https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Global vertically resolved reconstruction of cloud droplet number concentration combining satellite retrievals, reanalysis data, and in situ observations

Donifan Barahona[1], Travis Aerenson[2], Edward Gryspeerdt[3], and Daniel McCoy[2]

[1]Global Modeling and Assimilation Office. NASA, Goddard Space Flight Center, Greenbelt, MD, USA
[2]Department of Atmospheric Science, University of Wyoming, Laramie, WY, USA
[3]Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom

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

Abstract.

This work develops a global estimate of vertically resolved cloud droplet number concentration (N_d) by combining reanalysis data, satellite retrievals, ground-based observations, and field campaign measurements. This is accomplished by developing a neural-network estimator, termed NDnet, that uses satellite-retrieved cloud properties to infer three-dimensional N_d fields. Training proceeds in two stages, with a prior network first emulating a reanalysis and then being refined using observations to impose a global observational constraint. Probabilistic objective functions account for observational and representativeness errors and provide an explicit estimate of uncertainty in N_d. Applied over the satellite record, NDnet produces a more gradual decrease of N_d with altitude than suggested by the reanalysis data and higher N_d at high latitudes. The resulting fields also show substantial vertical heterogeneity, particularly in tropical and subtropical regions and at high altitude. The three-dimensional N_d dataset can be compared directly with model output and provides a new constraint for evaluating aerosol–cloud interactions in atmospheric models.

1 Introduction

Aerosols influence clouds throughout their life cycle. As cloud condensation nuclei, they affect cloud droplet number concentration and size (Seinfeld et al., 2016), while also influencing cloud thermodynamics, optical properties, phase partitioning, and precipitation (Murray et al., 2012; Gordon et al., 2023). These aerosol–cloud interactions (ACIs) remain an important source of uncertainty in weather and climate models (Bangert et al., 2012; Ekman, 2014; Zhang et al., 2016; Bellouin et al., 2019). N_d is particularly important because it provides a direct link between aerosol activation and cloud properties, controlling the magnitude of the cloud response to aerosol perturbations (Grosvenor et al., 2018b; Seinfeld et al., 2016; Song et al., 2025). Modern atmospheric models represent this link using two-moment and higher-order microphysics schemes that connect droplet activation to aerosol properties and prognostically represent cloud particles and precipitation (e.g., Barahona et al., 2014; Danabasoglu et al., 2020; Guo et al., 2021).

Global observational estimates of N_d, however, remain poorly constrained. The standard approach derives N_d from passive-sensor retrievals of cloud optical depth and effective radius, under assumptions about the vertical structure and evolution of the

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

cloud (Grosvenor and Wood, 2014; Bennartz and Rausch, 2017). These estimates are sensitive to cloud heterogeneity, retrieval errors, viewing geometry, and the criteria used to select valid pixels (Grosvenor et al., 2018b; Arola et al., 2022; Gryspeerdt et al., 2022; Loveridge and Di Girolamo, 2024; Goren et al., 2025). Sampling is particularly important because the conditions that produce the most reliable retrievals are not representative of all clouds. As a result, different filtering choices can change the retrieved N_d climatology and the inferred strength of ACIs. Since many observational estimates of aerosol–cloud radiative forcing rely on satellite-derived N_d, these limitations propagate into estimates of ACI forcing (Jia et al., 2021).

A more fundamental limitation is that clouds are three-dimensional, whereas passive satellite retrievals provide a two-dimensional view of cloud-top and column properties. Effective-radius retrievals are radiatively weighted toward the upper part of the cloud (Platnick, 2000; Grosvenor et al., 2018a), while conventional N_d retrievals assume a simplified, approximately adiabatic cloud structure and provide a single value representative of the cloud layer. This makes direct comparison with model profiles ambiguous, because there is no unique pressure or height at which the retrieved N_d should be placed. In situ aircraft observations, however, show substantial vertical variability in droplet concentration and related cloud and aerosol properties (Cai et al., 2022; Mace, 2024). The interpretation of satellite-derived N_d is therefore uncertain outside relatively homogeneous cloud regimes, limiting its direct use for model evaluation (Song et al., 2025). A vertically resolved estimate of N_d would help overcome this limitation by providing a more direct basis for comparison with model output.

Machine learning provides a way to combine information that is difficult to represent with a single retrieval equation. Recent studies have used machine learning to separate aerosol and meteorological controls on N_d (Jia et al., 2023; Chen et al., 2022; Irfan et al., 2025; Yu et al., 2022) and to improve retrievals of N_d using a broader set of passive remote sensing channels (Gonzalez et al., 2025). These approaches can reduce some of the assumptions of conventional retrievals. However a model trained on simulations inherits information and biases from the physical model used for training. This suggests that neither simulations nor satellite observations alone provide a complete description of N_d. Instead, they provide complementary constraints that can be combined with direct observations.

Here we develop NDnet, a probabilistic estimator of the global three-dimensional distribution of N_d. NDnet combines a physically based prior derived from the GiOcean coupled aerosol–cloud reanalysis (Song et al., 2025) with MODIS (Moderate Resolution Imaging Spectroradiometer, Salomonson et al., 1989) retrievals and direct constraints from ground-based sites and aircraft campaigns. The reanalysis provides vertical and spatial structure that cannot be observed directly from passive satellite measurements, MODIS constrains the reconstruction with the observed cloud state, and field observations provide independent information on N_d. By treating the reanalysis-derived state as a prior rather than as ground truth, this framework allows the available sources of information to be combined while retaining an estimate of the remaining uncertainty (Barahona et al., 2024).

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 1. Notation.

Symbol	Meaning
t	Time index (e.g., daily).
$\mathbf{x}$	Horizontal grid location (lat–lon cell).
z	Vertical coordinate / level (e.g., pressure or model level).
$k \in \mathcal{K}$	Ground-site index k in the set of sites $\mathcal{K}$.
$\mathbf{x}_k$	Grid location corresponding to site k.
$N_d(t, \mathbf{x}, z)$	Cloud droplet number concentration profile from reanalysis (training target).
$N_d^{\mathrm{obs}}(t, k, z)$	Observed (ground-site) N_{d} profile at site k.
$\overline{N}_d(z)$	Long-term mean N_{d} profile (climatology) used to define anomalies.
$N'_d(t, \mathbf{x}, z)$	Anomaly profile: $N'_d = N_d - \overline{N}_d$.
$\mathrm{EOF}_j(z)$	jth vertical empirical orthogonal function (EOF) basis function.
J	Number of retained EOF modes.
$\mathbf{A}(t, \mathbf{x}) \in \mathbb{R}^J$	EOF coefficient vector representing $N'_d(t, \mathbf{x}, z)$.
$A_j(t, \mathbf{x})$	jth component of $\mathbf{A}(t, \mathbf{x})$.
$\overline{\mathbf{A}}$	Long-term mean EOF coefficient vector (training climatology).
$\Delta\mathbf{A}(t, \mathbf{x})$	Residual coefficients: $\Delta\mathbf{A} = \mathbf{A} - \overline{\mathbf{A}}$.
$\mathbf{s}_{\mathrm{COSP}}(t, \mathbf{x})$	Simulated satellite retrieval features (COSP) paired with reanalysis for training.
$\mathbf{s}_{\mathrm{MODIS}}(t, \mathbf{x})$	Actual satellite retrieval features (MODIS) used for long-term inference.
$f_\theta(\cdot)$	Prior neural network ($\mathrm{NDnet}_{\mathrm{prior}}$) with parameters θ, predicting $\Delta\mathbf{A}$.
$\mathcal{L}_{\mathrm{xx}}$	Training loss (i.e., MSE in coefficient space, Gaussian, gradient penalty) against "xx" target.
∇	Gradient operator used in the gradient-penalty term (applied to outputs and targets consistently).
λ_{xx}	Weight for the xx-penalty term.
$\mathbf{A}_{\mathrm{base}}(t, \mathbf{x})$	Baseline EOF coefficients inferred from MODIS using the prior: $\overline{\mathbf{A}} + f_\theta(\mathbf{s}_{\mathrm{MODIS}})$.
$\mathbf{A}_{\mathrm{site}}(t, k)$	EOF coefficients projected from site observations at site k.
$\Delta\mathbf{A}_{\mathrm{site}}(t, k)$	Site residual relative to baseline: $\mathbf{A}_{\mathrm{site}}(t, k) - \mathbf{A}_{\mathrm{base}}(t, \mathbf{x}_k)$.
$g_\phi(\cdot)$	Correction head ($\mathrm{NDnet}_{\mathrm{corr}}$) with parameters ϕ.
$\boldsymbol{\mu}_{\Delta A}(t, \mathbf{x})$	Predicted mean correction in coefficient space.
$\boldsymbol{\sigma}_{\Delta A}(t, \mathbf{x})$	Predicted standard deviation of the correction in coefficient space (component-wise, positive).
$\mathcal{N}(\cdot; \mu, \Sigma)$	Gaussian distribution with mean μ and covariance Σ.
$\mathrm{diag}(\cdot)$	Diagonal matrix formed from a vector (diagonal covariance assumption).
$\hat{\mathbf{A}}(t, \mathbf{x})$	Final EOF coefficients: $\hat{\mathbf{A}} = \mathbf{A}_{\mathrm{base}} + \boldsymbol{\mu}_{\Delta A}$.
$\hat{\boldsymbol{\sigma}}_{\mathbf{A}}(t, \mathbf{x})$	Final coefficient uncertainty ($\boldsymbol{\sigma}_{\Delta A}$).
$\hat{N}_d(t, \mathbf{x}, z)$	Reconstructed mean N_{d} profile from $\hat{\mathbf{A}}$ and EOFs.
$\hat{\sigma}_{N_d}(t, \mathbf{x}, z)$	Reconstructed uncertainty of N_{d} propagated from $\hat{\boldsymbol{\sigma}}_{\mathbf{A}}$.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

2 Methods and Data

We use machine learning methods to synthesize several data sources into a coherent representation N_d. The data sources included a climate reanalisys, ground-based observations, satellite retrievals and field campaign data, as described below.

2.1 Data Sources

2.1.1 The GiOcean reanalysis

GiOcean is a coupled atmosphere–ocean reanalysis built around the NASA GEOS-S2S prediction system (Molod et al., 2020). Its coupled framework includes the GEOS atmospheric model (Molod et al., 2015), the MOM5 ocean model (Griffies et al., 2005), the CICE-4 sea ice model (Hunke, 2008), and a catchment land-surface model (Koster et al., 2005). The atmospheric component is run at approximately 50 km horizontal resolution, while the ocean uses a resolution of about 25 km. Ocean observations are assimilated with a Local Ensemble Transform Kalman Filter (Penny et al., 2013). Atmospheric and aerosol observations are assimilated every 6 h, whereas ocean observations are assimilated every 5 d over the period beginning in 1998. GiOcean follows a weakly coupled assimilation strategy in which the ocean and aerosol states are actively assimilated, while the atmospheric state is replayed from the GEOS-IT retrospective forecast (Gelaro et al., 2017). Aerosol assimilation is performed with the Goddard Aerosol Assimilation System (Buchard et al., 2017), with aerosol transport represented by the GOCART model (Colarco et al., 2010).

Cloud microphysics in GiOcean is represented with a two-moment scheme (Barahona et al., 2014) that predicts both condensate mass and droplet number concentration. This allows aerosol number and composition to affect cloud formation and evolution directly and provides an explicit representation of aerosol indirect effects. The scheme follows the microphysical framework of Morrison and Gettelman (2008), with droplet activation described by Abdul-Razzak and Ghan (2000), ice formation by Barahona and Nenes (2009), and vertical velocity constraints derived from high-resolution simulations (Barahona et al., 2017). The representation of the droplet size distribution allows the shape parameter (k) to vary with liquid water content and droplet number concentration following Liu et al. (2008). GiOcean was additionally run with the COSP MODIS simulator enabled (Bodas-Salcedo et al., 2008), providing simulated cloud optical depth, cloud fraction and droplet effective radius that can be consistently compared with MODIS retrievals.

For this study, GiOcean provides daily mean N_d fields on 48 pressure levels and a 1° horizontal grid for 1998–2025. Previous work has shown that its simulated N_d is broadly consistent with current satellite-derived estimates and reproduces observed long-term variations in N_d (Song et al., 2025). Although N_d itself is not assimilated and remains a product of parameterized cloud microphysics, the surrounding atmospheric and aerosol state is constrained by observations. The combination of model physics, data assimilation, and MODIS-like simulated observations makes GiOcean particularly useful as an initial training dataset. It provides a first-order, physically based estimate of N_d that can serve as a feature-extracting prior and subsequently be refined using direct observations. A similar approach has been successfully used in previous work to derive subgrid-scale dynamics (Barahona et al., 2025).

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

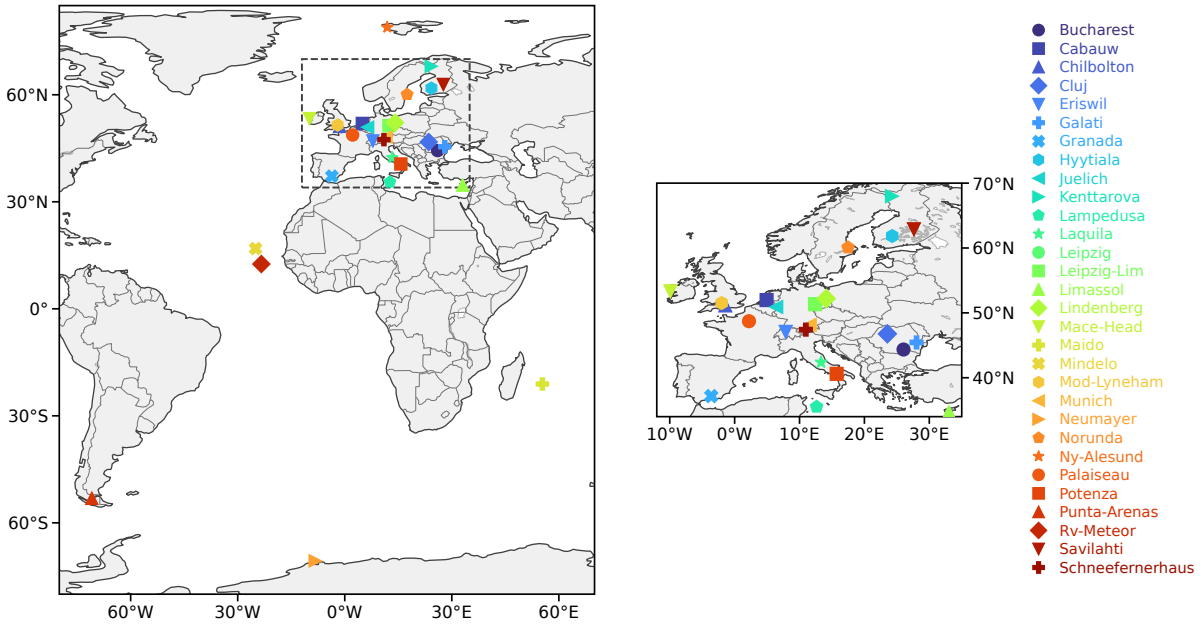


Figure 1. Ground-based sites used for refinement.

2.1.2 Satellite Retrievals

The Moderate Resolution Imaging Spectroradiometer MODIS (Salomonson et al., 1989) is the longest continuous record of space-based observations of cloud properties, and has been used in several studies to estimate N_{d}. We use the daily Level 3 gridded MODIS COSP-aligned cloud product MCD06COSP_D3_MODIS as the primary satellite input stream. This dataset aggregates Terra and Aqua MODIS cloud-mask, cloud-top, and cloud optical retrievals on a $1^{\circ} \times 1^{\circ}$ grid in a framework intended to facilitate comparisons to climate models and COSP outputs (NASA LAADS DAAC, 2022; NASA Earthdata, 2022; Pincus et al., 2023). Therefore, it is closely aligned with the outputs and retrieval framework of the COSP simulator implemented in GiOcean (Song et al., 2025).

