Evaluation of Four Ground-based Retrievals of Cloud Droplet Number Concentration in Marine Stratocumulus with Aircraft *In Situ* Measurements

Damao Zhang[1], Andrew M. Vogelmann[2], Fan Yang[2], Edward Luke[2], Pavlos Kollias[2, 3], Zhien Wang[4], Peng Wu[1], William I. Gustafson Jr.[1], Fan Mei[1], Susanne Glienke[1], Jason Tomlinson[1], and Neel Desai[5]

[1]Pacific Northwest National Laboratory, Richland, Washington, USA
[2]Brookhaven National Laboratory, Upton, New York, USA
[3]School of Marine and Atmospheric Sciences, Stony Brook University, New York, USA
[4]College of Arts and Sciences, University of Colorado Boulder, CO, USA
[5]Department of Meteorology and Climate Science, San Jose State University, CA, USA

Correspondence to: Damao Zhang (damao.zhang@pnnl.gov)

Abstract. Cloud droplet number concentration (N_d) is crucial for understanding aerosol-cloud interactions (ACI) and associated radiative effects. We present evaluations of four ground-based N_d retrievals based on comprehensive datasets from the Atmospheric Radiation Measurements (ARM) Aerosol and Cloud Experiments in the Eastern North Atlantic (ACE-ENA) field campaign. The N_d retrieval methods use ARM ENA observatory ground-based remote sensing observations from a Micropulse lidar, Raman lidar, cloud radar, and the ARM NDROP Value-added Product (VAP), all of which also retrieve cloud effective radius (r_e). The retrievals are compared against aircraft measurements from the Fast-Cloud Droplet Probe (FCDP) and the Cloud and Aerosol Spectrometer (CAS) obtained from low-level marine boundary layer clouds on 12 flight days during summer and winter seasons. Additionally, the *in situ* measurements are used to validate the assumptions and characterizations used in the retrieval algorithms. Statistical comparisons of the probability distribution function (PDF) of the N_d and cloud r_e retrievals with aircraft measurements demonstrate that these retrievals align well with *in situ* measurements for overcast clouds, but they may substantially differ for broken clouds or clouds with low liquid water path (LWP). The retrievals are applied to four years of ground-based remote sensing measurements of overcast marine boundary layer clouds at the ARM ENA observatory to find that N_d (r_e) values exhibit seasonal variations, with higher (lower) values during the summer season and lower (higher) values during the winter season. The ensemble of various retrievals using different measurements and retrieval algorithms such as those in this paper can help to quantify N_d retrieval uncertainties and identify reliable N_d retrieval scenarios. Of the retrieval methods, we recommend using the Micropulse lidar-based method. This method has good agreement with *in situ* measurements, less sensitivity to issues arising from precipitation and low cloud LWP/optical depth, and broad applicability by functioning for both day and nighttime conditions.

1 Introduction

Clouds play a crucial role in regulating the energy balance and water cycle of the Earth (Stephens et al., 2012). By reflecting incoming solar radiation back to space (the 'albedo effect') and trapping outgoing longwave radiation (the 'greenhouse effect'), they cause both cooling and warming effects on Earth's climate. On a global scale, clouds have a net cooling effect of approximately 20 Wm^{-2}, which is more than five times greater than the warming effect caused by doubling the concentration of atmospheric CO_2 (IPCC 2021). Hence, even small changes in cloud properties, such as those induced by anthropogenic activities like aerosol emissions, can significantly impact Earth's climate sensitivity (Zelinka et al., 2017). Aerosols indirectly affect cloud properties by serving as cloud condensation nuclei (CCN) or ice nucleation particles. Such effects can increase the concentration of cloud droplets (N_d) and decrease their sizes, which can substantially alter cloud radiative properties and precipitation efficiency (Twomey 1977; Albrecht 1989). Recent studies have also revealed that aerosol-cloud interactions (ACI) are strongly influenced by atmospheric dynamics and thermodynamic conditions, as well as the physical properties and chemical compositions of aerosols (Chen et al., 2016; Fan et al., 2016). The uncertainty in the magnitude of ACI remains the largest source of uncertainty in estimates of climate forcing (IPCC 2021; Regayre et al., 2014). N_d, which is a direct link between cloud properties and aerosol concentrations, is of utmost importance in improving our understanding of ACI processes and quantifying their effective radiative forcing (Rosenfeld et al., 2019).

To improve the representation of clouds in weather and climate models, it is essential to validate modeled N_d against observations (Storelvmo et al., 2006; Moore et al., 2013; Gryspeerdt et al., 2017). Although aircraft *in situ* instruments can measure N_d directly, these measurements are limited to specific regions and time periods during field campaigns. Collecting a large N_d database from these measurements is a challenging task, making it difficult to statistically study factors that influence the spatial and temporal variations of N_d and ACI processes across different climate zones and atmospheric thermodynamic conditions. Ground-based and space-borne remote sensing techniques provide continuous observations of clouds and aerosols across different regions, and the latter includes global scales. Remote sensing measurements have been widely used to retrieve aerosol and cloud properties including N_d. Grosvenor et al. (2018) comprehensively reviewed passive satellite remote sensing retrievals of N_d from the retrieved cloud optical depth, cloud droplet effective radius (r_e), and cloud-top temperature. They concluded that satellite N_d retrievals could achieve a relative uncertainty of 78% at the pixel-level for single-layer warm stratiform and optically thick clouds. Ground-based N_d retrievals have higher temporal and spatial resolutions than satellite measurements. By taking advantage of more reliable retrievals of liquid water path (LWP) from passive microwave radiometers, ground-based N_d retrievals usually use cloud optical depth and LWP instead of r_e in the retrieval algorithms. These remote sensing data provide invaluable information for statistically studying ACI processes and have been used to validate and improve cloud representations in climate models (McComiskey et al., 2009; Rosenfeld et al., 2019; McCoy et al., 2020).

Passive remote sensing N_d retrievals, such as noted above, commonly rely on reflected or transmitted sunlight measured from spaceborne and ground-based remote sensors, respectively. Therefore, these retrievals are limited to single-layer and optically thick clouds under conditions when the Sun is high in the sky. These limitations can be alleviated by using active remote sensing measurements. Active remote sensors transmit electromagnetic waves at a specific visible, infrared, or microwave wavelength and receive reflected signals from the atmosphere in a narrow field-of-view. Therefore, active remote sensing measurements can be used for cloud property retrievals anytime (i.e., including nighttime) and under much broader atmospheric conditions (e.g., beneath cirrus cloud decks).

Ground-based active remote sensing N_d retrievals use either the cloud radar reflectivity factor (Z) or lidar extinction coefficient (β_e) profiles together with microwave radiometer-retrieved LWP. A monomodal droplet size distribution (DSD) is usually assumed to connect these measured quantities. Radar-based N_d retrievals use the relationships between Z, liquid water content (LWC), DSD, and N_d (Dong et al., 1998; Mace and Sassen 2000; Wu et al., 2020). Since Z is proportional to the sixth power of the DSD, radar-based N_d retrievals are very sensitive to the assumed DSD, and it is challenging to retrieve N_d under drizzling conditions. Recently, lidar-based N_d retrievals have been developed by synergizing multiple instruments similar to the radar-based retrievals (Boers et al., 2006, Martucci and O'Dowd, 2011; Snider et al., 2017; Zhang et al., 2019) by using dual-field-of-view lidar extinction profiles (Schmidt et al., 2013) or by using lidar multiple scattering measurements (Donovan et al., 2015). Since lidar measurements are proportional to the second moment of cloud DSD, lidar-based N_d retrievals are more sensitive to N_d than radar-based methods and have the potential to provide more accurate retrievals.

In the past decade, there has been significant progress in developing N_d retrieval algorithms; however, the validation of these algorithms against *in situ* measurements is still inadequate. Most N_d retrieval methods were developed and tested under specific conditions, making it crucial to evaluate their performance against *in situ* measurements from different locations and cloud conditions to understand better their uncertainties and to confidently extend these algorithms. The Department of Energy (DOE) Atmospheric Radiation Measurements (ARM) Aerosol and Cloud Experiments in the Eastern North Atlantic (ACE-ENA) field campaign offers an excellent opportunity to validate different N_d retrieval algorithms under the same range of cloud conditions. The ACE-ENA campaign (Wang et al., 2022) collected comprehensive data sets from the ARM Eastern North Atlantic (ENA) site, where the ARM Aerial Facility (AAF) research aircraft made *in situ* measurements over the Azores where the ENA atmospheric observatory routinely makes measurements from state-of-the-art remote sensing instruments. The flights during ACE-ENA were designed to take full advantage of the synergy between aircraft *in situ* measurements and ARM ground-based remote sensing observations. In this study, four N_d retrievals are evaluated, considering their potential for operational applications and ease of use across different locations. These methods cover major ground-based N_d retrieval algorithms including two lidar-based retrievals similar to Snider et al. (2017), a radar-based retrieval similar to Wu et al. (2020a), and the N_d retrieval from the ARM Droplet Number Concentration (NDROP) Value-Added Product (VAP) available at https://www.arm.gov/capabilities/vaps/ndrop. We did not include lidar-based N_d retrievals

that either utilize dual-field-of-view lidar extinction profiles or rely on depolarization measurements from lidar multiple scattering. This is due to the specific requirements of the dual-field-of-view lidar configuration and the substantial calibration efforts needed for lidar depolarization measurements. This study evaluates the N_d retrieval algorithms against *in situ* data to enhance our understanding of their uncertainties and extend their application to other locations.

The paper is organized as following: Section 2 presents a brief introduction of the ARM ENA site, the lidar- and radar-based retrieval algorithms, the ARM NDROP VAP, and the ACE-ENA field campaign measurements; Section 3 shows evaluations of N_d retrievals with *in situ* probe measurements during the ACE-ENA field campaign, and a four-year climatology of overcast marine boundary layer (MBL) cloud N_d climatology based on retrievals at the ENA observatory; and Section 4 presents the summary and conclusions.

2 Ground-based N_d Retrievals and ACE-ENA Measurements

The lidar-based retrievals, radar-based retrieval, and the ARM NDROP VAP use different remote sensing measurements and algorithms to retrieve N_d. Brief descriptions of these methods are presented in sections 2.2-2.4. These retrieval methods use both passive and active remote sensing measurements. We expect the ensemble of these peer-reviewed retrievals for the same cloud to indicate a reasonable range of the retrieved N_d. We refined the lidar-based retrieval method in section 2.2. Then we evaluated assumptions in each retrieval method, and for the first time, compared four different N_d retrievals with *in situ* measurements to evaluate the robustness of their performances.

2.1 The ARM ENA Atmospheric Observatory

Established in October 2013, the ARM ENA atmospheric observatory is located on Graciosa Island in the Azores, Portugal, at 39° 5′ 29.76″ N, 28° 1′ 32.52″ W. This region of the northeastern Atlantic Ocean is characterized by the presence of marine stratocumulus clouds and is subject to diverse meteorological and aerosol conditions (Wood et al. 2015). Consequently, the ARM ENA site presents an ideal opportunity to study the properties of clouds and precipitation in a remote marine environment as well as the response of low clouds to natural and anthropogenic aerosols and meteorological conditions. Facilitating these studies, the ARM ENA atmospheric observatory has been equipped with a large array of advanced instruments, capable of providing high spatial and temporal resolution measurements of the atmospheric state, aerosols, clouds, precipitation, and radiation budget. These instruments include a variety of aerosol instrumentation, lidars (Muradyan and Richard 2020; Newsom et al., 2022), radars (Johnson et al., 2022), radiometers (Cadeddu 2021; Hodges and Michalsky 2016), as well as the balloon-borne sounding (SONDE) system(Holdridge 2020). Table 1 lists the key ground-based instruments and their measurements which were used for N_d retrievals in this study.

2.2 Lidar-based N_d Retrieval

In this study, Raman Lidar (RL) and Micropulse lidar (MPL) data are used in separate lidar-based retrievals. The method for retrieving N_d employs the interrelationships among N_d, β_e, LWC, and cloud DSD, where β_e is the extinction coefficient (Snider et al., 2017). At an altitude z above the cloud base, N_d, β_e, and LWC can be expressed as functions of the cloud DSD:

$$N_{d,z} = \int_0^\infty n_{d,z} dr \quad (1)$$

$$\beta_{e,z} = Q_{ext}\pi \int_0^\infty n_{d,z} r^2 dr \quad (2)$$

$$LWC_z = \frac{4}{3}\pi\rho_w \int_0^\infty n_{d,z} r^3 dr \quad (3)$$

where Q_{ext} is the extinction efficiency, r is the cloud droplet radius, $n_{d,z}$ is the droplet number concentration within the size range between r and $r+dr$, and ρ_w is the density of liquid water. Since water droplet sizes are much larger than the lidar laser wavelength, $Q_{ext} \approx 2$. The cloud droplet effective radius $r_{e,z}$ is defined as:

$$r_{e,z} = \frac{\int_0^\infty n_{d,z} r^3 dr}{\int_0^\infty n_{d,z} r^2 dr} = \frac{3Q_{ext}}{4\rho_w}\frac{LWC_z}{\beta_{e,z}} \quad (4)$$

To establish a connection between the properties that are a function of the second and third moment of the cloud DSD, respectively $\beta_{e,z}$ and LWC_z, previous research has made the assumption that the cloud DSD follows either a Gamma distribution or lognormal distribution and has a constant spectrum width (Martucci and O'Dowd, 2011; Snider et al., 2017). Drawing inspiration from the passive remote sensing retrieval algorithms outlined by McComiskey et al. (2009), an empirical parameter k is introduced to link $\beta_{e,z}$ and LWC_z that is a measure of the width of the cloud DSD. This parameter represents the cube of the ratio between the volume radius and the effective radius:

$$k = \frac{1}{N_{d,z}} \int_0^\infty n_{d,z} r^3 dr / r_{e,z}^3 \quad (5)$$

To determine N_d, the k parameter is assumed to remain constant vertically within the cloud (Brenguier et al., 2011). Through the analysis of aircraft *in situ* probe measurements from five distinct field experiments, Brenguier et al. (2011) demonstrated that the k parameter values range from 0.7-0.9, with uncertainties between 10% and 14% across different cloud systems and various atmospheric conditions. By integrating equations (2), (3), and (5), $N_{d,z}$ can be derived as a function of $\beta_{e,z}$ and $\mathrm{LWC_z}$:

$$N_{d,z} = \frac{2\rho_w^2}{9\pi k}\frac{\beta_{e,z}^3}{LWC_z^2} \quad (6)$$

The newly derived Equation (6) eliminates the need for assuming a specific DSD shape (e.g., Gamma or lognormal distribution) that was necessary in previous studies.

