Deep learning representation of the aerosol size distribution

Donifan Barahona[1], Katherine H. Breen[1, 2], Karoline Block[3], and Anton Darmenov[1]

[1]NASA, Goddard Space Flight Center, Greenbelt, MD, USA
[2]Morgan State University, Baltimore, MD, USA
[3]Leipzig Institute for Meteorology, Faculty of Physics and Earth Sciences, University of Leipzig, Leipzig, Germany

Correspondence: Donifan Barahona (donifan.o.barahona@nasa.gov)

Abstract. Aerosols influence Earth's radiative balance via the scattering and absorbing of solar radiation, affect cloud formation, and play important roles on precipitation, ocean seeding and human health. Accurate modeling of these effects requires knowledge of the the chemical composition and size distribution of aerosol particles present in the atmosphere. Computationally intensive applications like remote sensing and weather forecasting commonly use simplified representations of aerosol microphysics, prescribing the aerosol size distribution (ASD), introducing uncertainty in climate predictions and aerosol retrievals. This work develops a neural network model, termed MAMnet, to predict the ASD and mixing state using the bulk mass of aerosol and the meteorological state. MAMnet can be driven by the output of single moment, mass-based, aerosol schemes or using reanalysis products. We show that MAMnet is able to accurately reproduce the predictions of a two-moment microphysics aerosol model as well as field measurements. Our model paves the way to improve the physical representation of aerosols in physical models while maintaining the versatility and efficiency required in large scale applications.

1 Introduction

Aerosols play a crucial role in the Earth's climate system by influencing radiative forcing (Forster et al., 2007; Bender, 2020), cloud formation and lifetime (Christensen et al., 2020), and precipitation patterns (Stier et al., 2024). Aerosol particle size and composition determine their atmospheric lifetime (Seinfeld and Pandis, 2016), impact on human health (Arfin et al., 2023), long range transport (Uno et al., 2009), and their ability to become cloud droplets and ice crystals (Seinfeld et al., 2016). The size and composition of atmospheric aerosols are critical parameters determining the concentration of cloud condensation nuclei (CCN) in the atmosphere (Seinfeld and Pandis, 2016). Understanding the distribution and composition of atmospheric aerosols is thus essential for accurate climate and weather simulations (Seinfeld et al., 2016).

The representation of the aerosol size distribution (ASD) and mixing state is at the center of the ability of climate models to accurately simulate the transport and chemical evolution of aerosol species (Aquila et al., 2011; Bender et al., 2019). Atmospheric models represent the ASD using approximations with different degrees of sophistication. The bulk mass approach aggregates aerosol species into a limited number of prescribed bins based on their mass and composition (Jones et al., 1994; Langner and Rodhe, 1991; Ginoux et al., 2001; Chin et al., 2000). This approach inherently assumes externally mixed components, that is, all particles of the same size have the same composition (Seinfeld and Pandis, 2016). Bulk schemes offer low computational expense, but they fail to distinguish between processes affecting number concentration and mass, hence cannot

explicitly predict the ASD. However, they are amenable to data assimilation methods (Randles et al., 2017), and are often used in forecasting systems, satellite retrievals, and reanalysis products (Chu et al., 2002; Gelaro et al., 2017; Inness et al., 2019). For example, the Goddard Chemistry, Aerosol, Radiation, and Transport model, GOCART (Colarco et al., 2010b), explicitly calculates the mass of dust, black carbon, organic material, sea salt, and sulfate using a bulk, externally-mixed approach, and

30 forms the basis of aerosol assimilation in the MERRA-2 climate reanalysis (Gelaro et al., 2017; Randles et al., 2017).

In contrast to bulk methods, modal aerosol schemes estimate both the number concentration and mass of atmospheric aerosol, hence two-moments of the ASD (e.g., Whitby and McMurry, 1997; Wilson et al., 2001; Stier et al., 2005; Liu et al., 2012). Because they predict the number concentration and mass independently, modal schemes can handle internally-mixed aerosol species, where aerosol species partition between different modes, better approximating the ASD (Wilson et al., 2001; Her-

35 zog et al., 2004; Stier et al., 2005). This leads to a better representation of aerosol-cloud interactions (Adams and Seinfeld, 2002), the variability in net radiative effects (Herzog et al., 2004), and the effects of alterations to emissions on a global scale (Wei et al., 2022). More sophisticated aerosol schemes either calculate additional modes of the ASD (Zhang et al., 2020) or explicitly resolve it by using a binned approach (e.g., Adams and Seinfeld, 2002). These models offer the most physically consistent representation of the ASD, but often are too computationally expensive for operational forecast and long term climate

40 predictions.

In computationally intensive applications it is desirable to maintain the efficiency and simplicity of the bulk schemes, however with a realistic representation of the ASD. To address these challenges, there is growing interest in leveraging machine learning (ML) techniques to develop more efficient and accurate aerosol models (e.g., Gong et al., 2021; Rasp et al., 2018; Harder et al., 2022; Silva et al., 2021). ML models, we can in principle capture complex nonlinear relationships between

45 aerosol properties and environmental variables with reduced computational costs. For example, Harder et al. (2022) developed a surrogate of the Modal Aerosol Module (MAM; Liu et al., 2012) to predict the mass and number tendencies of aerosol species, with the aim to improve computational performance. The emulator replaces computationally intensive parts of MAM, however does not map the ASD to the mass of the aerosol species, a requirement to many assimilation and remote sensing algorithms (Randles et al., 2017; Buchard et al., 2017).

50 Here we present a novel ML-based approach for predicting the ASD and aerosol mixing state in atmospheric models that run bulk aerosol schemes. This is accomplished by developing a ML-based parameterization that emulates the ASD predicted by the MAM model, using as input the total mass of aerosol species from relatively fast single-moment bulk aerosol models like GOCART. By combining the strengths of machine learning and physical principles, the parameterization maps the bulk mass model into the ASD that would be predicted by the modal approach, enhancing the former. Our method offers a promising

55 avenue for advancing aerosol representation in climate predictions, data assimilation and remote sensing applications.

2 Methods and Data

We developed a neural network (NN) to estimate the ASD using as input the total mass of aerosol species and the meteorological state. This was accomplished by training the NN on simulated data using the MAM model implemented on the NASA's Global

Earth Observing System (GEOS), combining the modal aerosol species predicted by MAM into their total mass across all
60   modes, thus resembling the externally-mixed species predicted by the GOCART scheme. This section details the modeling
components as well as the development and evaluation approach of the NN.

2.1   Modeling components

The NASA Goddard Earth Observing System (GEOS), consists of a set of components that numerically represent different
aspects of the Earth system (atmosphere, ocean, land, sea-ice, and chemistry), coupled following the Earth System Model-
65   ing Framework (https://gmao.gsfc.nasa.gov/GEOS_systems/). In GEOS-AGCM configuration, atmospheric transport of water
vapor, condensate and other tracers, and associated land-atmosphere exchanges, is computed explicitly, whereas sea-ice and
sea surface temperature (SST) are prescribed as time-dependent boundary conditions (Reynolds et al., 2002; Rienecker et al.,
2008). Cloud microphysics in the operational version of GEOS uses a single moment microphysics scheme for short-term
weather forecast (Molod et al., 2015), and a two-moment cloud scheme in subseasonal and seasonal prediction (Barahona
70   et al., 2014; Molod et al., 2020). GEOS constitutes the modeling base of MERRA-2 (Modern Era Retrospective analysis for
Research and Applications, version 2), the first multidecadal reanalysis to integrate both aerosol and meteorological obser-
vations (Gelaro et al., 2017; Randles et al., 2017). In MERRA-2 aerosol fields are described using GOCART. Aerosols are
interactive and radiatively active, hence MERRA-2 has a representation of the aerosol direct effect. Aerosol assimilation uses
the Goddard Aerosol Assimilation System (GAAS), and the overall assimilation cycle is controlled by the meteorology.

2.1.1   Aerosol transport schemes

GEOS implements two aerosol schemes to interactively calculate the evolution of aerosol and gaseous tracers. GOCART (Chin
et al., 2000; Colarco et al., 2010a) is used operationally on weather forecast and data assimilation applications. GOCART is a
mass-based aerosol model that explicitly calculates the transport and evolution of dust, black carbon, organic material, sea salt,
and sulfate. Aerosol species are assumed externally mixed. Dust and sea salt are represented in five mass bins whereas a single
80   bin is assumed for other species. Both organics and black carbon are split into hydrophilics and hydrophobic components. Dust
and sea salt emissions are prognostic whereas sulfate and biomass burning emissions are obtained from the MERRA-2 dataset
(Randles et al., 2017).

GEOS also implements the MAM model (Liu et al., 2012), as an alternative aerosol scheme for research applications. MAM
is a modal aerosol scheme that predicts the mass and number concentration of Aitken (AIT), accumulation (ACC), primary
85   carbon (PCM), fine dust (FDU) and sea salt (FSS), and coarse dust (CDU) and sea salt (CSS) aerosol modes. The aerosol
representation is internally mixed with modal composition as detailed in Table 2. The total number of simulated tracers in
MAM is 31: 24 modal mass components and seven aerosol number concentrations. The size distributions for each mode is
assumed to follow a lognormal distribution, with prescribed geometric standard deviation for each mode (Liu et al., 2012).

Table 1. Aerosol species considered in this work; κ is the hygroscopicity parameter (Kreidenweis et al., 2005).

Abbreviation	Description	κ	Density Kg m^{-3}
SU	Sulfates, includes ammonium	0.64	1600
SS	Sea salt	1.3	2200
OG	Primary and secondary organics	0.25	900
POM	primary Organic Matter	0.25	900
BC	Black carbon	0.01	1600
DU	Dust	0.1	1700

Table 2. Aerosol modes predicted by MAM and MAMnet; σ_{g} is the geometric standard deviation.

Abbreviation	Mode	σ_{g}	Species in mode
ACC	Accumulation	1.8	SU, OG, POM, BC, SS
AIT	Aitken	1.6	SU, OG, SS
CDU	Coarse dust	1.8	SU, DU
CSS	Coarse sea salt	2.0	SU, SS
FDU	Fine dust	1.8	SU, DU
FSS	Fine sea salt	2.0	SU, SS
PCM	Primary carbon matter	1.6	POM, BC

2.2 Development of the deep learning model

90   We built a neural network, termed "MAMnet", to estimate the aerosol number concentration and composition mimicking the MAM model (Table 2), using as input the total mass mixing ratios for dust, sulfates, organics, black carbon and sea salt, and the atmospheric state (temperature, T and air density, ρ_{air}), for a total of 31 predicted tracers as shown in Fig. 1. This section describes the development of the NN.

2.2.1 Data generation

95   The AGCM configuration of GEOS, running MAM (referred to as "GEOS+MAM"), was used to develop a robust dataset to train the neural network. We ran a 5-year simulation at 1-degree horizontal resolution and 72 vertical levels, with diurnal, instantaneous outputs at UTC 9:00:00 and 21:00:00. From these simulation, 25 output files were randomly selected without replacement for training, and 10 for the testing of the trained model. Temperature and horizontal winds were "replayed" to MERRA-2. The replay technique is a form of nudging that combines analysis increments with the model results, to correct

100   the model state every six hours (Takacs et al., 2018). This ensures that the simulation reproduces the observed meteorological state.