2.1.3 Ground-based measurements

Ground-based retrievals of N_{d} are commonly designed to estimate a representative value for the cloud layer rather than its vertical profile. For example, the ARM NDROP product derives a layer-constant N_{d} from cloud optical depth and liquid water path, while the lidar-based retrieval evaluated by Zhang et al. (2023) estimates N_{d} close to cloud base. This assumption performs well for the single-layer marine stratiform clouds but that does not provide an independent constraint on $N_{\mathrm{d}}(z)$. Because our objective is to constrain the vertical structure of N_{d}, we instead use Cloudnet sites (Illingworth et al., 2007). These use a combination of radar, lidar, and microwave radiometry to estimate cloud microphysical properties (Rusli et al.,

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

2017; Frisch et al., 2002). N_{d} is then inferred from other microphysical variables, i.e., liquid water content, lwc, and droplet effective radius r_{e}, in the form:

$$N_{\mathrm{d}}(z) = \frac{3}{4\pi\rho_w k}\frac{lwc(z)}{r_{\mathrm{e}}(z)^3} \tag{1}$$

where ρ_w is the water density, z the altitude, and k the droplet dispersion factor. The profiles $lwc(z)$ and $r_{\mathrm{e}}(z)$ are considered pseudo-adiabatic, as an initial estimate is done using an adiabatic approximation which is then scaled to match the observed liquid water path (LWP), inferred from the microwave radiometry (Frisch et al., 2002). We assume $k = 0.8$, typical of stratocummulus regions (Grosvenor et al., 2018b). There is some variability in k which may introduce some uncertainty N_{d}. This is attenuated as k varies within a narrow range $(0.8 - 1.2)$ even in very diverse cases like convective and stratocumulus clouds. The resulting profiles remain retrieval-based and retain uncertainties associated with the assumed droplet size distribution and retrieved cloud properties, but they do not explicitly impose a vertically constant N_{d} and are therefore better suited to constrain the three-dimensional reconstruction, particularly when taken as an ensemble (Section 2.1.5).

Radar back-scattering is sensitive to the presence of drizzle (Frisch et al., 2002; Zhang et al., 2023) introducing uncertainty in the estimation of N_{d}. Compounded with the fact that r_{e} is cubed in Eq. 1, this makes N_{d} overly sensitive to errors in r_{e}. To account for this and minimize the effect of drizzle, only reliable retrievals, that is where the scene was identified as liquid water with no rain nor ice present, were used. Liquid water content was also limited to below 1 g m^{-3} to target persistent, non-drizzling clouds. Further error can result from uncertainty in the remote sensing retrievals and the estimation of cloud top height (Albrecht et al., 1990). These are reported as one standard deviation error around the observed mean. Assuming that errors are independent, and using Eq. 1 they can be propagated to error in N_{d} in the form,

$$\sigma_{\mathrm{N,exp}}(z) = N_{\mathrm{d}}(z)\sqrt{\left[\frac{\sigma_{lwc}(z)}{lwc(z)}\right]^2 + \left[3\frac{\sigma_{r_e}(z)}{r_e(z)}\right]^2} \tag{2}$$

where $\sigma_{\mathrm{N,exp}}$, σ_{lwc}, and σ_{r_e} denote the experimental uncertainties in N_{d}, lwc, and r_e, respectively. Uncertainty associated with assumed retrieval parameters, such as the coefficient k, is not treated as an independent measurement error, but rather as part of the structural uncertainty of the retrieval representation. Uncertainty in N_{d} is explicitly accounted for during training (Section 2.2).

We selected 30 sites from the CloudNet network (Illingworth et al., 2007), representing a diverse set of meteorological conditions and cloud formation regimes, as listed in Table 2 and depicted in Fig. 1. Measurements were reported on 30 s frequency. In-cloud data were aggregated into daily means. Each vertical profile was represented by an averaged of 10 vertical levels, and each daily mean included on average, 50-100 vertical N_{d} profiles. Data was interpolated on 21 pressure levels, from 250 to 1000 hpa using the collocated pressure and height fields from the MERRA-2 reanalysis (Gelaro et al., 2017). Each site was assigned to the closest GiOcean grid cell on the 1° grid. Data for most sites ranged between a few months to a few years. Experimental error was typically of the order of mean $0.5\bar{N}_{\mathrm{d}}$ signaling significant dispersion on the data (Table 2).

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 2. Site coverage diagnostics and N_{d} statistics for $p > 700$ hPa (Nd, error, ϵ, in cm^{-3}). Shown for each site are the number of valid days (Ndays), mean droplet concentration ($\overline{N_d}$), the total ($\bar{\epsilon}_{\mathrm{tot}}$, includes subgrid-scale variability) and experimental error $\bar{\epsilon}_{\mathrm{exp}}$, Eq. 2) relative to the mean, respectively, the average number of vertical profiles per 111-km segment (AvgProfiles111km), and the average number of vertical profiles used per valid day (AvgProfilesDay). Schneefernerhaus is omitted because there are no valid retrieval days for $p > 700$ hPa.

ID	Site	Ndays	Period	$\overline{N_d}$	$\bar{\epsilon}_{\mathrm{tot}}/\overline{N_d}$	$\bar{\epsilon}_{\mathrm{exp}}/\overline{N_d}$	AvgProfiles111km	AvgProfilesDay
BU	bucharest	614	19/05-25/10	161.2	0.47	0.43	25.6	52.0
CB	cabauw	618	21/05-25/10	135.2	0.49	0.41	26.2	99.0
CH	chilbolton	991	20/02-25/02	133.1	0.54	0.46	15.7	68.1
CJ	cluj	144	24/05-25/10	159.9	0.44	0.39	24.9	72.1
ER	eriswil	77	22/02-24/02	133.6	0.47	0.43	17.4	37.9
GA	galati	442	22/03-25/10	172.5	0.49	0.44	27.0	73.9
GR	granada	323	18/09-25/10	143.7	0.45	0.39	17.6	52.3
HY	hyytiala	1136	17/11-25/10	191.2	0.48	0.42	18.2	62.2
JU	juelich	2265	11/02-25/10	162.7	0.53	0.46	13.8	52.0
KT	kenttarova	276	22/08-25/10	153.0	0.45	0.40	22.4	62.7
LA	lampedusa	253	24/01-25/10	136.3	0.49	0.44	14.0	49.3
AQ	laquila	153	24/05-25/10	126.9	0.46	0.40	15.6	28.3
LP	leipzig	787	11/09-25/10	156.3	0.50	0.44	18.1	55.3
LL	leipzig-lim	361	20/12-25/02	153.8	0.48	0.42	23.0	79.9
LI	limassol	73	16/10-25/10	188.3	0.56	0.52	5.0	9.3
LB	lindenberg	1588	12/10-25/10	151.8	0.54	0.49	15.9	54.0
MH	mace-head	290	20/10-24/02	162.1	0.62	0.53	9.2	30.9
MD	maido	135	22/11-25/10	148.9	0.57	0.50	5.8	9.7
MN	mindelo	549	21/09-25/10	119.1	0.42	0.37	17.3	60.8
ML	mod-lyneham	34	23/06-23/09	132.3	0.64	0.57	3.5	10.5
MU	munich	1105	17/01-25/10	152.5	0.49	0.44	17.4	43.2
NE	neumayer	293	23/01-25/10	135.7	0.44	0.41	17.1	37.5
NO	norunda	932	19/03-25/10	151.8	0.46	0.41	22.0	67.5
NY	ny-alesund	1200	17/05-25/10	155.2	0.44	0.39	20.9	48.2
PA	palaiseau	2047	12/04-25/10	148.4	0.62	0.54	8.3	31.5
PO	potenza	257	23/08-25/10	169.9	0.49	0.46	11.6	27.7
PU	punta-arenas	333	18/11-21/11	164.6	0.51	0.47	8.5	20.5
RM	rv-meteor	37	24/08-24/09	179.7	0.69	0.60	11.3	36.0
SA	savilahti	52	20/09-20/11	204.0	0.45	0.38	28.3	121.5

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 3. Aircraft campaigns used in this study.

Campaign	Period	Reference	Cloud regime	Location
E-PEACE	07/11–08/11	Russell et al. (2013)	Marine stratocumulus; ship-track aerosol perturbations	Eastern Pacific off California
FASE	07/16	Sorooshian et al. (2018)	Fog and marine stratocumulus	Central California coast / NE Pacific
MACAWS	06/18–07/18	Zeider et al. (2024)	Marine low clouds with wildfire influence	Central California coast / NE Pacific
MASE1	07/05	Lu et al. (2007)	Marine stratocumulus	Off Monterey, California
MASE2	07/07	Lu et al. (2009)	Marine stratocumulus	Off Monterey, California
VOCALS-REx	10/08–11/08	Wood et al. (2011)	Subtropical stratocumulus deck	SE Pacific off Chile/Peru
ORACLES	09/16, 10/18	Redemann et al. (2021)	Stratocumulus under biomass-burning aerosol	SE Atlantic off Namibia
NAAMES	11/15–09/17	Behrenfeld et al. (2019)	Marine boundary-layer clouds	North Atlantic
ACTIVATE	2020	Sorooshian et al. (2019)	Shallow cumulus and postfrontal stratocumulus	W North Atlantic off the eastern USA
SOCRATES	01/18–02/18	McFarquhar et al. (2021)	Southern Ocean low clouds; mixed-phase	Southern Ocean south of Tasmania
COPE	07/13–08/13	Leon et al. (2016)	Convective clouds	Southwest England, UK

2.1.4 Field Campaign Data

Data collected during field campaigns was used to further inform the reconstruction of N_{d}. These are described in detail in Gryspeerdt et al. (2022), and summarized in Table 3. Unlike remote sensing data, aircraft measurements estimate N_{d} from direct sampling of the droplet size distribution, providing independent validation. Most of the field data correspond to marine stratocumulus decks in mid latitude regions, although campaigns like NAAMES (Behrenfeld et al., 2019) and SOCRATES (McFarquhar et al., 2021) cover the North Atlantic and the Southern Ocean, respectively. Similarly to Gryspeerdt et al. (2022) all the data was daily aggregated, screened for water content ($lwc > 0.1$ g m^{-3}) to avoid cloud edges and regions heavily influenced by entrainment. As with the ground-based data, measured N_{d} was interpolated to pressure levels, and assigned to the closest GiOcean grid cell.

2.1.5 On the representativeness of observational data

Representativeness is an inherent challenge in data-driven approaches based on sparse observations. Experimental measurements are episodic, often collected at high frequency, and require assumptions about how the sampled states relate to larger spatial scales. Ground-based measurements sample one location repeatedly in time, while field campaigns sample selected clouds along an aircraft track. Neither is directly equivalent to a model grid cell. In addition, pseudo-adiabatic retrievals of N_{d}

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

assume that droplet number concentration is approximately constant within a liquid-cloud layer. Although Cloudnet observations identify the vertical extent of the cloud and provide the quantities needed to estimate N_{d} at a given altitude, they should not be interpreted as exact instantaneous profiles of a single cloud column.

We treat each accepted retrieval as one sample of the cloudy air present at a given height. Over time, a fixed site samples many different clouds and cloud parcels passing over the instrument. These samples can therefore be combined to describe the typical in-cloud N_{d} at each level. We refer to this collection of samples as a local pseudo-adiabatic cloud ensemble (LPCE). This interpretation is consistent with how cloud properties are represented within a large model grid cell. A simulated grid cell does not correspond to a single cloud, but represents many cloudy subcolumns with different depths and properties. This view is also related to the motivation behind superparameterization, in which unresolved cloud-scale variability is represented within a large-scale model grid (Randall et al., 2003). Under the LPCE interpretation, localized observations can therefore be used to refine the reanalysis-trained neural network ("NN") as long as enough samples are available to represent the local cloud population (Section 2.2.2).

For the ground-based sites, we reduce the mismatch in scale by using the MERRA-2 horizontal wind speed to estimate the time required for an air mass to travel 100 km, approximately the scale of the model grid. Observations are first averaged over this period and then combined into daily means. Each grid-cell-scale average contains, on average, about 20 valid samples, while each site provides roughly 100 valid profiles per day (Table 2). Campaign data are treated in the same way conceptually, but the aircraft already samples different locations directly. Individual in-cloud measurement segments are assigned to the nearest model grid cell and vertical level and are then averaged within each grid cell and day. Measurements at different heights are not assumed to belong to the same cloud profile, but are instead treated as separate samples of the clouds encountered during the campaign. Because aircraft measurements target particular clouds and conditions, the resulting averages represent the sampled cloud conditions rather than all possible clouds within the grid cell.

2.2 Reconstruction Method

A neural network estimator, termed "NDnet", was trained to reconstruct N_{d} globally, synthesizing the datasets outlined in Section 2.1. First, reanalysis data was used to train NDnet to predict N_{d} as a function of MODIS-like variables from the COSP satellite simulator. Then, ground-based observations and in-situ measurements were used to refine the network accounting for error through a probabilistic approach. This refined network was used in inference mode to estimate the global distribution N_{d} using MODIS retrievals as input. This procedure is summarize in Table 4, and detailed in this section.

2.2.1 Development of a prior EOF-based neural network from reanalysis data: NDnet-prior

Simulated N_{d} and MODIS-like variables from the GiOcean reanalysis were used to train a neural network as a preliminary input to NDnet, termed "NDnet-prior". We selected four variables retrieved by MODIS (all for liquid clouds): optical thickness, τ_c, effective radius r_e, liquid water path (LWP), and cloud top pressure p_{cldtop}, all on daily means, interpolated to the GiOcean grid. Targets corresponded to the in-cloud N_{d} from the GiOcean reanalysis. MODIS cloud properties are available as daily, Level-3, two-dimensional (referred as "2D") horizontal gridded products. In contrast, N_{d} in GiOcean is resolved vertically and

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 4. N_{d} reconstruction workflow.

Step	Operation	Description
1	Data assembly	Assemble reanalysis–simulator pairs $N_d(t,\mathbf{x},z), \mathbf{s}_{\mathrm{COSP}}(t,\mathbf{x})$, MODIS inputs $\mathbf{s}_{\mathrm{MODIS}}(t,\mathbf{x})$, fixed-site profiles, and field-campaign observations.
2	EOF representation	Represent $N_d \approx \overline{N}_d + \sum_j A_j \mathrm{EOF}_j$, with $\mathbf{A} = \overline{\mathbf{A}} + \Delta\mathbf{A}$. The network predicts the time-varying residual coefficients $\Delta\mathbf{A}$.
3	Train the prior network	Train f_θ to predict residual coefficients from simulated satellite fields, $\boldsymbol{\mu}0 = f\theta(\mathbf{s}_{\mathrm{COSP}}) \approx \Delta\mathbf{A}$. The loss combines coefficient error, spatial-gradient and Laplacian terms, and error in reconstructed N_{d}.
4	Generate MODIS prior	Apply the trained prior to MODIS, $\boldsymbol{\mu}0 = f_\theta(\mathbf{s}_{\mathrm{MODIS}})$, giving $\mathbf{A}_{\mathrm{base}} = \overline{\mathbf{A}} + \boldsymbol{\mu}_0$. The residual fields $\boldsymbol{\mu}_0$ are stored for refinement.
5	Prior correction and uncertainty	Combine $\boldsymbol{\mu}_0$ with learned MODIS features and predict $\widehat{\Delta\mathbf{A}} = \boldsymbol{\mu}_0 + \boldsymbol{\delta}$ and $\boldsymbol{\sigma}_A = \mathrm{softplus}(\boldsymbol{\rho}) + \sigma_{\min}$. Refinement modifies the MODIS prior.
6	Fixed-site sampling	At each site k, sample $\widehat{\Delta\mathbf{A}}_k$ and $\boldsymbol{\sigma}_{A,k}$ at the corresponding grid cell. Observed profiles are projected onto the EOF basis to obtain coefficient targets $\mathbf{y}_k^A$.
7	Site coefficient loss	Use a Gaussian coefficient loss, $\mathcal{L}_{\mathrm{site}} = \mathrm{mean}_k[-\log\mathcal{N}(\mathbf{y}_k^A; \widehat{\Delta\mathbf{A}}_k, \mathrm{diag}(\boldsymbol{\sigma}_{A,k}^2))]$. Losses are normalized within sites, weighted by EOF mode, and averaged equally across sites.
8	Site profile losses	Reconstruct $\widehat{N}_{d,k} = \overline{N}_d + \sum_j(\overline{A}_j + \widehat{\Delta A}_{j,k})\mathrm{EOF}_j$, with $\sigma^2_{d,k} = \sum_j \sigma^2_{A_j,k}\mathrm{EOF}_j^2$. Use profile MSE and Gaussian losses, with $\sigma^2_{\mathrm{eff}} = \sigma^2_{d,k} + E_k^2$.
9	Dense regularization	Regularize the correction $\mathbf{C} = \widehat{\Delta\mathbf{A}} - \boldsymbol{\mu}_0$ using $\mathcal{L}_{\mathrm{anc}} = \mathrm{MSE}\,w(\mathbf{C},0)$. Site uncertainty is additionally constrained by $\mathcal{L}_\sigma = \mathrm{mean}(\sigma^2_{A_j}/V_j)$.
10	Campaign sampling	For each campaign observation, sample the predicted coefficients and uncertainties at its day and grid location. Reconstruct N_{d} and its variance and interpolate both to the observed pressure linearly in $\log p$.
11	Campaign losses	Use campaign MSE and Gaussian losses with $\sigma^2_{\mathrm{eff}} = \sigma_d^2 + E^2$. Losses are averaged within flights, then within campaigns, and finally equally across campaigns. Campaign uncertainty is also constrained using $\mathcal{L}_{\mathrm{camp},\sigma} = \mathrm{mean}(\sigma {A_j}^2/V_j)$.
12	Combined training loss	Train with $\mathcal{L} = \lambda_{\mathrm{site}}\mathcal{L}_{\mathrm{site}} + \lambda_{N_d}\mathcal{L}_{N_d} + \lambda_{N_d,\mathrm{G}}\mathcal{L}_{N_d,\mathrm{G}} + \lambda_{\mathrm{camp}}\mathcal{L}_{\mathrm{camp}} + \lambda_{\mathrm{camp,G}}\mathcal{L}_{\mathrm{camp,G}} + \lambda_{\mathrm{anc}}\mathcal{L}_{\mathrm{anc}} + \lambda_\sigma\mathcal{L}_\sigma + \lambda_{\mathrm{camp},\sigma}\mathcal{L}_{\mathrm{camp},\sigma}$. Unavailable observation terms contribute zero.
13	Final reconstruction	The final coefficients are $\widehat{\mathbf{A}} = \overline{\mathbf{A}} + \widehat{\Delta\mathbf{A}}$. Reconstruct $\widehat{N}_d = \overline{N}_d + \sum_j \widehat{A}_j\mathrm{EOF}_j$ and $\widehat{\sigma}_{N_d} = [\sum_j \sigma^2_{A_j}\mathrm{EOF}_j^2]^{1/2}$, assuming independent EOF-coefficient errors.

