Synergistic retrievals of ice in high clouds from lidar, Ku-band radar and submillimeter wave radiometer observations

Mircea Grecu[a,b] and John Yorks [a]

[a] *NASA GSFC*

[b] *Morgan State University*

Corresponding author: Mircea Grecu, mircea.grecu-1@nasa.gov

1

ABSTRACT: In this study, we investigate the synergy of lidar, Ku-band radar, and sub-millimeter-wave radiometer measurements in the retrieval of the ice from satellite observations. The synergy is analyzed through the generation of a large dataset of IWC profile and simulated lidar, radar and radiometer observations. The characteristics of the instruments e. g. frequencies, sensitivities, etc. are set based on the expected characteristics of instruments of the AOS mission. A cross-validation methodology is used to assess the accuracy of the retrieved IWC profiles from various combinations of observations from the three instruments. Specifically, the IWC and associated observations is randomly divided into two datasets, one for the training and the other for the evaluation. The training dataset is used to train the retrieval algorithm, while the evaluation dataset is used to assess the retrieval performance. The dataset of IWC profiles is derived from CloudSat reflectivity observations.

The retrieval of the ice water content IWC profiles from the computed observations is achieved in two steps. In the first step, a class, out of 18 potential classes characterized by different vertical distribution of IWC, is estimated from the observations. The 18 classes are predetermined based on k-Means clustering algorithm. In the second step, the IWC profile is estimated using a Ensemble Kalman Smoother (EKS) algorithm that uses the estimated class as a priori information.

The results of the study show that the synergy of lidar, radar, and radiometer observations is significant in the retrieval of the IWC profiles. The inclusion of the lidar observations in the retrieval process has a larger impact on the retrieval performance than the inclusion of the radar observations.

1. Introduction

The future NASA Atmospheric Observing System (AOS) mission (Braun 2022) is expected to feature new combinations of observations that may be used to quantify the amounts of ice in high clouds and characterize the microphysical properties of ice particles. These observations include lidar backscatter, Ku-band radar reflectivity, and submillimeter wave radiometer brightness temperature measurements. While not optimal for cloud ice estimation, but for the characterization of a broader spectrum of cloud and precipitation processes, these observations are nevertheless synergistic from the characterization of ice clouds perspective. That is, despite the fact that lidar observations attenuate quickly in thick ice clouds and the Ku-band radar will not be able to detect clouds characterized by an echo weaker than 8.0 dBZ, the active observations are expected to provide context that may be incorporated into the radiometer retrievals. Herein, term retrieval is defined as the process of estimating geophysical variables from remote sensing observations. In this study, we investigate the impact of incorporating lidar and radar observations into the radiometer retrieval of ice clouds. Because the existent amount of coincident backscatter lidar, Ku-band radar, and submillimeter-wave radiometer observations is rather insufficient to derive conclusive results, we employ accurate physical models to simulate lidar, radar and radiometer observations and use a cross-validation methodology to characterize the retrieval accuracy. As estimates from passive instrument observations strongly depend on "a priori" information (Rodgers 2000), for results to be relevant in real applications, it is necessary to base them on realistic vertical distributions of ice properties. Such distributions may be derived from cloud-resolving-model (CRM) simulations (Pfreundschuh et al. 2020; Liu and Mace 2022) or directly from observations. In this study, we employ the latter approach, as CRMs may still be deficient in properly reproducing the vertical distribution of ice clouds and their associated microphysical properties. Specifically, we use observations and products from the CloudSat (CS) mission (Stephens et al. 2002) to derive a database of ice microphysical properties and associated simulated lidar, radar, and radiometer observations. This database is used to investigate the accuracy of estimated ice cloud properties from the simulated observations. Another major difference relative to previous studies is the unique combination of instruments investigated herein. The article is organized as follows. In Section 2, we describe the approach used to derive the ice properties and the associated simulated

observations, the retrieval and the evaluation methodology. In Section 3, we present the results of the evaluation methodology. We conclude in Section 4.

2. Methodology

As previously mentioned, we use CloudSat (CS) observations (Stephens et al. 2002) to derive the vertical distributions of ice properties needed in the investigation. Although research quality CS cloud ice products exist, to maximize the physical consistency of the approach, we do not use them but derive ice amounts and associated properties directly from CS reflectivity observations. This ensures the consistency between the particle distribution assumptions and the electromagnetic scattering properties used in the CS reflectivity processing and those used the simulation of the lidar, Ku-band radar and radiometer observations. Lidar, Ku-band radar, and submillimeter-wave radiometer observations are simulated from CS observations using accurate physical models and realistic assumptions consistent with the most recent knowledge in the field of ice cloud microphysics, and a non-parametric estimation methodology based on the k-Means clustering algorithm MacKay (2003) is used to investigate the instrument synergy. Details of the methodology are presented below.

a. Assumptions and forward models

To quantify the number of ice particles in an elementary atmospheric volume as a function of their size, we use normalized gamma functions (Bringi et al. 2003). The benefit of normalized gamma functions is that they encapsulate the variability of Ice Water Content (IWC) - reflectivity relationship into a single parameter, i.e. the normalized Particle Size Distribution (PSD) intercept (Testud et al. 2001; Bringi et al. 2003). The normalized PSD intercept is defined as $N_w = \frac{4^4}{\pi \rho_w} \frac{IWC}{D_m^4}$, where IWC is the ice water content associated with the PSD, and D_m is the mass weighted mean diameter. Testud et al. (2001) showed that the variability in IWC reflectivity (Z) relationships may be fully explained by variability in N_w, and that a formula of the type

$$IWC = N_w^{1-b} a Z^b \tag{1}$$

perfectly explains the relationships between IWC and Z calculated from observed PSDs. Equation (1) is not sufficient to derive accurate, unbiased estimates of ice water contents, because N_w