We combined the internally-mixed, modal mass components parameterized by MAM across 5 different species including sulfate, sea salt, dust, primary and secondary organic matter, and black carbon (Table 1 and Fig. 1) to derive the total mass
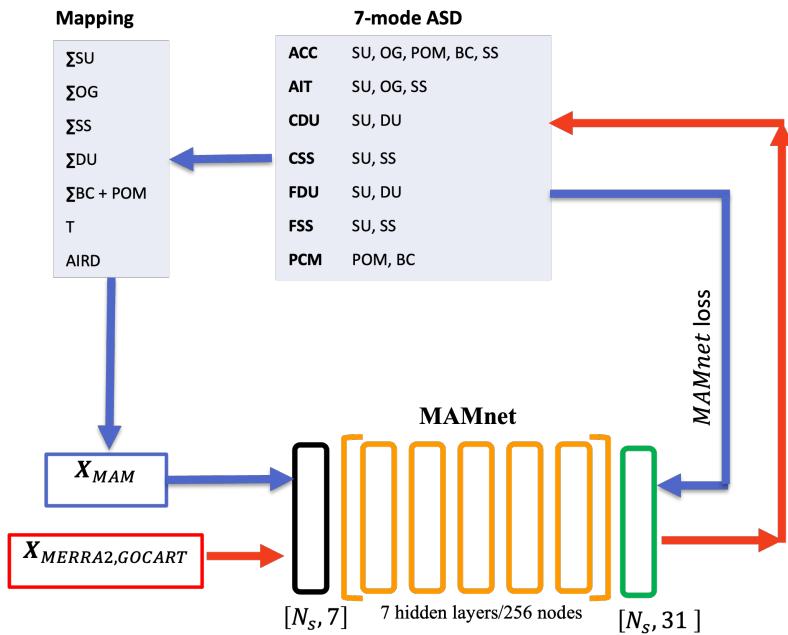
Figure 1. Neural network development workflow. Blue arrows represent the training steps, while red arrows correspond to the inference process. During training, the "mapping" step aggregates aerosol species across modes from the MAM output to construct the input X_{MAM}. The MAM output is then used to calculate the MAMnet loss. During inference, input from MERRA-2 or GOCART is used to predict the aerosol size distribution and mixing state, mimicking the MAM output. MAMnet consists of a single input layer (black), seven hidden layers (orange), and one output layer (green). T and ρ_{air} represent the temperature and air density, respectively. N_s is the number of samples.

mixing ratios for the input features. The input mass features were $\log_{10}$-transformed and all input variables (including T and ρ_{air}) were standardized by removing the mean and dividing by the standard deviation. Statistics used for normalization were calculated using 100 random instantaneous output files not used during training. Target variables included the mass of each of the MAM species and the number concentration for each mode. Because aerosol mass and number concentration vary over several orders of magnitude, we $\log_{10}$-transformed the targets, and filtered out values less than minimum threshold values, $10^{-20} \mu\mathrm{g\ Kg^{-1}}$ and $10^{-4}\ \mathrm{mg^{-1}}$ for mass and number, respectively. The modal aerosol dry diameter (hereafter D_{pg}) was not directly included as a a target of MAMnet. Instead it was used to check for mass conservation, that is, matching the predicted D_{pg} against the target values indicates that the mass and number concentration remain consistent in the prediction. D_{pg} was derived for each ith mode in the form (Seinfeld and Pandis, 2016),

$$D_{\mathrm{pg,i}} = \left(\frac{6}{\pi N_{\mathrm{i}}} \sum_{j=1}^{N_{\mathrm{sp,i}}} \frac{M_{\mathrm{j,i}}}{\rho_{\mathrm{j,i}}} \right)^{1/3} \exp\left(-\frac{3\ln^2 \sigma_{\mathrm{g,i}}}{2} \right) \tag{1}$$

where $\sigma_{g,i}$ and N_i are the geometric standard deviation and number concentration for the ith mode, respectively. $M_{j,i}$ and $\rho_{j,i}$ are the mass and density of the jth species in the ith mode, respectively, and $N_{sp,i}$ is the number of species present in the mode.

During training data was fed to the NN using a "single-level" approach where the 4-dimensional geospatial fields are flattened into one-dimensional arrays with a total number of samples given by:

$$N_{samples} = N_{lev} \times N_{time} \times N_{lat} \times N_{lon} \tag{2}$$

where for each of N_{time} time steps, N_{lev} is the number of model levels (72), N_{lat} is the number of latitudinal grid points (181), and N_{lon} is the number of longitudinal grid points (360). The single-level approach makes the parameterization resolution-independent facilitating its potential incorporation into GCMs with various resolutions. It also provides a large volume of training data (> 100 M samples) capturing a wide array of physically plausible instances in the training data for MAMnet. It however has the disadvantage that potentially significant spatial relations may be missing in the input. Tests using a whole-column approach suggested that given the large number of samples such caveat did not result in loss of accuracy (not shown).

2.2.2 Model architecture

Various levels of complexity were tested for the MAMnet architecture, including Multilayer Perceptrons (MLPs) and Convolutional Neural Networks (CNNs) (Bengio et al., 2017). These architectures have demonstrated success in capturing multi-scale behaviors of GCMs for different physical properties (e.g., Brenowitz and Bretherton, 2019; Rasp et al., 2018; Barahona et al., 2024). MLPs extract global patterns from the entirety of the input feature vector simultaneously, resulting in a greater number of model parameters for optimization. This approach compels the NN to make localized decisions, considering what occurs at an individual model level within each grid cell and time step, utilizing global information encompassing all grid cells and time steps. In contrast, CNNs extract features from smaller spatiotemporal blocks, enabling local decisions to be influenced by nearby information where the receptive field of each sample is a hyperparameter. Testing of both architectures showed that the MLP configuration exhibited superior performance and was easier to optimize. The final architecture is shown in Fig. 1.

The MAMnet model was trained using the Keras library with Tensorflow backend (Chollet et al., 2015). Optimization was carried out with the Adam algorithm (Kingma and Ba, 2014) using the minimum mean square error (MSE) as the loss function. Hyperparameter optimization for MAMnet was performed using the Keras Tuner software (O'Malley et al., 2019). Approximately 1500 optimization trials were performed using random configurations of the hyperparameters in Table 3, using a subset of the training data as in Yu et al. (2024). All trials used the same subset of the training/validation data (5 output files for training, 2 for validation). For each parameter set, a new model was built and trained for up to 100 epochs with the same early stopping criteria used during the training of MAMnet. For each trial, a custom metric was recorded at the end of each epoch, the convergence loss ($\mathcal{L}_{conv}$), defined as the absolute difference between the training and validation losses. Using this custom metric allowed us to select the model that generalizes the best over both the training and validation data sets. The best set of hyperparameters was selected by choosing the configuration that minimized the MSE on the validation set and had the lowest $\mathcal{L}_{conv}$.

Table 3. Parameters used during hyperparameter tuning for MAMnet. Optimal hyperparameters are shown in bold.

Hyperparameter	Values Interrogated
Number of dense layers	$1, 2, 3, 4, 5, 6, \mathbf{7}, 8, 9, 10, 15, 20$
Number of nodes per layer	$32, 64, 128, \mathbf{256}, 512$
Batch normalization	True, **False**,
Dropout	**True**, False,
Dropout rate	$\mathbf{0.1}, 0.2, 0.3, 0.5$
Initial learning rate	$1e{-}3, 1e{-}4, \mathbf{1e{-}5}, 1e{-}6$
Activation function	ReLU, ELU, **Leaky ReLU**
Activation α	$0.1, 0.2, \mathbf{0.3}$
Optimizer	**Adam**, SGD, RMSprop
Batch size	$\mathbf{64 \times 72}, 128 \times 72, 256 \times 72, 512 \times 72,$

2.3 Observational data

Besides synthetic data the neural network was evaluated on its ability to reproduce observations when driven by the MERRA-2 reanalysis output. This was important for testing the reliability of MAMnet when applied outside of the purely simulated environment. Near-surface aerosol number concentrations ranging from 30 nm to 500 nm, compiled by Asmi et al. (2011), were used for model evaluation. The dataset includes two years (2008–2009) of hourly measurements from 24 sites across Western Europe, as detailed in Table 4. These measurements were collected from two major monitoring networks: the European Supersites for Atmospheric Aerosol Research (EUSAAR) project, part of the Sixth Framework Programme of the European Commission (Philippin et al., 2009), and the German Ultrafine Aerosol Network (GUAN) (Birmili et al., 2009). The data are reported as cumulative number concentrations for four aerosol size ranges, N_{30}, N_{50}, N_{100}, and N_{250}, defined as,

$$N_X = \sum_{D_p=X}^{Y} N(D_p) \tag{3}$$

where the subscript X represents the aerosol number concentration for size range defined by threshold X, and $Y = 500$ nm for $X = 50, 100$ and 250 nm and $Y = 50$ nm for $X = 30$ nm. Equivalently, these can be calculated from the predicted ASD in the form (Seinfeld and Pandis, 2016),

$$N_X = \sum_{i=1}^{N_{\mathrm{mod}}} \frac{N_{\mathrm{i}}}{2} \left[\mathrm{erf}\left(\frac{\ln Y - \ln D_{\mathrm{pg,i}}}{\sqrt{2}\ln\sigma_{\mathrm{g,i}}} \right) - \mathrm{erf}\left(\frac{\ln X - \ln D_{\mathrm{pg,i}}}{\sqrt{2}\ln\sigma_{\mathrm{g,i}}} \right) \right] \tag{4}$$

where $N_{\mathrm{mod}} = 7$, is the number of lognormal modes. For each site, MERRA-2 derived aerosol mass concentration, temperature and air density are interpolated at the location and time of the measurements, then used in MAMnet to predict the ASD. Using Eqs. 1 and 4, the predicted N_{X} can be compared against the observations.

Table 4. Datasets for the period 2008–2009 used for comparison with surface aerosol size distributions predicted by MAMnet. The original data reference is given, although all data sets used in this work were curated by Asmi et al. (2011).

Station Name	Station Code	Altitude (m.a.s.l.)	Reference
Aspvreten	ASP	30	Tunved et al. (2004)
Birkenes	BIR	190	Amunsen et al. (1992)
Pallas	PAL	560	Lihavainen et al. (2008)
Preila	PLA	5	Ulevicius et al. (2010)
SMEAR II	SMR	181	Hari et al. (2013)
Vavihill	VHL	172	Kristensson et al. (2008)
Bösel	BOE	16	Birmili et al. (2009)
K-Puszta	KPO	125	Kiss et al. (2002)
Melpitz	MPZ	87	Engler et al. (2007)
Kosetice	OBK	534	Červenkova and Vá ňa (2010)
Hohenpeissenberg	SMPS	988	Birmili et al. (2003)
Waldhof	WAL	70	Birmili et al. (2009)
Cabauw	CBW	60	Russchenberg et al. (2005)
Harwell	HWL	60	Charron et al. (2007)
Mace Head	MHD	5	Jennings et al. (1991)
Finokalia	FKL	250	Mihalopoulos et al. (1997)
JRC-Ispra	ISP	209	Gruening et al. (2009)
Zeppelin	ZEP	474	Ström et al. (2003)
Puy de Dôme	PDD	1465	Venzac et al. (2009)
Schauinsland	SCH	1210	Birmili et al. (2009)
Zugzpitze	ZSF	2650	Birmili et al. (2009)
Jungfraujoch	JFJ	3580	Jurányi et al. (2011)
BEO Moussala	BEO	2971	Nojarov et al. (2009)
Monte Cimone	CMN	2165	Marinoni et al. (2008)

Further evaluation was performed by comparing the concentration of cloud condensation nuclei, N_{CCN}, derived from MAM-net against global datasets. Watson-Parris et al. (2019) utilized the Global Aerosol Synthesis and Science Project (GASSP) dataset (Reddington et al., 2017) to assess the vertical distributions of aerosol number and N_{CCN} in the global aerosol–climate model ECHAM-HAM. The GASSP database contains aerosol measurements from 37 field campaigns and over 1000 flights, largely concentrated over North America and Western Europe. Choudhury and Tesche (2022) estimated N_{CCN} from spaceborne CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) lidar measurements. Block et al. (2024) derived N_{CCN} based on the latest Copernicus Atmosphere Monitoring Service (CAMS) reanalysis provided by the European Centre for Medium-Range Weather Forecast (ECMWF) (Block, 2023). Finally, we compare MAMnet results against the GiOcean atmosphere-ocean-aerosol reanalysis (Song et al., 2025), derived from the NASA GEOS-S2S system (Molod et al., 2020). Unlike MERRA-2 which only assimilates the atmospheric state, GiOcean is a coupled atmosphere-ocean reanalysis that incorporates an updated model framework. This framework includes two-moment cloud microphysics, enabling the explicit calculation of N_{CCN}. The evaluation of N_{CCN} also serves as a test for the estimation of particle size, as CCN concentrations are highly sensitive to the aerosol size (Lee et al., 2013). We used the approach of Fountoukis and Nenes (2005) to estimate N_{CCN} from the derived 7-modal size distribution and modal composition. Hygroscopicity parameters, κ, for each mode were obtained by volume-weighting the values for each aerosol species as listed in Table 1.

3 Results and Discussion

We evaluated the MAMnet model for both its ability to reproduce the original MAM model when driven by the testing data set, and to reproduce observations when driven by aerosol concentrations derived from the MERRA-2 reanalysis. We assessed whether MAMnet reproduces the spatial distribution of aerosol variables in GEOS+MAM using the the mean Pearson's spatial correlation coefficient (R). We also calculated the mean log-bias, i.e.,

$$\mathrm{MLB} = \frac{\sum_{n=1}^{N_{\mathrm{samples}}} \log_{10}(\hat{\mathbf{Y}}) - \log_{10}(\mathbf{Y})}{N_{\mathrm{samples}}}, \tag{5}$$

where $\hat{\mathbf{Y}}$ and $\mathbf{Y}$ correspond to the predicted variables by MAMnet and GEOS+MAM, respectively. In general MLB in the range $[-0.5, 0, 5]$ indicates a prediction within an order of magnitude of the target value. These metrics are summarized in Fig. 2 for each pressure level and output variable.