the target field is three-dimensional (referred as "3D"). This implies that to avoid imposing an explicit assumption about cloud vertical structure, it is necessary to infer a 3D field from 2D inputs. Such a mapping is possible in principle with architectures that lift 2D features into volumetric representations using 3D transposed convolutions, but they may be difficult to train robustly and could lead to nonphysical representations (Xie et al., 2019; Antun et al., 2020).

To avoid these shortcomings we represent $N_{\mathrm{d}}(t,\mathbf{x},z)$, where t and $\mathbf{x}$ are time and horizontal location, respectively, in terms of low-dimensional empirical orthogonal functions (i.e., EOF basis) derived from the GiOcean reanalysis profiles, and learn EOF coefficient fields from satellite predictors (Table 4, step 2),

$$N'_d(t,\mathbf{x},z) \approx \sum_{j=1}^{J} A_j(t,\mathbf{x})\,\mathrm{EOF}_j(z). \tag{3}$$

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

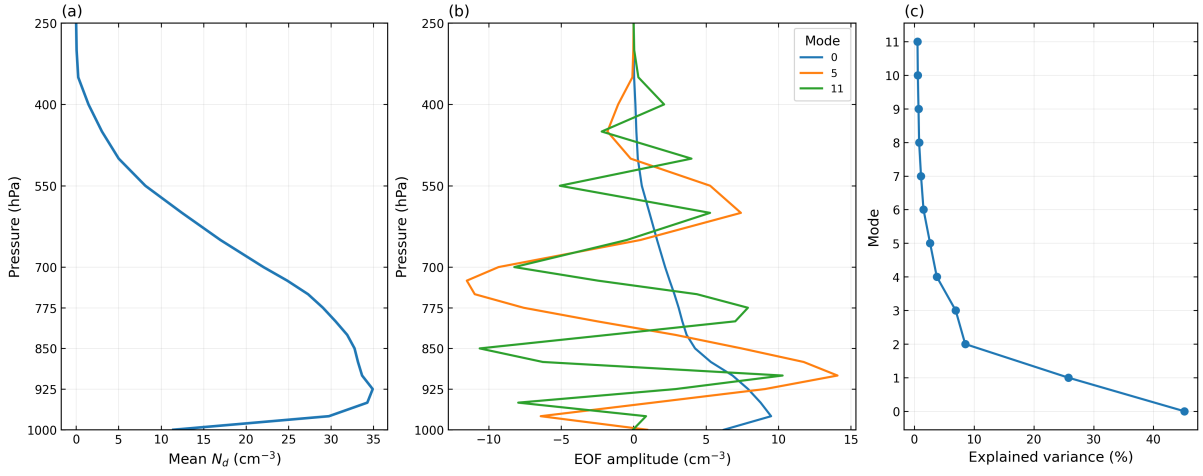


Figure 2. Statistics of the GiOcean data used to train NDnet-prior. (a) Global annual mean vertical profile of N_{d}. (b) Empirical Orthogonal Function Amplitude. (c) Explained variance for different EOF modes.

Where $N_d'(t, \mathbf{x}, z)$ is the anomaly with respect to the long-term mean, $A_j(t, \mathbf{x})$ are the projected EOF coefficients, $\mathrm{EOF}_j(z)$ is the j^{th} EOF basis, and J is the number of modes. Since the EOFs remain fixed, the network can be trained to predict the 2D $A_j(t, \mathbf{x})$ coefficients hence removing the need for map-lifting the 2D MODIS features to 3D N_{d} fields, with the added value that $N_d'(t, \mathbf{x}, z)$ inherits the general global spatial distribution imprinted by GiOcean, making the network more robust. The EOFs and mean N_{d} can be directly calculated by projecting $N_d'(t, \mathbf{x}, z)$ from the GiOcean dataset onto the EOF basis, using Python's Scipy library (Virtanen et al., 2020), and are depicted in Fig. 2. We retain the 12 main EOF modes which explain nearly 100% of the variance in N_{d}. While the lower modes carry most of the variance, hence nudging the reconstructed N_{d} to the mean, progressively higher modes explain variability at higher levels (Fig. 2b) and therefore allow for a full 3D reconstruction.

With the EOF decomposition the NN training is akin to image-to-image translation over J channels, for which U-Net family architectures are a suitable choice (Ronneberger et al., 2015). The NN (f_θ, termed "NDnet-prior") was thus built as a two-dimensional encoder–decoder convolutional model mapping MODIS-like gridded input ($\mathbf{s}_{\mathrm{COSP}}(t, \mathbf{x})$, Table 4, step 3) variables to the $A_j(t, \mathbf{x})$ coefficients of the EOF basis at each horizontal grid point, as depicted in Fig. S1. The encoder applies successive 3×3 convolutions and downsampling operations to extract spatial features over progressively larger scales, while the decoder restores the original resolution through learned upsampling. Skip connections link encoder and decoder features at matching spatial scales, which typically allow the model to retain fine-scale information. Each convolutional block uses Group Normalization and ReLU activations; the former is advantageous for small batch sizes because it does not depend on batch-wide statistics (Ronneberger et al., 2015; Wu and He, 2018). A final 1×1 convolution projects the decoded feature maps onto the retained EOF modes (Fig. S1).

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Data for the years 2000-2019 (7034 daily maps) were used for training and validation and from the year 2020 for testing. For stability, the network was trained on the EOF coefficient anomalies (Table 4, step 3) using a linear combination of minimum square error (MSE) and laplacian loss functions to guide optimization, the later to improve the contrast between pixels preventing the network for over-smoothing the $A_j(t, \mathbf{x})$ maps (Li et al., 2017).

2.2.2 Probabilistic refinement against ground and field observations

Application of NDnet-prior to MODIS retrievals produces EOF coefficients that can be used to reconstruct the global 3D distribution of N_d (Table 4, step 4). Replacing the simulated COSP inputs with observed MODIS retrievals allows the reconstructed field to respond directly to the observed spatial variability of cloud properties. This can provide a more realistic representation of N_d than the original GiOcean field, particularly because N_d is not directly constrained by the GiOcean data-assimilation system. The vertical structure of the reconstruction, however, remains dependent on GiOcean. MODIS does not directly constrain the vertical distribution of N_d, so the mapping from 2D satellite retrievals to 3D N_d profiles is inherited from the relationships learned during training. As a result, systematic biases in the simulated aerosol and cloud environment can propagate into the reconstructed profiles. Previous studies suggest that GiOcean may underestimate near-surface CCN concentrations, which could introduce corresponding biases in reconstructed N_d (Song et al., 2025; Barahona et al., 2026).

A refinement step is introduced to reduce remaining biases using ground-based retrievals and aircraft campaign measurements. The MODIS-based prior generated in Step 4 (Table 4) is retained as the baseline solution, and the refinement model then learns a correction to that prior, together with its uncertainty (Table 4, Step 5). Thus, NDnet does not attempt to learn the full 3D N_d distribution from the sparse observations alone. This distinction is important because the observations provide strong local constraints but limited spatial and temporal coverage. The prior preserves the large-scale spatial and vertical structure learned from the reanalysis, while the refinement adjusts that structure toward the observed conditions.

The probabilistic refinement is designed as a retrospective synthesis rather than a purely out-of-sample prediction problem and therefore uses the full set of ground-based and campaign observations to constrain the final N_d reconstruction. Network validation is still performed on a separate 10% subset during training, but no independent test set is reserved in the final model for evaluating skill at individual sites or campaigns. To provide an additional check on out-of-sample behavior, we also train a separate model in which 20% of the campaign segments are withheld from refinement and used only for further evaluation. This experiment tests generalization to unseen campaign segments, but not to entirely unseen campaigns or sites. Accordingly, comparisons using the final reconstruction are interpreted primarily as consistency with the observational constraints rather than as fully independent estimates of predictive skill.

Training against observations raises two related issues: experimental uncertainty and representativeness. NDnet predicts daily mean fields, so the observations are aggregated to a comparable temporal scale before evaluation. Fixed-site retrievals are averaged over each day, while aircraft measurements are averaged within campaign segments and then combined by flight and day. For aircraft data, this does not constitute a true daily mean, but rather a daily constraint that reduces sensitivity to short-lived variability. Together with the LPCE approach described in Section 2.1.5, this reduces sampling mismatch with the daily NDnet representation, while residual differences contribute to representativeness uncertainty.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

NDnet does not resolve the cloud diurnal cycle. However, NDnet-prior is trained on the continuous reanalysis N_{d} field, while COSP provides the satellite-sampling representation of the inputs, allowing the prior to interpolate beyond the MODIS overpass sampling before refinement against the observations. Variability within the observational averages is retained in the uncertainty assigned to each constraint. The refinement workflow is shown in Fig. S2 and detailed below.

Let $\boldsymbol{\mu}_0(t,\mathbf{x})$ denote the residual EOF coefficients predicted by NDnet-prior from the MODIS fields. The corresponding baseline absolute coefficients are

$$\mathbf{A}_{\mathrm{base}}(t,\mathbf{x}) = \overline{\mathbf{A}}(\mathbf{x}) + \boldsymbol{\mu}_0(t,\mathbf{x}), \tag{4}$$

where $\overline{\mathbf{A}}(\mathbf{x})$ is the mean EOF coefficient field. These prior coefficients are computed before refinement training. NDnet receives the complete two-dimensional MODIS maps, $\mathbf{s}_{\mathrm{MODIS}}(t,\mathbf{x})$, together with $\boldsymbol{\mu}_0(t,\mathbf{x})$. The MODIS variables are passed through three 3×3 convolutional layers with ReLU activations (Fig. S2). These layers transform the original MODIS variables into learned spatial feature maps while retaining the horizontal dimensions of the input grid. The feature maps are then concatenated with the prior coefficient fields. Two parallel convolutional heads predict an additive coefficient correction, $\boldsymbol{\delta A}$, and a positive coefficient uncertainty, $\boldsymbol{\sigma}_A$,

$$\widehat{\Delta\mathbf{A}} = \boldsymbol{\mu}_0 + \boldsymbol{\delta A}, \qquad \boldsymbol{\sigma}_A = \mathrm{softplus}(\boldsymbol{\rho}) + \sigma_{\min}, \tag{5}$$

where $\boldsymbol{\rho}$ is the network raw uncertainty output. The final absolute coefficients are therefore

$$\widehat{\mathbf{A}} = \overline{\mathbf{A}} + \widehat{\Delta\mathbf{A}}. \tag{6}$$

This residual formulation stabilizes the refinement because the network only needs to learn departures from the prior estimate (He et al., 2016). The second output head provides an input-dependent, estimate of the unresolved variability in each EOF coefficient (Kendall and Gal, 2017).

Although the observational constraints are sparse, the inputs and outputs of NDnet remain global two-dimensional fields. The network predicts $\boldsymbol{\delta A}$ and $\boldsymbol{\sigma}_A$ at every horizontal grid point. The ground-based and campaign observations are associated with their timestamp, horizontal and vertical coordinates, and are represented through validity masks. They therefore act as sparse samples of a global three-dimensional target field. Losses are evaluated only where valid observations are available, but the convolutional model, together with the prior and spatial regularization, distributes the learned correction over the full domain. The refinement is thus not a collection of independent site corrections; it is a global adjustment constrained locally by observations.

2.2.3 Ground-based site constraints.

At each fixed ground-based site s, the predicted residual coefficients and their uncertainties are sampled from the global output fields. The retrieved site profiles are projected onto the same EOF basis used by NDnet-prior, producing observed

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

residual coefficients $\Delta A_m^{\mathrm{obs}}(t,s)$. A validity mask excludes times for which a site retrieval is unavailable. The coefficient-space constraint is expressed as a diagonal Gaussian negative log-likelihood in the form,

$$\mathcal{L}_{\mathrm{site}} = \frac{1}{|\mathcal{S}_v|} \sum_{s\in\mathcal{S}_v} \frac{1}{\sum_{m=1}^{M} w_m} \sum_{m=1}^{M} w_m \left[\log \sigma_{A,m}(s) + \frac{\left(\widehat{\Delta A}_m(s) - \Delta A_m^{\mathrm{obs}}(s)\right)^2}{2\sigma_{A,m}^2(s)} \right], \tag{7}$$

where $\mathcal{S}_v$ is the set of sites containing valid observations and w_m accounts for differences in variance among the EOF modes. The loss is first normalized within each site and then averaged over sites. This equal-site weighting prevents sites with longer records from dominating the refinement.

The predicted site profile is reconstructed directly from the mean coefficients and the refined residual coefficients,

$$\widehat{N}_d(z,s) = \overline{N}_d(z) + \sum_{m=1}^{M} \left[\overline{A}_m(s) + \widehat{\Delta A}_m(s) \right] \mathrm{EOF}_m(z). \tag{8}$$

Assuming that the uncertainties of the EOF coefficients are independent, the model variance is propagated into N_{d}-space as

$$\sigma_{N_d,\mathrm{model}}^2(z,s) = \sum_{m=1}^{M} \mathrm{EOF}_m^2(z)\, \sigma_{A,m}^2(s). \tag{9}$$

The effective variance used in the profile-space Gaussian loss includes both the predicted variance and the experimental retrieval uncertainty,

$$\sigma_{N_d,\mathrm{eff}}^2(z,s) = \sigma_{N_d,\mathrm{model}}^2(z,s) + \sigma_{N_d,\mathrm{exp}}^2(z,s). \tag{10}$$

Two complementary losses are applied after reconstruction. The first is a mean-squared profile error,

$$\mathcal{L}_{N_d} = \frac{1}{|\mathcal{S}_v|} \sum_{s\in\mathcal{S}_v} \frac{1}{|\mathcal{Z}_s|} \sum_{z\in\mathcal{Z}_s} \omega_{z,s} \left[\widehat{N}_d(z,s) - N_d^{\mathrm{obs}}(z,s) \right]^2, \tag{11}$$

where $\mathcal{Z}_s$ is the set of valid vertical levels at site s. The bounded weight $\omega_{z,s}$ gives slightly more influence to observations with smaller reported error while preventing individual points from dominating the loss.

