Validation of GMI Snowfall Observations by Using a Combination of Weather Radar and Surface Measurements

ANNAKAISA VON LERBER,a,b DMITRI MOISSEEV,c,a DAVID A. MARKS,d WALTER PETERSEN,e ARI-MATTI HARRI,a AND V. CHANDRASEKARf,a,c

a Finnish Meteorological Institute, Helsinki, Finland
b School of Electrical Engineering, Aalto University, Finland
c University of Helsinki, Helsinki, Finland
d NASA GSFC Wallops Flight Facility, Wallops Island, Virginia
e NASA MSFC Earth Science Office, Huntsville, Alabama
f Colorado State University, Fort Collins, Colorado

(Manuscript received 22 June 2017, in final form 27 October 2017)

ABSTRACT

Currently, there are several spaceborne microwave instruments suitable for the detection and quantitative estimation of snowfall. To test and improve retrieval snowfall algorithms, ground validation datasets that combine detailed characterization of snowfall microphysics and spatial precipitation measurements are required. To this endpoint, measurements of snow microphysics are combined with large-scale weather radar observations to generate such a dataset. The quantitative snowfall estimates are computed by applying event-specific relations between the equivalent reflectivity factor and snowfall rate to weather radar observations. The relations are derived using retrieved ice particle microphysical properties from observations that were carried out at the University of Helsinki research station in Hyytiälä, Finland, which is about 64 km east of the radar. For each event, the uncertainties of the estimate are also determined. The feasibility of using this type of data to validate spaceborne snowfall measurements and algorithms is demonstrated with the NASA GPM Microwave Imager (GMI) snowfall product. The detection skill and retrieved surface snowfall precipitation of the GPROF detection algorithm, versions V04A and V05A, are assessed over southern Finland. On the basis of the 26 studied overpasses, probability of detection (POD) is 0.90 for version V04A and 0.84 for version V05A, and corresponding false-alarm rates are 0.09 and 0.10, respectively. A clear dependence of detection skill on cloud echo top height is shown: POD increased from 0.8 to 0.99 (V04A) and from 0.61 to 0.94 (V05A) as the cloud echo top altitude increased from 2 to 5 km. Both versions underestimate the snowfall rate by factors of 6 (V04A) and 3 (V05A).

1. Introduction

At high latitudes and polar regions, more than 25% of precipitation falls in the form of snow (Field and Heymsfield 2015). In these regions, snow has a significant environmental, economic, and societal impact through, for example, spring floods and challenges in ground transportation and aviation (Juga et al. 2012; Venäläinen and Kangas 2003). Furthermore, to improve our understanding of the global hydrological circulation and its connection to Earth’s energy budget, accurate global observations of snowfall are needed (Kulie and Bennartz 2009; Hou et al. 2014).

Satellite-based remote sensing observations provide the needed coverage for global snowfall observations. Ground measurements, such as networks of precipitation gauges and ground-based radars, are both spatially and temporally limited, especially for rural areas and oceans (Kidd and Levizzani 2011). Even if the technology and algorithms for space-based observations of moderate and heavy rain are more or less established now, the snowfall measurements are still challenging for both active and passive instruments. This limitation arises because of weak and complex signatures of snow at microwave frequencies (e.g., Kidd and Huffman 2011; Ferraro et al. 2005; Liu 2008; Liu and Seo 2013; Skofronick-Jackson et al. 2013; Munchak and Skofronick-Jackson 2013). However, over the last 20 years, significant improvements in the development of satellite microwave snowfall algorithms have taken

Corresponding author: Annakaisa von Lerber, annakaisa.von.lerber@fmi.fi

DOI: 10.1175/JAMC-D-17-0176.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
place (e.g., Liu and Curry 1997; Staelin and Chen 2000; Kongoli et al. 2003; Skofronick-Jackson et al. 2004; Noh et al. 2006; Kim et al. 2008; Marchand et al. 2008; Noh et al. 2009; Haynes et al. 2009; Liu and Seo 2013; Kongoli et al. 2015; Kummerow et al. 2015). Also, instruments suitable for observing snowfall have been launched with various satellites: for example, Advanced Microwave Sounding Unit B (AMSU-B) on NOAA-15, NOAA-16, and NOAA-17; Microwave Humidity Sounder (MHS) on NOAA-18 and NOAA-19, MetOp-A, and MetOp-B; Advanced Technology Microwave Sounder (ATMS) on Suomi National Polar-Orbiting Partnership (SNPP); and Cloud Profiling Radar (CPR) on CloudSat. Most recently, the Core Observatory satellite of the Global Precipitation Measurement (GPM) mission was launched in 2014 with the multichannel microwave radiometer GPM Microwave Imager (GMI) and the Dual-Frequency Precipitation Radar (DPR) (Hou et al. 2014).

Passive instruments measure electromagnetic radiation emitted and scattered by Earth’s surface and atmospheric particles integrated over a vertical column. Snowfall is identified from the decrease of radiation caused by the scattering of ice particles (Kidd and Huffman 2011), which is more prominent at high-frequency channels above 100 GHz (Bennartz and Bauer 2003; Skofronick-Jackson et al. 2004; Kongoli et al. 2015; You et al. 2017). However, one of the challenges in detecting snowfall is the emission signal of supercooled liquid water, which can increase the detected radiation and thus weaken the depression caused by the snow particles, as stated in You et al. (2017) and the references therein. Passive snowfall measurements can also be masked by the ground emissivity, especially from the snow-covered surfaces (Munchak and Skofronick-Jackson 2013), but at higher frequencies, the surface emission component should be screened by the water vapor absorption (Skofronick-Jackson et al. 2004; You et al. 2017). Skofronick-Jackson et al. (2013) concluded, based on their sensitivity study, that the snow detection threshold for a passive radiometer utilizing the frequency channel of 166 GHz would realistically be close to 0.5–1.0 mm h$^{-1}$ and in ideal conditions around 0.2 mm h$^{-1}$. This is confirmed by the observations of Munchak and Skofronick-Jackson (2013) showing that the detectable precipitation rate over the snow-covered surface was 0.89 mm h$^{-1}$, and adding a priori knowledge of the snow cover combined with surface emissivity database, this could be reduced to 0.44 mm h$^{-1}$.

Active sensors not only provide observations of the vertical structure of clouds and precipitation, but also, the measurements are carried out at higher spatial resolution if compared with passive sensors. There are still relatively few space-based active sensors that are suitable for snowfall detection. The 94-GHz nadir-looking CPR on CloudSat (Stephens et al. 2008) has been utilized for snowfall surface measurements (e.g., Hudak et al. 2008; Liu 2008; Kulie and Bennartz 2009; Hiley et al. 2011; Behrangi et al. 2012, 2014; Norin et al. 2015; Kulie et al. 2016; Chen et al. 2016). The sensitivity of CPR with minimum detectable signal level of $<30\text{dBZ}$ (Tanelli et al. 2008) is high enough to detect light precipitation including drizzle and snowfall (Kulie and Bennartz 2009). The polar-orbiting satellite CloudSat covers the high-latitude and polar regions, but because of the narrow swaths, its utilization for climate monitoring is somewhat limited (Liu and Seo 2013; Maahn et al. 2014). The disadvantages of the relatively high frequency of CPR is the significant attenuation of the signal and the issue of multiple scattering, which are mainly affecting observations of intense snowfall (Hudak et al. 2008). Active sensors also suffer from surface clutter contamination, and thus, the closest clutter-free observations to the ground can be as high as 1.3 km above the surface (Kulie and Bennartz 2009). The DPR on board of GPM Core Observatory has a minimum detectable signal level between 12 and 14 dBZ for the Ku-band (13.6 GHz) radar (Toyoshima et al. 2015) and 12 dBZ for the Ka-band (35 GHz) radar (Hou et al. 2014), which are high for detecting the majority of the snowfall events (Kulie and Bennartz 2009).