MAMnet demonstrates strong performance in reproducing the modal number concentrations ("NUM" variables in Fig. 2) from the GEOS+MAM simulations, with high spatial correlations $R > 0.9$ and mean log-bias (MLB) within ± 0.1 across most pressure levels. However, performance slightly degrades at pressures below $p < 100$ hPa, particularly for the Aitken (NUM_AIT) and coarse dust modes (NUM_CDU), where correlations drop slightly ($R > 0.7$) and MLB increases to ± 0.3. The largest discrepancies occur near the surface ($p > 900$ hPa) and in the upper troposphere ($150 - 400$ hPa). Specifically, NUM_AIT shows underprediction between $150 - 400$ hPa while it overpredicts from $700 - 850$ hPa, indicating that MAMnet tends to underestimate fine particles near the tropopause and overestimate them in the mid-to-lower troposphere. Similarly, NUM_PCM exhibits negative biases near $150 - 400$ hPa, suggesting underprediction of particle number in the primary carbon
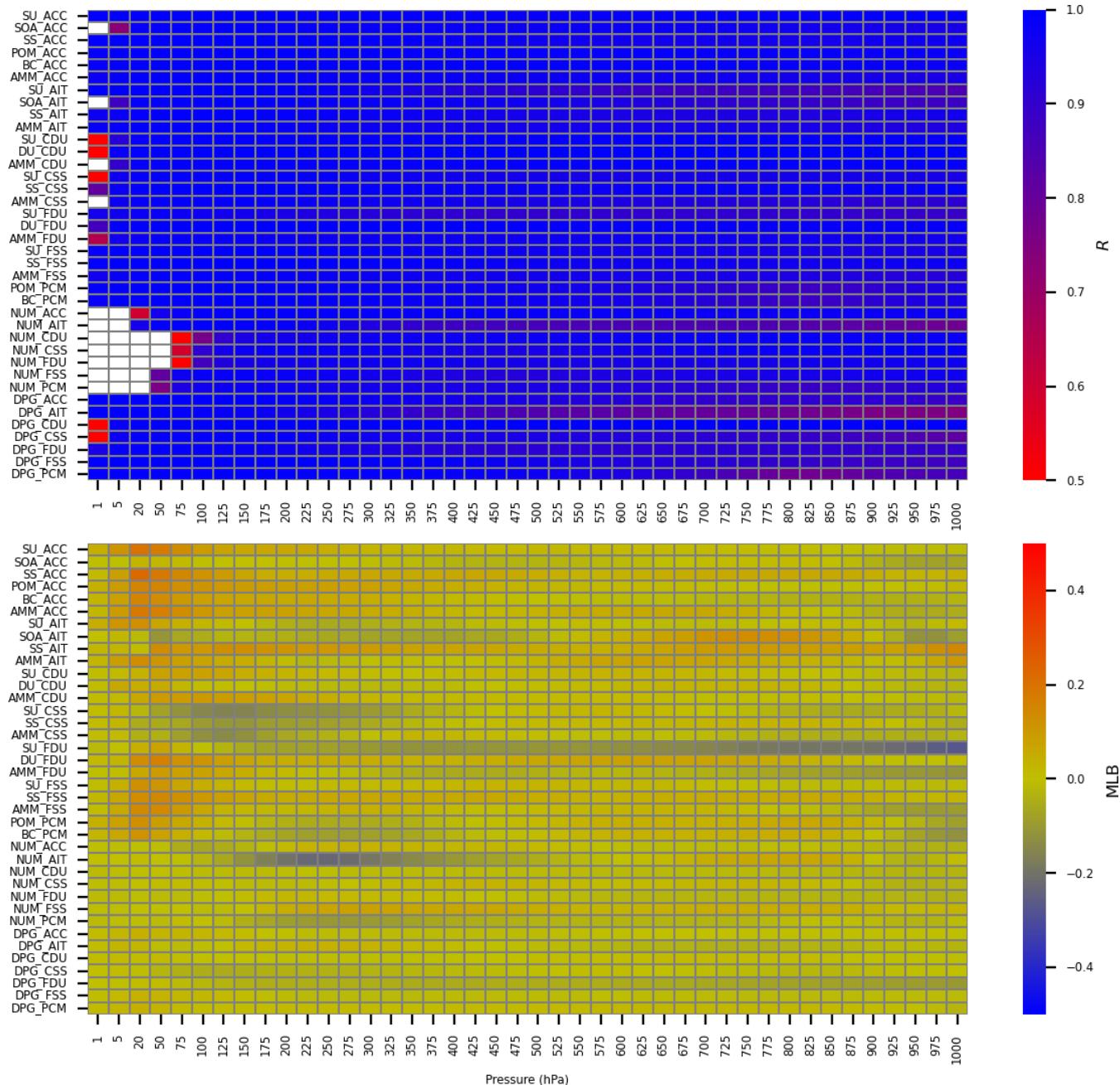
Figure 2. Pearson's spatial correlation (R) (top) and mean log-bias (Eq. 5; bottom) predicted by MAMnet calculated on the reserved test set, against the GEOS+MAM simulation. Results are shown for mass (Table 1 and 2) and number (NUM) concentration for each mode, as well as for the derived modal diameter (DPG).

195   mode at higher altitudes. The MLB patterns (Fig. 2, bottom panel) reveal localized biases at specific pressure ranges. Positive MLBs (orange shading) occur in NUM_PCM and NUM_ACC, indicating slight overestimation at mid-to-lower pressure levels. In contrast, NUM_CDU displays small negative MLB (blue shading) around $500 - 700$ hPa, suggesting underprediction of coarse dust particles in the mid-troposphere. In summary, while MAMnet captures the overall trends in the modal number concentrations with high accuracy, systematic errors emerge for specific modes and pressure ranges. Aitken mode number

200   concentrations show the largest deviations, particularly near the surface and upper troposphere, while biases in primary carbon and coarse dust modes are localized to mid-tropospheric levels.

Figure 3 shows the zonal mean profiles of modal aerosol number concentration, comparing GEOS+MAM outputs (left column), MAMnet predictions (center column). Consistent with Fig. 2, the Aitken mode exhibits the largest biases, characterized by underestimation above 400 hPa and overestimation in the lower troposphere, mostly below 700 hPa. Unlike other aerosol

205   modes, the vertical distribution of Aitken mode particles is unique, with higher concentrations found in the upper troposphere and lower stratosphere ($p < 400$ hPa) compared to lower altitudes, where the other modes exhibit the highest concentrations near the surface. The underestimation in the upper troposphere and overestimation in the lower troposphere may be influenced by a "dilution" effect, as Aitken particles contribute relatively little mass compared to other modes. The accumulation mode (ACC), coarse sea salt (CSS), and fine sea salt (FSS) modes show generally strong agreement between true and predicted val-

210   ues, with minimal biases across most pressure levels, with MLB values close to zero. However, localized biases are apparent for FSS and PCM near the surface, suggesting slight overestimation in these regions. The coarse dust mode (CDU) and fine dust mode (FDU) exhibit minimal errors overall, with MLB values near zero across most pressure levels. MAMnet accurately predicts the aerosol number concentration for most modes, with systematic biases for the Aitken and primary carbon modes, particularly near the tropopause and in the lower troposphere, suggest that smaller and less massive particles are more difficult

215   to predict accurately due to their unique vertical distribution and sparse representation in the data.

MAMnet accuately reproduces the spatial patterns of the aerosol mass, with accumulation mode mass variables such as SU_ACC, SS_ACC, and SOA_ACC showing near-perfect correlations ($R > 0.9$) across the entire pressure range (Fig. 2). This is also the case for most other variables with only DU_FDU and AMM_FSS showing slight reduction in correlation near 1000 hPa, indicating slightly worse performance in the lower atmosphere. Biases shown in Fig. 2 (bottom) indicate that all

220   but six mass tracers (SOA_ACC, SU_AIT, SOA_AIT, SU_CSS, SS_A_CSS, AMM_CSS) are systematically overpredicted for $p < 200$ hPa where mass values are very small ($\sim 10^{-20}$ kg kg^{-1}). These errors tend to be exacerbated in logarithmic space but remain negligible in absolute terms. Negative biases are also notable for SU_FDU and AMM_FDU, which become increasingly negative towards the surface. This is explained by the low mass of sulfates in the fine dust mode leading to a "dilution" of sulfate in fine dust mode relative to other aerosol modes. Overall, the model demonstrates robust predictive skill

225   for most aerosol types, with minor discrepancies concentrated near the surface and in sparse aerosol regimes.