To derive N_d, the $\mathrm{LWC_z}$ in stratiform cloud is typically assumed to be a constant fraction (f_{ad}) of its adiabatic value ($\mathrm{LWC_{z,ad}}$): $LWC_z = f_{ad}\, LWC_{z,ad}$. The $LWC_{z,ad}$ profile can be determined from cloud-base temperature and pressure measurements. By analyzing two years of ground-based remote sensing data set at Leipzig, Germany, Merk et al. (2016)

shows that f_{ad} values are 0.63±0.22. In this study, f_{ad} is calculated as the ratio of the retrieved LWP from the MWRRETv2 VAP (https://www.arm.gov/capabilities/science-data-products/vaps/mwrretv2) to the LWP calculated from the adiabatic LWC profile. The MWRRETv2 VAP retrieves LWP from microwave radiometer brightness temperature measurements at 23.8, 31.4, and 90 GHz using the retrieval algorithm developed by Turner et al. (2007). The third channel at 90 GHz provides additional sensitivity to liquid water enabling an LWP uncertainty of ± 10-15 g/m^2 (Cadeddu et al., 2013).

Advanced lidar systems, such as the RL and High Spectral Resolution Lidar, are absolutely calibrated by referencing to molecular scattering. These systems offer reliable estimates of particulate backscatter and extinction coefficients by solving the lidar equation (Thorsen and Fu, 2015; Marais et al., 2016). Our RL retrieval use the RL-estimated $\beta_{e,z}$ from the ARM Raman Lidar Profiles – Feature detection and Extinction (RLPROF-FEX) VAP (https://www.arm.gov/capabilities/science-data-products/vaps/rlprof-fex), which computes $\beta_{e,z}$ using the algorithms developed by Thorsen et al. (2015) and Thorsen and Fu (2015). However, due to the weak strength of the Raman scattering compared to the elastic scattering, noise poses a considerable challenge for the extinction coefficient retrieval. To enhance the signal-to-noise (SNR) ratio, the fine-resolution RL data at 10 s temporal and 7.5 m vertical resolution are aggregated coarser resolutions of 2 min and 30 m, respectively. While enhancing the SNR, this coarser resolution RLPROF-FEX $\beta_{e,z}$ may introduce additional uncertainty in N_d retrievals for broken clouds. It is important to note that advanced lidar systems are more costly and, as a result, are not widely available.

In contrast, elastic-scattering lidars, such as the MPL and ceilometer, are available at all ARM observatories and numerous locations worldwide including the MPLNET and Cloudnet (Welton et al., 2001; Illingworth et al., 2007). These instruments provide high temporal and vertical measurements of the strong elastic scattering from atmospheric particles. However, elastic-scattering lidar measurements cannot be directly used to derive particulate backscatter and extinction coefficients since there is only one lidar equation (measurement) for these two variables, i.e., one equation two unknowns. This issue is often addressed using the lidar extinction-to-backscatter ratio (S), which represents the relationship between particulate backscatter and extinction coefficients. Once S is determined, the lidar $\beta_{e,z}$ can be inverted from the MPL backscatter intensity measurements by analytically solving the lidar equation using the inversion method developed by Klett (1981) and Fernald (1984). For liquid cloud droplets, S is approximately 18.8 (O'Connor et al., 2004; Thorsen and Fu 2015). To account for multiple scattering from liquid droplets, a multiple-scattering correction scheme developed by Hogan (2008) is applied. Sarna et al. (2021) demonstrated that, after all corrections to elastic-scattering lidar signals, the inversion method could obtain $\beta_{e,z}$ with an error less than 5% within 90 m above cloud base at the lidar wavelength of 355 nm. We assess the sensitivity and reliability of lidar-based N_d retrievals using both RL- and MPL-estimated $\beta_{e,z}$.

It should be noted that N_d retrievals at cloud base (Z_{cb}) are adversely impacted by noise introduced by turbulent mixing. Entrainment mixing may cause LWC_{cb} to deviate significantly from the adiabatic value, resulting in considerable differences between the retrieved $N_{d,cb}$ and $N_{d,z}$ above the Z_{cb}. Furthermore, lidar can only penetrate the low portion of the liquid cloud due to the strong attenuation by liquid droplets. The signal becomes fully attenuated when the optical depth reaches ~3 which corresponds to 100 to 300 m above the Z_{cb}. Consequently, our retrievals use $\beta_{e,z}$ and LWC_z only within the range between Z_{cb} + 30 m and Z_{cb} + 90 m. For lightly drizzling maritime stratocumulus clouds, such as those with the column maximum radar reflectivity (Z_e) < 0 dBZ, the contribution of drizzle particles to lidar extinction is negligible compared to that from liquid droplets; thus, the lidar-based retrievals can still be employed. Based on equation (4) and the assumptions that the cloud maintains a constant fraction of its adiabatic value and N_d remains vertically constant, r_e for the rest of the cloud layer can be estimated.

Using a similar lidar-based N_d retrieval approach, Snider et al. (2017) discovered that, in general, the lidar-based N_d retrievals were smaller than *in situ* probe measurements during the VAMOS Ocean–Cloud–Aerosol–Land Study (VOCALS) Regional Experiment over the southeastern Pacific (Wood et al. 2011). It is worth noting that Snider et al. (2017) used the adiabatic LWC lapse rate without considering the subadiabaticity, which results in an overestimation of LWC_z and consequently an underestimation of N_d based on equation (6). Therefore, in the present study, the bias of the N_d retrieval should not be as large since we consider cloud subadiabaticity.

2.3 Radar-based N_d Retrieval

Obtaining N_d values from radar reflectivity poses challenges due to the frequent presence of drizzle within MBL clouds, which subsequently contributes significantly to the measured radar reflectivity (Zhu et al., 2022). Wu et al. (2020a) recently developed a method to separate drizzle and cloud droplet contributions to the measured radar reflectivity while simultaneously retrieving cloud and drizzle microphysical properties, including N_d, in precipitating MBL clouds. They distinguish between drizzle and cloud droplet contributions by identifying the height where Z_e exceeds -15 dBZ, when moving downward from the cloud top. This height marks the initiation of drizzle where, above this point, the measured Z_e is solely attributed to cloud droplets.. While it is convenient to use this threshold, it should be noted that a number of recent studies demonstrate that drizzle having significantly lower reflectivity that -15 dBZ can be observed within stratocumulus clouds (Kollias et al., 2011; Luke at al., 2013; Zhu et al., 2022). The cloud contribution to Z_e at the cloud base is calculated as the difference of Z_e from the radar range gates above and below cloud base. Subsequently, they construct the cloud radar reflectivity (Z_c) by assuming a linear increase in cloud liquid water content (LWC_c) with height above cloud base (and thus a linear increase in $\sqrt{Z}$ if N_d is invariant with height). By assuming that the cloud droplet particle size distribution follows a lognormal distribution with a logarithmic width of σ_x, the relationship between N_d, LWC_c, and Z_c can be expressed as:

$$LWC_c = \frac{\pi}{6}\rho_w \exp(-4.5\sigma_x^2)\sqrt{N_d Z_c} \quad (7)$$

The logarithmic width σ_x is set to 0.38 from Miles et al. (2000). However, under the assumption of a lognormal DSD, a value of 0.38 for σ_x is equivalent to a k value of 0.65. Martin et al. (1994) showed that k ranges from 0.67 ± 0.07 in continental air masses to 0.80 ± 0.07 in the marine ones. Consequently, we adopt σ_x of 0.23, which equates to a k value of 0.86 under a lognormal DSD condition. This is in line with the k value utilized in lidar-based retrievals and the NDROP VAP. ρ_w is the liquid water density. To determine N_d, equation (7) is further constrained by the cloud LWP, derived from the difference between the MWRRETv2 (total) LWP and the calculated drizzle water path, which is obtained from the retrieved drizzle water content profile. Subsequently, the r_e profile is derived from N_d and the LWC_c profile.

To mitigate the impact of MWRRETv2 LWP uncertainties, cloud microphysical property retrievals were smoothed to a temporal resolution of 1 min. A sensitivity analysis conducted by Wu et al. (2020a) revealed that the retrieved N_d values are not sensitive to the selection of the radar reflectivity threshold of -15 dBZ. Using aircraft measurements from the ACE-ENA field campaign as a benchmark, the median N_d retrieval error is approximately ~ 35%.

2.4 The ARM NDROP VAP

The N_d retrieval method employed by the ARM NDROP VAP uses the relationship between LWP, cloud optical depth (τ), cloud DSD, and N_d. Following Lim et al. (2016), the layer-mean N_d can be expressed as:

$$N_d = \left[\frac{2^{-5/2}}{k^*}\right]\left[\frac{3\pi}{5}Q_{ext}\right]^{-3}\left[\frac{3}{4\pi\rho_w}\right]^{-2}\tau^3 LWP^{-5/2}(f_{ad}c_w)^{1/2} \quad (8)$$

k^* is the cloud system k parameter, which is the cube of the ratio between the layer-mean volume radius and the layer-mean effective radius. As both τ and LWP represent vertical integrals through the entire cloud layer, Brenguier et al. (2011) propose using the cloud system k^* parameter in equation (8). Consequently, the NDROP VAP retrievals utilize the cloud system k* parameter, while other methods deploy the local mean k parameter. In the case of a linearly stratified cloud with constant k within the cloud, k^* can be derived as $k^* = 0.864\,k$. The NDROP VAP adopts a k^* value of 0.74, as recommended by Brenguier et al. (2011) (Riihimaki et al., 2021).

The adiabatic LWC lapse rate, c_w, can be calculated using cloud-base temperature and pressure from the ARM INTERPSONDE VAP (https://www.arm.gov/capabilities/vaps/interpsonde) (Jensen and Toto 2016). τ is available from the ARM MFRSRCLDOD VAP (https://www.arm.gov/capabilities/vaps/mfrsrcldod), which retrieves τ for overcast liquid clouds from multifilter rotating shadowband radiometer (MFRSR) measurements using the retrieval algorithm developed by Min and Harrison (1996) (Turner et al., 2021). The MFRSR measures both global and diffuse components of solar irradiance at multiple narrowband channels with a hemispheric viewing geometry. The retrieval algorithm employs the transmitted irradiance at 415 nm from the MFRSR, so the retrieved τ is available only during daytime. The retrieval assumes a single cloud layer comprised of liquid water drops and assumes the surface is not covered with snow or ice. Analyses show that τ

from the MFRSRCLDOD VAP has uncertainties ranging from 0.5 to 2.5 (Turner et al., 2021). LWP is available from the ARM MWRRETv2 VAP as mentioned in section 2.2. The MWR3C has a field-of-view of between 5 and 6º. Since both the τ and LWP retrievals have significant relative uncertainties for optically thin clouds, this retrieval approach should be applied for overcast, optically thick liquid clouds.

Lim et al. (2016) evaluated N_d values retrieved using this approach by comparing them to aircraft *in situ* probe measurements obtained during the Routine ARM Aerial Facility (AAF) Clouds with Low Optical Water Depths (CLOWD) Optical Radiative Observations (RACORO) field campaign at the ARM Southern Great Plains (SGP) site (Vogelmann et al., 2012). Their findings indicate that the retrieved N_d values are substantially larger than the *in situ* measurements. This discrepancy may be attributed to the fact that clouds sampled during the RACORO campaign often exhibited small LWPs. Consequently, NDROP retrievals still require evaluation under optically thick cloud conditions.