varies considerably in time and space. Nevertheless, multiple studies showed that it is beneficial to parameterize N_w as a function of various variables, such as temperature (Hogan et al. 2006; Delanoë and Hogan 2008; Deng et al. 2010), rather than using N_w independent relations. In this study, we parameterize N_w as a function of temperature based on the CloudSat 2C-ICE product (Deng et al. 2010, 2013). Specifically, we cluster, based on similarity, a large set 2C-ICE profiles into 18 classes using a k-Means procedure. The mean IWC profiles associated with the 18 classes are shown in continuous lines in Figure 1. Alternative estimates, derived using PSD assumptions and electromagnetic scattering calculations that enable accurate and physically consistent simulations of radar observations at Ku-band and radiometer observations of submillimeter-wave frequencies are also shown in Figure 1. These estimates are based on the self-similarity Rayleigh-Gans approximation (SSRGA) of Hogan et al. (2017). Details regarding the estimation process are provided in the following paragraphs. As apparent in Figure 1, the CS and SSRGA estimates are in good agreement. Some discrepancies due to differences between the SSRGA N_w parameterization and the CS 2C-ICE "a priori assumptions" are also apparent, but they are not deemed critical in this study, whose objective is the investigation of synergistic lidar, Ku-band radar and submillimeter-wave radiometer retrievals, because the outcome is not likely to be sensitive to such details. One may notice that the average IWC profiles in Figure 1 are characterized by different peak values and heights. This facilitates a simple way to reverse-engineer to (some extent) the "a priori" assumptions used in the CS 2C-ICE product and use them in formulation of the type described in Equation (1). Specifically, the derivation of relationships of the type $IWC = a_i Z^{b_i}$ for every class i may be used to study a_i as a function of height. Shown in Figure 2 is a representation of the class multiplicative coefficient a_i as a function of relative height scatter plot. As apparent in the figure, and as expected, a_i exhibits a strong variation with the relative height. Coefficient b_i exhibits a height dependency as well (not shown), but the range of variation is significantly smaller, almost zero relative to the mean value of b. Given that any deviation of the multiplicative coefficient in an IWC-Z relation from an average is equivalent to a deviation of the associated N_w from its mean value (Testud et al. 2001), the variation of a as a function of relative-height may be converted into a N_w as a function of relative-height relationship. We, therefore, use the data in Fig. 2 to parameterize N_w as a function of the relative height.
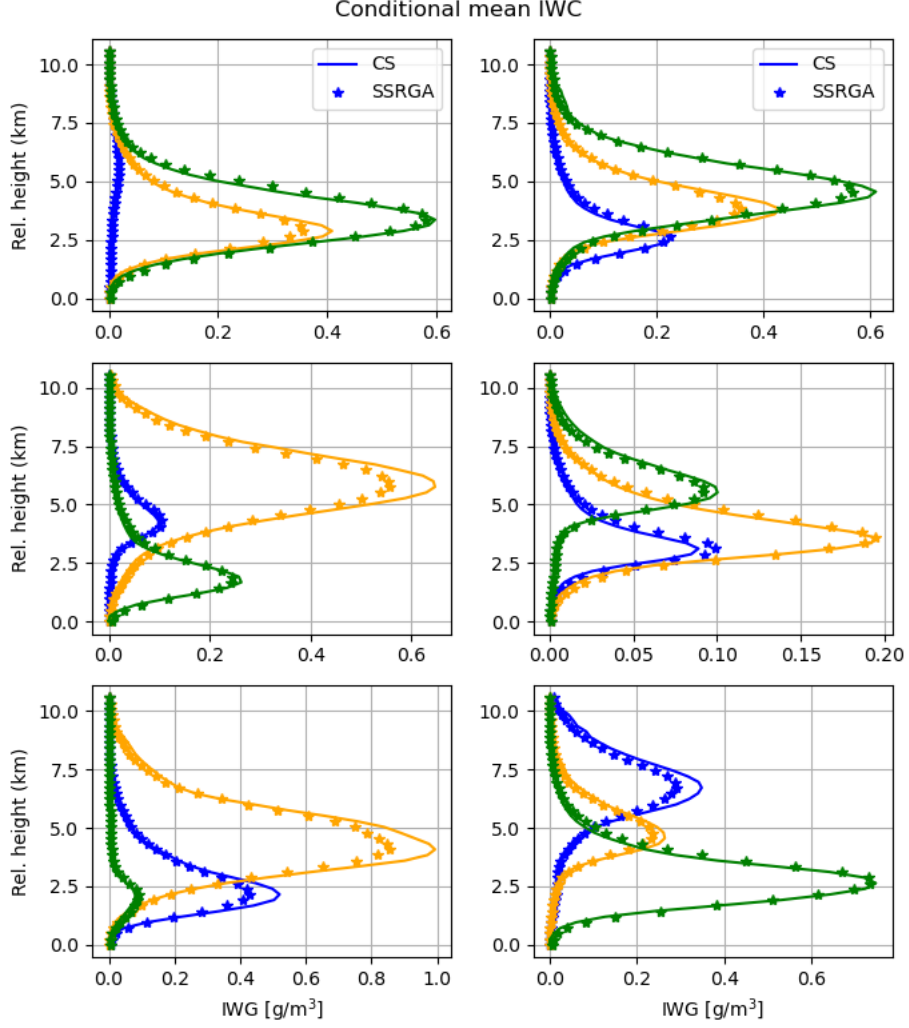
Conditional mean IWC

FIG. 1. Mean CS IWC profiles for 18 classes derived using the k-Means clustering algorithm. Associated mean profiles derived from CS reflectivity observations derived using SSRGA scattering calculations and N_w parameterization developed in this study are shown using symbol *. The vertical coordinate is defined relative to the freezing level

For the determination of reference a and b values to be used with Equation (1), we assume that PSDs are normalized gamma distributions with $N_w = 0.08 cm^{-4}$ and $\mu = 2$ and calculate

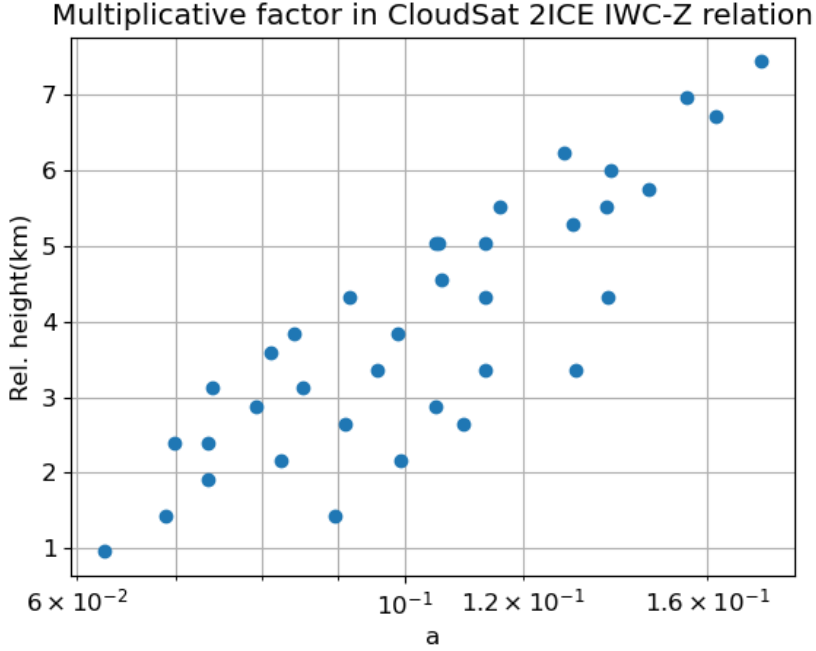$$Z = \frac{\lambda^4}{\pi^5 |K_w|^2} \int_0^\infty N(D, D_m) \sigma_b(D) dD \qquad (2)$$

