Water Resources Research

RESEARCH ARTICLE

10.1029/2018WR023830

Special Section:

Big Data & Machine Learning in
Water Sciences: Recent
Progress and Their Use in
Advancing Science

Key Points:

- Passive microwave precipitation was estimated by training deep neural networks to spaceborne radar precipitation at high latitudes
- Deep neural networks improve high-latitude rain and snow estimation compared to GPM's passive microwave algorithm
- Performance of various passive microwave and infrared channels and the contribution of environmental data were explored

Correspondence to:

D. Long and Y. Hong,
hongyang@tsinghua.edu.cn;
dlong@tsinghua.edu.cn

Citation:

Tang, G., Long, D., Behrangi, A., Wang, C., & Hong, Y. (2018). Exploring deep neural networks to retrieve rain and snow in high latitudes using multisensor and reanalysis data. *Water Resources Research*, 54, 8253–8278. <https://doi.org/10.1029/2018WR023830>

Received 7 AUG 2018

Accepted 4 OCT 2018

Accepted article online 11 OCT 2018

Published online 25 OCT 2018

Exploring Deep Neural Networks to Retrieve Rain and Snow in High Latitudes Using Multisensor and Reanalysis Data

Guoqiang Tang^{1,2} , Di Long¹ , Ali Behrangi^{2,3,4} , Cunguang Wang¹, and Yang Hong^{1,5}

¹State Key Laboratory of Hydrosience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, China, ²Joint Institute for Regional Earth System Science and Engineering and Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA, USA, ³Department of Hydrology and Atmospheric sciences, University of Arizona, Tucson, AZ, USA, ⁴Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, ⁵School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, OK, USA

Abstract Satellite remote sensing is able to provide information on global rain and snow, but challenges remain in accurate estimation of precipitation rates, particularly in snow retrieval. In this work, the deep neural network (DNN) is applied to estimate rain and snow rates in high latitudes. The reference data for DNN training are provided by two spaceborne radars onboard the Global Precipitation Measurement (GPM) Core Observatory and CloudSat. Passive microwave data from the GPM Microwave Imager (GMI), infrared data from MODerate resolution Imaging Spectroradiometer and environmental data from European Centre for Medium-Range Weather Forecasts are trained to the spaceborne radar-based reference precipitation. The DNN estimates are compared to data from the Goddard Profiling Algorithm (GPROF), which is used to retrieve passive microwave precipitation for the GPM mission. First, the DNN-based retrieval method performs well in both training and testing periods. Second, the DNN can reveal the advantages and disadvantages of different channels of GMI and MODerate resolution Imaging Spectroradiometer. Additionally, infrared and environmental data can improve precipitation estimation of the DNN, particularly for snowfall. Finally, based on the optimized DNN, rain and snow are estimated in 2017 from orbital GMI brightness temperatures and compared to ERA-Interim and Modern-Era Retrospective analysis for Research and Applications Version 2 reanalysis data. Evaluation results show that (1) the DNN can largely mitigate the underestimation of precipitation rates in high latitudes by GPROF; (2) the DNN-based snowfall estimates largely outperform those of GPROF; and (3) the spatial distributions of DNN-based precipitation are closer to reanalysis data. The method and assessment presented in this study could potentially contribute to the substantial improvement of satellite precipitation products in high latitudes.

Plain Language Summary Snow has a significant influence on the hydrological cycle and water resource availability. Compared to ground gauges and radars with limited coverage, satellite remote sensing can provide global rain and snow observations from space. However, traditional satellite precipitation retrieval methods are prone to errors in snow estimation at high latitudes. In this study, we developed a new rain and snow estimation method at high latitudes using deep neural networks. The reference data sets were from two spaceborne radars that provide the most direct precipitation observations from space. Passive microwave, infrared, and environmental data were trained to the reference data sets for rain and snow estimation. Results show that the neural network-based method can largely reduce the underestimation of rain and snow rates in high latitudes in many prior algorithms. Statistical indices for snow estimation are notably improved. Furthermore, combining data from passive microwave, infrared, and environmental data sets contribute to better precipitation estimation than a single source. We suggest that deep neural networks could potentially contribute to the improvement of satellite precipitation at high latitudes, which is valuable for expanding the spatial coverage of current satellite products.

1. Introduction

Precipitation is a critical component in the water and energy cycles (Kidd & Huffman, 2011). The phase of precipitation, which can be categorized into rain, snow, and a mix of both, has an important effect on hydrological processes (Harpold et al., 2017; Skofronick-Jackson et al., 2015). Rain has widely known effect on nature and society, whereas snow is the dominant precipitation type in high-latitude regions (Skofronick-Jackson et al., 2015). As the climate is getting warmer, less precipitation may fall as snow that may cause a profound impact on snowpack and streamflow (Bales et al., 2006), and subsequently on the water

availability and food security in regions where rivers are mainly fed by snow and ice melt (Chen et al., 2017, 2018; Immerzeel et al., 2010). Due to the limited coverage of meteorological stations, ocean buoys, and weather radars, satellite remote sensing is the main observational source to acquire global rain and snow data (Hong et al., 2007).

Globally gridded precipitation products rely mainly on data from microwave (MW) sensors onboard low Earth orbiting satellites and infrared (IR) sensors onboard geostationary (GEO) satellites. Compared to IR, MW provides more direct and accurate precipitation measurements because of its sensitivity to rain/ice particles, whereas IR-based algorithms are often built on empirical and indirect links between cloud top brightness temperature (T_b) and surface precipitation (Huffman et al., 2007; Kidd et al., 2003; Sorooshian et al., 2000). Active microwave sensors, that is, spaceborne radars, often provide more direct and accurate precipitation estimates than passive microwave (PMW) sensors (Behrangi et al., 2014; Gao et al., 2017; Tang et al., 2017; Tang, Long, et al., 2018). So far, there have been three popular spaceborne radars for observing precipitation, that is, the Ku-band precipitation radar onboard the Tropical Rainfall Measuring Mission (TRMM) satellite operating from 1997 to 2015, the W-band Cloud Profiling Radar (CPR) onboard the CloudSat operating from 2006 to the present, and the Dual-frequency Precipitation Radar (DPR) onboard the Global Precipitation Measurement (GPM) Core Observatory operating from 2014 to the present. The CPR is capable of observing snow due to its high frequency (Liu, 2008) with good performance for light snow but degraded performance for moderate to heavy snow due to signal attenuation (Cao et al. 2014). The DPR also has the capability of detecting snow, whereas related algorithms and products are still immature (Harpold et al., 2017; Skofronick-Jackson et al., 2015). Due to the high sensitivity of CPR to light rain and snow, it has been used individually or in combination with other sensors to quantify global oceanic rain (Behrangi et al., 2012), assess reanalysis rain and snow in high latitudes (Behrangi et al., 2016; Palermé et al., 2017), evaluate the performance of other satellite sensors (Behrangi et al., 2014; Tang et al., 2017), and to study the mass change on the Antarctic ice sheet (Boening et al., 2012).

Modern multisatellite precipitation products, such as the latest Integrated Multi-satellite Retrievals for GPM (IMERG; Huffman et al., 2017), TRMM Multi-satellite Precipitation Analysis (Huffman et al., 2007), and the Climate Prediction Center morphing technique (CMORPH; Joyce et al., 2004) have been widely applied in scientific and societal fields (Hou et al., 2014). These products take advantage of data combination between multiple MW/IR sensors and ground-based observations. IMERG is built on the advances in previous algorithms and produces Level-3 rain and snow estimates at 0.1° and 30-min resolutions within 60°S–N since 2014 (Huffman et al., 2017). Many recent studies have shown that IMERG products outperform previous products such as TRMM Multi-satellite Precipitation Analysis (Asong et al., 2017; Tang et al., 2016). IMERG adopts the latest Goddard Profiling Algorithm (GPROF V5; Kidd et al., 2016; Kummerow et al., 2001, 2015) to retrieve Level-2 precipitation for all PMW sensors in the GPM constellation. GPM Microwave Imager (GMI) and DPR carried by the GPM Core Observatory play a specific role in GPROF, providing the basis for the a priori database to constrain possible solutions of all radiometers.

While most satellite retrieval algorithms perform reasonably well for rain estimation, they are not satisfactory for snow estimation, particularly over snow-covered or iced surfaces (Noh et al., 2009). High-frequency (e.g., >80 GHz) bands of PMW can be used to detect snowfall based on the scattering effect of ice hydrometers (Liu & Seo, 2013), and several retrieval algorithms have been developed (Kongoli et al., 2003, 2015; Liu & Seo, 2013; Skofronick-Jackson et al., 2004). However, some factors limit the accuracy of snow estimation based on high-frequency PMW data. First, the emission of cloud liquid water could mask the scattering signature of ice hydrometers (Liu & Seo, 2013). Second, frozen land surfaces are characterized by increasing emissivity, which could obscure the depression of the microwave signal caused by scattering (Harpold et al., 2017; Yong et al., 2014). Another problem of satellite precipitation retrieval is the poor quality in high latitudes where there is a stronger response to climate change than other regions (Behrangi et al., 2016; Long et al., 2017). The large fractions of the snow-covered surface and frequent snowfall mainly result in deteriorated precipitation estimation from space, which poses considerable challenges to satellite-only precipitation products that target cold regions, including polar areas. In contrast, reanalysis precipitation data sets could perform better than satellite products in high latitudes (Beck et al., 2017; Serreze et al., 2005; Yong et al., 2014).

Recent development of deep learning techniques brings great opportunities for the scientific community of water sciences (Hamshaw et al., 2018). The utilization of deep learning methods is shown to be useful in

resolving some traditionally difficult problems. For example, Fang et al. (2017) developed a system that can predict Soil Moisture Active Passive soil moisture using the Long Short-Term Memory network. Zhang et al. (2018) applied the Long Short-Term Memory network to the simulation and prediction of the water level of the combined sewer overflow structure. Sahoo et al. (2017) utilized three-layer Multilayer Perceptron networks to modeling groundwater level changes in agricultural regions of the United States. Laloy et al. (2018) used a deep neural network (DNN) of the generative adversarial network type to address the problem of probabilistic inversion for complex geologic media. The DNN is also applied to IR-based rainfall estimation (Tao et al., 2016, 2017, 2018), gaining better occurrence detection and intensity estimation compared to Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System (PERSIANN-CCS) developed by Hong et al. (2004). Compared to the traditional shallow artificial neural network (ANN), the DNN is characterized by more hidden layers, providing stronger classification and regression capability.

ANN techniques have been applied to rain estimation for decades. For example, IR-based products such as PERSIANN (Hsu et al., 1997; Sorooshian et al., 2000), PERSIANN-CCS (Hong et al., 2004), and PERSIANN-Multispectral Analysis (MSA; Behrangi et al., 2009), all utilize neural network techniques to train GEO IR data for direct rainfall estimation. Among these products, precipitation provided by PERSIANN-CCS is used by the IMERG algorithm (Huffman et al., 2017) to fill the gap caused by insufficient sampling of PMW sensors. To acquire global IR-based precipitation estimates, GEO IR T_b data are trained using PMW precipitation estimates and various neural networks. However, one major limitation of these products is that PMW and IR data are not essentially integrated as a whole to estimate precipitation, and therefore, the products are mainly based on IR information which curbs further improvement of precipitation accuracy.