The second profile-space term is a Gaussian negative log-likelihood, i.e.,

$$\mathcal{L}_{N_d\mathrm{G}} = \frac{1}{|\mathcal{S}_v|} \sum_{s\in\mathcal{S}_v} \frac{1}{|\mathcal{Z}_s|} \sum_{z\in\mathcal{Z}_s} \left[\log \sigma_{N_d,\mathrm{eff}}(z,s) + \frac{\left[\widehat{N}_d(z,s) - N_d^{\mathrm{obs}}(z,s) \right]^2}{2\sigma_{N_d,\mathrm{eff}}^2(z,s)} \right]. \tag{12}$$

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Both Gaussian losses, Eqs. 7 and 12, are introduced to account for experimental and representativeness error during training, and constraint the uncertainty output head. The coefficient-space and profile-space losses provide complementary constraints. The coefficient loss preserves consistency with the EOF representation, whereas the reconstructed-profile losses constrain N_d directly.

2.2.4 Field-campaign constraints.

Aircraft campaign measurements differ from the fixed-site retrievals because the aircraft move through the model domain and sample N_d at different horizontal locations and pressures. However, NDnet produces one mean field per day and cannot represent individual, instantaneous aircraft samples. The campaign measurements are therefore aggregated to daily means before they are used for refinement, keeping the unresolved variability as a measure of prediction uncertainty. The resulting daily campaign constraints are matched to the corresponding daily MODIS input and daily NDnet prediction.

Let n denote a valid daily campaign constraint rather than an individual instantaneous measurement. Each constraint is defined by its day, horizontal grid cell, representative pressure p_n, daily mean droplet number concentration $N_{d,n}^{\mathrm{obs}}$, and observational uncertainty $\sigma_{n,\mathrm{exp}}$. The latter is estimated from the variability of the measurements contributing to the daily mean. Campaign and flight identifiers are retained to construct the loss weights so that campaigns containing more flights or valid days do not disproportionately influence the loss. For each campaign day, the refined coefficient and uncertainty fields are sampled from the corresponding daily global output at the grid cell containing the observation. A complete daily-mean N_d profile and its model variance are then reconstructed on the EOF pressure grid, i.e.,

$$\widehat{N}_{d,n}(z) = \overline{N}_d(z) + \sum_{m=1}^{M} \left[\overline{A}_m(\mathbf{x}_n) + \widehat{\Delta A}_m(t_n, \mathbf{x}_n)\right] \mathrm{EOF}_m(z), \tag{13}$$

$$\sigma^2_{N_d,\mathrm{model},n}(z) = \sum_{m=1}^{M} \mathrm{EOF}^2_m(z)\, \sigma^2_{A,m}(t_n, \mathbf{x}_n). \tag{14}$$

These daily profiles are then log-interpolated linearly to the representative pressure of the daily campaign. The campaign mean-squared loss is

$$\mathcal{L}_{\mathrm{camp}} = \left\langle \left[\widehat{N}_{d,n}(p_n) - N_{d,n}^{\mathrm{obs}}\right]^2 \right\rangle, \tag{15}$$

and the corresponding Gaussian loss is

$$\mathcal{L}_{\mathrm{camp,G}} = \left\langle \log \sigma_{n,\mathrm{eff}} + \frac{\left[\widehat{N}_{d,n}(p_n) - N_{d,n}^{\mathrm{obs}}\right]^2}{2\sigma^2_{n,\mathrm{eff}}} \right\rangle, \tag{16}$$

where

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

$$\sigma^2_{n,\mathrm{eff}} = \sigma^2_{N_d,\mathrm{model},n}(p_n) + \sigma^2_{n,\mathrm{exp}}. \tag{17}$$

The campaign losses are combined using a campaign-balanced average, denoted by $\langle . \rangle$. Losses from the valid daily mean constraints are first averaged within each flight. The resulting group means are then averaged within each campaign, and the campaign means are finally averaged across campaigns. Thus, each flight has equal weight within its campaign, and each campaign has equal weight in the total loss, regardless of the number of measurements or valid days it contains.

The campaign and fixed-site data provide complementary daily mean constraints. Fixed sites provide repeated vertical-profile information at a small number of locations, whereas the campaigns extend the observational coverage across a wider range of regions, cloud regimes, and pressure levels. In both cases, the observations constrain daily global output fields rather than instantaneous cloud states.

2.2.5 Regularization

Because the observational constraints remain sparse, dense regularization is applied over the full domain to limit large departures from the base state calculated from NDnet-prior caused by single sites/campaigns. The refinement correction in EOF-coefficient space as

$$\mathbf{C}(t,\mathbf{x}) = \widehat{\Delta\mathbf{A}}(t,\mathbf{x}) - \boldsymbol{\mu}_0(t,\mathbf{x}) \tag{18}$$

Rather than penalizing each coefficient independently, the correction is reconstructed in N_d space,

$$\Delta N_\mathrm{d}(p,\mathbf{x}) = \sum_m C_m(\mathbf{x})\mathrm{EOF}_m(p). \tag{19}$$

where m indexes the EOF modes. The correction is then penalized through a whole-column anchor,

$$\mathcal{L}_\mathrm{anchor} = \left\langle w_\mathrm{lat}, w_p \left(\frac{\Delta N_\mathrm{d}}{N_*}\right)^2 \right\rangle, \tag{20}$$

where w_lat accounts for grid-cell area, w_p provides the vertical weighting for integration in $\ln p$, and $N_* = 100\ \mathrm{cm}^{-3}$ is a fixed reference concentration. This constraint acts throughout the atmospheric column and over the full globe. It favors NDnet-prior where the observations do not provide sufficient evidence for a correction, while allowing departures where they are supported by observations. Because observational support becomes particularly weak in the upper troposphere, an additional constraint is applied to the large-scale upper-level correction. The correction is separated into its zonal mean and longitude anomaly, and the zonal-mean component is increasingly penalized from 700 to 500 hPa, reaching full strength above 500 hPa. The longitude-anomaly penalty is set to zero in the final configuration, allowing localized upper-level corrections where they are supported by observations. Regularization acts only during training and it is not applied during inference.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

The Gaussian likelihoods can reduce their penalty by predicting excessively large uncertainties. We therefore also regularize the predicted EOF-coefficient uncertainty relative to the characteristic scale S_m of each EOF mode,

$$\mathcal{L}_{\sigma,\mathrm{site/camp}} = \left\langle \left(\frac{\sigma_{A_m}}{S_m} \right)^2 \right\rangle. \tag{21}$$

where S_m is the latitude-area-weighted root mean square of the corresponding NDnet-prior residual coefficient over the training period. This penalty is evaluated separately at fixed-site and campaign locations, using equal-site and campaign averaging, respectively.

2.2.6 Total refinement loss

The final refinement objective combines a dense satellite-informed prior, sparse ground-based profiles, moving aircraft measurements, explicit experimental uncertainty, and spatial regularization, i.e.,

$$\mathcal{L} = \lambda_{N_d}\mathcal{L}_{N_d} + \lambda_{N_d\mathrm{G}}\mathcal{L}_{N_d\mathrm{G}} + \lambda_{\mathrm{camp}}\mathcal{L}_{\mathrm{camp}} + \lambda_{\mathrm{camp,G}}\mathcal{L}_{\mathrm{camp,G}} + \lambda_{\mathrm{anchor}}\mathcal{L}_{\mathrm{anchor}} + \lambda_{\sigma,\mathrm{site}}\mathcal{L}_{\sigma,\mathrm{site}} + \lambda_{\sigma,\mathrm{camp}}\mathcal{L}_{\sigma,\mathrm{camp}}. \tag{22}$$

Terms associated with fixed sites or campaigns are omitted for training samples that do not contain the corresponding observations. Before combination, each loss term is expressed in dimensionless form using fixed normalization scales. Regularization terms are normalized by a fixed reference concentration, $N_* = 100$ cm^{-3}, while coefficient-space uncertainty terms are normalized by the training-period RMS of the corresponding prior EOF coefficient. Gaussian likelihoods use the corresponding normalized uncertainty scale.

The coefficients λ_i are dimensionless hyperparameters that control the relative contribution of each term to the total refinement loss. Most terms were assigned $\lambda_i = 1$, so that no additional rescaling was introduced. The two Gaussian N_d likelihood terms were given smaller weights, with $\lambda_{N_d\mathrm{G}} = 0.25$ for the fixed sites and $\lambda_{\mathrm{camp,G}} = 0.20$ for the aircraft campaigns. These terms complement the corresponding deterministic N_d losses by accounting for predicted and observational uncertainty, but their smaller weights prevent the probabilistic likelihood from dominating agreement with the observations. The slightly lower campaign value reflects the more intermittent sampling of the campaign observations.

3 Results

Figure 3, shows the annual mean N_d over the entire MODIS record (2003-2024) predicted by the network at different stages of development. The left panels correspond to the prediction of NDnet-prior on the GiOcean COSP data, essentially emulating the GiOcean reanalysis. The center panels show N_d predicted by the same network but using MODIS retrievals as input. The right panels, correspond to the final network, after refinement. We evaluated the network at these different stages, that is, on its ability to reproduce the reanalysis data (NDnet-prior), and then to correct for biases against ground-based observations and campaign measurements (NDnet). NDnet-prior was evaluated on a separate test set (year 2020) not used during training.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

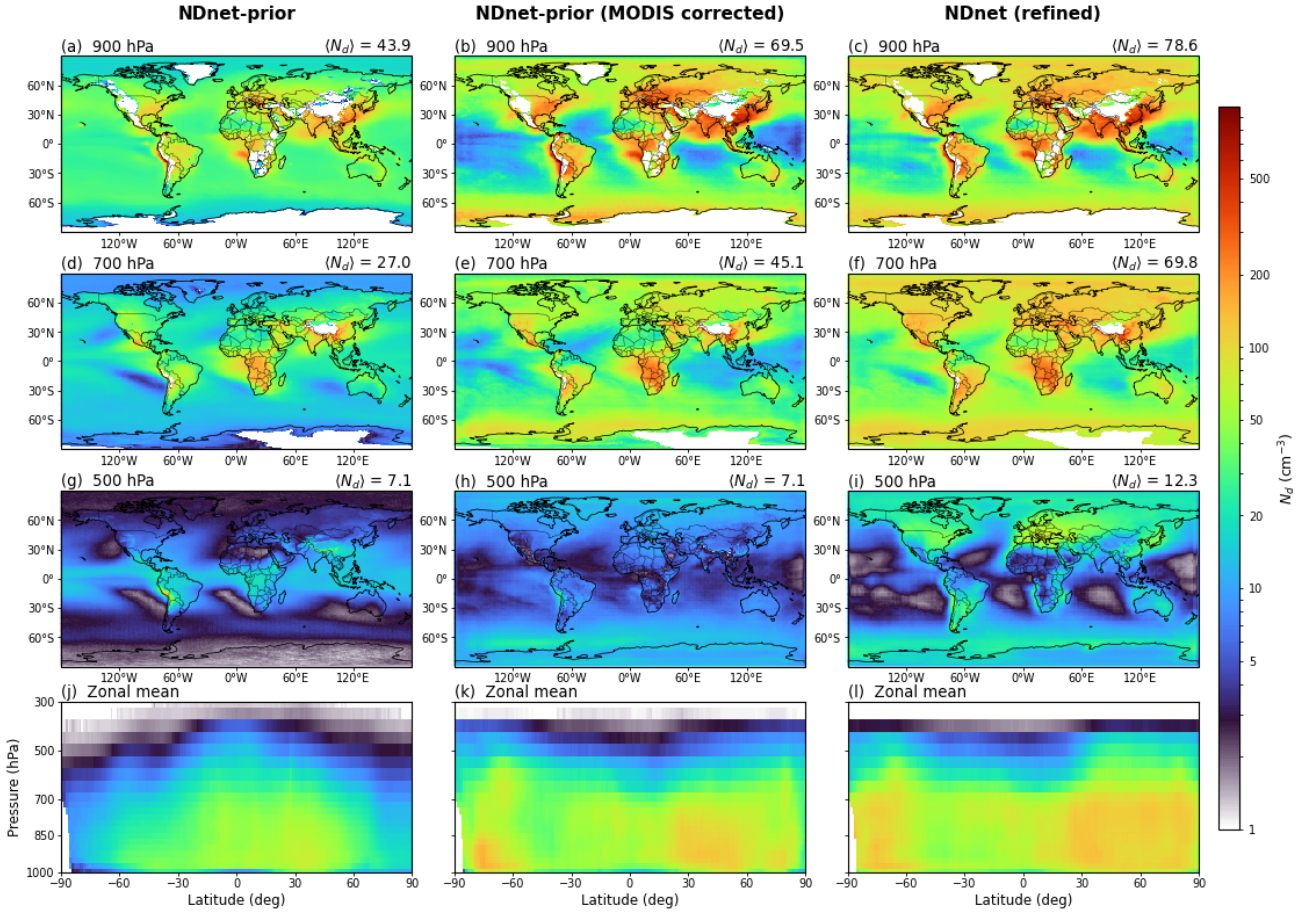


Figure 3. Global mean cloud droplet number concentration for 2003–2024. Left panels: NDnet-prior emulation of the GiOcean reanalysis. Middle panels: NDnet-prior driven by MODIS retrievals. Right panels: refined NDnet driven by MODIS retrievals.

The final N_{d} product is treated as a retrospective synthesis using the full observational record, while a separate withheld-data experiment is used to provide an additional check on out-of-sample behavior (Section 3.1).

NDnet-prior is able to emulate the reanalysis to high accuracy, with a mean bias about (1%), slightly increasing to with altitude as N_{d} decreases (Fig. S3). This is an important test of the EOF approach described in Section 2.2.1, as it demonstrates
370 that the network effectively lifts retrieval-like 2D features (optical depth, liquid water path, cloud fraction and effective radius) into 3D N_{d} fields, while ensuring a physically-based prediction guided by the reanalysis. Locally, the bias can grow up to 50% in some regions with low N_{d}, particularly in the extra tropics and at high altitudes. Low N_{d} tends to exacerbated relative bias due to spatial variability in N_{d}. In absolute terms, the error remains low, and even regions with high relative error represent only about $1-5$ cm^{-3} discrepancy in N_{d} (Fig. S3). No additional training was attempted to reduce these discrepancies as it

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

is important that the network learns the large scale features of N_d instead of the inherent variability of the reanalysis, hence making it more suitable for refinement.

Figure 3 compares the N_d distributions obtained using NDnet-prior with simulated COSP inputs from GiOcean and with actual MODIS retrievals (hereafter MODIS-corrected). Both reconstructions reproduce the main spatial features commonly found in N_d retrievals and simulations (Bennartz and Rausch, 2017; Grosvenor et al., 2018b; Zeng et al., 2014; McCoy et al., 2018). Near the surface (900 hPa), N_d reaches approximately 1000 cm^{-3} over major industrial regions, including Southeast Asia and Europe, consistent with enhanced aerosol emissions. Marine environments generally exhibit lower concentrations ($\sim$100 cm^{-3}), except in persistent stratocumulus regions off the coasts of North and South America and West Africa, where continental outflow may enhance aerosol concentrations. Some extratropical regions exhibit particularly low N_d ($\sim$20 cm^{-3}), potentially associated with limited cloudiness and aerosol removal by precipitation. Despite these common features, replacing the simulated COSP inputs with MODIS retrievals produces substantial differences in both the horizontal and vertical distributions of N_d.

The MODIS-corrected reconstruction generally produces higher N_d over continental regions, particularly at high latitudes (Fig. S6). Below 700 hPa, the global mean difference is approximately 20 cm^{-3}, reaching 100 cm^{-3} over some industrial regions, while differences become less pronounced at higher altitudes. Whereas N_d in the GiOcean reanalysis decreases steeply toward the poles and with altitude, the MODIS-corrected reconstruction shows a slight poleward increase and a relatively more homogeneous vertical distribution up to approximately 700 hPa. These differences are consistent with changes in the satellite input fields (Fig. S4). Compared with GiOcean, MODIS generally exhibits higher cloud optical depth and cloud fraction at high latitudes and slightly lower values in the tropics, where the reconstructed N_d is also generally lower. The network appears to translate these differences into changes in the predicted EOF coefficients, increasing the contribution of higher-order modes to the reconstructed vertical structure (Fig. 2). Mode-ablation experiments, in which individual EOF modes are excluded during inference, confirm that higher-order modes contribute to the enhanced N_d in the MODIS-corrected reconstruction (Fig. S5).