For passive remote sensing retrievals, the layer-mean r_e (r_{em}) between the cloud layer top and base can be determined using the relationship among r_{em}, τ, and LWP:

$$r_{em} = \frac{3LWP}{2\rho\tau} \quad (9)$$

r_{em} is available from the ARM Cloud Optical Properties from the Multifilter Shadowband Radiometer (MFRSRCLDOD) VAP (https://www.arm.gov/capabilities/science-data-products/vaps/mfrsrcldod). To compare r_e retrievals from different approaches and with *in situ* measurements, we use r_{em} for comparisons for the rest of the discussion following previous studies (Chiu et al., 2012; Grosvenor et al.,2018). r_{em} is derived by averaging r_e at each layer between the cloud top and base from lidar- and radar-based retrievals, and *in situ* measurements.

2.5 ACE-ENA *In Situ* Measurements

The ACE-ENA field campaign deployed the ARM AAF Gulfstream-159 (G-1) research aircraft over the Azores during the two intensive operational periods (IOPs) in early summer 2017 (June to July) and winter 2018 (January to February). The G-1 was equipped with a range of *in situ* sensors, enabling comprehensive measurements of aerosol particles, cloud droplets, precipitation, and atmospheric conditions. Cloud probes particularly relevant to N_d measurements include the Fast-Cloud Droplet Probe (FCDP) and the Cloud and Aerosol Spectrometer (CAS). The FCDP measures cloud droplets in the diameter size range of 1.5-50 μm with temporal resolution of 1 or 0.1 s. The CAS provides measurements of aerosol or cloud droplets in the 0.5-50 μm diameter size range with a temporal resolution of 1 s. Given the different particle size ranges measured by the various probes, we used *in situ* N_d data for the particle size between 3-50 μm. It is noted that although *in situ* probes provide reliable N_d measurements, they also have uncertainties ranging from 10-30% as presented by Baumgardner et al. (2017). Therefore, we include both FCDP and CAS measurements for evaluating N_d retrievals.

During the ACE-ENA campaign, the G-1 aircraft conducted both vertical profiling flights and horizontal flights at physically important levels, such as near the ocean surface, just below clouds, within clouds, as well as at and above the cloud top (Wang et al., 2022). These flights were specifically designed to maximize synergy between G-1 aircraft measurements and ENA ground-based remote sensing observations, offering an ideal dataset for evaluating N_d retrievals. Most G-1 flights employed an L-shaped pattern, including both upwind and crosswind legs at different altitudes, with the L "corner" over the ENA site. Additionally, four G-1 flights used a "Lagrangian drift" pattern, starting upwind of the ENA site and performing crosswind measurements while drifting with the prevailing boundary layer winds (Wang et al., 2022). In total, 39 flights were conducted during the two IOPs.

Among those flights, 12 flight days featuring multiple in-cloud flight legs under single-layer stratiform cloud conditions were selected for this study. Each selected day had at least one complete traversal from the cloud base to the cloud top. Heavily drizzling stratocumulus flight days were excluded. Table 2 provides the date, in-cloud flight time, and cloud conditions for the 12 flight days. In-cloud measurements are defined as those when the FCDP-measured N_d values are larger than 10 cm^{-3}. Approximately 11 total hours of in-cloud flight measurements were used to evaluate the N_d retrievals. Figure 1 illustrates cloud properties of the 12 selected flight days, including fractional sky cover, cloud-base height, cloud depth, LWP, and column-maximum Z_e (Z_{e_max}) derived from the ENA ground-based remote sensing observations. The box-and-whisker plots display the 5th, 25th, 50th, 75th, and 95th percentiles. Of the 12 selected flight days, nine have overcast cloud conditions and three have broken cloud conditions (6/28/2017, 7/8/2017, 2/12/2018). These three broken cloud days had among the smallest LWPs, as shown in Figure 1d. Cloud-base heights ranged from 0.5 to 1.5 km with variations often smaller than 0.3 km on a given flight day. Overall, these clouds had LWPs less than 200 g/m^2 and Z_{e_max} values smaller than 0 dBZ, which are typical for marine low-level clouds.

3 Results and Discussions

3.1 Evaluation of Retrieval Assumptions

In the lidar-based, radar-based, and NDROP VAP N_d retrievals, several assumptions are made regarding the vertical N_d variation, the k and k^* parameter, and the LWC profile, as described in sections 2.2 and 2.3. We test these assumptions in this section. Figure 2 presents statistics of these cloud properties from *in situ* and ground-based measurements during the 12 selected flight days. For example, Figure 2a shows that the mean N_d normalized by the flight average is close to 1, with standard deviations of approximately 0.4 though the cloud layer, which supports the assumption that N_d can be treated as constant within the cloud layer.

The k parameter is assumed to be vertically constant. Some previous studies find that the k parameter increases with height (Brenguier et al., 2011), while others suggest that the k parameter can either increase or decrease with height (Pawlowska et

al., 2006; Painemal and Zuidema 2010). Our analysis shows that the mean k parameter remains essentially constant with height (Figure 2b). The probability distribution functions (PDF) of the k and k^* parameters (Figure 2c) reveal that k (k^*) ranges between 0.6 (0.5) and 1.0 (0.86), with a mean value of 0.86 (0.74) and a standard deviation of 0.10 (0.09). Since N_d is inversely proportional to k (k^*) as shown in equation (6) and (8), an uncertainty of 0.10 in k value alone could cause an uncertainty of ~12% in the retrieved N_d value from the uncertainty propagation analysis. The lidar-based N_d retrievals in this study use a k value of 0.86. We note that the k^* value of 0.74 used in the NDROP VAP is well justified. As the k value of 0.86 corresponds exactly to the recommended k^* value of 0.74 by Brenguier et al. (2011), where their value is based on data from five field program locations, it suggests that a k value of 0.86 might be more broadly applicable for lidar-based N_d retrievals of boundary layer clouds at other locations.

As aircraft *in situ* probes are unable to provide continuous cloud-base height measurements, and the LWC_{ad} or c_w profile is sensitive to cloud-base height, it is challenging to determine f_{ad} and its vertical variations within a cloud. Instead, we use the ratio of the MWRRETv2 LWP to the computed adiabatic LWP to calculate f_{ad}. As seen in Figure 2d, The LWP from WMRRET and the adiabatic LWP show a strong correlation, evidenced by a Pearson correlation coefficient of 0.85. Adiabatic LWPs are generally larger than MWRRETv2 LWPs, especially when the LWP is above 150 g/m^2, but they correlate well. The PDF of f_{ad} shows that f_{ad} has a mean value of 0.76, with a standard deviation of 0.42 (Figure 2e). Since cloud LWP should not exceed the adiabatic LWP, f_{ad} is set to 1 when it is larger than 1. Those values, and possibly those at the extreme lower end of f_{ad}, appear to be affected primarily by uncertainties in cloud thickness for thin clouds (< 200 m) and by uncertainties in low MWRRETv2 LWPs (< 75 g/m^2), based on scatter plots of f_{ad} vs. these respective properties (not shown). On the other end, when the LWP is above ~150 g/m^2, the cloud could contain drizzle. The current MWRRETv2 retrieval does not account for drizzle scattering effects at frequencies above 90 GHz, which could cause overestimation of the LWP by 10-15%, as outlined in the study by Cadeddu et al. (2020). We did not implement corrections to this bias due to two reasons: firstly, there are currently no reliable methods to correct such bias; secondly, we removed strong drizzling stratocumulus cases by excluding clouds with Z_{e_max} larger than 0 dBZ as discussed in section 2.

3.2 Evaluation of N_d Retrievals

For convenience, we label N_d (r_{em}) retrievals from the MPL, RL, KAZR radar measurements, and from the NDROP (MFRSRCLDOD) VAP as N_{d_mpl} (r_{em_mpl}), N_{d_rl} (r_{em_rl}), N_{d_radar} (r_{em_radar}), N_{d_vap}(r_{em_vap}), respectively, and *in situ* measured N_d (r_{em}) from FCDP and CAS as N_{d_FCDP} (r_{em_FCDP}), N_{d_cas} (r_{em_CAS}). Figure 3 shows an example of ground-based remote sensing measurements and N_d and r_e retrievals on January 26, 2018. The cloud is a typical stratiform MBL cloud with a cloud-base height of ~1.1 km, and a cloud-top height of ~1.5 km. The cloud system persisted for more than 55 hours from ~5:00 UTC January 25 to ~12 UTC January 27 (full period not shown in Figure 3). From the mean sea level pressure distribution (Fig. S1), the Azores high was located to the northeast of the Azores. Near-surface winds were south to southeast across the ENA

observatory. The synoptic environment created a strong stable boundary layer condition that was favorable for the maintenance of marine boundary layer stratocumulus. From Figure 3a and b, large β_e and its rapid attenuation indicate the presence of the liquid layer. Figure 3c shows radar reflectivity up to -20 dBZ below the liquid layer, indicating that drizzle frequently forms and falls out of the liquid layer. The mean MWRRETv2 LWP (LWP_{mwr}) and calculated adiabatic cloud LWP (LW_{ad}) are 107 g/m^2 and 119 g/m^2, respectively. Figure 3d shows that LWP_{ad} and LWP_{mwr} correlate very well and are close in magnitude, indicating the cloud is nearly adiabatic. Retrieved N_{d_mpl}, N_{d_rl}, N_{d_radar}, N_{d_vap} are shown in Figure 3e. For this case, N_{d_mpl} and N_{d_radar} have a similar magnitude at ~ 50 cm^{-3} but are smaller than N_{d_rl} and N_{d_vap}. Derived r_{em_mpl}, r_{em_rl}, r_{em_radar}, and r_{em_vap} are very close at ~ 10.5 μm as shown in Figure 3f.

This case is one of the four "Lagrangian drift" flights during the entire ACE-ENA field campaign. The prevailing boundary layer winds were south- to southeast ward. Boundary layer wind speeds were generally less than 10 m/s, based on radiosonde measurements at 11:30 UTC at the ENA observatory (Figure 4a). The G-1 aircraft took off at approximately 11:05 UTC upwind of the ENA observatory and landed at around 15:00 UTC (Figure 4). During the four-hour flight, the G-1 aircraft made about 1 hour and 37 minutes of in-cloud measurements, including several horizontal legs just below cloud top, within the cloud layer, and just above cloud base, as well as several spirals. Satellite imagery from the Moderate Resolution Imaging Spectroradiometer (MODIS) shows that closed-cellular stratocumulus clouds dominated the region (Figure 4b).

Due to the continuous movement of the G-1 aircraft near the ENA observatory, establishing direct one-to-one comparisons between ground-based retrievals and aircraft *in situ* measurements is challenging. Instead, we evaluate the PDFs of N_d retrievals against those from the aircraft *in situ* measurements. Figure 5a shows the comparison of ground-based N_d retrievals N_{d_mpl}, N_{d_rl}, N_{d_radar}, and N_{d_vap} during the time of the concurrent aircraft flight against *in situ* FCDP (N_{d_FCDP}) and CAS (N_{d_CAS}) measurements for the case on January 26, 2018. It should be noted that measurements from the two *in situ* probes show slightly different N_d distributions. N_{d_CAS} is generally less than N_{d_FCDP} with a median of 74 cm^{-3} and a narrower distribution with a standard deviation of 24 cm^{-3}, while N_{d_FCDP} has a median of 97 cm^{-3} and a standard deviation of 34 cm^{-3}. Among the four N_d retrievals, N_{d_mpl} shows a very similar distribution to N_{d_cas}, with a median of 73 cm^{-3} and a standard deviation of 31 cm^{-3}. N_{d_rl} shows a broader distribution with a median of 114 cm^{-3} and a standard deviation of 71 cm^{-3}, probably because the retrieved RL β_e has a larger random noise than that of MPL β_e. N_{d_radar} has the narrowest distribution, with a median of 62 cm^{-3} and a standard deviation of 13 cm^{-3}. N_{d_vap} retrieval exhibits the highest values, with a median of 127 cm^{-3} and a standard deviation of 46 cm^{-3}.

As r_e changes with distance above cloud base and it is challenging to know instantaneous cloud-base height from aircraft measurements, it is more difficult to conduct one-to-one comparisons between ground-based r_e retrievals and aircraft *in situ* measurements. Therefore, we compare PDFs of the r_{em} retrievals against aircraft *in situ* measurements during all in-cloud penetrations. Figure 5b shows that the median r_{e_FCDP} and r_{e_CAS} are almost the same, around 10.4 μm. The median (standard

deviation) of r_{em_mpl}, r_{em_rl}, r_{em_radar}, and r_{e_vap} are 11.5 μm (1.9 μm), 10.2 μm (2.6 μm), 9.3 μm (0.7 μm), and 10.5 μm (1.1 μm), respectively. Although median r_{em} values from different retrieval methods are very close, r_{em_mpl} *and* r_{em_rl} have broader distributions than r_{em_radar} and r_{e_vap}.

Figure 6 presents the comparison of ground-based N_d and r_{em} retrievals against *in situ* FCDP and CAS measurements for the 12 selected flight days. Table 3 displays the median N_d of the 12 selected flight days and their relative differences with respect to N_{d_FCDP}. In accordance with prior studies of cloud microphysical properties (Yeom et al., 2021; Zhang et al., 2021), we consider FCDP measurements as the benchmark. The median N_{d_FCDP} for the 12 days ranges from 33 to 125 cm^{-3}. There are substantial variations in N_d from *in situ* measurements among the 12 days, with generally higher N_d observed on summer IOP days, and lower N_d on winter IOP days. This agrees with the analysis in Wang (2022) of all *in situ* N_d measurements during the ACE-ENA field campaign, which reveals that the flight-mean N_d ranges from 20 – 50 cm^{-3}, and that summer IOP N_d is generally larger than that of the winter IOP. Encouragingly, ground-based N_d retrievals generally follow the same seasonal variation trend as, shown in Figure 6a.

