: Vertical profiles of (from left to right): aerosol number concentration, mass concentration of sulfate, mass concentration of total organic matter and CCN number concentration under the supersaturation in the parentheses for HI-SCALE (top) IOP1 and (bottom) IOP2. The percentile box represents 25% and 75% percentiles, and the bar represents 5% and 95% percentiles.

The differences in sulfate and organic matter aloft is consistent with longer term surface measurement differences shown in Figure 8, suggesting this is a model bias. Note that near-surface measurements by aircraft are not always consistent with ground measurements (e.g., total organic matter in IOP1), which reflects the large spatial variability in aerosol properties associated with the aircraft flight paths up to a few hundred kilometers around the ARM site. The greater fraction of sulfate in EAMv1 suggests that the simulated aerosol hygroscopicity is likely higher than observed. Currently only these two species are available in both EAMv1 and AMS/ACSM observations for comparison purpose. Zaveri et al. (2021)

347 recently added chemistry associated with NO_3 formation in MAM4, which is expected to be implemented
348 in a future version of EAM.

349

350 **Figure 8: Bar plots of the surface average aerosol composition during HI-SCALE IOP1 (top) and**
351 **IOP2 (bottom). Observations are obtained from an ACSM. Dust and black carbon (BC) are not**
352 **measured in the observation. NO_3 and NH_4 are not predicted in EAMv1 and RRM.**

353   Ongoing developments in E3SM will soon permit regional-refined meshes with grid spacings as small
354 as ~ 3 km as well as global convection-permitting simulations (Δx ~ 3 km); therefore, this diagnostics
355 package is designed to be flexible in scale to take advantage of higher-resolution ESM simulations that
356 are more compatible with high-resolution in-situ aerosol observations. This study demonstrates this
357 ability by using a 0.25° RRM simulation. Overall, the RRM analyzed here has similar biases as EAMv1,
358 with differences that vary seasonally. The inter-quartile range in Figure 7 show that the variability of
359 organic aerosols and CCN from the EAMv1 and RRM simulations are similar. However, the variability of
360 sulfate in RRM is larger than EAMv1 and observations during the spring IOP (IOP1). During the summer
361 IOP (IOP2), the variabilities of sulfate in EAMv1, RRM, and observations are similar, and the sulfate
362 concentrations from RRM are closer to observed than EAMv1. Individual timeseries from the RRM
363 simulation are still too smooth to capture the fine scale variability of aerosols in observations (not shown).

364   We expect E3SM to capture more fine scale variabilities related to urban and point sources of aerosols

365   and their precursors when the simulation grid spacing is further reduced to ~ 3 km. A sensitivity study

366   will be conducted when this high-resolution version of E3SM simulation becomes available.

367   Figure 9 shows the vertical variation in percentiles of aerosol properties for ACE-ENA. The observed

368   aerosol number concentrations, composition masses, and CCN number concentrations are much smaller

369   than those for HI-SCALE, representing a cleaner ocean environment. EAMv1 produces larger mean

370   values than the observations for all these quantities. The overall variabilities in predicted aerosol number

371   and concentrations of sulfate and organic matter are also greater than observed. Note that the observed

372   variabilities for HI-SCALE are much larger than for ACE-ENA, indicating that EAMv1 has smaller

373   location variation on aerosol variabilities. The observed total organic concentration shows a peak aloft

374   between 1.6 and 2.2 km, corresponding to the level of CCN number concentration peak. This implies a

375   major source of aerosols or precursors is free tropospheric transport (Zawadowicz et al., 2021). This peak

376   of total organic concentration aloft is also captured by the model.

377   The bar plots in Figure 10 of aerosol composition at the surface during ACE-ENA from the ACSM

378   instrument and EAMv1 illustrate a similar bias in sulfate and organic mass as aloft. While the surface

379   sulfate measurements are like those from the aircraft at the lowest altitudes, the observed surface organic

380   matter is much higher than aloft, particularly during IOP2. The differences in these measurements may be

381   due to local effects or possible contamination from aircraft since the surface station is located near an

382   airport on an island.

383

384

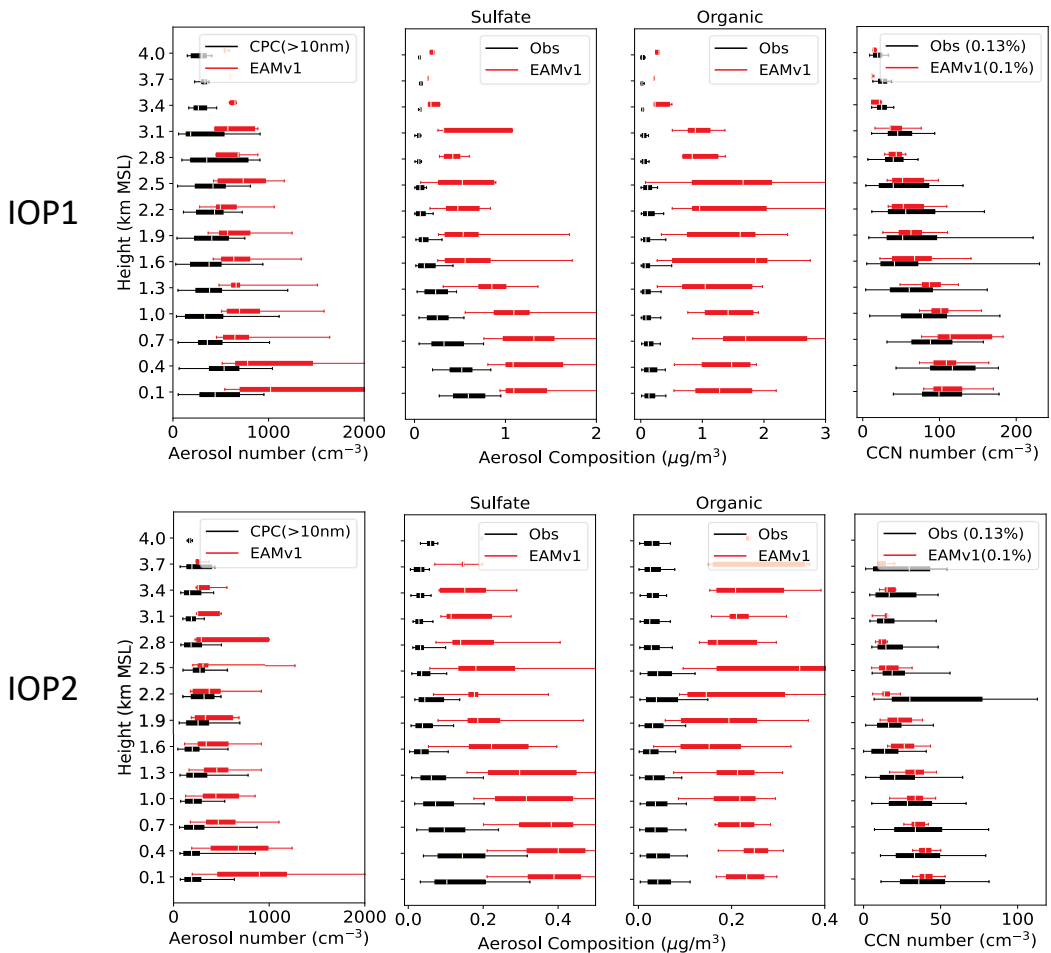



Figure 9: Same as Figure 7 but for ACE-ENA.

386

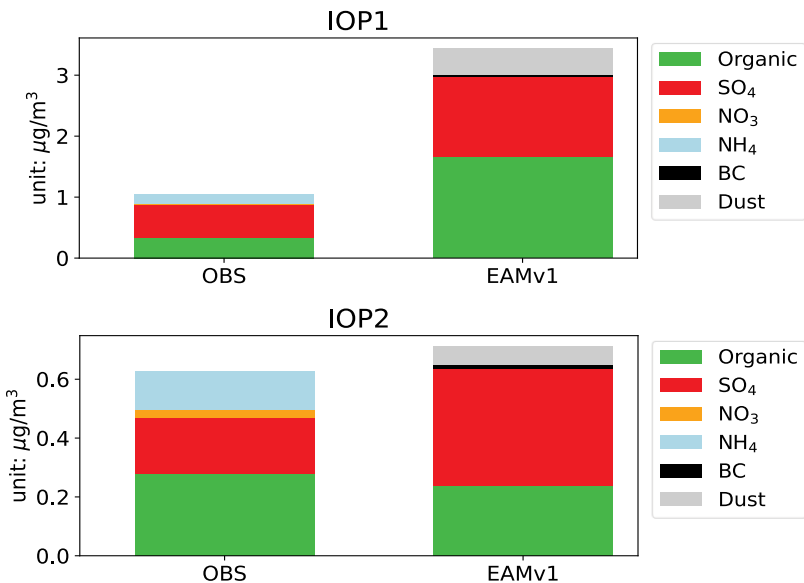



Figure 10: Same as Figure 8 but for ACE-ENA.

