

ESM Aerosol-Cloud Diagnostics Package (ESMAC Diags)

Version 2.1

Shuaiqi Tang

Adam C. Varble

Jerome D. Fast

Joseph C. Hardin

Po-Lun Ma

Contact: Shuaiqi Tang (Shuaiqi.tang@pnnl.gov)

This diagnostics package is 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.

Abstract

This document describes the **version 2.1** of Earth System Model (ESM) aerosol-cloud diagnostics package (**ESMAC Diags**) that facilitate routine evaluation of aerosols, clouds and aerosol-cloud interactions simulated by the Department of Energy's (DOE) Energy Exascale Earth System Model (E3SM). This package focuses on comparing simulated **aerosols, clouds and aerosol-cloud interactions** with in-situ and remote-sensing measurements from surface, aircraft, ship and satellite platforms. Various types of diagnostics and metrics are performed to assess how well E3SM represents observed aerosol properties and aerosol-cloud interactions across spatial scales. This document describes the evaluation datasets, structure of the diagnostics package, as well as the instruction of how to run or modify this package. The diagnostics package is coded and organized in a way that can be easily adapted to other model simulations and are flexible to add more field campaigns in the future.

Content

- Introduction
- Observation data
 - Field campaign and measurements
 - Long-term measurements at ARM SGP and ENA sites
- Preparation of observation data
 - Data quality controls
 - Different data sources or retrieval algorithms
- E3SM simulation and preparation
- ESMAC Diags structure
- Installation and execution of ESMAC Diags

1. Introduction

This document describes the Earth System Model (ESM) aerosol-cloud diagnostics package (ESMAC Diags) that facilitates routine evaluation of aerosols, clouds and aerosol-cloud interactions simulated by the Department of Energy's (DOE) Energy Exascale Earth System Model (E3SM). Recent field campaign measurements from surface, aircraft, ship and satellite platforms from the U.S. DOE Atmospheric Radiation Measurements (ARM) and National Science Foundation (NSF) NCAR user facilities and long-term routine measurements at ARM fixed sites are utilized to evaluate ESM performance and facilitate new parameterization development.

2. Observation data

The diagnostics package initially focuses 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). Six field campaigns from these four testbeds are selected in the initial phase of ESMAC Diags. Besides these month-long field campaign data, multi-year long-term measurements at two ARM sites: SGP and ENA, are also included. The location and platforms of the six field campaigns and two ARM sites are shown in Figure 1. This section introduces the measurements used in ESMAC Diags.

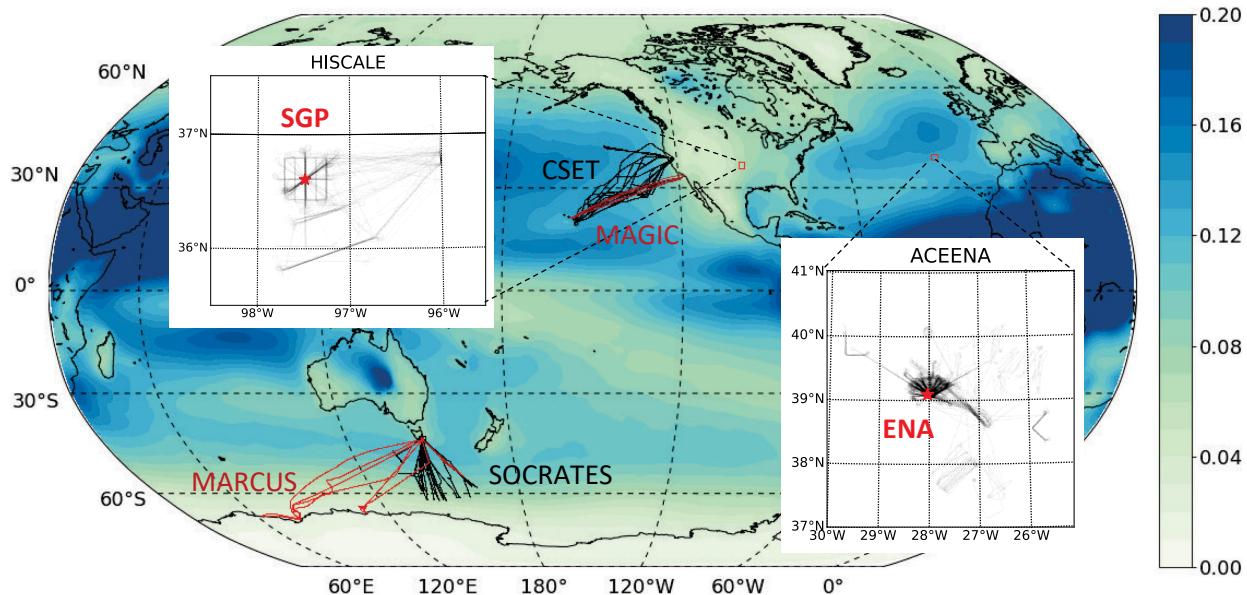


Figure 1. Aircraft (black) and ship (red) tracks for the six field campaigns. Two ARM fixed sites: SGP and ENA, are shown as red stars in the enlarged map. Overlaid is aerosol optical depth at 550nm averaged from 2014 to 2018 simulated in EAMv1.

2.1 Field campaign measurements

Information of the six field campaigns is shown in Table 1. HI-SCALE and ACE-ENA are based on long-term ARM ground sites with aircraft field campaigns sampling below, within, and above convective and marine boundary layer clouds, respectively, within a few hundred kilometers around the sites. CSET and MAGIC are field campaigns with aircraft and ship platforms, respectively, sampling transects between California and Hawaii characterized by a transition between stratocumulus and trade cumulus dominated regions. SOCRATES and MARCUS are field campaigns with aircraft and ship platforms, respectively, based out of Hobart, Australia. Aircraft transects during SOCRATES extended south to around 60°S, while ship transects during MARCUS extended southwest from Hobart to Antarctica. The aircraft (black) and ship (red) tracks for these field campaigns are shown in Figure 1.

Table 1. Descriptions of the field campaigns used in this study. Numbers after aircraft or ship 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 flights, IOP2: 21 flights)	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 flights, IOP2: 19 flights)	Marine stratocumulus with low aerosol loading	(Wang et al., 2021)
MAGIC	Oct 2012 – Sep 2013	Ship (18 legs)	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 flights)	Same as above	(Albrecht et al., 2019)
MARCUS	Oct 2017 – Apr 2018	Ship (4 legs)	Marine liquid and mixed phase clouds with low aerosol loading	(McFarquhar et al., 2021)
SOCRATES	15 Jan – 24 Feb, 2018	Aircraft (14 flights)	Same as above	(McFarquhar et al., 2021)

* Full names of the listed field campaigns:

HI-SCALE: Holistic Interactions of Shallow Clouds, Aerosols and Land Ecosystems

ACE-ENA: Aerosol and Cloud Experiments in the Eastern North Atlantic

MAGIC: Marine ARM GCSS Pacific Cross-section Intercomparison (GPCI) Investigation of Clouds

CSET: Cloud System Evolution in the Trades

MARCUS: Measurements of Aerosols, Radiation and Clouds over the Southern Ocean

SOCRATES: Southern Ocean Cloud Radiation and Aerosol Transport Experimental Study

The instruments and measurements used in ESMAC Diags are listed in Tables S1 to S6 at the end of this document. Each table corresponds to each field campaigns. Note that some instruments are only available for certain periods, so that model evaluation is limited by the availability of collected data.

2.2 long-term datasets at ARM SGP and ENA sites

ARM has long-term ground-based aerosol and cloud measurements at its permanent sites at the Southern Great Plains (SGP) and Eastern North Atlantic (ENA), which allows evaluating aerosols, clouds and aerosol-cloud interactions in a statistical point of view. It also archives VISST products from geostationary satellite, provided by NASA/Langley, within a few degrees near the ARM sites. To allow maximum availability of key measurements at the same time of ensuring long enough period for statistical evaluation, we select the periods of 1 Jan 2011 – 31 Dec 2020 for SGP and 23 Jun 2016 – 30 Jul 2018 for ENA for analysis. Tables S7 and S8 list the information of all aerosol, cloud, radiation and meteorological datasets at SGP and ENA, respectively.