In summary, current satellite precipitation products have several common weaknesses including inaccurate snow estimates, low-quality precipitation retrieval in high latitudes, and imperfect merging between PMW and IR data. Deep learning techniques provide new choices to resolve these problems. By using the intersections/coincidences between GMI, DPR, and CPR provided by the 2B-CSATGPM (Turk, 2017), this study is designed to explore whether the DNN is able to mitigate the aforementioned problems. Some environmental variables and MODerate resolution Imaging Spectroradiometer (MODIS) IR data are also attached to 2B-CSATGPM, enabling us to explore the performance of different PMW and IR channels. It is valuable to experiment the extent to which the DNN framework can extract information from PMW and IR to produce a robust joint precipitation estimation tool in high latitudes. Considering the relatively high accuracy of DPR and CPR and their complementary roles, their precipitation data are merged to train PMW and IR T_b , which has not been performed in previous studies.

Therefore, the objectives of this study are twofold: (1) to explore the extent to which the DNN can develop a retrieval method by extracting information from various GMI and MODIS channels as well as environmental variables in high latitudes (defined as $>50^\circ\text{S/N}$ in this study) and (2) to apply the well-trained networks to estimate high-latitude rain and snow in an independent period and compare the results to GPROF, IMERG, and reanalysis products.

2. Data Sets

All products used in this study are summarized in Table 1. Detailed descriptions are given as follows.

2.1. CloudSat-GPM Coincidence Data Set

The 2B-CSATGPM (V03B) extracts coincident samples between CloudSat and the GPM Core Observatory within ± 15 -min time differences (Turk, 2017). Data from GMI, DPR, and CPR as well as some auxiliary data sets are included. In addition, the Microwave Humidity Sounder (MHS) and the Advanced Technology Microwave Sounder also have some coincident samples with CloudSat and the GPM Core Observatory which are included in 2B-CSATGPM. However, we only focus on GMI in this study because samples of MHS and Advanced Technology Microwave Sounder are relatively few.

This study utilized 2B-CSATGPM coincidences from 2014 to 2016 to train and test the DNN. The spatial distribution of the coincident sample numbers between CPR and GMI is shown in Figure 1. The study area was restricted to high-latitude regions between 50° and 70°S/N because the samples are inadequate at low latitudes and the GPM Core Observatory can only reach 70°S/N . The sample number around 60°N was much

Table 1
Summary of Brightness Temperature, Precipitation, and Environmental Data Sets Used in the Study

Data set	Resolution	Period	Description
Coincident data extracted from 2B-CSATGPM for model training			
1C-GMI (V05A)	15 × 9 km at 36.5 GHz ^a	Mar 2014 to Dec 2016	Microwave T_b
2A-GPROF-GMI (V05A)	15 × 9 km at 36.5 GHz ^a	Mar 2014 to Dec 2016	Rain and snow
2A-DPR (V05A)	~5 km	Mar 2014 to Dec 2016	Precipitation
2C-PRECIP-COLUMN (P2_R04_E06)	1.4 × 1.7 km	Mar 2014 to Dec 2016	Rain
2C-RAIN-PROFILE (P1rc1_R04_E06)	1.4 × 1.7 km	Mar 2014 to May 2016	Rain
2C-SNOW-PROFILE (P_R04_E06)	1.4 × 1.7 km	Mar 2014 to Dec 2016	Snow
MODIS-AUX	1.4 × 1.7 km	Mar 2014 to Dec 2016	Infrared T_b
ECMWF-AUX	1.4 × 1.7 km	Mar 2014 to Dec 2016	Environmental data
Precipitation data for independent validation			
ERA-Interim	0.75° × 0.75°/3 hr	Jan–Dec 2017	Rain and snow
MERRA2	0.5° × 0.625°/1 hr	Jan–Dec 2017	Rain and snow
IMERG Final run (V05B)	0.1° × 0.1°/0.5 hr	Jan–Dec 2017	Rain and snow

^aGMI resolutions vary with channels

higher than that around 60°S because CloudSat only collects data during the sunlit portion of the orbit since 2011 due to the battery anomaly. Although the problem may affect the regional performance of the DNN, rain and snow estimation based on the DNN was proved successful in this study. It should be noted that since the horizontal resolutions of GMI, DPR, and CPR are different, there are often more than one CPR or DPR pixel corresponding to a specific GMI pixel. Data for these CPR or DPR pixels are averaged to match the data for the corresponding GMI pixel. Details of GMI, DPR, and CPR data are introduced separately as follows.

2.1.1. GMI Brightness Temperature and Precipitation

GMI Level 1C T_b is provided by the product 1C-GPM (V05A) with 13 channels divided into two swaths. Swath S1 has nine channels, that is, 10V, 10H, 19V, 19H, 23V, 37V, 37H, 89V, and 89H (unit: GHz). Swath S2 has four channels, that is, 166V, 166H, 183 ± 3V, and 183 ± 8V (unit: GHz). Different channels have different spatial resolutions. Swath S2 has slightly different pixel centers with Swath S1 due to the effect of incidence angles. To address this issue, 2B-CSATGPM first locates the closest S1 pixels to each CPR beam location and then finds the nearest S2 pixels to those of S1 (Turk, 2017).

GMI Level 2A precipitation and frozen precipitation based on the GPROF algorithm are from 2A-GPROF-GMI (V05A). Snow is provided in the form of frozen precipitation. Rain is calculated by subtracting snow from total precipitation. The GPROF algorithm was developed in the 1990s for PMW precipitation estimation and has experienced numerous upgrades since then. The GPM mission uses the fully parametric GPROF 2014 with great improvement compared to the pseudo-parametric GPROF 2010 used in the TRMM era (Kummerow

Figure 1. Spatial distribution of the number of coincident samples between Cloud Profiling Radar and GPM Microwave Imager (GMI) at 1° resolution from 2014 to 2016. GPM = Global Precipitation Measurement.

et al., 2015). The GPROF builds a priori databases of precipitation profiles and the associated T_b signals based on model outputs or matchups between PMW sensors and observations from surface gauges or spaceborne radars. Precipitation estimation from both cross-track and conically scanning PMW sensors in the GPM constellation is realized in a unified framework based on a priori databases in conjunction with Bayesian inversion techniques.

In addition to the segmental GMI data included in 2B-CSATGPM, complete orbital data of GMI T_b and precipitation in 2017 were also downloaded for independent application and validation.

2.1.2. DPR Precipitation

DPR precipitation data are from 2A-DPR (V05A) with three scanning modes, that is, the Ku-band normal scan (NS), the Ka-band matched scan (MS), and the Ka-band high-sensitivity scan (HS; Iguchi et al., 2010). NS is adopted in this study because of its larger swath (245 km) compared to MS and HS and the stable quality due to the mature algorithm. The spatial resolution is around 5 km (Table 1). Compared to CPR, DPR shows less skill in estimating light precipitation and snow due to its lower sensitivity, but it is expected to provide more reasonable heavy precipitation estimates due to the smaller effect from signal attenuation (Tang et al., 2017).

2.1.3. CloudSat Precipitation and Auxiliary Data

CloudSat products used in this study include two rain data sets, that is, 2C-PRECIP-COLUMN (henceforth 2C-PRECIP, Version P2_R04_E06; Haynes et al., 2009), 2C-RAIN-PROFILE (henceforth 2C-RAIN, Version P1rc1_R04_E06; Lebsock & L'Ecuyer, 2011; L'Ecuyer & Stephens, 2002; Mitrescu et al., 2010), and one snow data set, that is, 2C-SNOW-PROFILE (henceforth 2C-SNOW, Version P_R04_E06; Wood et al., 2013; Wood et al., 2014). The 2C-PRECIP estimates precipitation intensity over oceans as well as global precipitation phase (rain, snow, or mixed) and likelihood (certain, probable, and possible). The 2C-RAIN calculates rain intensity when the likelihood flag is certain for rain and assigns zero values to rain rates if the melted fraction is less than 0.1. The 2C-SNOW calculates snow intensity when the phase flag is snow, mixed certain, or mixed possible if the melted fraction is under 0.1. The 2C-PRECIP and 2C-RAIN only provide rain estimates over oceans because land surface reflectivity is very complex, resulting in large uncertainties in rain rate retrieval. The 2C-SNOW provides snow estimates for both land and ocean areas. The spatial resolution of the three products is $1.4 \text{ km} \times 1.7 \text{ km}$.

Furthermore, CloudSat products are accompanied by two auxiliary data sets from MODIS and the European Centre for Medium-Range Weather Forecasts (ECMWF), that is, MODIS-AUX and ECMWF-AUX which are also contained in 2B-CSATGPM. MODIS provides IR data of middle infrared or thermal infrared channels 20 and 27–36 with wavelengths from 3 to 15 μm . MODIS channels 27–36 were used in this study because they can be used during day and night. ECMWF forecast analysis provides environmental variables such as 2-m air temperature (T_{2m}), vertical temperature profile, pressure, and specific humidity that are also helpful for precipitation retrieval (e.g., Behrangi et al., 2015).

2.2. Reanalysis Precipitation

Previous studies have shown that reanalyses have relatively good skill in estimating rain and snow rates in high latitudes (Beck et al., 2017; Serreze et al., 2005; Yong et al., 2014). Since high-quality ground precipitation observations with high temporal resolutions are sparse in high latitudes, we compared ERA-Interim (Dee et al., 2011) and Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2; Gelaro et al., 2017) precipitation in 2017 to precipitation estimates from the GPROF algorithm (V5) and the DNN (referring to section 3.3). ERA-Interim is at 3-hourly and 0.75° resolutions. MERRA2 is at hourly and $0.5^\circ \times 0.625^\circ$ resolutions. Both provide rain and snow estimates. The two products were resampled to a uniform resolution ($1^\circ \times 1^\circ$) using the bilinear interpolation method for comparison with GMI estimates, which will also be resampled to the same resolution.

2.3. IMERG Precipitation

IMERG (V05B) precipitation was also compared to GPROF and DNN estimates in 2017. IMERG produces quasi-global precipitation through two steps. First, GPROF-based PMW precipitation estimates from ~ 10 satellites in the GPM constellation are intercalibrated and morphed following the cloud motion vectors derived from GEO IR images which is used in the CMORPH algorithm (Joyce et al., 2004). Second, GEO-IR estimates are employed to fill the gap of morphed precipitation fields caused by the much sparser sampling of PMW sensors using a

Kalman filter method (Huffman et al., 2017). IMERG generates three types of products, that is, the near-real-time Early and Late runs, and the post-real-time Final run. The Final run is released about 2.5 months after the observation time because it merges gauge observations provided by the Global Precipitation Climatology Centre. We adopted the Final run product of IMERG due to its higher quality than the Early and Late runs. IMERG provides snow information in the form of the probability of liquid/solid precipitation. In this study, the snow rate is calculated by multiplying the solid probability with the precipitation rate. Another method is to choose a probability threshold to determine precipitation phase. For example, if the solid probability is lower than the threshold (e.g., 50%), the snow rate is set to be zero; otherwise the total precipitation is regarded as snow. We tested thresholds of 0%, 25%, 50%, 75%, and 100% and found that when the solid probability threshold was set to 50%, the spatial distribution of IMERG snow is very close to that derived from the multiplication method. Lower thresholds would result in too small snow, whereas higher thresholds would result in too large snow compared to ERA-Interim and MERRA2.