Differences in the zonal distribution between aerosol modes can contribute to errors, especially due to the uneven representation of smaller particles, such as those in the Aitken and organic modes, in the training dataset, typically referred to as "class imbalance" (Japkowicz and Stephen, 2002; Buda et al., 2018). For instance, the mass of sulfate particles in the accumulation mode is often at least ten times greater than in the Aitken mode, whereas the opposite is true for the number concentration. As
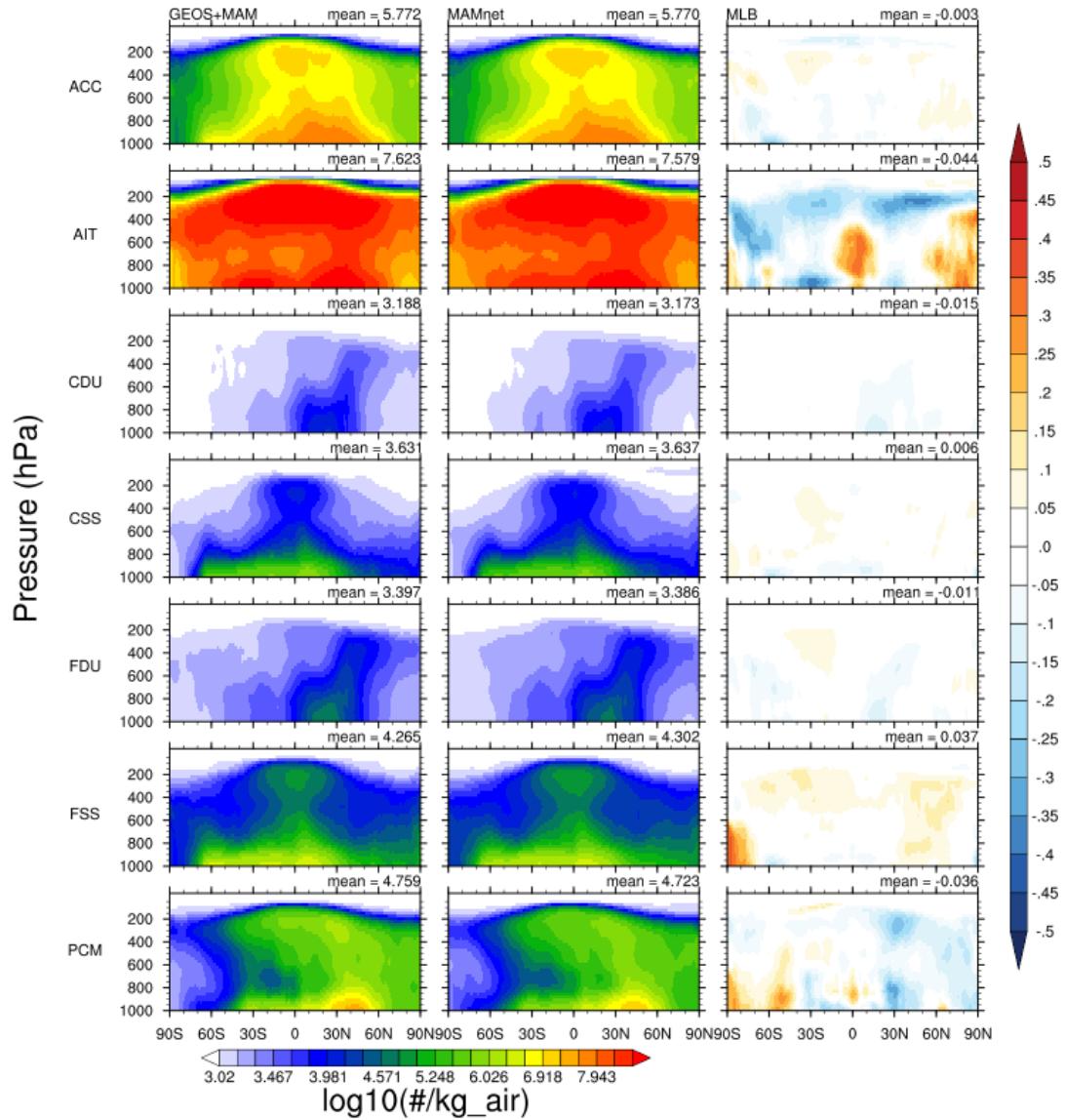
Figure 3. Zonal profiles for modal aerosol number concentration. Left: GEOS+MAM reserved test set. Middle: MAMnet prediction. Right: Mean Log-Bias. From top: Accumulation (ACC), Aitken (AIT), coarse dust (CDU), coarse sea salt (CSS), fine dust (FDU), fine sea salt (FSS), primary carbon matter (PCM).
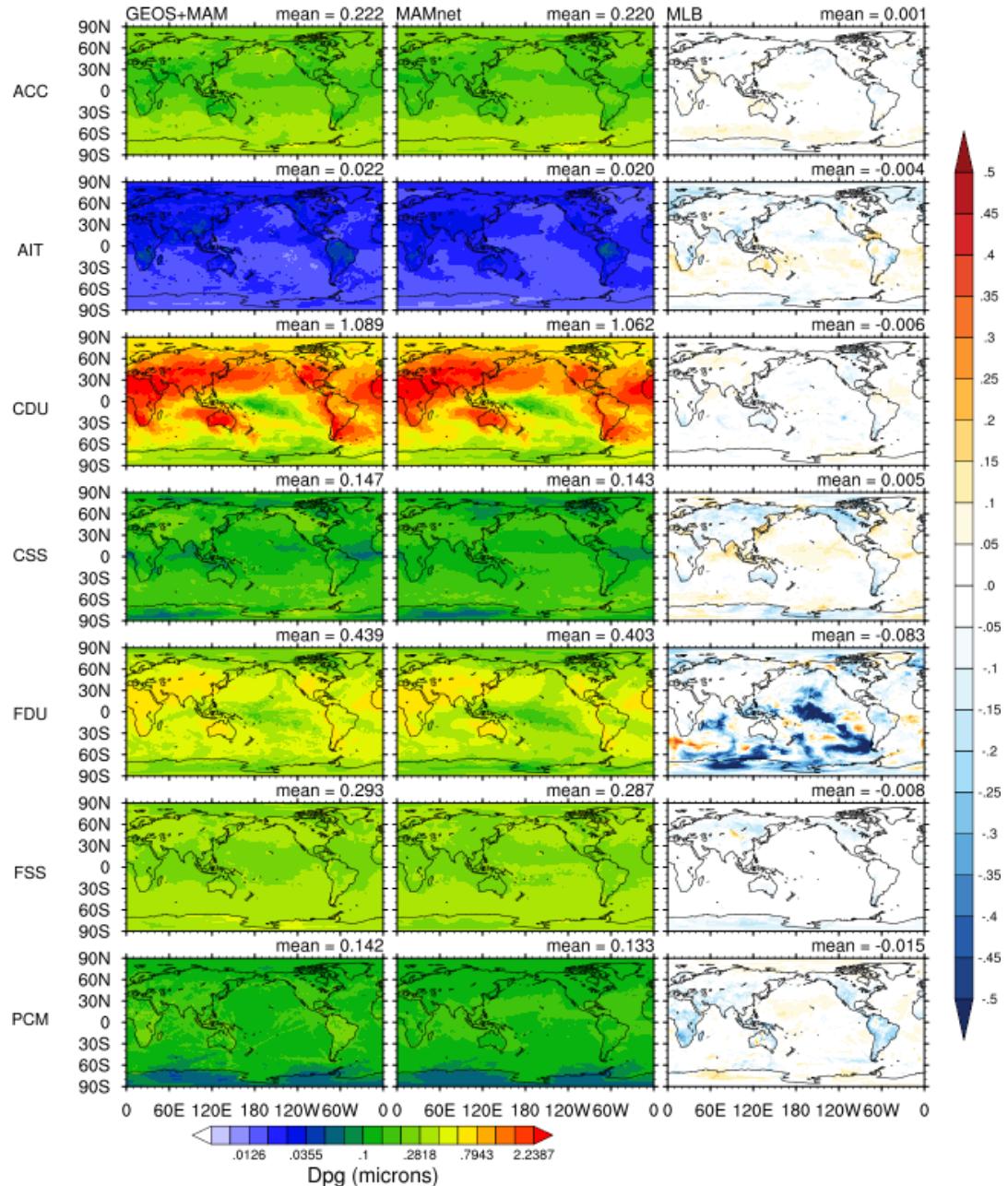
Figure 4. Modal geometric diameter, D_{pg} at 950 hPa. Left: GEOS+MAM reserved test set. Middle: MAMnet prediction. Right: Residual ($\log_{10}$(MAMnet) - $\log_{10}$(GEOS+MAM)). From top: Accumulation (ACC), Aitken (AIT), coarse dust (CDU), coarse sea salt (CSS), fine dust (FDU), fine sea salt (FSS), primary carbon matter (PCM).

230 a result, the variability in sulfate mass is primarily driven by the accumulation mode, causing the Aitken mode to be underrepresented in the neural network's input data. Despite this, the residual differences between predicted and true values for Aitken mode aerosol number concentrations are small compared to other modes, and the mean global error remains well below an order of magnitude, highlighting the neural network's accuracy.

Figures 4 and 5 show the MAMnet-derived D_{pg} in remarkable agreement with GEOS+MAM. MLBs between the two
235 datasets are typically below 0.01, indicating consistency in the predicted modal mass and number concentrations, and suggesting that MAMnet conserves mass. This is consistent with the correlation coefficients for DPG ($R > 0.9$), shown in Fig. 2, highlighting strong agreement with the test dataset across most pressure levels. The global distribution of D_{pg} for the different aerosol modes closely matches the spatial patterns of the modal number concentrations. Larger residuals are observed near the surface in the tropics and the Southern Hemisphere, particularly for fine dust (FDU). This discrepancy arises primarily due to
240 the very low number concentrations of fine dust over the oceans, making it challenging for the neural network to accurately predict values close to zero. However, such inaccuracies are unlikely to affect climate simulations, as regions with negligible aerosol concentrations typically do not contribute significantly to atmospheric processes. Residuals for D_{pg} in the coarse dust (CDU) mode are also slightly larger compared to other modes. This is evident in the zonal mean profiles (Fig. 5), where biases are most prominent in the tropics and near the Arctic. Additionally, MAMnet tends to underestimate D_{pg} in the Southern
245 Hemisphere around 30S, particularly in the free troposphere. This underestimation likely results from class imbalance, as dust concentrations in this region are very low, making it difficult for the neural network to learn accurate predictions. Overall MAMnet shows remarkable skill at reproducing the ASD from GEOS+MAM, with no evidence of systematic biases.

3.1   Evaluation against observations

The ability of a neural network to generalize to new data is a key measure of its effectiveness and reliability in real-world
250 applications. While the neural network may perform well reproducing simulated data, it is important to test whether MAMnet is able to reproduce patterns observed in nature. To accomplish this, we take advantage of the ability of MAMnet to work with reanalysis data, that is, using as input the assimilated fields of MERRA-2.

We first tested whether MAMnet had learned the physical relationships underlying the ASD or simply memorized the training data. To investigate this, we evaluated whether MAMnet, when driven with MERRA-2 inputs, would reproduce the biases of
255 the GEOS+MAM simulation or instead align with the MERRA-2 fields. Figure 6 compares the total aerosol mass column from MERRA-2 (left), MAMnet driven by MERRA-2 inputs (center), and GEOS+MAM (right). Since aerosol concentrations in GEOS+MAM are not assimilated, they are expected to differ from MERRA-2, which incorporates observational constraints. This discrepancy is evident in Fig. 6, where GEOS+MAM tends to underestimate black carbon (BC) and sea salt (SS) over the ocean. If MAMnet had merely memorized the GEOS+MAM outputs, these same biases would persist when MERRA-2 inputs
260 were used. Instead, MAMnet accurate reproduces the MERRA-2 aerosol concentrations when driven by MERRA-2 inputs. It is important to note that in this analysis, the aerosol outputs were mapped from the seven modes produced by MAMnet. This result highlights the internal consistency of the model and demonstrates that MAMnet generalizes to new, unseen data.
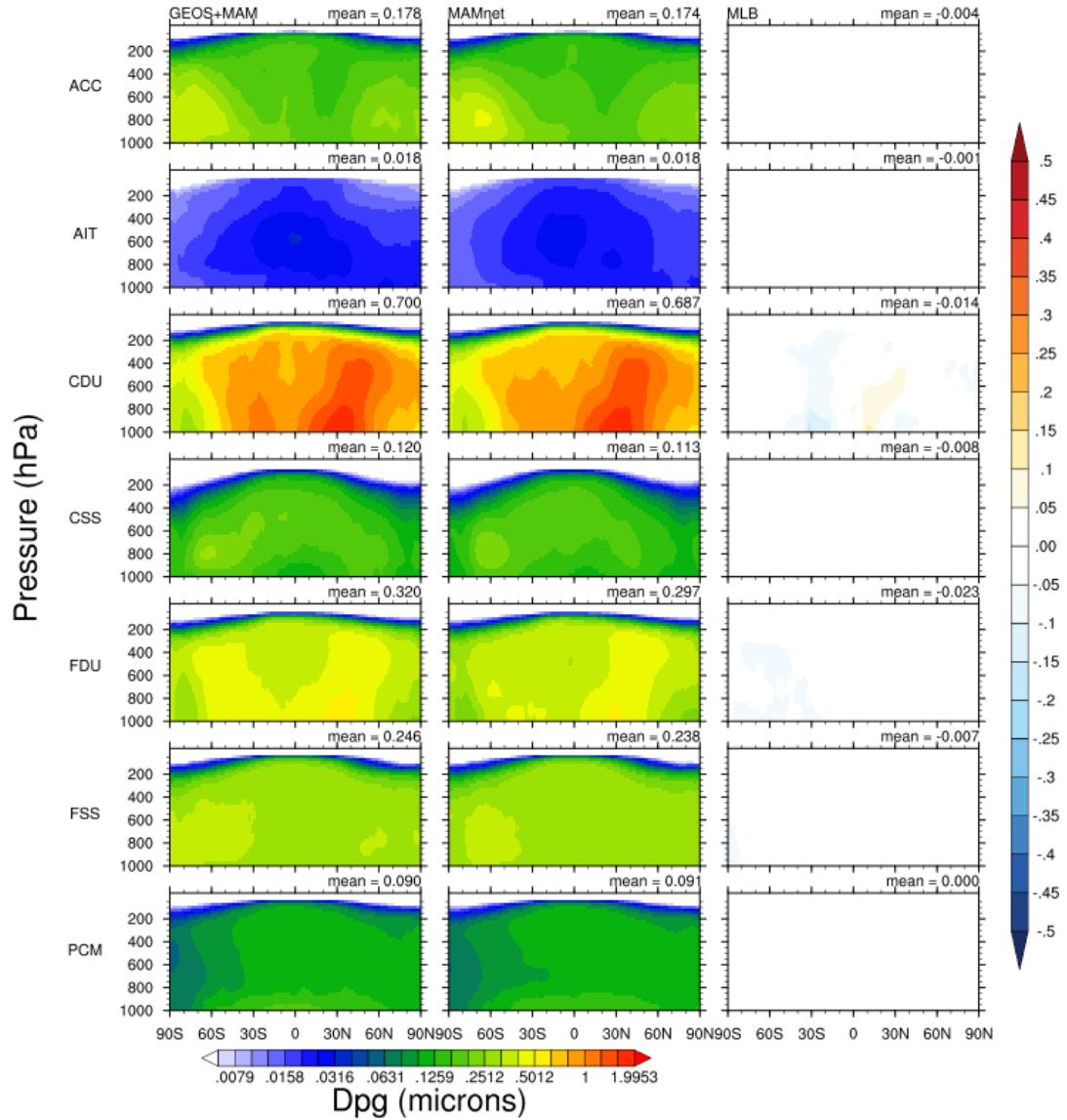
Figure 5. As in Figure 3, but for geometric mean diameter (D_{pg}) per mode.