    ### 3.3 New particle formation events

389   Aerosol number concentrations and size distributions are highly impacted by NPF events (Kulmala et al.,

390   2004), which further influence CCN concentration (e.g., Kuang et al., 2009; Pierce and Adams, 2009) and

391   ultimately cloud properties. NPF and subsequent particle growth are frequently observed in the CUS

392   region (Hodshire et al., 2016). As described by Fast et al. (2019) and shown in Figure 11a, several NPF

393   events were observed during the HI-SCALE spring IOP (IOP1). Large concentrations of aerosols smaller

394   than 10 nm were observed, with the size growing larger over the next few hours. The average diurnal

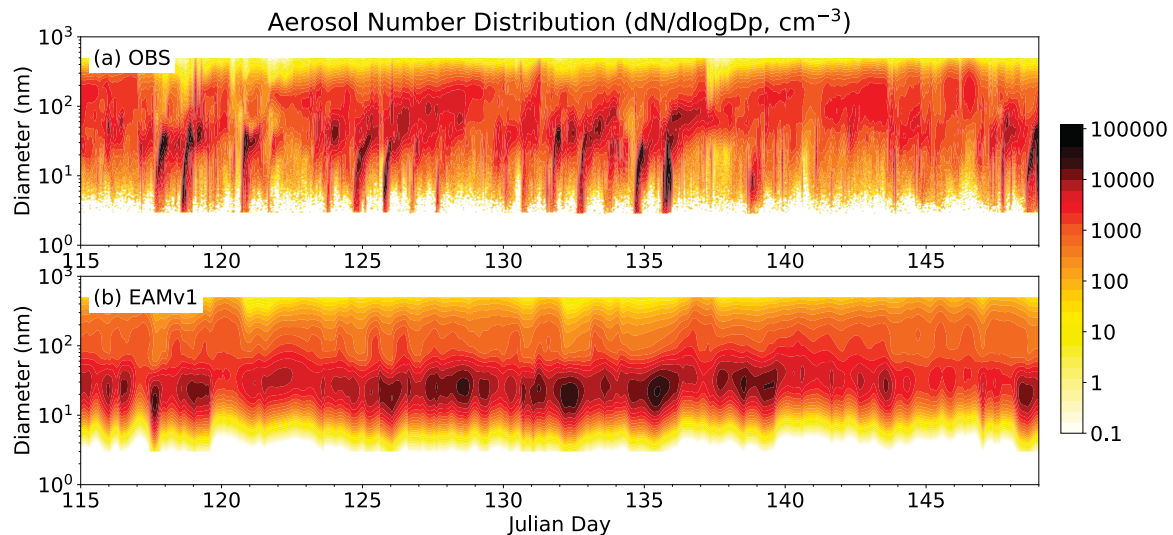
395   variation in aerosol number distribution in Figure 12a shows that NPF events usually occur during the

396   morning between 12 and 15 UTC (6 – 9 am local time), followed by particle growth during the rest of the

397   morning and afternoon. This variation is also seen in the diurnally averaged CPC measurements of

398   aerosol diameters > 3 nm and > 10 nm (Figure 12c) but diurnal changes in CCN number concentrations

399   (Figure 12d) are more modest.

400

401   **Figure 11: Time series of (a) observed and (b) simulated surface aerosol number distribution**

402   **during HI-SCALE IOP1. The observed aerosol number distribution is from merged nanoSMPS**

403   **and SMPS. Model data is cut off at 500 nm to compare with observation.**

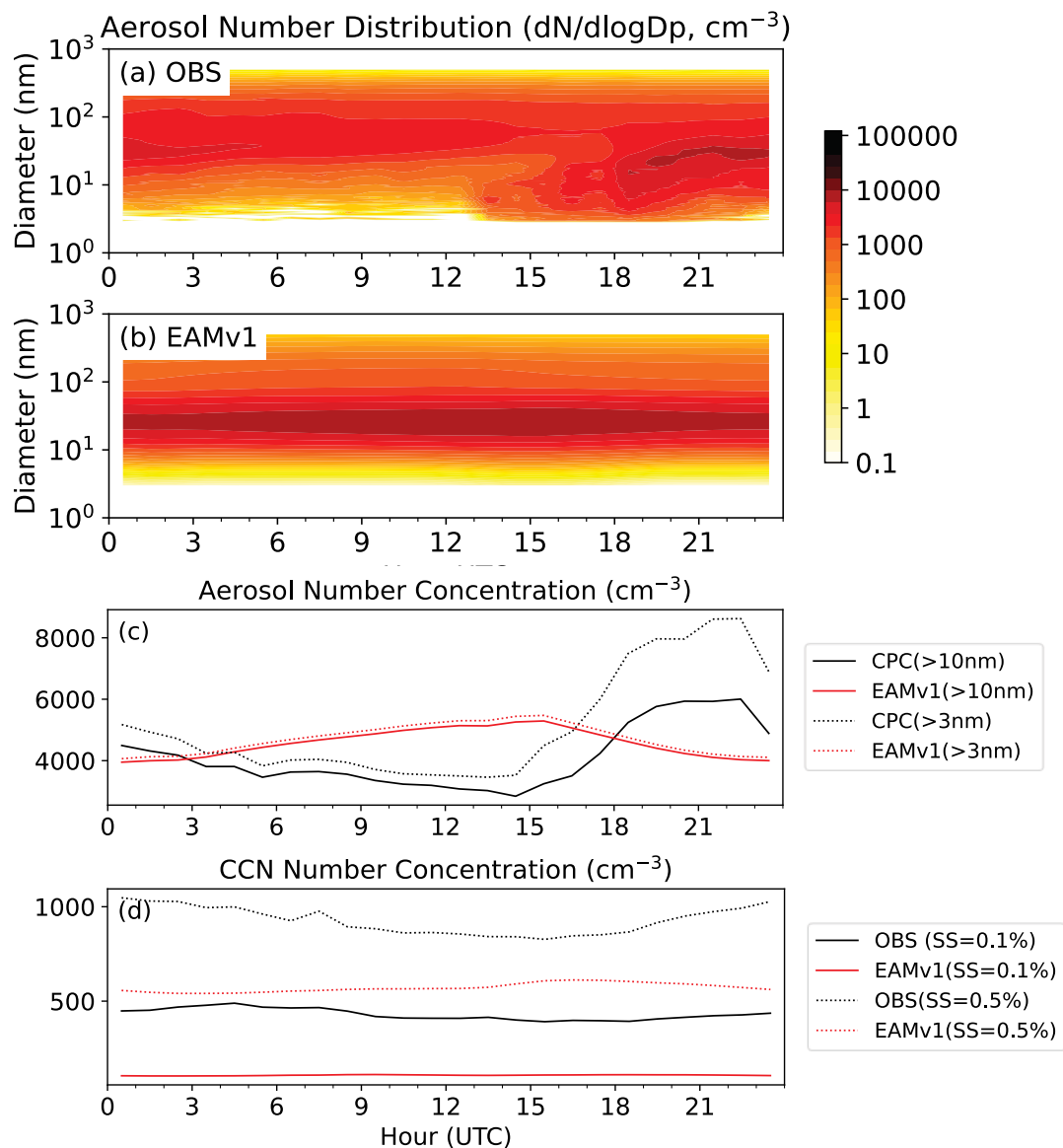



Figure 12: Average diurnal cycle of surface (a) observed aerosol number distribution, (b) simulated aerosol number distribution, (c) aerosol number concentration for diameters > 10 nm and > 3 nm, and (d) CCN number concentration for supersaturations of 0.1% and 0.5% for HI-SCALE IOP1.

Various NPF pathways associated with different chemical species have been proposed and implemented in models. Two NPF pathways are considered in MAM4 in EAMv1: a binary nucleation pathway and a PBL cluster nucleation pathway. However, the current simulation does not reproduce the observed large day-to-day variability of small particle concentrations due to NPF. Instead, the model produces high aerosol concentrations between 10 and 100 nm almost all the time. It also fails to reproduce

414   the large diurnal variability of aerosol and CCN number concentration with a peak seen in the morning

415   near 15 UTC (9 am local time), 7 hours earlier than the observed 22 UTC (4 pm local time) afternoon

416   peak. Its overestimation of aerosol number concentration for particle diameter >10 nm and

417   underestimation of CCN number concentration is consistent with that shown in Figure 5. Several efforts

418   are underway to improve the simulation of NPF by adding a nucleation mode in MAM4 to explicitly

419   resolve ultrafine particles and implementing new chemical pathways to simulate NPF following Zhao et

420   al. (2020). ESMAC Diags is being used to evaluate these new model developments.

