A Dual-Frequency Radar Retrieval of Two Parameters of the Snowfall Particle Size Distribution Using a Neural Network

RANDY J. CHASE, a STEPHEN W. NESBITT, a AND GREG M. MCFARQUHAR b,c

a Department of Atmospheric Sciences, University of Illinois at Urbana–Champaign, Urbana, Illinois
b Cooperative Institute of Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma
c School of Meteorology, University of Oklahoma, Norman, Oklahoma

(Manuscript received 12 August 2020, in final form 18 November 2020)

ABSTRACT: With the launch of the Global Precipitation Measurement Dual-Frequency Precipitation Radar (GPM-DPR) in 2014, renewed interest in retrievals of snowfall in the atmospheric column has occurred. The current operational GPM-DPR retrieval largely underestimates surface snowfall accumulation. Here, a neural network (NN) trained on data that are synthetically derived from state-of-the-art ice particle scattering models and measured in situ particle size distributions (PSDs) is used to retrieve two parameters of the PSD: liquid equivalent mass-weighted mean diameter Dml and the liquid equivalent normalized intercept parameter Nwl. Evaluations against a test dataset showed statistically significantly improved ice water content (IWC) retrievals relative to a standard power-law approach and an estimate of the current GPM-DPR algorithm. Furthermore, estimated median percent errors (MPE) on the test dataset were $-0.7\%$, $+2.6\%$, and $+1\%$ for $D_{ml}$, $N_{wl}$, and IWC, respectively. An evaluation on three case studies with collocated radar observations and in situ microphysical data shows that the NN retrieval has MPE of $-13\%$, $+120\%$, and $+10\%$ for $D_{ml}$, $N_{wl}$, and IWC, respectively. The NN retrieval applied directly to GPM-DPR data provides improved snowfall retrievals relative to the default algorithm, removing the default algorithm’s ray-to-ray instabilities and recreating the high-resolution radar retrieval results to within $15\%$ MPE. Future work should aim to improve the retrieval by including PSD data collected in more diverse conditions and rimed particles. Furthermore, different desired outputs such as the PSD shape parameter and snowfall rate could be included in future iterations.

KEYWORDS: Snowfall; Ice particles; Aircraft observations; Radars/Radar observations; Remote sensing; Neural networks

1. Introduction

Despite being confined to high latitudes and altitudes when occurring at the surface, snow can be related to approximately 50% by mass accumulation (Heymsfield et al. 2020) of 50% by number (Field and Heymsfield 2015) and approximately occurring at the surface, snow can be related to approximately models (e.g., Duncan and Eriksson 2018). Agreement between observational retrievals and numerical atmosphere’s ice water path remains large, and there is poor quantification of the hydrologic cycle. Furthermore, the evaluation of atmospheric numerical model simulations and their parameterized ice-phase microphysics (e.g., Delanoë et al. 2011; Stein et al. 2015; Ori et al. 2020). Despite many advances in satellite remote sensing techniques and sensors in the past few decades, the uncertainty in the estimate of the atmosphere’s ice water path remains large, and there is poor agreement between observational retrievals and numerical models (e.g., Duncan and Eriksson 2018).

The best way to retrieve global snowfall properties is to use spaceborne microwave radars since ground-based observations are limited to easily accessible locations and passive spaceborne sensors have additional ambiguity in determining the vertical distribution of hydrometeors. Currently, there exist two NASA missions with spaceborne radars designed to sample hydrometeors. The first mission, launched in 2006, is CloudSat (Stephens et al. 2002), which consists of a highly sensitive 94 GHz nonscanning cloud radar in a 98° inclination orbit. The second mission, starting in 2014, is the Global Precipitation Measurement mission (GPM; Hou et al. 2014), which operates the Dual-Frequency Precipitation Radar (DPR; 13.5 and 35.5 GHz) in a 65° inclination orbit. Both satellites have collected observations of equivalent radar reflectivity factor Z_e in a variety of snowfall over their missions.

To retrieve snowfall properties, namely snowfall rate S, the most common method has been the use of a prescribed power-law relation between $Z_e$ and S that takes the form

$$Z_e = \alpha S^\beta.$$ (1)

The parameters in Eq. (1), $\alpha$ and $\beta$, are typically derived from direct comparisons of observed $Z_e$ and S (e.g., Langille and Than 1951; Boucher and Wieler 1985; Fujiyoshi et al. 1990), theoretical particle scattering models (e.g., Liu 2008; Kulie and Bennartz 2009), or a combination of both (e.g., Matrosov 1992;
von Lerber et al. 2017). The direct comparison of observed Ze and S is sensitive to the radar’s calibration and operating frequency, as well as intrinsic particle properties including the large observed variability in particle size distributions, crystal habits, and degrees of riming, making it difficult to find a relation that can be applied universally. The use of theoretical scattering models (e.g., Mie theory; T matrix; discrete dipole approximation) is largely dependent on the microphysical assumptions made within them, leading uncertainties in the estimate of S up to a factor of 10 (Hiley et al. 2011). Thus, it should be clear that in order to improve the estimate of global snowfall properties, methods beyond using a single standard power-law fit must be explored.

Currently, the operational retrieval method used in CloudSat uses optimal estimation (Rodgers 2000), where the observed Ze and a priori constraints are used to retrieve snowfall properties (Wood et al. 2013; Wood and L’Ecuyer 2021). Since this method has shown good agreement with ground-based radar retrievals in the United States (Cao et al. 2014; Chen et al. 2016), Sweden (Norin et al. 2015), and Antarctica (Souverijns et al. 2018), CloudSat is currently considered the best estimate of global snowfall properties and has been used in numerous snowfall investigations (e.g., Palerme et al. 2014; Kulie et al. 2016; Palerme et al. 2017; Milani et al. 2018; Kulie and Milani 2018; Kulie et al. 2020). The uncertainty in CloudSat’s retrieval of S has improved with the optimal estimation technique relative to a single power-law fit, although the retrieval has nontrivial uncertainties estimated to be between 50% (Palerme et al. 2014) and 160% (Kulie et al. 2020).

To retrieve snowfall properties from GPM-DPR, an algorithm is used that prescribes a relation between the precipitation rate, mass-weighted mean diameter Dm, and Ze to simultaneously retrieve the properties of hydrometeors regardless of phase (Kozu et al. 2009; Iguchi et al. 2018). Direct comparisons between GPM-DPR snowfall retrievals and surface-based references have yet to be conducted, but the snowfall retrievals have still been used to investigate the global distribution of snowfall (e.g., Adhikari et al. 2018; Adhikari and Liu 2019). However, a comparison using CloudSat as a reference has shown that GPM-DPR’s retrieval of global average snowfall rate is underestimated by approximately 43% even after considering measurement differences between CloudSat and GPM-DPR (Skofronick-Jackson et al. 2019). Furthermore, an investigation of the GPM-DPR retrieval microphysical assumptions by Chase et al. (2020) showed that the current GPM-DPR algorithm is likely inappropriate for snowfall retrievals and thus other retrieval methods should be investigated.