Between the two *in situ* probe measurements, N_{d_FCDP} and N_{d_CAS} show good agreement. The median N_d relative differences of N_{d_CAS} with respect to N_{d_FCDP} are smaller than 10% for most flights (Table 3). However, significant differences are observed for several flights, such as on 01/19/2018 and 02/07/2018, when the median N_d relative differences are larger than 40%. N_{d_mpl} compares well with *in situ* probe measurements, with the median N_d relative differences in N_{d_mpl} with respect to N_{d_FCDP} ranging from 9% to 89%. Interestingly, Figure 6a reveals that N_{d_mpl} overestimates N_d during the summer IOP but underestimates N_d during the winter IOP, partially because the k parameter values were smaller (larger) during the summer (winter) IOP than the default value of 0.86 used in the retrieval algorithms (Figure S2). N_{d_rl} compares well with *in situ* probe measurements for overcast clouds but significantly underestimates N_d for broken clouds (06/28/2017, 07/08/2017, and 02/12/2018), which is likely due to the coarse temporal resolution of RLPROF-FEX extinction data. Similar to the 01/26/2018 case, N_{d_radar} values for other flight days consistently have a narrower range and are generally smaller than *in situ* probe measurements. N_{d_vap} considerably overestimates N_d for either broken clouds or when clouds have low LWPs, such as on 06/21/2017, 06/28/2017, and 07/08/2017. For overcast clouds with LWPs greater than ~25 g/m^2, N_{d_vap} compares well with *in situ* probe measurements. Overall, retrieved N_d have a larger spread and poorer comparison with *in situ* probe measurements during the summer IOP than that of the winter IOP, likely because more broken low-level clouds are present during summer at the ENA observatory (Figure 1a).

Figure 6b reveals significant differences in r_e between the two IOPs, with smaller r_e values during the summer IOP and larger r_e values during the winter IOP. This is in line with the differences in Z_{e_max} between two IOP as shown in Figure 1d, since Z_{e_max} is highly sensitive to the presence of large particles. *In situ* probe-derived r_e values are very close to each other, with differences between r_{em_FCDP} and r_{em_CAS} being less than 1 μm for all the 12 selected flight days (Table S1). Retrieved

r_{em_mpl}, r_{em_rl}, r_{em_radar}, and r_{em_vap} all correspond well with the r_{em_FCDP} variations. r_{em_mpl} values are slightly larger than r_{em_FCDP}, with absolute differences usually within 2 μm. This is likely because r_{em_mpl} is calculated assuming a constant subadiabatic LWC profile, leading to a linear increase in r_e from cloud base to cloud top. In reality, cloud r_e increases above cloud base but decreases slightly at cloud top due to entrainment mixing of dry air (Wang et al., 2022). r_{em_rl} compares well with r_{em_FCDP} for most cases but is significantly larger than r_{em_FCDP} for flight days when the retrieved N_{d_rl} values are considerably smaller than N_{d_FCDP} due to broken clouds and the coarse temporal resolution of the RL extinction data. r_{e_radar} values are also within 2 μm of r_{e_FCDP}, which can be either larger or smaller. The values of r_{e_vap} are also slightly greater than those of r_{e_FCDP} in general. This is primarily because r_{e_vap} is calculated from measured LWP and τ, both of which are more heavily influenced by the cloud's upper regions where larger droplet particles are prevalent.

3.3 Implementing N_d Retrievals to multiple years of ENA Data

A significant advantage of ground-based N_d retrievals is their applicability to long-term, continuous, and high temporal resolution remote sensing measurements, facilitating process-level understanding of cloud microphysical properties and their climatology. The N_d retrievals are applied to four years of ground-based remote sensing measurements of overcast MBL clouds at the ENA observatory between 2016 and 2019. MBL clouds are identified as those with base heights lower than 4 km above sea level (ASL). Considering the limitation of RL and NDROP retrievals, we selected single-layer overcast MBL cloud systems that persist longer than 20 minutes with a concurrent total sky imager (TSI) fractional sky cover greater than 95% and an LWP greater than 25 g/m^2. To avoid heavily precipitating cloud systems, we excluded clouds with Z_{e_max} larger than 0 dBZ. Since the RL data and retrievals have the coarsest temporal resolution of 2 min, other retrievals were subsampled to the same temporal resolution as RL data. In total, approximately 245,000 retrieved N_d and r_e data samples were collected.

Figure 7a displays the monthly occurrence of overcast MBL clouds at the ENA observatory meeting the above-stated criteria. The annual mean occurrence of these clouds is approximately 0.26 with higher monthly mean occurrences in June and July, and a lower occurrence during December. The mean MBL cloud occurrence and its seasonal variations align closely with those of the low-level cloud presented in Wu et al. (2020b), which used a similar dataset to study MBL cloud and drizzle properties at the ENA observatory but for cloud top height below 3 km. Monthly N_d statistics are shown in Figure 7b. The annual median N_{d_mpl}, N_{d_rl}, N_{d_radar}, and N_{d_vap} are approximately 79.7, 75.9, 54.4, and 116.9 cm^{-3}, respectively. As with the evaluations for the ACE-ENA field campaign, N_{d_vap} at the ENA observatory are consistently larger than other retrievals. Lim et al. (2016) suggested that unrealistically high N_{d_vap} over 2000 cm^{-3} generally occur when LWP is low. By limiting retrievals to only MBL systems with LWP greater than 25 g/m^2, we do not find N_{d_vap} larger than 500 cm^{-3}. However, the systematically larger N_{d_vap} compared to other retrievals indicates that cloud optical depth retrievals might also be biased by off-zenith clouds, which are not considered in the cloud optical depth retrievals. N_{d_mpl} and N_{d_rl} are generally

very close to each other, suggesting that cloud droplet particulate extinction inversion using either the Fernald method or RL data is reasonably reliable. N_{d_radar} compares well with lidar-based retrievals and has the narrowest distributions each month and the smallest monthly variations. All retrievals show slightly seasonal N_d variations with higher N_d during the summer season and lower N_d during the winter season, consistent with the N_d differences between the summer IOP and winter IOP during the ACE-ENA field campaign, as discussed in section 3.2. Wang et al. (2022) suggested that N_d is positively correlated to the boundary layer accumulation mode aerosol concentration, but the ratio of summer to winter N_d is smaller than the seasonal variations of accumulation mode aerosol concentration. Figure 7c shows the monthly distributions of cloud condensation nuclei (N_{ccn}) at the supersaturation of 0.1% from the ARM CCN counter (CCN-100) of the surface aerosol observing system (AOS). N_{ccn} has similar seasonal variations as N_d with larger values in June and July, and smaller values in December, but its seasonal variations are much larger than those of N_d, consistent with the finding of Wang et al. (2022).

Figure 7d presents the monthly distributions of retrieved r_{em} values. The annual median $r_{em\_mpl,}$ $r_{em\_rl,}$ $r_{em\_radar,}$ and r_{em_vap} are 14.5, 13.8, 10.4, and 11.7 μm, respectively. Both r_{em_mpl} and r_{em_rl} are slightly larger than r_{em_radar} and r_{em_vap}, due to the assumption of a constant subadiabatic LWC profile when calculating r_{em_mpl} and r_{em_rl}, as discussed in section 3.2. Also note that the Wu et al. (2020a) method retrieves cloud and drizzle drop size separately and that r_{em_radar} is the effective radius solely for cloud droplet and does not account for drizzle particle size. Thus, the smaller r_{em_radar} with respect to other retrievals is expected. While r_{em_radar} and r_{em_vap} do not exhibit significant monthly variations, r_{em_mpl} and r_{em_rl} are slightly smaller in June and July and slightly larger in November and December, displaying an opposite seasonal variation pattern compared to that of the N_{d_mpl} and N_{d_rl} in Figure 7b. Figure 7e illustrates the monthly statistics of Z_{e_max}, which shares a similar seasonal variation pattern with r_{em_mpl} and r_{em_rl} in Figure 7d, reinforcing the observed r_{em_mpl} and r_{em_rl} seasonal variation pattern.

4 Summary

Remote sensing techniques offer extensive cloud properties for studying ACI processes and validating climate model simulations. Validating N_d retrieval algorithms against *in situ* probe measurements is needed to understand their uncertainties. The ARM ACE-ENA field campaign offers a unique opportunity to validate four different N_d ground-based retrieval algorithms, which use ENA atmospheric observatory data, against G-1 research aircraft observations, which flew over the Azores during intensive IOPs in early summer 2017 and winter 2018. Twelve flight days under single-layer stratiform low-level cloud conditions were selected, with six days in the summer IOP and six days in the winter IOP. Approximately 11 total hours of in-cloud flight measurements were used to evaluate N_d retrievals.

Several assumptions used in the retrieval algorithms were assessed or characterized.

- *Cloud DSD Shape*: For the lidar-based N_d retrieval, we demonstrate in equation (6) that using the k parameter can eliminate the need to assume a shape of the cloud DSD (e.g., Gamma or lognormal distribution). The k parameter is the cube of the ratio of the volume radius to the effective radius (r_e) of the cloud droplets, representing the width of cloud DSD.
- *Constant N_d with height:* Aircraft *in situ* measurements confirm that N_d can be treated as constant through the cloud layer for stratiform MBL clouds, with the mean k parameter remaining constant with height. The k value ranges between 0.6-1.0 with a mean of 0.86, which is very close to k values at other geographic locations (Brenguier et al., 2011).
- *Treating Subadiabatic LWC:* The ratio of the retrieved LWP from the MWRRETv2 VAP divided by the LWP calculated from the adiabatic LWC profile is used to estimate the subadiabaticity fraction, f_{ad}. The mean value of f_{ad} is 0.76 at the ENA observatory during the ACE-ENA campaign period.

Retrieved N_d (N_{d_mpl}, N_{d_rl}, N_{d_radar}, N_{d_vap}) and cloud-layer-mean r_e (r_{em_mpl}, r_{em_rl}, r_{em_radar}, r_{em_vap}) are evaluated against aircraft *in situ* probe measurements of N_d (N_{d_FCDP}, N_{d_CAS}) and r_{em} (r_{em_FCDP}, r_{em_CAS}). To manage the challenge of direct one-to-one comparisons between ground-based retrievals and aircraft *in situ* measurements, we compare the PDFs of the retrievals with aircraft measurements. Analysis of the *in situ* measurements and retrievals for the 12 flight days reveal the following.

1) There is good agreement in the N_d *in situ* probe measurements, N_{d_FCDP} and N_{d_CAS}, with the relative differences in the median N_d often being smaller than 10% for most flights (albeit with larger differences in some cases).
2) Ground-based N_d retrievals generally follow the same day-to-day variation of the *in situ* measurements.
3) The assessment of the N_d retrievals with the *in situ* measurements reveals:
    a. N_{d_mpl} compares well overall with the aircraft measurements, but it overestimates N_d during the summer IOP and underestimates it during the winter IOP;
    b. N_{d_rl} compares well for overcast clouds but underestimates N_d for broken clouds;
    c. N_{d_radar} values are consistently smaller and have a narrower range than *in situ* measurements;
    d. N_{d_vap} overestimates N_d for broken clouds or clouds with low LWPs.
4) There is good agreement in the r_{em} *in situ* probe measurements, r_{em_FCDP} and r_{em_CAS}. The evaluations of r_{em} show that the retrievals following the variations in r_{em_FCDP}. There is a tendency for r_{em_mpl} to be slightly larger than r_{em_FCDP}.

These retrieval algorithms are further applied to four years of continuous ground-based remote sensing measurements of overcast MBL clouds at the ENA observatory. Monthly statistics of N_d (r_{em}) show slightly seasonal variations with a tendency towards higher (lower) values during the summer season and lower (higher) values during the winter season. N_{d_mpl} and N_{d_rl} are generally very close to each other. The N_{d_vap} is found to be systematically larger than other retrievals, which

might arise from the dissimilar fields of view (FOVs) for the cloud optical depth and LWP retrievals, where the former is a hemispheric FOV while the latter is a zenith radiance. N_{d_radar} compares well with lidar-based retrievals and has the narrowest distributions each month with the smallest monthly variations. r_{em_mpl} and r_{em_rl} are found to be slightly larger than r_{em_radar} and r_{em_vap}.

N_d retrievals evaluated in this study use various remote sensing measurements and employed different retrieval algorithms. Consequently, the ensemble of these retrievals for the same cloud can help us to quantify N_d retrieval uncertainties and identify reliable retrievals, such as when the ensemble of all retrievals has a narrow range (Zhao et al., 2012). Out of the four retrieval methods, we recommend using the MPL lidar-based method given its good agreement with *in situ* measurements, it has less sensitivity to issues arising from precipitation and low cloud LWP/optical depth, and it has broad applicability by functioning for both day and nighttime conditions. Ground-based N_d retrievals can be used to enhance our understanding of local cloud microphysical processes and can provide long-term verification of spaceborne N_d retrievals that can provide a global dataset needed for validating and improving global climate model simulations of clouds (Bennartz and Rausch 2017)

Data availability

The ARM ground-based measurements and ACE-ENA field campaign data used in this study can be downloaded from the ARM data archive site: https://www.archive.arm.gov/discovery/. Ground-based N_d and r_e retrievals are available upon request.