421   Using aircraft measurements from ACE-ENA, Zheng et al. (2021) recently found evidence of NPF

422   events occurring in the upper part of marine boundary layer between broken clouds following the passage

423   of a cold front. 16 February 2018 is identified as a typical NPF day in Zheng et al. (2021). The vertical

424   profiles of aerosol number and CCN concentrations measured by aircraft on 16 February 2018 are shown

425   in Figure 13. The NPF event and particle growth happened in the upper boundary layer is shown by the

426   large mean and variance of aerosol number concentration just below the base of the marine boundary

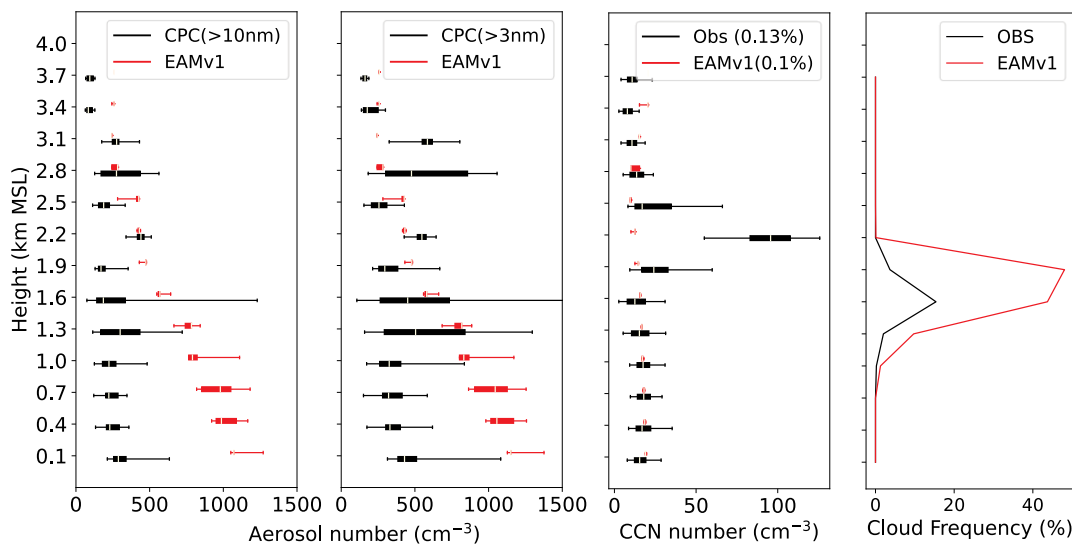
427   layer clouds. EAMv1 could not simulate NPF events in the upper marine boundary layer on this day and

428   other days during ACE-ENA, likely due to the lack of NPF mechanisms related to dimethyl sulfide

429   (DMS) oxidation, and/or missing part in parameterizations to deal with the processes related to broken

430   marine boundary layer clouds and sub-grid circulation. Similarly, the sharp increase of CCN number just

431   above the level of marine boundary layer clouds is not simulated.

432

433   **Figure 13: Vertical profiles of aerosol number concentration for diameters >10 nm and >3 nm,**

434   **CCN number concentration, and cloud frequency measured by the 16 February 2018 flight in**

435   **ACE-ENA. The percentile box represents 25% and 75% percentiles, and the bar represents 5%**

436   **and 95% percentiles.**

437

3.4 Latitudinal dependence of aerosols and clouds

Unlike some field campaigns (i.e., HI-SCALE and ACE-ENA) where aircraft missions were conducted over a relatively localized region with limited spatial variability of the meteorological conditions, ship and/or aircraft measurements over the NEP and SO testbed regions span regions > 1500 km (i.e., from California to Hawaii and from Tasmania to the far Southern Ocean, respectively). As shown in Figure 3, there are large spatial gradients in EAMv1 simulated aerosol optical depth along these ship/aircraft tracks. In ESMAC Diags version 1, we include composite plots of aerosol and cloud properties binned by latitude to assess model representation of synoptic-scale variations.

446

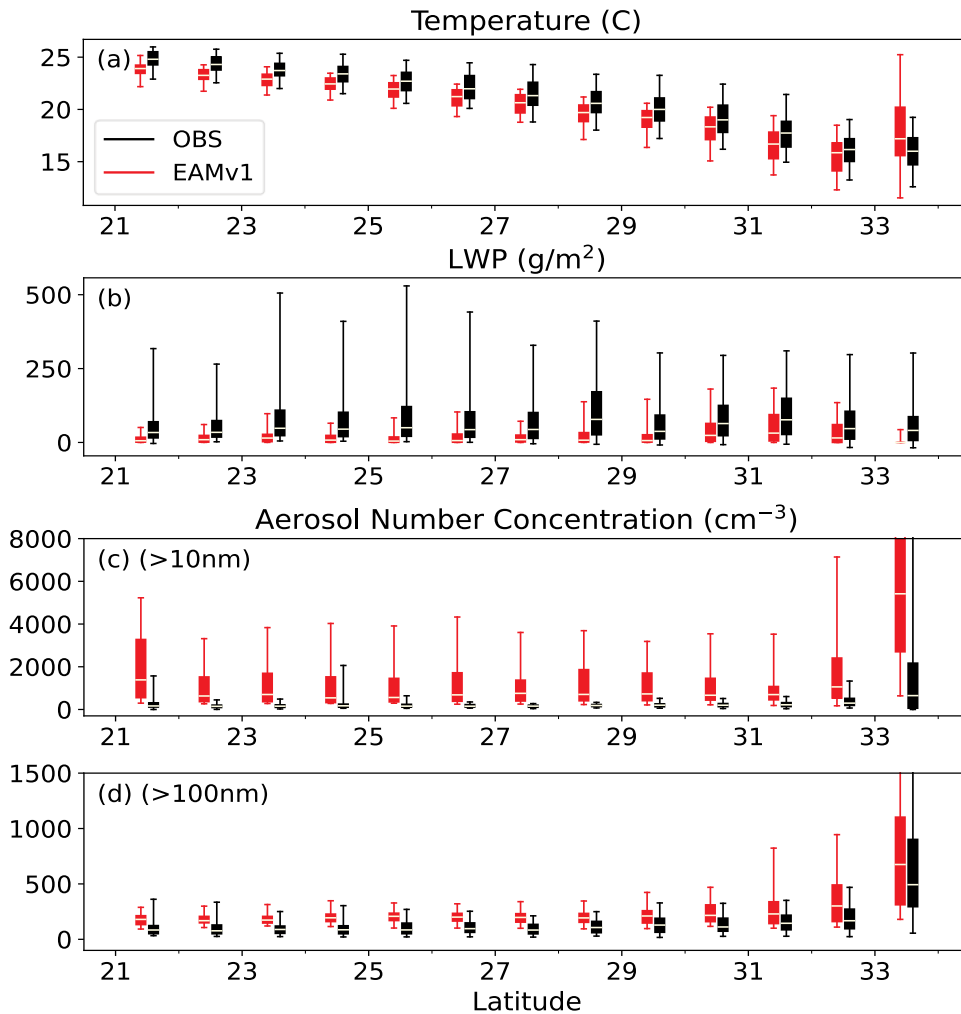



Figure 14: Percentiles of (a) air temperature, (b) grid-mean liquid water path (LWP), (c) aerosol number concentration for diameter >10 nm, and (d) aerosol number concentration for diameter >100 nm for all ship tracks in MAGIC binned by 1° latitude bins. The percentile box

453

454

Figure 15: Percentiles of (a) cloud fraction, (b) aerosol number concentration for diameter >10 nm,

and (c) aerosol number concentration for diameter >100 nm for all aircraft measurements between

0-3 km in CSET binned by 1° latitude bins. The percentile box represents 25% and 75%

percentiles, and the bar represents 5% and 95% percentiles. The observed aerosol number

concentrations for diameters >10 nm and >100 nm are obtained from CNC and UHSAS,

respectively.

461

462   The research ship (aircraft) from the MAGIC (CSET) field campaign in the NEP testbed travelled

463   between California and Hawaii, where there is frequently a transition between marine stratocumulus

464   clouds near California and broken trade cumulus clouds near Hawaii (e.g., Teixeira et al., 2011).

465   Although ESMAC Diags v1 focuses primarily on aerosols, we show some basic meteorological and cloud

466   fields here since they are important to illustrate the transition of cloud regimes along the ship (aircraft)

467   tracks. Additional cloud properties derived from surface and satellite measurements are not included in

468   the current analysis, but are being implemented in ESMAC Diags v2. Some of the meteorological, cloud,

469   and aerosol properties along the ship (aircraft) tracks binned by latitude are shown in Figure 14 (Figure

470   15). Note that cloud fraction in Figure 15 is calculated as cloud frequency in aircraft observation and from

471   grid-mean cloud fraction in model along the flight track. This is different from the classic definition of

472   cloud fraction usually used for satellite measurements or models and is subject to aircraft sampling

473   strategy. As the surface temperature increases from California to Hawaii (Figure 14a), the cloud fraction

474   (Figure 15a) shows an decreasing trend southwestward, indicating the transition from stratocumulus to

475   cumulus clouds. However, ship-measured LWP (Figure 14b) has no trend related to latitude, possibly

476   because cumulus clouds at lower latitudes have smaller cloud fraction but larger LWP when clouds exist.

477   EAMv1 shows decreasing trends of both cloud fraction and LWP from high to low latitudes along these

478   tracks. It generally underestimates LWP and overestimates cloud fraction to the north of 30° N. For

479   aerosol number concentrations, EAMv1 produces too many aerosols compared to measurements both at

480   the surface (ship) and aloft (aircraft), consistent with the aerosol size distribution in Figure 5 and total

481   number concentration in Table 3. However, EAMv1 does reproduce the increase trend in accumulation

482   mode aerosol concentration approaching the California coast.