Observational refinement produces smaller global changes in N_d than the switch from COSP to MODIS inputs (Fig. 3). Relative to the MODIS-corrected prior, refinement generally increases N_d over marine regions and at high latitudes, while reducing near-surface concentrations over several continental regions (Fig. S6). These corrections are spatially coherent and extend well beyond the observational locations, indicating that refinement adjusts the global relationship between the MODIS retrievals and the predicted EOF coefficients rather than introducing localized corrections. The global-mean adjustment is positive, reaching 9.1, 24.7, and 5.2 cm^{-3} at 900, 700, and 500 hPa, respectively (Fig. S6). The largest mean response occurs at 700 hPa, reflecting differences in the vertical redistribution of N_d. Over land, refinement generally decreases near-surface N_d while increasing concentrations aloft, whereas marine regions exhibit a more uniform increase with altitude. The resulting profiles are also more vertically homogeneous below 700 hPa. Some of the largest regional adjustments occur in areas influenced by substantial aerosol emissions, suggesting potential implications for estimates of aerosol–cloud interactions. At lower pressures, the adjustments gradually diminish as the reconstruction transitions toward the prior, limiting the extrapolation of observational corrections into regions of weaker observational constraint. This effect is particularly evident in the Southern Hemisphere, where observational coverage is more limited.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

These features are in line with theoretical expectations (Khain et al., 2019) as well as experimental studies showing that enough reservoir of aerosol and persistent dynamical forcing exist in the upper troposphere to reach N_d of the order of 100 cm^{-3} (McCoy et al., 2020; Barahona et al., 2025). It is also consistent with observations showing unexpectedly large summertime droplet concentrations toward Antarctica and systematic low biases in simulated Southern Ocean N_d, which can increase with altitude. Elevated polar liquid-bearing clouds can retain N_d of order 50–150 cm^{-3} several kilometers above the surface. Such behavior may reflect aerosol sources poorly represented in large-scale models, including marine-biogenic particle production and free-tropospheric aerosol reservoirs (McCoy et al., 2020), or an overactive ice formation and growth at high latitudes (Vergara-Temprado et al., 2018).

3.1 Evaluation against observations

Figure 4 show the comparison of the predictions of the NDnet and NDnet-prior networks against observations, both using the global MODIS retrievals as input and then interpolated at individual sites (Fig. 1). The fidelity of the comparison in each case is evaluated by calculating the bias against the observations and the weighted Earth movers distance W, the latter given by,

$$W_1 = \sum_{i=1}^{n-1} \frac{|F_i - G_i| + |F_{i+1} - G_{i+1}|}{2} \left(\ln p_{i+1} - \ln p_i\right), \tag{23}$$

where F_i and G_i are the cumulative distributions of the observed and modeled daily N_d profiles, respectively, after each profile has been normalized over common pressure levels. The reported W_1 and bias are the medians across all valid daily profiles at each site; smaller values indicate more similar vertical distributions. For both NDnet-prior (calculated on MODIS inputs) and NDnet, the whole global distribution of N_d is calculated for each day and then interpolated at each time, site, and location.

In general, NDnet-prior and NDnet produce similar vertical structure of N_d at each site, with relatively good agreement with the observations near the surface, around 900 hPa, followed by decreasing N_d with altitude. NDnet-prior, however, tends to decrease more rapidly and consequently underestimates N_d at higher altitudes. This behavior is reflected in the generally larger Wasserstein distance and integrated bias of NDnet-prior relative to NDnet. At some sites, including Galati, Maido, and Mindelo, NDnet still underestimates the retrieved N_d at higher altitudes. Exact agreement with every site was not imposed during refinement. Instead, the site constraints are balanced against the field-campaign observations and the regularization toward NDnet-prior, particularly where the vertical observational support becomes sparse.

The pseudo-adiabatic retrieval assumes that N_d is approximately constant within a liquid-cloud layer, so repeated retrievals at different cloud heights can produce an apparently persistent or weakly varying vertical profile. Forcing the network to reproduce this structure at individual sites could therefore transfer assumptions of the retrieval method into the global N_d profile. The time-mean site profiles also represent an ensemble of clouds sampled over a fixed location rather than a directly observed grid-cell-mean vertical profile. The refinement was therefore designed to improve agreement with the observational ensemble without requiring a perfect fit to individual site profiles, allowing the large-scale vertical structure inherited from NDnet-prior to be retained where the observations do not provide sufficient independent constraint.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

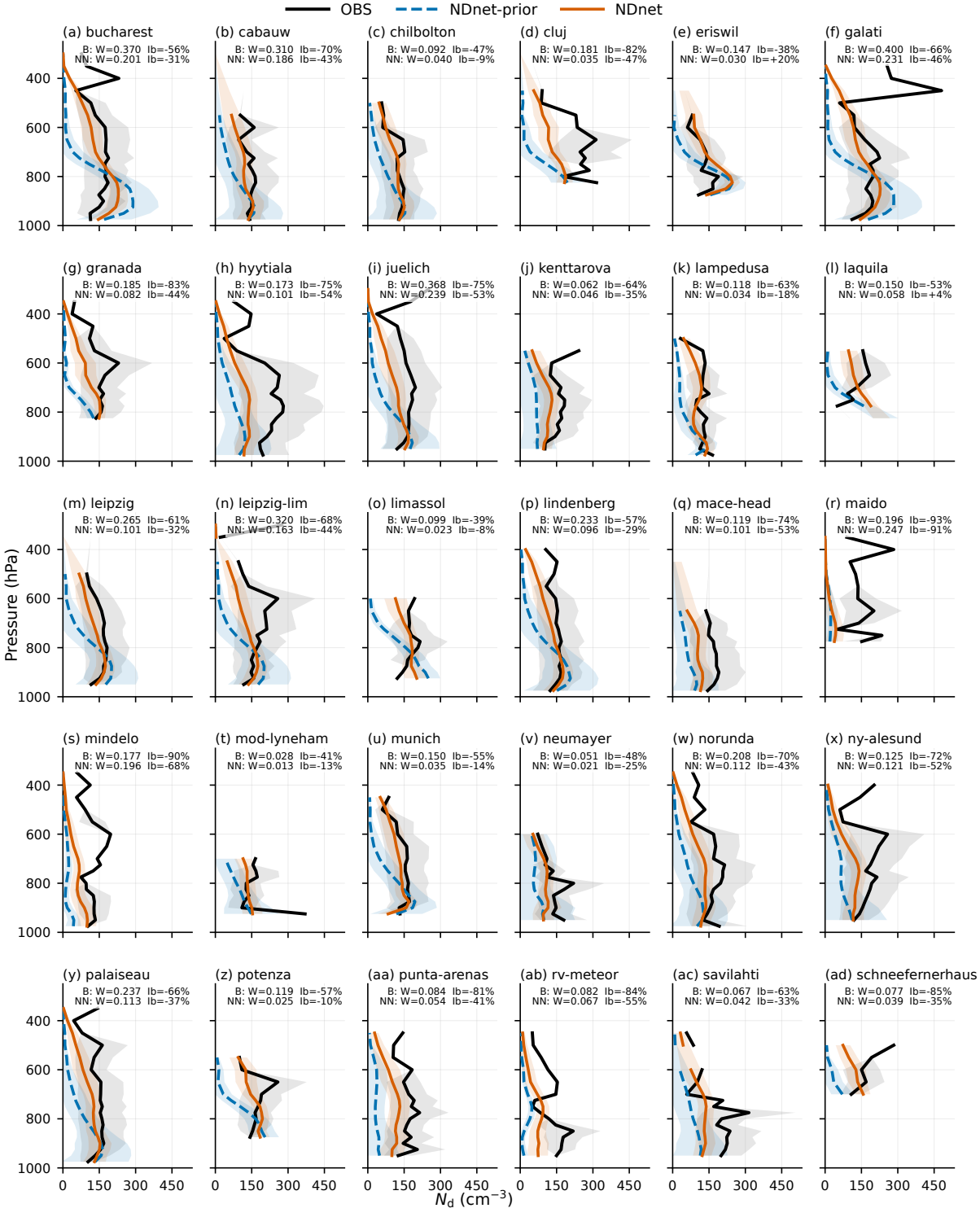


Figure 4. Performance of NDnet and NDnet-prior for the ground-based observation sites listed in Table 2. Each panel reports the Earth Mover's Distance (W) and relative bias (I_b) with respect to observations for NDnet (NN) and NDnet-prior (B).

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

In contrast to the ground-based retrievals, the aircraft campaigns provide direct in situ measurements of N_{d} across a range of altitudes and cloud environments. Figure 5 compares these observations with NDnet-prior and NDnet, both using MODIS as input, along the sampled campaign segments. As in the site evaluation, the predictions are not generated independently at each observation location. Instead, the network produces a global 3D N_{d} field from the corresponding global satellite inputs; this field is then sampled at the campaign location and day and interpolated to the pressure of each observation. Although all the campaign observations are used in the final refinement, a separate test withholding 20% of the campaign segments during training shows that performance remains comparable out of sample (Fig. S7), supporting the ability of the model to generalize beyond the individual samples used for training.

NDnet-prior generally underestimates the campaign observations, with the largest discrepancies occurring in campaigns that sample higher altitudes, such as ACTIVATE and COPE. Nevertheless, it captures a substantial fraction of the observed variability, with correlations of about 0.4–0.6 across most campaigns, particularly when individual segments are aggregated into daily grid-cell means (Fig. S8). Observational refinement improves both the magnitude and variability of the predictions, increasing correlations to 0.67–0.95 across the campaigns. This indicates that NDnet retains the large-scale structure learned by the prior while using the campaign data to correct systematic biases and better constrain the vertical distribution of N_{d}. Time series of observed, NDnet-prior and NDnet N_{d} for each campaign are shown in Fig. S8.

3.2 Uncertainty in N_{d}

An important feature of NDnet is that the refinement provides an estimate of uncertainty together with the predicted N_{d}. The second network head (Fig. S2) predicts the standard deviation of the EOF coefficients, σ_A, which is propagated through the EOF representation to obtain σ_{N_d} (Eq. 9). The observational errors are treated separately in the Gaussian likelihood, where they are combined with the predicted uncertainty (Eq. 10). Thus, σ_{N_d} represents the uncertainty in the reconstructed N_{d} for a given set of satellite inputs, rather than the experimental uncertainty of the observations. It may account for unresolved variability, observational representativeness, and limitations of the learned relationships between MODIS cloud properties and N_{d}. An explicit penalty on large σ_A prevents the network from improving the likelihood simply by inflating the predicted uncertainty.

The global mean distribution of σ_{N_d} for the years 2003-2024 is shown in Fig. 6. Mean predicted uncertainty is typically around $30\ \mathrm{cm}^{-3}$ near the surface and decreases to approximately $10\ \mathrm{cm}^{-3}$ above 500 hPa. Spatially, larger uncertainties occur over the continents than over marine regions likely due to the larger N_{d} over land. Although the pattern reverses when considered relative to the predicted N_{d} where large relative variability is found in the marine regions of the Southern Hemisphere the Tropics. Interestingly regions of persistent stratocumulus off the west coast of south and north America and Africa exhibit the lowest σ_{N_d}, in absolute and realtive sense, reflecting the homogeneous and persistent nature of the clouds in these regions. These patterns provide a measure of where the reconstructed N_{d} field is less well constrained. The magnitude of σ_{N_d} is broadly comparable to uncertainties reported for satellite-based N_{d} retrievals and to the experimental variability of the observations (Grosvenor et al., 2018b; Gryspeerdt et al., 2022).

The predicted uncertainty can also be evaluated along the aircraft trajectories. As for N_{d}, the global σ_{N_d} field is generated from the MODIS inputs and subsequently sampled at the location and pressure of each campaign segment. Figure 7 compares

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

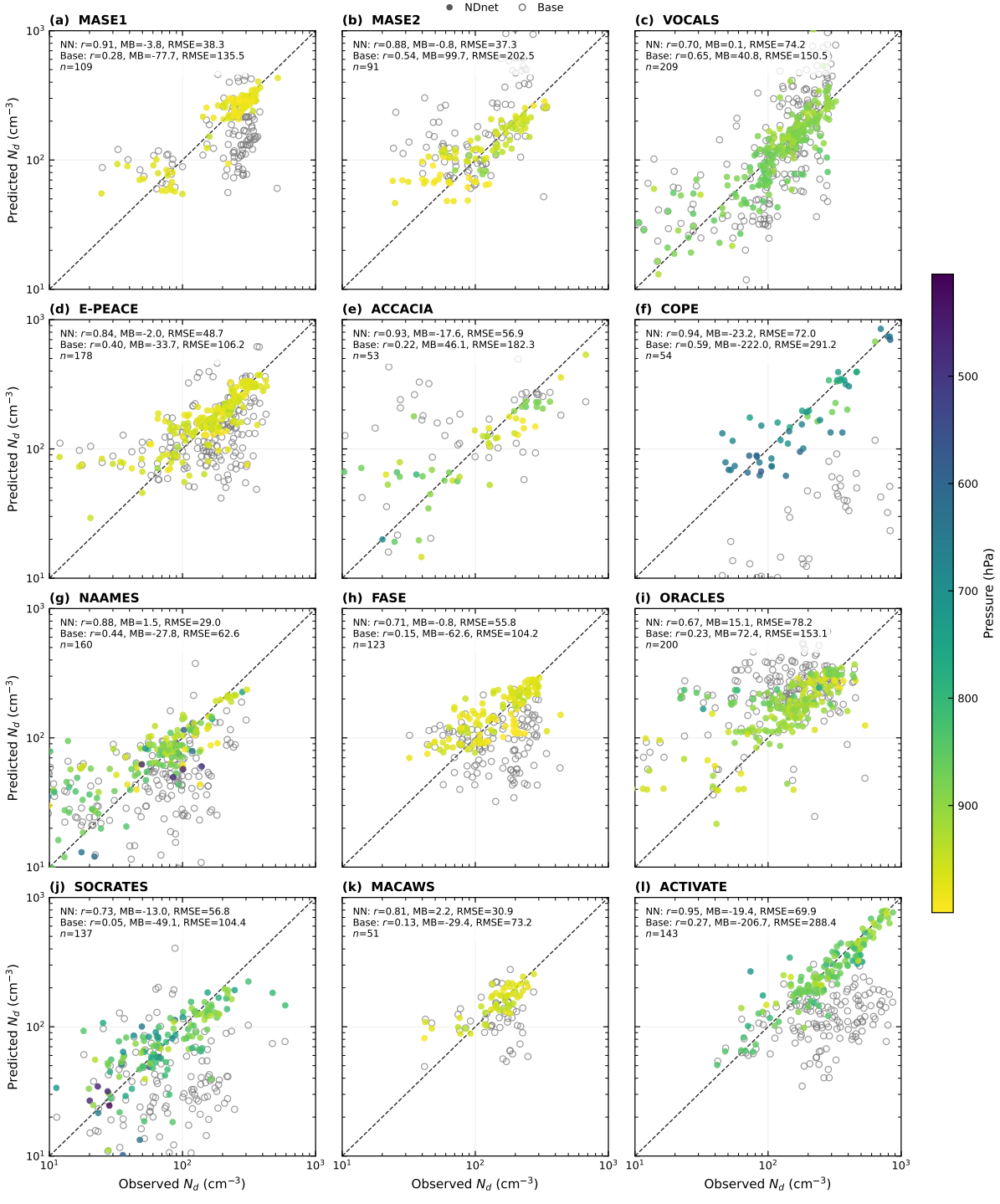


Figure 5. Performance of NDnet and NDnet-prior for the field campaigns listed in Table 3. Each panel reports the time series correlation (r) and root mean square error (RMSE) with respect to observations, for NDnet (NN) and NDnet-prior (Base).