One alternative retrieval method for GPM-DPR is to adopt the same optimal estimation technique as CloudSat but at GPM-DPR frequencies. This has been shown to work well for triple-frequency observations from field campaigns (Grecu et al. 2018; Leinonen et al. 2018; Tridon et al. 2019) and could potentially be applied to dual-frequency observations. A second method, explored in this paper, would be to use an artificial neural network (NN) to retrieve snowfall properties.

Neural networks have been widely used in remote sensing (Mas and Flores 2008), including the classification of clouds from passive satellite irradiances (e.g., Key et al. 1989) as well as the retrieval of snowfall properties (e.g., Xiao et al. 1998; Sekelsky et al. 1999). Xiao et al. (1998) trained several NNs using the vertical column of single-frequency radar measurements and ground based measured snowfall accumulations, showing that the NN can perform better than the standard power-law approach. In Sekelsky et al. (1999), an NN was trained on scattering simulations produced by a T-matrix code (Mishchenko and Travis 1998) at three frequencies (S, Ka, and W bands) and a range of theoretical negative exponential distributions of particles in order to retrieve the volume weighted mean diameter D0 and the negative exponential intercept parameter N0. Since the study by Sekelsky et al. (1999), comparisons with more accurate numerical modeling of ice particle scattering, namely the discrete dipole approximation (DDA; Yurkin and Hoekstra 2011), have shown that T matrix predicted scattering properties cannot fully reproduce 94 GHz scattering properties of aggregates (Kniefel et al. 2011). Furthermore, the results from Sekelsky et al. (1999) were never evaluated against observations and thus the uncertainties of the retrieval are unquantified and unknown. Both Xiao et al. (1998) and Sekelsky et al. (1999) show that NNs contain potential for accurate retrievals of snowfall properties. Thus, a renewed investigation of NNs with the latest results from scattering models and ice particle observations (i.e., in situ and radar) is motivated.

Here, an NN for retrieving two parameters of the normalized gamma distribution (Testud et al. 2001; Delanoë et al. 2014), namely the normalized intercept parameter N0 and the mass-weighted mean diameter Dm, from radar observables is motivated. Specifically, scattering results from numerous DDA and generalized multiparticle Mie method (GMM; Xu 1995) simulations of a wide variety of unrimed particle types are used in conjunction with measured particle size distributions (PSDs) from NASA field campaigns to synthesize a database of snowfall properties and their associated ze at GPM-DPR frequencies. This database is then used to train an NN for the simultaneous retrieval of N0 and Dm, which is evaluated against a standard power-law retrieval as well as an estimate of the current GPM-DPR retrieval. Then the NN retrieval is evaluated on coincident observations of Ze and in situ snow properties from three case studies obtained from NASA Ground Validation campaigns. The paper is structured as follows: section 2 describes that data used in this study and how the synthetic database is generated. Section 3 contains the results of the evaluation of the retrievals on the synthetic database as well as three case studies. Section 4 discusses how the NN retrieval compares to the operational GPM-DPR algorithm and how it could be implemented on the GPM-DPR record. Section 5 summarizes the results and conclusions.

2. Data and methods

a. Description of PSD parameters

The features used for retrieving N0 and Dm within the NN are Ze at Ku band, dual-frequency ratio (DFR_{Ku-Ka}), and temperature (T). Using in situ aircraft observations of PSDs sorted into different bin sizes, Ze can be calculated by
\[ Z_{r,A} = \frac{\Lambda^4}{\pi^2 |K|^2} \sum_{i=0}^{n_{\text{bins}}} \sigma_{\text{bsc}}(D_r, \Lambda) N(D_r) \Delta D_r, \]  

(2)

where \( \lambda \) is the radar wavelength, \( |K|^2 \) is the dielectric factor (0.93 and 0.92 for Ku and Ka band, respectively), \( \sigma_{\text{bsc}}(D_r, \Lambda) \) is the backscatter cross section, and \( N(D_r) \) is the number distribution function for particles with maximum dimension of \( D_r \). The \( i \) subscript indicates the \( i \)th bin of the PSD, and \( \Delta D_r \) is the bin width of the \( i \)th bin. The DFRKu-Ka is defined as the ratio of \( Z_r \) at Ku band (\( Z_{r,Ku} \)) to \( Z_r \) at Ka band (\( Z_{r,Ka} \)):

\[ \text{DFRKu-Ka} = 10 \log_{10} \left( \frac{Z_{r,Ku}}{Z_{r,Ka}} \right). \]  

(3)

The output targets for the retrieval are adapted versions of the three-parameter normalized gamma distribution (Testud et al. 2001; Delanoë et al. 2014). Specifically, the solid-phase mass-weighted mean diameter \( D_{m_s} \), the liquid equivalent mass-weighted mean diameter \( D_{ml} \), and the liquid equivalent normalized intercept parameter \( N_{wl} \). Adapted versions of the normalized gamma distribution are used to supply consistency with the current GPM-DPR retrieval that outputs \( D_{m_s} \) and \( N_{wl} \). The \( D_{m_s} \) is defined as

\[ D_{m_s} = \frac{\sum_i m(D_r) N(D_r) \Delta D_r}{\sum_i m(D_r) N(D_r) \Delta D_r}, \]  

(4)

where \( m(D_r) \) is the mass of a particle with maximum dimension \( D_r \). The actual assumption of the mass of each particle is discussed more in sections 2e and 2g. To calculate \( D_{m_s} \), the liquid equivalent dimension \( D_{\text{melted}} \) must be known, and for the \( i \)th bin it is defined as

\[ D_{\text{melted}} = \left[ \frac{6m(D_r)}{\pi \rho_l} \right]^{1/3}, \]  

(5)

where \( \rho_l \) is the density of liquid water. Then \( D_{m_s} \) is defined as

\[ D_{m_s} = \frac{\sum_i m(D_r) D_{\text{melted}} N(D_r) \Delta D_r}{\sum_i m(D_r) N(D_r) \Delta D_r}. \]  

(6)

Chase et al. (2020) provides more details and justification for Eq. (6). The liquid equivalent normalized intercept parameter is defined following Delanoë et al. (2014), except using the \( D_{\text{melted}} \) rather than \( D_r \), expressed as

\[ N_{wl} = \frac{\left[ \sum_i D_{\text{melted}}^4 N(D_r) \Delta D_r \right]^6}{6 \left[ \sum_i D_{\text{melted}}^4 N(D_r) \Delta D_r \right]^2}. \]  

(7)

Ice water content (IWC) can then be calculated from \( D_{m_s} \) and \( N_{wl} \) by

\[ \text{IWC} = \frac{N_{wl} D_{m_s} \rho_l \pi}{4^3}. \]  

(8)

The retrieval of \( S \) was not performed herein because of the added uncertainty from the myriad of particle terminal fall velocities associated with snowfall, but one could calculate \( S \) with the retrieved \( D_{m_s} \) and \( N_{wl} \), assuming a value for \( \mu \) (shape parameter in the three-parameter gamma distribution; Borue et al. (2019) provides the guidance on the distribution of \( \mu \) from the GPM Cold Season Precipitation Experiment (GCPEX)] and assuming values for the terminal fall velocity as a function of \( D_r \). Furthermore, there is no current method to observe \( S \) from an airborne platform, which is the primary evaluation method used here.