3. Preparation of observation data

All the observational data are quality controlled with their time resolution re-scaled to that suitable for evaluating E3SM. Currently, ground, ship and satellite measurements are re-scaled to 1-hour frequency which is approximately consistent with 1-degree resolution E3SM output. As E3SM can run at higher resolution and may soon run with kilometer scale grid spacing, the resolution of observation data can also be modified using rescaling. Rescaling consists of computing either the median, mean or interpolated value depending on the original data frequency and variable properties. For most aerosol and cloud microphysics measurements, the median value is computed to remove occasional spikes or zeros resulting from data contamination or measurement error. For some bulk cloud properties (e.g., cloud fraction, liquid water path (LWP)), the mean value is computed to be consistent with grid-mean E3SM output. Interpolation is only used when the input frequency is equal to or coarser than the frequency of model output. For aircraft measurements, 1-minute resolution is used to retain high variability and allow matching samples of aerosol and cloud at the same time. To compare with high-frequency aircraft data, E3SM output is down-scaled to 1-minute resolution using the nearest grid cell and time slice. All processed data are saved in a standardized NETCDF format (NETCDF, 2022).

3.1 data quality issues and treatments

Many observation data used in ESMAC Diags are ARM level-b (quality-controlled) and level-c (value-added) products, which include quality flags to indicate data quality issues. For most datasets, a quality-control (QC) treatment is applied to remove all data with questionable flags. However, there are certain datasets or circumstances in which a quality flag is overly strict and removes too much good data or not strict enough such that some bad data are not removed. Here we document some of these situations and how we handle them in our data processing.

3.1.1 ARM CPC measurements

ARM CPCs data have several QC bit values representing failure of different quality checks. One of them checks if the concentration is greater than a maximum allowable value, which is $8,000 \text{ cm}^{-3}$ for CPC (model 3010, size detection limit 10 nm), $10,000 \text{ cm}^{-3}$ for CPCF (model 3772, size detection limit 10 nm) and $50,000 \text{ cm}^{-3}$ for CPCU (model 3776, size detection limit 3 nm). At SGP, new particle formation (NPF) events occur on many days, during which CPC and CPCF measurements can be up to $30,000 \text{ cm}^{-3}$. This is much higher than their maximum allowable value but physically reasonable. Simply removing these large values results in an underestimation of aerosol number concentration and exhibits an unrealistic diurnal cycle as they usually occur in the daytime (Tang et al., 2022). By consulting the ARM instrument mentor Ashish Singh <asingh@bnl.gov>, we remove data with a few critical QC bit values, while keep data with other QC bit values that can be overly restrictive. Only the following qc_bit is applied:

- CPC and CPCF – bit value 7, 8, 11, 12, 13, 14
- CPCu and CPCuf – bit value 4, 5, 6, 15, 16

3.1.2 NCAR research flight aerosol number concentration (CN) measurements

NCAR research flight (RF) data used in ESMAC Diags do not include quality control flags but occasionally show suspiciously large or negative aerosol counts. The following minimum and maximum thresholds are applied to remove suspicious data:

- total CN from CNC (CONCN): minimum = 0, maximum = $25,000 \text{ cm}^{-3}$.
- total CN from UHSAS (UHSAS100): minimum = 0, maximum = $5,000 \text{ cm}^{-3}$.

aerosol number size distribution from UHSAS (CUHSAS_RWOOU or CUHSAS_LWII): minimum = 0, maximum = 500 cm^{-3} per size bin.

3.1.3 Ship-measured aerosol properties

Aerosol instruments on ships are occasionally contaminated by ship emissions, which present as large spikes in aerosol and CCN number concentrations. For MARCUS, Humphries (2020) published reprocessed CN and CCN data from ARM MARCUS measurements to remove ship exhaust contamination using method described in Humphries et al. (2019). This data is used in this diagnostics package. For MAGIC, we could not find any ship exhaust contamination information. Therefore, a simple maximum threshold ($25,000 \text{ cm}^{-3}$ for CPC, $5,000 \text{ cm}^{-3}$ for UHSAS100, $2,000 \text{ cm}^{-3}$ for CCN at 0.1% supersaturation and $4,000 \text{ cm}^{-3}$ for CCN at 0.5% supersaturation) is applied to remove likely contamination from ship emissions.

3.1.4 CCN

There are a few different supersaturation (SS) setting strategies for CCN measurements. Some aircraft campaigns measured CCN in constant SS (ACE-ENA, HI-SCALE). Some other campaigns measured CCN in scanning SS (SOCRATES, surface CCN counters at SGP and ENA). However, the actual supersaturation in a scanning SS strategy has fluctuations that are different than the target supersaturation. For the latter, CCN for each SS (0.1%, 0.2%, 0.3% and 0.5%) are obtained by selecting CCN measured within +/- 0.05% SS range.

For long-term measurements at SGP and ENA, near-hourly CCN spectra data are available and a quadratic polynomial is fit to the spectra such that CCN number concentration can be estimated at any SS between the measured minimum and maximum SS values:

$$\text{CCN} = \text{coefs}[0] + \text{coefs}[1]*\text{SS} + \text{coefs}[2]*\text{SS}^2$$

We calculate and output CCN number concentration at three target supersaturations (0.1%, 0.2% and 0.5%). The fitted spectra data provides CCN number concentration at the exact target supersaturations, but the sample number is slightly smaller due to occasional failure of polynomial fitting.

3.1.5 Contamination in surface aerosol measurements at ENA

The ARM ENA site is located at a local airport. Aerosol measurements at ENA are sometimes contaminated by aircraft and vehicle emissions, rendering the measurements not representative of the background environment. Gallo et al. (2020) identified periods when CPC measurements were likely contaminated from localized emissions (Figure 2a). Their aerosol mask data has 1-min resolution. When we rescale the data to 1-hr resolution and apply the mask on other coarse-resolution aerosol measurements (e.g., ACSM, Figure 2c), we mask hours in which more than half of the hour is flagged by the aerosol mask. A sensitivity analysis was performed, showing that 50% is a reasonable threshold to balance removal of contamination with keeping reasonable data (not shown).

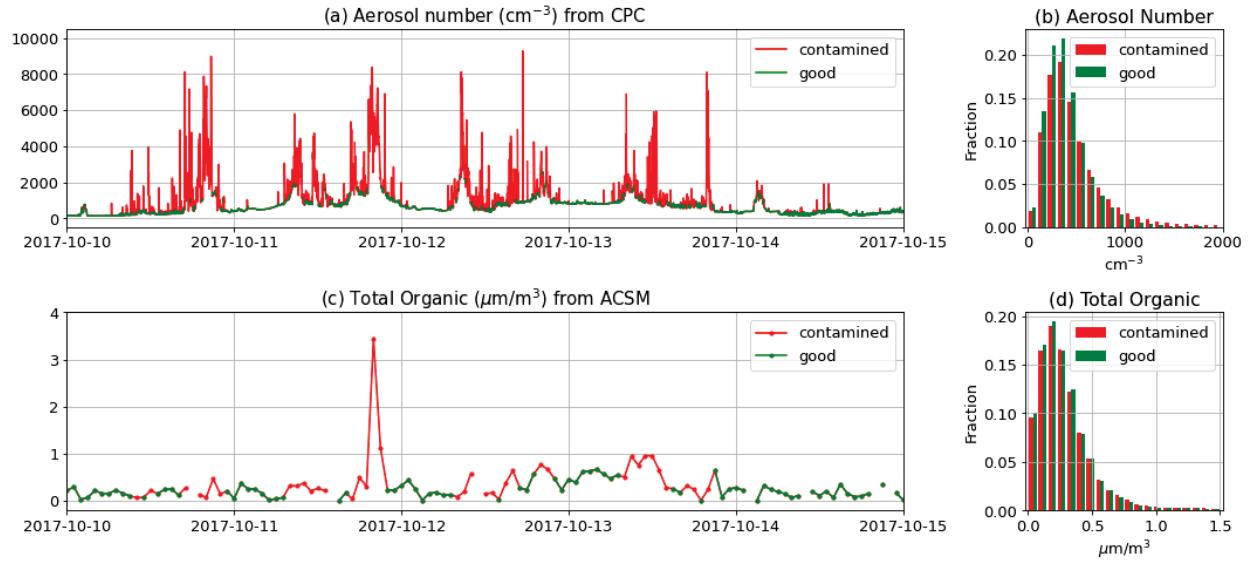


Figure 2. (a) CPC-measured CN from 10 to 15 October 2017 (1-minute resolution) with local contamination flagged by Gallo et al. (2020). (b) histogram of CPC-measured CN for all data from 2016-2018. (c) ACSM measured total organic matter from 10 to 15 October 2017 (1-hour resolution). Hours with more than half or the hour flagged in 1-minute CPC data are masked as contaminated. (d) histogram of ACSM-measured total organic matter for all data from 2016-2018.