483   Similar latitudinal gradients of aerosol and CCN number concentrations along ship tracks from

484   MARCUS and aircraft tracks from SOCRATES are shown in Figures 16 and 17, respectively. Over the

485   SO region, NPF frequently occurs during austral summer when ample biogenic precursor gases (e.g.,

486   DMS) are released and rise into the free troposphere (McFarquhar et al., 2021; McCoy et al., 2021). Large

487   values of ship-measured aerosol and CCN number concentration are observed near Antarctica

488   corresponding to the coastal biological emissions of aerosol precursors, and also occur to the north of

489   45°S, indicating impacts from continental and anthropogenic sources. This is consistent with other studies

490   (Sanchez et al., 2021; Humphries et al., 2021). EAMv1 underestimates aerosol and CCN number

491   concentration near Antarctica. This bias, which may be related to too strong wet scavenging or

492   insufficient NPF and growth, is commonly seen in many other ESMs (e.g., McCoy et al., 2020; McCoy et

493   al., 2021). Aircraft flight paths during SOCRATES (Figure 17) do not extent as far south as the ship

494   measurements (Figure 16). The observed aerosol properties have little latitudinal variation in general.

495   EAMv1 underestimates aerosol number concentration for size > 10nm and CCN number concentration

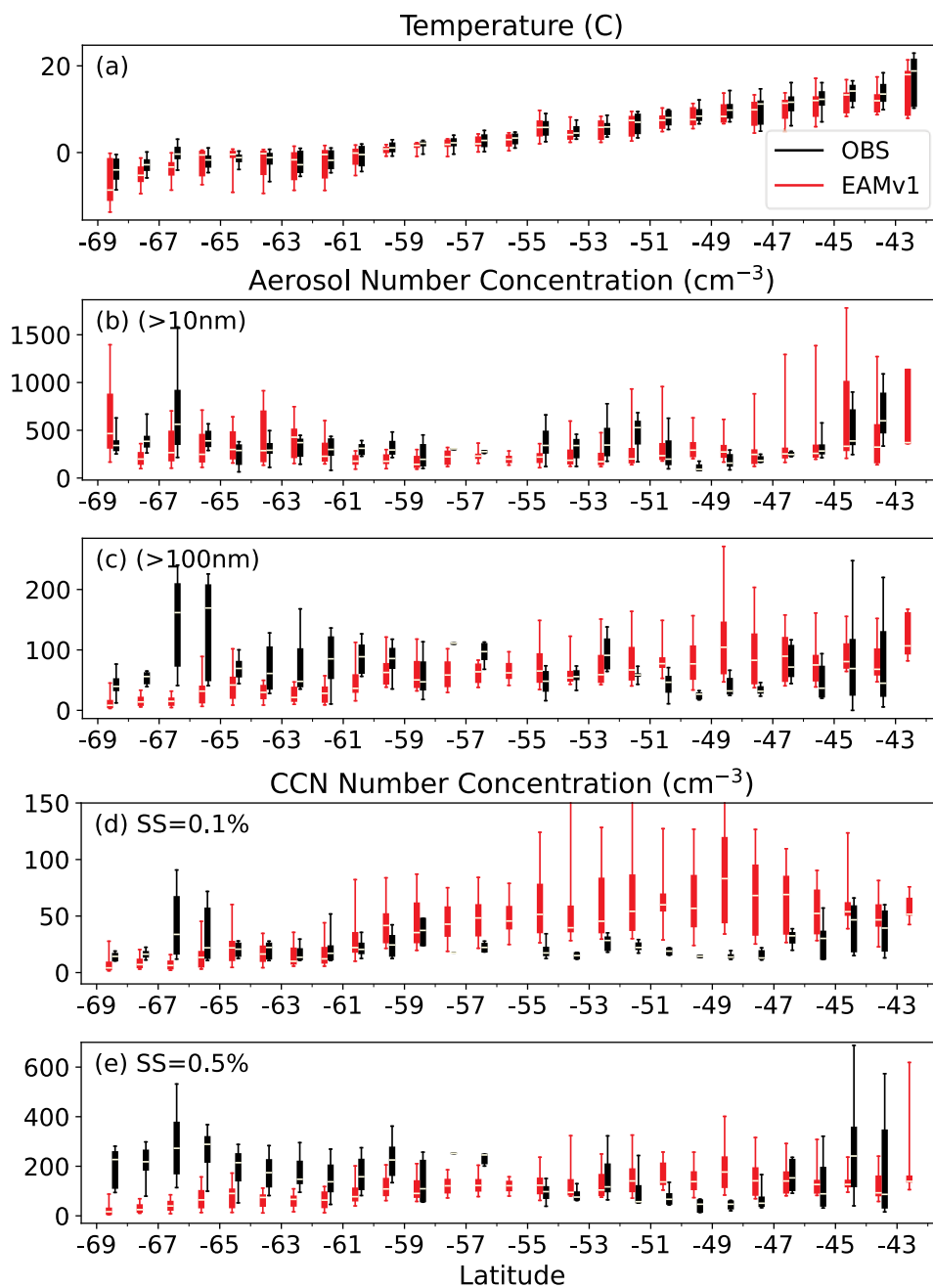
496   with SS=0.5%, but the predictions are closer to observed for aerosol size > 100 nm and CCN with

497   SS=0.1% (Figure 17), consistent with the mean aerosol size distribution in Figure 5. This indicates that

498   the model performs better in simulating accumulation mode than Aitken mode particles over SO. These

499   model aerosol biases are highly relevant when considering their interaction with clouds and radiations,

500   which will be included in version 2 of ESMAC Diags.

501

502

503   **Figure 16: Percentiles of (a) air temperature, (b) aerosol number concentration for diameter >10**

504   **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.

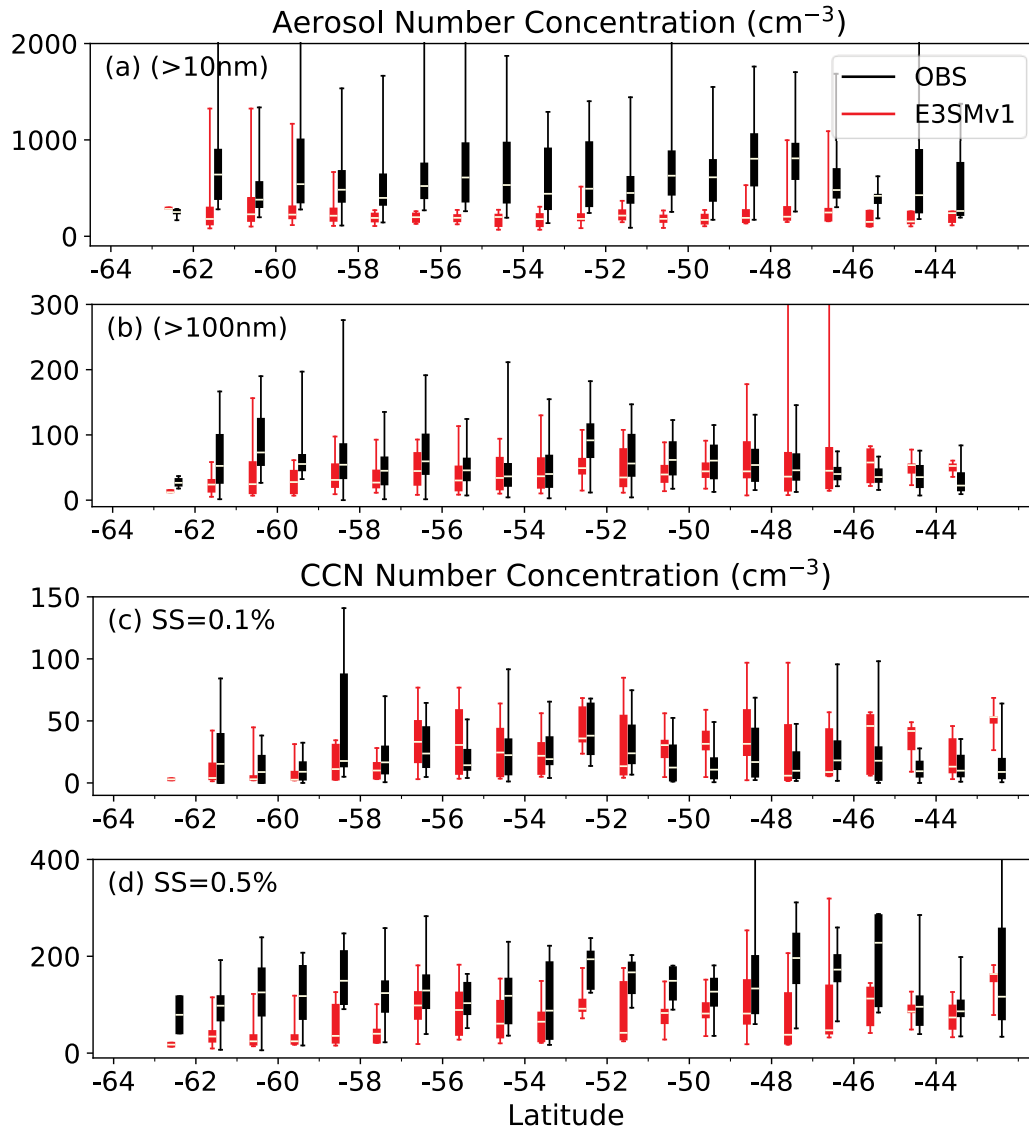



Figure 17: Percentiles of (a) aerosol number concentration for diameter >10 nm, (b) aerosol number concentration for diameter >100 nm, (c) CCN number concentration for supersaturation SS=0.1%, and (d) CCN number concentration for supersaturation SS=0.5% for all aircraft measurements between 0-3 km in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

4. Summary

515   A Python-based ESM aerosol-cloud diagnostics (ESMAC Diags) package is developed to quantify the

516   performance of the DOE's E3SM atmospheric model using ARM and NCAR field campaign

517   measurements. The first version of this diagnostics package focuses on aerosol properties. The

518   measurements include aerosol number, size distribution, chemical composition, and CCN collected from

519   surface, aircraft, and ship platforms needed to assess how well the aerosol lifecycle is represented across

520   spatial and temporal scales which will subsequently impact uncertainties in aerosol radiative forcing

521   estimates. Currently, the diagnostics cover the field campaigns of ACE-ENA, HI-SCALE,

522   MAGIC/CSET, and MARCUS/SOCRATES over Northeastern Atlantic, Continental U.S., Northeastern

523   Pacific, and Southern Ocean, respectively. The code structure is designed to be flexible and modular for

524   future extension to other field campaigns or additional datasets. Since there is no one instrument that can

525   measure the entire aerosol size distribution, we have constructed merged aerosol size distributions from

526   two or more ARM instruments to better assess predicted size distributions. An "aircraft simulator" is used

527   to extract aerosol and meteorological model variables along flight paths that vary in space and time.