b. DDA-GMM database

The results of numerous investigations of snowfall scattering of unrimed crystals at microwave wavelengths (Leinonen and Moisseev 2015; Leinonen and Szyrmer 2015; Lu et al. 2016; Kuo et al. 2016; Eriksson et al. 2018) are combined into a single database. Particle habits within these studies include pristine monomer shapes that occur in the atmosphere such as bullet rosettes, dendrites, plates, columns, and different aggregates consisting of these habits. The range of \( D_r \) within the combined dataset is from 13 \( \mu \)m to 6.3 cm. The dependence of \( \sigma_{\text{bsc}} \) on mass \( m \) for the Ku and Ka bands and the dependence of \( m \) on \( D_r \) for all particles used in the database are shown in Figs. 1a–c, respectively. Interestingly, even though only unrimed particles are included, a power-law fit between \( m \) and \( D_r \) to the amalgamation of all particles results in similar power-law relationship of moderately rimed particles reported in Leinonen and Szyrmer (2015) (Fig. 1c).

c. NASA GV observations

Observations collected as part of the NASA Ground Validation (GV) field campaigns (Petersen et al. 2020) were used in both the formulation of the NN and its evaluation. The formulation of the NN requires a large dataset of the three aforementioned input features (\( Z_r \), DFRKu-Ka, and \( T \)), which in turn requires the use of an estimate of \( N(D_r) \). In situ data collected on the University of North Dakota’s Citation Aircraft (Delene et al. 2019) during the Midlatitude Continental Convective Clouds Experiment (MC3E; Jensen et al. 2016), GCPEX (Skofronick-Jackson et al. 2015) and the Olympic Mountains Experiment (OLYMPEX; Houze et al. 2017) are used here. The \( N(D_r) \) is derived from measurements by two optical array probes (OAP) that capture silhouetted images of cloud and precipitation particles. In MC3E, the two-dimensional cloud probe (2DC) and the high-volume precipitation spectrometer, version 3 (HVPS3), were used for \( N(D_r) \) for size ranges between 175 \( \mu \)m and 1 mm and between 1 mm and 3 cm, respectively. In GCPEX, the cloud imaging probe (CIP) and the HVPS3 were used, while for OLYMPEX the two-dimensional stereo probe (2DS) and the HVPS3 were used for the same size ranges as MC3E, respectively. All OAP data were processed using the University of Illinois–University of Oklahoma Optical array Probe Software (UOOPS; McFarquhar et al. 2017; Jackson et al. 2014) to remove scattered artifacts and reconstruct both hollow images and images of particles whose edges touched one of the sides of the photodiode array. The 1-s PSDs were then averaged to 10 s to allow for better sampling statistics of large particles (McFarquhar et al. 2007) and to have similar horizontal spatial scales to that of the airborne radar.
To evaluate the trained retrieval, dual-frequency radar measurements collected by the Airborne Precipitation Radar (APR), versions 2 and 3, were used. The APR is a scanning radar that collects beam matched measurements of Ze at 13.4 and 35.6 GHz (Ku and Ka band) from nadir through 24 scans (Sadowy et al. 2003; Durden et al. 2019). The nominal vertical resolution of the radar is 30 m, while the along-track resolution is approximately 1 km. The APR was flown on NASA’s DC8 aircraft, which flew mostly constant altitude flight legs above precipitation echoes. The APR, version 2, was used in GCPEX and had a nominal sensitivity of 0 and −20 dBZ for the Ku band and Ka band, respectively, while the APR, version 3, was used in OLYMPEX with a sensitivity of 10 dBZ and −20 dBZ for Ku band and Ka band, respectively. The only difference between versions 2 and 3 was the addition of 94 GHz (W band) measurements in the APR, version 3. Since this work is GPM-DPR centric and attenuation from snowfall at W band is nontrivial, up to 1 dB km⁻¹ whereas at Ku and Ka band are estimated to be around 0.1 dB km⁻¹ or less (Kneifel et al. 2011), only the Ku- and Ka-band reflectivities are used.

To ensure correct absolute calibration of the radars, the Ku-band radar is calibrated by considering surface echoes of a water body in nonprecipitating conditions (GCPEX: Lake Huron and Lake Ontario; OLYMPEX: Pacific Ocean; Tanelli et al. 2006). Then the Ka band is calibrated against the Ku band by considering low reflectivity regions of the echoes where there is likely scattering in the Rayleigh regime (Durden et al. 2019). The uncertainty in this calibration is estimated to be approximately 1 dB for Ku and Ka band.

Several steps of processing of the APR data are required before its use in the retrieval. While attenuation from O₂ and H₂O vapor is small, they are corrected for using the gaspl package in MATLAB (Radiocommunication Sector of International Telecommunication Union 2013; https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.676-10-201309-S!!PDF-E.pdf) and a thermodynamic sounding collected near the time of the radar data collection. The mean values for two-way correction from gaseous attenuation is 0.15 and 0.6 dB for Ku and Ka band, respectively at the surface. Since the focus of this analysis is on only solid-phase hydrometeors, liquid-phase echoes, melting echoes (i.e., the bright band), surface echoes, and the radar returns from the in situ aircraft itself were all removed prior to running the retrieval. Liquid-phase echoes where determined as the echoes found at altitudes lower than the melting level.
that was determined by considering the peak in linear depolarization ratio (LDR). Furthermore, echoes found near or below the minimum sensitivity of the Ku-band were removed (10 and 0 dBZ for OLYMPEX and GCPEX, respectively).

To execute the NN retrieval, a \( T \) profile is required. For the APR, the \( T \) and altitude measured by the University of North Dakota (UND) Citation are used from each mission day to construct a mean temperature profile. This profile is then linearly interpolated to the APR vertical resolution. The Citation \( T \) is used as opposed to a radiosonde observation because it flew in a region more representative temporally and spatially of the environment sampled by the radar. A sensitivity test using the observed sounding as opposed to the Citation derived sounding resulted in an absolute mean percent difference of 5%, 20%, and 14% for \( D_{\text{ml}}, N_{\text{wl}}, \) and \( \text{IWC} \), respectively.

d. Collocation of in situ and radar measurements

To quantitatively evaluate the output of the NN retrieval, collocated in situ and radar data are required. Collocated points are identified following a technique similar to Chase et al. (2018) and Ding et al. (2020), who used a kd-tree searching algorithm from Scipy (Oliphant 2007) to efficiently search the APR sample volume for 30 of the closest gates within 1 km of the in situ aircraft location. One difference from Chase et al. (2018), where the weighted mean of the 30 closest gates were used, is that the closest gate is chosen from the 30 closest located by the kd-tree algorithm. The closest gate method was chosen to prevent auto correlations in the radar data from influencing the performance statistics. Testing the sensitivity between the 30-gate average used in Chase et al. (2018) and the closest gate used here results in a \( Z_r \) change of less than the calibration uncertainty (<1 dB). The average spatial error of collocation is approximately 370 m. Another difference from Chase et al. (2018) is that observations were considered coincident for this study when they were collected within 5 min temporally, while Chase et al. (2018) used a 10-min temporal threshold.