The refinement and testing of the space-based algorithms require high-quality ground validation datasets in snowing environments (Hiley et al. 2011; Skofronick-Jackson et al. 2013, 2015). However, the ground-based measurements are limited by measurement errors and retrieval uncertainties. Gauge measurements are often treated as the ground truth, although gauges are known to suffer from marked undercatchment in solid precipitation, the value of which depends on wind speed and used wind shield (Goodison et al. 1998). Snowfall estimates based on ground radars are affected by uncertainties in the relation between equivalent reflectivity factor $Z_e$ measurements and snowfall rate $S$ (Rasmussen et al. 2003). Ground radar measurements are challenged by overshooting shallow precipitation events, ground clutter, and beam broadening with the increasing distance from the radar (Norin et al. 2015).

Smalley et al. (2014) compared an hourly precipitation accumulation of the NCEP Stage IV product in a 4.7-km-sized grid with the CloudSat estimated snowfall rate. The Stage IV product is constructed from a combination of the U.S. National Weather Surveillance Radar-1988 Doppler (WSR-88D) network of ground radars and surface gauges. Norin et al. (2015) compared the snowfall estimates of CPR over the Swedish ground-based radar network (Swerad) for winter observations.
during 2008–10. The intercomparison showed that observations of precipitation rate in the range of 0.01–1.0 mm h\(^{-1}\) with both instruments are similar, and the best agreement is achieved when measurement volumes are located 46–82 km from the nearest radar. The skill scores, probability of detection (POD) and false-alarm rate (FAR), demonstrated that both instruments have difficulties with shallow snow events: CPR because of the filtered lower bins to avoid ground clutter and Swerad because of the lower sensitivity at larger distances from the radar and vertical profile of reflectivity effects (Koistinen and Pohjola 2014). The performance of CPR retrievals for detection and estimation of snowfall is also compared with the Multi-Radar Multi-Sensor (MRMS) system quantitative precipitation estimation (QPE, hereinafter Q3) (Chen et al. 2016). The Q3 precipitation product is presented in Zhang et al. (2011). In Q3, if the dry surface temperature is below 2°C and the wet-bulb temperature is below 0°C, the precipitation is classified as snow, and \(Z_e = 75S^{2.0}\) is applied for snowfall estimate (Zhang et al. 2016). Because of the S-band radars in the NEXRAD network and the quality-control threshold of 5 dBZ, light snowfall is not detected by the Q3 (Chen et al. 2016).

In this study, point measurements of snow microphysics are combined with large-scale weather radar observations to generate a dataset suitable for ground validation of satellite products. The key point of this study is the derivation of the event-specific \(Z_e-S\) relations from video disdrometer and gauge observations (von Lerber et al. 2017). Since the event-specific \(Z_e-S\) relations are best suited for estimating storm total snowfall accumulation, and not the instantaneous precipitation rate estimates needed for the comparison with satellite observations, uncertainties of using \(Z_e-S\) relations are also computed. The method to derive these uncertainties is introduced and discussed in section 3c. Using the radar-based snowfall rates and associated uncertainties, the space-based snowfall estimates of the GMI Goddard profiling algorithm (GPROF), versions V04A and V05A, are evaluated using 26 GPM overpasses in southern Finland. Both the detection skill and quantitative estimation of surface snowfall rate are compared. The detection skills, critical success index (CSI), FAR, and POD are studied as a function of cloud echo top and snow depth on the ground.

2. Measurement setup and data

a. Surface observations of Hyytiälä measurement site

The landscape in southern Finland can be classified as lowland, mostly lying below 200 m above mean sea level (MSL) (Tikkanen 1994), and there are no large variations in topography. The region is hilly, forest-covered land with marshland and ridges; typically in Finland about 10% of the surface area is covered by lakes. The Hyytiälä Forestry Field Station of the University of Helsinki is located 230 km north of Helsinki (61.845°N, 24.287°E; 150 m MSL). The measurement site is sheltered by the forest. Hence, the local wind conditions are moderate, and the challenges of blowing snow or undercatchment often encountered in snow measurements (Goodison et al. 1998) are not usually severe. The event-specific \(Z_e-S\) relations utilized in this study are obtained by combining observations of a video disdrometer with a collocated precipitation gauge. The video disdrometer, particle imaging package (PIP), and weighing gauge, OTT Hydromet GmbH Pluvio\(^2\) 200, are deployed at this site as a part of the National Aeronautics and Space Administration (NASA) GPM ground validation (GV) program. The gauge is placed inside a wind fence similar to WMO standard double fence intercomparison reference (DFIR) wind protection (Goodison et al. 1998), and the PIP is located 10 m away from the fence (Petijä et al. 2016).

The PIP provides particle size distribution (PSD), fall velocity, and particle shape measurements. It is an optical disdrometer, constructed from a video camera with a high frame rate (380 frames per second) and lamp providing background light (Newman et al. 2009; Tiira et al. 2016). As particles fall between the lamp and the camera, 2D-shadowed images are recorded. From consecutive image frames, fall velocity is computed. The fall velocity–dimensional \(v(D)\) relation is fitted to the dataset (Tiira et al. 2016). Here, the factors \(a_v\) and \(b_v\) of the power-law format \((v = a_vD^{b_v})\) are given as a function of a maximum diameter \(D\), which is the diameter of a circumscribing circle around the particle.

The other particle characteristics measured with PIP include, for example, disk-equivalent diameter (diameter of a disk with the area of the particle image), total area as an area of shadowed pixels \((A_{tot})\), bounding box dimensions, and particle orientation with respect to the horizontal axis. To estimate the maximum dimension \(D\) of the particle, an ellipse is fitted inside a given bounding box with a given orientation. The area ratio is defined as a ratio of the measured total area of the shadowed pixels and area of the circumscribing circle, that is, \(A_r = A_{tot}/(\pi D^2/4)\), and is limited to be smaller or equal to 1. Given multiple observations for a single particle, mean values of particle parameters are computed, except for \(D\), where the largest recorded value is always used.

PSD is calculated every minute as a function of disk-equivalent diameter \(D_{deq}\). Here, version 1403 of the data-processing software is used, and in this version, the
Table 1. AWS gauges with WMO identification number (ID), internal FMI code, locations [in World Geodetic System 1984 (WGS84)], distances to Ikaalinen weather radar, and estimated beam height over the gauge with 0.3° elevation angle.