6

FIG. 2.

where λ is the radar frequency, $|K_w|$ is the dielectric factor of water, $N(D, D_m)dD$ is the number of ice particles of diameter with D and D+dD per unit volume, D_m is the mass weighted mean diameter of the distribution, and $\sigma_b(D)$ is the backscattering cross-section of ice particle of diameter D. The mass weighted mean diameter is equidistantly sampled to span the entire range of IWC values in the CS 2C-ICE dataset. The assumed mass-size relation is that of Brown and Francis (1995) because it works well with the SSRGA scattering calculations (Heymsfield et al. 2022). The open source software scatter-1.1 of (Hogan 2019) is used to provide the actual scattering properties. The SSRGA theory was developed for millimeter and submillimeter-wave calculations and may not be applicable at lidar's wavelength. Therefore, for lidar calculations, we use the Mie solution included in the scatter-1.1 package. Although more accurate calculations based on more realistic ice particle shapes exist, they are rather incomplete and not readily available. Moreover, Wagner and Delene (2022) compared lidar backscatter observations with backscatter calculations based on coincident PSD observations and the Mie solution and found good agreement, which suggests that electromagnetic properties derived from Mie calculations are adequate for practical applications. The lidar molecular backscatter and extinction are calculated using the lidar module of the CFMIP Observation Simulator Package (COSP; Bodas-Salcedo et al. (2011)). To account for multiple-

7
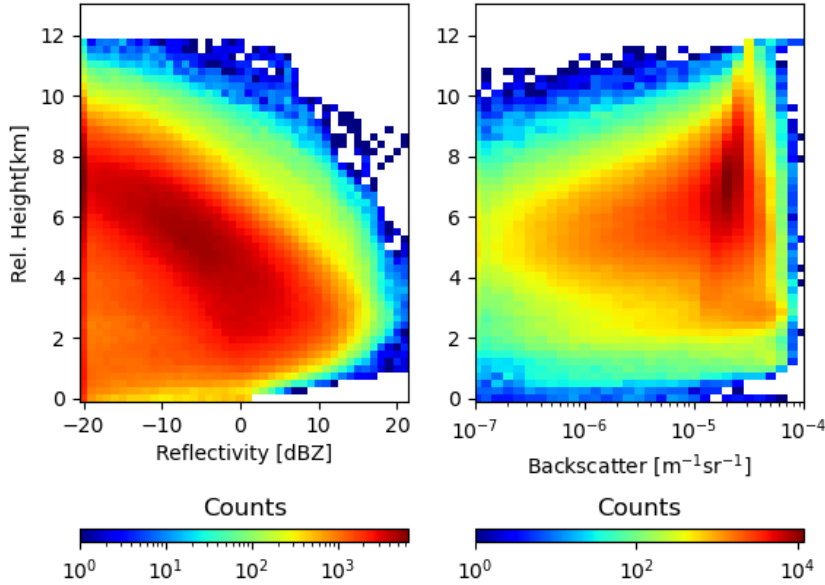
FIG. 3. Simulated distributions of Ku-band radar reflectivity (left) and lidar bacscatter (right) as function of height above the freezing level

scattering in the lidar observations, we are using the multiscatter-1.2.11 model (Hogan 2015) of Hogan and Battaglia (2008). Shown in Figure 3 are the distributions of simulated Ku-band radar reflectivity and lidar bacscatter as function of height above the freezing level.

The radiometer observations are calculated using a one-dimensional efficient, but accurate, radiative transfer solver based on Eddington's approximation (Kummerow 1993). The Eddington's approximation has been found to work well in cloud and precipitation retrieval application despite its simplicity relative to more general (but also computationally intensive) approaches such as the Monte Carlo radiative transfer solvers (Liu et al. 1996) It should be noted tough that the phase functions of ice particles tend to be highly asymmetric at sub-millimeter wave frequencies. For radiative transfer solutions based on the Eddington's approximation to be accurate it is necessary that the delta-scaling approach (Joseph et al. 1976) be employed. The delta-scaling approach transforms the initial radiative transfer equation into an equivalent one characterized by a less asymmetric scattering function and more extinction, which makes the solution Eddington approximation more stable and accurate. The absorption due to water vapor and other gases is quantified using the

8

Rosenkranz model (Rosenkranz 1998). The water vapor, temperature, and pressure distributions are derived based on a WRF simulation of summer convection over the United States. Specifically, the water vapor, temperature, and pressure profiles associated with times and areas where the model produces anvils are selected and clustered into 40 classes using the k-Means approach. The mean extinction profiles at the radiometer frequencies are calculated for every class and used in process of calculating the brightness temperatures from the estimated ice profiles using a simple Monte Carlo procedure. That is, given a retrieved ice profile and its scattering property, an anvil class and its associated absorption, temperature, and pressure profiles are randomly selected and attached to the ice scattering properties. To make the procedure physically meaningful, temperature rather than height is used in the ice scattering-gas absorption collocation process. The emissivities are randomly chosen between 0.8 and 1.0 and assumed constant for all radiometer frequencies. Brightness temperatures are calculated at 89-, 183.31 ± 1.1, and 325.15 ± 1.5 GHz, which correspond to three of the 10 channels of the SAPHIR-NG radiometer envisioned to be deployed in the AOS mission (Brogniez et al. 2022). The other channels are centered on the same water vapor absorption lines and are not likely to offer additional information in this rather controlled experiment. Nevertheless, the other channels are expected to be useful in reducing the uncertainties caused by variability in the vertical distribution of water vapor, which may be greater in real life than in the simulated environment.

The processing steps used to process the CS reflectivity observations and calculate the lidar, Ku-band and submillimeter-wave radiometer observations may be summarized as follows:

1. Derivation of physically consistent radar and radiometer lookup tables to relate basic radar and radiometer properties (e.g. reflectivity, attenuation, extinction, scattering-albedo, etc.) to PSD parameters such as IWC and D_m. The tables are derived for a single of N_w, but are usable with any value of N_w using the "normalization" operations described in (Grecu et al. 2011).