Author contributions

Conceptualization, D.Z. and A.M.V.; methodology, D.Z., Z.W., P.W., A.M.V.; software, D.Z.; validation, D.Z.; formal analysis, D.Z., Z.W., P.W., A.M.V.; investigation, D.Z.; resources, D.Z.; data curation, D.Z. and P.W.; writing—original draft preparation, D.Z.; writing—review and editing, all co-authors; visualization, D.Z.; supervision, D.Z.; project administration, D.Z.; funding acquisition, W.I.G. and A.M.V. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank the ACE-ENA field campaign team for their data collection during challenging operations. Data were obtained from the ARM user facility, a U.S. DOE Office of Science user facility managed by the Biological and Environmental Research (BER) program. This research was supported by the DOE ARM and the LASSO program.

References

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness. *Science*, **245**, 1227–1230, doi:https://doi.org/10.1126/science.245.4923.1227, 1989.

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.

Boers, R., Acarreta, J. R., and Gras, J. L.: Satellite monitoring of the first indirect aerosol effect: Retrieval of the droplet concentration of water clouds, *J. Geophys. Res.*, 111, D22208, doi:10.1029/2005JD006838, 2006.

Brenguier, J., Burnet, F., & Geoffroy, O.: Cloud optical thickness and liquid water path-does the k coefficient vary with droplet concentration? Atmospheric Chemistry and Physics, 11, 9771–9786. https://doi.org/10.5194/acp-11-9771-2011, 2011.

Cadeddu, M. P., Liljegren, J. C., and Turner, D. D.: The Atmospheric radiation measurement (ARM) program network of microwave radiometers: instrumentation, data, and retrievals, Atmos. Meas. Tech., 6, 2359–2372, https://doi.org/10.5194/amt-6-2359-2013, 2013.

Cadeddu, M. P., Ghate, V. P., and Mech, M.: Ground-based observations of cloud and drizzle liquid water path in stratocumulus clouds, Atmos. Meas. Tech., 13, 1485–1499, https://doi.org/10.5194/amt-13-1485-2020, 2020.

Cadeddu, M. P.: *Microwave Radiometer – 3-Channel (MWR3C) Instrument Handbook*. United States: N. p. Web. doi:10.2172/1039668, 2021.

Chen, J., Liu, Y., Zhang, M., and Peng, Y.: New understanding and quantification of the regime dependence of aerosol-cloud interaction for studying aerosol indirect effects, *Geophys. Res. Lett.*, 43, 1780– 1787, doi:10.1002/2016GL067683, 2016.

Desai, N., Liu, Y., Glienke, S., Shaw, R. A., Lu, C., Wang, J., & Gao, S.: Vertical variation of turbulent entrainment mixing processes in marine stratocumulus clouds using high-resolution digital holography. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD033527. https://doi.org/10.1029/2020JD033527, 2021.

Dong, X., Ackerman, T. P., and Clothiaux, E. E. 1998: Parameterizations of the microphysical and shortwave radiative properties of boundary layer stratus from ground-based measurements, *J. Geophys. Res.*, 103(D24), 31681– 31693, doi:10.1029/1998JD200047, 1998.

Donovan, D. P., Klein Baltink, H., Henzing, J. S., de Roode, S. R., & Siebesma, A. P.: A depolarisation lidar-based method for the determination of liquid-cloud microphysical properties. Atmospheric Measurement Techniques, 8, 237–266. https://doi.org/10.5194/amt-8-237-2015, 2015

Fan, J., Y. Wang, D. Rosenfeld, and X. Liu: Review of Aerosol–Cloud Interactions: Mechanisms, Significance, and Challenges. *J. Atmos. Sci.,* **73**, 4221–4252, https://doi.org/10.1175/JAS-D-16-0037.1, 2016.

Fernald, F. G.: Analysis of atmospheric lidar observations: some comments, Appl. Opt., 23, 652–653, 1984.

Grosvenor, D. P.,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, 2018.

Gryspeerdt, E., Quaas, J., Ferrachat, S., Gettelman, A., Ghan, S., Lohmann, U., Morrison, H., Neubauer, D., Partridge, D. G., Stier, P., Takemura, T., Wang, H., Wang, M., and Zhang, K.: Constraining the instantaneous aerosol influence on cloud albedo, P. Natl. Acad. Sci. USA, 114, 4899–4904, https://doi.org/10.1073/pnas.1617765114, 2017.

Hogan, R. J.: Fast lidar and radar multiple-scattering models. Part I: Small-angle scattering using the photon variance–covariance method. *Journal of the Atmospheric Sciences*, **65**, 3621– 3635. https://doi.org/10.1175/2008JAS2642.1, 2008.

Hodges, G. B., and Michalsky, J. J.: *Multifilter Rotating Shadowband Radiometer (MFRSR) Handbook With subsections for derivative instruments: Multifilter Radiometer (MFR) Normal Incidence Multifilter Radiometer (NIMFR)*. United States: N. p. Web. doi:10.2172/1251387, 2016.

Holdridge, D.: *Balloon-Borne Sounding System (SONDE) Instrument Handbook*. United States: N. p. Web. doi:10.2172/1020712, 2020.

Jensen, M. P., and Toto, T.: "Interpolated Sounding and Gridded Sounding Value-Added Products". United States. https://doi.org/10.2172/1326751, 2016.

Johnson, K. L, Giangrande, S. E, and Zhou, A.: *Ka-Band ARM Zenith Radar (KAZR) Active Remote Sensing of Clouds (ARSCL) CloudSat Calibration (KAZRARSCL-CLOUDSAT) (Value-Added Product Report)*. United States: N. p.. Web. doi:10.2172/1847644, 2022.

Klett, J. D.: Stable analytical inversion solution for processing lidar returns., Appl. Opt., 20, 211–220, https://doi.org/10.1364/AO.20.000211, 1981.

Kollias, P., Rémillard, J., Luke, E., and Szyrmer, W.: Cloud radar Doppler spectra in drizzling stratiform clouds: 1. Forward modeling and remote sensing applications, *J. Geophys. Res*., 116, D13201, doi:10.1029/2010JD015237, 2011.

Illingworth, A. J., Hogan, R. J., O'Connor, E.J., Bouniol, D., Brooks, M. E., Delanoé, J., Donovan, D. P., Eastment, J. D., Gaussiat, N., Goddard, J. W. F., Haeffelin, M., Baltink, H. K., Krasnov, O. A., Pelon, J., Piriou, J.-M., Protat, A., Russchenberg, H. W. J., Seifert, A., Tompkins, A. M., van Zadelhoff, G.-J., Vinit, F., Willén, U., Wilson, D. R., & Wrench, C. L. Cloudnet, *Bulletin of the American Meteorological Society*, *88*(6), 883-898. DOI: https://doi.org/10.1175/BAMS-88-6-883, 2007.

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2391 pp. doi:10.1017/9781009157896. 2021.

Lim, K.-S. S., L. Riihimaki, J. M. Comstock, B. Schmid, C. Sivaraman, Y. Shi, and G. M. McFarquhar: Evaluation of long-term surface-retrieved cloud droplet number concentration with in situ aircraft observations, *J. Geophys. Res. Atmos*., 121, 2318–2331, doi:10.1002/ 2015JD024082, 2016.

Luke, E. P., & Kollias, P.: Separating Cloud and Drizzle Radar Moments during Precipitation Onset Using Doppler Spectra, *Journal of Atmospheric and Oceanic Technology*, *30*(8), 1656-1671. doi: https://doi.org/10.1175/JTECH-D-11-00195.1, 2013.

Mace, G. G., and Sassen, K.: A constrained algorithm for retrieval of stratocumulus cloud properties using solar radiation, microwave radiometer, and millimeter cloud radar data, *J. Geophys. Res.*, 105(D23), 29099– 29108, doi:10.1029/2000JD900403., 2000.

Marais, W. J., Holz, R. E., Hu, Y. H., Kuehn, R. E., Eloranta, E. E., and Willett, R. M.: Approach to simultaneously denoise and invert backscatter and extinction from photon-limited atmospheric lidar observations, *Appl. Opt.* **55**, 8316-8334, https://doi.org/10.1364/AO.55.008316, 2016.

Martin, G. M., Johnson, D. W., and Spice, A.: The Measurement and Parameterization of Effective Radius of Droplets in Warm Stratocumulus Clouds. *J. Atmos. Sci.*, 51, 1823–1842, https://doi.org/10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2., 1994.

Martucci, G. and O'Dowd, C. D.: Ground-based retrieval of continental and marine warm cloud microphysics, Atmos. Meas. Tech., 4, 2749-2765, https://doi.org/10.5194/amt-4-2749-2011, 2011.

McComiskey, A., Feingold, G., Frisch, A. S., Turner, D. D., Miller, M. A., Chiu, J. C., Min, Q., and Ogren, J. A.: An assessment of aerosol-cloud interactions in marine stratus clouds based on surface remote sensing, *J. Geophys. Res.*, 114, D09203, doi:10.1029/2008JD011006, 2009.

McCoy, I. L., McCoy, D. T., Wood, R., Regayre, L., Watson-Parris, D., Grosvenor, D. P., et al.: The hemispheric contrast in cloud microphysical properties constrains aerosol forcing. *Proceedings of the National Academy of Sciences of the United States of America*, **117**(32), 18998– 19006. https://doi.org/10.1073/pnas.1922502117, 2020.

Merk, D., Deneke, H., Pospichal, B., and Seifert, P.: Investigation of the adiabatic assumption for estimating cloud micro- and macrophysical properties from satellite and ground observations, Atmos. Chem. Phys., 16, 933–952, https://doi.org/10.5194/acp-16-933-2016, 2016.

Miles, N. L., Verlinde, J., and Clothiaux, E. E.: Cloud Droplet Size Distributions in Low-Level Stratiform Clouds. *J. Atmos. Sci.*, **57**, 295–311, https://doi.org/10.1175/1520-0469(2000)057<0295:CDSDIL>2.0.CO;2., 2000.

Min, Q, and LC Harrison.: "Cloud properties derived from surface MFRSR measurements and comparison with GOES results at the ARM SGP site." Geophysical Research Letters 23(13):1641−1644, https://doi.org/10.1029/96GL01488, 1996.

Moore, R. H., Karydis, V. A., Capps, S. L., Lathem, T. L., and Nenes, A.: Droplet number uncertainties associated with CCN: an assessment using observations and a global model adjoint, Atmos. Chem. Phys., 13, 4235–4251, https://doi.org/10.5194/acp-13-4235-2013, 2013.

Muradyan, P., and Coulter, R.: *Micropulse Lidar (MPL) Handbook*. United States: N. p., Web. doi:10.2172/1020714, 2020.

Newsom, R. K., Bambha, R., and Chand, D.: *Raman Lidar (RL) Instrument Handbook*. United States: N. p. Web. doi:10.2172/1020561., 2022.

O'Connor, E. J., Illingworth, A. J., & Hogan, R. J.: A Technique for Autocalibration of Cloud Lidar. *Journal of Atmospheric and Oceanic Technology*, **21**(5), 777– 786. https://doi.org/10.1175/1520-0426(2004)021<0777:ATFAOC>2.0.CO;2, 2004.

Painemal, D. and Zuidema, P.: Microphysical variability in southeast Pacific Stratocumulus clouds: synoptic conditions and radiative response, Atmos. Chem. Phys., 10, 6255–6269, https://doi.org/10.5194/acp-10-6255-2010, 2010.

Pawlowska, H., Grabowski, W. W., & Brenguier, J.-L.: Observations of the width of cloud droplet spectra in stratocumulus. Geophysical Research Letters, 33, L19810. https://doi.org/10.1029/2006GL02684, 2006.

Regayre, L. A., Pringle, K. J., B. Booth, B. B., Lee, L. A., Mann, G. W., Browse, J., Woodhouse, M. T., Rap, A., Reddington, C. L., and Carslaw, K. S.: Uncertainty in the magnitude of aerosol-cloud radiative forcing over recent decades, *Geophys. Res. Lett.*, 41, 9040– 9049, doi:10.1002/2014GL062029, 2014.

Riihimaki, L., McFarlane, S., and Sivaraman, C. "Droplet Number Concentration Value-Added Product". United States. https://doi.org/10.2172/1237963, 2021.

Rosenfeld, D., et al.: Satellite retrieval of cloud condensation nuclei concentrations by using clouds as CCN chambers. Proceedings of the National Academy of Sciences, 113(21), 5828–5834. https://doi.org/10.1073/pnas.1514044113, 2016.

Rosenfeld, D., Y. Zhu, M. Wang, Y. Zheng, T. Goren, S. Yu: Aerosol-driven droplet concentrations dominate coverage and water of oceanic low-level clouds. *Science,* **363**, eaav0566 (2019).

Sarna, K., Donovan, D. P., and Russchenberg, H. W. J.: Estimating the optical extinction of liquid water clouds in the cloud base region, Atmos. Meas. Tech., 14, 4959–4970, https://doi.org/10.5194/amt-14-4959-2021, 2021.