<table>
<thead>
<tr>
<th>WMO ID</th>
<th>FMI code</th>
<th>Site</th>
<th>Lat</th>
<th>Lon</th>
<th>Distance (km)</th>
<th>Height (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>02937</td>
<td>1140</td>
<td>Kokemäki, Tulkila</td>
<td>61.253°N</td>
<td>22.350°E</td>
<td>69.3</td>
<td>784</td>
</tr>
<tr>
<td>02763</td>
<td>1222</td>
<td>Tampere, Härmälä</td>
<td>61.466°N</td>
<td>23.750°E</td>
<td>48.9</td>
<td>477</td>
</tr>
<tr>
<td>02708</td>
<td>2103</td>
<td>Karvia, Alikka</td>
<td>62.183°N</td>
<td>22.803°E</td>
<td>48.6</td>
<td>402</td>
</tr>
<tr>
<td>02753</td>
<td>2123</td>
<td>Kankaanpää, Niinisalo</td>
<td>61.838°N</td>
<td>22.467°E</td>
<td>33.3</td>
<td>280</td>
</tr>
<tr>
<td>02735</td>
<td>2218</td>
<td>Virrat, Äjänneva</td>
<td>62.328°N</td>
<td>23.546°E</td>
<td>67.1</td>
<td>657</td>
</tr>
<tr>
<td>02770</td>
<td>2303</td>
<td>Juupajoki, Hyytiälä</td>
<td>61.845°N</td>
<td>24.291°E</td>
<td>64.4</td>
<td>605</td>
</tr>
</tbody>
</table>

diameter bin size is 0.2 mm. PSD is divided into 120 bins, and the highest bin value is for particles with a diameter larger than 26.0 mm. The resolution of PIP is 0.1 mm, and the field of view is 48 mm × 64 mm at the focal length. In practice, the smallest \( D_{\text{deq}} \) diameter used in the computations is \( \approx 0.2 \) mm. The measurement volume of PIP is described in Newman et al. (2009). The sampling uncertainty of PSD and its influence on the defined \( Z - S \) relation are discussed in section 3a and in von Lerber et al. (2017). For computing the \( Z - S \) relation described in section 3a, the mass–dimensional \( m(D) \) and \( u(D) \) relations are given as function of the maximum diameter \( D \) and PSD as function of the \( D_{\text{deq}} \). Therefore, for each studied snow event, the linear ratio between the observed \( D \) and the \( D_{\text{deq}} \) is defined: the scale factors of the linear relations are deviating between 1.28 and 1.49, and the mean value is 1.38.

The weighing gauge, Pluvio\(^2\) 200, inside DFIR is protected by an additional Tretyakov wind shield. The measurement orifice is 200 cm\(^2\). The gauge provides liquid water equivalent (LWE) accumulation measurement every minute, and for comparing with PIP-estimated accumulation, the 1-min observations are summed for the whole event.

b. **FMI operational precipitation gauges and snow depth sensors**

The precipitation gauges and snow depth sensors of the automatic weather station (AWS) network of the Finnish Meteorological Institute (FMI) utilized in this study are located within 30–70 km from the Ikaalinen weather radar. The list of gauges with their locations and the estimated height of the radar beam (elevation angle 0.3°) at the gauge location calculated assuming a 4/3 Earth radius model (Doviak and Zrnić 1993) is given in Table 1. These gauges are also OTT Pluvio\(^2\) model but with a larger orifice of 400 cm\(^2\). They have Alter wind shields, and the sensor orifices are set at 1.5 m above ground level. The average precipitation rate every 10 min is utilized in the validation of radar-based precipitation estimates and summed when the LWE accumulation of the whole event is studied. Next to the gauges, the snow depth is measured by an ultrasonic snow depth sensor and given as a point measurement every 10 min.

c. **FMI Ikaalinen weather radar**

The Ikaalinen weather radar (IKA) is located 64 km west of Hyytiälä, at 61.7673°N, 23.0764°E. The radar antenna height on a water tower is 153 m MSL. It is a dual-polarization Doppler radar operating at 5.5 GHz. It is one of the 10 radars of FMI’s operational weather radar network (Saltikoff et al. 2010). The minimum detectable \( Z_e \) of IKA is \(-46.81\) dBZ at 1-km range, which corresponds approximately to minimum detectable precipitation rate of 0.01 mm h\(^{-1}\) at 80 km. It performs plan position indicator (PPI) scans every 5 min. The lowest elevation angle is 0.3°, and with this angle, there are five specified directions (156°, 207°, 237°, 241°, 271°, 273°, and 292°) with severe partial beam blockage; these are shown in Fig. 1. The clutter-contaminated radar bins in the vicinity of the radar are removed from the analysis by utilizing a stationary statistical ground clutter map (Lakshmanan et al. 2012). The clutter map for each elevation angle is defined from clear-air echoes during the period of 1 December 2014–1 February 2015 (Pulkkinen et al. 2016). Here, the threshold value of removing clutter pixels is set for observing a persistent echo over \(-5\) dB more than 60% of the time. Removing the pixels will result in holes in the reflectivity factor field. The error in the snowfall rate at the locations of the AWS gauges is estimated to be small as none of the gauges are located in clutter-contaminated areas, as shown in Fig. 1. In the comparison with the space-based snowfall rate, the radar-estimated snowfall rates are averaged spatially over four elevation angles, temporally over 20 min, and to a larger sampling grid of 15 km × 15 km to resemble the columnar observations of the GMI instrument. Therefore, clutter removal is not likely to produce a significant error in the retrieved radar snowfall estimates.

d. **Cloud echo top product**

Cloud echo top is computed from the volume scans of IKA (elevation angles between 0.3° and 45°) using
0 dBZ as the reflectivity factor threshold. The threshold is selected conservatively as the radar minimum detectable reflectivity at the edge of the studied area (at 80 km) is approximately −10 dBZ. The cloud echo top product describes the maximum altitude at which the threshold value is exceeded. The altitude is the top of the radar beam, which is defined from the half beamwidth above the nominal elevation angle. The values between the observed altitudes are linearly interpolated from the neighboring altitudes to a grid size of 0.6 km × 0.6 km. Given the multiple elevation angles, there are several altitude estimates to each grid point, and the used altitude is a quality-weighted mean value of the estimates.

e. Snow depth proxy

Snow depth proxy is estimated with the radar-based snowfall LWE accumulation adjusted by the AWS snow depth measurements. At the beginning of the event, the snow depth on the studied area is interpolated from the measured AWS snow depth values. Data from approximately 45 different stations are used with a distance of 200 km around IKA. There is a gap in the AWS observations on the west side of the area, and in this region, snow depth is not evaluated. Every 5 min, the LWE snowfall accumulation is derived with the consecutive PPI scans utilizing the event-specific $Z_e$–$S$ relation. This radar-estimated LWE accumulation is then converted to a change in snow depth according to the snow depth measurements of the five AWS stations showed in Table 1. Only the scan of lowest elevation angle (0.3°) is used, and therefore, some artifacts of the clutter contamination and partial beam blockage can be noticed. However, for studying the effect of a change in snow depth influencing the detection skill of the GMI snowfall product, the snow depth is averaged over the larger sampling grid, and this diminishes the underestimation caused by the clutter.

f. GPM Microwave Imager

GPM Core Observatory has an orbit with a 65° inclination angle covering mid- and high latitudes extending close to the Arctic and Antarctic Circles. The revisit time over southern Finland between northern latitudes of 60°–62° can be as frequent as every 1.5 h. The GMI is a conical-scanning, multifrequency passive radiometer. The cone sweeps with an off-nadir angle of 48.5° corresponding to an incidence angle of 52.8° on Earth’s surface and a swath width of 140° representing an arc of 885 km on the surface (Hou et al. 2014). GMI has 13 microwave frequency channels ranging from 10 to 183 GHz. The lower channels are more sensitive to liquid precipitation (10, 19, and 37 GHz) and water vapor emission (21 GHz) (Hou et al. 2014), while the higher-frequency bands above 89 GHz can be utilized in detecting scattering signals of ice particles (Bennartz and Bauer 2003; Skofronick-Jackson et al. 2013; Munchak and Skofronick-Jackson 2013). The footprint size is dependent on the used frequency; at 183 GHz, it is 4.4 km × 7.3 km, whereas at the lowest frequency of 10 GHz, it is 19.4 km × 32.2 km (Hou et al. 2014).