2. Derivation of Nw-relative height parameterization using the 2C-ICE product.

3. Estimation of IWC and related PSD parameters from CS W-band radar observations, using the tables constructed in Step 1 and parameterization derived in Step 2.

9

177 4. Calculation of lidar, Ku-band radar and radiometer observations from the estimates derived
178   in Step 3 and the tables obtained in Step 1.

179 The application of these steps produces a large dataset of approximately 200,000 cloud ice
180 profiles and associated lidar, radar and radiometer observations that may be used to investigate the
181 synergy of the three sensors. Details are provided in the next section.

b. Estimation and evaluation

183 Given that the lidar observations may attenuate quickly in thick clouds, while the Ku-band radar
184 will not detect clouds with an echo weaker than 8.0 dBZ, the radiometer is the instrument likely
185 to provide by itself the most complete information about the total amount of ice in its observing
186 volume. However, the vertical distribution of ice is difficult to quantify from radiometer-only
187 observations, because significantly different ice vertical distributions may lead to very similar
188 radiometer observations. This makes radiometer-only retrievals highly dependent on the "a priori"
189 information on the distribution of ice clouds in the atmosphere. As previously mentioned, this
190 is the reason why CS-based IWC retrievals were preferred to CRM simulations, as retrievals are
191 expected to result in more natural and less biased distributions.

192 We employ a two-step estimation methodology similar to that of Grecu et al. (2018). In the
193 first step, we estimate the IWC class, out of the 18 classes of shown in Figure 1, to which the
194 estimated IWC profile is most likely to belong. The class is estimated directly from the synthetic
195 observations. In the second step, we estimate the IWC profile, using a class specific ensemble
196 Kalman Smoother (EKS) methodology similar to that of Grecu et al. (2018). The EKS algorithm
197 updates the estimated IWC relative to the mean IWC of the class to which the profile belongs. The
198 differences between the actual active and passive observations and their mean class values are used
199 in the update. The second step of this procedure is formally identical to the one used in Grecu
200 et al. (2018), but the first step is different. In Grecu et al. (2018), the first step was based on a
201 simple distance-based evaluation. That strategy is likely to be suboptimal in this study, because the
202 joint distribution of IWC profiles and associated observations are significantly more complex. We
203 therefore use a more complex classification methodology based on the TensorFlow library (Abadi
204 et al. 2016). The class estimation model is defined as a TensorFlow Model with two dense layers
205 of 30 neurons each, followed by a softmax layer (Goodfellow et al. 2016). The class estimation

10

model is trained using the 70% of the simulated observations and the corresponding IWC profiles, the remaining 30% of the data being used for evaluation.

$$\mathbf{X} = \bar{\mathbf{X}}_{\mathbf{i}} + \mathbf{Cov}(\mathbf{X}_{\mathbf{i}}, \mathbf{Y}_{\mathbf{i}}) \mathbf{Cov}(\mathbf{Y}_{\mathbf{i}}, \mathbf{Y}_{\mathbf{i}})^{-1} (\mathbf{Y} - \bar{\mathbf{Y}}_{\mathbf{i}}) \qquad (3)$$

where $\mathbf{X}$ is the state variable describing the IWC profile, $\mathbf{Y}$ is the vector containing the variation, $\mathbf{X}_{\mathbf{i}}$ is the set of state variables for profiles in class i, and $\mathbf{Y}_{\mathbf{i}}$ is the set of associated observations. Variables $\bar{\mathbf{X}}_{\mathbf{i}}$ and $\bar{\mathbf{Y}}_{\mathbf{i}}$ are the mean values of the state variables and observations in class i, respectively. The covariance matrices between $\mathbf{X}_{\mathbf{i}}$ and $\mathbf{Y}_{\mathbf{i}}$ are denoted by $\mathbf{Cov}(\mathbf{X}_{\mathbf{i}}, \mathbf{Y}_{\mathbf{i}})$. In step 1, the class is estimated using the TensorFlow model, while in step 2, the IWC profile is estimated using the EKS algorithm summarized in Equation 3.

As already mentioned, a cross-validation methodology is used for evaluation, with 70% of the data used for training and the remaining 30% of the data used for validation. The partition of the data into training and evaluation subsets is done randomly. Usually, the partition, training and evaluation steps are repeated several times. However, given the fact that differences in the relationships between the ice property and their associated simulated observations are functions of the meteorological context, and that all regimes are well-sampled in both the training and testing subsets (e.g. out of every 10 pixels in a scene, about 7 end-up in the training dataset, while the others in the testing dataset), the repetition of the partition, training, and evaluation steps multiple times is not necessary. Therefore, in our evaluation, we partition the data into training and evaluation only once and perform all the evaluation for a single partition. The evaluation criteria include the correlation coefficient, the bias, and visual inspections of graphical representations of the estimated properties relative to their references.

3. Results

a. Radiometer-only retrievals

As previously mentioned, submillimeter-wave radiometers are likely to provide by themselves more complete information about the total amount of ice in their observing volumes than lidars or Ku-band radars with limited sensitivity. However, radiometers observations are an integrated measure of radiative process in clouds that provide little information about the vertical distribution

11

of ice. From this perspective, an evaluation in terms of the ice water path (IWP) defined as the vertical integral of the IWC, i.e. $IWP = \int_0^{Z_{top}} IWC(z)dz$ is insightful. Shown in Figure 4 is the frequency of IWP estimated from radiometer-only observations as a function of its true value. As apparent in the figure, there is good correlation between the retrieved and the true IWP values. The numerical value of the correlation coefficient is 0.92, and there is no-overall bias. That is, the mean values of retrieved IWP and true IWP values are equal. However, conditional biases are apparent, with overestimation of IWP for values smaller than 100 g/m^2 and some underestimation for values larger than 1000 g/m^2. The biases at the low end of the IWP range are not surprising, given that the impact caused by ice scattering on the total radiometric signal is small for low values of IWP and hard to distinguish from other sources of variability in radiometer observations. Saturation effects are most likely responsible for underestimation at the high end. It should be noted that in this evaluation, only atmospheric profiles that exhibit ice detectable by the CS radar are used. Therefore, a radiometer-only estimation procedure derived from this training dataset is likely to result in significant overestimation if not used in conjunction with a discrimination procedure. However, such procedure is not critical in this study, as the lidar observations may be used to discriminate between clear skies and ice clouds. Although the radiometer-only estimation procedure is able to estimate the integrated amount of ice in clouds fairly wells, its ability to characterize the vertical distribution of ice in clouds is limited. Figure 5 shows the conditional vertical distributions of the estimated and true IWC for the 18 classes described in Section 2a and shown in Figure 1. As apparent in the figure, there are significant differences between the estimated and true IWC profiles.