528   Similarly, the aircraft simulator is applied to ship tracks in which the altitude remains fixed at sea level.

529   The version 1 of ESMAC Diags package provides various types of diagnostics and metrics, including

530   timeseries, diurnal cycles, mean aerosol size distribution, pie charts for aerosol composition, percentiles

531   by height, percentiles by latitude, mean statistics of aerosol number concentration, and more. This allows

532   quantification of model performance predicting aerosol number, size, composition, vertical distribution,

533   spatial distribution (along ship tracks or aircraft tracks) and new particle formation events. A full set of

534   diagnostics plots and metrics for simulations used in this paper are available at

535   https://portal.nersc.gov/project/m3525/sqtang/ESMAC_Diags_v1/forGMD/webpage/ and are archived as

536   an electronic supplement of this paper. This paper shows some examples to demonstrate the capability of

537   ESMAC Diags to evaluate EAMv1 simulated aerosol properties. The diagnostics package also allows

538   multiple simulations in one plot to compare different models or model versions. It can also be applied to

539   evaluate other ESMs with necessary modifications to fit different model output formats.

540   Because in-situ aerosol measurements are usually collected at high temporal frequency (typically 1

541   second to a minute) over fine spatial volumes, there is a spatiotemporal scale mismatch with the standard

542   climate model resolution (usually 1-degree grid spacing with hourly output). This is a limitation that

543   cannot be completely overcome and must be accepted to perform model-observation comparisons

544   necessary for identifying shortcomings in model representation of aerosol, cloud, and aerosol-cloud

545   interaction processes that are the primary source for uncertainties in prediction of future climate. As new

546   versions of E3SM become available that has grid spacings as small as a few kilometers via regional-

547   refined and convection-permitting global domains (e.g., Caldwell et al., 2021), spatiotemporal

548 variabilities of aerosols at finer scales should be captured and be more compatible with fine resolution
549 observations such that resolution impacts on statistical differences can be quantified. The diagnostics
550 package will be applied to diagnose high resolution model output when the data are available.

551 While the current version focuses on aerosol properties, a version 2 of ESMAC Diags is being
552 developed to include more diagnostics and metrics for cloud, precipitation, and radiation properties to
553 facilitate the evaluation of aerosol-cloud interactions. These include inversion strength, above cloud
554 relative humidity, cloud-surface coupling, cloud fraction, depth, LWP, optical depth, effective radius,
555 droplet number concentration, adiabaticity, and albedo, precipitation rate, and more. Long-term surface-
556 based and satellite retrievals will also be used to provide better statistics in model evaluation and to
557 address limitations related to data coverage and uncertainty. Analyses are being designed to quantify
558 relationships between these variables and relate them to effective radiative forcing, which will be used to
559 assess and improve model parameterizations. In the future, this diagnostics package may also be extended
560 to include other field campaigns that provide valuable data on aerosol properties and cloud-aerosol
561 interactions, such as the ARM Layered Atlantic Smoke Interactions with Clouds (LASIC, Zuidema et al.,
562 2018), NASA ObseRvations of Aerosols above CLouds and their intEractionS (ORACLES, Redemann et
563 al., 2021), or NASA Atmospheric Tomography Mission (ATom, Brock et al., 2019) campaigns. As an
564 open-source package, ESMAC Diags can also be applied by any user to other ESMs with small
565 modifications on model preprocessing.

566 While there are other efforts to develop model diagnostics packages, this diagnostics package provides
567 a unique capability for detailed evaluation of aerosol properties that are tightly connected with
568 parameterized processes. Together with other commonly used diagnostics packages such as the ARM
569 diagnostics package (Zhang et al., 2020), the DOE E3SM diagnostics package, and the PCMDI's metrics
570 package (Gleckler et al., 2016), we expect to better understand the strengths and weaknesses of E3SM or
571 other ESMs and provide insights into model deficiencies to guide future model development. This
572 includes studies that develop a better understanding of how various processes contribute to uncertainties
573 in aerosol number and composition predictions and subsequent representation of CCN and aerosol
574 radiative forcing estimates.

575

Code availability:

The current version of ESMAC Diags is publicly available through GitHub (https://github.com/eagles-project/ESMAC_diags) under the new BSD license. The exact version (1.0.0-beta.2) of the code used to produce the results used in this paper is archived on Zenodo (https://doi.org/10.5281/zenodo.6371596). The model simulation used in this paper is version 1.0 of E3SM (https://doi.org/10.11578/E3SM/dc.20180418.36). The model configuration and execution scripts are uploaded as electronic supplement to this paper.

Data availability:

Measurements from the HI-SCALE, ACE-ENA, MAGIC, and MARCUS campaigns as well as the SGP and ENA sites are supported by the DOE Atmospheric Radiation Measurement (ARM) user facility and available at https://adc.arm.gov/discovery/. Measurements from the CSET and SOCRATES campaigns are supported by National Science Foundation (NSF) and obtained from NCAR Earth Observing Laboratory at https://data.eol.ucar.edu/master_lists/generated/cset/ and https://data.eol.ucar.edu/master_lists/generated/socrates/, respectively. DOI numbers or references of individual instruments are given in Table 2. All the above observational data and preprocessed model data used to produce the results used in this paper is archived on Zenodo (https://doi.org/10.5281/zenodo.6369120).

Author contribution:

ST, JDF and PM designed the diagnostics package; ST wrote the code and performed the analysis; JES, FM and MAZ processed the field campaign data; KZ contributed to the model simulation; JCH and ACV contributed to the package design and setup; ST wrote the original manuscript; all authors reviewed and edited the manuscript.

Competing interests:

Po-Lun Ma is a Topical Editor of Geoscientific Model Development. Other authors declare that they have no conflict of interest.

Acknowledgements:

This study was supported by the Enabling Aerosol-cloud interactions at GLobal convection-permitting scalES (EAGLES) project (74358), funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Earth System Model Development (ESMD) program area. We thank the numerous instrument mentors for providing the data. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. Pacific Northwest National Laboratory (PNNL) is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830.

References:

Albrecht, B., Ghate, V., Mohrmann, J., Wood, R., Zuidema, P., Bretherton, C., Schwartz, C., Eloranta, E., Glienke, S., Donaher, S., Sarkar, M., McGibbon, J., Nugent, A. D., Shaw, R. A., Fugal, J., Minnis, P., Paliknoda, R., Lussier, L., Jensen, J., Vivekanandan, J., Ellis, S., Tsai, P., Rilling, R., Haggerty, J., Campos, T., Stell, M., Reeves, M., Beaton, S., Allison, J., Stossmeister, G., Hall, S., and Schmidt, S.: Cloud System Evolution in the Trades (CSET): Following the Evolution of Boundary Layer Cloud Systems with the NSF–NCAR GV, Bull. Amer. Meteor. Soc., 100, 93-121, https://doi.org/10.1175/bams-d-17-0180.1, 2019.

Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, Science, 245, 1227-1230, https://doi.org/10.1126/science.245.4923.1227, 1989.

AMWG Diagnostic Package: https://www.cesm.ucar.edu/working_groups/Atmosphere/amwg-diagnostics-package/, last access: 2 November 2021. 2021.

ARM Intensive Operational Period (IOP) Data Browser: https://iop.archive.arm.gov/arm-iop/2012/mag/magic/reynolds-marmet/, last access: 2 November 2021. 2014.

ARM Intensive Operational Period (IOP) Data Browser: https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/matthews-wcm, last access: 2 November 2021. 2016a.

ARM Intensive Operational Period (IOP) Data Browser: https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/mei-cpc, last access: 16 September 2021. 2016b.

ARM Intensive Operational Period (IOP) Data Browser: https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/wang-fims, last access: 2 November 2021. 2017.

ARM Intensive Operational Period (IOP) Data Browser: https://iop.archive.arm.gov/arm-iop/2017/ena/aceena/tomlinson-pcas, last access: 2 November 2021. 2018.

ARM Intensive Operational Period (IOP) Data Browser: https://iop.archive.arm.gov/arm-iop/2017/ena/aceena/wang-fims/, last access: 2 November 2021. 2020.