In this study, the sampling grid of 15 km × 15 km is chosen, which is an intermediate between the different footprint sizes.

The surface snowfall rate estimate of the GPM GMI instrument is retrieved with the GPROF. The GPROF is one of the physically based schemes to interpret the observed brightness temperature by utilizing a Bayesian database to determine the vertical cloud structure and thereby to estimate the surface precipitation (Kidd and Huffman 2011; Kummerow et al. 2015). For the presented analysis, four data fields are used: surface precipitation $S$ (mm h⁻¹), column water vapor index (mm), probability of precipitation (PoP; %), and surface class. In the latest version of GPM GPROF, the large a priori database of physical profiles is collected from the combined observations of GPM and NEXRAD radars and GMI radiometer (Kummerow et al. 2015). To reduce the computational time when searching for the suitable profiles from the database, the profiles have been clustered to subsets (Elsaesser and Kummerow 2015). The land surface temperature and total precipitable water are utilized as one of the parameters to determine the searched subset from the database (Kummerow et al. 2015).
and hence, the detection skills of GPROF are investigated here as a function of the column water vapor index. Similarly, given that the uncertainty of the snowfall rate is higher on the snow-covered surface (Munchak and Skofronick-Jackson 2013), the detection skills are studied relative to the surface class product. In this study, versions V04A and V05A of the GPROF algorithm are both investigated. The significant difference between the versions is that, for surface snowfall retrievals over snow-covered surfaces, in V05A, a priori database relies on ground estimates using the MRMS dataset (Zhang et al. 2011, 2016) with coincident radiometer observations of GPM constellation members (Kummerow et al. 2015). Also, in V05A, an intrinsic threshold based on the probability of precipitation is applied to determine the nonprecipitating pixels.

3. Method

The backbone of the proposed GV approach is the ground-based radar snowfall LWE rate estimate. To compute the S, two types of cases are identified for which different approaches to convert Z_r to S are used. In both cases, the power-law form \( Z_r = a_r Z_{eq}^{\alpha} \) is used. In the first case, a snowstorm moves over the measurement site, and microphysical properties of snow are considered to be well sampled. In this case, the event-specific factors of \( Z_r-S \) are applied to the PPI scans for an accurate estimate of the snowfall accumulation (von Lerber et al. 2017), but because, for the comparison with GMI observations, the instantaneous snowfall rate is needed, uncertainties associated with applying the \( Z_r-S \) relation have to be estimated. For this purpose, instantaneous \( Z_r-S \) relations are derived. Since the exponent of the relation does not vary much (Rasmussen et al. 2003; von Lerber et al. 2017), only the variation of the prefactor needs to be investigated. The derived distribution of the \( Z_r-S \) prefactor is then used to compute the uncertainty of the radar-based snowfall estimate. Kirstetter et al. (2015) presented a somewhat similar method to account for quantitative precipitation estimation uncertainties, where probability distributions of precipitation rates are determined for different precipitation types based on coinciding radar and gauge observations.

In the second case, the main part of the snowstorm (the highest accumulation) does not move over the measurement site, and only its fraction is sampled by the ground-based precipitation sensors. In this case, the derived \( Z_r-S \) relation is not representative, and an average relation computed from long-term observations is applied instead. The uncertainty of the precipitation estimate is derived in a similar manner to the event-specific relation: it is computed from the prefactor values of 5-min intervals during all the recorded events.

a. Event-specific \( Z_r-S \) relation

The detailed description for retrieving the event-specific and the instantaneous \( Z_r-S \) relations from PIP and gauge measurements can be found in von Lerber et al. (2017), and here, only a summary of the retrieval method is given.

The equivalent reflectivity factor in a Rayleigh regime can be stated as (Battan 1973)

\[
Z_r(t) = \frac{36 \times 10^8}{\pi^2 \rho_{ice}^2} \left| \frac{K_{r_{ice}}}{K_{r_w}} \right|^2 \int_{d_{min}}^{d_{max}} m(D, t)^2 N(D_{deq}, t) dD_{deq},
\]

where density of ice is \( \rho_{ice} = 0.917 \text{ g cm}^{-2} \), and at C band, the dielectric constant values of ice and water are \( |K_{r_{ice}}|^2 = 0.17 \) and \( |K_{r_w}|^2 = 0.93 \), respectively. PSD \( N(D_{deq}) \) is given in inverse meters cubed per millimeter for disk-equivalent diameter, and \( m(D) \) is given in grams for the maximum dimension and described as the power law \( m(D) = a_mD^{b_m} \). The integral limits are the minimum \( d_{min} \) and maximum \( d_{max} \) disk-equivalent diameters observed during the examined time period \( t \). The precipitation rate \( (\text{mm h}^{-1}) \) is

\[
S(t) = \frac{3.6}{\rho_w} \int_{d_{min}}^{d_{max}} m(D, t) v(D, t) N(D_{deq}, t) dD_{deq}.
\]

The difference between the diameter definitions and scaling factor is explained in section 2a. For each studied snow event, both \( Z_r \) and \( S \) are defined every 5 min with the observed PSD, the fitted \( v(D) \) relation, and the retrieved \( m(D) \) relation. It is assumed that, during the period of 5 min, the dominant particle type is the same and can be characterized by the single mass and velocity relations. The sample size is typically for every 5 min on the order of \( 10^9 \) particles; hence, the moment estimators of PSD can be applied with sufficient accuracy for the \( Z_r-S \) relation (Smith and Kliche 2005; von Lerber et al. 2017).

The mass of falling snow particles can be estimated with hydrodynamic theory (e.g., Böhm 1989; Mitchell and Heymsfield 2005), where mass can be retrieved by equating drag force with the difference between gravitational force and buoyancy force. For determining drag force, the measurements of terminal fall velocity, maximum diameter, and effective particle area are needed. As discussed more thoroughly in, for example, Schefold (2004), Szyrmer and Zawadzki (2010), and von Lerber et al. (2017), the particle area and
dimensions perpendicular to the airflow define drag force. However, the ground-based disdrometers observe particles on one or more projections from the side. For nonspherical particles with complex structures, such as ice crystals or snowflakes, the true dimensions must be estimated from the side measurements, which causes uncertainty in mass retrievals. Here, a simple correction factor is applied [explained in von Lerber et al. (2017)] to estimate the error between the observed maximum diameter from the side relative to the true maximum diameter (Wood et al. 2013). The value of the correction factor is determined by comparing the LWE accumulation of the event estimated by PIP with accumulation measured by the gauge. The area ratio from the side is assumed to be same as it is in the direction perpendicular to flow (Szyrmer and Zawadzki 2010). Similarly, hydrodynamic theory has earlier been applied for retrieving mass, for example, in Szyrmer and Zawadzki (2010), Wood et al. (2014), and Huang et al. (2015). The retrieved mass relation $m(D) = a_mD^{b_m}$ is given as a function of maximum diameter, and linear regression fit is performed in log scale. Note that the diameter in $u(D)$ and $m(D)$ relations in Eqs. (1) and (2) is corrected also with selected correction factor between the observed and the true diameter.