3.1.1 Comparison against ground observations

Figure 7 compares the cumulative ASD predicted by MAMnet against surface observations from different European sites.
265 These are mostly coastal sites with composition typical of clean and polluted continental origin, that is, mostly of sulfates, dust, organics and sea salt (Asmi et al., 2011). Altitude ranges from a few meters to about 3 km providing a good overview of the lower troposphere. Although representing a limited set, the range of aerosol compositions, sources, and altitudes offers a meaningful assessment of the model's ability to generalize to different atmospheric states. To carry out the comparison, MAMnet was run using collocated aerosol concentrations and meteorological fields obtained from MERRA-2 at each site, and
270 using Eq. 4.

Except for high altitude sites (Fig. 7, bottom row), MAMnet tends to predict slightly lower median values compared to observations, with this discrepancy becoming more pronounced as particle size increases (N100 and N250). This is particularly noticeable at locations such as PAL, PLA, OBK, MHD, FKL, and JFJ, where the model underestimates values consistently. The pattern is reversed at high-altitude sites (PDD, SCH, ZSF, JFJ, BEO, and CMN), where median N100 and N250 are gener-
275 ally overestimated by MAMnet, although the observations themselves display significant variability. Some locations like SMR, WAL, CBW, and SCH exhibit better agreement, with overlapping medians and interquartile ranges. Additionally, the spread of values for MAMnet is typically narrower than for observations, indicating that the model underestimates variability. Observations also show more outliers, whereas MAMnet predictions are more constrained. There are a few exceptions where MAMnet slightly overestimates values, such as VHL (N250) and SMR (N100). Overall, systematic bias exists in MAMnet-predicted
280 particle concentrations, particularly for larger size bins, capturing less variability compared to observations. MERRA-2 data, which provide MAMnet's input, may not resolve local emissions, terrain, or small-scale meteorology, such as boundary layer height and humidity. This can lead to biases, particularly in the larger particle size categories that depend on aerosol growth processes. Moreover, the training data likely lacks sufficient diversity, particularly for remote or high-altitude sites, which are probably underrepresented in the training set. Additionally, aerosol evolution involve complex, nonlinear interactions that are
285 not explicitly modeled by MAMnet. These factors likely contribute to the model's challenges in capturing the magnitude and variability of aerosol concentrations observed in the real world. Additionally, it is important to note that retrievals of ASD are inherently complex, and experimental errors can be significant, particularly for larger particle sizes (Asmi et al., 2011). Nevertheless, the consistent results across many sites indicate that MAMnet is capable of reasonably capturing the ASD on regional scales when driven by reanalysis data.

290 ### 3.1.2 Comparison against global CCN datasets

Figure 8 illustrates the global mean distribution of cloud condensation nuclei (CCN) at 0.2% supersaturation at 900 hPa, derived from MAMnet driven by MERRA-2 (shown as MAMnet-MERRA2), GiOcean (Song et al., 2025), the CAMS aerosol reanalysis (Block et al., 2024), and CALIOP satellite retrievals (Choudhury and Tesche, 2022). Data were averaged over the period $2006 - 2021$. All datasets reveal similar spatial patterns, with lower CCN concentrations over oceans, particularly in
295 polar regions, and higher concentrations over central and eastern Asia, Europe, and the Americas. However, large differences in
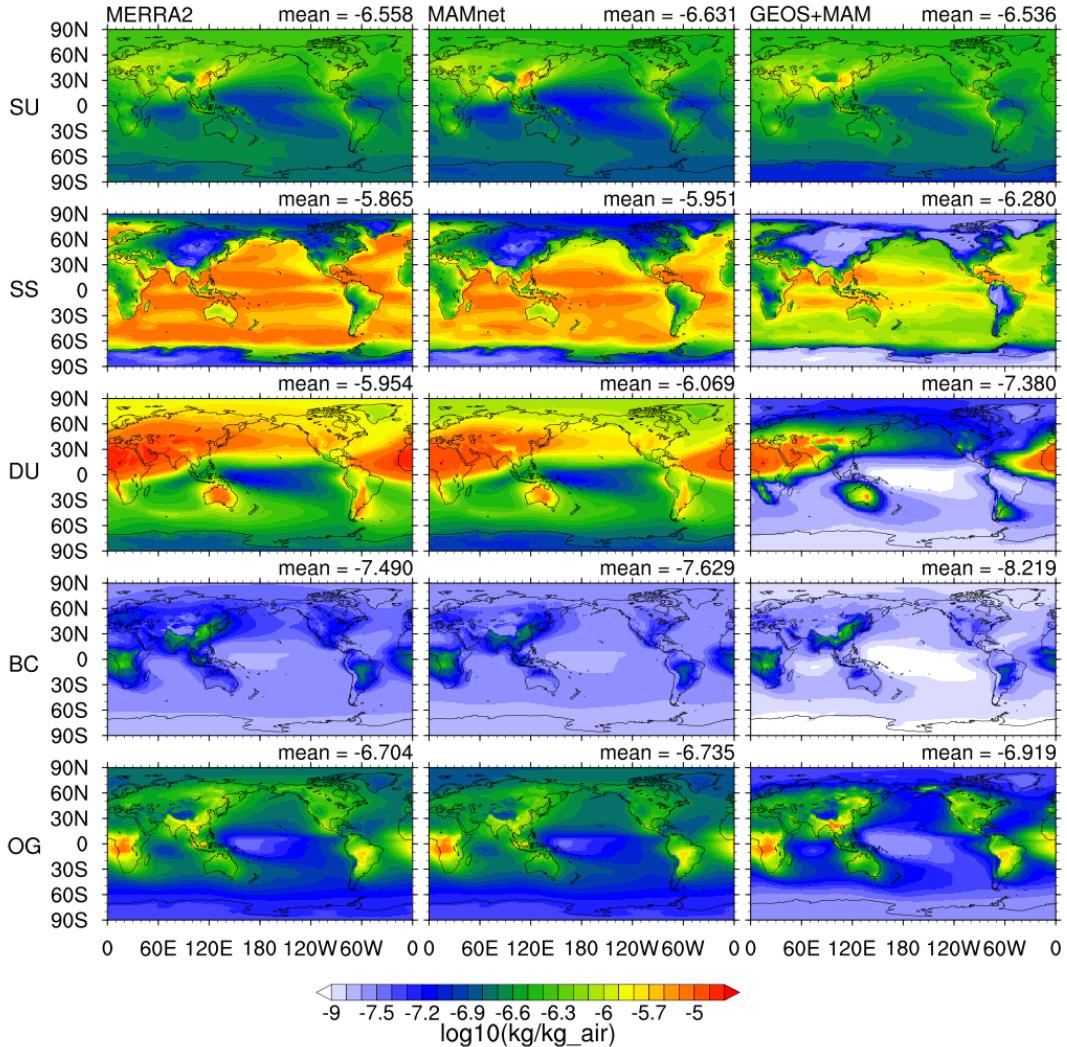
Figure 6. Comparison of column-integrated bulk aerosol (from top: sulfates, sea salt, dust, black carbon, organic matter) represented by (from left) MERRA-2, the trained MAMnet model applied to MERRA-2 inputs, and the reserved GEOS+MAM test data.
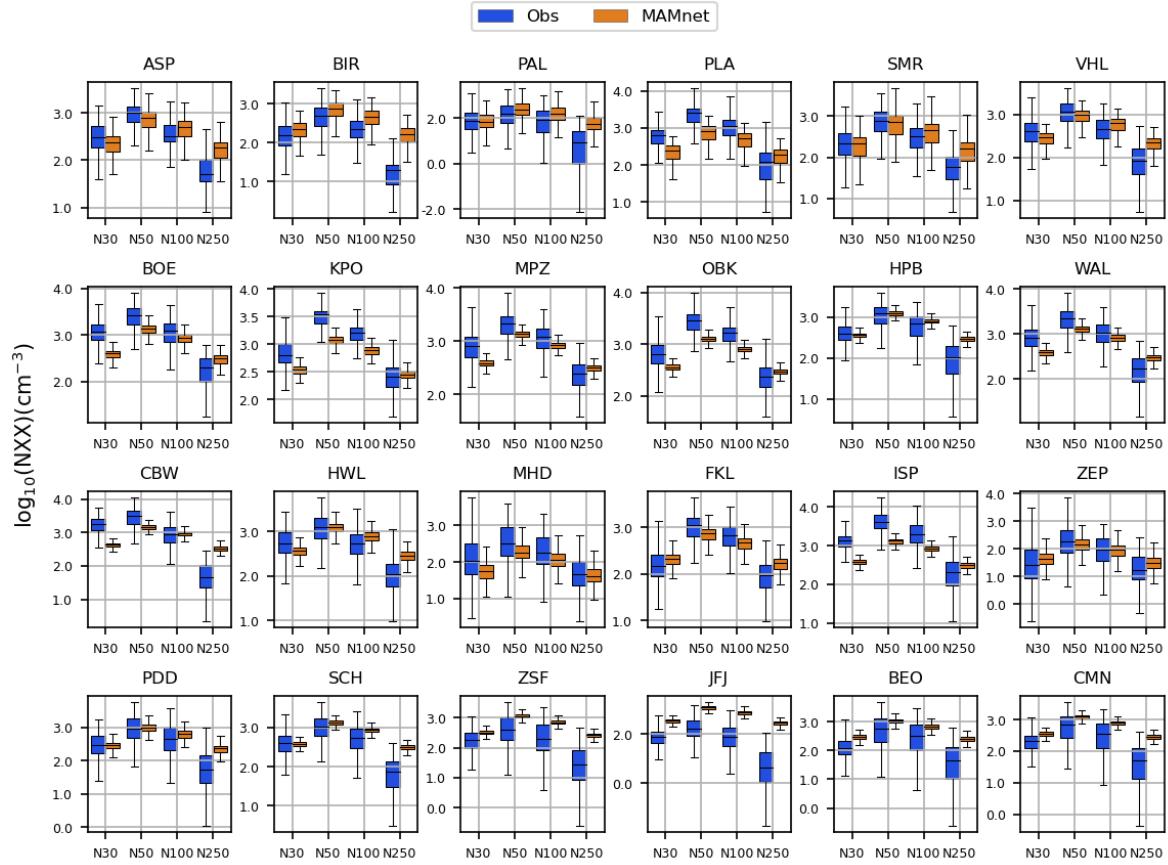
Figure 7. Cumulative size distribution comparison of the trained MAMnet model applied to MERRA-2 inputs against surface measurements (Asmi et al., 2011). The sites in the bottom row (PDD, SCH, ZSF, JFJ, BEO, CMN) are characterized as high altitude sites, with altitudes between 1200 and 3600 m.a.s.l.

absolute values are evident with MAMnet-MERRA2 consistently showing the lowest N_{CCN}, and CALIOP-derived the highest. These discrepancies likely arise from differing assumptions in estimation methods. GiOcean and CAMS estimate CCN based on aerosol mass, prescribing the ASD, and assuming externally-mixed aerosols, which may double-count CCN as organics and sulfates are typically internally mixed (Adachi and Buseck, 2008; Kirpes et al., 2018). MAMnet-MERRA2 avoids this issue but underpredicts N_{CCN} over oceanic regions, likely due to low sea salt concentrations in MERRA-2, stemming from uncertainties in the aerosol assimilation system (Buchard et al., 2017). In contrast, CALIOP may overestimate N_{CCN} from the assumption of CCN as all soluble aerosols above 50 nm, some of which may not activate as CCN at 0.2% supersaturation.