Schmidt, J., Wandinger, U., & Malinka, A.: Dual-field-of-view Raman lidar measurements for the retrieval of cloud microphysical properties. Applied Optics, 52, 2235. https://doi.org/10.1364/AO.52.002235, 2013.

Snider, J. R., Leon, D., & Wang, Z.: Droplet Concentration and Spectral Broadening in Southeast Pacific Stratocumulus Clouds, *Journal of the Atmospheric Sciences*, *74*(3), 719-749. https://doi.org/10.1175/JAS-D-16-0043.1, 2017.

Stephens, G. L., et al.: An update on Earth's energy balance in light of the latest global observations, Nat. Geosci., 5, 691–696, 2012.

Storelvmo, T., Kristjánsson, J. E., Ghan, S. J., Kirkevåg, A., Seland, Ø., and Iversen, T.: Predicting cloud droplet number concentration in Community Atmosphere Model (CAM)-Oslo, *J. Geophys. Res.*, 111, D24208, doi:10.1029/2005JD006300, 2006.

Thorsen, T. J., Fu, Q., Newsom, R. K., Turner, D. D., & Comstock, J. M.: Automated Retrieval of Cloud and Aerosol Properties from the ARM Raman Lidar. Part I: Feature Detection, *Journal of Atmospheric and Oceanic Technology*, *32*(11), 1977-1998, DOI: https://doi.org/10.1175/JTECH-D-14-00150.1, 2015.

Thorsen, T. J., & Fu, Q.: Automated Retrieval of Cloud and Aerosol Properties from the ARM Raman Lidar. Part II: Extinction, *Journal of Atmospheric and Oceanic Technology*, *32*(11), 1999-2023, DOI: https://doi.org/10.1175/JTECH-D-14-00178.1, 2015.

Turner, D. D., S. A. Clough, J. C. Liljegren, E. E. Clothiaux, K. Cady-Pereira, and K. L. Gaustad.: "Retrieving Liquid Water Path and Precipitable Water Vapor From the Atmospheric Radiation Measurement (ARM) Microwave Radiometers." *IEEE Transactions on Geoscience and Remote Sensing*, 45(11), DOI: 10.1109/TGRS.2007.903703, 2007.

Turner, D. D., McFarlane, S. A., Riihimaki, L., Shi, Y., Lo, C., and Min, Q.: "Cloud Optical Properties from the Multifilter Shadowband Radiometer (MFRSRCLDOD). An ARM Value-Added Product". United States. https://doi.org/10.2172/1237958, 2014.

Twomey, S.: Influence of pollution on shortwave albedo of clouds. *J. Atmos. Sci.*, **34**, 1149–1152, doi: https://doi.org/10.1175/15200469(1977)034<1149:TIOPOT>2. 0.CO;2, 1977.

Vogelmann, A. M., et al.: RACORO Extended-Term Aircraft Observations of Boundary Layer Clouds, *Bulletin of the American Meteorological Society*, *93*(6), 861-878. DOI: https://doi.org/10.1175/BAMS-D-11-00189.1, 2012.

Wang, Z.: A refined two-channel microwave radiometer liquid water path retrieval for cold regions by using multiple-sensor measurements*, IEEE Geoscience & remote sensing letters,* **4**, 591-595*,* 2007.

Wang, J., Daum, P. H., Yum, S. S., Liu, Y., Senum, G. I., Lu, M.-L., et al.: Observations of marine stratocumulus microphysics and implications for processes controlling droplet spectra: Results from the marine stratus/stratocumulus experiment. *Journal of Geophysical Research: Atmospheres*, *114*(D18). https://doi.org/10.1029/2008jd011035, 2009.

Wang, J., et al.: Aerosol and Cloud Experiments in the Eastern North Atlantic (ACE-ENA), *Bulletin of the American Meteorological Society*, *103*(2), E619-E641, doi: https://doi.org/10.1175/BAMS-D-19-0220.1, 2022.

Welton, E. J., James R. Campbell, James D. Spinhirne, and V. Stanley Scott III: "Global monitoring of clouds and aerosols using a network of micropulse lidar systems", Proc. SPIE 4153, Lidar Remote Sensing for Industry and Environment Monitoring, (13 February 2001), https://doi.org/10.1117/12.417040, 2001.

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

Wood, R., et al.: Clouds, aerosol, and precipitation in the marine boundary layer: An ARM Mobile Facility deployment. *Bull. Amer. Meteor. Soc.*, **96**, 419–439, https://doi.org/10.1175/BAMS-D-13-00180.1, 2015.

Wu, P., Dong, X., Xi, B., Tian, J., & Ward, D. M.: Profiles of MBL cloud and drizzle microphysical properties retrieved from ground-based observations and validated by aircraft in situ measurements over the Azores. Journal of Geophysical Research: Atmospheres, 125, e2019JD032205. https://doi.org/10.1029/2019JD032205, 2020 a.

Wu, P., X. Dong, and B. Xi: A Climatology of Marine Boundary Layer Cloud and Drizzle Properties Derived from Ground-Based Observations over the Azores. *J. Climate*, **33**, 10133–10148, https://doi.org/10.1175/JCLI-D-20-0272.1, 2020 b.

Yeom, J. M., Yum, S. S., Shaw, R. A., La, I., Wang, J., Lu, C., et al.: Vertical variations of cloud microphysical relationships in marine stratocumulus clouds observed during the ACE-ENA campaign. *Journal of Geophysical Research: Atmospheres*, 126, e2021JD034700. https://doi.org/10.1029/2021JD034700, 2021.

Zelinka, M. D., Randall, D. A., Webb, M. J., and Klein, S. A.: Clearing Clouds of Uncertainty, Nat. Clim. Change, 7, 674–678, https://doi.org/10.1038/nclimate3402, 2017.

Zhang, D., Vogelmann, A., Kollias, P., Luke, E., Yang, F., Lubin, D., & Wang, Z.: Comparison of Antarctic and Arctic single-layer stratiform mixed-phase cloud properties using ground-based remote sensing measurements. *Journal of Geophysical Research: Atmospheres*, 124, 10186– 10204. https://doi.org/10.1029/2019JD030673, 2019.

Zhang, Z., Song, Q., Mechem, D. B., Larson, V. E., Wang, J., Liu, Y., Witte, M. K., Dong, X., and Wu, P.: Vertical dependence of horizontal variation of cloud microphysics: observations from the ACE-ENA field campaign and implications for warm-rain simulation in climate models, Atmos. Chem. Phys., 21, 3103–3121, https://doi.org/10.5194/acp-21-3103-2021, 2021.

Zhu, Z., Kollias, P., Luke, E., and Yang, F.: New insights on the prevalence of drizzle in marine stratocumulus clouds based on a machine learning algorithm applied to radar Doppler spectra, Atmos. Chem. Phys., 22, 7405–7416, https://doi.org/10.5194/acp-22-7405-2022, 2022.

Table 1: Ground-based instruments and measurements at the ENA site used in this study.

Instrument	Temporal/vertical resolutions	Measured or derived quantities
Micropulse Lidar (MPL)	10 s/15 m	Lidar backscatter intensity, linear depolarization ratio
Raman Lidar (RL)	10 s/7.5 m	Particulate lidar backscatter and extinction coefficient, linear depolarization ratio
Ka-band ARM Zenith Radar (KAZR)	2 s/30 m	Radar reflectivity, Doppler velocity, spectral width
Microwave Radiometer 3-Channel (MWR3C)	30 s/column	Brightness temperatures, LWP
Multifilter Rotating Shadowband Radiometer (MFRSR)	20 s/column	Narrowband irradiance at 415, 500, 615, 673, 870, and 940 nm, aerosol optical depth, cloud optical depth
Balloon-Borne Sounding System (SONDE)	2 times per day	Atmospheric pressure, temperature, and moisture profiles

Table 2: The selected 12 flight days and their descriptions. * Indicates broken cloud conditions.

Date	Time (UTC)	In-cloud time	Cloud Conditions	Mean Distance between the ENA observatory and G1
2017/06/21	11:34-15:17	14 min	Stratocumulus cloud layer	23.3 km
***2017/06/28**	9:02-12:34	10 min	Low-level stratus (broken conditions)	13.4 km
2017/06/30	9:27-13:16	1 hour 8 min	Persistent stratus cloud layer with top near 1 km	13.1 km
2017/07/06	8:22-11:58	1 hour 1 min	Stratocumulus cloud with embedded drizzle patches	9.8 km
***2017/07/08**	8:34-12:44	37 min	Low-level stratus with cloud top near 1 km (broken conditions)	147. 7 km
2017/07/18	8:31-12:04	1 hour 32 min	Drizzling stratocumulus clouds	14.8 km
2018/01/19	12:10-16:06	48 min	Drizzling stratocumulus clouds	3.6 km
2018/01/25	11:02-14:49	1 hour 25 min	Overcast stratocumulus clouds	11.9 km
2018/01/26	11:05-15:00	1 hour 37 min	Overcast stratocumulus clouds	134.8 km
2018/01/30	9:34-13:50	1 hour 34 min	Solid stratocumulus cloud deck	18.5 km
2018/02/07	17:28-19:22	44 min	Overcast stratocumulus clouds	17.9 km
***2018/02/12**	11:05-15:07	26 min	Low-level stratus (broken conditions)	2.8 km

730 **Table 3: Median N_d values for the selected 12 flight days. The percentages in parentheses represent the relative difference of the N_d retrievals compared to N_{d_FCDP}. * Indicates broken cloud conditions.**

Date	N_{d_FCDP} cm^{-3}	N_{d_CAS} cm^{-3}	N_{d_mpl} cm^{-3}	N_{d_rl} cm^{-3}	N_{d_radar} cm^{-3}	N_{d_vap} cm^{-3}
2017/06/21	66	75 (13%)	101 (53%)	n/a	81 (22%)	129 (94%)
***2017/06/28**	58	63 (8%)	94 (62%)	24 (-59%)	61 (5%)	44 (-24%)
2017/06/30	115	125 (9%)	217 (89%)	136 (19%)	77 (-33%)	314 (174%)
2017/07/06	95	96 (2%)	127 (34%)	106 (12%)	64 (-33%)	87 (-8%)
***2017/07/08**	76	83 (9%)	102 (34%)	16 (-79%)	79 (4%)	123 (61%)
2017/07/18	67	61 (-9%)	61 (-9%)	34 (-49%)	55 (-18%)	62 (-8%)
2018/01/19	33	50 (52%)	38 (15%)	37 (13%)	54 (64%)	54 (65%)
2018/01/25	57	62 (8%)	33 (-43%)	52 (-10%)	52 (-9%)	46 (-20%)
2018/01/26	94	72 (-23%)	69 (-27%)	98 (4%)	60 (-37%)	120 (27%)
2018/01/30	80	74 (-7%)	54 (-32%)	60 (-25%)	61 (-23%)	57 (-29%)
2018/02/07	125	72 (-42%)	76 (-39%)	133 (7%)	63(-50%)	90 (-28%)
***2018/02/12**	105	71 (-32%)	77 (-26%)	47 (-55%)	65 (-38%)	96 (-9%)