The event-specific $Z_e-S$ relation is computed with the total least squares method (Petráš and Bednárová 2010) from the values determined every 5 min shown as an example in Fig. 2a for the snow event on 6–7 November 2014. The finite integral limits induce error to the estimated $Z_e-S$. The effect of the limited observations of small particles is assumed to cause underestimation in the precipitation rate. For snow events with small-sized ice particles (median volume diameter $D_{deq,0}$ smaller than 1 mm), the underestimation can be significant, approximately 10%–20% (Tiira et al. 2016), and with larger $D_{deq,0}$ values, the error is close to 5%. Partly, this underestimation is compensated with the correction factor applied to the mass estimate as the value is defined by equalizing the PIP-estimated and the gauge-measured accumulation. The influence of truncation from the small particle size on higher moments of PSD, such as $Z_e$, is assumed to be small. The truncation from the side of large particles can be estimated similarly as in Ulbrich and Atlas (1998).

b. **Instantaneous $Z_e-S$ relation**

In the event-specific $Z_e-S$ relation, it is assumed that a single prefactor $a_{zs}$ and exponent $b_{zs}$ describe the event, albeit it is known that factors are varying depending on crystal type, particle density, particle terminal velocity, and PSD (Rasmussen et al. 2003; von Lerber et al. 2017).
which in snowstorms can change from one to another on a scale of minutes. Rewriting Eq. (1),

\[ Z_e = \frac{36}{\pi^3 \rho_s^2 \rho_i^2} \left( \frac{K_{ic}^2}{K_{iw}^2} \right) (a_m D_m^3) \sum N_0 D^\mu \exp(-\Delta D) dD \]

\[ = \frac{36}{\pi^3 \rho_s^2 \rho_i^2} \left( \frac{K_{ic}^2}{K_{iw}^2} \right) \sum N_0 \Lambda^{-2(\lambda + \mu + 1)} \Gamma(2b_m + \mu + 1) \]

\[ = F_e \Lambda^{-2(\lambda + \mu + 1)} \]  

(3)

and Eq. (2),

\[ S = \frac{1}{\rho_w} \sum a_m D_m^3 a_v D_v^3 N_0 D^\mu \exp(-\Delta D) dD \]

\[ = \frac{1}{\rho_w} \sum a_m a_v N_0 \Lambda \left( b_m + b_v + \mu + 1 \right) \]

\[ = F_s \Lambda^{b_m + b_v + \mu + 1}, \]  

(4)

when gamma PSD in the form of \( N(D) = N_0 D^\mu \exp(-\Delta D) \) is assumed with the intercept parameter \( N_0 \) in inverse millimeters per meter cubed and shape parameter \( \mu \) and slope parameter \( \lambda \) in inverse meters. Therefore, the relation between \( Z_e \) and \( S \) is

\[ Z_e = F_e F_s S \]

\[ = a_{zs} S^b_{zs}, \]  

(5)

where the prefactor \( a_{zs} \) can be written as

\[ a_{zs} = \frac{36}{\pi^3 \rho_s^2 \rho_i^2} \left( \frac{K_{ic}^2}{K_{iw}^2} \right) \sum 2b_m + \mu + 1 \}

\[ \times \Gamma(2b_m + \mu + 1) \Gamma(b_m + b_v + \mu + 1) \]  

(6)

For analytic solution, the integral limits in Eqs. (3) and (4) are assumed to be infinite. It can be seen that the instantaneous factors \( a_{zs} \) and \( b_{zs} \) vary as function of the \( m(D) \) and \( v(D) \) relations and PSD parameters. Factors can be utilized to assess uncertainties in quantitative precipitation estimates.

c. Error limits of the \( Z_e-S \) relation

The exponent \( b_{zs} \) in Eq. (5) is dependent on the exponents of the \( m(D) \) and \( v(D) \) relation and \( \mu \) of PSD. Inserting table values of different \( m(D) \) and \( v(D) \) relations from literature (e.g., Locatelli and Hobbs 1974; Mitchell et al. 1990; Barthazy and Schefold 2006), for \( b_m \), the values vary from 1.4 for aggregates to 3.0 for graupel, and for \( b_v \) from 0.12 for aggregates to 0.66 for graupel. Assuming an exponential PSD with \( \mu = 0 \), the exponent \( b_{zs} \) has a value of approximately 1.5. Estimating \( \mu \)

### Table 2. The contingency table for skill scores.

<table>
<thead>
<tr>
<th>IKA</th>
<th>True (( S \geq 0.1 \text{ mm h}^{-1} ))</th>
<th>False (( S &lt; 0.1 \text{ mm h}^{-1} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMI</td>
<td>True [( S_{V04A}(\text{POP} &gt; 9%) ) &gt; a]</td>
<td>False [( S_{V04A}(\text{POP} &gt; 9%) = 0 )]</td>
</tr>
<tr>
<td></td>
<td>False [( S_{V04A}(\text{POP} &gt; 9%) = 0 )]</td>
<td>b</td>
</tr>
</tbody>
</table>

### Table 3. Overview of the meteorological conditions during the studied snowfall events at the Hyttiälä measurement site and the amount and times of the GPM overpasses.