Figure 9 compares global mean vertical profiles of N_{CCN} from the datasets in Figure 8 and from in-situ observations (Watson-Parris et al., 2019). Vertical distributions vary significantly. GiOcean, CALIOP, and in-situ profiles exhibit similar shapes with peak concentrations around 950 hPa, while the MAMnet-MERRA2 and CAMS profiles show a monotonic de-
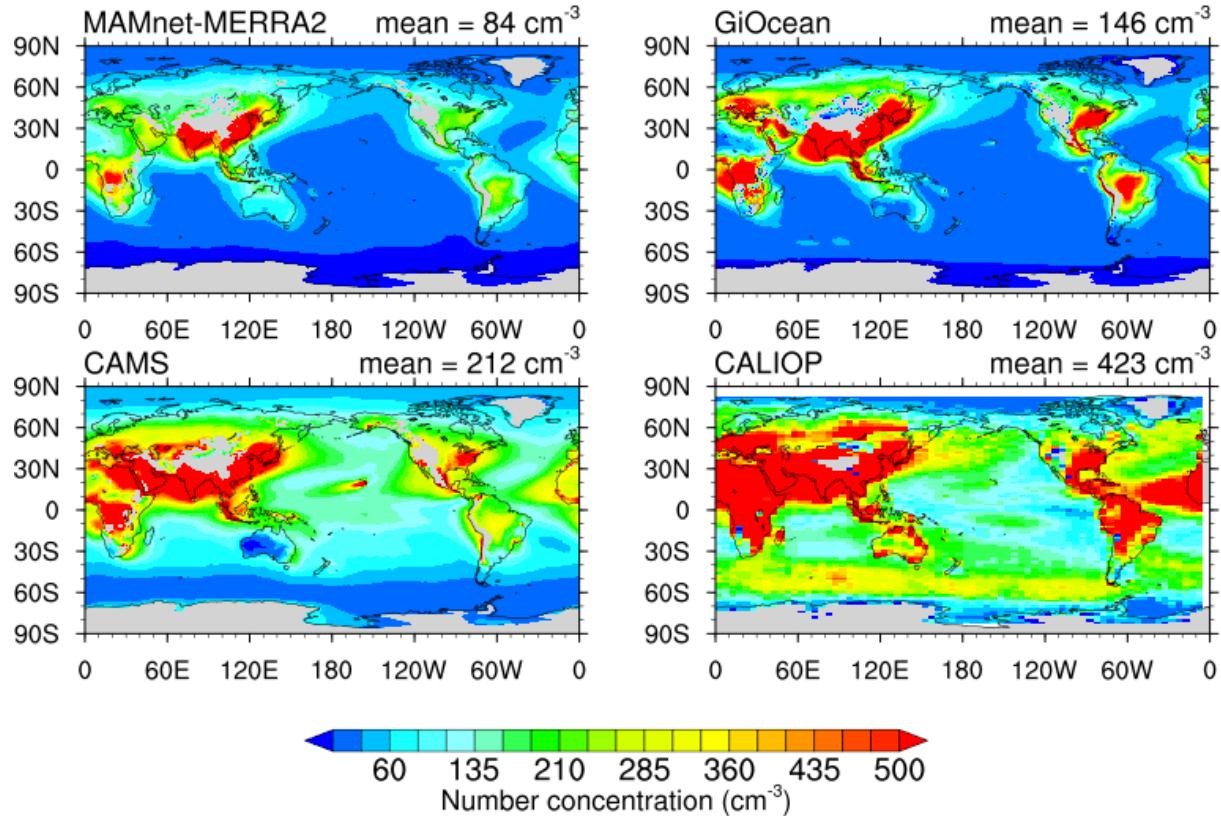
Figure 8. MAMnet-derived CCN at 0.2% supersaturation at 900 hPa using the MERRA2 reanalysis (MAMnet-MERRA2) against global CCN datasets. Also shown are results from the GiOcean reanalysis (Song et al., 2025), CALIOP- (Choudhury and Tesche, 2022), and CAMS- (Block et al., 2024) derived CCN.

crease with altitude. The peak in N_{CCN} at 950 hPa may result from more efficient aerosol scavenging near the surface, better represented by two-moment cloud microphysics in GiOcean (Song et al., 2025; Barahona et al., 2014). In contrast, MAMnet-MERRA2 and CAMS rely on single-moment cloud microphysics, which may explain the smoother decrease in N_{CCN} with height. The reliance on single-moment microphysics may also explain the more gradual decrease in N_{CCN} with height in

310   CAMS and MAMnet-MERRA2 than in the other data sets, noticeable over the ocean. In the free troposphere, MAMnet-MERRA2 tends align more closely with GiOcean, CALIOP and the in situ data.

3.2 Explainable machine learning analysis

Shapley values were used to elucidate the relative influence of input features on the model's decision-making process (Winter, 2002). Figure 10 is a summary plot of Shapley values calculated for each input feature relative to predicted targets for aerosol

315   modal number concentration (left) and mass (right). Each row represents a specific feature. The x-axis represents SHAP values,
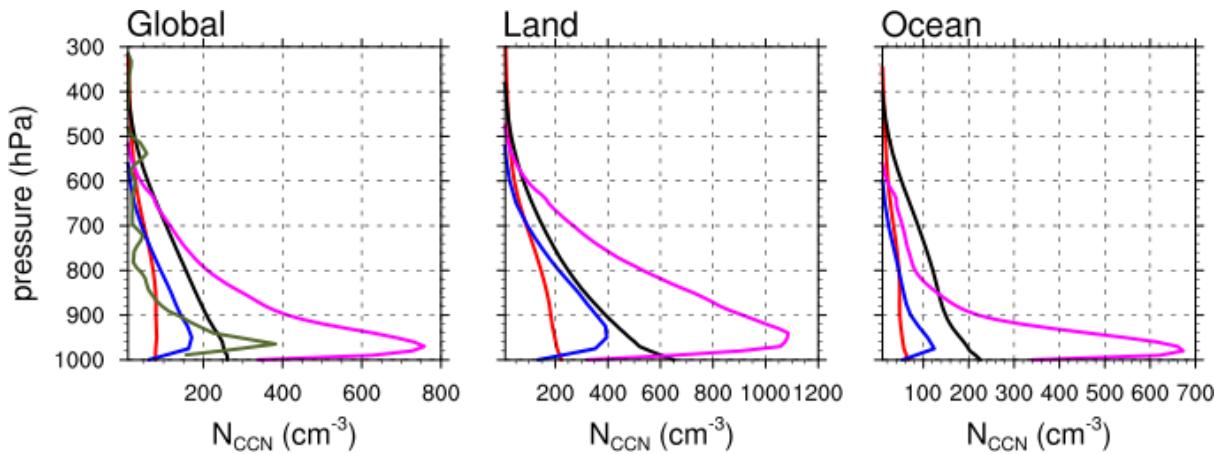
Figure 9. Annual mean profile of CCN concentration derived from MERRA2 using MAMnet (red). Also shown are CALIOP-derived CCN (magenta; Choudhury and Tesche, 2022), the GiOcean reanalysis (blue; Song et al., 2025), CCN derived from field campaign data around the globe (green; Watson-Parris et al., 2019), and CAMS-derived CCN (black; Block et al., 2024).

indicating the impact (positive or negative) of each feature on the model's prediction, so that the features with larger SHAP values contribute more significantly to the model output. Red dots represent high feature values, while blue dots indicate low feature values.

Features such as sulfate (SU), sea salt (SS), dust (DU), temperature (T), and air density (AIRD) are consistently ranked as dominant contributors, with their relative importance varying across aerosol modes. Fine-mode outputs, such as ACC and AIT, are strongly influenced by sulfate and temperature, where higher feature values positively impact predictions. In contrast, coarse-mode outputs like CDU and FDU are heavily driven by dust, with significant positive contributions observed for high dust concentrations. Intuitively, this makes sense because as SU, SS, and DU are the largest components of accumulation mode aerosols. However the relation is non-linear as high SU values correspond to a strong positive impact on aerosol number, particularly for dust (CDU) and sea salt (CSS) modes, whereas low values lead to neutral or negative contributions. The SHAP values further highlight the critical role of air density and temperature in for CSS and FSS, which may be related to the aerosol activation processes.

For aerosol mass, the SHAP plots show significant influence from DU, SS, BC, SU. The interplay between feature importance and values is evident as, high sea salt concentrations (SS) are positively correlated with increased mass in CSS, while low values lead to neutral or negative contributions. One significant characteristic of the mass SHAP plots is the broader range of SHAP values compared to number concentration, indicating greater variability in the importance of input features for predicting mass. For instance, black carbon (BC) has a consistently positive influence on ACC and AIT mass predictions, but its impact is less pronounced for other modes. Additionally, temperature at low values negatively impacts aerosol mass, while at higher values, it positively influences mass, possibly by promoting secondary aerosol formation. In some cases however a SHAP value for a feature may have no obviously interpretable significance to the target prediction, or may be so strongly

correlated with another feature that the individual contribution is negligible with respect to the feedbacks between a pair of features or more (Aas et al., 2020). Overall, the SHAP analysis identifies key physical drivers that govern aerosol behavior across fine and coarse scales in MAMnet.

4 Conclusions

340 This study develops a neural network, termed MAMnet, to predict the aerosol size distribution and mixing state using as input the bulk mass of different aerosol species, temperature and density. MAMnet is oriented towards allowing a better estimation of the ASD and the aerosol physicochemical properties in cases where computational cost considerations prevent the usage two and higher moment aerosol microphysics schemes, for instance, weather forecast, or where limited information is available to constraint the ASD as in remote sensing and data assimilation.

345 MAMnet was designed to reproduce the output of the MAM model (Liu et al., 2012). The neural network was optimized for performance taking into account the model architecture, training parameters and the rank of the data used as input. We performed a comprehensive evaluation of the NN model against simulated data and observations, showing that MAMnet is a robust and accurate model over a wide set of conditions. Importantly, MAMnet reproduces MERRA-2 aerosol concentrations when driven by MERRA-2 inputs, demonstrating that it has learned physical relationships rather than memorizing the training

350 data. Explainable machine learning analysis showed that MAMnet identifies key physical drivers and the non-linear behavior governing the aerosol distribution across fine and coarse scales.

Comparison of MAMnet predictions against a reference dataset from GEOS+MAM simulations exhibited remarkable agreement, with log-mean residuals typically below 0.1 and spatial correlation typically exceeding 0.9 for all aerosol modes. The greatest discrepancies were observed near the surface and in regions with low aerosol concentrations i.e., fine dust over oceans

355 and coarse dust in the Southern Hemisphere free troposphere. These discrepancies are primarily attributed to challenges in the prediction of concentrations near zero and class imbalance. Notably, biases in number and mass concentrations do not significantly influence the prediction of geometric mean diameter, which tends to compare exceedingly well against the GEOS+MAM simulations, indicating that MAMnet captures the physical relationship between aerosol mass and number, inherently conserving mass.

360 We took advantage of the fact that MAMnet can be driven by output from reanalysis data to evaluate its performance against observations. When driven using collocated MERRA-2 fields, MAMnet reasonably reproduced the measured aerosol size distribution at different ground observation sites, representing a variety of aerosol composition, origin and meteorological conditions. The median values of the predicted concentrations were generally consistent with observations. However the range of values predicted by MAMnet is in general smaller than observed, indicating that the model underestimates variability.

365 It is likely that coarse reanalysis inputs, limited training data diversity, and the complexities of aerosol evolution, not modeled explicitly by MAMnet, contributed to the observed discrepancies. CCN concentrations derived from MAMnet using the MERRA2 dataset were within the range of reported values but showed discrepancy near the surface and in regions with high variability, which may originate from uncertainty in the MERRA-2 aerosol fields. As MAMnet reproduces well the training
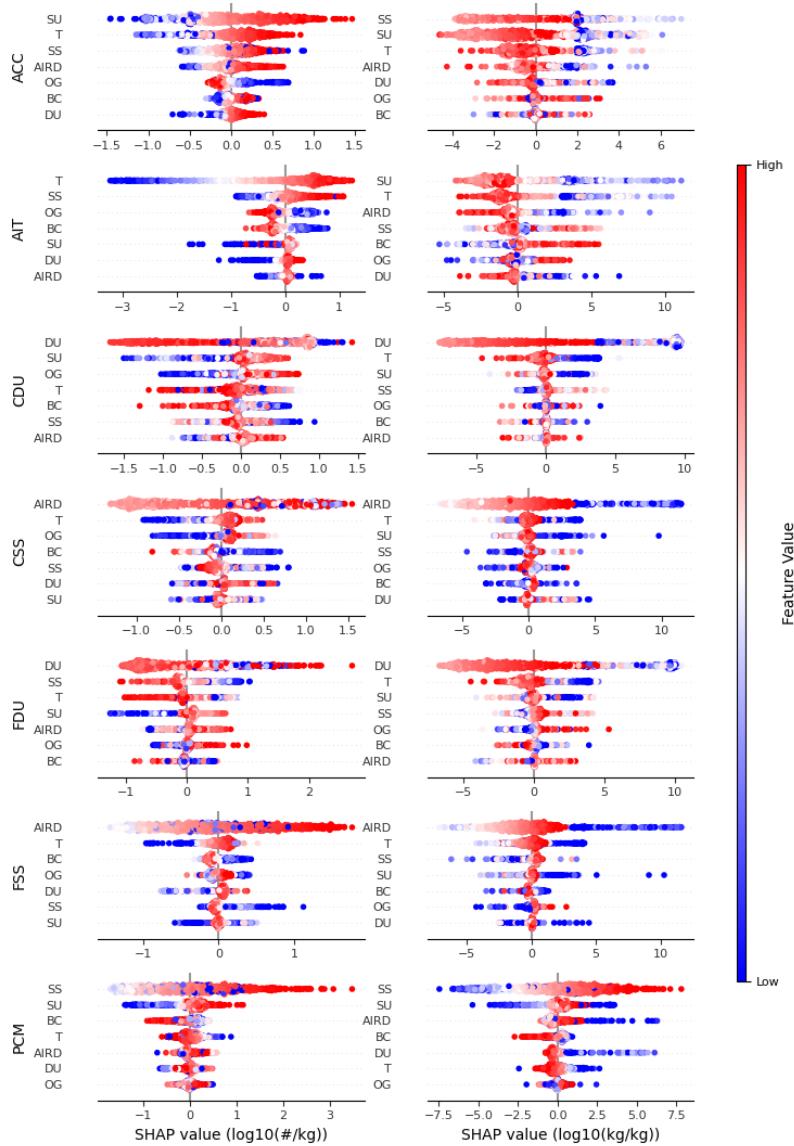
Figure 10. SHAP analysis for aerosol number concentration (left) and total mass (right) across different modes in the troposphere, based on 1,000 randomly selected samples from the test set. Modes are displayed from top to bottom: Accumulation (ACC), Aitken (AIT), coarse dust (CDU), coarse sea salt (CSS), fine dust (FDU), fine sea salt (FSS), and primary carbon matter (PCM). The color gradient (red for high values, blue for low values) indicates the relative value of each feature, with features ordered top-to-bottom by their importance to the prediction (most sensitive at the top). The x-axis represents SHAP values, quantifying how much each feature contributes to deviations from the mean prediction.

dataset it is likely that the biases against observations result from biases in the input and complex physics not modelled by MAM. Despite such biases, the comparison against observations indicate that MAMnet is able to capture the aerosol size distribution on regional and global scales.