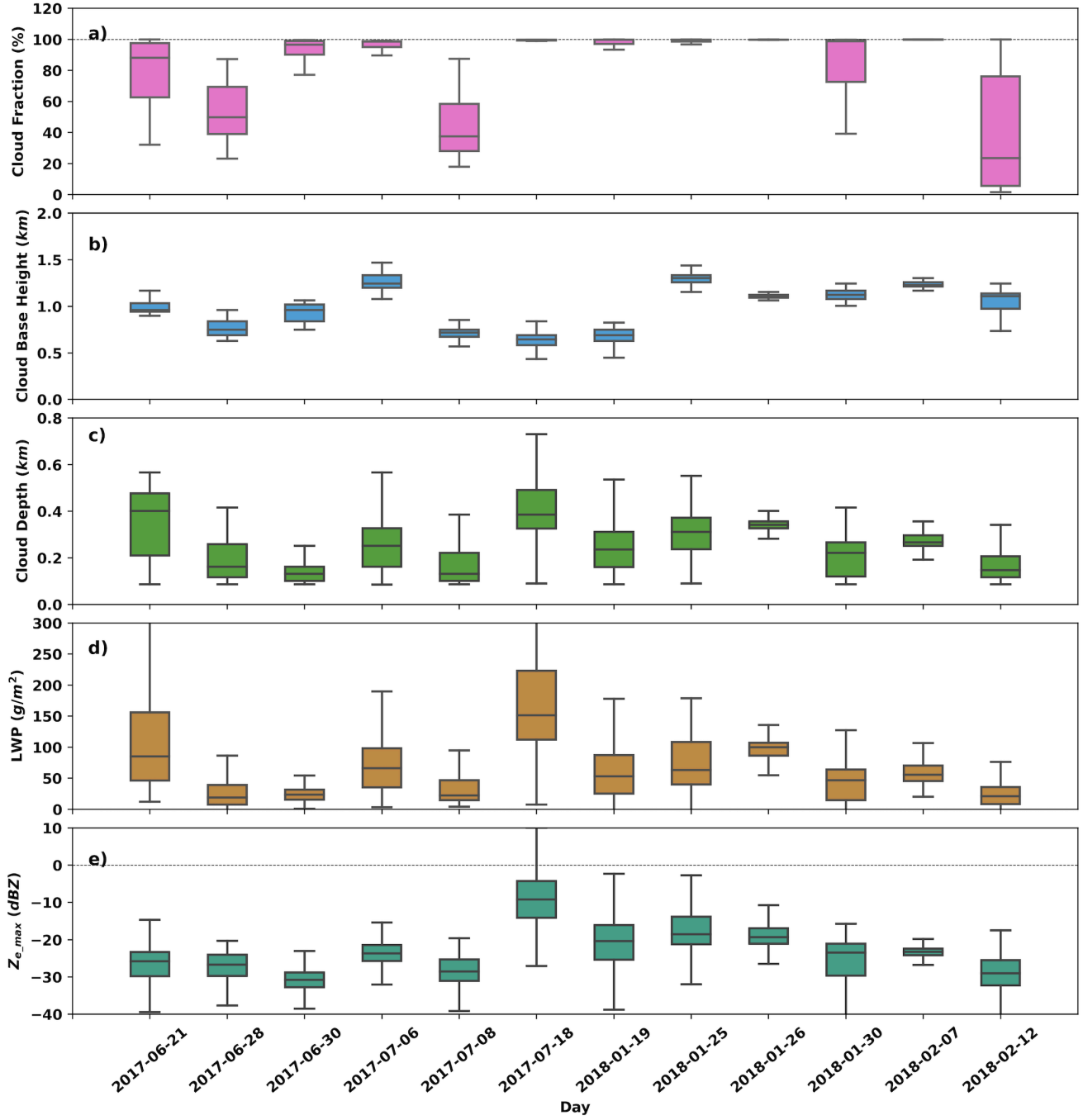
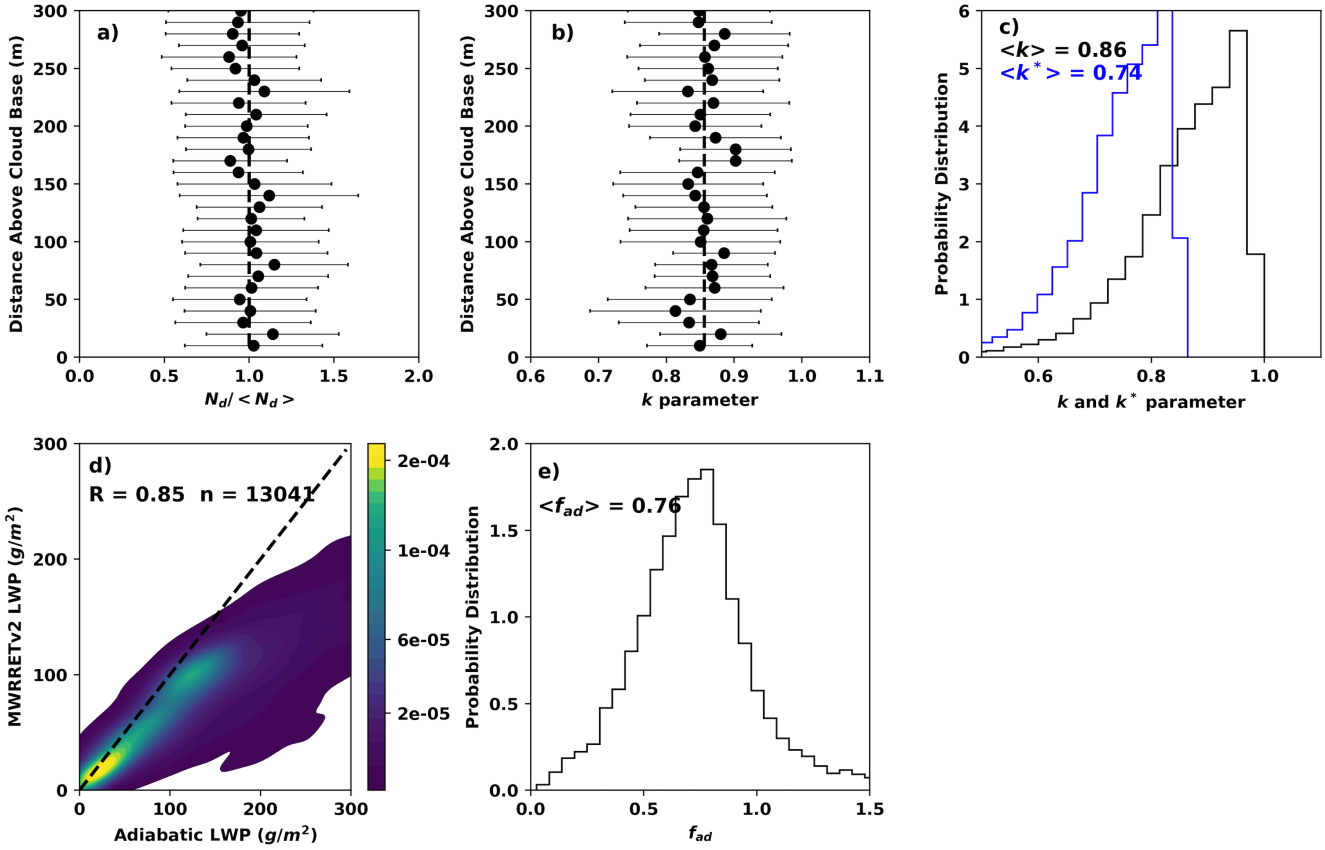


Figure 1: Cloud properties of the 12 selected flight days derived from the ENA ground-based remote sensing observations during aircraft measurement periods: a) Fractional sky cover obtained from total sky imager (TSI) observations; b) Cloud-base height determined from MPL measurements; c) Cloud depth; d) LWP obtained from the MWRRETv2 VAP; and e) column maximum Z_e (Z_{e_max}) from KAZR measurements. The box-and-whisker plots display the 5th, 25th, 50th, 75th, and 95th percentiles. Dashed lines in a) and e) represent 100% cloud fraction and Z_{e_max} of 0 dBZ, respectively.



740 **Figure 2: Statistics of cloud properties used in the retrieval algorithms from *in situ* and ground-based measurements during the 12 selected flight days: a) The mean and standard deviation of the FCDP-measured N_d normalized by the flight average; b) The mean and standard deviation of the derived *k* parameter profile within clouds; c) The PDFs of the *k* and k^* parameters; d) The regression between LWPs from MWRRETv2 retrievals and those calculated assuming an adiabatic cloud, where *R* is the Pearson correlation coefficient and *n* the total number of profiles; and e) PDF of f_{ad}.**

745

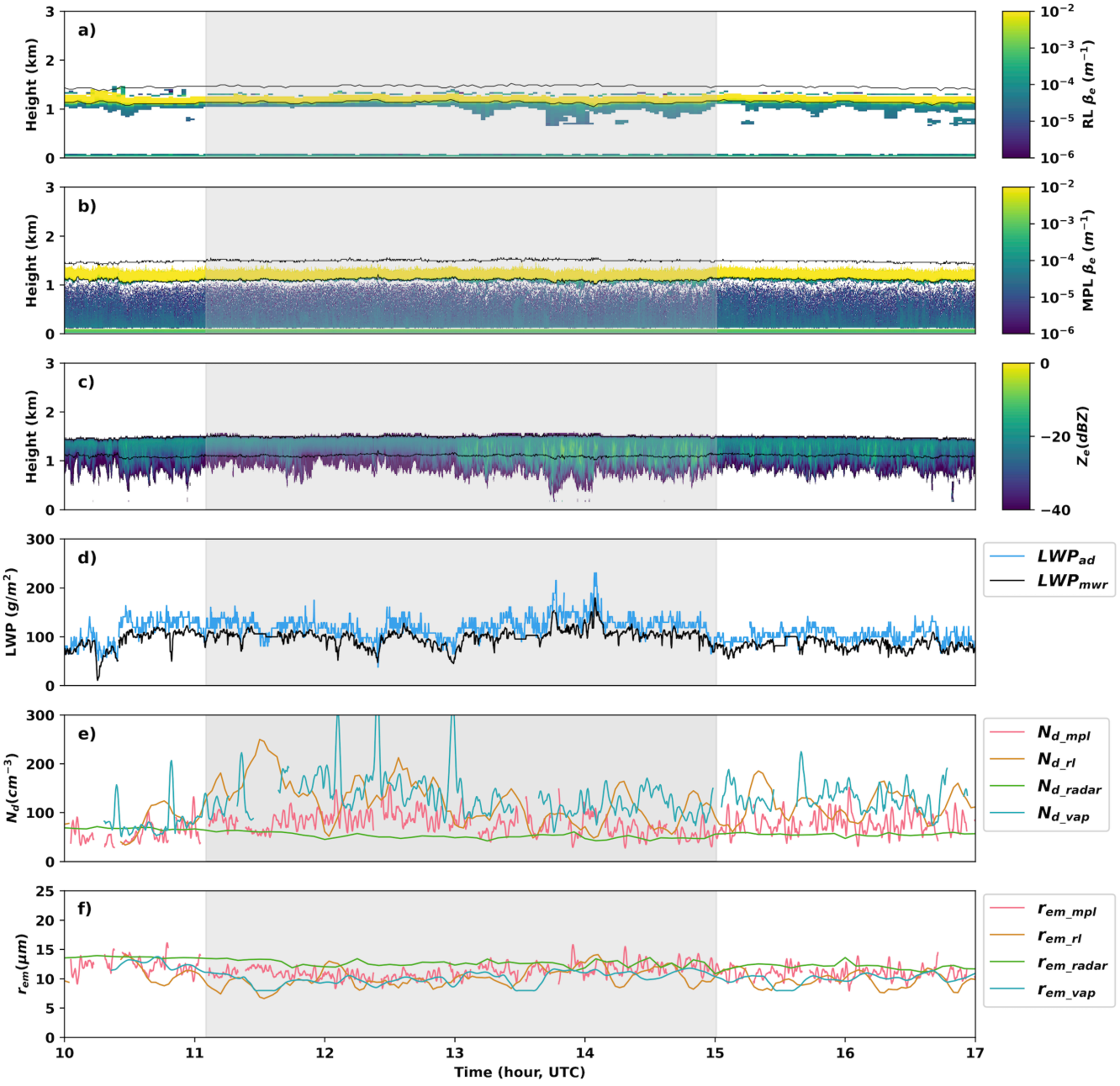


Figure 3: An example of ground-based remote sensing measurements and N_d retrievals on January 26, 2018: a) RL extinction coefficient (β_e) profiles from the RLPROF-FEX VAP; b) MPL β_e profiles; c) KAZR radar reflectivity profiles; d) LWPs from MWRRETv2 retrievals (LWP_{mwr}) and calculated assuming an adiabatic cloud liquid water content vertical profile (LWP_{ad}); e) retrieved N_{d_mpl}, N_{d_rl}, N_{d_radar}, N_{d_vap}; f) derived layer mean r_e (r_{em}) per retrieval,

r_{em_mpl}, r_{em_rl}, r_{em_radar}, r_{em_vap}. Black lines in a), b), c) are cloud top and base detected with combined lidar and radar measurements. The grey zone indicates the time of concurrent aircraft *in situ* measurements.