<table>
<thead>
<tr>
<th>Date</th>
<th>Accumulation LWE (mm)</th>
<th>Min temp (°C)</th>
<th>Max temp (°C)</th>
<th>Min RH (%)</th>
<th>Max RH (%)</th>
<th>Wind mean (m s(^{-1}))</th>
<th>No. of overpasses</th>
<th>Time of overpass (UTC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1300–2355 UTC</td>
<td>6.3</td>
<td>-4.3</td>
<td>-0.5</td>
<td>67</td>
<td>97</td>
<td>1.8</td>
<td>4</td>
<td>1346:52, 1521:37, 1657:14, and 1832:26</td>
</tr>
<tr>
<td>20 Mar 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1800 UTC 6 Nov–1700 UTC 7 Nov 2014</td>
<td>10.7</td>
<td>-2.4</td>
<td>-1.2</td>
<td>84</td>
<td>97</td>
<td>0.8</td>
<td>4</td>
<td>1816:18, 1952:33, 2127:26, and 2302:35</td>
</tr>
<tr>
<td>0900–1400 UTC</td>
<td>1.4</td>
<td>-10.6</td>
<td>-8.9</td>
<td>90</td>
<td>92</td>
<td>0.8</td>
<td>1</td>
<td>0916:04</td>
</tr>
<tr>
<td>24 Dec 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0200–1500 UTC</td>
<td>6.4</td>
<td>-9.4</td>
<td>-0.4</td>
<td>92</td>
<td>98</td>
<td>1.4</td>
<td>4</td>
<td>0224:24, 0359:18, 0534:43, and 0710:15</td>
</tr>
<tr>
<td>30 Dec 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000–2100 UTC</td>
<td>5.6</td>
<td>-7.6</td>
<td>-0.8</td>
<td>92</td>
<td>97</td>
<td>1.3</td>
<td>4</td>
<td>0002:58, 0139:04, 0314:11, and 0448:57</td>
</tr>
<tr>
<td>7 Jan 2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1200–2355 UTC</td>
<td>7.1</td>
<td>-2.0</td>
<td>-0.4</td>
<td>92</td>
<td>97</td>
<td>1.1</td>
<td>4</td>
<td>1620:12, 1755:29, 1931:31, and 2105:59</td>
</tr>
<tr>
<td>31 Jan 2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1300–2355 UTC</td>
<td>4.7</td>
<td>-0.5</td>
<td>0</td>
<td>96</td>
<td>97</td>
<td>1.5</td>
<td>1</td>
<td>2329:50</td>
</tr>
</tbody>
</table>
to vary between -1 and 3, the exponent has values of 1.84–1.23. Hence, the changes in $b_{zs}$ are small, and an average value of the instantaneous $b_{zs}$ ($\mu = 0$) of each snow event is used as a fixed value for defining the error limits as shown in Fig. 2b. As seen in Eq. (6), the prefactor $a_{zs}$ is strongly dependent on $N_0$ and on both prefactors of $m(D)$ and $v(D)$ relations, and therefore the uncertainty in the $Z_e-S$ relation is defined by the uncertainty of the prefactor. Error limits for $a_{zs}$ are determined from the cumulative distribution function (CDF) at percentiles of 25% and 75% of different values of $a_{zs}$ gained from the observed 5-min values of $Z_e$ and $S$, keeping $b_{zs}$ to its fixed value. This is demonstrated in Fig. 2c.

d. Comparison of spaceborne and ground-based snowfall rate

Validation of the GMI GPROF snowfall surface precipitation rate is performed by first comparing precipitation rate values directly between GMI-estimated and ground radar–estimated values in the rectangular 15 km × 15 km sampling grid. Second, the detection skills in the grid points of the same sampling grid are evaluated with three traditionally used metrics: CSI, POD, and FAR (Wilks 2011). The contingency table with the criteria thresholds used is presented in Table 2 with hits $a$, false detection $b$, misses $c$, and correct rejections $d$. The CSI is a metric presenting the skill to detect correctly snowfall relative to all the correctly or falsely detected snowfall, but it excludes the analysis of detecting the nonsnowing situations correctly. The CSI is defined as

$$CSI = \frac{a}{a+b+c}. \quad (7)$$

POD is also called the hit rate; it describes the ratio of correctly detected snowfall grid points to all snowfall points:

$$POD = \frac{a}{a+c}. \quad (8)$$

FAR is the amount of false detected no-snowfall grid points divided by the total amount of no-snowfall grid points, which can be stated as

$$FAR = \frac{b}{b+d}. \quad (9)$$

In this study, the selected overpass cases include those in which both GMI and ground radar have observed snowfall. Therefore, this dataset is already biased relative to detecting precipitation, and for example, Heidke skill score is not meaningful to compute, as no-snowfall observations in the sampling grid points are underrepresented. Also when $c = 0$, POD automatically equals 1 even though this would not describe the true detection skills. Similarly, if $d = 0$, then FAR = 1. Both values in a grid point are neglected from the analysis.

### TABLE 4. The prefactors and exponents of the $Z_e-S$ relations. The event-specific factors, mean exponent from the instantaneous relations, and upper and lower limit for the prefactor at the percentiles of 25% and 75% of the defined CDF are shown. (The unit for $a_{zs}$ is mm$^{-0.5}$ mm$^{-2}$ h$^{-1}$.)

<table>
<thead>
<tr>
<th>Date</th>
<th>Event $a_{zs}$</th>
<th>Event $b_{zs}$</th>
<th>Mean $b_{zs}$</th>
<th>Lower limit $a_{zs}$ (75%)</th>
<th>Upper limit $a_{zs}$ (25%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1300–2355 UTC 20 Mar 2014</td>
<td>73.3</td>
<td>1.61</td>
<td>1.55</td>
<td>104.8</td>
<td>45.5</td>
</tr>
<tr>
<td>1800 UTC 6 Nov–1700 UTC 7 Nov 2014</td>
<td>233.3</td>
<td>1.55</td>
<td>1.55</td>
<td>358.6</td>
<td>134.9</td>
</tr>
<tr>
<td>0900–1400 UTC 24 Dec 2014; long-term average</td>
<td>115.5</td>
<td>1.36</td>
<td>1.55</td>
<td>276.4</td>
<td>72.6</td>
</tr>
<tr>
<td>0200–0915 UTC 30 Dec 2014</td>
<td>44.6</td>
<td>1.22</td>
<td>1.57</td>
<td>114.9</td>
<td>32.4</td>
</tr>
<tr>
<td>0130–2100 UTC 7 Jan 2015</td>
<td>38.6</td>
<td>1.20</td>
<td>1.57</td>
<td>83.2</td>
<td>45.3</td>
</tr>
<tr>
<td>2100 UTC 12 Jan–0300 UTC 13 Jan 2015</td>
<td>83.7</td>
<td>1.34</td>
<td>1.55</td>
<td>128.8</td>
<td>57.4</td>
</tr>
<tr>
<td>1200–2355 UTC 31 Jan 2015</td>
<td>134.6</td>
<td>1.53</td>
<td>1.56</td>
<td>238.1</td>
<td>60.7</td>
</tr>
<tr>
<td>1300–2100 UTC 30 Mar 2015; long-term average</td>
<td>115.5</td>
<td>1.36</td>
<td>1.55</td>
<td>276.4</td>
<td>72.6</td>
</tr>
</tbody>
</table>
FIG. 4. Precipitation rate estimated with IKA compared with gauge-measured rate for the snowfall event on 6–7 Nov 2014 at each gauge location. The dark-blue solid line shows the radar-estimated precipitation rate utilizing the event-specific factors of the $Z_e-S$ relation averaged over 3 km x 3 km area at the precipitation gauge location, the light-blue shaded area presents the error limits determined with the upper and lower values of prefactor $a_{Z_e}$ of the instantaneous $Z_e-S$ relation, and red diamonds show the 10-min average precipitation rate measured with gauges except in (f) with the PL200 in Hyytiälä, where the rate is averaged over 5 min.
4. Results

Eight snowfall events with 26 GMI overpasses are selected for studying the performance of the \( Z_e - S \) estimates with ground-based radar and for validating the space-based GMI snowfall estimates. An overview of the meteorological conditions of the cases is shown in Table 3. Most of the snow events are covered with four overpasses every 1.5 h. The focus is on testing the potential of the retrieved \( Z_e - S \) relations in snowfall, and thus, the cases are selected to have temperatures below \( 0^\circ C \) (measured in Hyytiälä at the FMI automatic weather station), and no wet-snow events are studied. For each event, the sufficient precipitation rate and accumulation for a reliable retrieval of \( Z_e - S \) are needed. The accumulated LWE measured by the Pluvio\(^2\) 200 gauge is shown in Table 3. The highest LWE accumulation of 10.7 mm occurred on 6–7 November 2014. During this event, several particle types were observed. In the 5-min time series, the median volume disk-equivalent diameter \( D_{\text{deq},0} \) had a maximum value of 6.26 mm associated with \( m(D) = 3.7 \times 10^{-5} D^{2.07} \), indicating large aggregates at 1150 UTC, whereas at 0015 UTC, the maximum value of \( D_{\text{deq},0} = 2.19 \text{ mm with } m(D) = 5.8 \times 10^{-5} D^{2.75} \), suggesting the presence of highly rimed particles.