Further insight into the radiometer-only estimation performance may be derived by defining the ice profile gravity center (GC) as $z_{GC} = \frac{\int_0^{Z_{top}} zIWC(z)dz}{\int_0^{Z_{top}} IWC(z)dz}$, where z is the distance relative to the freezing level, the Z_{top} is the distance from the top of the atmosphere to the freezing level. Shown in Figure 6 is the frequency of IWC gravity center estimated from radiometer-only observations as a function of its true value. It may be observed in the figure that while the true IWC gravity center exhibits quite a broad distribution, the one retrieved from the radiometer-only observations exhibits a multimodal narrow distribution. Moreover, there is almost no correlation between the retrieved and the true IWC gravity center. This is another indication that, while the total amount
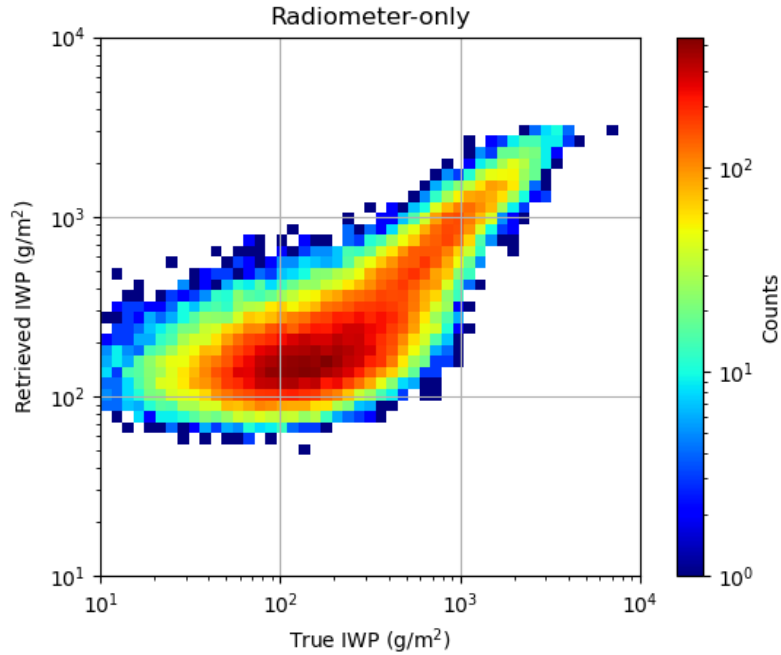
FIG. 4. Frequency plot of estimated IWP derived radiometer-observations as a function of the true IWP used in observations synthesis

of ice may be reasonably estimated from radiometer-only observations, its vertical distribution can not be determined from radiometer-only observations.

b. Synergistic retrievals

The synergy of the instrument on the estimates may be investigating by simply incorporating lidar and radar observations into the retrieval process and comparing the results with the radiometer-only estimates. Although the lidar observations are subject to attenuation, they are able to provide information about the vertical distribution of ice in clouds, especially at the top of the clouds. The radar observations, on the other hand, are able to provide information in the bottom part of the clouds, where the lidar signal is below the noise level due to attenuation. Therefore, the combined used of lidar and radar observations is expected to provide a more complete characterization of the vertical distribution of ice in clouds and enable the derivation of more specific estimates than those derived from radiometer-only observations.
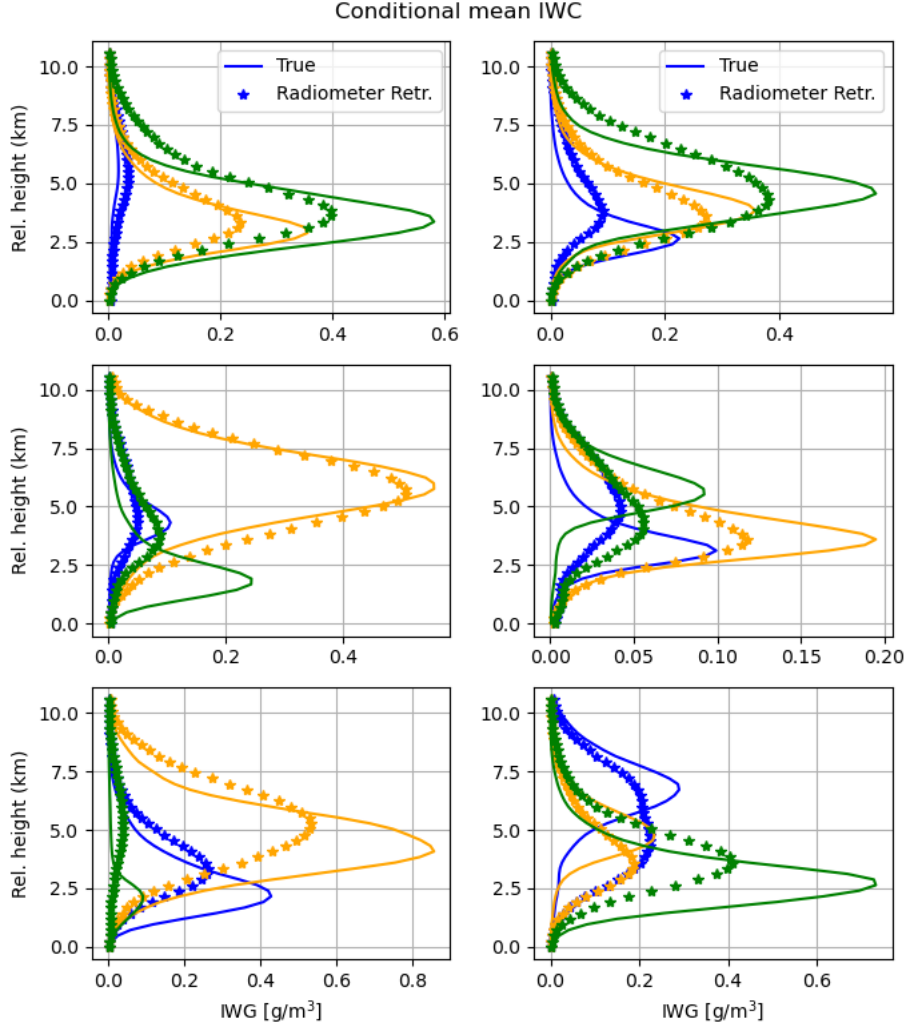
FIG. 5. True and radiometer-only retrieved conditional mean IWC for the 18 classes described in Figure 1.