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

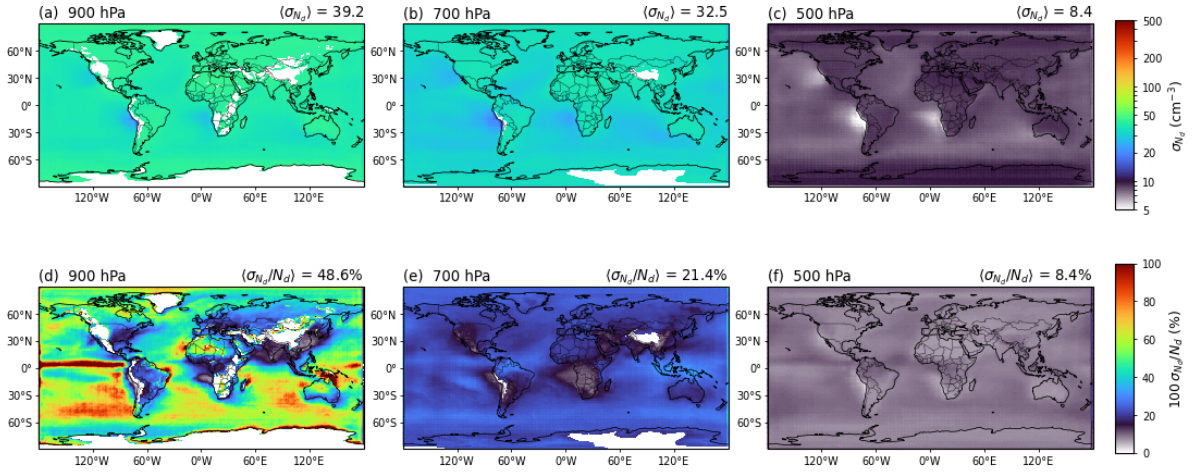


Figure 6. Global mean Gaussian uncertainty in cloud droplet number concentration, σ_{N_d}, for 2003–2024. Bottom panels show the uncertainty relative to N_d.

this quantity with the observed standard deviation of N_d within each campaign segment. The two quantities are not expected to agree point by point since σ_{N_d} represents the uncertainty of a grid-cell-scale daily prediction, whereas the campaign standard deviation measures variability among the observations within an individual segment. Nevertheless, their magnitudes are generally comparable across the campaigns. NDnet tends to underestimate the observed segment variability in some campaigns with higher sampling altitudes, like COPE and SOCRATES. Correlations between predicted and observed standard deviations are generally weak, indicating that the model reproduces the characteristic magnitude of the variability more readily than its segment-to-segment fluctuations. This comparison therefore provides a test of the scale of the predicted uncertainty rather than reproducing exactly the observed variability within individual campaigns.

4 Discussion

NDnet combines the vertical information learned from the reanalysis with satellite retrievals and direct observations to produce a globally consistent N_d field, while its predicted uncertainty provides an indication of where the inferred vertical structure is more or less strongly constrained. The three-dimensional nature of the reconstructed N_d provides an opportunity to examine its vertical structure. We quantify vertical heterogeneity over a pressure layer using H, defined as the mean square deviation of N_d within the layer relative to its layer-mean value,

$$H = \frac{1}{\langle N_\mathrm{d} \rangle} \left[\alpha_p \int_{p_\mathrm{top}}^{p_\mathrm{bottom}} \left(N_\mathrm{d}(p) - \langle N_\mathrm{d} \rangle \right)^2 d\ln p \right]^{1/2}, \tag{24}$$

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

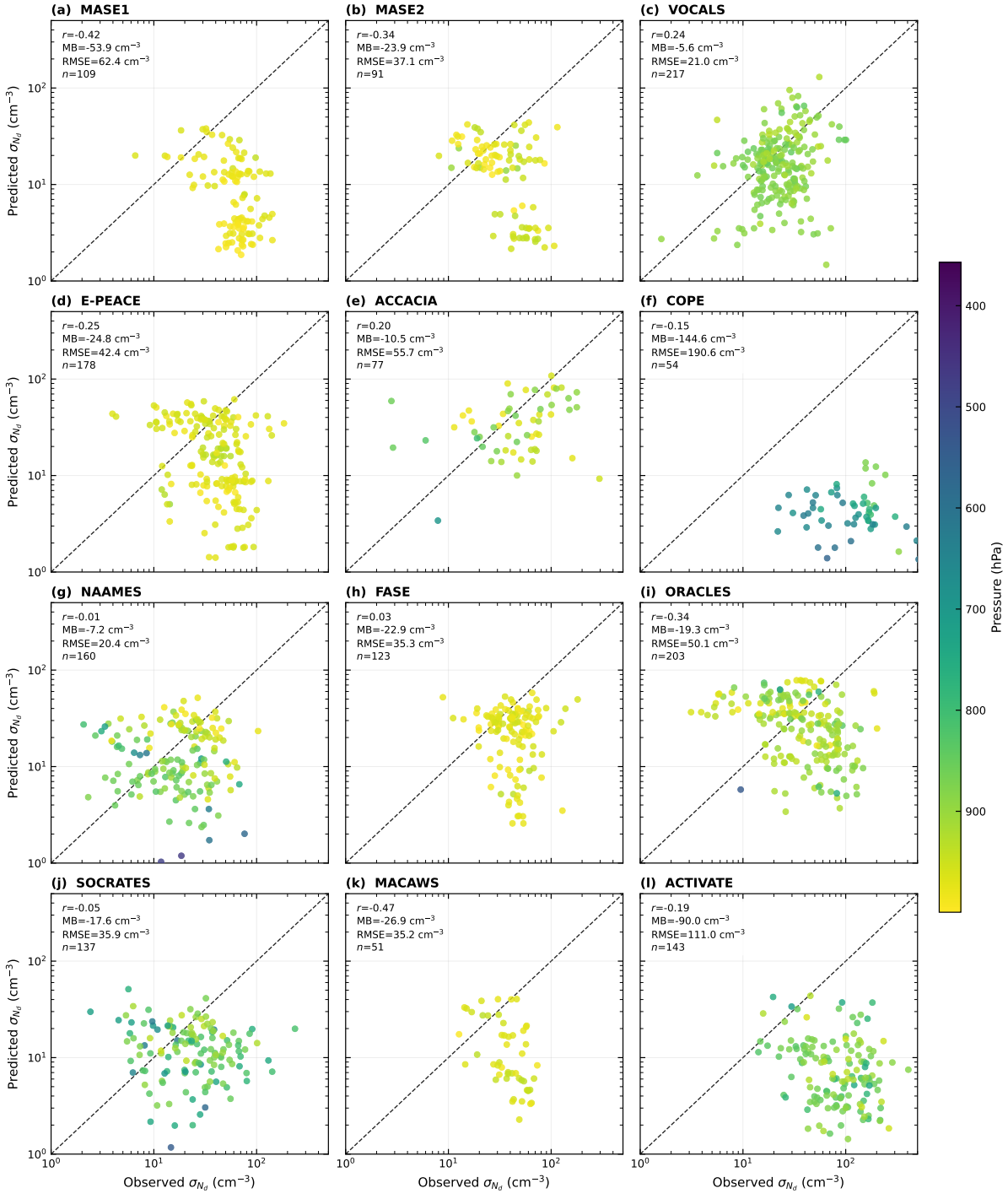


Figure 7. Gaussian uncertainty in N_{d} predicted by NDnet for the field campaigns listed in Table 3. Each panel reports the correlation coefficient (r), mean bias (MB), and root-mean-square error (RMSE) relative to observations, together with the number of daily constraints included in the comparison (n).

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

where, $\alpha_p = \ln(p_{\rm top}/p_{\rm bottom})$, and, $\langle N_{\rm d}\rangle$ is the mean $N_{\rm d}$ over the same pressure interval. $H = 0$ corresponds to a vertically homogeneous profile, while increasing H indicates progressively stronger vertical structure relative to the mean $N_{\rm d}$. We additionally define R as the ratio between the vertical variability of $N_{\rm d}$ and the root-mean-square predicted uncertainty over the same layer,

$$R = \left[\frac{\left\langle \left(N_{\rm d} - \langle N_{\rm d}\rangle\right)^2 \right\rangle_{\ln p}}{\left\langle \sigma^2_{N_{\rm d}} \right\rangle_{\ln p}} \right]^{1/2} . \tag{25}$$

Values of $R > 1$ indicate that vertical variation exceeds the predicted uncertainty over the layer. Thus, H describes how strongly $N_{\rm d}$ varies vertically, while R indicates how significant that structure is relative to the uncertainty of the estimate. Figure 8 shows the median H (left panels) over the the period 2003-2024 obtained from the final NDnet estimate after refinement. The top panels reflect variability between $p_{\rm top} = [800, 700, 600]$ hPa relative to the $p_{\rm bottom} = 950$ hPa. The "zonal" mean panel at the bottom uses $p_{\rm bottom} = 1000$ hPa and calculates H and R for any level with $p_{\rm top} <= 950$ hPa.

The reconstructed $N_{\rm d}$ exhibits substantial vertical variability, with strong spatial differences (Fig. 8). Median H exceeds 50% in tropical and subtropical regions for $p_{\rm top} = 800$ hPa and generally increases with altitude, approaching 100% for $p_{\rm top} < 600$ hPa. The frequency of days with $H > 25\%$ follows a similar spatial pattern (middle panels), indicating that this variability is persistent rather than episodic. Moreover, $R > 1$ across most regions, indicating that the inferred vertical variability exceeds the predicted NDnet uncertainty. Exceptions occur in some subtropical marine regions, where $R < 1$ and the variability is comparable to the reconstruction uncertainty.

Pressure boundaries may influence these estimates, particularly where cloud tops below the selected pressure, as in some subtropical stratocumulus regions (Fig. S4). However, using $p_{\rm bottom}$ defined as the minimum pressure for which $N_{\rm d}$ exceeds 1 cm^{-3} starting from the surface, and $p_{\rm top}$ from daily MODIS retrievals, produces even larger and more persistent H (Fig. 9). Both approaches indicate that regions with relatively homogeneous horizontal cloud properties can exhibit substantial vertical variability in $N_{\rm d}$. This result however must be interpreted within the local ensemble framework (Section 2.1.5); a model grid cell represents an ensemble of subgrid-scale cloud realizations rather than an individual cloud. Thus, the inferred variability does not necessarily contradict the relatively homogeneous structure of individual stratocumulus clouds. Nevertheless, it challenges the applicability of a vertically homogeneous $N_{\rm d}$ assumption at the grid-cell scale, even in stratocumulus regimes.

4.0.1 Comparison against other datasets

Except for some tropical regions, Fig. 8 suggests that both H and R are smallest near the surface, particularly for $p > 800$ hPa, and vertical homogeneity could be a more reasonable approximation. Taking this into account we select the $800 - 950$ hPa layer for comparison with conventional MODIS-derived $N_{\rm d}$ retrievals. These satellite retrievals typically apply several filtering criteria targeting horizontal and vertical homogeneity and pseudo-adiabatic cloud structure (Grosvenor et al., 2018b; Grosvenor and Wood, 2014; Bennartz and Rausch, 2017). These include restricting the retrievals to high cloud fractions, cloud-top temperatures above 268 K, and cloud-top height below 3.2 km. Such filtering may alter the resulting $N_{\rm d}$ climatology, but

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

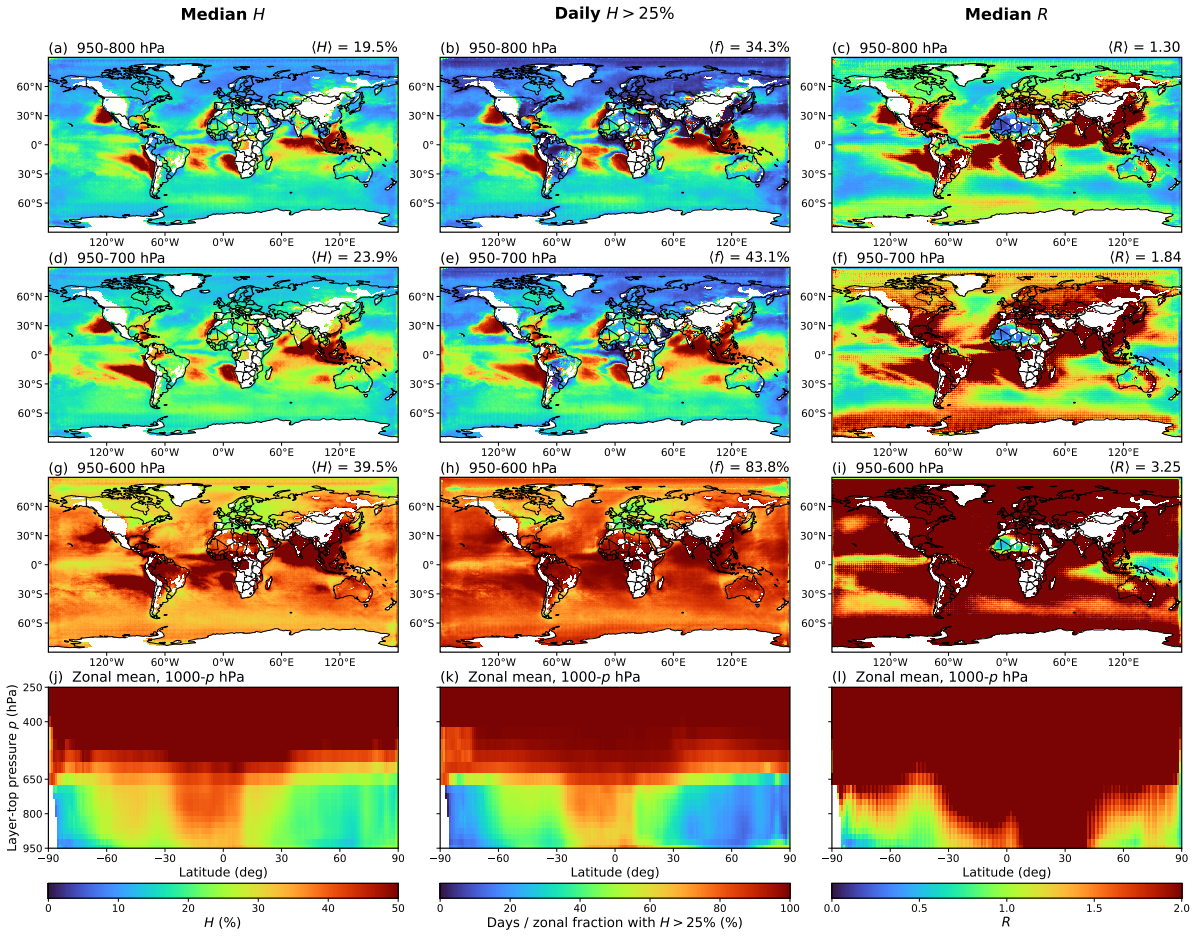


Figure 8. Vertical heterogeneity in cloud droplet number concentration for 2003–2024. Left panels: median relative vertical variability, H, in N_{d} (Eq. 24). Middle panels: percentage of days with $H > 25\%$. Right panels: vertical variability relative to the predicted uncertainty, R (Eq. 25).

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

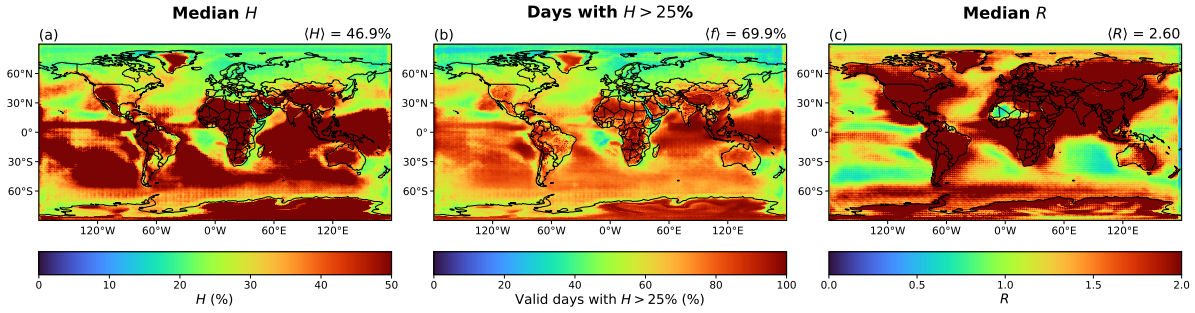


Figure 9. Similar to Fig. 8, but with cloud base defined at $N_{\mathrm{d}} \sim 1$ cm^{-3} and cloud top pressure from the MODIS retrieval.

focus on isolating conditions for which the retrieval assumptions are better satisfied (Gryspeerdt et al., 2022; Song et al., 2025). For consistency, we first calculate an 800–950 hPa layer-mean N_{d} from NDnet, weighted by the GiOcean liquid water content (denoted as $N_{\mathrm{d,sfc}}$), and then apply the same filtering criteria used for the 2D retrievals.