e. PSD parameter reference from in situ

While an independent measure of bulk water content within clouds was made during GCPEX and OLYMPEX, previous studies have suggested the Nevzorov probe, used in both field campaigns, underestimates mass in comparison with other bulk water content probes (e.g., Korolev et al. 2013; Abel et al. 2014). Since an estimate of particle mass is required to calculate \( D_{\text{ml}}, D_{\text{ms}}, N_{\text{wl}}, \) and \( \text{IWC} \), a parameterization given by

\[
m = aD^b \tag{9}
\]

is used, where \( a \) and \( b \) can vary depending on the conditions where the measurements were obtained or on the probes used to collect the data (Finlon et al. 2019). To constrain the choice of \( a \) and \( b \), the collocated \( Z_r \) from the APR and the \( Z_r \) calculated from the PSD are compared. Specifically, \( Z_r \) is calculated from the PSD using Eq. (2) and the \( \sigma_{\text{bsc}} \) from 6 different degrees of riming categories reported in Leinonen and Szyrmer (2015). Then, the \( a \) and \( b \) parameters associated with the \( \sigma_{\text{bsc}} \) that provides least error between measured and calculated \( Z_r \) at both Ku and Ka band are chosen. From there, \( D_{\text{ml}}, D_{\text{ms}}, N_{\text{wl}}, \) and \( \text{IWC} \) are all calculated using the selected \( a \) and \( b \).

Table 1. RMSE of the retrieved IWC (g m\(^{-3}\)) using the power-law fits on the synthetic test dataset. Sassen (1987) and Liu and Illingworth (2000) only had Ka-band fits. For convenience, the \( \alpha \) and \( \beta \) parameters of each of the power laws are included in the table. Ku-band parameters are in parentheses, and Ka-band parameters are in square brackets.

<table>
<thead>
<tr>
<th></th>
<th>Ku band</th>
<th>Ka band</th>
<th>( (\alpha_{\text{ku}}, \beta_{\text{ku}}); [\alpha_{\text{ka}}, \beta_{\text{ka}}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDA/GMM</td>
<td>0.29</td>
<td>0.25</td>
<td>(0.024, 0.570); [0.026, 0.638]</td>
</tr>
<tr>
<td>2ADPR</td>
<td>0.47</td>
<td>0.47</td>
<td>(0.003, 0.469); [0.002, 0.548]</td>
</tr>
<tr>
<td>Sassen (1987)</td>
<td>—</td>
<td>0.53</td>
<td>(—, —); [0.037, 0.7]</td>
</tr>
<tr>
<td>Liu and Illingworth (2000)</td>
<td>—</td>
<td>0.87</td>
<td>(—, —); [0.097, 0.59]</td>
</tr>
</tbody>
</table>

FIG. 3. Joint distributions of \( Z_r - \text{IWC} \) for the synthetic database generated from the DDA/GMM particles and measured PSDs. (a) Ku-band and (b) Ka-band results. The solid red line is the new power-law fit to the data plotted, and the dashed red line is the version-6 GPM-DPR power-law fit (see appendix A). Additional yellow lines in (b) are two power-law fits from Liu and Illingworth (2000) (solid line) and Sassen (1987) (dashed line).
The average error of $Ze$ from this method is 0.20 and 1.3 dB for Ku and Ka band, respectively.

f. GPM-DPR observations

The current GPM-DPR retrieval process is described in the algorithm theoretical basis document (Iguchi et al. 2018). To compare the current GPM-DPR retrieval of $Nwl$ and $Dml$ to the NN retrieval developed herein, the NN retrieval is applied directly to the GPM-DPR data, which are the version-6 level-2 (2ADPR) data product. As with the APR data, an estimate of $T$ at each radar gate is required. Thus, to allow for consistency between the NN retrieval and the GPM-DPR retrieval, the NN when applied to the GPM-DPR data uses the same $T$ data used in the GPM-DPR algorithm that are provided by the Japan Meteorological Agency numerical weather prediction model.

g. Snowfall properties database

The process of creating the training and test dataset for the NN is described here and is shown graphically in Fig. 2. The PSD dataset is split into training and testing groups by randomly selecting 333 PSDs (approximately 10% of the PSD data) without replacement from each field campaign and labeling them as the test dataset. The remaining 15,487 PSDs are labeled as the training dataset. Then the datasets are upsampled by randomly sampling with replacement. For the training dataset, 33,333 PSDs from each campaign are randomly sampled, whereas for the test dataset 3333 PSDs are sampled, providing a 90%/10% split between the two datasets. For each bin characterizing a PSD, one DDA-GMM particle is randomly sampled with replacement that has a maximum dimension $D_p$ such that $D_{i} - \langle D_D \rangle / 2 \leq D_p \leq D_{i} + \langle D_D \rangle / 2$. This creates a random distribution of simulated particle types that match the prescribed dimensions of $D_{i}$ and subsequently $N(D_{i})$. Once a random collection of particles is assigned to each PSD, $Ze$, $Dml$, and $Nwl$, and IWC are calculated using the particle type information provided by the DDA-GMM simulations and the equations shown in section 2a. Note that, although there is no additional information gained from upscaling the PSD data, there are over 20,000 particles that create numerous unique combinations of particles and PSDs. The unique combinations provide a wider variety of plausible microphysical/microwave properties, which should allow the NN to become more generalized.

To provide context of $Ze$ and IWC that have been synthetically generated, the joint distribution of $Ze$ and IWC is shown in Fig. 3. Additional information of the range of different radar scattering properties of the training database can be found in appendix B and Fig. B1. It is not surprising that the range of IWC calculated is larger than that measured by the bulk water probes used in MC3E, OLYMPEX and GCPEX (e.g., Nevzorov) because the probes are known to underestimate IWC. Thus, the data are kept for training in order to provide a spectrum of plausible values to the NN.

A new power law between $Ze$ and IWC (solid red line in Fig. 3) of the form

$$IWC = aZ_e^b$$

was obtained from the training dataset, and the coefficients are listed in Table 1. This relationship provides a baseline retrieval to compare with the NN. Herein this will be referred to as the legacy method retrieval. A comparison of the new legacy method with previous relations (e.g., Sassen 1987; Liu and Illingworth 2000) shows that the new fit generally estimates lower IWC for the same $Ze$ by approximately 300% and 50%
relative to the Liu and Illingworth (2000) and Sassen (1987) relations, respectively. For a second comparison, an average relation of the form in Eq. (10) is obtained from output of the current GPM-DPR algorithm (see appendix A). The IWC obtained from the GPM-DPR relation are approximately a factor of 10 smaller than the IWC generated here. Since $S$ is proportional to IWC, it is consistent with Skofronick-Jackson et al.’s (2019) study showing that GPM-DPR is underestimating $S$. To quantify the error in retrieved IWC for all retrievals, the root-mean-square error (RMSE),

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

is used, where $x_i$ is the vector of retrieved values, $\bar{x}$ is the vector of true values, and $n$ is the number of data samples in the test set. The RMSE for each legacy method calculated on the test dataset is shown in Table 1.