3. Methods

The framework of the study is shown in Figure 2. First, DNN input and reference data sets were produced by combining CloudSat, GPM, MODIS, and ECMWF data (section 3.1). Second, DNNs were trained from 2014 to 2016 based on various scenarios (section 3.2). Finally, high-latitude GMI rain and snow were estimated based on the best-performing DNN and the results were compared to ERA-Interim and MERRA2 data (section 3.3).

3.1. Training Data Sets

3.1.1. Inputs of Passive Microwave, Infrared, and Environmental Data

Inputs to the DNN comprise PMW data from GMI, IR data from MODIS, and environmental data from ECMWF. The combination of GMI and MODIS data is helpful to explore if cloud top temperature observations from IR data can provide complementary information to PMW observations that penetrate clouds. The answer to this question may recommend joint PMW-IR retrieval in the future. The inclusion of ECMWF data might also be helpful to improve the accuracy of precipitation estimates. Two environmental variables, that is, T_{2m} and total precipitable water (TPW) were employed in this study. TPW is not directly provided in 2B-CSATGPM but can be calculated from the vertical profiles of temperature, pressure, and specific humidity. The two variables were chosen as representative variables because they are also used in the GPROF algorithm and have a close relationship with precipitation rates and phases.

Various input scenarios were designed to explore the advantages and disadvantages of different GMI channels, and the effectiveness of combining PMW, IR, and environmental data (Group 1 in Table 2). These scenarios include (1) GMI1–5, (2) GMI6–9, (3) GMI10–13, (4) GMI, (5) GMI + MODIS, (6) GMI + ECMWF, and (7) GMI + MODIS+ECMWF. Only MODIS channels 31–32 were considered in Group 1 because their wavelengths (10.780–11.280 μm and 11.770–12.270 μm) are close to those of infrared imagers onboard GEO satellites.

In addition, we also tested the combinations between different MODIS channels and ECMWF environmental variables because IR-only precipitation is sometimes used to fill the gaps between PMW sensor overpasses or as a replacement over snow-covered areas where PMW retrievals are problematic. Moreover, because of the wide spectrum of GMI, some GMI channel combinations can imitate other PMW sensors and enable a preliminary investigation into their performance of precipitation estimation based on the DNN. These scenarios are referred to as Group 2 in Table 2, including (1) GMI1–9 which is similar to TMI channels, (2) GMI3–9 which is similar to Special Sensor Microwave Imager (SSM/I) channels, (3) MODIS (10–12 μm), (4) MODIS_E (6–14 μm), and (5) MODIS+ECMWF. The way of naming MODIS and MODIS_E is to simplify the usage in this paper. Given that the Geostationary Operational Environmental Satellite (GOES)–16 launched in November 2016 has a wider IR spectrum compared to previous GEO satellites, the scenario of MODIS_E helps explore the benefits brought by the additional IR channels of GOES–16.

The numbers of rain samples are ~57 and ~178 thousands over land and ocean areas, respectively. For snow, the sample numbers are ~183 and ~259 thousands, respectively. Although different scenarios have varying input channels, the sample number does not change.

3.1.2. Rain and Snow References

The CPR provides relatively reliable rain and snow data at high latitudes. However, the CPR may be subjected to signal saturation during heavy rain. Additionally, the CPR lacks an official rain rate product over land. The

Figure 2. Flowchart describing the production of training data sets combining CloudSat and GPM data (section 3.1), the training, and testing of deep neural networks during the period from 2014 to 2016 (section 3.2), and the application of well-trained DNN to GMI rain and snow estimation with the comparison to various reference precipitation data sets (section 3.3). Rain and snow were trained separately. Differences in the inputs between training and testing phases for the period 2014–2016 and the application in 2017 are explained in section 3.1 and 3.3.

complementarity between the CPR and DPR makes it possible to produce a complete precipitation training data set based on the two radars at high latitudes. The procedure for producing rain and snow references is given as follows.

- (1). Merging CPR and DPR rain rates over high-latitude oceans. Negative CPR rain rates indicate that the radar signal has been saturated and the actual rates should not be smaller than the absolute values. For such events, we adopted the larger one between the coincident DPR rain and the absolute value of CPR. Moreover, the CPR rain rates larger than 5 mm/hr could have a high risk of signal saturation and were therefore replaced with the DPR rain rates. It should be noted that although those saturated events only account for ~1% of all the positive events, they contribute to more than 10% of the total rain amount. This procedure was executed for both 2C-RAIN and 2C-PRECIP.
- (2). Excluding low-quality pixels (using available flags) that often correspond to large snow rates (>5 mm/hr) from 2C-SNOW in Greenland and Antarctic which are very few in number. For example, some of the low-quality pixels over the northern part of Greenland are due to complex topography and steep slopes.

As a result, rain over the oceans was from merged CPR and DPR rain data, rain over land was from DPR rain data, and snow over both land and ocean areas was from CPR snow data.

Table 2
Two Groups of Input Data From GMI, MODIS, and ECMWF for DNN Training

Group 1	Input	Group 2	Input
GMI1–5	GMI low-frequency channels 1–5 from 10 to 24 GHz	GMI1–9	GMI channels 1–9 similar to TMI channels
GMI6–9	GMI midfrequency channels 6–9 from 36 to 89 GHz	GMI3–9	GMI channels 3–9 similar to SSM/I channels
GMI10–13	GMI high-frequency channels 10–13 from 165 to 183 GHz	MODIS	MODIS channels 31–32 (10–12 μm)
GMI	GMI full channels 1–13	MODIS_E	Extended MODIS channels 27–36 (6–14 μm)
GMI + MODIS	GMI and MODIS channels 31–32 (10–12 μm)	MODIS+ECMWF	MODIS and two ECMWF variables
GMI + ECMWF	GMI and two ECMWF variables (T_{2m} and TPW)		
GMI + MODIS+ ECMWF	GMI, MODIS, and ECMWF

Figure 3. Structure of the DNN used in the study. There are 12 input scenarios as shown in Table 2 based on different combinations between GMI, MODIS, and ERA-Interim. Rain and snow were trained and tested separately. DNN = deep neural network; GMI = GPM Microwave Imager; GPM = Global Precipitation Measurement; MODIS = MODerate resolution Imaging Spectroradiometer; ERA-Interim = ECMWF Re-analysis Interim; ECMWF = European Centre for Medium-Range Weather Forecasts.

3.2. Deep Neural Networks for Rain and Snow Estimation

The DNN is characterized by deeper layers than traditional ANNs, which enables the DNN to automatically extract more complicated information from inputs, whereas ANNs tend to rely on manually designed features (Tao et al., 2016). Please refer to Bengio (2009) for detailed descriptions of the DNN. The architecture of the DNN used in this study is shown in Figure 3.

DNNs were trained separately for land and ocean areas in this study because of their different background emissivity characteristics. Land and ocean areas were defined using surface flags provided by the GPROF algorithm. Coasts were treated as part of land rather than a separate surface type considering that sample numbers in coastal regions are limited. Future studies should explore training individual networks over coasts for better precipitation estimates. To acquire the best rain/snow estimation, we tested different machine learning frameworks, transfer/activation functions, and optimization methods. The final network structure is given as follows.

A four-layer fully connected DNN was used in this study. The input layer receives inputs as introduced in section 3.1.1. For different scenarios, the numbers of input neurons range between 2 and 17. The two hidden layers both have 30 neurons. The number of coincident samples between CloudSat and GPM limits the complexity of the DNN because a too wide or deep structure has a high risk of overfitting due to the too many parameters and inadequate samples. The output layer only has one neuron sending out rain or snow data that are trained to the reference data sets in section 3.1.2.

The transfer function is tangent sigmoid, and the training method is the Levenberg–Marquardt algorithm (Moré, 1978), which blends the steepest descent method and the Gauss–Newton algorithm (Hao & Wilamowski, 2011). Actually, the widely used gradient descent method and its derivatives perform much worse than the Levenberg–Marquardt algorithm concerning precipitation estimation in this study. For small- and medium-sized networks like the DNN in this study, the Levenberg–Marquardt algorithm is more effective than the steepest descent and Gauss–Newton methods. However, for large-sized networks, the Levenberg–Marquardt algorithm would fail because the Hessian matrix inversion calculation will be highly time-consuming and the Jacobian matrix will occupy too much memory space (Hao & Wilamowski, 2011). Mean squared error (MSE) is taken as the cost function. The learning rate is set to 0.05. The initial μ value is 0.001, with an increasing factor of 10 and a decreasing factor of 0.1.

The data sets were randomly divided into the training set (60%), the validation set (20%), and the testing set (20%) to avoid overfitting. We also tested the use of the first 60% of the series as the training set, the second

20% as the validation set and the last 20% as the testing set, which does not show significant differences with the random division. All scenarios (see section 3.1.1) adopted the same random seeds for data set division and DNN parameter initialization to ensure the repeatability of the training results. The training is terminated when the error rate of the validation set increases continuously for more than six epochs, which means that further training would likely result in overfitting. The time of network training is generally within 2 hr on a laptop (i7-6,600 U CPU, 2.60 GHz, 16 GB RAM). Precipitation estimation based on the trained networks is very quick with almost negligible time cost.

These aforementioned hyperparameters were determined by trial and error until most input scenarios showed satisfying performance (see section 3.1.1). To facilitate a fair comparison between different training scenarios, we adopted fixed hyperparameters for all DNNs in this study, which may not be adaptive to alteration in input feature numbers. For practical applications, a specific DNN can be further optimized to achieve better performance.

We also tested regular neural networks that has only one hidden layer (30 neurons) with the other settings the same as the DNN. In the training and testing periods, regular networks were worse than the DNN in many cases. For example, the correlation coefficient (CC) of the DNN for snow training over land increased by 13% compared to the regular neural network. However, in some cases, some metrics of the DNN were slightly worse. To demonstrate the differences between deep and regular networks thoroughly, we applied the two types of networks to GMI rain/snow estimation in 2017 (see section 3.3). The validation results show that over the oceans, CC values of the DNN increases by 15% and 9% for rain and snow, respectively. However, the improvement is less obvious over land probably because (1) the DNN suffers more from overfitting due to the large number of parameters and limited samples and (2) the initial conditions cause uncertainties in the comparison between the regular and deep networks. Overall, the DNN is a better choice than the regular network in this study, although the improvement is not very obvious in some cases. It is expected that future studies can obtain better results if more samples are employed to train more complex networks.

3.3. Application to GMI Rain and Snow Estimation

The well-trained networks were further applied to high-latitude rain and snow estimation using GMI T_b from all the 5,677 orbits in 2017. We selected two representative networks, which are based on GMI channels 1–13 (henceforth DNN1), and GMI + ECMWF (henceforth DNN2), respectively. MODIS IR data were not included because MODIS only has some coincidences with GMI, which are not enough for practical applications. For DNN2, ERA-Interim environmental data at 3-hourly and 0.75° resolutions were matched to orbital GMI T_b data using linear interpolation. ERA-Interim data were first interpolated to all GMI pixel locations and then interpolated to GMI observation time. Then, the matched GMI T_b and environmental data were taken as the inputs to DNN2.