Figure 10 compares $N_{\mathrm{d,sfc}}$ with the retrieval climatology reported by Grosvenor and Wood (2018), hereafter referred to as the 2D retrieval. The two datasets reproduce the same broad spatial structure, including a pronounced land–ocean contrast and enhanced N_{d} in the subtropical stratocumulus regions. In general, however, $N_{\mathrm{d,sfc}}$ is lower than the 2D retrieval, except in the subtropical regions and parts of the North Atlantic. Globally, the mean difference is approximately -30 cm^{-3} and is generally larger than the predicted NDnet uncertainty, although it may remain within the uncertainty associated with the 2D retrieval itself (Grosvenor et al., 2018b).

In the zonal mean (Fig. 10d), positive and negative regional differences partly compensate, and the two datasets therefore appear more similar. Both show enhanced N_{d} near 40°N and at high southern latitudes. This agreement, however, is strongly influenced by the extensive filtering applied to both datasets. Comparison with the unfiltered NDnet climatology (Fig. S9) shows larger overall differences (~ -50 cm^{-3}), while much of the pronounced enhancement over the subtropical regions of the Northern Hemisphere is reduced. More importantly, the latitudinal structure becomes more gradual, with less pronounced peaks and a clear decrease toward the poles. This suggests that filtering for high cloud fraction and approximately adiabatic conditions preferentially samples high-N_{d} clouds at very high latitudes, thereby producing an apparent enhancement in the retrieved climatology that may not be representative of the broader cloud population.

5 Conclusions

This study develops a new global estimate of cloud droplet number concentration, N_{d}, by combining climate reanalysis, satellite retrievals, ground-based observations, and field campaign measurements. Unlike conventional satellite retrievals, the approach presented here does not require an assumed vertical cloud structure and explicitly represents uncertainty throughout the reconstruction. This is achieved using a neural-network estimator, termed NDnet, that maps two-dimensional MODIS cloud

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

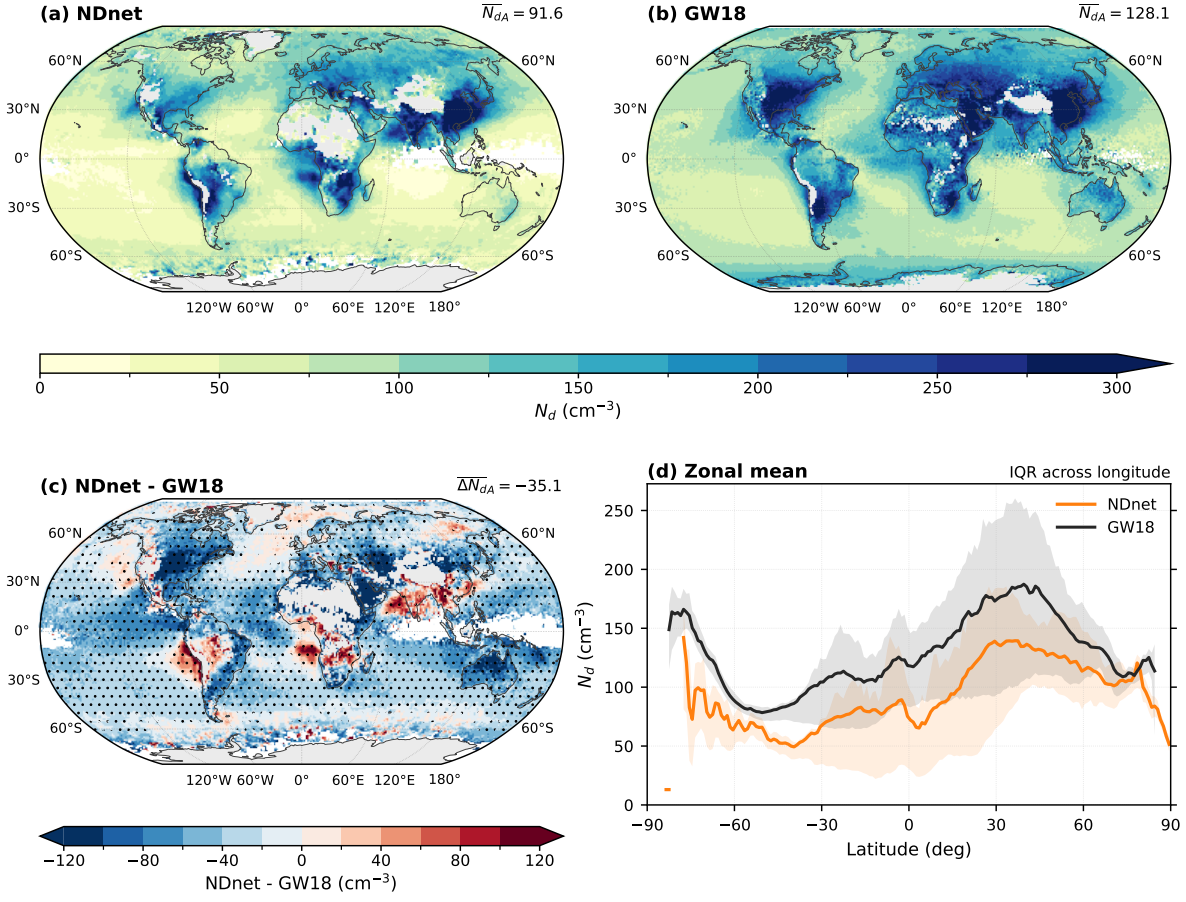


Figure 10. Layer mean $(800 - 950$ hPa) N_{d} compared to the data reported by Grosvenor and Wood (2018) (shown as GW18). Hatching indicates differences that are statistically significant at more than two standard errors.

properties to three-dimensional N_{d} fields. NDnet is applied over the MODIS record (2003-2024) to produce daily N_{d} estimates at 1° spatial resolution and 22 pressure levels from the surface to 250 hPa.

NDnet predicts the coefficients of an Empirical Orthogonal Function (EOF) representation of N_{d} rather than predicting the full three-dimensional field directly. This dimensionality reduction avoids the need for a complex neural network that lifts two-dimensional inputs directly into three-dimensional volumes, which can lead to unstable or physically unrealistic solutions. Because the EOF basis is derived from the GiOcean reanalysis, reconstructed fields remain within a physically plausible range of vertical structures. The approach also makes the problem computationally tractable for long-term training, refinement, and global inference.

NDnet is trained in two stages. First, a prior network is trained to emulate the GiOcean reanalysis using simulated MODIS variables as input. This network is then applied to actual MODIS retrievals to generate a baseline set of EOF coefficients.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

In the second stage, these coefficients and the MODIS cloud retrievals are combined with field campaign and ground-based observations to refine the reconstruction and reduce biases inherited from the reanalysis. The observations were selected to provide broad spatial, vertical, and temporal coverage. The refinement is not intended to optimize individual sites or campaigns, but rather to introduce observational constraints that act on the global N_{d} field.

Because both satellite retrievals and in situ observations contain substantial uncertainty, NDnet is formulated probabilistically. Observational and representativeness errors are included directly in the training objective, and a second network output estimates the Gaussian uncertainty associated with reconstructed N_{d}. This provides an estimate of uncertainty arising from both the intrinsic reconstruction error and the limited representativeness of the observational constraints. Predicted relative uncertainties are typically about 10–40%, generally larger over ocean than over land, and decrease with altitude.

Analysis of the approximately 20-year NDnet record shows enhanced N_{d} over major continental industrial regions and in the subtropical stratocumulus regions of both hemispheres, consistent with regions of elevated aerosol influence. These enhancements remain evident up to approximately 600 hPa, suggesting that aerosol-related variability in N_{d} extends well above the low level cloud layer. Relative to the GiOcean reanalysis, NDnet produces a more gradual vertical decrease in N_{d} and substantially larger values at high southern latitudes. The latter is consistent with evidence for significant aerosol concentrations at high latitudes (McCoy et al., 2020), suggesting that treating these regions as uniformly pristine may bias estimates of aerosol radiative forcing.

The reconstructed fields also show substantial vertical variability. Relative vertical differences in N_{d} are approximately 40% near the surface and increase toward lower pressures, often exceeding the predicted reconstruction uncertainty. Tropical and subtropical regions show particularly strong vertical variability in daily grid-cell N_{d}. This does not necessarily imply strong vertical variation within individual clouds, but rather differences among the cloud populations sampled at different heights. The persistence of this structure suggests that assuming vertically uniform N_{d} at the grid-cell scale may miss important vertical variability and affect estimates of cloud radiative effects. Nevertheless, comparison between conventional two-dimensional MODIS retrievals and an 800–950 hPa layer-mean N_{d} from NDnet shows reasonable agreement when both datasets are subjected to the extensive filtering normally required by the conventional satellite retrieval.

The three-dimensional gridded N_{d} dataset developed here can be compared directly with model output, reducing ambiguities associated model-sampling when evaluated against retrievals that assume vertically homogeneous N_{d}. The NDnet framework could also be adapted to other satellite observations (Wang et al., 2025) or developed as an observation operator for data assimilation systems, for which the uncertainty estimate could provide an explicit representation of reconstruction error within Bayesian inference.

Further improvements will require additional observational constraints, particularly at higher altitudes, where existing measurements remain sparse. Vertically resolved N_{d} retrievals and in situ measurements would be especially valuable for evaluating and refining reconstruction methods such as NDnet. Future work should also examine how vertical heterogeneity in N_{d} affects estimates of aerosol–cloud interactions and the covariability among N_{d}, cloud water path, and cloud fraction. By combining complementary observations within a physically constrained and probabilistic framework, NDnet provides a more comprehensive representation of N_{d} that can support cloud-model evaluation, climate analysis, and data assimilation.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Code and data availability. The NDnet model and training and representative test data can be downloaded at https://github.com/dbarahon/NDnet. The entire reconstructed N_d data set is available at https://portal.nccs.nasa.gov/datashare/Wstd-MERRA2/NDnet. Data from the GiOcean reanalysis was downloaded from https://portal.nccs.nasa.gov/datashare/gmao/geos-s2s-3/GiOCEAN_e1/. MODIS data was obtained from NASA LAADS DAAC (2022); NASA Earthdata (2022). CloudNet data was obtained from https://cloudnet.fmi.fi/. AI tools (ChatGSFC) were utilized to generate the workflow sketches in Figures S1 and S2 and to debug python code.

Author contributions. D.B. conceived and developed the work. E.G. curated the campaign data. T.A. and D.M. contributed to the conceptual development of the study and to scientific discussion.

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. Pytorch libraries were obtained from https://github.com/pytorch/pytorch. Donifan Barahona by the NASA Modeling, Analysis and Prediction program, Grant NNH20ZDA001N-MAP.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Abdul-Razzak, H. and Ghan, S.: A parameterization of aerosol activation, 2. Multiple aerosol types., J. Geophys. Res., 105, 6837–6844, https://doi.org/10.1029/1999JD901161, 2000.

Albrecht, B. A., Fairall, C. W., Thomson, D. W., White, A. B., Snider, J. B., and Schubert, W. H.: Surface-based remote sensing of the observed and the adiabatic liquid water content of stratocumulus clouds, Geophysical Research Letters, 17, 89–92, https://doi.org/10.1029/GL017i001p00089, 1990.

Antun, V., Renna, F., Poon, C., Adcock, B., and Hansen, A. C.: On instabilities of deep learning in image reconstruction and the potential costs of AI, Proceedings of the National Academy of Sciences, 117, 30 088–30 095, https://doi.org/10.1073/pnas.1907377117, 2020.

Arola, A., Lipponen, A., Kolmonen, P., Virtanen, T. H., Bellouin, N., Grosvenor, D. P., Gryspeerdt, E., Quaas, J., and Kokkola, H.: Aerosol effects on clouds are concealed by natural cloud heterogeneity and satellite retrieval errors, Nature Communications, 13, 7357, https://doi.org/10.1038/s41467-022-34948-5, 2022.

Bangert, M., Nenes, A., Vogel, B., Vogel, H., Barahona, D., Karydis, V. A., Kumar, P., Kottmeier, C., and Blahak, U.: Saharan dust event impacts on cloud formation and radiation over Western Europe, Atmospheric Chemistry and Physics, 12, 4045–4063, https://doi.org/10.5194/acp-12-4045-2012, 2012.

Barahona, D. and Nenes, A.: Parameterizing the competition between homogeneous and heterogeneous freezing in cirrus cloud formation - polydisperse ice nuclei, Atmospheric Chemistry & Physics, 9, 5933–5948, 2009.

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), Geosci. Model Dev., 7, 1733–1766, 2014.

Barahona, D., Molod, A., and Kalesse, H.: Direct estimation of the global distribution of vertical velocity within cirrus clouds, Scientific Reports, 7, 6840, 2017.

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.

Barahona, D., Breen, K. H., Ngo, D., Maciel, F. V., Patnaude, R., and Diao, M.: Trends in vertical wind velocity variability reveal cloud microphysical feedback, Nature Communications, 16, https://doi.org/10.1038/s41467-025-67541-7, 2025.

Barahona, D., Breen, K. H., Block, K., and Darmenov, A.: Deep learning representation of the aerosol size distribution, Geoscientific Model Development, 19, 2437–2459, https://doi.org/10.5194/gmd-19-2437-2026, 2026.

Behrenfeld, M. J. et al.: The North Atlantic Aerosol and Marine Ecosystem Study (NAAMES): Science Motive and Mission Overview, Frontiers in Marine Science, 6, 122, https://doi.org/10.3389/fmars.2019.00122, 2019.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A.-L., Dufresne, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Reviews of Geophysics, 58, https://doi.org/10.1029/2019rg000660, 2019.

Bennartz, R. and Rausch, J.: Global and regional estimates of warm cloud droplet number concentration based on 13 years of AQUA-MODIS observations, Atmos. Chem. Phys., 17, 9815–9836, https://doi.org/10.5194/acp-17-9815-2017, 2017.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Bodas-Salcedo, A., Webb, M., Brooks, M., Ringer, M., Williams, K., Milton, S., and Wilson, D.: Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities, Journal of Geophysical Research: Atmospheres, 113, publisher: Wiley Online Library, 2008.

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.

Cai, Z., Li, Z., Li, P., Li, J., Sun, H., Gao, X., Peng, Y., Wang, Y., Zhang, D., and Ren, G.: Vertical Distributions of Aerosol and Cloud Microphysical Properties and the Aerosol Impact on a Continental Cumulus Cloud Based on Aircraft Measurements From the Loess Plateau of China, Frontiers in Environmental Science, 9, https://doi.org/10.3389/fenvs.2021.808861, 2022.

Chen, Y., Haywood, J., Wang, Y., Malavelle, F., Jordan, G., Partridge, D., Fieldsend, J., De Leeuw, J., Schmidt, A., Cho, N., Oreopoulos, L., Platnick, S., Grosvenor, D., Field, P., and Lohmann, U.: Machine learning reveals climate forcing from aerosols is dominated by increased cloud cover, Nature Geoscience, 15, 609–614, https://doi.org/10.1038/s41561-022-00991-6, 2022.

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, 2010.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D., DuVivier, A., Edwards, J., Emmons, L., Fasullo, J., Garcia, R., Gettelman, A., et al.: The community earth system model version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, e2019MS001 916, 2020.

Ekman, A. M. L.: Do sophisticated parameterizations of aerosol-cloud interactions in CMIP5 models improve the representation of recent observed temperature trends?, Journal of Geophysical Research: Atmospheres, pp. n/a–n/a, https://doi.org/10.1002/2013JD020511, 2014.