Brock, C. A., Williamson, C., Kupc, A., Froyd, K. D., Erdesz, F., Wagner, N., Richardson, M., Schwarz, J. P., Gao, R. S., Katich, J. M., Campuzano-Jost, P., Nault, B. A., Schroder, J. C., Jimenez, J. L., Weinzierl, B., Dollner, M., Bui, T., and Murphy, D. M.: Aerosol size distributions during the Atmospheric Tomography Mission (ATom): methods, uncertainties, and data products, Atmos. Meas. Tech., 12, 3081-3099, https://doi.org/10.5194/amt-12-3081-2019, 2019.

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., Beydoun, H., Taylor, M., Bertagna, L., Bradley, A. M., Clevenger, T. C., Donahue, A. S., Eldred, C., Foucar, J., Golaz, J.-C., Guba, O., Jacob, R., Johnson, J., Krishna, J., Liu, W., Pressel, K., Salinger, A. G., Singh, B., Steyer, A., Ullrich, P., Wu, D., Yuan, X., Shpund, J., Ma, H.-Y., and Zender, C. S.: Convection-Permitting Simulations With the E3SM Global Atmosphere Model, Journal of Advances in Modeling Earth Systems, 13, e2021MS002544, https://doi.org/https://doi.org/10.1029/2021MS002544, 2021.

Campbell, J. R., Hlavka, D. L., Welton, E. J., Flynn, C. J., Turner, D. D., Spinhirne, J. D., Scott , V. S., III, and Hwang, I. H.: Full-Time, Eye-Safe Cloud and Aerosol Lidar Observation at Atmospheric Radiation Measurement Program Sites: Instruments and Data Processing, J. Atmos. Ocean. Technol., 19, 431-442, https://doi.org/10.1175/1520-0426(2002)019<0431:FTESCA>2.0.CO;2, 2002.

648 Carslaw, K. S., Lee, L. A., Regayre, L. A., and Johnson, J. S.: Climate Models Are Uncertain, but We Can
649 Do Something About It, Eos, 99, https://doi.org/10.1029/2018EO093757, 2018.

650 E3SM Diagnostics: https://e3sm.org/resources/tools/diagnostic-tools/e3sm-diagnostics/, last access: 2
651 November 2021. 2021.

652 Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I.,
653 Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K. D., Hagemann, S.,
654 Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Mäkelä, J., Martin, G., Mason,
655 E., Phillips, A. S., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., van Ulft, L. H., Walton, J., Wang,
656 S., and Williams, K. D.: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for
657 routine evaluation of Earth system models in CMIP, Geosci. Model Dev., 9, 1747-1802,
658 https://doi.org/10.5194/gmd-9-1747-2016, 2016.

659 Fast, J. D., Gustafson, W. I., Chapman, E. G., Easter, R. C., Rishel, J. P., Zaveri, R. A., Grell, G. A., and
660 Barth, M. C.: The Aerosol Modeling Testbed: A Community Tool to Objectively Evaluate Aerosol Process
661 Modules, Bull. Amer. Meteor. Soc., 92, 343-360, https://doi.org/10.1175/2010bams2868.1, 2011.

662 Fast, J. D., Berg, L. K., Alexander, L., Bell, D., D'Ambro, E., Hubbe, J., Kuang, C., Liu, J., Long, C.,
663 Matthews, A., Mei, F., Newsom, R., Pekour, M., Pinterich, T., Schmid, B., Schobesberger, S., Shilling, J.,
664 Smith, J. N., Springston, S., Suski, K., Thornton, J. A., Tomlinson, J., Wang, J., Xiao, H., and Zelenyuk,
665 A.: Overview of the HI-SCALE Field Campaign: A New Perspective on Shallow Convective Clouds, Bull.
666 Amer. Meteor. Soc., 100, 821-840, https://doi.org/10.1175/bams-d-18-0030.1, 2019.

667 Gates, W. L., Boyle, J. S., Covey, C., Dease, C. G., Doutriaux, C. M., Drach, R. S., Fiorino, M., Gleckler,
668 P. J., Hnilo, J. J., Marlais, S. M., Phillips, T. J., Potter, G. L., Santer, B. D., Sperber, K. R., Taylor, K. E.,
669 and Williams, D. N.: An Overview of the Results of the Atmospheric Model Intercomparison Project
670 (AMIP I), Bull. Amer. Meteor. Soc., 80, 29-56, https://doi.org/10.1175/1520-
671 0477(1999)080<0029:Aootro>2.0.Co;2, 1999.

672 Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C., Darmenov, A.,
673 Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V.,
674 Conaty, A., da Silva, A., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka,
675 G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-
676 Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Climate, 30, 5419-
677 5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

678 Gleckler, P. J., Doutriaux, C., Durack, P. J., Taylor, K. E., Zhang, Y., Williams, D. N., Mason, E., and
679 Servonnat, J.: A more powerful reality test for climate models, Eos, Trans. Amer. Geophys. Union, 97,
680 https://doi.org/10.1029/2016EO051663, 2016.

681 Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G.,
682 Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter,
683 M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter,
684 R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B.
685 E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D.,
686 Klein, S. A., Larson, V. E., Leung, L. R., Li, H.-Y., Lin, W., Lipscomb, W. H., Ma, P.-L., Mahajan, S.,
687 Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch,
688 P. J., Reeves Eyre, J. E. J., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G.,
689 Shaheen, Z., Shi, X., Singh, B., Tang, J., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M.,
690 Wan, H., Wang, H., Wang, S., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S., Yang, Y., Yoon, J.-

H., Zelinka, M. D., Zender, C. S., Zeng, X., Zhang, C., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, J. Adv. Model. Earth Syst., 11, 2089-2129, https://doi.org/10.1029/2018ms001603, 2019.

Hodshire, A. L., Lawler, M. J., Zhao, J., Ortega, J., Jen, C., Yli-Juuti, T., Brewer, J. F., Kodros, J. K., Barsanti, K. C., Hanson, D. R., McMurry, P. H., Smith, J. N., and Pierce, J. R.: Multiple new-particle growth pathways observed at the US DOE Southern Great Plains field site, Atmos. Chem. Phys., 16, 9321-9348, https://doi.org/10.5194/acp-16-9321-2016, 2016.

Howie, J. and Kuang, C.: Scanning Mobility Particle Sizer (AOSSMPS) [dataset], https://doi.org/10.5439/1476898, Accessed 14 January 2020.

Humphries, R. S., Keywood, M. D., Gribben, S., McRobert, I. M., Ward, J. P., Selleck, P., Taylor, S., Harnwell, J., Flynn, C., Kulkarni, G. R., Mace, G. G., Protat, A., Alexander, S. P., and McFarquhar, G.: Southern Ocean latitudinal gradients of cloud condensation nuclei, Atmos. Chem. Phys., 21, 12757-12782, https://doi.org/10.5194/acp-21-12757-2021, 2021.

IPCC, Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M. (Eds.): Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp., https://doi.org/10.1017/CBO9781107415324, 2013.

Koontz, A. and Kuang, C.: Scanning mobility particle sizer (AOSNANOSMPS) [dataset], https://doi.org/10.5439/1242975, Accessed 14 January 2020.

Koontz, A. and Uin, J.: Ultra-High Sensitivity Aerosol Spectrometer (AOSUHSAS) [dataset], https://doi.org/10.5439/1333828, Accessed 3 February 2021.

Kuang, C., McMurry, P. H., and McCormick, A. V.: Determination of cloud condensation nuclei production from measured new particle formation events, Geophys. Res. Lett., 36, https://doi.org/10.1029/2009GL037584, 2009.

Kuang, C., Salwen, C., and Singh, A.: Condensation Particle Counter (AOSCPCF) [dataset], https://doi.org/10.5439/1046184, Accessed 24 November 2020.

Kuang, C., Andrews, E., Salwen, C., and Singh, A.: Condensation Particle Counter (AOSCPC) [dataset], https://doi.org/10.5439/1025152, Accessed 16 September 2020.

Kuang, C., Salwen, C., Boyer, M., and Singh, A.: Condensation Particle Counter (AOSCPCF1M) [dataset], https://doi.org/10.5439/1418260, Accessed 3 February 2021.

Kulmala, M., Vehkamäki, H., Petäjä, T., Dal Maso, M., Lauri, A., Kerminen, V. M., Birmili, W., and McMurry, P. H.: Formation and growth rates of ultrafine atmospheric particles: a review of observations, J. Aerosol Sci., 35, 143-176, https://doi.org/10.1016/j.jaerosci.2003.10.003, 2004.

Kyrouac, J. and Shi, Y.: Surface Meteorological Instrumentation (MET) [dataset], https://doi.org/10.5439/1786358, Accessed 11 November 2020.

Lewis, E. R. and Teixeira, J.: Dispelling clouds of uncertainty, Eos, Trans. Amer. Geophys. Union, 96, https://doi.org/10.1029/2015eo031303, 2015.

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.

Liu, X., Ma, P. L., Wang, H., Tilmes, S., Singh, B., Easter, R. C., Ghan, S. J., and Rasch, P. J.: Description and evaluation of a new four-mode version of the Modal Aerosol Module (MAM4) within version 5.3 of the Community Atmosphere Model, Geosci. Model Dev., 9, 505-522, https://doi.org/10.5194/gmd-9-505-2016, 2016.