Rain and snow estimates from DNN1, DNN2, and the GPROF algorithm, which are all at the original GMI resolution, were resampled to 1° by averaging all GMI pixels contained in 1° pixels. Then, ERA-Interim and MERRA2 precipitation were linearly interpolated to the GMI observation time. Finally, GPROF and DNN estimates were compared to the matched reanalysis precipitation.

ECMWF data from the ECMWF-AUX product (made available by the CloudSat team) were originally used in network training. However, it was noted that data from ECMWF-AUX are slightly different from those obtained from ERA-Interim used in the full-swath precipitation estimation in 2017. This problem is probably caused by (1) the differences between the two environmental data sources and (2) the different spatial downscaling and time matching methods used by ECMWF-AUX and this study. Such differences could lead to overestimation of precipitation based on DNN2. To reduce the impact of this issue, we retrained networks that replaced ECMWF-AUX provided in 2B-CSATGPM with ERA-Interim data. As a result, metrics in training and testing periods showed little differences, whereas the accuracy of DNN2-based precipitation was notably improved. Therefore, it is highly recommended to use unified input data sets to networks in practical applications.

3.4. Metrics

To comprehensively evaluate the performance of GPROF- and DNN-based precipitation estimates, several metrics were utilized. These include CC, mean error (ME), relative bias (BIAS), and root-mean-square error (RMSE). In addition, four contingency metrics were used including the probability of detection (POD), false

alarm ratio (FAR), critical success index (CSI), and Heidke skill score (HSS). Larger POD, CSI, and HSS and smaller FAR indicate better performance. Please refer to Tang, Behrangi et al. (2018) and Behrangi et al. (2018) for the formulas and detailed descriptions of these metrics.

4. Results

4.1. Intercomparison Between CPR, DPR, and GMI Precipitation Estimates

The precipitation from the CPR, DPR, and GMI was intercompared, which is helpful to understand the similarities and differences between these sensors before DNNs were trained. GMI rain and snow were provided by GPROF. Figure 4 shows the scatter density plot between precipitation of the CPR, DPR, and GMI. The two rain products of the CPR, that is, 2C-PRECIP and 2C-RAIN, were examined separately. The 2C-RAIN has fewer samples than the 2C-PRECIP due to the shorter time period (Table 1). With regard to rain over oceans, both the DPR and GMI agree well with the CPR, but the GMI shows better statistics. The CC between the DPR and 2C-PRECIP is 0.50, and the CC between GMI and 2C-PRECIP increases to 0.61. When compared to 2C-RAIN, the CC of the DPR is 0.48, whereas the CC of the GMI is 0.57. Both GMI and DPR underestimate the rain rates benchmarked by the CPR. The BIAS of the DPR and GMI is around -50% when compared to 2C-PRECIP and around -40% when compared to 2C-RAIN. While both the DPR and GMI underestimate moderate to heavy rain, they overestimate light rain.

Concerning snow comparison, the DPR is not included because 2B-CSATGPM does not contain DPR snow information. The CC between the GMI and CPR for snow is 0.30 over oceans and 0.26 over land, indicating that GPROF-based snow estimation is not as consistent with CPR as rain estimation. The BIAS of GMI snow is -54% over oceans and -37% over land with overestimation at light snow but underestimation at moderate to heavy snow, which is similar to the error characteristics of rain.

The occurrence frequency of DPR and GMI precipitation was also studied (Figure 5). Because the GPROF algorithm produces probabilities of precipitation, we adopted different probability thresholds to determine GMI precipitation occurrence. In the Northern Hemisphere (NH), the DPR and GMI show similar frequency curves for moderate to heavy precipitation except for the warm season during Jun–Aug (JJA) and Sep–Nov (SON) when the DPR observes more heavy precipitation. At light precipitation, different thresholds of probabilities strongly affect the shape of the GMI curves. The occurrence frequency of the GMI generally reaches the peak at 0.1 mm/hr and then drops sharply as the precipitation rate declines. As the probability threshold increases from 0 to 0.5, the over detection of the GMI at light precipitation mitigates and even turns into under detection compared to the frequency of the DPR. In the Southern Hemisphere (SH), the frequency curve shapes are similar with those in NH, while the sample number is smaller than that in NH as stated in section 2.1. In addition, frequency curves in SH show narrower tails over intense precipitation bins than those in NH.

4.2. Critical Issues of Rain and Snow Estimation

There are some critical issues about precipitation estimation based on the DNN that need careful considerations, including (1) which CloudSat rain product (2C-RAIN or 2C-PRECIP) should be used to train the DNN over the oceans, (2) does DNN-based precipitation method overestimate light precipitation and underestimate heavy precipitation, similar to GPROF? If so, how can we resolve this issue? (3) How much improvement is gained by merging CPR and DPR rain prior to training? and (4) Is it more effective to train the DNN separately for rain and snow or is it better to train the DNN by total precipitation and determine phases through post processing? The analyses in this section are based on DNN1 forced by GMI channels 1–13.

Concerning the first issue, we compared the performance of the DNN based on 2C-RAIN and 2C-PRECIP over high-latitude oceans from March 2014 to May 2016 as shown in Figure 6. CC of the DNN based on 2C-RAIN and 2C-PRECIP is similar to each other (~ 0.68). BIAS (8%) of the DNN based on 2C-PRECIP is slightly larger than that (-3%) based on 2C-RAIN. Both products perform similarly except that 2C-PRECIP has much more rain events between 0.2 and 2 mm/hr (Figure 6b) than 2C-RAIN. We further applied DNNs to GMI rain estimation in 2017 as introduced in section 3.3. Results show that the DNN based on 2C-PRECIP resulted in large overestimation compared to reanalysis products probably due to the too many moderate rain samples. In contrast, 2C-RAIN performed much better, probably because it implements more realistic assumptions on the vertical distribution of rainwater and the rain drop size distribution, making it more suited to warm rain retrieval compared to 2C-PRECIP (Behrangi et al., 2012). Therefore, 2C-RAIN was adopted in the following sections.

Figure 4. Scatter density plots between rain of the CloudSat CPR and GPM DPR/GMI from 2014 to 2016 in high latitudes (50–70°S/N) using (a, b) CPR 2C-PRECIP and (c, d) CPR 2C-RAIN as references, respectively. Rain was compared only over oceans. Panels (e) and (f) are the comparison of snow between the GMI and CPR 2C-SNOW over ocean and land areas, respectively. CPR = Cloud Profiling Radar; GPM = Global Precipitation Measurement; DPR = Dual-frequency Precipitation Radar; GMI = GPM Microwave Imager.

With regard to the second issue, neural networks often generate many tiny precipitation estimates (e.g., <0.01 mm/hr), which should be zero in most cases and may underestimate the frequency of heavy precipitation like the GPROF algorithm. To correct this problem, we adopted a probability matching method based on the cumulative density function (CDF), which was first suggested by Miller (1972) and is commonly used for satellite precipitation adjustment (Karbalaei et al., 2017; Yang et al., 2016). The CDF based on the reference data in the training period was used to adjust the DNN estimates in the testing period for each season and hemisphere separately. Figure 7 shows the probability density function (PDF) curves of rain and snow from radars, GPROF, the DNN before and after adjustment in the testing period. For the moderate to heavy rain and snow, the DNN agrees better with the reference data sets compared to GPROF. However, the frequency of the preadjusted DNN light rain and snow is higher than that of GPROF and the reference data set. After the

Figure 5. Occurrence frequencies of DPR and GMI precipitation at high latitudes ($>50^{\circ}\text{S/N}$) in the Northern Hemisphere (NH) and Southern Hemisphere (SH) in the four seasons. The dark and gray lines correspond to different thresholds of the probability of precipitation provided by the GMI. Precipitation is assigned zero values if the probability is lower than the threshold. In the legend, numbers behind “GMI-” represent thresholds. DJF = December–February; MAM = March–May; JJA = June–August; SON = September–November.

CDF adjustment, the curves of DNN rain and snow almost overlap the curves of the reference data sets. To implement the adjustment method in real applications, it is necessary to establish a spatially distributed CDF library based on long-term radar data because the CDF is regionally dependent. However, this study does not implement the CDF adjustment due to the short period.

For the third issue, the merging between CPR 2C-RAIN and DPR rain was conducted over high-latitude oceans according to section 3.1.2. CPR and DPR have almost equal contributions to the events with the saturation problem. For example, there are 241 events with rain rates larger than 10 mm/hr in the merged product, among which 178 suffer from the signal saturation problem according to the signs of CPR rain values. About half of the 178 events adopt DPR rain, and the other half adopt CPR rain. The merging process also affects DNN training performance. The DNN based on the merged CPR-DPR data set elevates the ratio of moderate to heavy rain compared to that based merely on 2C-RAIN.

About the fourth issue, the preprocessing of precipitation means that rain and snow are trained separately using the DNN and thus the well-trained DNN can directly produce rain and snow from T_b inputs. For

Figure 6. Scatter density plots between rain of CloudSat CPR and two deep neural networks based on (a) 2C-RAIN and (b) 2C-PRECIP over high-latitude oceans from March 2014 to May 2016. CC and BIAS are shown at the top left corner. CC = correlation coefficient; BIAS = relative bias; GMI = GPM Microwave Imager; DNN = deep neural network.

Figure 7. Probability density function of (a) rain and (b) snow over high-latitude oceans. The black lines represent reference data sets from spaceborne radars. The green lines represent GMI GPROF rain and snow estimates. The blue lines represent DNN rain and snow estimates. The red lines represent DNN estimates after cumulative density function adjustment using reference data sets. GPROF = Goddard Profiling Algorithm; DNN = deep neural network.

postprocessing, total precipitation is trained first and the DNN outputs are partitioned into rain and snow using environmental data such as temperature (Behrangi et al., 2018; Harpold et al., 2017; Sims & Liu, 2015). The study used a critical threshold of T_{2m} to implement rain-snow partitioning for the postprocessing. A sensitivity study was conducted with the threshold ranging between -3 and 3 °C. Figure 8 compares the metrics of rain and snow using preprocessing and postprocessing strategies over high-latitude oceans. Similar figures were produced over land that are not shown for simplification. Metrics of preprocessing do not change with temperature thresholds, because snow is estimated without using temperature data. CC and RMSE of rain are insensitive to thresholds probably because moderate and heavy snow, which has a stronger effect on the two metrics than light snow, is less sensitive to temperature thresholds. On the whole, the two methods show similar performance concerning rain estimation. This is probably because most rain events occur at relatively high temperatures and those occurring at a relatively low temperatures (e.g., <3 °C) are small and have limited effect on the overall performance. With regard to snow, postprocessing performs best when the temperature threshold is set to ~ 1 °C. However, preprocessing is always better at snow estimation than postprocessing in terms of all the metrics. Therefore, the preprocessing strategy was adopted in this study.

The success of the preprocessing strategy in snow estimation suggests that GMI PMW data may contain sufficient information for snow retrieval without the need for auxiliary data such as the temperature.

4.3. Performance of Deep Neural Networks

This section addresses three issues related to the performance of the DNN in precipitation estimation, including (1) the comparison between metrics of the DNN and the GPROF algorithm, (2) the effect of combining PMW, IR, and environmental data, and (3) the advantages and disadvantages of different GMI and MODIS channels.