Strategies to address the remaining biases include applying physical constraints via transfer learning, as well as including observational data during the training process (Barahona et al., 2024). Class imbalance can potentially be addressed by reconfiguring the NN such that each mode would be predicted by a separate layer, or even individual NNs. The latter option is less desirable because it would require developing, constraining, and maintaining multiple NNs as opposed to one. Future work would focus on applying MAMnet to elucidate long-term trends in the ASD as well as on its implementation with GCMs (Ott et al., 2020). The model developed here provides a versatile foundation to improve the physical representation of aerosols in weather forecasting, remote sensing and data assimilation, potentially enhancing our undesrtanding of their role in the climate system.

Code and data availability. The MERRA-2 Reanalysis is publicly available from https://disc.gsfc.nasa.gov/ (GMAO, 2015). The GEOS source code is available under the NASA Open Source Agreement at https://github.com/GEOS-ESM (GMAO, 2025a). The MAMnet model and sample test data can be downloaded at https://github.com/dbarahon/MAMnet (Barahona and Breen, 2025). CAMS data was obtained from Block (2023). Data from the GiOcean reanalysis was downloaded from https://portal.nccs.nasa.gov/datashare/gmao/geos-s2s-3/GiOCEAN_e1/ (GMAO, 2025b). CALIOP data was obtained from CALIPSO (2023).

Author contributions. D.B. conceived and directed the work. K.H.B. co-developed of the neural network model. A.D. implemented the MAM model within GEOS. K.B. Provided CCN data for comparison.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center. Keras and Tensorflow libraries were obtained from https://keras.io/. Maps were created using the NCAR Command Language (Version 6.6.2) Software. (2019). Boulder, Colorado: UCAR/NCAR/CISL/TDD. http://dx.doi.org/10.5065/D6WD3XH5. The SHAP python package was used to conduct the explainable machine learning analysis as described in https://shap-lrjball.readthedocs.io/en/latest/index.html. This work was supported by the NASA Modeling, Analysis and Prediction program, Grant NNH20ZDA001N-MAP.

References

395   Aas, K., Jullum, M., and Løland, A.: Explaining individual predictions when features are dependent: More accurate approximations to Shapley values, 2020.

Adachi, K. and Buseck, P.: Internally mixed soot, sulfates, and organic matter in aerosol particles from Mexico City, Atmospheric Chemistry and Physics, 8, 6469–6481, https://doi.org/10.5194/acp-8-6469-2008, 2008.

Adams, P. J. and Seinfeld, J. H.: Predicting global aerosol size distributions in general circulation models, Journal of Geophysical Research:
400   Atmospheres, 107, AAC–4, 2002.

Amunsen, C., Hanssen, J., Semb, A., and Steinnes, E.: Long-range atmospheric transport of trace elements to southern Norway, Atmospheric Environment. Part A. General Topics, 26, 1309–1324, 1992.

Aquila, V., Hendricks, J., Lauer, A., Riemer, N., Vogel, H., Baumgardner, D., Minikin, A., Petzold, A., Schwarz, J., Spackman, J., et al.:
MADE-in: A new aerosol microphysics submodel for global simulation of insoluble particles and their mixing state, Geoscientific Model
405   Development, 4, 325–355, 2011.

Arfin, T., Pillai, A. M., Mathew, N., Tirpude, A., Bang, R., and Mondal, P.: An overview of atmospheric aerosol and their effects on human health, Environmental Science and Pollution Research, 30, 125 347–125 369, 2023.

Asmi, A., Wiedensohler, A., Laj, P., Fjaeraa, A.-M., Sellegri, K., Birmili, W., Weingartner, E., Baltensperger, U., Zdimal, V., Zikova, N.,
et al.: Number size distributions and seasonality of submicron particles in Europe 2008–2009, Atmospheric Chemistry and Physics, 11,
410   5505–5538, 2011.

Barahona, D. and Breen, K.: MAMnet, https://github.com/dbarahon/MAMnet, 2025.

Barahona, D., Molod, A., Bacmeister, J., Nenes, A., Gettelman, A., Morrison, H., Phillips, V., and Eichmann, A.: Development of two-moment cloud microphysics for liquid and ice within the NASA Goddard Earth Observing System Model (GEOS-5), Geosc. Model Dev., 7, 1733–1766, https://doi.org/10.5194/gmd-7-1733-2014, 2014.

415   Barahona, D., Breen, K. H., Kalesse-Los, H., and Röttenbacher, J.: Deep Learning Parameterization of Vertical Wind Velocity Variability via Constrained Adversarial Training, Artificial Intelligence for the Earth Systems, 3, e230 025, https://doi.org/10.1175/AIES-D-23-0025.1, 2024.

Bender, F. A.-M.: Aerosol forcing: Still uncertain, still relevant, AGU advances, 1, e2019AV000 128, 2020.

Bender, F. A.-M., Frey, L., McCoy, D. T., Grosvenor, D. P., and Mohrmann, J. K.: Assessment of aerosol–cloud–radiation correlations in
420   satellite observations, climate models and reanalysis, Climate Dynamics, 52, 4371–4392, 2019.

Bengio, Y., Goodfellow, I., and Courville, A.: Deep learning, vol. 1, MIT press Cambridge, MA, USA, 2017.

Birmili, W., Berresheim, H., Plass-Dülmer, C., Elste, T., Gilge, S., Wiedensohler, A., and Uhrner, U.: The Hohenpeissenberg aerosol formation experiment (HAFEX): a long-term study including size-resolved aerosol, H 2 SO 4, OH, and monoterpenes measurements, Atmospheric Chemistry and Physics, 3, 361–376, 2003.

425   Birmili, W., Weinhold, K., Nordmann, S., Wiedensohler, A., Spindler, G., Müller, K., Herrmann, H., Gnauk, T., Pitz, M., Cyrys, J., et al.:
Atmospheric aerosol measurements in the German ultrafine aerosol network (GUAN), Gefahrst. Reinhalt. L, 69, 137–145, 2009.

Block, K.: Cloud condensation nuclei (CCN) numbers derived from CAMS reanalysis EAC4 (Version 1),
https://doi.org/10.26050/WDCC/QUAERERE_CCNCAMS_v1, 2023.

Block, K., Haghighatnasab, M., Partridge, D. G., Stier, P., and Quaas, J.: Cloud condensation nuclei concentrations derived from the CAMS
430   reanalysis, Earth System Science Data, 16, 443–470, https://doi.org/10.5194/essd-16-443-2024, 2024.

Brenowitz, N. D. and Bretherton, C. S.: Spatially extended tests of a neural network parametrization trained by coarse-graining, Journal of Advances in Modeling Earth Systems, 11, 2728–2744, 2019.

Buchard, V., Randles, C., Da Silva, A., Darmenov, A., Colarco, P., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A., Ziemba, L., et al.: The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies, Journal of Climate, 30, 6851–6872, 2017.

435 Buda, M., Maki, A., and Mazurowski, M. A.: A systematic study of the class imbalance problem in convolutional neural networks, Neural networks, 106, 249–259, 2018.

CALIPSO: Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observation Lidar Level 2 Aerosol Profile V4-20, https://doi.org/10.5067/CALIOP/CALIPSO/LID_L2_05KMAPRO-STANDARD-V4-20, 2023.

Červenkova, J. and Vá ňa, M.: Trend Assessment of deposition, throughfall and runoff water chemistry at the ICP-IM station Kosetice, Czech
440 Republic, IAHS-AISH publication, 336, 103–108, 2010.

Charron, A., Birmili, W., and Harrison, R. M.: Factors influencing new particle formation at the rural site, Harwell, United Kingdom, Journal of Geophysical Research: Atmospheres, 112, 2007.

Chin, M., Rood, R. B., Lin, S.-J., Müller, J.-F., and Thompson, A. M.: Atmospheric sulfur cycle simulated in the global model GOCART: Model description and global properties, Journal of Geophysical Research: Atmospheres, 105, 24 671–24 687, 2000.

445 Chollet, F. et al.: Keras, https://github.com/fchollet/keras, 2015.

Choudhury, G. and Tesche, M.: Estimating cloud condensation nuclei concentrations from CALIPSO lidar measurements, Atmospheric Measurement Techniques, 15, 639–654, 2022.

Christensen, M. W., Jones, W. K., and Stier, P.: Aerosols enhance cloud lifetime and brightness along the stratus-to-cumulus transition, Proceedings of the National Academy of Sciences, 117, 17 591–17 598, 2020.

450 Chu, D., Kaufman, Y., Ichoku, C., Remer, L., Tanré, D., and Holben, B.: Validation of MODIS aerosol optical depth retrieval over land, Geophysical research letters, 29, MOD2–1, 2002.

Colarco, P., da Silva, A., Chin, M., and Diehl, T.: Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, J. Geophys. Res., 115, D14 207–, http://dx.doi.org/10.1029/2009JD012820, 2010a.

455 Colarco, P., da Silva, A., Chin, M., and Diehl, T.: Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, Journal of Geophysical Research: Atmospheres, 115, 2010b.

Engler, C., Rose, D., Wehner, B., Wiedensohler, A., Brüggemann, E., Gnauk, T., Spindler, G., Tuch, T., and Birmili, W.: Size distributions of non-volatile particle residuals (D p< 800 nm) at a rural site in Germany and relation to air mass origin, Atmospheric Chemistry and Physics, 7, 5785–5802, 2007.

460 Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D. W., Haywood, J., Lean, J., Lowe, D. C., Myhre, G., et al.: Changes in atmospheric constituents and in radiative forcing. Chapter 2, in: Climate change 2007. The physical science basis, 2007.

Fountoukis, C. and Nenes, A.: Continued development of a cloud droplet formation parameterization for global climate models, Journal of Geophysical Research: Atmospheres, 110, 2005.

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

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S.-J.: Sources and distributions of dust aerosols simulated
470   with the GOCART model, Journal of Geophysical Research: Atmospheres, 106, 20 255–20 273, 2001.

GMAO: inst3$_3$$d_a$$sm_C$$p$: $MERRA-23DIAUState, MeteorologyInstantaneous3-hourly(p-coord, 0.625x0.5L42), https : //doi.org/10.5067/WWQSXQ8IVFW8, 2015.$

GMAO: GEOS, https://github.com/GEOS-ESM, 2025a.

GMAO: GiOcean Coupled Reanalysis, https://portal.nccs.nasa.gov/datashare/gmao/geos-s2s-3/GiOCEAN_e1/, 2025b.

Gong, X., Wex, H., Müller, T., Henning, S., Voigtländer, J., Wiedensohler, A., and Stratmann, F.: An unsupervised machine-learning-based
    classification of aerosol microphysical properties over 10 years at Cabo Verde, Atmospheric Chemistry and Physics Discussions, pp. 1–27,
475   2021.

Gruening, C., Adam, M., Cavalli, F., Cavalli, P., Dell'Acqua, A., Martins Dos Santos, S., Pagliari, V., Roux, D., and Putaud, J.: JRC Is-
    pra EMEP–GAW Regional Station for Atmos. Res, Tech. Rep. JRC55382, European Commission, http://publications.jrc.ec.europa.eu/
    repository/handle/111111111/538, 2009.