Shown in Figure 7 is the distribution of the synergistic IWP estimates as a function of their true values. As apparent in the figure, the synergistic IWP estimates are more accurate than the radiometer-only estimates. At the same time, as apparent in Figure 8, the retrieved conditional mean IWC for the 18 classes described section 2a and shown in Figure 1 are in significantly better (almost perfect) agreement with the true IWC profiles than those derived from radiometer-only observations. Furthermore, as seen in Figure 8 the synergistic IWC gravity center estimates are in very good agreement with the true IWC gravity center.

While the estimates based on all instruments are significantly more accurate than those based on radiometer-only observations, it is useful to investigate how the two active instruments (lidar and
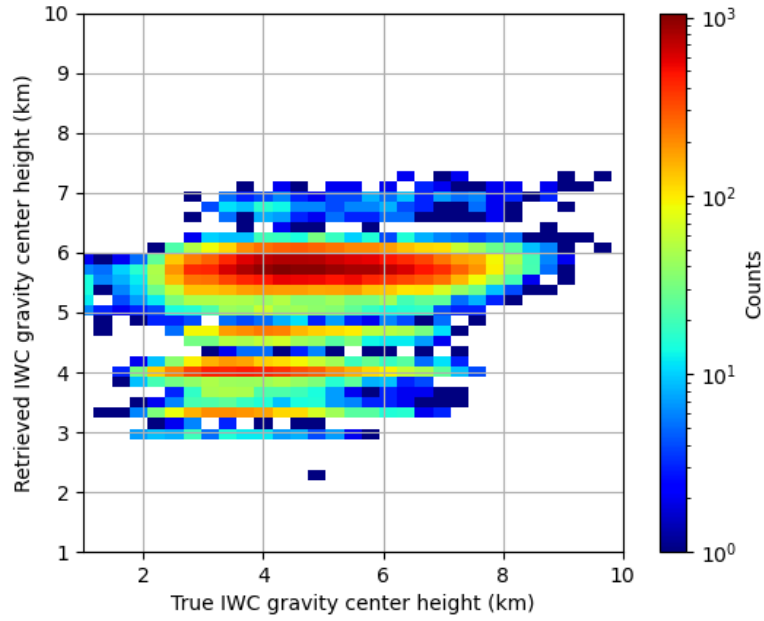
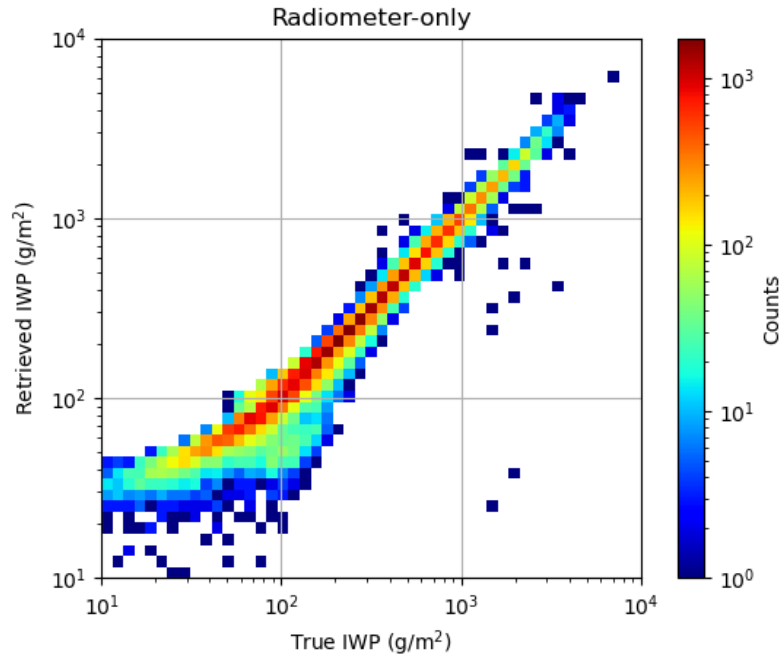FIG. 6. Same as in Figure 4, but for the *IWC* gravity center.

FIG. 7. Same as in Figure 4, but with the lidar and radar observations incorporated in the retrievals.
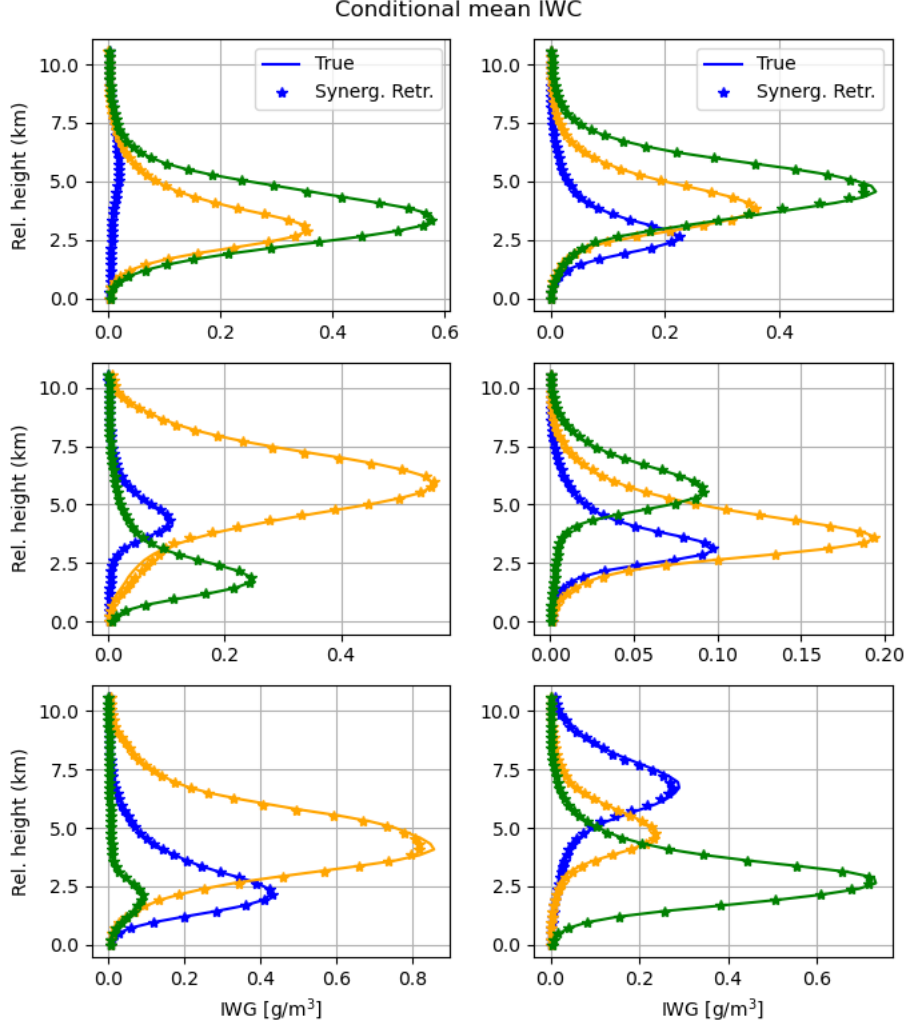
Conditional mean IWC