Frisch, S., Shupe, M., Djalalova, I., Feingold, G., and Poellot, M.: The Retrieval of Stratus Cloud Droplet Effective Radius with Cloud Radars, Journal of Atmospheric and Oceanic Technology, 19, 835–842, https://doi.org/10.1175/1520-0426(2002)019<0835:TROSCD>2.0.CO;2, 2002.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, 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.

Gonzalez, J., Dipu, S., Jimenez, G., Camps-Valls, G., and Quaas, J.: Machine Learning-Based Retrieval of Cloud Droplet Number Concentration and Liquid Water Path From Satellite Spectral Data, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 18, 21 910–21 922, https://doi.org/10.1109/jstars.2025.3601981, 2025.

Gordon, H., Glassmeier, F., and T. McCoy, D.: An Overview of Aerosol-Cloud Interactions, in: Clouds and their Climatic Impacts, Geophysical Monograph Series, pp. 13–45, https://doi.org/10.1002/9781119700357.ch2, 2023.

Goren, T., Choudhury, G., Kretzschmar, J., and McCoy, I.: Co-variability drives the inverted-V sensitivity between liquid water path and droplet concentrations, Atmospheric Chemistry and Physics, 25, 3413–3423, https://doi.org/10.5194/acp-25-3413-2025, 2025.

Griffies, S. M., Gnanadesikan, A., Dixon, K. W., Dunne, J. P., Gerdes, R., Harrison, M. J., Rosati, A., Russell, J. L., Samuels, B. L., Spelman, M. J., Winton, M., and Zhang, R.: Formulation of an ocean model for global climate simulations, Ocean Science, 1, 45–79, 2005.

Grosvenor, D. P. and Wood, R.: The effect of solar zenith angle on MODIS cloud optical and microphysical retrievals within marine liquid water clouds, Atmospheric Chemistry and Physics, 14, 7291–7321, https://doi.org/10.5194/acp-14-7291-2014, 2014.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.



Grosvenor, D. P. and Wood, R.: Daily MODIS (MODerate Imaging Spectroradiometer) derived cloud droplet number concentration global dataset for 2003–2015, https://catalogue.ceda.ac.uk/uuid/cf97ccc802d348ec8a3b6f2995dfbbff, published 23 May 2018, 2018.

Grosvenor, D. P., Painemal, D., et al.: The impact of the effective radius penetration depth on satellite retrievals of cloud droplet number concentration, Atmospheric Measurement Techniques, 11, 4273–4307, https://doi.org/10.5194/amt-11-4273-2018, 2018a.

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., et al.: 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/10.1029/2017RG000593, 2018b.

Gryspeerdt, E., McCoy, D. T., Crosbie, E., Moore, R. H., Nott, G. J., et al.: The impact of sampling strategy on the cloud droplet number concentration estimated from satellite data, Atmospheric Measurement Techniques, 15, 3875–3892, https://doi.org/10.5194/amt-15-3875-2022, 2022.

Guo, H., Ming, Y., Fan, S., Zhou, L., Harris, L., and Zhao, M.: Two-Moment Bulk Cloud Microphysics With Prognostic Precipitation in GFDL's Atmosphere Model AM4.0: Configuration and Performance, Journal of Advances in Modeling Earth Systems, 13, e2020MS002 453, https://doi.org/10.1029/2020MS002453, publisher: John Wiley & Sons, Ltd, 2021.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778, 2016.

Hunke, E. C.: Sea ice component of the Community Ice CodE (CICE), Version 4, Los Alamos National Laboratory Tech. Rep. LA-CC-06-012, 2008.

Illingworth, A., Hogan, R., O'connor, E., Bouniol, D., Brooks, M., Delanoë, J., Donovan, D., Eastment, J., Gaussiat, N., Goddard, J., et al.: Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, Bulletin of the American Meteorological Society, 88, 883–898, 2007.

Irfan, M., Lipponen, A., Kühn, T., Romakkaniemi, S., Calderón, S. M., Holopainen, E., Virtanen, A., Arola, A., and Kokkola, H.: Using a Machine Learning Approach for Isolating Aerosol Effects on Cloud Droplet Number Concentration in Marine Stratocumulus Clouds, Journal of Geophysical Research: Atmospheres, 130, https://doi.org/10.1029/2025jd043926, 2025.

Jia, H., Ma, X., Yu, F., and Quaas, J.: Significant underestimation of radiative forcing by aerosol–cloud interactions derived from satellite-based methods, Nature Communications, 12, https://doi.org/10.1038/s41467-021-23888-1, 2021.

Jia, Y., Andersen, H., and Cermak, J.: Analysis of cloud fraction adjustment to aerosols and its dependence on meteorological controls using explainable machine learning, EGUsphere, 2023, 1–25, 2023.

Kendall, A. and Gal, Y.: What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?, in: Advances in Neural Information Processing Systems, vol. 30, 2017.

Khain, P., Heiblum, R., Blahak, U., Levi, Y., Muskatel, H., Vadislavsky, E., Altaratz, O., Koren, I., Dagan, G., Shpund, J., and Khain, A.: Parameterization of Vertical Profiles of Governing Microphysical Parameters of Shallow Cumulus Cloud Ensembles Using LES with Bin Microphysics, Journal of the Atmospheric Sciences, 76, 533–560, https://doi.org/10.1175/jas-d-18-0046.1, 2019.

Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., and Kumar, P.: A catchment-based approach to modeling land surface processes in a GCM, Part 1, Model structure, J. Geophys. Res., 105, 24 809–24 822, 2005.

Leon, D. C. et al.: The Convective Precipitation Experiment (COPE), Bulletin of the American Meteorological Society, 97, 1003–1020, https://doi.org/10.1175/BAMS-D-14-00157.1, 2016.

Li, S., Xu, X., Li, J., and Wang, R.: Laplacian-Steered Neural Style Transfer, arXiv preprint arXiv:1707.01253, 2017.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.

Liu, Y., Daum, P., Guo, H., and Y., P.: Dispersion bias, dispersion effect, and the aerosol-cloud conundrum, Env. Res. Lett., 3, 045 021, http://stacks.iop.org/1748-9326/3/i=4/a=045021, 2008.

Loveridge, J. and Di Girolamo, L.: Do Subsampling Strategies Reduce the Confounding Effect of Errors in Bispectral Retrievals on Estimates of Aerosol Cloud Interactions?, ESSOAr, https://doi.org/10.22541/essoar.172656954.43296343/v1, preprint (posted 2024-09-17), submitted to JGR: Atmospheres, 2024.

Lu, M. et al.: The Marine Stratus/Stratocumulus Experiment (MASE): Aerosol-cloud relationships in marine stratocumulus, Journal of Geophysical Research, 112, D10 209, https://doi.org/10.1029/2006JD007985, 2007.

Lu, M. et al.: Marine stratocumulus aerosol-cloud relationships in the MASE-II experiment: Precipitation susceptibility in eastern Pacific marine stratocumulus, Journal of Geophysical Research, 114, D24 203, https://doi.org/10.1029/2009JD012774, 2009.

Mace, G. G.: Deriving cloud droplet number concentration from surface-based remote sensors with an emphasis on lidar measurements, Atmospheric Measurement Techniques, 17, 3679–3695, https://doi.org/10.5194/amt-17-3679-2024, 2024.

McCoy, D. T., Bender, F. A. M., Grosvenor, D. P., Mohrmann, J. K., Hartmann, D. L., Wood, R., and Field, P. R.: Predicting decadal trends in cloud droplet number concentration using reanalysis and satellite data, Atmospheric Chemistry and Physics, 18, 2035–2047, https://doi.org/10.5194/acp-18-2035-2018, 2018.

McCoy, I. L., McCoy, D. T., Wood, R., Regayre, L., Watson-Parris, D., Grosvenor, D. P., Mulcahy, J. P., Hu, Y., Bender, F. A. M., Field, P. R., Carslaw, K. S., and Gordon, H.: The hemispheric contrast in cloud microphysical properties constrains aerosol forcing, Proceedings of the National Academy of Sciences, p. 201922502, https://doi.org/10.1073/pnas.1922502117, 2020.

McFarquhar, G. M. et al.: Observations of Clouds, Aerosols, Precipitation, and Surface Radiation over the Southern Ocean: An Overview of CAPRICORN, MARCUS, MICRE, and SOCRATES, Bulletin of the American Meteorological Society, 102, E894–E928, https://doi.org/10.1175/BAMS-D-20-0132.1, 2021.

Molod, A., Takacs, L., Suarez, M., and Bacmeister, J.: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, Geoscientific Model Development, 8, 1339–1356, 2015.

Molod, A., Hackert, E., Vikhliaev, Y., Zhao, B., Barahona, D., Vernieres, G., Borovikov, A., Kovach, R. M., Marshak, J., Schubert, S., Li, Z., Lim, Y.-K., Andrews, L. C., Cullather, R., Koster, R., Achuthavarier, D., Carton, J., Coy, L., Friere, J. L. M., Longo, K. M., Nakada, K., and Pawson, S.: GEOS-S2S Version 2: The GMAO High-Resolution Coupled Model and Assimilation System for Seasonal Prediction, J. Geophys. Res. Atmos., 125, 2020.

Morrison, H. and Gettelman, A.: A New Two-Moment Bulk Stratiform Cloud Microphysics Scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I: Description and Numerical Tests, J. Clim., 21, 3642–3659, https://doi.org/10.1175/2008JCLI2105.1, 2008.

Murray, B. J., O'Sullivan, D., Atkinson, J. D., and Webb, M. E.: Ice nucleation by particles immersed in supercooled cloud droplets, Chemical Society Reviews, 41, 6519–6554, https://doi.org/10.1039/c2cs35200a, 2012.

NASA Earthdata: MODIS (Aqua/Terra) Cloud Properties Level 3 daily, 1x1 degree grid (MCD06COSP_D3_MODIS) catalog entry, Online catalog entry, https://www.earthdata.nasa.gov/data/catalog/laads-mcd06cosp-d3-modis-6.2, accessed 2026-02-27, 2022.

NASA LAADS DAAC: MCD06COSP_D3_MODIS: MODIS (Aqua/Terra) Cloud Properties Level 3 daily, 1x1 degree grid (Collection 6.2) product description, Online product page, https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MCD06COSP_D3_MODIS/, accessed 2026-02-27, 2022.

Penny, S. G., Miyoshi, T., and Kalnay, E.: Localization of the ensemble transform Kalman filter, Monthly Weather Review, 141, 1601–1621, https://doi.org/10.1175/MWR-D-12-00091.1, 2013.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.



Pincus, R. et al.: Updated observations of clouds by MODIS for global model assessment, Earth System Science Data, 15, 2483–2510, https://doi.org/10.5194/essd-15-2483-2023, 2023.

Platnick, S.: Vertical photon transport in cloud remote sensing problems, Journal of Geophysical Research: Atmospheres, 105, 22 919–22 935, https://doi.org/10.1029/2000JD900333, 2000.

Randall, D., Khairoutdinov, M., Arakawa, A., and Grabowski, W.: Breaking the parameterization deadlock, Bull. Am. Meteorol. Soc., 84, 1547–1564, 2003.

Redemann, J. et al.: An overview of the ORACLES (ObseRvations of Aerosols above CLouds and their intEractionS) project: aerosol-cloud-radiation interactions in the southeast Atlantic basin, Atmospheric Chemistry and Physics, 21, 1507–1563, https://doi.org/10.5194/acp-21-1507-2021, 2021.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, arXiv preprint arXiv:1505.04597, 2015.

Rusli, S. P., Donovan, D. P., Russchenberg, H. W. J., et al.: Simultaneous and synergistic profiling of cloud and drizzle properties using ground-based observations, Atmospheric Measurement Techniques, 10, 4777–4799, https://doi.org/10.5194/amt-10-4777-2017, 2017.

Russell, L. M. et al.: Eastern Pacific Emitted Aerosol Cloud Experiment, Bulletin of the American Meteorological Society, 94, 709–729, https://doi.org/10.1175/BAMS-D-12-00015.1, 2013.

Salomonson, V., Barnes, W., Maymon, P., Montgomery, H., and Ostrow, H.: MODIS: advanced facility instrument for studies of the Earth as a system, IEEE Transactions on Geoscience and Remote Sensing, 27, 145–153, https://doi.org/10.1109/36.20292, 1989.

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

Song, C., McCoy, D., Molod, A., Aerenson, T., and Barahona, D.: Signatures of aerosol-cloud interactions in GiOcean: a coupled global reanalysis with two-moment cloud microphysics, Atmospheric Chemistry and Physics, 25, 15 567–15 592, https://doi.org/10.5194/acp-25-15567-2025, 2025.

Sorooshian, A. et al.: A multi-year data set on aerosol-cloud-precipitation-meteorology interactions for marine stratocumulus clouds, Scientific Data, 5, 180 026, https://doi.org/10.1038/sdata.2018.26, 2018.

Sorooshian, A. et al.: Aerosol-cloud-meteorology interaction airborne field investigations: Using lessons learned from the U.S. West Coast in the design of ACTIVATE off the U.S. east coast, Bulletin of the American Meteorological Society, 100, 1511–1528, https://doi.org/10.1175/BAMS-D-18-0100.1, 2019.

Vergara-Temprado, J., Miltenberger, A. K., Furtado, K., Grosvenor, D. P., Shipway, B. J., Hill, A. A., Wilkinson, J. M., Field, P. R., Murray, B. J., and Carslaw, K. S.: Strong control of Southern Ocean cloud reflectivity by ice-nucleating particles, Proc. Nat. Acad. Sci., 115, 2687–2692, https://doi.org/10.1073/pnas.1721627115, 2018.

Virtanen, P., Gommers, R., and Oliphant, T. E.: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Wang, Y., Zhu, Y., Rosenfeld, D., Wang, M., and Lu, X.: Satellite-Retrieved Cloud Base Droplet Number Concentration Improved by In-Cloud Adiabatic Fraction, Journal of Geophysical Research: Atmospheres, 130, https://doi.org/10.1029/2025jd043789, 2025.

https://doi.org/10.5194/egusphere-2026-6341
Preprint. Discussion started: 10 November 2026
© Author(s) 2026. CC BY 4.0 License.



Wood, R. et al.: The VAMOS Ocean-Cloud-Atmosphere-Land Study Regional Experiment (VOCALS-REx): goals, platforms, and field operations, Atmospheric Chemistry and Physics, 11, 627–654, https://doi.org/10.5194/acp-11-627-2011, 2011.

Wu, Y. and He, K.: Group Normalization, arXiv preprint arXiv:1803.08494, 2018.

Xie, H., Yao, H., Sun, X., Zhou, S., and Zhang, S.: Pix2Vox: Context-Aware 3D Reconstruction From Single and Multi-View Images, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 2690–2698, 2019.

Yu, F., Luo, G., Nair, A. A., Tsigaridis, K., and Bauer, S. E.: Use of Machine Learning to Reduce Uncertainties in Particle Number Concentration and Aerosol Indirect Radiative Forcing Predicted by Climate Models, Geophysical Research Letters, 49, https://doi.org/10.1029/2022gl098551, 2022.

Zeider, K. et al.: Constraining cloud droplet number concentration in marine boundary layer clouds from in situ observations and MODIS retrievals, Atmospheric Chemistry and Physics, 24, 9059–9083, https://doi.org/10.5194/acp-24-9059-2024, 2024.

Zeng, S., Riedi, J., Trepte, C. R., Winker, D. M., and Hu, Y. X.: Study of global cloud droplet number concentration with A-Train satellites, Atmos. Chem. Phys., 14, 7125–7134, https://doi.org/10.5194/acp-14-7125-2014, 2014.

Zhang, D., Vogelmann, A. M., Yang, F., Luke, E., Kollias, P., Wang, Z., Wu, P., Gustafson, W. I., et al.: Evaluation of four ground-based retrievals of cloud droplet number concentration in marine stratocumulus with aircraft in situ measurements, Atmospheric Measurement Techniques, 16, 5827–5846, https://doi.org/10.5194/amt-16-5827-2023, 2023.

Zhang, J., Reid, J. S., Christensen, M., and Benedetti, A.: An evaluation of the impact of aerosol particles on weather forecasts from a biomass burning aerosol event over the Midwestern United States: observational-based analysis of surface temperature, Atmospheric Chemistry and Physics, 16, 6475–6494, https://doi.org/10.5194/acp-16-6475-2016, 2016.