3.1.6 SMPS and nanoSMPS for HI-SCALE

Surface SMPS and nanoSMPS are used in HI-SCALE to measure aerosol number size distribution. However, nanoSMPS systematically overcounts aerosol numbers and the reason is still under exploration. We divided the nanoSMPS measurements by 3.8 to ensure a smooth transition between nanoSMPS and SMPS when merging aerosol size distribution from these two instruments. Note that the factor of 3.8 is not suitable for longer period at SGP. A better calibration (e.g., a time-dependent rescaling factor) is needed to combine long-term SMPS and nanoSMPS at SGP.

3.1.7 Solar zenith angle for satellite retrievals

Satellite retrievals are known to have large biases when the solar zenith angle is large (Grosvenor et al., 2018). The solar zenith angle data is outputted but not pre-filtered. Users can read the solar zenith angle data and apply filter to their diagnostics.

3.2 variables from different data source and potential uncertainties

Some variables are obtained from different data sources, or retrieved using different algorithms. The differences of the same variable among different data sources (using different algorithms) indicate the reliability and uncertainty of this variable in observation. This section lists these variables from different sources (algorithms) and briefly discuss their potential uncertainties.

3.2.1 Liquid water path

Observed liquid water path (LWP) are obtained in mfrsrclod, armbeclrad and VISST data. Mfrsrclod and armbeclrad both use surface microwave radiometer (MWR) measurements. However, there are several algorithms to calculate LWP from raw MWR measurements resulting several different products, and the two datastreams have different strategies of choosing which product is used. Mfrsrclod chooses LWP datasets from mwrret1liljclou.c2, mwrret1liljclou.c1, mwrret2turn.c1, and mwrlos.b1, with decreasing priority. Armbeclrad uses mwrret1liljclou.s2 at SGP and mwr3c.b1 at ENA. This difference may result in slightly different diurnal cycle and distribution of LWP (Figure 3).

VISST data use MWR measurements on geostationary satellites. The specific satellite depends on the location and time period. For example, GOES satellite measurements are used around SGP region, with GOES-13 is used during 2011 to 2017 and GOES-16 is used during 2018 to 2020. Around ENA region, METEOSAT satellite measurements are used.

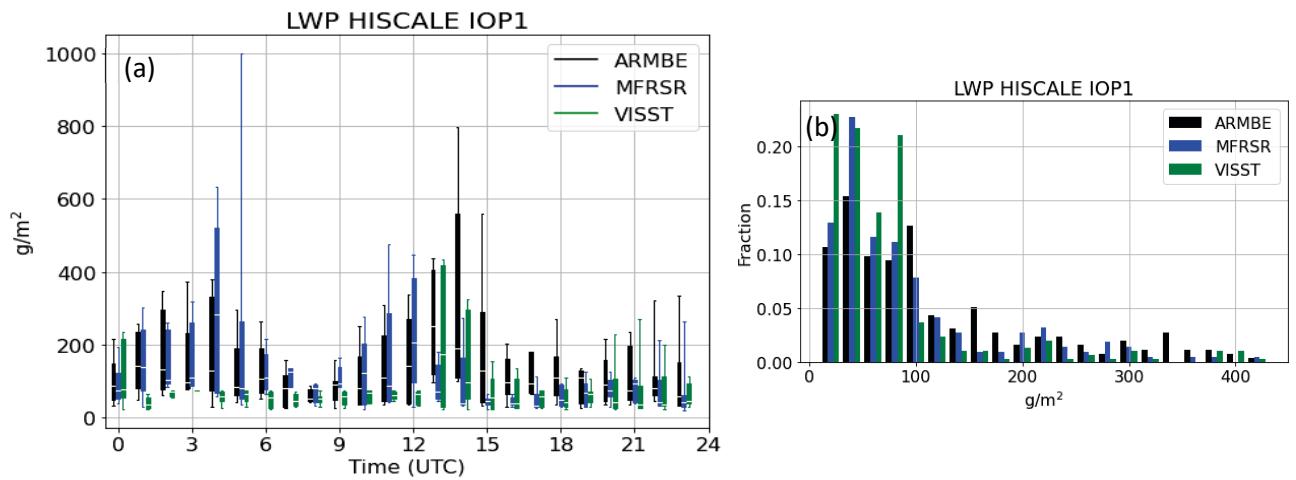


Figure 3. (a) diurnal cycle and (b) histogram of LWP obtained from different datasets for HI-SCALE IOP1 (2016-04-25 to 2016-05-22).

3.2.2 Cloud microphysics retrievals

Cloud microphysical properties (e.g., droplet number concentration N_d , effective radius R_{eff}) are important variables that connect clouds to other aspects in the climate system such as aerosols and radiation. Except in field campaigns where in-situ aircraft measurements are available, remote sensing retrieval algorithms have to be used to derive these quantities. Several cloud retrieval products from ground and satellite measurements with different algorithms are used in ESMAC Diags. This section compares these cloud retrievals with in-situ aircraft measurements to assess retrieval limitations, uncertainties, and utility.

Table 2 lists R_{eff} and N_d retrieval products used in ESMAC Diags. We retrieved Nd_sat with input data from VISST products using the algorithms described in Bennartz (2007). Other datasets are all available as released products. All retrievals assume a homogeneous single-layer liquid phase cloud. However, retrieval algorithms are usually run for all conditions whenever they return valid values. When assumptions are not satisfied, retrieved properties will contain large errors and likely alter statistics such as increasing the occurrence frequency of small N_d as will be shown next.

Table 2: Cloud droplet effective radius R_{eff} and number concentration N_d retrievals

Variable	Dataset	Platform	Campaign/site	Retrieved from	Reference
R_{eff}	MFRSRCLDOD	Ground	HI-SCALE, ACE-ENA, SGP, ENA	SW diffuse flux, LWP	(Min and Harrison, 1996; Turner et al., 2021)
	VISST	satellite	HI-SCALE, ACE-ENA, MAGIC, MARCUS, SGP, ENA	Brightness temperature	(Minnis et al., 2011)
	Wu_etal	Ground	ACE-ENA, MAGIC, ENA	Radar reflectivity, LWP	(Wu et al., 2020)
N_d	Ndrop	Ground	HI-SCALE, ACE-ENA, SGP, ENA	LWP, COD, cloud height	(Riihimaki et al., 2021; Lim et al., 2016)
	Nd_sat (calculated from VISST)	satellite	HI-SCALE, ACE-ENA, MAGIC, MARCUS, SGP, ENA	LWP, COD, CTT	(Bennartz, 2007)
	Wu_etal	Ground	ACE-ENA, MAGIC, ENA	Radar reflectivity, LWP	(Wu et al., 2020)

MFRSRCLDOD: Cloud Optical Properties from the MultiFilter Shadowband Radiometer (MFRSR)

SW: shortwave

COD: cloud optical depth

CTT: cloud top temperature

Note that these retrievals all assume single-layer overcast low-level liquid clouds. The algorithms will have large errors or fail for broken shallow cumulus, multi-layer clouds or ice clouds. Cloud type filter is needed to ensure reasonable physical meaning and reliable evaluation of model results.

4. E3SM simulation and preparation

This diagnostic package is currently used to evaluate E3SM model, so this section shows how to get required E3SM output and how to prepare them. For other ESMs, users can create their own pre-processed model files similar to the model file format.

4.1 E3SM simulation

High-frequency E3SM output of aerosol and cloud variables over the field campaign regions are needed for evaluation. An example of namelist in `user_nl_cam` in the E3SM running script is given in Appendix A, which includes hourly output variables and region domains. Detailed information on how to run E3SM can be found at <https://github.com/E3SM-Project/E3SM>. Typically, model will cover the same period of the observational data (with a few months of spin-up prior) and be nudged to reanalysis to ensure direct comparison with observations. In a successful run, hourly variables for each field campaign domain should be seen in the E3SM output file `*.eam.h1.yyyy-mm-dd-00000.nc`. (e.g., “`PS_260e_to_265e_34n_to_39n`” for PS at SGP region).

4.2 E3SM data preparation

The model output variables needed to be evaluated are extracted or calculated (if they are not directly output) at the surface site, along the flight tracks or along the ship tracks depending on what measurements are available for each field campaign. For field campaigns with surface or ship measurements, E3SM output are interpolated into designed frequency. The default output preparation output frequency is 1 hour, same as the current standard E3SM output frequency, so the data should not be impacted. For aircraft campaigns, E3SM output are rescaled to higher frequency aircraft measurements (1-min), by saving data at the nearest grid, level, and time of the aircraft sample. Note that although E3SM are rescaled to higher frequency at the exact height of the aircraft, it does not mean that the model data can resolve the higher temporal and spatial variability that aircraft measures.

Aerosol and cloud droplet number size distributions are calculated based on the spectrum information output from E3SM’s aerosol and cloud microphysics schemes: MAM4 and MG2 respectively. Aerosol size distribution is calculated from 1 to 3001 nm in 1nm increment, while cloud droplet size distribution is calculated from 1 to 1000 μm in 1 μm increment. The size distributions are mainly evaluated respect to aircraft measurements. They are also used to calculate particle number concentration to compare with bulk particle counters such as CPC.