a. Event-specific \( Z_e - S \) relations with error limits

The fit of the event-specific \( Z_e - S \) relation and the defined error limits on 6–7 November 2014 is illustrated in Fig. 2. The values for the other events are shown in Table 4. The exponent of the event-specific \( Z_e - S \) relation varies from 1.2 to 1.61, whereas the mean \( b_{zs} \) from the instantaneous relations is between 1.55 and 1.57. The prefactor \( a_{zs} \), on the other hand, varies significantly (i.e., between 38.6 and 233.3 mm\(^{b_{zs}}\) m\(^{-3}\) h\(^{-1}\)), and therefore, it is the main contributor to the error limits, as also discussed in Rasmussen et al. (2003) and von Lerber et al. (2017). As a first check for the retrieved \( Z_e - S \) relations, they are applied to IKA PPI scans with a 0.3° elevation angle. The LWE accumulation for the whole event is summed from the 5-min scans and compared with the measurements of AWS gauges, as shown for the 6–7 November 2014 event in Fig. 3. There are several limitations in this comparison. The artifacts of the ground clutter and partial beam blocking are visible as well as the underestimation in accumulation with increasing distance from the radar caused by the beam broadening and the increased height of the radar beam. Also, the snow advection (Lauri et al. 2012) is not considered. Nevertheless, at the gauge locations, the radar-estimated LWE accumulation seems to be reasonably accurate and compares well to gauge measurements.

A more detailed image of the radar-estimated snowfall rate for this snow event is shown in Fig. 4, where the radar-estimated precipitation rate is averaged between the two 5-min observations and depicted together with the gauge-measured accumulation rate. Error limits are applied to this comparison. As expected, the best correspondence between the estimated and observed rates is found in Hyytiälä. However, the comparison appears accurate also in the other locations. Altogether, comparing estimated and observed rates of 10 min linearly together, the coefficient of determination

![Fig. 5. Example case of two clearly different snow types and the corresponding \( Z_e - S \) relations. (a) The \( Z_e - S \) fits for the two separate time periods on 30 Dec 2014. (b) LWE accumulation map, where the radar-estimated precipitation rate is calculated with the \( Z_e - S \) relation defined for the 0200–0915 UTC time period, compared with the gauge-measured accumulation.](http://journals.ametsoc.org/jamc/article-pdf/57/4/797/3602500/jamc-d-17-0176_1.pdf)
In Fig. 4, the error limits are shown as a shaded area around the estimate. Occasionally, the error limits have a wide spread around the expected estimate, but it is following the large variations occurring in the snow events. For a more uniform snowfall event, the error limits are narrower.

Generally, the defined event-specific relation with error limits describe the snow rate well; one-to-one linear correlation of an hourly accumulation for 10 snow cases was presented in von Lerber et al. (2017), with the coefficient of determination of $r^2 = 0.76$ and RMSE of 0.38. Occasionally, the presented method has difficulties estimating the snowfall rate (e.g., with shallow clouds). One example is from 0000 to 2100 UTC 7 January 2015. The event-specific prefactor $a_{zs}$ value is the lowest of all the studied cases here. The LWE accumulation during the event is 5.6 mm, but the precipitation rate is low. The cloud echo top is below

![Diagram](http://journals.ametsoc.org/journals/jamc/article-pdf/57/4/797/3602500/jamc-d-17-0176_1.pdf)
5 km, with a mean altitude of 2.5 km. Precipitation consisted mostly of small particles with mean $D_{deq,0} = 0.84$ mm. Given that the cloud height is low, the particles are presumably single crystals. The underestimation of precipitation rate and the possible limitation in observing the true area ratio are compensated with the selected correction factor value used in the mass estimate as described in section 3a.

The snow event recorded on 30 December 2014 is a good example where the use of a single $Z_e-S$ relation is not ideal. In Fig. 5a, one can see that two separate relations can be identified. During the morning period of the event between 0200 and 0915 UTC, the particles are smaller and denser as the mean $D_{deq,0}$ for the period is 1.18 mm, and the mean value for the exponent of the $m(D)$ relation is $b_m = 2.12$. During the second period, between 0915 and 1500 UTC, the particle sizes increase with the mean and maximum values of $D_{deq,0}$ being 2.02 and 5.37 mm, respectively, and the mean value of $b_m$ decreased to 1.92, indicating a presence of aggregates. Two separate fits were performed as depicted in Fig. 5a, and the relation of the morning period is used in the comparison of GMI snowfall estimate as all four overpasses occurred during this period. The radar-estimated LWE accumulation map relative to gauge measurements for the morning period is shown in Fig. 5b.

In snowfall events, where the main precipitation system did not pass over the Hyttiälä measurement site, the derived event-specific $Z_e-S$ relations are not representative. Instead, the long-term average relation is applied. The relation is defined for 24 snow events measured during the winters of 2014 and 2015 from the 5-min calculated $Z_e$ and $S$ values, and the error limits are determined as described in section 3c. In this study, these are applied to the snow events on 24 December 2014 and 30 March 2015 (Fig. 6).

b. Bias of GMI surface snowfall estimates

For the event on 31 January 2015 in Fig. 7, the estimated snowfall rate of GMI GPROF is compared with the radar and gauge observations. The GMI values are taken from the closest pixel to a gauge. It shows that the GPROF algorithm detects most of the precipitation but underestimates the precipitation rate.