### h. NN details

The type of NN used here is a fully connected feed forward multilayered perceptron NN. The network has three input features, six hidden layers of eight neurons and three output labels (Fig. 4). The input features were chosen based on what is available operationally from GPM-DPR, namely the Ku-band $Z_e$ in logarithmic units (dBZ), the DFRKu–Ka in logarithmic units (dB) and the $T$ in degrees Celsius. The structure of the network was determined by systematically retraining the network with 2–128 neurons and 2–10 layers and choosing the network with the least amount of error on the test dataset before showing a signal of overfitting. All networks were trained on graphical processing units (GPUs) provided by Google’s freely available computation platform (Google Colaboratory) and the open source software package Tensorflow (Abadi et al. 2016).

Before training, the data require a transformation to prevent weighting any specific input feature unfairly based on its absolute magnitude and range. Thus, all input features were scaled to have a mean of 0 and a variance of 1. In addition, taking the logarithm of the output labels and scaling them to have mean 0 and variance 1 provided the least RMSE. All parameters described above, as well as a few other specific details, are noted in Table 2. In the event the reader would like to use the trained NN, an example of loading and running the network is shown in a Jupyter notebook with the associated data included with this paper (see data availability statement).

### 3. Results

#### a. Evaluation of the retrieval methods on the synthetic dataset

Any empirical method can be overfit intentionally to provide optimal results on a training dataset. Thus, here the NN retrieval is evaluated on the test dataset. The median percent error (MPE), defined as

$$\text{MPE} = \frac{\text{median} \left( \frac{y_i - \bar{y}}{y_i} \right) \times 100}{100}$$

where $y_i$ is the vector of retrieved values and $\bar{y}$ is the vector of true values, is chosen as the metric of evaluation. Figure 5 shows the MPE as a function of the retrieved parameter of interest’s magnitude. For retrieved $D_{ml}$ and $D_{ms}$ there is little bias, with MPE around 0% and an interquartile range (IQR) on average of 6\% and 17\% for $D_{ml}$ and $D_{ms}$, respectively (Fig. 5a). Similarly, $N_w$ has MPE of 2\% but has much larger

<table>
<thead>
<tr>
<th>Parameter</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{ml}$</td>
<td>0.12 (0.41) mm</td>
</tr>
<tr>
<td>$N_w$</td>
<td>$1.27 \times 10^6$ m$^{-3}$ mm$^{-1}$</td>
</tr>
<tr>
<td>IWC</td>
<td>0.24 g m$^{-3}$</td>
</tr>
</tbody>
</table>
average IQR than does $D_m$, with the mean first quartile being $-51\%$ and the mean third quartile being $+114\%$.

Using Eq. (8), IWC can be calculated so that the NN retrieval can be compared directly to the legacy methods. The NN (red in Fig. 5c) performs best, showing a MPE around $+1\%$, and a mean first quartile of $-25\%$ and a mean third quartile of $+43\%$. As the magnitude of the retrieved IWC increases, the IQR decreases, implying that there is less relative uncertainty when the retrieval is retrieving IWC $\geq 0.5 \text{ g m}^{-2}$. Meanwhile, the legacy fit provided in section 2g (blue in Fig. 5c) provides the second-best method. There is a clear high bias of $+35\%$ at IWC $\leq 0.01 \text{ g m}^{-2}$ and of $+61\%$ IWC $\geq 1.0 \text{ g m}^{-2}$, while underestimating between $0.01 \text{ g m}^{-2} < \text{IWC} < 1.0 \text{ g m}^{-2}$. The legacy fit also shows considerably more uncertainty than the NN, showing a larger IQR. The estimation of the current GPM-DPR algorithm (yellow in Fig. 5c) shows a constant underestimation for all IWC of approximately $86\%$.

Comparing the RMSE of the NN retrieved IWC (Table 3) with the RMSE of the legacy methods (Table 1), the NN
outperforms all other legacy methods. To test statistical significance, a two-sample Student’s t test is used to compare the square error of the legacy methods to the square error of the NN retrieval. The result shows that the NN has a significantly ($p < 0.05$) lower square error than all methods except the Ka-band legacy power-law fit. Thus, the NN retrieval provides estimates of IWC that are equal to or better than legacy method, while additionally retrieving two parameters of the PSD.

### b. Case studies

While the evaluation of the NN on the test dataset provides an initial quantification of uncertainty and errors, the test dataset may not be truly independent from the training dataset. The PSDs used to synthesize the training and test dataset were measured in close spatial and temporal proximity to each other and thus are potentially correlated. The aforementioned implicit correlation could have skewed the evaluation done in section 3a. Furthermore, the stochastic nature of particle selection in the synthetic database could be unphysical, leading to training and evaluation of the NN on potentially unphysical Ze, Dml, and Nml triplets. Thus, to gain more confidence in the NN retrieval and more physical insight, three case studies are used to further evaluate the retrieval.

#### 1) 31 JANUARY 2012

The first case study is a synoptic snowfall event from GCPEX that occurred on 31 January 2012 (Fig. 6). This case represents a typical continental cyclone, with warm advection leading to widespread light snowfall over the GCPEX domain. Both aircraft flew a coordinated flight around 0016 UTC with the Citation flying an oval pattern northwest of the King City (Ontario, Canada) radar and the DC-8, carrying the APR, flying straight overt the Citation (Fig. 6a). The sounding from the start of the mission (2218 UTC 30 January 2012) shows temperatures from the surface to 750 hPa are between $-5^\circ$ and $-10^\circ$C, with the dendritic growth zone located between 700 and 600 hPa (Fig. 6b). The King City ground-based C-band radar plan position indicator (PPI) scan shows widespread snow, with localized bands of higher Zr (Fig. 6a).

Considering the cross section at near nadir of the swath shown in Fig. 6a, meteorological radar echoes are found up to the altitude of the DC-8 (8 km; Fig. 7a). However, echoes that are sufficiently above the minimum sensitivity of Ku-band radar only extend to 4 km. A clear fall-streak echo pattern shows up in the Ku and Ka band as well as the DFFKu-Ka (Fig. 7c) between 20 and 60 km horizontally and between 0 and 3 km vertically. The results of applying the NN retrieval to the APR data are shown in Figs. 7d-f. Only Dml is shown because Dml is highly correlated to Dml. Within the fall streak, Dml is 0.75–1.00 mm, while outside the fall streak it is $\leq 0.5$ mm (Fig. 7d) consistent with the increase in particle size in fall streaks noted by Plummer et al. (2015). Inversely, Nml within the fall streak is lower than the surrounding cloud (Fig. 7e). Calculating IWC shows peak values found within the fall streak of approximately 0.1 g m$^{-3}$.