Cloud optical depth (`TOT_CLD_VISTAU`) and effective radius (`REL`) are output in each vertical level. To compare with surface and satellite retrievals, cloud optical depth are integrated in the entire atmospheric column while effective radius are averaged weighted by cloud fraction. Layer-mean cloud droplet number concentration (`Nd`) is also calculated by dividing column-integrated droplet number concentration (`CDNUMC`) with vertically weighted cloud fraction. Again, evaluation of `Nd` with surface and satellite retrievals should be limited to single-layer liquid clouds that retrievals are more accurate.

The steps of how to prepare E3SM output is given in Section 6.

5. ESMAC Diags structure

The directory structure of ESMAC Diags is shown in Figure 4. The structure has fundamental change from ESMAC Diags version 1. In version 1 all observation data are read from their original format. All quality controls and resolution change are applied along with diagnostic process. In version 2, all observation data are preprocessed and output to a common NETCDF format for more convenient usage. The “scripts” directory contains executable scripts. The “testcase” directory provides a small amount of data and verify figures that can be used to verify the package is successfully installed and code is working correctly. All observational and model data are stored in the “raw_data” directory, while after preprocessing, the preprocessed data are output in the “prep_data” directory, both are organized by field campaigns. The “src” directory contains all source code including code used to read files, control data qualities, rescale time resolution, calculate required variables, preprocess model and observation data, make diagnostics plots and calculate statistics. The diagnostic plots and statistics are put in the “figures” directory, also organized by field campaign.

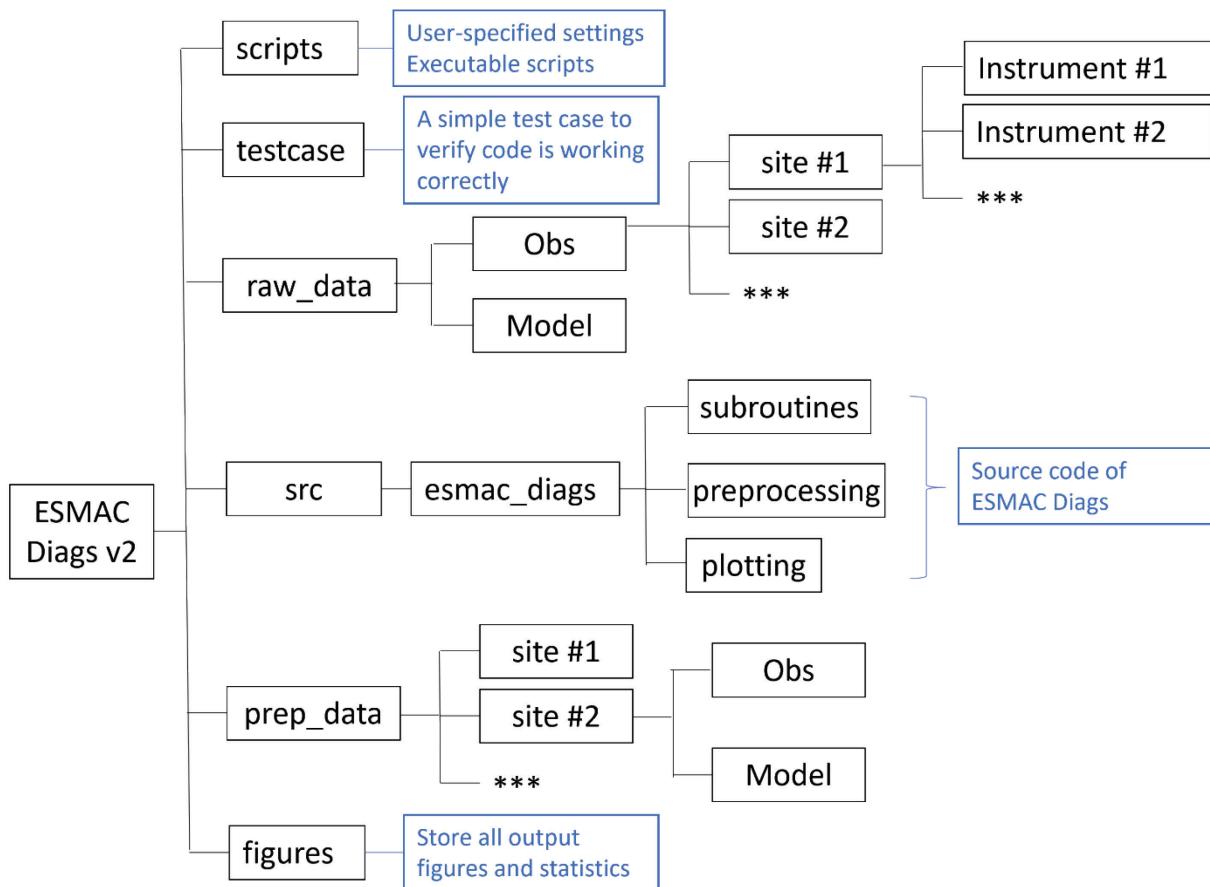


Figure 4. directory structure of ESMAC Diags version 2.

6. Installation and execution

6.1 Installation

This code is best run using a conda virtual environment. To install the required environment one can do

```
conda env create -f environment.yml
```

to set up an esmac_diags environment. Note if running this on a HPC system (e.g., DOE's NERSC supercomputer), you may need to load the appropriate module for anaconda.

Once the environment has been created you can activate it with

```
conda activate esmac_diags
```

and then this code can be installed with

```
pip install -e .
```

Which will install the code as editable allowing you to make changes to the codebase and it be reflected in the installed package.

6.2 Testrun

To verify the package, enter scripts/ directory and run

```
python run_testcase.py
```

Then go to the directory in testcase/figures/. Compare the output figure with the plot testcase/figures_verify/timeseries_organic_HISCALE.png. If the two figures look the same, the testcase is successfully run.

6.3 Preprocess model data

- a. Run your model and collect output data. See Section 4 for more information on E3SM run.
- b. Enter *scripts/* directory
- c. Edit the preprocessing script you want to run (take prep_HISCALE_E3SM.py as example)

Edit data path where E3SM output is stored and data path of aircraft location files (IWG path) if you want to extract E3SM output along aircraft tracks. Edit input_filehead (filename before .cam.h* for E3SMv1 or .eam.h* for E3SMv2) and output_filehead.

- d. Run the script:
 - i. (option 1) Interactive run:
Active ESMAC Diags:

```
conda activate esmac_diags
```

Run the script on a terminal:

```
python prep_HISCALE_E3SM.py
```

- ii. (option 2) Submit job on U.S. DOE's NERSC supercomputer:

Active ESMAC Diags:

```
conda activate esmac_diags
```

Edit the job submission parameters in scripts_jobsubmit.csh and submit it by:

```
sbatch scripts_jobsubmit.csh
```

6.4 Download observation data

Download the prepared observation data from Zenodo: <https://doi.org/10.5281/zenodo.7478657>. These data can be directly used by the diagnostics package. If users have their own measurements and would like to add into the diagnostics package, they can revise the data preprocessing code (at `src/esmac_diags/preprocessing/`) and apply to their data. The Zenodo data bundle also includes prepared model data for E3SMv1 and E3SMv2.

6.5 run diagnostics

We offer a set of example scripts for customized diagnostics that users can choose evaluation data, apply additional data treatments, and specify plotting parameters. These example scripts are labeled as “`example_*.py`” in scripts/ directory.

The steps to run the designed diagnostics are similar as running preprocessing code described in the previous section. We also provide the scripts for full single-variable diagnostics (i.e., timeseries, diurnal cycle, mean statistics, histograms, vertical or horizontal variations; excluding multi-variable relations such as scatter plot, joint histograms and heatmaps) for each field campaign. See `plot_HISCALE_flight.py` and `plot_HISCALE_sfc_toa.py` as examples.

7. Update history

V1.0.0-alpha (November 28, 2021): first release of ESMAC diags for the initial submission of GMD paper: “*Earth System Model Aerosol-Cloud Diagnostics Package (ESMAC Diags) Version 1: Assessing E3SM Aerosol Predictions Using Aircraft, Ship, and Surface Measurements*”
<https://doi.org/10.5194/gmd-2021-350>.