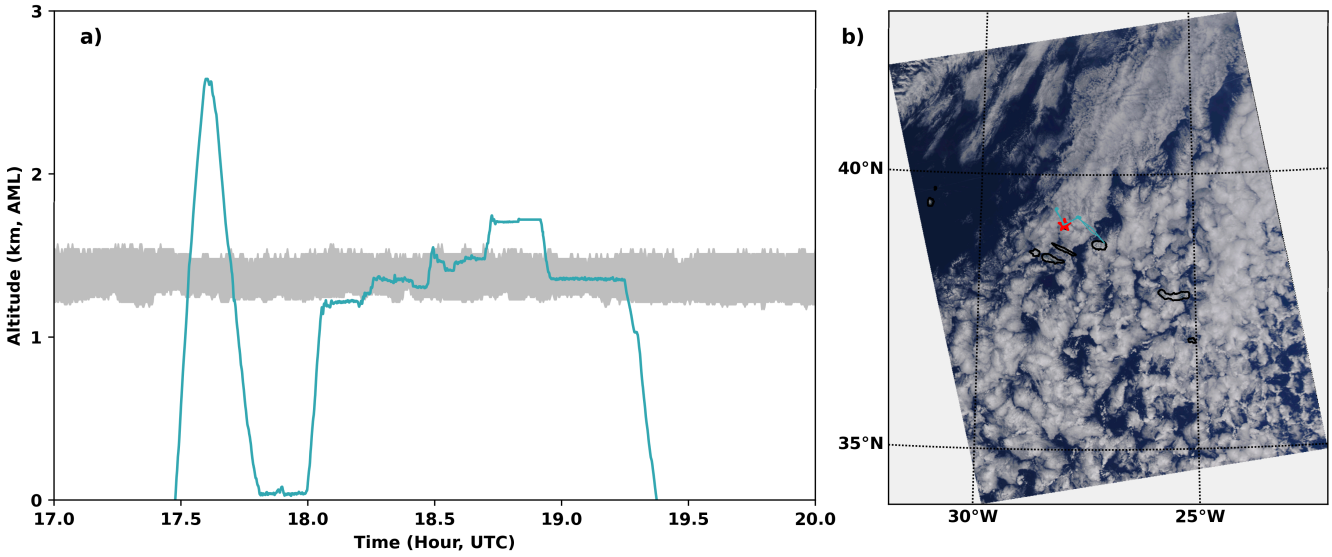
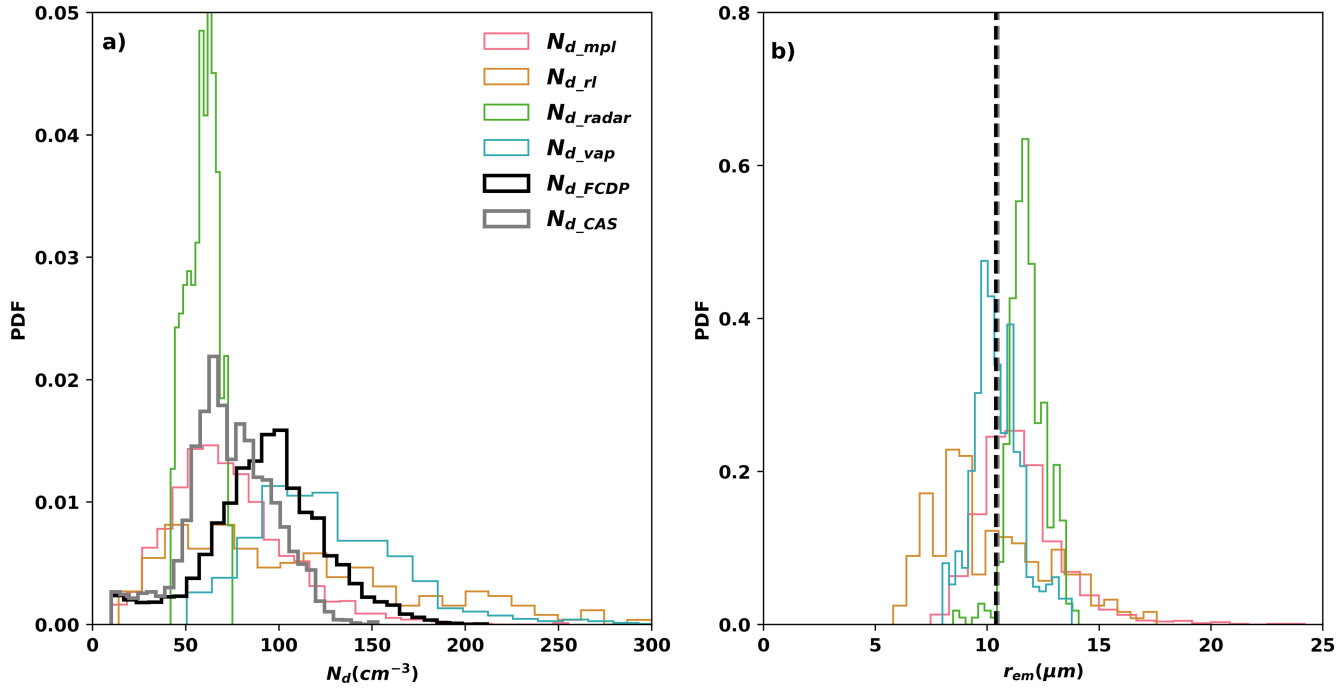


Figure 4: a) The G-1 aircraft flight track on January 26, 2018. The grey zone represents the cloud layer. Wind barbs are from the ARM radiosonde measurements at 11:30 UTC at the ENA observatory; b) Moderate Resolution Imaging Spectroradiometer (MODIS) true color image of clouds between 13:00-13:10 UTC on January 26, 2018. The red star indicates the location of the ENA observatory. The black circular regions represent islands. The blue lines in a) and b) represent the aircraft flight track.



765

Figure 5: Evaluation of ground-based retrievals of N_d and r_e with aircraft *in situ* measurements for the case on January 26, 2018. a) PDFs of N_{d_mpl}, N_{d_rl}, N_{d_radar}, N_{d_vap}, N_{d_FCDP}, and N_{d_CAS} during the time of concurrent aircraft measurements; b) PDFs of r_{em_mpl}, r_{em_rl}, r_{em_radar}, r_{em_vap}, and derived r_{em_FCDP} and r_{em_CAS} from *in situ* probe measurements. The colors of r_{em} lines in b) correspond to those given in a). Dashed lines in b) are mean r_e from FCDP

770 **and CAS measurements during all in-cloud penetrations.**

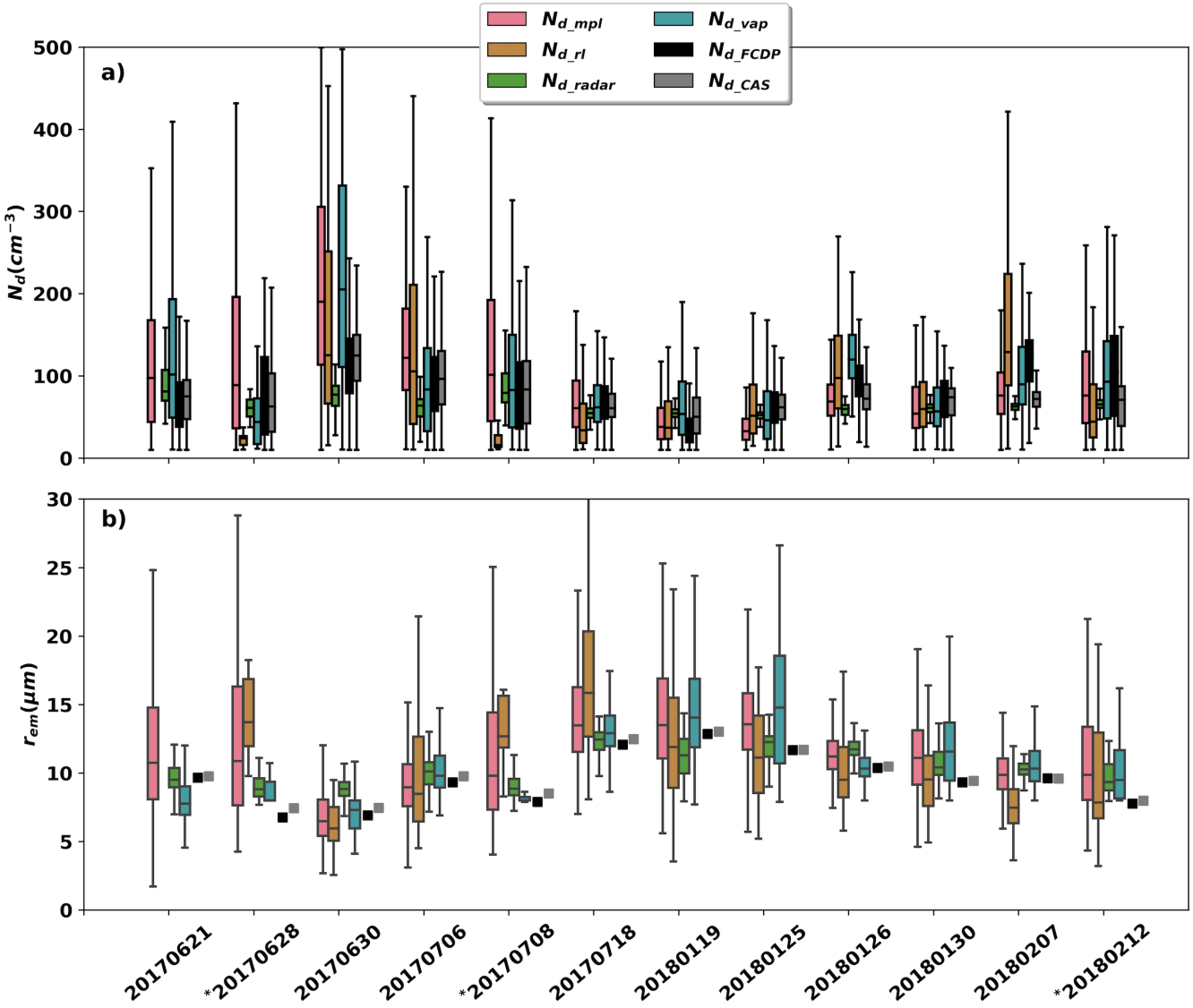
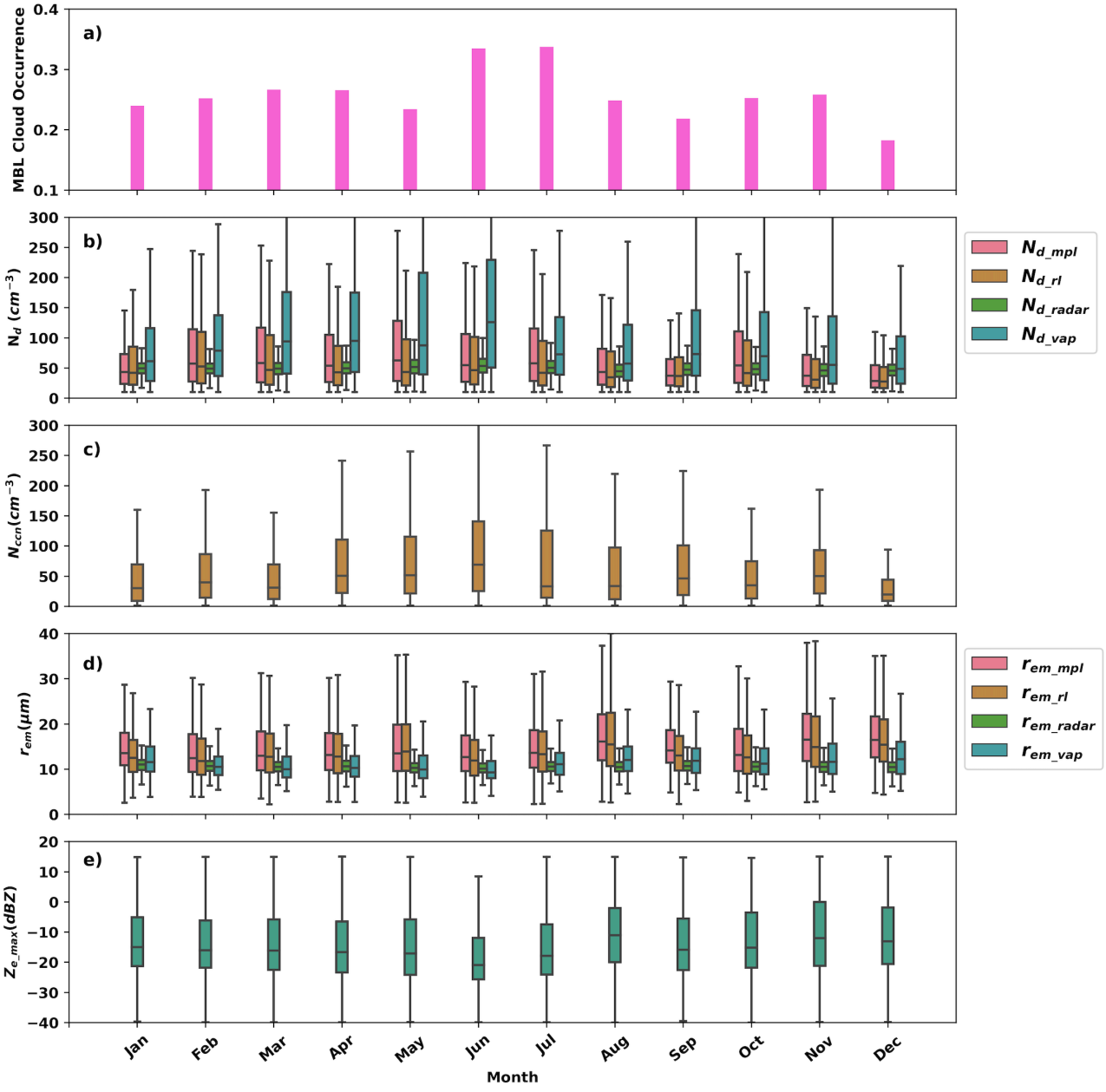


Figure 6: Evaluation of ground-based retrievals of N_d and r_e with aircraft *in situ* measurements for the 12 selected flight days. a) Boxplots of N_{d_mpl}, N_{d_rl}, N_{d_radar}, N_{d_vap}, N_{d_FCDP}, and N_{d_CAS} during the time of concurrent aircraft measurements; b) Boxplots of r_{em_mpl}, r_{em_rl}, r_{em_radar}, r_{e_vap}, and derived r_{em_FCDP} and r_{em_CAS} from *in situ* probe measurements. The colors of r_{em} boxplots in b) correspond to those given in a). The * indicates broken cloud conditions.



780

Figure 7: Monthly variations of overcast marine boundary layer cloud N_d and r_e using four years of ground-based measurements at the ENA observatory. a) Occurrence of MBL clouds; b) retrieved N_{d_mpl}, N_{d_rl}, N_{d_radar}, and N_{d_vap}; c) cloud condensation nuclei (N_{ccn}) from the ARM CCN counter (CCN-100) of aerosol observing system (AOS) at a supersaturation of 0.1%; d) derived r_{em_mpl}, r_{em_rl}, r_{em_radar}, and r_{em_vap}; and e) Z_{e_max}.