Figure 9 shows the CC, RMSE, and HSS of rain and snow estimates from the GPROF algorithm and the DNN using radar precipitation as the reference. The DNN scenarios include GMI, GMI + MODIS, GMI + ECMWF, and GMI + MODIS + ECMWF (section 3.1.1). The DNN outperforms GPROF for almost all the metrics over land and ocean areas in training and testing periods. Taking metrics over oceans as an example, the CC of GPROF rain is 0.50 and 0.55 in two periods, whereas DNNs' CC ranges between 0.7 and 0.8. The RMSE of DNNs (<0.48 mm/hr) is also smaller than that of GPROF (~ 0.56 mm/hr). For HSS, GPROF is better than the DNN based on the GMI and GMI + MODIS but worse than the other two scenarios. The improvement of the DNN in snow estimation is much more significant. For instance, over oceans, the CC was improved from 0.28 for GPROF to more than 0.7 for DNN scenarios, and the HSS of DNNs (>0.5) largely exceeds that of GPROF (~ 0.19). Similar conclusions were also found over land. On the whole, the DNN substantially improves rain and snow estimation compared to GPROF at high latitudes. The relatively limited improvement of the

Figure 8. Comparison of (a) CC, (b) RMSE, (c) ME, and (d) HSS of two rain-snow partitioning strategies, that is, preprocessing and postprocessing, over high-latitude oceans. The X axis represents the critical thresholds of temperature used to separate rain and snow for postprocessing. CC = correlation coefficient; RMSE = root-mean-square error; ME = mean error; HSS = Heidke skill score.

DNN on rain estimation over land is mainly because DPR data, which are close to GMI as shown in Figure 5, were used as the training reference.

The addition of MODIS to GMI improves the quality of rain estimation over oceans and snow estimation over both land and ocean areas, indicating that merging PMW and IR data is better than PMW only in precipitation

Figure 9. CC, RMSE and HSS of rain and snow estimates from the GPROF algorithm and deep neural networks over high-latitude land and ocean areas. Various combinations between GMI, MODIS and ECMWF data are used as network inputs. CC = correlation coefficient; RMSE = root-mean-square error; HSS = Heidke skill score; GMI = GPM Microwave Imager; GPROF = Goddard Profiling Algorithm; MODIS = MODerate resolution Imaging Spectroradiometer; ECMWF = European Centre for Medium-Range Weather Forecasts.

Figure 10. Taylor diagrams displaying the performance of rain and snow estimates from the GPROF algorithm and deep neural networks over high-latitude land and ocean areas. Various combinations between GMI, MODIS, and ECMWF data are used as network inputs. GMI = GPM Microwave Imager; GPROF = Goddard Profiling Algorithm; MODIS = MODerate resolution Imaging Spectroradiometer; ECMWF = European Centre for Medium-Range Weather Forecasts.

estimation (Figure 9). However, over land GMI + MODIS is slightly worse than GMI, which could be partly due to the false information related to precipitation events derived from IR window channels over the frozen surface. The addition of ECMWF can further improve the accuracy of rain and snow estimation. GMI + ECMWF and GMI + MODIS+ECMWF are similar to each other and better than GMI + MODIS in most cases. Therefore, environmental data seem more useful than IR data in assisting PMW precipitation retrieval. Nevertheless, the overall statistics shown in Figure 9 suggest that a relatively marginal improvement is obtained by adding IR and environmental data to GMI channels compared to the improvement that is otherwise obtained by applying the DNN to only GMI channels.

To demonstrate the differences between scenarios comprising various GMI and MODIS channels, the Taylor diagram was utilized, which is based on CC, standard deviation and centered root-mean-square difference (Taylor, 2001). Figure 10 compares the performance of scenarios in Group 1 (see section 3.1.1) based on data in the testing period. The point on the horizontal axis represents the observation data set, which was used as the reference in the evaluation. The distance between the other points in the diagram and the reference point is inversely proportional to the quality of the estimated rain or snow. The low-frequency GMI1–5 is good for rain estimation over the oceans but is worst for snow estimation over oceans and rain/snow estimation over land, which all rely heavily on information from high-frequency channels. The midfrequency GMI6–9 is good at both rain and snow estimation over both land and ocean areas. The high-frequency GMI10–13 is best at snow estimation but worst at rain estimation over the oceans, because a large fraction of rain over the oceans is from shallow systems with no ice aloft and thus little scattering signal. GMI10–13 is also not as

Figure 11. Similar with Figure 10 but with different network inputs.

good as GMI6–9 for rain estimation over land. GMI1–13 is constantly better than any other scenarios with incomplete GMI channels. GMI + MODIS, GMI + ECMWF, and GMI + MODIS+ECMWF are close to each other and generally better than GMI1–13, which is similar to the results in Figure 9.

Figure 11 is similar to Figure 10 but displays the results based on the scenarios in Group 2. MODIS_E (6–15 μm) is better than MODIS (10–12 μm) at rain and snow estimation, indicating that IR channels beyond the commonly used 10–12 μm can also contribute to improvements in precipitation retrieval. The value of water vapor channels ($\sim 6.5 \mu\text{m}$) has particularly been shown in the past (Ba & Gruber, 2001; Behrangi et al., 2009; Kurino, 1997). These results support joint retrieval using a full range of GOES–16 and GMI channels, although it may not be simple due to differences in observation time. By combining MODIS and ECMWF data, better precipitation estimates were obtained as expected.

Moreover, GMI channel combinations were used to imitate other sensors. GMI1–9 can represent TMI and GMI3–9 can represent SSM/I. The two scenarios performed similarly to each other. They were even better than GMI1–13 at rain estimation over oceans, possibly because the weak relationship between high-frequency channels (GMI10–13) and rain may confuse the DNN. In other cases, GMI1–13 is generally better, implying that high-frequency channels are beneficial to improving high-latitude precipitation estimation.

In summary, the trained DNN is superior to GPROF V5 at high-latitude rain and snow estimation. It is important to note that the value of DNNs is not only to establish the relationships between PMW/IR data and precipitation but also to reveal some hidden physical mechanisms (Tapiador et al., 2004). An interesting application is the use of DNNs to explore optimal channel combinations for estimating rain and snow under different surface types.

Figure 12. Spatial distributions between GPROF, DNN, ERA-Interim, and MERRA2 rain and snow for an event on 17 May, 2017. A segment of the GMI orbit is presented where both GPM and CloudSat have samples and rain and snow coexist. GPROF = Goddard Profiling Algorithm; DNN = deep neural network; ERA = ECMWF Re-Analysis; ECMWF = European Centre for Medium-Range Weather Forecasts; MERRA2 = Modern-Era Retrospective analysis for Research and Applications Version 2.

4.4. Application in GMI Rain and Snow Estimation

4.4.1. Evaluation Based on Reanalysis Data Sets

GMI rain and snow were estimated based on the well-trained DNN as introduced in section 3.3. Figure 12 compares precipitation estimates of the GPROF algorithm, DNN, ERA-Interim, and MERRA2 for an event on 17 May 2017 with rain and snow at the same time. The DNN was trained based on T_b from GMI channels 1–13. On the whole, the precipitation extents of the DNN and GPROF are similar with two separate rain bands in the southern and central parts of the orbit. ERA-Interim and MERRA2 show similar precipitation patterns, whereas the intensity is much larger than the GPROF algorithm and DNN. The DNN enhances both rain and snow rates compared to GPROF and is closer to ERA-Interim and MERRA2. For example, for this event, the CC values between GPROF and ERA-Interim are 0.21 for rain and 0.31 for snow, which increase to 0.37 and 0.43 for the DNN rain and snow estimates, respectively.

In addition to the above event comparison, for each GMI orbit in 2017, rain and snow were estimated using two types of DNNs, that is, DNN1 based on GMI and DNN2 based on GMI + ECMWF defined in section 3.3. Precipitation estimates from GPROF, DNN1, and DNN2 were compared to ERA-Interim data over land, ocean, and land plus ocean areas (LO), respectively. Boxplots of four typical metrics for rain and snow are shown in Figures 13 and 14, respectively. The left and right edges of each box indicate the 25th and 75th percentiles, respectively. The central black mark indicates the median. The dashed lines are whiskers extending from the interquartile with a length of 1.5 times that of the box width. The dots are outliers. Table 3 summarizes the metrics of GPROF, two DNNs, and IMERG over LO.

With regard to rain, the CC of DNN1 and DNN2 is similar to that of GPROF over both land and ocean areas (Figure 13). For instance, the CC of GPROF, DNN1, and DNN2 over LO is 0.49, 0.47, and 0.50, respectively. The RMSE of the DNNs and GPROF over LO are all 0.16 mm/hr in Table 3 due to rounding off. The DNN greatly reduces the underestimation by GPROF. The BIAS of GPROF is reduced from -52% to -24% for DNN1 and -26% for DNN2. Furthermore, DNN2 has much better CC than DNN1 over land. This is not observed over the oceans, indicating that environmental data may be more effective for land precipitation retrieval partly

Figure 13. Boxplots of (a) CC, (b) RMSE, (c) ME, and (d) HSS of GPROF and DNN rain estimates using ERA-Interim as the benchmark at high latitudes in 2017. Metrics were calculated over land, ocean, and land plus oceans (LO) areas, respectively. DNN1 uses GMI brightness temperature as inputs. DNN2 uses GMI and ECMWF data as inputs. CC = correlation coefficient; RMSE = root-mean-square error; ME = mean error; HSS = Heidke skill score; GPROF = Goddard Profiling Algorithm; DNN = deep neural network; GMI = GPM Microwave Imager; ERA-Interim = ECMWF Re-Analysis Interim; ECMWF = European Centre for Medium-Range Weather Forecasts; LO = land plus ocean areas.

Figure 14. Same as Figure 13 but for snow evaluation.

Table 3*Metrics of Rain and Snow Estimates From the GPROF Algorithm, Deep Neural Networks, and IMERG at High Latitudes in 2017 Compared to ERA-Interim Data*

	Rain				Snow			
	GPROF	DNN1	DNN2	IMERG	GPROF	DNN1	DNN2	IMERG
CC	0.49	0.47	0.50	0.45	0.28	0.45	0.53	0.23
RMSE (mm/hr)	0.16	0.16	0.16	0.97	0.08	0.08	0.08	0.45
ME (mm/hr)	−0.02	−0.01	−0.01	0.02	−0.01	0.01	0.01	0.05
BIAS (%)	−52.27	−23.66	−26.33	2.24	−32.83	11.42	19.84	54.18
POD	0.30	0.35	0.36	0.32	0.19	0.45	0.52	0.27
FAR	0.33	0.47	0.47	0.29	0.71	0.64	0.56	0.63
CSI	0.24	0.24	0.25	0.28	0.12	0.24	0.30	0.18
HSS	0.33	0.32	0.32	0.28	0.16	0.32	0.40	0.19

Note. DNN1 was forced with GMI data. DNN2 was forced with GMI and ECMWF environmental data. Metrics were calculated over land and ocean areas, respectively. GPROF = Goddard Profiling Algorithm; DNN = deep neural network; IMERG = Integrated Multi-satellite Retrievals for GPM; GPM = Global Precipitation Measurement; CC = correlation coefficient; RMSE = root-mean-square error; ME = mean error; BIAS = relative bias; POD = probability of detection; FAR = false alarm ratio; CSI = critical success index; HSS = Heidke skill score.

because microwave retrievals tend to be worse over land. With regard to contingency metrics, DNNs have better POD but worse FAR, resulting in similar CSI and HSS compared to GPROF.