The retrieval data matched to the in situ plane are shown in Fig. 8. Both the retrieval and in situ measurements of Dml have good agreement showing an MPE of less than 15%. The retrieval of Nml performs worse than that of the Dml, showing a constant low bias of approximately 0.5 in units of log(m$^{-3}$ mm$^{-1}$). This low bias propagates into the calculation of IWC that is biased low

### Table 4. RMSE and MPE (in parentheses) from the 31 Jan 2012 case study. Asterisks indicate that the exponential relations do not explicitly predict any other parameter except IWC.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>NN</th>
<th>DDA/GMM Ku, Ka</th>
<th>2ADPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dml, Dms</td>
<td>0.08, 0.25 mm; (+11%, −7%)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Nml</td>
<td>$1.87 \times 10^4$ m$^{-3}$ mm$^{-1}$; (−61%)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>IWC</td>
<td>0.05 g m$^{-3}$; (−28%)</td>
<td>0.06, 0.05 g m$^{-3}$; (−43%, −34%)</td>
<td>0.13 g m$^{-3}$; (−94%)</td>
</tr>
</tbody>
</table>
by approximately 0.05 g m\(^{-3}\). The RMSE and MPE between the retrieved parameters and the in situ estimates are found in Table 4. For this relatively uniform low IWC case, the NN retrieval performs better than suggested in Table 3.

2) 12 FEBRUARY 2012

The second case analyzed evaluates the NN retrieval on a shallow convective snowfall event (i.e., lake effect snow) from 12 February 2012 during GCPEX. The PPI scans from the King City radar show a narrow band of snowfall emanating from Georgian Bay and impacting the local Southern Ontario region (Fig. 9a). A radiosonde observation taken 40 min after the flights in Fig. 9b, show cold 850 hPa temperatures at about \(-15^\circ\)C providing ample conditional instability for lake-effect snow given an average lake surface temperature of Lake Huron of 2°C. The DC-8 and Citation flew multiple legs between Georgian Bay and the King City radar site, transecting the lake effect snowbands.

The cross section shown from 12 February 2012 shows echo tops much shallower than 31 January 2012, peaking at approximately 2.5–3 km MSL (Fig. 10). Individual convective elements are seen in the \(Z_e\) field, diagnosed from the pockets of \(Z_e \approx 30\) dBZ (Fig. 10a), which contained observed Doppler velocities of 1.0–1.5 m s\(^{-1}\) upward (not shown). There are no clear fall streaks in this case, but DFR\(_{K_u-K_a}\) is locally increased to \(>5\) dB at 40–60 km (Fig. 10c). The retrieved \(D_m\) is increased in the larger DFR\(_{K_u-K_a}\) regions to 1–1.5 mm (Fig. 10d) while being near 0.5 mm outside. The same larger DFR\(_{K_u-K_a}\) regions...
are characteristic of $N_w$ around 3.5 $\log(\text{m}^{-3}\text{mm}^{-1})$ and IWC around 1 $\text{g m}^{-2}$.

In comparing the retrieved $D_m$ with in situ data, it is seen that there is an initial poor agreement that appears along the flight track, with the retrieval overestimating the measured values by 0.75 mm, but then agreement improves along the track by 35 km (Fig. 11a). This could be due to the poor temporal matchup between the APR and the Citation (5 mins at the start) and the large spatial heterogeneity. Interestingly, unlike the previous case where the performance of the retrieved $D_m$ was similar to the performance of the $D_m$, retrieved $D_m$ shows a constant low bias of a factor of 2 (Fig. 11a) at the same location where the retrieval of $D_m$ is performing well (with MPE of 8%). The retrievals of $N_w$ and IWC show the same poor agreement at the start, but then agree within 33%-43% (MPE) at distances greater than 35 km. Although the $N_w$ retrieval has less error (by 13%) relative to the 31 January 2012 case, the retrieval of IWC is still low when compared with the in situ measurement but is performing better than the other legacy methods. The retrieved RMSE of parameters for 12 February 2012 are worse than the RMSE on the synthetic data and the previous case, but the NN retrieval outperforms all legacy methods in retrieving IWC with a RMSE of 0.27 $\text{g m}^{-2}$ (Table 5). It should be noted that during the 12 February 2012 case, the Citation was detecting the presence of supercooled liquid water from the Rosemount icing detector probe (Baumgardner and Rodi 1989), which implies that rimed particle types are plausible and could partially explain why the retrieval performs worse on the 12 February 2012 case than the 31 January 2012 case and the synthetic dataset.

3) 3 DECEMBER 2015

The last case presented herein considers a synoptically forced event that occurred over complex topography on 3 December 2015 during OLYMPEX. The 1.45° PPI scan from the National Weather Service Langley Hill (Washington) S-band radar shows widespread precipitation echoes, with enhancements of reflectivity over the terrain located NE of the radar location (Fig. 12a). A radiosonde from approximately the same time and location as the ground-based radar scan shows a moist environment, with the melting level located at approximately 750 hPa (or 2 km MSL). Both airplanes were flying a coordinated stacked leg pattern from northwest to southeast over the Olympic Mountains (dashed and solid line in Fig. 12a).

The cross section of raw $Z_e$ and the retrieved products are shown in Fig. 13. Echo tops at Ku and Ka band on 3 December 2015 extend up to about 8 km MSL, with an apparent melting level at 2 km (i.e., bright band; Figs. 13a, b). Interestingly, there is also a secondary level of enhancement of $Z_e$ found at 3.5–4.5 km that shows up well in DFRKu–Ka (Fig. 13c). The secondary enhancement is characteristic of larger retrieved $D_m$, lower $N_w$, and larger IWC than the immediate surrounding areas. This region of enhancement is likely associated with a mountain wave, forced by a stable layer lifted by synoptic flow over the terrain, leading to locally increased upward vertical velocities within the layer. Doppler velocities within the second layer of enhancement showed pockets of 1–2 m s$^{-1}$ updrafts (not shown). Considering the in situ measured temperature and the sounding (Fig. 12b), the secondary enhancement is close to the dendritic growth zone, which could be leading to quickly growing particles by vapor deposition and aggregation and thus the enhanced $Z_e$ and DFRKu–Ka. Below the layer of enhancement, retrieved IWC increases above 1 $\text{g m}^{-2}$.

There is good agreement between the retrieval and in situ observations initially, with MPE of −8%, +9%, and −8% for $D_m$, $N_w$, and IWC, respectively, up to 65 km. After 65 km, the retrieval underestimates $D_m$ and $N_w$ (−17% and −28% MPE, respectively; Fig. 14a), and overestimates $N_w$ (+163% MPE; Fig. 14b) leading to an overestimate of IWC (+31% MPE; Fig. 14e). The RMSE of all parameters shows less error than 12 February 2012.

![Fig. 11. As in Fig. 8, but for 12 Feb 2012.](image-url)
but larger errors than the case on 31 January 2012 (Table 6). As in the 12 February 2012 case, there is supercooled liquid water sampled by the Citation aircraft on 3 December 2015, which was noted by Chase et al. (2018) and could be leading to the observed errors in the NN retrieval. Despite the potential presence of rimed particles, the retrieval of IWC shows similar RMSE (0.16 g m\(^{-3}\)) to the legacy methods (0.22 and 0.17 g m\(^{-3}\)) and improved results relative to GPM-DPR average relation (0.42 g m\(^{-3}\)).