FIG. 8. Same as in Figure 5, but with the lidar and radar observations incorporated in the retrievals.

radar) impact the estimates. For conciseness, we used two statistical scores, namely, the normalized root mean square (NRMS) and the classification accuracy, to summarize the performance of the estimates. The NRMS is defined as

$$NRMS = \frac{\sqrt{\frac{\sum_{i=1}^{N}(IWC_i - IWC_{true,i})^2}{N}}}{\sqrt{\frac{\sum_{i=1}^{N}(IWC_{true,i} - \overline{IWC})^2}{N}}} \tag{4}$$

where IWC_i is the estimated IWP for the i-th sample, $IWC_{true,i}$ is the true IWC for the i-th sample, $\overline{IWC}$ is the IWC mean, and N is the size of the estimation dataset. The classification accuracy is
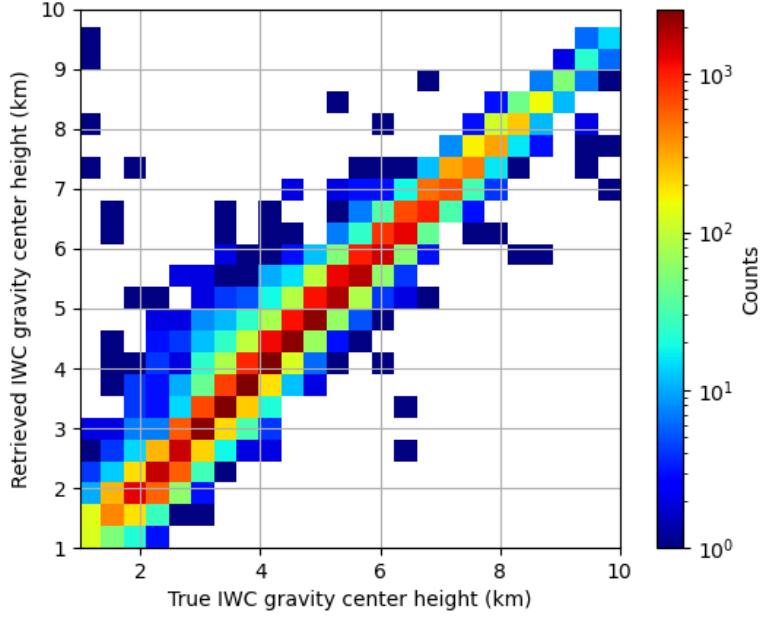
16

FIG. 9. Same as in Figure 6, but with the lidar and radar observations incorporated in the retrievals.

defined as

$$CA = \frac{\sum_{i=1}^{N} \delta_i}{N} \tag{5}$$

where δ_i is a binary variable that is equal to 1 if the estimated IWC class for the i-th sample is equal to the true IWC class for the i-th sample, and 0 otherwise. The performance summary is shown in Table 1 for several combinations of instruments. It may be observed in the table that the performance of the estimates based on all instruments is significantly better than those based on radiometer-only observations. Furthermore, the inclusion of the lidar observations in the retrieval process has a larger impact on the retrieval performance the inclusion of the radar observations. This is expected since the lidar observations are able to provide information about the top of the clouds, where the radar observations are above the noise level only occasionally. Nevertheless, the inclusion of the radar observations in the retrieval process has a notable impact on the accuracy of the IWC estimates relative to radiometer-only retrievals.

17

TABLE 1. Performance summary.

Instruments / Score	Radiometer	Radar-Radiometer	Lidar-Radiometer	Radar-Lidar-Radiometer
NRMS	0.73	0.59	0.32	0.22
Class. Accurracy	0.39	0.48	0.92	0.94

4. Conclusions

In this study, we investigate the synergy of lidar, Ku-band radar, and sub-millimeter-wave radiometer measurements in the retrieval of the ice from satellite observations. The synergy is analyzed through the generation of a large dataset of IWC profile and the calculation of lidar, radar and radiometer observations using realistic models. The characteristics of the instruments (e.g. frequencies, sensitivities, etc.) are set based on the expected characteristics of instruments of the AOS mission. A cross-validation methodology is used to assess the accuracy of the retrieved IWC profiles from various combinations of observations from the three instruments. Specifically, the IWC and associated observations is randomly divided into two datasets, one for the training and the other for the evaluation. The training dataset is used to train the retrieval algorithm, while the evaluation dataset is used to assess the retrieval performance.

To ensure the self-consistency of results and their relevance to practical applications, the dataset of IWC profiles is derived from CloudSat reflectivity observations. Although subject to potential biases and uncertainties due to deficiencies in the retrieval models, these profiles are deemed to be more realistic than those derived from cloud resolving model simulations. Moreover, they are roughly consistent with the 2C-ICE CloudSat product (Deng et al. 2015), while relying on assumptions and parameterizations that enable the accurate computation of backscatter lidar, Ku-band radar, and sub-millimiter-wave radiometer observations.

The retrieval of the ice water content (IWC) profiles from the computed observations is achieved in two steps. In the first step, a class, out of 18 potential classes characterized by different vertical distribution of IWC, is estimated from the observations. The 18 classes are predetermined based on k-Means clustering algorithm. In the second step, the IWC profile is estimated using an Ensemble Kalman Smoother (EKS) algorithm that uses the estimated class as a priori information.

The results of the study show that the synergy of lidar, radar, and radiometer observations is significant in the retrieval of the IWC profiles. The inclusion of the lidar observations in the retrieval

18

process has a larger impact on the retrieval performance than the inclusion of the radar observations. Although results are not directly comparable to those from other studies (Pfreundschuh et al. 2020; Liu and Mace 2022), given the differences between the instruments in this study relative to those from other studies, it may be concluded that they are not inconsistent with previous studies. Specifically, previous studies show some skills in radiometer-only retrievals and significant improvements in the retrieval performance when the active observations are incorporated. From this perspective, our findings are consistent with previous studies.