With regard to snow, DNN1 and DNN2 are always better than GPROF concerning almost all metrics over land and ocean areas, except that the RMSE of DNN1 is slightly larger than that of GPROF over land. The CC over LO is improved from 0.28 for GPROF to 0.45 and 0.53 for DNN1 and DNN2, respectively (Table 3). The BIAS over LO is reduced from −33% for GPROF to 11% and 20% for DNN1 and DNN2, respectively. For all the four contingency metrics, DNN2 is better than DNN1 followed by GPROF. For instance, the CSI increases from 0.12 for GPROF to 0.30 for DNN2, and the corresponding HSS increases from 0.16 to 0.40.

In addition, metrics for IMERG rain and snow are also displayed in Table 3. IMERG first estimates total precipitation from Level-2 PMW and IR products and separates rain and snow following the method proposed by Sims and Liu (2015). Although the GPROF algorithm uses the same method as IMERG to determine precipitation phases, the total precipitation rate of GPROF is different from that of IMERG. Therefore, IMERG and GPROF snow data are different, leading to the large metric difference between them particularly for RMSE (Table 3). In addition, IMERG data only cover 60°S to N while GPROF and DNN data reach 70°S to N resulting in the difference in study regions. We also calculated GPROF and DNN metrics over the same region with IMERG and found that the metrics changed little. Concerning rain, IMERG is slightly worse than GPROF and DNN. For example, CC decreases from 0.49 for GPROF to 0.45 for IMERG. For snow, IMERG shows worse CC, RMSE, and BIAS but better contingency metrics than GPROF. A possible explanation is that IMERG merges data from other PMW and IR sensors, which are not as good as the GMI but can improve precipitation detectability. On the whole, IMERG snow is worse than DNN snow, indicating that DNN-based rain and snow estimates are beneficial to the improvement of the quality of IMERG at high latitudes.

MERRA2 data with higher temporal resolutions than ERA-Interim were also used to evaluate GPROF and DNN estimates. The general conclusions are similar to those based on ERA-Interim. The largest difference is that the MERRA2 snow rate is much larger than that of ERA-Interim, which reduces the overestimation of DNN snow whereas exacerbates the underestimation of GPROF snow as shown in Table 3.

4.4.2. Evaluation Based on the Triple Collocation Method

Although reanalysis data sets were used to assess the performance of DNN and GPROF precipitation, they also contain uncertainties. Therefore, we adopted the Triple Collocation (TC) method to analyze the errors of the precipitation estimates. The TC method was originally introduced by Stoffelen (1998) and utilized by McColl et al. (2014), Alemohammad et al. (2015), and Pan et al. (2015), which can obtain relative error estimation (i.e., CC and RMSE) of three independent products without the need for the true value. Previous studies have shown the effectiveness of the TC method in precipitation evaluation (Li et al., 2018). TC-based CC and RMSE of GPROF were calculated for each 1° grid pixel based on the triplet comprising GPROF, ERA-Interim, and MERRA2. Metrics of the DNN were acquired in the same way. Although ERA-Interim and MERRA2 were not entirely independent, the TC-based metrics should be useful to judge the relative performance of the DNN and GPROF. Figure 15 shows that the TC-based CC of the DNN is better than that of GPROF and the

Figure 15. Spatial distributions of correlation coefficient for GPROF, DNN1, and DNN2 rain and snow using the triple collocation method at high latitudes in 2017. GPROF = Goddard Profiling Algorithm; DNN1 = deep neural network based on GMI channels 1–13; DNN2 = deep neural network based on GMI+ECMWF.

patterns are similar to the analyses in section 4.1. Similar conclusions can be drawn for TC-based RMSE. The TC-based validation verifies the conclusions in this study.

4.4.3. High-Latitude Rain and Snow Distribution

Mean spatial distributions of high-latitude rain from GPROF, DNN1, DNN2, IMERG, ERA-Interim, and MERRA2 for 2017 are shown in Figure 16. Reanalysis data sets, that is, ERA-Interim and MERRA2, show similar rain patterns. Compared with the two reanalysis data sets, GPROF notably underestimates rain over the oceans whereas IMERG significantly overestimates rain over the oceans. DNN1 and DNN2 agree better with reanalysis data sets than GPROF and IMERG. For example, the DNN mean rain rates in the northern Pacific Ocean, the western coast of North America, Greenland, and the oceans around the 50°S show values that are more comparable to reanalysis than GPROF. DNN2 produces stronger rain rates than DNN1 over the oceans and Europe. In contrast, the DNN2 rain rate is weaker than DNN1 in the mountainous region along the western coast of North America, probably because the quality of environmental data is worse in complex terrain and more precipitation is identified as snow as discussed below. It should be noted that similar to CloudSat and reanalysis data, the DNN shows a clear precipitation gradient around 70°N over the Atlantic Ocean, a feature that was not detected by the Global Precipitation Climatology Project based on a study by Behrangi et al. (2016). GPROF also captures this gradient, but with lower intensities.

Snow distributions are exhibited in Figure 17. The GPROF mean snow rate is very low at high latitudes, although the spatial pattern agrees with that of the reanalysis data sets. IMERG produces higher snow intensity than GPROF but overestimates snow in regions around 60°S and 0°E/W. The DNN enhances snow intensity particularly over the oceans in the SH compared to GPROF and is much closer to ERA-Interim and MERRA2. In addition, in Greenland where snow is frequent, the DNN produces much heavier snow than adjacent areas similar to reanalysis data sets. This feature is not noticeable in GPROF. Particularly, DNN2 produces heavy snowfall along the western coast of North America similar to reanalysis data sets, which partly explains

Figure 16. Spatial distributions of rain from various products and two deep neural networks at high latitudes in 2017. DNN1 uses GMI brightness temperature as inputs. DNN2 uses GMI and ECMWF data as inputs. The dark dashed lines outline the boundary of the IMERG coverage. GPROF = Goddard Profiling Algorithm; DNN1 = deep neural network based on GMI channels 1–13; DNN2 = deep neural network based on GMI+ECMWF; IMERG = Integrated Multi-satellite Retrievals for GPM; GMI = GPM Microwave Imager; GPM = Global Precipitation Measurement; ERA-Interim = ECMWF Re-Analysis Interim; ECMWF = European Centre for Medium-Range Weather Forecasts; MERRA2 = Modern-Era Retrospective analysis for Research and Applications Version 2.

its low rain rates in this region. On the whole, the spatial distributions of DNN snow are more reasonable and closer to reanalysis data sets.

5. Discussion

This study uses optimized DNNs to establish relationships between precipitation data from merged spaceborne radars and PMW, IR, and environmental data in high latitudes. However, the CPR faces the challenge of measuring heavy precipitation due to signal saturation/loss, which could result in precipitation underestimation. Compared to CPR data, GPROF-based GMI estimates and DPR precipitation data are even lower in precipitation intensity (Figure 4), implying that the actual underestimation of GPROF and DPR could be more severe. Despite its disadvantages, the CPR should be better than other satellite sensors in precipitation estimation in high latitudes according to the findings in many published studies. The independent validation in section 4.4 also proves that the use of CPR precipitation data to train GMI T_b data seems effective. On the whole, the CloudSat CPR is the least bad choice for this study.

Ground precipitation observations can also be used to train regional DNNs. For example, the Multi-Radar Multi-Sensor system in the United States utilizes data measured by the Weather Surveillance Radar-1988 Doppler (WSR-88D) Next Generation Weather Radar (NEXRAD), automated rain gauges, and numerical model outputs providing precipitation products at the resolution of 1 km and 2 min (Zhang et al., 2016). Other regions such as Europe and China are also equipped with dense radars and/or weather stations that can provide high resolution and high-accuracy precipitation observations. It is possible that DNNs trained using ground observations would have better quality than those based on spaceborne radars only. However, the

Figure 17. Same as Figure 16 but for snow.

challenge lies in the transferability of the network trained in a specific region to other areas to facilitate global application.

The proposed precipitation estimation method can also be applied at low latitudes. However, the coincident sample number of spaceborne radar precipitation and GMI data at low latitudes is largely reduced as we move to lower latitudes. Furthermore, the method can also be applied to other PMW or IR sensors to improve their precipitation retrieval accuracy. MHS and Advanced Very High Resolution Radiometer are on National Oceanic and Atmospheric Administration platforms, creating a great opportunity to utilize PMW-IR data for precipitation estimation based on the DNN.

It should be noted that different IR sensors may have notable differences. For example, the experience in the TRMM era showed that IR data from the Visible Infrared Radiometer sensor carried by the TRMM satellite was different enough from geosynchronous IR data that the calibrations developed with Visible Infrared Radiometer could not be directly applied to geosynchronous data. The study utilized MODIS IR data, which are provided along with CloudSat CPR data and easy to access. For other IR sensors, networks should be retrained and similar results may be expected.

6. Conclusions

Satellite precipitation retrieval is confronted with difficulties in estimating high-latitude rain and snow. The study takes advantage of precipitation estimates from spaceborne radars (CPR and DPR) to train PMW (GMI) and IR (MODIS) T_b data as well as reanalysis environmental variables (ECMWF) based on the coincident samples between GPM DPR/GMI and CloudSat CPR. Various scenarios were designed to assess the performance of different combinations of PMW and IR channels as well as environmental data. Finally, optimized DNNs were applied to estimate GMI rain and snow at high latitudes in 2017. The DNNs-based rain and

snow estimates were then compared to GPROF-based GMI estimates and reanalysis precipitation data sets (ERA-Interim and MERRA2) for detailed analysis. Main conclusions are given as follows.

The original GMI and DPR rain estimates show CC from 0.47 to 0.61 relative to CPR data over high-latitude oceans, which is, however, accompanied by large underestimation with BIAS from -35% to -41% . Snow estimates of the GMI exhibit low CC (0.25–0.28) and large BIAS (-38% to -53%) benchmarked by CPR data. The complementarity between CPR and DPR makes it possible to generate a training reference data set with good quality over both land and ocean areas. The signal saturation/loss problem of the CPR at heavy rain is mitigated by the merging process, which is proved effective in improving DNN-based moderate and heavy rain estimates. Between the two CPR rain products, 2C-RAIN is selected for training the DNN because 2C-PRECIP results in notable overestimation of DNN-based rain based on the evaluation performed using reanalysis data sets.

In the training and testing phases of networks, DNN-based rain is constantly better than GPROF-based rain over high-latitude land and ocean areas. For snowfall retrieval, DNN outperformed GPROF by far. Combining IR and environmental data with PMW T_b contributes to better precipitation estimation. Environmental data are particularly useful in improving snow estimation accuracy. The investigation into different GMI channels reveals that the low-frequency channels are good at estimating rain over the oceans but is inefficient for estimating ocean snow and land rain. The midfrequency channels are good at both rain and snow estimation in all cases. The high-frequency channels are best at estimating land and ocean snow but worst at estimating ocean rain. The results on MODIS channel combinations show that IR channels from 6 to $15\ \mu\text{m}$ perform better than the commonly used channels between 10 and $12\ \mu\text{m}$. In summary, the DNN is efficient in reflecting the general physical mechanisms of precipitation retrieval and could be used to explore the optimal channel combinations for precipitation estimation.