Maloney, E. D., Gettelman, A., Ming, Y., Neelin, J. D., Barrie, D., Mariotti, A., Chen, C. C., Coleman, D. R. B., Kuo, Y.-H., Singh, B., Annamalai, H., Berg, A., Booth, J. F., Camargo, S. J., Dai, A., Gonzalez, A., Hafner, J., Jiang, X., Jing, X., Kim, D., Kumar, A., Moon, Y., Naud, C. M., Sobel, A. H., Suzuki, K., Wang, F., Wang, J., Wing, A. A., Xu, X., and Zhao, M.: Process-Oriented Evaluation of Climate and Weather Forecasting Models, Bull. Amer. Meteor. Soc., 100, 1665-1686, https://doi.org/10.1175/bams-d-18-0042.1, 2019.

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, J. Geophys. Res. Atmos., 126, https://doi.org/10.1029/2020jd033529, 2021.

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, 117, 18998-19006, https://doi.org/doi:10.1073/pnas.1922502117, 2020.

McFarquhar, G. M., Bretherton, C. S., Marchand, R., Protat, A., DeMott, P. J., Alexander, S. P., Roberts, G. C., Twohy, C. H., Toohey, D., Siems, S., Huang, Y., Wood, R., Rauber, R. M., Lasher-Trapp, S., Jensen, J., Stith, J. L., Mace, J., Um, J., Järvinen, E., Schnaiter, M., Gettelman, A., Sanchez, K. J., McCluskey, C. S., Russell, L. M., McCoy, I. L., Atlas, R. L., Bardeen, C. G., Moore, K. A., Hill, T. C. J., Humphries, R. S., Keywood, M. D., Ristovski, Z., Cravigan, L., Schofield, R., Fairall, C., Mallet, M. D., Kreidenweis, S. M., Rainwater, B., D'Alessandro, J., Wang, Y., Wu, W., Saliba, G., Levin, E. J. T., Ding, S., Lang, F., Truong, S. C. H., Wolff, C., Haggerty, J., Harvey, M. J., Klekociuk, A. R., and McDonald, A.: Observations of Clouds, Aerosols, Precipitation, and Surface Radiation over the Southern Ocean: An Overview of CAPRICORN, MARCUS, MICRE, and SOCRATES, Bull. Amer. Meteor. Soc., 102, E894-E928, https://doi.org/10.1175/bams-d-20-0132.1, 2021.

Mei, F.: Condensation particle counter for ACEENA [dataset], https://doi.org/10.5439/1440985, Accessed 5 December 2020.

Müller, D., Hostetler, C. A., Ferrare, R. A., Burton, S. P., Chemyakin, E., Kolgotin, A., Hair, J. W., Cook, A. L., Harper, D. B., Rogers, R. R., Hare, R. W., Cleckner, C. S., Obland, M. D., Tomlinson, J., Berg, L. K., and Schmid, B.: Airborne Multiwavelength High Spectral Resolution Lidar (HSRL-2) observations during TCAP 2012: vertical profiles of optical and microphysical properties of a smoke/urban haze plume over the northeastern coast of the US, Atmos. Meas. Tech., 7, 3487-3496, https://doi.org/10.5194/amt-7-3487-2014, 2014.

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Chin, M., Diehl, T., Easter, R. C., Feichter, J., Ghan, S. J., Hauglustaine, D., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J. F., Lin, G., Liu, X., Lund, M. T., Luo, G., Ma, X., van Noije, T., Penner, J. E., Rasch, P. J., Ruiz, A., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., Wang, P., Wang, Z., Xu, L., Yu, H., Yu, F., Yoon, J. H., Zhang, K., Zhang, H., and Zhou, C.: Radiative forcing of the direct

772 aerosol effect from AeroCom Phase II simulations, Atmos. Chem. Phys., 13, 1853-1877,
773 https://doi.org/10.5194/acp-13-1853-2013, 2013.

774 NETCDF: Introduction and Overview: https://www.unidata.ucar.edu/software/netcdf/docs/index.html, last
775 access: 2 November 2021. 2021.

776 Petters, M. D. and Kreidenweis, S. M.: A single parameter representation of hygroscopic growth and cloud
777 condensation nucleus activity, Atmos. Chem. Phys., 7, 1961-1971, https://doi.org/10.5194/acp-7-1961-
778 2007, 2007.

779 Pierce, J. R. and Adams, P. J.: Uncertainty in global CCN concentrations from uncertain aerosol nucleation
780 and primary emission rates, Atmos. Chem. Phys., 9, 1339-1356, https://doi.org/10.5194/acp-9-1339-2009,
781 2009.

782 Rasch, P. J., Xie, S., Ma, P.-L., Lin, W., Wang, H., Tang, Q., Burrows, S. M., Caldwell, P., Zhang, K.,
783 Easter, R. C., Cameron-Smith, P., Singh, B., Wan, H., Golaz, J.-C., Harrop, B. E., Roesler, E., Bacmeister,
784 J., Larson, V. E., Evans, K. J., Qian, Y., Taylor, M., Leung, L. R., Zhang, Y., Brent, L., Branstetter, M.,
785 Hannay, C., Mahajan, S., Mametjanov, A., Neale, R., Richter, J. H., Yoon, J.-H., Zender, C. S., Bader, D.,
786 Flanner, M., Foucar, J. G., Jacob, R., Keen, N., Klein, S. A., Liu, X., Salinger, A. G., Shrivastava, M., and
787 Yang, Y.: An Overview of the Atmospheric Component of the Energy Exascale Earth System Model, J.
788 Adv. Model. Earth Syst., 11, 2377-2411, https://doi.org/10.1029/2019ms001629, 2019.

789 Reddington, C. L., Carslaw, K. S., Stier, P., Schutgens, N., Coe, H., Liu, D., Allan, J., Browse, J., Pringle,
790 K. J., Lee, L. A., Yoshioka, M., Johnson, J. S., Regayre, L. A., Spracklen, D. V., Mann, G. W., Clarke, A.,
791 Hermann, M., Henning, S., Wex, H., Kristensen, T. B., Leaitch, W. R., Pöschl, U., Rose, D., Andreae, M.
792 O., Schmale, J., Kondo, Y., Oshima, N., Schwarz, J. P., Nenes, A., Anderson, B., Roberts, G. C., Snider, J.
793 R., Leck, C., Quinn, P. K., Chi, X., Ding, A., Jimenez, J. L., and Zhang, Q.: The Global Aerosol Synthesis
794 and Science Project (GASSP): Measurements and Modeling to Reduce Uncertainty, Bull. Amer. Meteor.
795 Soc., 98, 1857-1877, https://doi.org/10.1175/bams-d-15-00317.1, 2017.

796 Redemann, J., Wood, R., Zuidema, P., Doherty, S. J., Luna, B., LeBlanc, S. E., Diamond, M. S., Shinozuka,
797 Y., Chang, I. Y., Ueyama, R., Pfister, L., Ryoo, J. M., Dobracki, A. N., da Silva, A. M., Longo, K. M.,
798 Kacenelenbogen, M. S., Flynn, C. J., Pistone, K., Knox, N. M., Piketh, S. J., Haywood, J. M., Formenti, P.,
799 Mallet, M., Stier, P., Ackerman, A. S., Bauer, S. E., Fridlind, A. M., Carmichael, G. R., Saide, P. E., Ferrada,
800 G. A., Howell, S. G., Freitag, S., Cairns, B., Holben, B. N., Knobelspiesse, K. D., Tanelli, S., L'Ecuyer, T.
801 S., Dzambo, A. M., Sy, O. O., McFarquhar, G. M., Poellot, M. R., Gupta, S., O'Brien, J. R., Nenes, A.,
802 Kacarab, M., Wong, J. P. S., Small-Griswold, J. D., Thornhill, K. L., Noone, D., Podolske, J. R., Schmidt,
803 K. S., Pilewskie, P., Chen, H., Cochrane, S. P., Sedlacek, A. J., Lang, T. J., Stith, E., Segal-Rozenhaimer,
804 M., Ferrare, R. A., Burton, S. P., Hostetler, C. A., Diner, D. J., Seidel, F. C., Platnick, S. E., Myers, J. S.,
805 Meyer, K. G., Spangenberg, D. A., Maring, H., and Gao, L.: An overview of the ORACLES (ObseRvations
806 of Aerosols above CLouds and their intEractionS) project: aerosol–cloud–radiation interactions in the
807 southeast Atlantic basin, Atmos. Chem. Phys., 21, 1507-1563, https://doi.org/10.5194/acp-21-1507-2021,
808 2021.

809 Sanchez, K. J., Roberts, G. C., Saliba, G., Russell, L. M., Twohy, C., Reeves, J. M., Humphries, R. S.,
810 Keywood, M. D., Ward, J. P., and McRobert, I. M.: Measurement report: Cloud processes and the transport
811 of biological emissions affect southern ocean particle and cloud condensation nuclei concentrations, Atmos.
812 Chem. Phys., 21, 3427-3446, https://doi.org/10.5194/acp-21-3427-2021, 2021.

813 Schulz, M., Textor, C., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O.,
814 Dentener, F., Guibert, S., Isaksen, I. S. A., Iversen, T., Koch, D., Kirkevåg, A., Liu, X., Montanaro, V.,

Myhre, G., Penner, J. E., Pitari, G., Reddy, S., Seland, Ø., Stier, P., and Takemura, T.: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations, Atmos. Chem. Phys., 6, 5225-5246, https://doi.org/10.5194/acp-6-5225-2006, 2006.