4) SUMMARY OF CASES

Three case studies representing different meteorological conditions were analyzed to compare the retrieved \(D_{ml}, D_{ms}, N_{wl},\) and IWC with the in situ measured value. Combining all three, the case studies show that the NN performs similarly to what was described in section 3a (Table 3), with statistically significant (\(p < 0.05\)) lower square error on IWC when compared with an estimate of the current GPM-DPR algorithm and a legacy power-law fit using Ku-band \(Z_e\). The resulting combined MPE is \(-13\%, -23\%, +120\%,\) and \(+10\%\) for \(D_{ml}, D_{ms}, N_{wl},\) and IWC, respectively. It is important to emphasize that for cases 2 and 3 the NN is performing better than legacy methods despite the presence of potentially rimed particles being sampled in situ. Thus, while the formulation of the training dataset is somewhat simple (i.e., no rimed particles), it is still general enough to provide improved retrieval results

![Fig. 12](image12.png)

**Fig. 12.** As in Fig. 6, but for 1509 UTC 3 Dec 2015 and the Langley Hill NEXRAD radar (KLGX). (a) A second set of lines (dotted) indicates the GPM-DPR inner swath where Ku- and Ka-band observations are made. (b) Radiosonde observation taken at 1516 UTC 3 Dec 2015 from near the marked radar location in (a).

![Fig. 13](image13.png)

**Fig. 13.** As in Fig. 7, but for 3 Dec 2015.
relative to a single-frequency power law and an estimate of the current GPM-DPR algorithm.

4. Neural network implementation on GPM-DPR data

One unique aspect of the 3 December 2015 case was that the GPM-DPR overpassed the region during the coordinated flights of the DC-8 and Citation (GPM-DPR orbit track in dotted lines on Fig. 12a). Thus, the GPM-DPR observations and the direct output of the operational algorithm can be compared against the APR data and the output from the NN retrieval. Since the sensitivity of the GPM-DPR is approximately 12 and 18 dBZ for the Ku- and Ka-band normal scan, respectively (Toyoshima et al. 2015), the first noticeable difference between the APR (Fig. 13) and GPM-DPR (Fig. 15) is in echo-top height. In the GPM-DPR data, the echo tops are found at about 5–6 km at Ku band (Fig. 15a) and 4–5 km at Ka band (Fig. 15b) while the APR showed echo tops to about 8 km (Figs. 13a,b). The decrease in along-track resolution and vertical resolution can also be gleaned by comparing Figs. 13a–c and 15a–c. Despite the differences, GPM-DPR measures similar enhancements in DFRKu–Ka at 4 km altitude and 25–100 km horizontally and from 100 to 125 km horizontally.

Note the current operational GPM-DPR algorithm has an along-track ray-to-ray variability that is consistent throughout the vertical column in the retrieved products (Figs. 15d–f). This likely occurs because the profiles of $Z_e$ are attenuation corrected by considering the surface echo as reference, and this value is used in the microphysical solver to estimate retrieved quantities such as $N_{wl}$ and $D_{ms}$ (Iguchi et al. 2018, see their section 3g). Since these observations were collected over complex topography, the estimation of path integrated attenuation using the surface reference technique likely contains large errors resulting in the unphysical retrieval result. Despite this, there does seem to be a local enhancement of $D_{ms}$ at 4 km between 30 and 50 km (Fig. 15d), which is consistent with the NN retrieval applied to the APR data (Fig. 13d). Looking beyond the ray-to-ray instability, the overall magnitude of $N_{wl}$ is a factor of 10 lower than what was retrieved by the NN on the APR data, which through Eq. (8) leads to a factor-of-10 lower IWC retrieved by GPM-DPR. This is consistent with the low bias reported by Casella et al. (2017), Heymsfield et al. (2018), Skofronick-Jackson et al. (2019), and Chase et al. (2020) and with the results shown in Figs. 8c, 11c, and 14c.

While it has been shown here that the NN retrieval performs better than the average GPM-DPR retrieval of IWC (section 3) using the APR data, the uncertainties caused by the resolution differences, radar sensitivity differences, and the source of environmental temperature information could impact the retrieval when applied to the GPM-DPR data (e.g., Pfitzenmaier et al. 2019). To investigate how the radar differences impact the retrieval, the NN is applied directly to the GPM-DPR data on the 3 December 2015 (Fig. 16). The first noticeable improvement is the correction of the ray-to-ray variability in retrieved parameters. As a result, the enhancement of $D_{ms}$ at 4 km becomes more pronounced, showing up continuously in along-track scans (Fig. 16a). The second improvement is the retrieval of $N_{wl}$ at 4 km, which has similar magnitude [approximately $5 \log (m^{-3} \text{mm}^{-1})$] as in the APR retrieval (Fig. 13e), leading to similar retrieved IWC values.

To directly compare the results from sections 3 and 4, the median profiles of $D_{ms}$, $N_{wl}$, and IWC for each retrieval are shown in Fig. 17. Note that the means are taken from all scans of the APR, which in areal coverage approximately equal the area of the GPM-DPR. Furthermore, the APR data were constrained to have the same minimum sensitivity of GPM-DPR, which is 12 and 18 dBZ for Ku and Ka band, respectively. Despite the resolution and sensitivity differences, the NN retrieval applied to the GPM-DPR data largely capture the same profile of the NN retrieval applied to the APR. The main differences that occur are that there is an overestimation (+0.03 mm) of $D_{ms}$ and an underestimation of $N_{wl}$ [$-0.04 \log (m^{-3} \text{mm}^{-1})$] between 2 and 4 km. Since the bias of the retrieval of $D_{ms}$ and $N_{wl}$ are of opposite in sign, they largely

<table>
<thead>
<tr>
<th>Parameter</th>
<th>NN</th>
<th>DDA/GMM Ku, Ka</th>
<th>2ADPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{ms}$, $D_{ms}$</td>
<td>0.17, 1.17 mm; (−13%, −25%)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$N_{wl}$</td>
<td>$5.56 \times 10^3$ m$^{-3}$ mm$^{-1}$; (+129%)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>IWC</td>
<td>0.16 g m$^{-3}$; (+13%)</td>
<td>0.22, 0.17 g m$^{-3}$; (+47%, +35%)</td>
<td>0.42 g m$^{-3}$; (−89%)</td>
</tr>
</tbody>
</table>

Fig. 14. As in Fig. 8, but for 3 Dec 2015.
offset in the calculated IWC. The median profiles from the current GPM-DPR algorithm show a constant high bias of $D_{ml} (0.35 \text{ mm})$, low bias of $N_{w, l} [2.1 \log (m^{-2} \text{ mm}^{-1})]$, and thus a low bias of IWC ($0.5 \text{ g m}^{-2}$). Thus, the NN applied to the GPM-DPR data can provide physically plausible results that are improved relative to the current GPM-DPR retrieval for solid-phase hydrometeor properties.