V1.0.0-beta (March 18, 2022): updated for the revision submission of GMD paper “*Earth System Model Aerosol-Cloud Diagnostics Package (ESMAC Diags) Version 1: Assessing E3SM Aerosol Predictions*

Using Aircraft, Ship, and Surface Measurements” (<https://doi.org/10.5194/gmd-2021-350>) addressing comments from reviewers and some other updates:

- Re-organize data quality-control module
- Fix a bug that uses incorrect temperature and pressure for FIMS measurements
- Update quality controls for surface CPC for HI-SCALE
- Use a ship exhaust-free CN and CCN data for MARCUS
- Some other minor bug fixings

V1.0.0 (May 24, 2022): updated for the final publication of GMD paper “*Earth System Model Aerosol-Cloud Diagnostics Package (ESMAC Diags) Version 1: Assessing E3SM Aerosol Predictions Using Aircraft, Ship, and Surface Measurements*” (<https://doi.org/10.5194/gmd-15-4055-2022>) with updated readme file.

V2.0.0 (August, 2022): major update with new code organization, adding new observations and new diagnostics:

- Code structure is re-organized to put preprocessing offline and save prepared data.
- Add cloud variables from ground and satellite measurements, such as cloud fraction, cloud top height, cloud optical depth, cloud droplet number concentration, cloud droplet effective radius, top of atmosphere albedo, etc.
- Add new diagnostics such as percentiles, histograms, joint histograms, scatter plots and heatmaps. The two-variable scatter plots, joint histograms, and three-variable heatmaps provides the functionality to study ACI-related relationships.

V2.1.0 (October, 2022): add long-term diagnostics at ARM SGP (year 2011-2020) and ENA (year 2016-2018) sites.

V2.1.1 (November 20, 2022): for long-term sites at SGP and ENA, preprocess satellite and E3SM data within 5x5 degree domain around SGP or ENA site. This gives larger sample size to study aerosol-cloud interactions.

Appendix A: namelist of E3SM script. *fincl2* defines output variables with the output frequency *nhtfrq* = -1 (means 1 hr), equivalent to output interval *mfilt* = 24 (24 per day). *Fincl2latlon* defines the latitude and longitude range of *fincl2* output.

```

nhtfrq      = 0,-1
mfilt       = 1,24
avgflag_pertape ='A','I',
fincl1      = 'FREQR',
              'PS'
fincl2      = 'PS',      !! dynamical fields
              'U',       !! ..
              'V',       !! ..
              'T',       !! ..
              'Q',       !! vapor (kg/kg)
              'CLDLIQ',  !! cloud hydrometeors (kg/kg)
              'CLDICE',   !! ..
              'CLDTOT',
              'CLDLLOW',
              'CLDMED',
              'CLDHGH',
              'NUMLIQ',   !! ..
              'NUMICE',   !! ..
              'RAINQM',   !! ..
              'SNOWQM',   !! ..
              'NUMRAI',   !! ..
              'NUMSNO',   !! ..
              'PBLH',     !! PBL height
              'LHFLX',    !! energy fluxes
              'SHFLX',   !! ..
              'FLDS',
              'FLNS',
              'FLNT',    !! ..
              'FLUT',
              'FSNS',
              'FSNT',    !! ..
              'FSUTOA',
              'FSDS',
              'FSDSC',
              'SOLIN',
              'LWCF',
              'SWCF',
              'TREFHT',  !! ..
              'Z3',       !! geopotential height
              'RELHUM',   !! relative humidity (RH)
              'RHW',      !! RH with respect to water
              'RHI',      !! RH with respect to ice
              'RHICE',    !! RH before nucleation
              'RHCFMIP',  !! RH with respect to water above 273 K, ice below 273 K
              'CLOUD',    !! cloud fraction
              'AWN1',     !! in-cloud values

```

```
'AWNC', !! Average cloud water number conc (1/m3)
'AQRAIN', !! Average rain mixing ratio (kg/kg)
'AQSNOW', !! Average snow mixing ratio (kg/kg)
'CCN1', !! CCN concentration at S=0.02% (#/cm3)
'CCN3', !! CCN concentration at S=0.1% (#/cm3)
'CCN4', !! CCN concentration at S=0.2% (#/cm3)
'CCN5', !! CCN concentration at S=0.5% (#/cm3)
'AREI', !! ..
'AREL', !! ..
'REL',
'REI',
'ACTREL',
'ACTREI',
'ACTNL',
'ACTNI',
'lambda_cloud',
'mu_cloud',
'FREQL', !! frequency of cloud appearance
'FREQI', !! ..
'FREQS', !! ..
'FREQR', !! ..
'PRECT', !! precipitation
'PRECC', !! ..
'PRECL', !! ..
'CDNUMC', !! vertically-integrated droplet concentration (m-2)
'CMELIQ', !! rate of cond-evap of liq within the cloud (kg/kg/s)
'DCQ', !! Q tendency due to moist processes (kg/kg/s)
'FICE', !! ice mass fraction
'IWC', !! grid box average ice water content (kg/m3)
'LWC', !! grid box average liquid water content (kg/m3)
'ICLDIWP', !! in-cloud ice water path
'ICLDTWP', !! in-cloud total water path
'TGCLDLWP', !! liquid water path (including convective clouds)
'TGCLDIWP', !! ice water path (including convective clouds)
'ICWNC', !! prognostic in-cloud water number conc (m-3)
'ICINC', !! prognostic in-cloud ice number conc (m-3)
'ICWMRST', !! Prognostic in-stratus water mixing ratio (kg/kg)
'ICIMRST', !! Prognostic in-stratus ice mixing ratio (kg/kg)
'AODVIS', !! AOD
'WP2_CLUBB', !! Vertical Velocity Variance (m2/s2)
'REFFCLWMODIS',
'TAUTMODIS',
'TAUWMODIS',
'TAUIMODIS',
'LWPMODIS',
'IWPMODIS',
'TOT_CLD_VISTAU',
'TOT_ICLD_VISTAU',
'LIQ_ICLD_VISTAU',
'ICE_ICLD_VISTAU',
'SO2', !!
'bc_a1', !! aerosols mass (kg/kg)
'bc_a3', !!
```

```

'bc_a4', !!
'dst_a1', !!
'dst_a3', !!
'mom_a1', !!
'mom_a2', !!
'mom_a3', !!
'mom_a4', !!
'ncl_a1', !!
'ncl_a2', !!
'ncl_a3', !!
'pom_a1', !!
'pom_a3', !!
'pom_a4', !!
'so4_a1', !!
'so4_a2', !!
'so4_a3', !!
!     'so4_a5', !!
'soa_a1', !!
'soa_a2', !!
'soa_a3', !!
!     'soa_a5', !!
'num_a1', !! aerosols number (#/kg)
'num_a2', !!
'num_a3', !!
'num_a4', !!
!     'num_a5', !!
'num_c1', !! aerosols number (#/kg)
'num_c2', !!
'num_c3', !!
'num_c4', !!
!     'num_c5', !!
'dgnd_a01', !! dry aerosol size
'dgnd_a02', !! ..
'dgnd_a03', !! ..
'dgnd_a04', !! ..
!     'dgnd_a05', !! ..
'dgnw_a01', !! wet aerosol size
'dgnw_a02', !! ..
'dgnw_a03', !! ..
'dgnw_a04', !! ..
!     'dgnw_a05', !! ..
'EXTINCT', !! Aerosol extinction (1/m)
'AODABS', !! Aerosol absorption optical depth 550 nm
'AODALL',
'ABSORB', !! Aerosol absorption (1/m)
fincl2lonlat =      '260e:265e_34n:39n', ! SGP (~5x5 degs)
!           '330e:335e_37n:42n', ! ENA
!           '202e:240e_19n:40n', ! CSET
!           '60e:160e_42s:70s', ! MARCUS
!           '133e:164e_42s:63s', ! SOCRATES
!           '202e:243e_20n:35n', ! MAGIC

```

References:

- Albrecht, B., Ghate, V., Mohrmann, J., Wood, R., Zuidema, P., Bretherton, C., Schwartz, C., Eloranta, E., Gienke, 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.
- Bennartz, R.: Global assessment of marine boundary layer cloud droplet number concentration from satellite, Journal of Geophysical Research: Atmospheres, 112, <https://doi.org/https://doi.org/10.1029/2006JD007547>, 2007.
- Fast, J. D., Berg, L. K., Alexander, L., Bell, D., D'Ambro, E., Hubbe, J., Kuang, C., Liu, J., Long, C., Matthews, A., Mei, F., Newsom, R., Pekour, M., Pinterich, T., Schmid, B., Schobesberger, S., Shilling, J., Smith, J. N., Springston, S., Suski, K., Thornton, J. A., Tomlinson, J., Wang, J., Xiao, H., and Zelenyuk, A.: Overview of the HI-SCALE Field Campaign: A New Perspective on Shallow Convective Clouds, Bull. Amer. Meteor. Soc., 100, 821-840, <https://doi.org/10.1175/bams-d-18-0030.1>, 2019.
- Gallo, F., Uin, J., Springston, S., Wang, J., Zheng, G., Kuang, C., Wood, R., Azevedo, E. B., McComiskey, A., Mei, F., Theisen, A., Kyrouac, J., and Aiken, A. C.: Identifying a regional aerosol baseline in the eastern North Atlantic using collocated measurements and a mathematical algorithm to mask high-submicron-number-concentration aerosol events, Atmos. Chem. Phys., 20, 7553-7573, <https://doi.org/10.5194/acp-20-7553-2020>, 2020.
- Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z., and Quaas, J.: Remote Sensing of Droplet Number Concentration in Warm Clouds: A Review of the Current State of Knowledge and Perspectives, Reviews of Geophysics, 56, 409-453, <https://doi.org/https://doi.org/10.1029/2017RG000593>, 2018.
- Humphries, R.: MARCUS ARM CN and CCN data reprocessed to remove ship exhaust influence (v2) [dataset], <https://doi.org/10.25919/ezp0-em87>, 2020. Accessed 8 March 2022.
- Humphries, R. S., McRobert, I. M., Ponsonby, W. A., Ward, J. P., Keywood, M. D., Loh, Z. M., Krummel, P. B., and Harnwell, J.: Identification of platform exhaust on the RV Investigator, Atmos. Meas. Tech., 12, 3019-3038, <https://doi.org/10.5194/amt-12-3019-2019>, 2019.
- Lewis, E. R. and Teixeira, J.: Dispelling clouds of uncertainty, Eos, Trans. Amer. Geophys. Union, 96, <https://doi.org/10.1029/2015eo031303>, 2015.
- Lim, K.-S. S., Riihimaki, L., Comstock, J. M., Schmid, B., Sivaraman, C., Shi, Y., and McFarquhar, G. M.: Evaluation of long-term surface-retrieved cloud droplet number concentration with in situ aircraft observations, Journal of Geophysical Research: Atmospheres, 121, 2318-2331, <https://doi.org/https://doi.org/10.1002/2015JD024082>, 2016.
- 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.

Min, Q. and Harrison, L. C.: Cloud properties derived from surface MFRSR measurements and comparison with GOES results at the ARM SGP Site, Geophysical Research Letters, 23, 1641-1644, <https://doi.org/https://doi.org/10.1029/96GL01488>, 1996.

Minnis, P., Sun-Mack, S., Young, D. F., Heck, P. W., Garber, D. P., Chen, Y., Spangenberg, D. A., Arduini, R. F., Trepte, Q. Z., Smith, W. L., Ayers, J. K., Gibson, S. C., Miller, W. F., Hong, G., Chakrapani, V., Takano, Y., Liou, K. N., Xie, Y., and Yang, P.: CERES Edition-2 Cloud Property Retrievals Using TRMM VIRS and Terra and Aqua MODIS Data—Part I: Algorithms, IEEE Transactions on Geoscience and Remote Sensing, 49, 4374-4400, <https://doi.org/10.1109/TGRS.2011.2144601>, 2011.

NETCDF: Introduction and Overview: <https://www.unidata.ucar.edu/software/netcdf/docs/index.html>, last access: 12 November 2022. 2022.

Riihimaki, L., McFarlane, S., and Sivaraman, C.: Droplet Number Concentration Value-Added Product, ARM Research Facility, Report number: DOE/SC-ARM-TR-140, 2021.

Tang, S., Fast, J. D., Zhang, K., Hardin, J. C., Varble, A. C., Shilling, J. E., Mei, F., Zawadowicz, M. A., and Ma, P. L.: Earth System Model Aerosol–Cloud Diagnostics (ESMAC Diags) package, version 1: assessing E3SM aerosol predictions using aircraft, ship, and surface measurements, Geosci. Model Dev., 15, 4055-4076, <https://doi.org/10.5194/gmd-15-4055-2022>, 2022.

Turner, D. D., Lo, C., Min, Q., Zhang, D., and Gaustad, K.: Cloud Optical Properties from the Multifilter Shadowband Radiometer (MFRSRCLDOD): An ARM Value-Added Product, ARM Research Facility, Report number: DOE/SC-ARM-TR-047, 2021.

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.

Wu, P., Dong, X., Xi, B., Tian, J., and Ward, D. M.: Profiles of MBL Cloud and Drizzle Microphysical Properties Retrieved From Ground-Based Observations and Validated by Aircraft In Situ Measurements Over the Azores, Journal of Geophysical Research: Atmospheres, 125, e2019JD032205, <https://doi.org/https://doi.org/10.1029/2019JD032205>, 2020.

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

Supplementary information

Table S1: data used for HI-SCALE

Platform	Instrument	Measurements	Datastream name	DOI or link
Ground	Aerosol chemical speciation monitor (ACSM)	Aerosol composition	sgpaosacsmC1.b1	DOI: 10.5439/1762267
	Scanning mobility particle sizer (SMPS)	Aerosol size distribution (20–700 nm)	sgpaossmmpsS01.a1 (IOP1) and shilling-smps (IOP2)	DOI: 10.5439/1476898
	Nano scanning mobility particle sizer (nanoSMPS)	Aerosol size distribution (2–150 nm)	sgpaosnosmmpsS01.a1	DOI: 10.5439/1242975
	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 – 1000 nm), number concentration	sgpaosuhisasS01.a1	DOI: 10.5439/1333828
	Condensation particle counter (CPC)	Aerosol number concentration (> 10 nm)	sgpaoscpcC1.b1	DOI: 10.5439/1025152
	Condensation particle counter – ultrafine (CPCU)	Aerosol number concentration (> 3 nm)	sgpaoscpcuS01.b1	DOI: 10.5439/1046186
	Cloud condensation nuclei (CCN) counter	CCN number concentration	sgpaosccn1colavgC1.b1	DOI: 10.5439/1342133
	Multifilter rotating shadowband radiometer (MFRSR)	Cloud optical depth and effective radius	sgpmfrsrclod1minC1.c1	DOI: 10.5439/1027296
	Cloud droplet number concentration retrieval (Ndop)	Cloud droplet number concentration	sgpndropmfrsrC1.c1	DOI: 10.5439/1131339
	Active remote sensing of clouds (ARSCL)	Cloud base height, cloud top height	sgparsclkazrbnd1kolliasC1.c0	DOI: 10.5439/1393438
	Surface meteorological station (MET)	Temperature, relative humidity, wind, pressure, precipitation	sgparmbeatmC1.c1	DOI: 10.5439/1333748
	Radiosonde	Temperature, relative humidity, wind		
	Cloud radar, lidar, ceilometer	Cloud fraction	sgparmbeclradC1.c1	DOI: 10.5439/1333228
	Total sky image (TSI)	Cloud fraction		
	Surface radiation measurement	Longwave and shortwave radiation		
	Microwave radiometer (MWR)	Liquid water path		
Satellite	Geostationary satellite-based retrievals using Visible Infrared	TOA shortwave and longwave radiation; cloud fraction; height, pressure and temperature at cloud	sgpvisstgridg13v4minnisX1.c1 sgpvisstpx2dg13minnisX1.c1	https://adc.arm.gov/discovery/#/results/s::sgpvisstgridg13v4minnisX1.c1

	Solar-Infrared Split Window Technique (VISST) algorithm	top; liquid water path; cloud optical depth; droplet effective radius		
Aircraft	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 – 1000 nm), number concentration	tomlinson-uhsas	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/tomlinson-uhsas
	Condensation particle counter (CPC)	Aerosol number concentration (> 10 nm)	mei-cpc	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/mei-cpc
	Condensation particle counter – ultrafine (CPCU)	Aerosol number concentration (> 3 nm)	mei-cpc	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/mei-cpc
	Cloud condensation nuclei (CCN) counter	CCN number concentration	mei-ccn	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/mei-ccn
	Interagency working group for airborne data and telemetry systems (IWG)	navigation information and basic atmospheric state parameters	mei-iwg1	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/mei-iwg1
	Fast integrated mobility spectrometer (FIMS)	Aerosol size distribution (10 – 425 nm)	wang-fims	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/wang-fims
	Passive cavity aerosol spectrometer (PCASP)	Aerosol size distribution (120 – 3000 nm)	tomlinson-pcasp	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/tomlinson-pcasp
	Best estimate aerosol size distribution (BEASD)	Aerosol size distribution combining FIMS, PCASP, CAS, and FCDP (~10nm - 10µm)	pekour-aafbe	DOI: 10.5439/1838448
	Cloud probe merged size distribution (mergedSD)	Cloud size distribution combining FCDP, 2-DS and HVPS (1.5µm - 9075µm)	mei-merged	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/mei-merged
	High-resolution time-of-flight aerosol mass spectrometer (AMS)	Aerosol composition	shilling-ams	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/shilling-ams
	Water content measuring system (WCM)	Cloud liquid and total water content	matthews-wcm	https://iop.archive.arm.gov/arm-iop/2016/sgp/hiscale/matthews-wcm