Seinfeld, J. H., Bretherton, C., Carslaw, K. S., Coe, H., DeMott, P. J., Dunlea, E. J., Feingold, G., Ghan, S., Guenther, A. B., Kahn, R., Kraucunas, I., Kreidenweis, S. M., Molina, M. J., Nenes, A., Penner, J. E., Prather, K. A., Ramanathan, V., Ramaswamy, V., Rasch, P. J., Ravishankara, A. R., Rosenfeld, D., Stephens, G., and Wood, R.: Improving our fundamental understanding of the role of aerosol−cloud interactions in the climate system, Proc. Natl. Acad. Sci. U.S.A, 113, 5781-5790, https://doi.org/10.1073/pnas.1514043113, 2016.

Sun, J., Zhang, K., Wan, H., Ma, P.-L., Tang, Q., and Zhang, S.: Impact of Nudging Strategy on the Climate Representativeness and Hindcast Skill of Constrained EAMv1 Simulations, J. Adv. Model. Earth Syst., 11, 3911-3933, https://doi.org/10.1029/2019MS001831, 2019.

Tang, Q., Klein, S. A., Xie, S., Lin, W., Golaz, J. C., Roesler, E. L., Taylor, M. A., Rasch, P. J., Bader, D. C., Berg, L. K., Caldwell, P., Giangrande, S. E., Neale, R. B., Qian, Y., Riihimaki, L. D., Zender, C. S., Zhang, Y., and Zheng, X.: Regionally refined test bed in E3SM atmosphere model version 1 (EAMv1) and applications for high-resolution modeling, Geosci. Model Dev., 12, 2679-2706, https://doi.org/10.5194/gmd-12-2679-2019, 2019.

Teixeira, J., Cardoso, S., Bonazzola, M., Cole, J., DelGenio, A., DeMott, C., Franklin, C., Hannay, C., Jakob, C., Jiao, Y., Karlsson, J., Kitagawa, H., Köhler, M., Kuwano-Yoshida, A., LeDrian, C., Li, J., Lock, A., Miller, M. J., Marquet, P., Martins, J., Mechoso, C. R., Meijgaard, E. v., Meinke, I., Miranda, P. M. A., Mironov, D., Neggers, R., Pan, H. L., Randall, D. A., Rasch, P. J., Rockel, B., Rossow, W. B., Ritter, B., Siebesma, A. P., Soares, P. M. M., Turk, F. J., Vaillancourt, P. A., Von Engeln, A., and Zhao, M.: Tropical and Subtropical Cloud Transitions in Weather and Climate Prediction Models: The GCSS/WGNE Pacific Cross-Section Intercomparison (GPCI), J. Climate, 24, 5223-5256, https://doi.org/10.1175/2011jcli3672.1, 2011.

Twomey, S.: The Influence of Pollution on the Shortwave Albedo of Clouds, J. Atmos. Sci., 34, 1149-1152, https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2, 1977.

Uin, J., Senum, G., Koontz, A., and Flynn, C.: Ultra-High Sensitivity Aerosol Spectrometer (AOSUHSAS) [dataset], https://doi.org/10.5439/1409033, Accessed 24 November 2020.

Wang, H., Easter, R. C., Zhang, R., Ma, P.-L., Singh, B., Zhang, K., Ganguly, D., Rasch, P. J., Burrows, S. M., Ghan, S. J., Lou, S., Qian, Y., Yang, Y., Feng, Y., Flanner, M., Leung, R. L., Liu, X., Shrivastava, M., Sun, J., Tang, Q., Xie, S., and Yoon, J.-H.: Aerosols in the E3SM Version 1: New Developments and Their Impacts on Radiative Forcing, J. Adv. Model. Earth Syst., 12, e2019MS001851, https://doi.org/10.1029/2019ms001851, 2020.

Wang, J., Wood, R., Jensen, M. P., Chiu, J. C., Liu, Y., Lamer, K., Desai, N., Giangrande, S. E., Knopf, D. A., Kollias, P., Laskin, A., Liu, X., Lu, C., Mechem, D., Mei, F., Starzec, M., Tomlinson, J., Wang, Y., Yum, S. S., Zheng, G., Aiken, A. C., Azevedo, E. B., Blanchard, Y., China, S., Dong, X., Gallo, F., Gao, S., Ghate, V. P., Glienke, S., Goldberger, L., Hardin, J. C., Kuang, C., Luke, E. P., Matthews, A. A., Miller, M. A., Moffet, R., Pekour, M., Schmid, B., Sedlacek, A. J., Shaw, R. A., Shilling, J. E., Sullivan, A., Suski, K., Veghte, D. P., Weber, R., Wyant, M., Yeom, J., Zawadowicz, M., and Zhang, Z.: Aerosol and Cloud Experiments in the Eastern North Atlantic (ACE-ENA), Bull. Amer. Meteor. Soc., 1-51, https://doi.org/10.1175/bams-d-19-0220.1, 2021.

857 Watson-Parris, D., Schutgens, N., Reddington, C., Pringle, K. J., Liu, D., Allan, J. D., Coe, H., Carslaw, K.
858 S., and Stier, P.: In situ constraints on the vertical distribution of global aerosol, Atmos. Chem. Phys., 19,
859 11765-11790, https://doi.org/10.5194/acp-19-11765-2019, 2019.

860 Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.:
861 Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, J. Atmos. Ocean. Technol.,
862 26, 2310-2323, https://doi.org/10.1175/2009jtecha1281.1, 2009.

863 Zaveri, R. A., Easter, R. C., Singh, B., Wang, H., Lu, Z., Tilmes, S., Emmons, L. K., Vitt, F., Zhang, R.,
864 Liu, X., Ghan, S. J., and Rasch, P. J.: Development and Evaluation of Chemistry-Aerosol-Climate Model
865 CAM5-Chem-MAM7-MOSAIC: Global Atmospheric Distribution and Radiative Effects of Nitrate
866 Aerosol, J. Adv. Model. Earth Syst., 13, e2020MS002346, https://doi.org/10.1029/2020MS002346, 2021.

867 Zawadowicz, M. A., Suski, K., Liu, J., Pekour, M., Fast, J., Mei, F., Sedlacek, A. J., Springston, S., Wang,
868 Y., Zaveri, R. A., Wood, R., Wang, J., and Shilling, J. E.: Aircraft measurements of aerosol and trace gas
869 chemistry in the eastern North Atlantic, Atmos. Chem. Phys., 21, 7983-8002, https://doi.org/10.5194/acp-
870 21-7983-2021, 2021.

871 Zhang, C., Xie, S., Tao, C., Tang, S., Emmenegger, T., Neelin, J. D., Schiro, K. A., Lin, W., and Shaheen,
872 Z.: The ARM Data-Oriented Metrics and Diagnostics Package for Climate Models: A New Tool for
873 Evaluating Climate Models with Field Data, Bull. Amer. Meteor. Soc., 101, E1619-E1627,
874 https://doi.org/10.1175/bams-d-19-0282.1, 2020.

875 Zhang, K., Wan, H., Liu, X., Ghan, S. J., Kooperman, G. J., Ma, P. L., Rasch, P. J., Neubauer, D., and
876 Lohmann, U.: Technical Note: On the use of nudging for aerosol–climate model intercomparison studies,
877 Atmos. Chem. Phys., 14, 8631-8645, https://doi.org/10.5194/acp-14-8631-2014, 2014.

878 Zhao, B., Shrivastava, M., Donahue, N. M., Gordon, H., Schervish, M., Shilling, J. E., Zaveri, R. A., Wang,
879 J., Andreae, M. O., Zhao, C., Gaudet, B., Liu, Y., Fan, J., and Fast, J. D.: High concentration of ultrafine
880 particles in the Amazon free troposphere produced by organic new particle formation, Proc. Natl. Acad.
881 Sci. U. S. A., 117, 25344-25351, https://doi.org/10.1073/pnas.2006716117, 2020.

882 Zheng, G., Wang, Y., Wood, R., Jensen, M. P., Kuang, C., McCoy, I. L., Matthews, A., Mei, F., Tomlinson,
883 J. M., Shilling, J. E., Zawadowicz, M. A., Crosbie, E., Moore, R., Ziemba, L., Andreae, M. O., and Wang,
884 J.: New particle formation in the remote marine boundary layer, Nat. Commun., 12, 527,
885 https://doi.org/10.1038/s41467-020-20773-1, 2021.

886 Zhou, X., Kollias, P., and Lewis, E. R.: Clouds, Precipitation, and Marine Boundary Layer Structure during
887 the MAGIC Field Campaign, J. Climate, 28, 2420-2442, https://doi.org/10.1175/jcli-d-14-00320.1, 2015.

888 Zuidema, P., Alvarado, M., Chiu, C., de Szoeke, S., Fairall, C., Feingold, G., Freedman, A., Ghan, S.,
889 Haywood, J., Kollias, P., Lewis, E., McFarquhar, G., McComiskey, A., Mechem, D., Onasch, T.,
890 Redemann, J., Romps, D., Turner, D., Wang, H., Wood, R., Yuter, S., and Zhu, P.: Layered Atlantic Smoke
891 Interactions with Clouds (LASIC) Field Campaign Report, U.S. Department of Energy, Technical Report,
892 DOE/SC-ARM-18-018, 2018.

893