Further work is necessary out to assess the impact of sources of uncertainties such as variability in the PSD intercept not captured by the current parameterization, differences in the instruments' footprint sizes, and non-uniform beam filling on the retrievals of the IWC profiles. Other sources of uncertainties that need be considered include the potential existence of supercooled liquid water in the clouds and uncertainties in the electromagnetic scattering properties used in the in instruments' forward models.

Data availability statement. The CloudSat data can be accessed at: https://www.cloudsat.cira.colostate.edu/.

References

Abadi, M., and Coauthors, 2016: Tensorflow: a system for large-scale machine learning. *Osdi*, Savannah, GA, USA, Vol. 16, 265–283.

Bodas-Salcedo, A., and Coauthors, 2011: Cosp: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, **92 (8)**, 1023–1043.

Bringi, V., V. Chandrasekar, J. Hubbert, E. Gorgucci, W. Randeu, and M. Schoenhuber, 2003: Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis. *Journal of the atmospheric sciences*, **60 (2)**, 354–365.

Brogniez, H., and Coauthors, 2022: Time-delayed tandem microwave observations of tropical deep convection: Overview of the c 2 omodo mission. *Frontiers in Remote Sensing*, **3**, 854 735.

Brown, P. R., and P. N. Francis, 1995: Improved measurements of the ice water content in cirrus using a total-water probe. *Journal of Atmospheric and Oceanic Technology*, **12 (2)**, 410–414.

Delanoë, J., and R. J. Hogan, 2008: A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer. *Journal of Geophysical Research: Atmospheres*, **113 (D7)**.

Deng, M., G. G. Mace, Z. Wang, and E. Berry, 2015: Cloudsat 2c-ice product update with a new ze parameterization in lidar-only region. *Journal of Geophysical Research: Atmospheres*, **120 (23)**, 12–198.

Deng, M., G. G. Mace, Z. Wang, and R. P. Lawson, 2013: Evaluation of several a-train ice cloud retrieval products with in situ measurements collected during the sparticus campaign. *Journal of applied meteorology and climatology*, **52 (4)**, 1014–1030.

Deng, M., G. G. Mace, Z. Wang, and H. Okamoto, 2010: Tropical composition, cloud and climate coupling experiment validation for cirrus cloud profiling retrieval using cloudsat radar and calipso lidar. *Journal of Geophysical Research: Atmospheres*, **115 (D10)**.

Goodfellow, I., Y. Bengio, and A. Courville, 2016: *Deep learning*. MIT press.

Grecu, M., L. Tian, G. M. Heymsfield, A. Tokay, W. S. Olson, A. J. Heymsfield, and A. Bansemer, 2018: Nonparametric methodology to estimate precipitating ice from multiple-frequency radar reflectivity observations. *Journal of Applied Meteorology and Climatology*, **57 (11)**, 2605–2622.

Grecu, M., L. Tian, W. S. Olson, and S. Tanelli, 2011: A robust dual-frequency radar profiling algorithm. *Journal of applied meteorology and climatology*, **50 (7)**, 1543–1557.

Heymsfield, A., A. Bansemer, G. Heymsfield, D. Noone, M. Grecu, and D. Toohey, 2022: Relationship of multiwavelength radar measurements to ice microphysics from the impacts field program. *Journal of Applied Meteorology and Climatology*.

Hogan, R. J., 2019: scatter-1.1. URL http://www.met.reading.ac.uk/clouds/ssrga/scatter-1.1.tar.gz.

Hogan, R. J., and A. Battaglia, 2008: Fast lidar and radar multiple-scattering models. part ii: Wide-angle scattering using the time-dependent two-stream approximation. *Journal of the Atmospheric Sciences*, **65 (12)**, 3636–3651.

Hogan, R. J., R. Honeyager, J. Tyynelä, and S. Kneifel, 2017: Calculating the millimetre-wave scattering phase function of snowflakes using the self-similar rayleigh–gans approximation. *Quarterly Journal of the Royal Meteorological Society*, **143 (703)**, 834–844.

Hogan, R. J., M. P. Mittermaier, and A. J. Illingworth, 2006: The retrieval of ice water content from radar reflectivity factor and temperature and its use in evaluating a mesoscale model. *Journal of Applied Meteorology and Climatology*, **45 (2)**, 301–317.

Joseph, J. H., W. Wiscombe, and J. Weinman, 1976: The delta-eddington approximation for radiative flux transfer. *Journal of Atmospheric Sciences*, **33 (12)**, 2452–2459.

Kummerow, C., 1993: On the accuracy of the eddington approximation for radiative transfer in the microwave frequencies. *Journal of Geophysical Research: Atmospheres*, **98 (D2)**, 2757–2765.

Liu, Q., C. Simmer, and E. Ruprecht, 1996: Three-dimensional radiative transfer effects of clouds in the microwave spectral range. *Journal of Geophysical Research: Atmospheres*, **101 (D2)**, 4289–4298.

Liu, Y., and G. G. Mace, 2022: Assessing synergistic radar and radiometer capability in retrieving ice cloud microphysics based on hybrid bayesian algorithms. *Atmospheric Measurement Techniques*, **15 (4)**, 927–944.

MacKay, D. J., 2003: *Information theory, inference and learning algorithms*. Cambridge university press.

Pfreundschuh, S., P. Eriksson, S. A. Buehler, M. Brath, D. Duncan, R. Larsson, and R. Ekelund, 2020: Synergistic radar and radiometer retrievals of ice hydrometeors. *Atmospheric Measurement Techniques*, **13 (8)**, 4219–4245.

Rodgers, C. D., 2000: *Inverse methods for atmospheric sounding: theory and practice*, Vol. 2. World scientific.

Rosenkranz, P. W., 1998: Water vapor microwave continuum absorption: A comparison of measurements and models. *Radio Science*, **33 (4)**, 919–928.

Stephens, G. L., and Coauthors, 2002: The cloudsat mission and the a-train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, **83 (12)**, 1771–1790.

Testud, J., S. Oury, R. A. Black, P. Amayenc, and X. Dou, 2001: The concept of "normalized" distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing. *Journal of Applied Meteorology and Climatology*, **40 (6)**, 1118–1140.

Wagner, S. W., and D. J. Delene, 2022: Technique for comparison of backscatter coefficients derived from in situ cloud probe measurements with concurrent airborne lidar. *Atmospheric Measurement Techniques*, **15 (21)**, 6447–6466.