Table S2: data used for ACE-ENA

Platform	Instrument	Measurements	Datastream name	DOI or link
Ground	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 – 1000 nm), number concentration	enaaosuhsasC1.a1	DOI: 10.5439/1409033
	Condensation particle counter (CPC)	Aerosol number concentration (> 10 nm)	enaaoscpcfC1.b1	DOI: 10.5439/1046184
	Cloud condensation nuclei (CCN) counter	CCN number concentration	enaaosccn1colavgC1.b1	DOI: 10.5439/1342133
	Aerosol chemical speciation monitor (ACSM)	Aerosol composition	enaaosacsmC1.b2	DOI: 10.5439/1762267
	Condensation particle counter (CPC)	Aerosol number concentration (> 10 nm)	mei-cpc	DOI: 10.5439/1440985
	Multifilter rotating shadowband radiometer (MFRSR)	Cloud optical depth and effective radius	enamfrsrclod1minC1.c1	DOI: 10.5439/1027296
	Cloud droplet number concentration retrieval (Ndrop)	Cloud droplet number concentration	enandropmfrsrC1.c1	DOI: 10.5439/1131339
	Cloud droplet number concentration retrieval by Wu et al (2020)	Cloud droplet number concentration, cloud effective radius		DOI: 10.1175/jcli-d-20-0272.1
	Active remote sensing of clouds (ARSCL)	Cloud base height, cloud top height	enaarsclkazrbnd1kolliasC1.c0	DOI: 10.5439/1393438
	Surface meteorological station (MET)	Temperature, relative humidity, wind, pressure, precipitation	enaarmbeatmC1.c1	DOI: 10.5439/1333748
	Radiosonde	Temperature, relative humidity, wind		
	Cloud radar, lidar, ceilometer	Cloud fraction	enaarmbeclradC1.c1	DOI: 10.5439/1333228
	Total sky image (TSI)	Cloud fraction		
	Surface radiation measurement	Longwave and shortwave radiation		
	Microwave radiometer (MWR)	Liquid water path		

Satellite	Geostationary satellite-based retrievals using Visible Infrared Solar-Infrared Split Window Technique (VISST) algorithm	TOA shortwave and longwave radiation; cloud fraction; height, pressure and temperature at cloud top; liquid water path; cloud optical depth; droplet effective radius	enavisstgridm10minnisX1.c1 enavisstpx2dm10minnisX1.c1	https://adc.arm.gov/discovery/#/results/s::enavisstgridm10minnisX1.c1 https://adc.arm.gov/discovery/#/results/s::enavisstpx2dm10minnisX1.c1
Aircraft	Condensation particle counter – ultrafine (CPCU)	Aerosol number concentration (> 3 nm)	mei-cpc	DOI: 10.5439/1440985
	Cloud condensation nuclei (CCN) counter	CCN number concentration	enaaafccn2colaF1.b1, enaaafccn2colbF1.b1	DOI: not assigned
	Interagency working group for airborne data and telemetry systems (IWG)	navigation information and basic atmospheric state parameters	mei-iwg1	https://iop.archive.arm.gov/arm-iop/2017/ena/aceena/mei-iwg1
	Fast integrated mobility spectrometer (FIMS)	Aerosol size distribution (10 – 425 nm)	wang-fims	https://iop.archive.arm.gov/arm-iop/2017/ena/aceena/wang-fims
	Passive cavity aerosol spectrometer (PCASP)	Aerosol size distribution (100 – 3000 nm)	tomlinson-pcas	https://iop.archive.arm.gov/arm-iop/2017/ena/aceena/tomlinson-pcas
	Optical particle counter (OPC)	Aerosol size distribution (390 – 15960 nm)	pekour-opc_iso	https://iop.archive.arm.gov/arm-iop/2017/ena/aceena/pekour-opc_iso
	Best estimate aerosol size distribution (BEASD)	Aerosol size distribution combining FIMS, PCASP, CAS, and FCDP (~10nm - 10μm)	pekour-asdbe	DOI: 10.5439/1867870
	Cloud probe merged size distribution (mergedSD)	Cloud size distribution combining FCDP, 2-DS and HVPS (1.5μm - 9075μm)	mei-2dsfcdphvps	https://iop.archive.arm.gov/arm-iop/2017/ena/aceena/mei-2dsfcdphvps
	High-resolution time-of-flight aerosol mass spectrometer (AMS)	Aerosol composition	shilling-hrfams	Doi: 10.5439/1468474
	Water content measuring system (WCM)	Cloud liquid and total water content	matthews-wcm	Doi: 10.5439/1465759

Table S3: Instruments for MAGIC

Platform	Instrument	Measurements	Datastream name	DOI or link
Ship	Meteorological station (MET)	Temperature, relative humidity, wind, pressure	raynolds-marmet	https://iop.archive.arm.gov/arm-iop/2012/mag/magic/reynolds-marmet/
	Microwave radiometer (MWR)	Liquid water path, precipitable water vapor	magmwrret1liljclouM1.s2	DOI: 10.5439/1027369
	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 – 1000 nm), number concentration	magaosuhssasM1.a1	DOI: 10.5439/1333828
	Condensation particle counter (CPC)	Aerosol number concentration (> 10 nm)	magaoscpcfM1.a1	DOI: 10.5439/1046184
	Cloud condensation nuclei (CCN) counter	CCN number concentration	magaosccn100M1.a1	DOI: 10.5439/1227964
	Cloud droplet number concentration retrieval by Wu et al (2020)	Cloud droplet number concentration, cloud effective radius		DOI: 10.1175/jcli-d-20-0272.1
Satellite	Geostationary satellite-based retrievals using Visible Infrared Solar-Infrared Split Window Technique (VISST) algorithm	TOA shortwave and longwave radiation; cloud fraction; height, pressure and temperature at cloud top; liquid water path; cloud optical depth; droplet effective radius	magvisstpxg15minnisX1.c1	https://adc.arm.gov/discovery/#/results/s::magic%20visst

Table S4: Instruments for MARCUS

Platform	Instrument	Measurements	Datastream name	DOI or link
Ship	Meteorological station (MET)	Temperature, relative humidity, wind, pressure	maraadmetX1.b1	DOI: 10.5439/1593144
	Microwave radiometer (MWR)	Liquid water path, precipitable water vapor	marmwrrret1liljclouM1.s2	DOI: 10.5439/1027369
	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 – 1000 nm), number concentration	maraosuhssasM1.a1	DOI: 10.5439/1333828
	Condensation particle counter (CPC)	Aerosol number concentration (> 10 nm)	maraoscpcf1mM1.b1	DOI: 10.5439/1418260
	Cloud condensation nuclei (CCN) counter	CCN number concentration	maraosccn1colavgM1.b1	DOI: 10.5439/1342133
	Reprocessed CN and CCN	CN and CCN number concentration	MARCUS ARM CN and CCN data reprocessed to remove ship exhaust influence. v1.	DOI: 10.25919/ezp0-em87
Satellite	Geostationary satellite-based retrievals using Visible Infrared Solar-Infrared Split Window Technique (VISST) algorithm	TOA shortwave and longwave radiation; cloud fraction; height, pressure and temperature at cloud top; liquid water path; cloud optical depth; droplet effective radius	marvisstgridh08minnisX1.c1	https://adc.arm.gov/discovery/#/results/s::marvisstgridh08minnisX1.c1

Table S5: Instruments for CSET

Platform	Instrument	Measurements	Datastream name	DOI or link
aircraft	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 – 1000 nm), number concentration	Low Rate (LRT - 1 sps) Navigation, State Parameter, and Microphysics Flight-Level Data. Version 1.3	DOI: 10.5065/D65Q4T96
	Condensation nuclei counter (CNC)	Aerosol number concentration (11-3000 nm)	Same as above	DOI: 10.5065/D65Q4T96
	Passive cavity aerosol spectrometer (PCASP)	Aerosol size distribution (120 – 3000 nm)	Same as above	DOI: 10.5065/D65Q4T96
	PMS-King Liquid Water Content (LWC)	Liquid water content	Same as above	DOI: 10.5065/D65Q4T96
	1DC	Cloud droplet number size distribution (12.5 – 1590 μm)	Same as above	DOI: 10.5065/D65Q4T96
	2DC	Cloud droplet number size distribution (12.5 – 1590 μm)	Same as above	DOI: 10.5065/D65Q4T96
	CDP	Cloud droplet number size distribution (2 – 50 μm)	Same as above	DOI: 10.5065/D65Q4T96