5. Summary and conclusions

An approach using neural networks (NNs) to retrieve snowfall properties was formulated with a database of more than 20,000 ice particles whose microwave scattering properties were simulated using the discrete dipole approximation and the generalized multiparticle Mie method. The pool of particles was paired with observed particle size distributions (PSDs) measured within ice clouds during NASA Ground Validation field campaigns to produce a synthetic database of snowfall properties and their effective radar reflectivity factor ($Z_e$) at two frequencies (Ku and Ka band). The synthetic database was used to train an NN to retrieve two parameters of the 3-parameter gamma particle size distribution: the liquid equivalent mass-weighted mean diameter $D_{ml}$ and the liquid equivalent normalized intercept parameter $N_{w, l}$, from which the ice water content (IWC) can be calculated following Eq. (8). An evaluation on a subset of the synthetic database not used in training showed the NN retrieval has an RMSE of 0.1 mm, $1.28 \times 10^6 \text{ mm}^{-2}$, and 0.24 g m$^{-2}$ with an average range of uncertainty of $[-10\%, +7\%]$, $[-51\%, +114\%]$, and $[-25\%, +43\%]$ for $D_{ml}$, $N_{w, l}$, and IWC, respectively. Furthermore, three simple power-law relations between $Z_e$ and IWC evaluated on the same test dataset show significantly ($p < 0.05$) worse square error on the IWC retrieval relative to the NN.

Three case studies from NASA GV field campaigns provided an independent evaluation of the NN retrieval on coincident observations of $Z_e$ and the PSD. The first case, collected on 31 January 2012, showed that the NN had a median percentage error (MPE) between the retrieval and the in situ

![Figure 15](image1.png)

**FIG. 15.** Similar to Fig. 13, but now using GPM-DPR data along the APR3 swath (scan 9 of the inner swath). (a) Measured Ku-band $Z_e$. (b) Measured Ka-band $Z_e$. (c) Measured DFRKu–Ka. (d) GPM-DPR version-6 retrieved $D_{ml}$. (e) GPM-DPR version-6 retrieved $N_{w, l}$. (f) IWC calculated from the GPM-DPR version-6 retrieval of $D_{ml}$ and $N_{w, l}$.

![Figure 16](image2.png)

**FIG. 16.** Neural network retrieval applied to GPM-DPR data for (a) $D_{ml}$, (b) $N_{w, l}$, and (c) IWC.
estimates of +13%, −61%, and −28% for $D_{ml}$, $N_{wl}$, and IWC, respectively. The second case, a shallow lake effect snowfall event (12 February 2012), showed better performance with MPE of +8% and −48% for $D_{ml}$ and $N_{wl}$, respectively, but worse performance on IWC (MPE of −38%). The third case within a Pacific Northwest cyclone over complex topography (3 December 2015) showed MPE of −13%, +129%, and +13% for $D_{ml}$, $N_{wl}$, and IWC, respectively. Analysis for all three case studies showed that the NN has significantly lower square error when compared with an estimate of the current GPM-DPR algorithm and a legacy power-law fit between $Z_e$ and IWC. The NN retrieval was applied directly to GPM-DPR observations and indicated that the NN can provide a physically plausible profile, which is able to capture the same microphysical structures of the higher-resolution APR retrieval and correct for undesired retrieval artifacts found in the current GPM-DPR algorithm (e.g., ray-to-ray discontinuity).

This was the first attempt at providing a viable solid-phase retrieval alternative for GPM-DPR. Thus, the retrieval in its current form has caveats that readers should be aware of. Currently, there are no rimed particle types included in the training database of particles despite rimed particles being available from the literature (Leinonen and Szyrmer 2015). That being said, the NN retrieval continues to outperform a simple power-law and the GPM-DPR algorithm estimate on case studies where riming is likely present (see the 12 February 2012 and 3 December 2015 cases). In future iterations of the NN retrieval, rimed particle types should be included. Another caveat is that the training data for the NN were informed from three field campaigns all located in North America. Thus, if the goal is to have a global snowfall retrieval, future iterations of the retrieval should look to included additional field campaign measurements of PSDs collected in other precipitation regimes across the globe. A sensitivity analysis using different sources for the temperature input to the retrieval does affect the retrieval by around 10–25%. Thus, for optimal performance of this NN retrieval, users should use the most accurate available temperature input (e.g., Sounding in close proximity spatially and temporally).

Future avenues of research could modify the NN to also predict the shape parameter of the PSD ($\mu$); that way, other PSD characteristics, such as precipitation rate, can be derived. Future work should also look to evaluate the NN retrieval on additional case studies from OLYMPEX, GCPEX and other campaigns (e.g., Petijä et al. 2016; Lubin et al. 2020) with high-quality coincident multifrequency and microphysical measurements. The additional evaluation would further inform users of the accuracy and potential biases associated with the NN retrieval. Last, the comparison with CloudSat and the operational algorithms therein would be beneficial for understanding whether the bias reported in Skofronick-Jackson et al. (2019) and Casella et al. (2017) has been improved with the NN retrieval.

Acknowledgments. Funding for this research was provided to the University of Illinois by NASA Precipitation Measurement Missions Grant 80NSSC19K0713 and NASA Earth System Science Fellowship 80NSSC17K0439. We thank all of the participants of the field campaigns used here for their tireless effort in collecting the data used in this study. Furthermore, we also thank Google Colaboratory for their open-source free computing platform that allowed for the simple implementation of the neural network presented here. We also thank the three anonymous reviewers for the insightful comments that enhanced this paper.

Data availability statement. The amalgamation of DDA/GMM particles used in the formulation of the synthetic data, the synthetic data themselves, and the notebook used to train the neural network are found with the data repository associated with this paper (https://doi.org/10.13012/B2IDB-0791318_V2). The NASA GV field campaign data used in the three case studies can be made available upon request to the corresponding author, or the interested party can obtain all NASA GV data from NASA’s Global Hydrology Resource Center (https://ghrc.msfc.nasa.gov/home). Version 6 of the GPM-DPR data can be found online (https://doi.org/10.5067/GPM/DPR/GPM/2A/05).

APPENDIX A

Derivation of GPM-DPR estimate

Since the GPM-DPR operational algorithm is not trivial to implement and is not open source, a method to estimate the
GPM-DPR algorithm is described here. A legacy power-law fit of the form in Eq. (9) is formed from direct operational output of the level-2 2A.DPR files (https://doi.org/10.5067/GPM/DPR/GPM/2A/05). Specifically, four orbital files are chosen from 2014, 2015, 2016, 2017, and 2018, resulting in 20 total files. From there, the data are curated by selecting precipitating profiles (flagPrecip $= 11$) and where the near-surface temperature, determined from the 2A.ENV files, is less than 5°C. The temperature threshold is to prevent strong convective instances in the data where riming would be likely. From there, radar gates are chosen in which the gate temperature is less than 0°C to isolate solid-phase hydrometeors. Figure A1 shows randomly selected snowfall echoes from the 2A.DPR files. After the above conditions have been selected, the parameters in Eq. (10) are fit using the sklearn python package linear regression between the log(IWC) and the log($Z_e$). The $\alpha$ and $\beta$ values can be found in Table 1.

**Appendix B**

**Histograms of training data**

To provide perspective of the range of data used to train the neural network, all scattering properties are summarized in normalized histograms in Fig. B1. The left column contains all of the ranges of reflectivity at Ku (Fig. B1a), Ka (Fig. B1b), and W (Fig. B1c) band. The corresponding dual-frequency ratios between Ku and Ka (Fig. B1d), Ka and W (Fig. B1e), and Ku and W (Fig. B1f) are found in the right column.

**References**