Harder, P., Watson-Parris, D., Stier, P., Strassel, D., Gauger, N. R., and Keuper, J.: Physics-Informed Learning of Aerosol Microphysics,
480   arXiv preprint arXiv:2207.11786, 2022.

Hari, P., Nikinmaa, E., Pohja, T., Siivola, E., Bäck, J., Vesala, T., and Kulmala, M.: Station for measuring ecosystem-atmosphere relations:
    SMEAR, Physical and physiological forest ecology, pp. 471–487, 2013.

Herzog, M., Weisenstein, D. K., and Penner, J. E.: A dynamic aerosol module for global chemical transport models: Model description,
    Journal of Geophysical Research: Atmospheres, 109, 2004.

485   Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flem-
    ming, J., et al.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, 2019.

Japkowicz, N. and Stephen, S.: The class imbalance problem: A systematic study, Intelligent data analysis, 6, 429–449, 2002.

Jennings, S., O'Dowd, C., O'Connor, T., and McGovern, F.: Physical characteristics of the ambient aerosol at Mace Head, Atmospheric
    Environment. Part A. General Topics, 25, 557–562, 1991.

490   Jones, A., Roberts, D., and Slingo, A.: A climate model study of indirect radiative forcing by anthropogenic sulphate aerosols, Nature, 370,
    450–453, 1994.

Jurányi, Z., Gysel, M., Weingartner, E., Bukowiecki, N., Kammermann, L., and Baltensperger, U.: A 17 month climatology of the cloud
    condensation nuclei number concentration at the high alpine site Jungfraujoch, Journal of Geophysical Research: Atmospheres, 116,
    2011.

495   Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

Kirpes, R. M., Bondy, A. L., Bonanno, D., Moffet, R. C., Wang, B., Laskin, A., Ault, A. P., and Pratt, K. A.: Secondary sulfate is in-
    ternally mixed with sea spray aerosol and organic aerosol in the winter Arctic, Atmospheric Chemistry and Physics, 18, 3937–3949,
    https://doi.org/10.5194/acp-18-3937-2018, 2018.

Kiss, G., Varga, B., Galambos, I., and Ganszky, I.: Characterization of water-soluble organic matter isolated from atmospheric fine aerosol,
500   Journal of Geophysical Research: Atmospheres, 107, ICC–1, 2002.

Kreidenweis, S., Koehler, K., DeMott, P., Prenni, A., Carrico, C., and Ervens, B.: Water activity and activation diameters from hygroscopicity
    data-Part I: Theory and application to inorganic salts, Atmospheric Chemistry and Physics, 5, 1357–1370, 2005.

Kristensson, A., Dal Maso, M., Swietlicki, E., Hussein, T., Zhou, J., Kerminen, V.-M., and Kulmala, M.: Characterization of new particle formation events at a background site in Southern Sweden: relation to air mass history, Tellus B: Chemical and Physical Meteorology, 60, 330–344, 2008.

Langner, J. and Rodhe, H.: A global three-dimensional model of the tropospheric sulfur cycle, Journal of Atmospheric Chemistry, 13, 225–263, 1991.

Lee, L., Pringle, K., Reddington, C., Mann, G., Stier, P., Spracklen, D., Pierce, J., and Karslaw, K.: The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei, Atmospheric Chemistry and Physics, 13, 8879—-8914, https://doi.org/10.5194/acp-13-8879-2013, 2013.

Lihavainen, H., Kerminen, V.-M., Komppula, M., Hyvärinen, A.-P., Laakia, J., Saarikoski, S., Makkonen, U., Kivekäs, N., Hillamo, R., Kulmala, M., et al.: Measurements of the relation between aerosol properties and microphysics and chemistry of low level liquid water clouds in Northern Finland, Atmospheric Chemistry and Physics, 8, 6925–6938, 2008.

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., et al.: Toward a minimal representation of aerosols in climate models: Description and evaluation in the Community Atmosphere Model CAM5, Geoscientific Model Development, 5, 709–739, 2012.

Marinoni, A., Cristofanelli, P., Calzolari, F., Roccato, F., Bonafè, U., and Bonasoni, P.: Continuous measurements of aerosol physical parameters at the Mt. Cimone GAW Station (2165 m asl, Italy), Science of the total environment, 391, 241–251, 2008.

Mihalopoulos, N., Stephanou, E., Kanakidou, M., Pilitsidis, S., and Bousquet, P.: Tropospheric aerosol ionic composition in the Eastern Mediterranean region, Tellus B, 49, 314–326, 1997.

Molod, A., Takacs, L., Suarez, M., and Bacmeister, J.: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, Geosc. Model Dev., 8, 1339–1356, https://doi.org/10.5194/gmd-8-1339-2015, 2015.

Molod, A., Hackert, E., Vikhliaev, Y., Zhao, B., Barahona, D., Vernieres, G., Borovikov, A., Kovach, R. M., Marshak, J., Schubert, S., et al.: GEOS-S2S version 2: The GMAO high-resolution coupled model and assimilation system for seasonal prediction, Journal of Geophysical Research: Atmospheres, 125, e2019JD031 767, https://doi.org/10.1029/2019JD031767, 2020.

Nojarov, P., Ivanov, P., Kalapov, I., Penev, I., and Drenska, M.: Connection between ozone concentration and atmosphere circulation at peak Moussala, Theoretical and applied climatology, 98, 201–208, 2009.

O'Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al.: Keras Tuner, https://github.com/keras-team/keras-tuner, 2019.

Ott, J., Pritchard, M., Best, N., Linstead, E., Curcic, M., and Baldi, P.: A Fortran-Keras deep learning bridge for scientific computing, Scientific Programming, 2020, 2020.

Philippin, S., Laj, P., Putaud, J.-P., Wiedensohler, A., LEEUW, G. D., FJAERAA, A. M., PLATT, U., BALTENSPERGER, U., and FIEBIG, M.: EUSAAR-An unprecedented network of aerosol observation in Europe, Journal of Aerosol Research (Earozoru Kenkyu), 24, 78–83, https://doi.org/10.11203/jar.24.78, 2009.

Randles, C. A., da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., Shinozuka, Y., and Flynn, C. J.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation, Journal of Climate, 30, 6823–6850, https://doi.org/10.1175/JCLI-D-16-0609.1, 2017.

Rasp, S., Pritchard, M. S., and Gentine, P.: Deep learning to represent subgrid processes in climate models, Proceedings of the National Academy of Sciences, 115, 9684–9689, 2018.

Reddington, C., Carslaw, K., Stier, P., Schutgens, N., Coe, H., Liu, D., Allan, J., Pringle, K., Lee, L., Yoshioka, M., et al.: The Global Aerosol
540   Synthesis and Science Project (GASSP): measurements and modeling to reduce uncertainty, Bulletin of the American Meteorological
       Society, 98, 1857–1877, 2017.

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., and Wang, W.: An improved in situ and satellite SST analysis for climate, J.
       Climate, 15, 1609–1625, 2002.

Rienecker, M., Suarez, M., Todling, R., Bacmeister, J., Takacs, L., Liu, H.-C., Gu, W., Sienkiewicz, M., Koster, R., Gelaro, R., Stajner, I.,
545   and Nielsen, J.: The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0., vol. 27 of *Technical Report
       Series on Global Modeling and Data Assimilation*, NASA Goddard Space Flight Center, Greenbelt, MD, USA, 2008.

Russchenberg, H., Bosveld, F., Swart, D., ten BRINK, H., de LEEUW, G., Uijlenhoet, R., Arbesser-Rastburg, B., van der MAREL, H.,
       LIGTHART, L., Boers, R., et al.: Ground-based atmospheric remote sensing in the Netherlands: European outlook, IEICE Transactions
       on Communications, 88, 2252–2258, 2005.

550 Seinfeld, J. H. and Pandis, S. N.: Atmospheric chemistry and physics: from air pollution to climate change, John Wiley & Sons, 2016.

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., et al.:
       Improving our fundamental understanding of the role of aerosol- cloud interactions in the climate system, Proceedings of the National
       Academy of Sciences, 113, 5781–5790, 2016.

Silva, S. J., Ma, P.-L., Hardin, J. C., and Rothenberg, D.: Physically regularized machine learning emulators of aerosol activation, Geoscien-
555   tific Model Development, 14, 3067–3077, 2021.

Song, C., McCoy, D., Molod, A., and Barahona, D.: Signatures of aerosol-cloud interactions in GiOcean: A coupled global reanalysis with
       two-moment cloud microphysics, EGUsphere, 2025, 1–30, https://doi.org/10.5194/egusphere-2024-4108, 2025.

Stier, P., Feichter, J., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Ganzeveld, L., Tegen, I., Werner, M., Balkanski, Y., et al.: The aerosol-
       climate model ECHAM5-HAM, Atmospheric Chemistry and Physics, 5, 1125–1156, 2005.

560 Stier, P., van den Heever, S. C., Christensen, M. W., Gryspeerdt, E., Dagan, G., Saleeby, S. M., Bollasina, M., Donner, L., Emanuel, K.,
       Ekman, A. M., et al.: Multifaceted aerosol effects on precipitation, Nature Geoscience, 17, 719–732, https://doi.org/10.1038/s41561-024-
       01482-6, 2024.

Ström, J., Umegård, J., Tørseth, K., Tunved, P., Hansson, H.-C., Holmén, K., Wismann, V., Herber, A., and König-Langlo, G.: One year of
       particle size distribution and aerosol chemical composition measurements at the Zeppelin Station, Svalbard, March 2000–March 2001,
565   Physics and Chemistry of the Earth, Parts A/B/C, 28, 1181–1190, 2003.

Takacs, L. L., Suárez, M. J., and Todling, R.: The stability of incremental analysis update, Monthly weather review, 146, 3259–3275, 2018.

Tunved, P., Ström, J., and Hansson, H.-C.: An investigation of processes controlling the evolution of the boundary layer aerosol size distri-
       bution properties at the Swedish background station Aspvreten, Atmospheric Chemistry and Physics, 4, 2581–2592, 2004.

Ulevicius, V., Byčenkienė, S., Remeikis, V., Garbaras, A., Kecorius, S., Andriejauskienė, J., Jasinevičienė, D., and Mocnik, G.: Character-
570   ization of pollution events in the East Baltic region affected by regional biomass fire emissions, Atmospheric Research, 98, 190–200,
       2010.

Uno, I., Eguchi, K., Yumimoto, K., Takemura, T., Shimizu, A., Uematsu, M., Liu, Z., Wang, Z., Hara, Y., and Sugimoto, N.: Asian dust
       transported one full circuit around the globe, Nature Geoscience, 2, 557–560, 2009.

Venzac, H., Sellegri, K., Villani, P., Picard, D., and Laj, P.: Seasonal variation of aerosol size distributions in the free troposphere and residual
575   layer at the puy de Dôme station, France, Atmospheric Chemistry and Physics, 9, 1465–1478, 2009.

Watson-Parris, D., Schutgens, N., Reddington, C., Pringle, K. J., Liu, D., Allan, J. D., Coe, H., Carslaw, K. S., and Stier, P.: In situ constraints on the vertical distribution of global aerosol, Atmospheric Chemistry and Physics, 19, 11 765–11 790, 2019.

Wei, L., Lu, Z., Wang, Y., Liu, X., Wang, W., Wu, C., Zhao, X., Rahimi, S., Xia, W., and Jiang, Y.: Black carbon-climate interactions regulate dust burdens over India revealed during COVID-19, Nature communications, 13, 1839, 2022.

580 Whitby, E. R. and McMurry, P. H.: Modal aerosol dynamics modeling, Aerosol Science and Technology, 27, 673–688, 1997.

Wilson, J., Cuvelier, C., and Raes, F.: A modeling study of global mixed aerosol fields, Journal of Geophysical Research: Atmospheres, 106, 34 081–34 108, 2001.

Winter, E.: The shapley value, Handbook of game theory with economic applications, 3, 2025–2054, 2002.

Yu, S., Ma, P.-L., Singh, B., Silva, S., and Pritchard, M.: Two-step hyperparameter optimization method: Accelerating hyperparameter search

585 by using a fraction of a training dataset, Artificial Intelligence for the Earth Systems, 3, e230 013, 2024.

Zhang, H., Sharma, G., Dhawan, S., Dhanraj, D., Li, Z., and Biswas, P.: Comparison of discrete, discrete-sectional, modal and moment models for aerosol dynamics simulations, Aerosol Science and Technology, 54, 739–760, https://doi.org/10.1080/02786826.2020.1723787, 2020.