Table S6: Instruments for SOCRATES

Platform	Instrument	Measurements	Datastream name	DOI or link
aircraft	Cloud condensation nuclei (CCN) counter	CCN number concentration	SOCRATES CCN measurements. Version 1.1	DOI: 10.5065/D6Z036XB
	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 – 1000 nm), number concentration	Low Rate (LRT - 1 sps) Navigation, State Parameter, and Microphysics Flight-Level Data. Version 1.3	DOI: 10.5065/D6M32TM9
	Condensation nuclei counter (CNC)	Aerosol number concentration (11-3000 nm)	Same as above	DOI: 10.5065/D6M32TM9
	PMS-King Liquid Water Content (LWC)	Liquid water content	Same as above	DOI: 10.5065/D6M32TM9
	1DC	Cloud droplet number size distribution (12.5 – 1590 µm)	Same as above	DOI: 10.5065/D6M32TM9
	2DC	Cloud droplet number size distribution (12.5 – 1590 µm)	Same as above	DOI: 10.5065/D6M32TM9
	2DS	Cloud droplet number size distribution (5 – 2565 µm)	Same as above	DOI: 10.5065/D6M32TM9
	CDP	Cloud droplet number size distribution (2 – 50 µm)	Same as above	DOI: 10.5065/D6M32TM9

Table S7: long-term measurements at SGP

Datastream name	Instrument	Measurements	Data period	DOI or link
sgpmetE13.b1	Surface meteorological station (MET)	Temperature, relative humidity, wind, pressure	20110101 – 20201231	DOI: 10.5439/1786358
sgpaoscpcC1.b1 sgpaoscpcfE13.b1	Condensation particle counter (CPC)	Aerosol number concentration (> 10 nm)	20110310 – 20160711 20161114 – 20200402	DOI: 10.5439/1025152 DOI: 10.5439/1046184
sgpaossmpsE13.b1	Scanning mobility particle sizer (SMPS)	Aerosol size distribution (20-700 nm)	20161115 – 20201231	DOI: 10.5439/1476898
sgpaosnanosmpsE13.b1	Nano scanning mobility particle sizer (nanoSMPS)	Aerosol size distribution (2-150 nm)	20161115 – 20201231	DOI: 10.5439/1635016
sgpaosuhssE13.b1	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 – 1000 nm), number concentration	20190109 – 20200731	DOI: 10.5439/1409033
sgptdmasizeC1.b1	Tandem Differential Mobility Analyzer (TDMA)	Aerosol size distribution (12 nm - 15 µm)	20110101 – 20141120	DOI: 10.5439/1025303
sgpaosccn1colspectraC1.b1 sgpaosccn2colspectraE13.b1	Cloud condensation nuclei (CCN) counter	CCN number concentration, spectral data	20110810 – 20170816 20170412 – 20201231	DOI: 10.5439/1342134 DOI: 10.5439/1323896
sgpaosacsmC1.b2 sgpaosacsmE13.b2	Aerosol chemical speciation monitor (ACSM)	Aerosol composition	20110101 – 20161003 20161129 – 20201231	DOI: 10.5439/1762267 DOI: 10.5439/1762267
sgpmfrsrcldod1minC1.c1	Cloud optical properties from the multifilter rotating shadowband radiometer (MFRSR)	Cloud optical depth, droplet effective radius, liquid water path, cloud fraction	20110101 – 20201231	DOI: 10.5439/1395157
sgparsclkazrbnd1kolliasC1.c1 sgparsclkazrbnd1kolliasC1.c0	Active remote sensing of clouds (ARSCL) product using Ka-band ARM zenith radars	Cloud base height, cloud top height	20110118 – 20140315 20140316 – 20201231	DOI: 10.5439/1228769 DOI: 10.5439/1393438
sgpinterpolatedsondeC1.c1	Interpolated radiosonde data	Temperature, relative humidity, wind	20110101 – 20201231	DOI: 10.5439/1095316
sgparmbeatmC1.c1 sgparmbeclradC1.c1	ARM best estimate (ARMBE) data for atmospheric variables (armbeatm) or cloud and radiation variables (armbeclrad)	Temperature, relative humidity, wind, pressure, precipitation; cloud fraction, liquid water path, radiative fluxes	20110101 – 20201231 20110101 – 20201231	DOI: 10.5439/1333748 DOI: 10.5439/1333228
sgvisstgridg13v4minnisX1.c1* sgvisstgridg16v4minnisX1.c1	Visible Infrared Solar-Infrared Split-Window Technique (VISSST)-derived 0.5°×0.5° gridded products from GOES13 or GOES16 satellite	cloud and radiative properties	20110101 – 20171231 20180101 – 20201231	https://adc.arm.gov/discriminatory/#/results/s::sgpvisstgridg13v4minnisX1.c1
sgvisstpx2dg13minnisX1.c1 sgvisstpx2dg16minnisX1.c1	VISSST-derived 2D pixel-level (4km) products from GOES13 or GOES16 satellite	cloud and radiative properties	20110101 – 20171231 20180101 – 20201231	https://adc.arm.gov/discriminatory/#/results/s::sgpvisstpx2dg13minnisX1.c1

*there are missing periods 20120923 – 20121015 and 20130522 – 20130530 which are filled by sgvisstgridg15v4minnisX1.c1

Table S8: long-term measurements at ENA

Datastream name	Instrument	Measurements	Data period	DOI or link
enametC1.b1	Surface meteorological station (MET)	Temperature, relative humidity, wind, pressure	20160101 – 20181231	DOI: 10.5439/1786358
enaaoscpcfC1.b1	Fine condensation particle counter (CPCF)	Aerosol number concentration (> 10 nm)	20160101 – 20181231	DOI: 10.5439/1046184
enaaosuhssasC1.b1	Ultra-High Sensitivity Aerosol Spectrometer (UHSAS)	Aerosol size distribution (60 - 1000 nm), number concentration	20160101 – 20181231	DOI: 10.5439/1409033
enaaosacsmC1.b2	Aerosol chemical speciation monitor (ACSM)	Aerosol composition	20160219 – 20181231	DOI: 10.5439/1762267
enaaoscnc1colspectraC1.b1	Cloud condensation nuclei (CCN) counter	CCN number concentration, spectral data	20160623 – 20181231	DOI: 10.5439/1342134
enamfrsrclodod1minC1.c1	Cloud optical properties from the multifilter rotating shadowband radiometer (MFRSR)	Cloud optical depth, droplet effective radius, liquid water path, cloud fraction	20160101 – 20181231	DOI: 10.5439/1027296
enaarsclkazrbnl1kolliasC1.c0	Active remote sensing of clouds (ARSCL) product using Ka-band ARM zenith radars	Cloud base height, cloud top height	20160101 – 20181231	DOI: 10.5439/1393438
enainterpolatedsondeC1.c1	Interpolated radiosonde data	Temperature, relative humidity, wind	20160101 – 20181231	DOI: 10.5439/1095316
enaarmbeatmC1.c1 enaarmbecldradC1.c1	ARM best estimate (ARMBE) data for atmospheric variables (armbeatm) or cloud and radiation variables (armbeclrad)	Temperature, relative humidity, wind, pressure, precipitation; cloud fraction, liquid water path, radiative fluxes	20160101 – 20181231	DOI: 10.5439/1333748 DOI: 10.5439/1333228
enavisstgridm10minnisX1.c1 enavisstgridm11minnisX1.c1	VISST-derived 0.5°×0.5° gridded products from Meteosat-10 or Meteosat-11 satellite	cloud and radiative properties	20160101 – 20180220 20180220 – 20181231	https://adc.arm.gov/discovery/#/results/s::enavisstgridm10minnisX1.c1
enavisstpx2dm10minnisX1.c1 enavisstpx2dm11minnisX1.c1	VISST-derived 2D pixel-level (4km) products from Meteosat-10 or Meteosat-11 satellite	cloud and radiative properties	20160101 – 20180220 20180220 – 20181231	https://adc.arm.gov/discovery/#/results/s::enavisstpx2dm10minnisX1.c1
Wu_et.al Retrieval	Stratus cloud retrieval from ground-based measurements using retrieval algorithm from Wu et al. (2020)	Cloud droplet number concentration, effective radius	20160101 – 20181231	DOI: 10.1029/2019JD032205
Aerosol Mask	Masking 1-min CPC data dominated by local aerosol sources using algorithm from Gallo et al. (2020)	Aerosol number concentration and mask flag	20160101 – 20181231	DOI: 10.5194/acp-20-7553-2020