GMI rain and snow in 2017 are estimated based on two types of DNNs using inputs from GMI T_b data and GMI T_b plus ECMWF environmental data, respectively. Benchmarked by ERA-Interim and MERRA2 data, rain from the two DNNs shows much less underestimation than GPROF rain with the BIAS reduced from -52% to around -25% . With regard to snow, the two DNNs produce significantly better merit scores than GPROF and IMERG snow, showing higher consistency with reanalysis products. Compared to the DNN based only on GMI T_b , the DNN based on GMI T_b and environmental data generates more accurate precipitation estimates, particularly for snow. With regard to the spatial distributions, the DNN produces annual intensity and spatial distributions that are more consistent with reanalysis data sets than GPROF and IMERG.

In conclusion, the precipitation estimation method presented in this study exhibits satisfactory performance at high latitudes and could be considered as a viable alternative for the GPROF algorithm used to generate the IMERG products. The method is easy to implement and quick in speed with satisfying snow estimation accuracy. The study also shows that IR and environmental data can enhance PMW-only retrievals, especially for snowfall. Advanced Very High Resolution Radiometer and geostationary platforms offer IR observations with good spatiotemporal sampling. Environmental data, with growing accuracy and spatiotemporal resolutions, are also available from reanalysis products. Therefore, the inclusion of these additional data sets in the next generation of physical or statistical precipitation retrieval methods is recommended.

Acknowledgments

This study was financially supported by the National Natural Science Foundation of China (grants 91547210, 71461010701, and 91437214) and the National Key Research and Development Program of China (2016YFE0102400). A. Behrangi is supported by NASA Energy and Water Cycle Study (NNH13ZDA001N-NEWS) and NASA weather (NNH13ZDA001N-Weather) programs. Additional support came from a scholarship from the China Scholarship Council (CSC). The 2B-CSATGPM product, GPM Level-1C GMI brightness temperatures, GPM Level-2A GMI GPROF precipitation, and Level-3 IMERG precipitation were downloaded from the Science Team On-Line Request Module (STORM; <https://storm.pps.eos-dis.nasa.gov/storm>). ERA-Interim data were downloaded from the ECMWF Public Data sets web interface (<http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc>). MERRA2 data were downloaded from Goddard Earth Sciences Data and Information Services Center (GES DISC; <https://disc.gsfc.nasa.gov/datasets>).

References

- Alemohammad, S. H., McColl, K. A., Konings, A. G., Entekhabi, D., & Stoffelen, A. (2015). Characterization of precipitation product errors across the United States using multiplicative triple collocation. *Hydrology and Earth System Sciences*, 19(8), 3489–3503. <https://doi.org/10.5194/hess-19-3489-2015>
- Asong, Z. E., Razavi, S., Wheeler, H. S., & Wong, J. S. (2017). Evaluation of integrated multisatellite retrievals for GPM (IMERG) over southern Canada against ground precipitation observations: A preliminary assessment. *Journal of Hydrometeorology*, 18(4), 1033–1050. <https://doi.org/10.1175/JHM-D-16-0187.1>
- Ba, M. B., & Gruber, A. (2001). GOES multispectral rainfall algorithm (GMSRA). *Journal of Applied Meteorology*, 40(8), 1500–1514. [https://doi.org/10.1175/1520-0450\(2001\)040<1500:GMRAG>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1500:GMRAG>2.0.CO;2)
- Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., & Dozier, J. (2006). Mountain hydrology of the western United States. *Water Resources Research*, 42, W084325. <https://doi.org/10.1029/2005WR004387>
- Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & et al. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21(1), 589–615. <https://doi.org/10.5194/hess-21-589-2017>
- Behrangi, A., Christensen, M., Richardson, M., Lebsock, M., Stephens, G., Huffman, G. J., et al. (2016). Status of high-latitude precipitation estimates from observations and reanalyses. *Journal of Geophysical Research: Atmospheres*, 121, 4468–4486. <https://doi.org/10.1002/2015JD024546>

- Behrangi, A., Hsu, K.-L., Imam, B., Sorooshian, S., Huffman, G. J., & Kuligowski, R. J. (2009). PERSIANN-MSA: A precipitation estimation method from satellite-based multispectral analysis. *Journal of Hydrometeorology*, 10(6), 1414–1429. <https://doi.org/10.1175/2009JHM1139.1>
- Behrangi, A., Lebsack, M., Wong, S., & Lambrechts, B. (2012). On the quantification of oceanic rainfall using spaceborne sensors. *Journal of Geophysical Research*, 117, D20105. <https://doi.org/10.1029/2012JD017979>
- Behrangi, A., Nguyen, H., Lambrechts, B., Schreier, M., & Dang, V. (2015). Investigating the role of multi-spectral and near surface temperature and humidity data to improve precipitation detection at high latitudes. *Atmospheric Research*, 163, 2–12. <https://doi.org/10.1016/j.atmosres.2014.10.019>
- Behrangi, A., Tian, Y., Lambrechts, B. H., & Stephens, G. L. (2014). What does CloudSat reveal about global land precipitation detection by other spaceborne sensors? *Water Resources Research*, 50, 4893–4905. <https://doi.org/10.1002/2013WR014566>
- Behrangi, A., Yin, X., Rajagopal, S., Stampoulis, D., & Ye, H. (2018). On distinguishing snowfall from rainfall using near-surface atmospheric information: Comparative analysis, uncertainties, and hydrologic importance. *Quarterly Journal of the Royal Meteorological Society*. <https://doi.org/10.1002/qj.3240>
- Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, 2(1), 1–127. <https://doi.org/10.1561/2200000006>
- Boening, C., Lebsack, M., Landerer, F., & Stephens, G. (2012). Snowfall-driven mass change on the East Antarctic ice sheet. *Geophysical Research Letters*, 39, L21501. <https://doi.org/10.1029/2012GL053316>
- Cao, Q., Hong, Y., Chen, S., Gourley, J. J., Zhang, J., & Kirstetter, P. E. (2014). Snowfall detectability of NASA's cloudsat: The first cross-investigation of its 2C-snow-profile product and national multi-sensor mosaic QPE (NMQ) snowfall data. *Progress in Electromagnetics Research*, 148, 55–61. <https://doi.org/10.2528/PIER14030405>
- Chen, X., Long, D., Hong, Y., Zeng, C., & Yan, D. H. (2017). Improved modeling of snow and glacier melting by a progressive two-stage calibration strategy with GRACE and multisource data: How snow and glacier meltwater contributes to the runoff of the upper Brahmaputra River basin? *Water Resources Research*, 53, 2431–2466. <https://doi.org/10.1002/2016WR019656>
- Chen, X., Long, D., Liang, S., He, L., Zeng, C., Hao, X., & et al. (2018). Developing a composite daily snow cover extent record over the Tibetan plateau from 1981 to 2016 using multisource data. *Remote Sensing of Environment*, 215, 284–299. <https://doi.org/10.1016/j.rse.2018.06.021>
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. <https://doi.org/10.1002/qj.828>
- Fang, K., Shen, C., Kifer, D., & Yang, X. (2017). Prolongation of SMAP to spatiotemporally seamless coverage of continental U.S. using a deep learning neural network. *Geophysical Research Letters*, 44, 11,030–011,039. <https://doi.org/10.1002/2017GL075619>
- Gao, J., Tang, G., & Hong, Y. (2017). Similarities and improvements of GPM dual-frequency precipitation radar (DPR) upon TRMM precipitation radar (PR) in global precipitation rate estimation, type classification and vertical profiling. *Remote Sensing*, 9(11), 1142. <https://doi.org/10.3390/rs9111142>
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30(14), 5419–5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>
- Hamshaw, S. D., Dewoolkar, M. M., Schroth, A. W., Wemple, B. C., & Rizzo, D. M. (2018). A new machine-learning approach for classifying hysteresis in suspended-sediment discharge relationships using high-frequency monitoring data. *Water Resources Research*, 54, 4040–4058. <https://doi.org/10.1029/2017WR022238>
- Hao, Y., & Wilamowski, B. M. (2011). Levenberg–Marquardt training. In *The Industrial Electronics Handbook* (chap. 12, pp. 12–1–12–15). CRC Press.
- Harpold, A. A., Kaplan, M. L., Klos, P. Z., Link, T., McNamara, J. P., Rajagopal, S., et al. (2017). Rain or snow: Hydrologic processes, observations, prediction, and research needs. *Hydrology and Earth System Sciences*, 21(1), 1–22. <https://doi.org/10.5194/hess-21-1-2017>
- Haynes, J. M., L'Ecuier, T. S., Stephens, G. L., Miller, S. D., Mitrescu, C., Wood, N. B., & et al. (2009). Rainfall retrieval over the ocean with spaceborne W-band radar. *Journal of Geophysical Research*, 114, D00A22. <https://doi.org/10.1029/2008JD009973>
- Hong, Y., Gochis, D., Cheng, J.-t., Hsu, K.-l., & Sorooshian, S. (2007). Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. *Journal of Hydrometeorology*, 8(3), 469–482. <https://doi.org/10.1175/JHM574.1>
- Hong, Y., Hsu, K. L., Sorooshian, S., & Gao, X. G. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology*, 43(12), 1834–1853. <https://doi.org/10.1175/JAM2173.1>
- Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., et al. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, 95(5), 701–722. <https://doi.org/10.1175/BAMS-D-13-00164.1>
- Hsu, K.-l., Gao, X., Sorooshian, S., & Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, 36(9), 1176–1190. [https://doi.org/10.1175/1520-0450\(1997\)036<1176:PEFRS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRS>2.0.CO;2)
- Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., et al. (2017). *NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG)*. Greenbelt, MD, USA: Algorithm Theoretical Basis Document (ATBD), NASA/GSFC. https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.1b.pdf
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., et al. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1), 38–55. <https://doi.org/10.1175/JHM560.1>
- Iguchi, T., Seto, S., Meneghini, R., Yoshida, N., Awaka, J., & Kubota, T. (2010). GPM/DPR level-2 algorithm theoretical basis document, NASA Goddard Space Flight Center, Greenbelt, MD, USA, Tech. Rep.
- Immerzeel, W. W., van Beek, L. P., & Bierkens, M. F. (2010). Climate change will affect the Asian water towers. *Science*, 328(5984), 1382–1385. <https://doi.org/10.1126/science.1183188>
- Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5(3), 487–503. [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2)
- Karbalaee, N., Hsu, K., Sorooshian, S., & Braithwaite, D. (2017). Bias adjustment of infrared-based rainfall estimation using passive microwave satellite rainfall data. *Journal of Geophysical Research: Atmospheres*, 122, 3859–3876. <https://doi.org/10.1002/2016JD026037>
- Kidd, C., & Huffman, G. (2011). Global precipitation measurement. *Meteorological Applications*, 18(3), 334–353. <https://doi.org/10.1002/met.284>
- Kidd, C., Kniveton, D. R., Todd, M. C., & Bellerby, T. J. (2003). Satellite rainfall estimation using combined passive microwave and infrared algorithms. *Journal of Hydrometeorology*, 4(6), 1088–1104. [https://doi.org/10.1175/1525-7541\(2003\)004<1088:SREUCP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1088:SREUCP>2.0.CO;2)

- Kidd, C., Matsui, T., Chern, J., Mohr, K., Kummerow, C., & Randel, D. (2016). Global precipitation estimates from cross-track passive microwave observations using a physically based retrieval scheme. *Journal of Hydrometeorology*, 17(1), 383–400. <https://doi.org/10.1175/JHM-D-15-0051.1>
- Kongoli, C., Meng, H., Dong, J., & Ferraro, R. (2015). A snowfall detection algorithm over land utilizing high-frequency passive microwave measurements-application to ATMS. *Journal of Geophysical Research: Atmospheres*, 120, 1918–1932. <https://doi.org/10.1002/2014JD022427>
- Kongoli, C., Pellegrino, P., Ferraro, R. R., Grody, N. C., & Meng, H. (2003). A new snowfall detection algorithm over land using measurements from the advanced microwave sounding unit (AMSU). *Geophysical Research Letters*, 30(14), 1756. <https://doi.org/10.1029/2003GL017177>
- Kummerow, C., Hong, Y., Olson, W. S., Yang, S., Adler, R. F., McCollum, J., et al. (2001). The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors. *Journal of Applied Meteorology*, 40(11), 1801–1820. [https://doi.org/10.1175/1520-0450\(2001\)040<1801:TEOTGP>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2)
- Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N.-Y., Ferraro, R., Joseph Munchak, S., & et al. (2015). The evolution of the Goddard profiling algorithm to a fully parametric scheme. *Journal of Atmospheric and Oceanic Technology*, 32(12), 2265–2280. <https://doi.org/10.1175/JTECH-D-15-0039.1>
- Kurino, T. (1997). A satellite infrared technique for estimating “deep/shallow” precipitation. *Advances in Space Research*, 19(3), 511–514. [https://doi.org/10.1016/S0273-1177\(97\)00063-X](https://doi.org/10.1016/S0273-1177(97)00063-X)
- Laloy, E., Hérault, R., Jacques, D., & Linde, N. (2018). Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resources Research*, 54, 381–406. <https://doi.org/10.1002/2017WR022148>
- Lebsack, M. D., & L'Ecuier, T. S. (2011). The retrieval of warm rain from CloudSat. *Journal of Geophysical Research*, 116, D20209. <https://doi.org/10.1029/2011JD016076>
- L'Ecuier, T. S., & Stephens, G. L. (2002). An estimation-based precipitation retrieval algorithm for attenuating radars. *Journal of Applied Meteorology*, 41(3), 272–285. [https://doi.org/10.1175/1520-0450\(2002\)041<0272:AEBPRA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2002)041<0272:AEBPRA>2.0.CO;2)
- Li, C., Tang, G., & Hong, Y. (2018). Cross-evaluation of ground-based, multi-satellite and reanalysis precipitation products: Applicability of the triple collocation method across mainland China. *Journal of Hydrology*, 562, 71–83. <https://doi.org/10.1016/j.jhydrol.2018.04.039>
- Liu, G. (2008). Deriving snow cloud characteristics from CloudSat observations. *Journal of Geophysical Research*, 113, D00A09. <https://doi.org/10.1029/2007JD009766>
- Liu, G., & Seo, E.-K. (2013). Detecting snowfall over land by satellite high-frequency microwave observations: The lack of scattering signature and a statistical approach. *Journal of Geophysical Research: Atmospheres*, 118, 1376–1387. <https://doi.org/10.1002/jgrd.50172>
- Long, D., Pan, Y., Zhou, J., Chen, Y., Hou, X. Y., Hong, Y., et al. (2017). Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models. *Remote Sensing of Environment*, 192, 198–216. <https://doi.org/10.1016/j.rse.2017.02.011>
- McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., & Stoffelen, A. (2014). Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. *Geophysical Research Letters*, 41, 6229–6236. <https://doi.org/10.1002/2014GL061322>
- Miller, J. (1972). A climatological Z–R relationship for convective storms in the northern Great Plains, paper presented at 15th Conf. on Radar Meteor.
- Mitrescu, C., L'Ecuier, T., Haynes, J., Miller, S., & Turk, J. (2010). CloudSatPrecipitation profiling algorithm—Model description. *Journal of Applied Meteorology and Climatology*, 49(5), 991–1003. <https://doi.org/10.1175/2009JAMC2181.1>
- More, J. J. (1978). The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*, (pp. 105–116). Berlin, Heidelberg: Springer.
- Noh, Y.-J., Liu, G., Jones, A. S., & Vonder Haar, T. H. (2009). Toward snowfall retrieval over land by combining satellite and in situ measurements. *Journal of Geophysical Research*, 114, D24205. <https://doi.org/10.1029/2009JD012307>
- Palmer, C., Claud, C., Dufour, A., Genthon, C., Wood, N. B., & L'Ecuier, T. (2017). Evaluation of Antarctic snowfall in global meteorological reanalyses. *Atmospheric Research*, 190, 104–112. <https://doi.org/10.1016/j.atmosres.2017.02.015>
- Pan, M., Fisher, C. K., Chaney, N. W., Zhan, W., Crow, W. T., Aires, F., et al. (2015). Triple collocation: Beyond three estimates and separation of structural/non-structural errors. *Remote Sensing of Environment*, 171, 299–310. <https://doi.org/10.1016/j.rse.2015.10.028>
- Sahoo, S., Russo, T. A., Elliott, J., & Foster, I. (2017). Machine learning algorithms for modeling groundwater level changes in agricultural regions of the U. S. *Water Resources Research*, 53, 3878–3895. <https://doi.org/10.1002/2016WR019933>
- Serreze, M. C., Barrett, A. P., & Lo, F. (2005). Northern high-latitude precipitation as depicted by atmospheric reanalyses and satellite retrievals. *Monthly Weather Review*, 133(12), 3407–3430. <https://doi.org/10.1175/MWR3047.1>
- Sims, E. M., & Liu, G. (2015). A parameterization of the probability of snow–rain transition. *Journal of Hydrometeorology*, 16(4), 1466–1477. <https://doi.org/10.1175/JHM-D-14-0211.1>
- Skofronick-Jackson, G., Hudak, D., Petersen, W., Nesbitt, S. W., Chandrasekar, V., Durden, S., et al. (2015). Global precipitation measurement cold season precipitation experiment (GCPEX): For Measurement's sake, let it snow. *Bulletin of the American Meteorological Society*, 96(10), 1719–1741. <https://doi.org/10.1175/BAMS-D-13-00262.1>
- Skofronick-Jackson, G. M., Min-Jeong, K., Weinman, J. A., & Dong-Eon, C. (2004). A physical model to determine snowfall over land by microwave radiometry. *IEEE Transactions on Geoscience and Remote Sensing*, 42(5), 1047–1058. <https://doi.org/10.1109/TGRS.2004.825585>
- Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, 81(9), 2035–2046. [https://doi.org/10.1175/1520-0477\(2000\)081<2035:EOPSS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSS>2.3.CO;2)
- Stoffelen, A. (1998). Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research*, 103(C4), 7755–7766. <https://doi.org/10.1029/97JC03180>
- Tang, G., Behrangi, A., Long, D., Li, C., & Hong, Y. (2018). Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products. *Journal of Hydrology*, 559, 294–306. <https://doi.org/10.1016/j.jhydrol.2018.02.057>
- Tang, G., Long, D., Hong, Y., Gao, J., & Wan, W. (2018). Documentation of multifactorial relationships between precipitation and topography of the Tibetan plateau using spaceborne precipitation radars. *Remote Sensing of Environment*, 208, 82–96. <https://doi.org/10.1016/j.rse.2018.02.007>
- Tang, G., Ma, Y., Long, D., Zhong, L., & Hong, Y. (2016). Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over mainland China at multiple spatiotemporal scales. *Journal of Hydrology*, 533, 152–167. <https://doi.org/10.1016/j.jhydrol.2015.12.008>
- Tang, G., Wen, Y., Gao, J., Long, D., Ma, Y., Wan, W., & et al. (2017). Similarities and differences between three coexisting spaceborne radars in global rainfall and snowfall estimation. *Water Resources Research*, 53, 3835–3853. <https://doi.org/10.1002/2016WR019961>

- Tao, Y., Gao, X., Hsu, K., Sorooshian, S., & Ihler, A. (2016). A deep neural network modeling framework to reduce bias in satellite precipitation products. *Journal of Hydrometeorology*, 17(3), 931–945. <https://doi.org/10.1175/JHM-D-15-0075.1>
- Tao, Y., Gao, X., Ihler, A., Sorooshian, S., & Hsu, K. (2017). Precipitation identification with Bispectral satellite information using deep learning approaches. *Journal of Hydrometeorology*, 18(5), 1271–1283. <https://doi.org/10.1175/JHM-D-16-0176.1>
- Tao, Y., Hsu, K., Ihler, A., Gao, X., & Sorooshian, S. (2018). A two-stage deep neural network framework for precipitation estimation from Bispectral satellite information. *Journal of Hydrometeorology*, 19(2), 393–408. <https://doi.org/10.1175/JHM-D-17-0077.1>
- Tapiador, F., Kidd, C., Hsu, K.-L., & Marzano, F. (2004). Neural networks in satellite rainfall estimation. *Meteorological Applications*, 11(1), 83–91. <https://doi.org/10.1017/S1350482704001173>
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, 106(D7), 7183–7192. <https://doi.org/10.1029/2000JD900719>
- Turk, J. (2017). CloudSat-GPM coincidence dataset, *NASA Technical Report* (accessed on 26 March 2018).
- Wood, N. B., L'Ecuyer, T. S., Bliven, F. L., & Stephens, G. L. (2013). Characterization of video disdrometer uncertainties and impacts on estimates of snowfall rate and radar reflectivity. *Atmospheric Measurement Techniques*, 6(12), 3635–3648. <https://doi.org/10.5194/amt-6-3635-2013>
- Wood, N. B., L'Ecuyer, T. S., Heymsfield, A. J., Stephens, G. L., Hudak, D. R., & Rodriguez, P. (2014). Estimating snow microphysical properties using collocated multisensor observations. *Journal of Geophysical Research: Atmospheres*, 119, 8941–8961. <https://doi.org/10.1002/2013JD021303>
- Yang, Z., Hsu, K., Sorooshian, S., Xu, X., Braithwaite, D., & Verbist, K. M. J. (2016). Bias adjustment of satellite-based precipitation estimation using gauge observations: A case study in Chile. *Journal of Geophysical Research: Atmospheres*, 121, 3790–3806. <https://doi.org/10.1002/2015JD024540>
- Yong, B., Chen, B., Gourley, J. J., Ren, L., Hong, Y., Chen, X., et al. (2014). Intercomparison of the version-6 and version-7 TMPA precipitation products over high and low latitudes basins with independent gauge networks: Is the newer version better in both real-time and post-real-time analysis for water resources and hydrologic extremes? *Journal of Hydrology*, 508, 77–87.
- Zhang, D., Lindholm, G., & Ratnaweera, H. (2018). Use long short-term memory to enhance internet of things for combined sewer overflow monitoring. *Journal of Hydrology*, 556, 409–418. <https://doi.org/10.1016/j.jhydrol.2017.11.018>
- Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., et al. (2016). Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bulletin of the American Meteorological Society*, 97(4), 621–638. <https://doi.org/10.1175/BAMS-D-14-00174.1>