A comparison between the radar- and GMI-estimated snowfall rates is performed on the rectangular grid with a pixel size of 15 km $\times$ 15 km in an area of 25,600 km$^2$ centered on the IKA radar, with the minimum and maximum latitudes being 60.90° and 62.61°N and the minimum and maximum longitudes being 21.33° and 24.93°E. The event-specific $Z_e-S$ relation is applied to the PPI scans closest to the overpass and to the two scans before and after. The elevation angles of 0.3°, 0.7°, 1.0°, and 1.5° are used if the midvolume height of the bin is below 1.5 km as estimated from the 4/3 Earth radius model (Doviak and Zrnić 1993). The precipitation rate value in each radar bin is projected to the ground and for each point in the sampling grid. The snowfall rate is averaged over all the observations that fall within a grid point. The GMI surface precipitation observations are also sampled to the rectangular grid, where precipitation rate values are computed as the area-weighted average of the original pixels. The original data of the different datasets and the corresponding values to the sampled grid are illustrated in Fig. 8 for the overpass on 0710 UTC 30 December 2014. In this case, the GPROF
FIG. 8. The studied datasets shown in (left panel of a two panel set) the original accuracy and (right panel of a two panel set) as newly sampled to the $15\text{ km} \times 15\text{ km}$ grid for the GMI overpass at 0710:15 UTC 30 Dec 2014. (a),(b) The average radar-estimated surface precipitation rate, (c),(d) cloud echo top, (e),(f) GMI-estimated precipitation rate of V04A, (g),(h) GMI-estimated precipitation rate of V05A, (i),(j) columnar water vapor index of V04A, (k),(l) columnar water vapor index of V05A, (m),(n) PoP of V04A, (o),(p) PoP of V05A, (q),(r) snow depth proxy, and (s),(t) surface class of V05A version.
algorithm seems to detect precipitation well. The radar image shows that precipitation is closely following the cloud echo top pattern, and the GPROF algorithm seems to be sensitive to the cloud height. However, both algorithm versions underestimate the rate. No strong dependence of the detection on column water vapor index data is observed, and only slight differences between the values of different versions are noticed. Both versions V04A and V05A also show a similar pattern of PoP as the observed precipitation, although version V04A has higher PoP values. Surface classes are slightly different for the algorithm versions; both are showing the snow-covered land, but the depth indicated with the snow-cover class is not following the snow depth proxy. Similar correspondence of the different products is also found in Fig. 9 on 30 March 2014, where the edge of the cloud seen in the cloud echo top product and the radar-estimated precipitation can be distinguished in most products of GMI GPROF. The V05A algorithm is adjusted to improve the precipitation estimates over...
snow-covered land. In some cases, the precipitation estimate appears to follow the snow depth proxy. But, for example, during the overpass at 2333 UTC 12 January 2015 in Fig. 10, a part of the precipitation is missing, and the threshold seems to better follow the cloud echo top product more than snow depth proxy. During this overpass, the GPROF algorithm is detecting the precipitation pattern but strongly underestimating the rate.

Overall, the GPROF snowfall estimate correlates with the precipitation thickness, which is defined from the 0-dBZ cloud echo top estimated by IKA. In some cases, for example, shown in Fig. 11, at 2105 UTC 31 January 2015, the relation between precipitation rate and precipitation thickness is less straightforward. The IKA observations show that not all high echo top areas result in heavier precipitation, as was also pointed out in a study of the relation between satellite-observed water paths and surface precipitation rates (You and Liu 2012). The GPROF assigns larger precipitation rates to high echo top area, although on this area, IKA is measuring lower precipitation rates. Figure 12 illustrates the dependency between the probability of precipitation and the echo top. An occurrence threshold for V05A
precipitation estimates based on the PoP can be identified. The majority of the observations are between cloud echo top values of 2 and 4 km with PoP around 20%. Version V04A has wider spread both in the PoP and cloud echo top values, but approximately in both versions, the PoP is increased when cloud echo top is higher than 3 km. This indicates that snowfalls originating from the lower-level clouds may not be identified by the GPROF. It is also possible that snow growth processes (e.g., aggregation and riming), which can occur at lower parts of the cloud and which change the precipitation rate, are not well represented in the a priori database, and thus, the effect of these processes are not captured by the algorithm.

Quantitatively, the bias between the GMI- and radar-estimated snowfall rate for all the studied events is depicted in Fig. 13. It shows the density scatterplots of the precipitation rate of both versions relative to the rate observed by IKA, the CDF of precipitation rates, and the histogram of a ratio between the both GPROF algorithm versions and the radar-estimated snowfall rate. In the Bayesian approach of GPROF, the constructed vertical profile is computed by weighting the database entries in comparison with the measured brightness temperature. Therefore, every pixel has some value for the precipitation rate, although sometimes the value is extremely small (Kummerow et al. 2015). This can be

Figure 11. As in Fig. 9, but for 2105:59 UTC 31 Jan 2015 with (c) the columnar water vapor index of V04A and (f) surface class of V05A version.
c. Detection of snowfall with GMI

Qualitative analysis of the overpass images suggests that the detection of GPROF is mostly dependent on the cloud echo top height and no strong correlation is seen concerning changes in snow depth proxy or an absolute value of snow depth, which would be most physically tied to the land surface emission character. Contingency metrics CSI, POD, and FAR are chosen to investigate the detection skills. The skill scores have been calculated for all the cases for both versions V04A (with threshold) and V05A and utilizing threshold of 0.1 mm h$^{-1}$ for the radar-estimated precipitation rate. Results are shown in Table 5. Version V04A obtains slightly higher CSI of 0.83 than version V05A with a value of 0.79, though the precipitation rates with V05A were closer to the observed values. It may be that the new detection threshold performs differently in certain conditions, as was shown in Fig. 10, and these removed grid cells decrease the CSI score. Another possible reason is, because of Bayesian approach, version V04A always has value for every grid cell, and in this analysis, the number of hits is statistically overrepresented; therefore, CSI of the V04A version has higher values. Although, in this respect, FAR should also increase for the V04A version. The FAR is 0.09 for version V04A and 0.10 for version V05A, and no notable difference is noticed. As was stated in section 3d, this studied dataset is biased to find more hits than correct rejections, which favors version V04A. The POD values for both versions V04A and V05A are similar, with values of 0.90 and 0.84, respectively.

The same skills have been calculated as a function of the cloud echo top, column water vapor index, snow depth proxy, and surface classes. The cloud echo top is divided into 500-m bins, and the skill scores have been defined for hits falling inside each height bin. In some cases, the skill scores will have no values, which is typically caused by lack of observations falling within a bin. Altogether for both algorithm versions, V04A and V05A, all scores are very close to each other, and the detection skill is improved as the cloud echo top height increases as shown in Fig. 14. At 3-km height, FAR of V05A has decreased from 1 to 0.08 and of V04A from 0.92 to 0.08, whereas POD has increased for V05A from 0.6 to 0.8 and for V04A from 0.7 to 0.88. Similar behavior of skills is also seen with column water vapor index shown in Fig. 15. These results verify the strong correlation of the GPROF algorithm to cloud thickness. And the detection skill with shallow clouds is smaller, as has been stated, for example, in Bennartz and Bauer (2003) and Skofronick-Jackson et al. (2013).

Scattering of ice particles in snowfall can be challenging to distinguish from scattering from the snow-covered
surface. In GPROF, the surface type is modified to a snow estimate from NOAA’s AutoSnow product (Aires et al. 2011; Romanov and Gutman 2000; Kummerow et al. 2015). The skill metrics are plotted as a function of the snow depth proxy every 2.5 cm in Fig. 16, and no similar strong dependence is observed on skill scores as with cloud height. There is a decrease in both POD and FAR, especially for the V05A version, around snow depth of 20–30 cm, but also, the number of data points is low and can explain the decreased skills. By visual inspection, GPROF surface class product seems to catch the snow conditions (e.g., on event 6–7 November); when the snow event started, the ground was bare, and this is also observable in the product. Interestingly, the four different snow classes do not seem to follow the snow depth estimated with the proxy, and the correspondence to skill scores is not strong (Fig. 17).

5. Conclusions

In this study, a method to estimate the snowfall rate from the combined measurements of a weather radar, video disdrometers, and precipitation gauges is utilized for validation of the GMI surface snowfall product. The \( Z_{cr} - S \) relation is determined for each studied case based on microphysical properties of falling snow including mean measured PSD and retrieved \( m(D) \) and fitted \( v(D) \) relations of every 5 min. An event-specific \( Z_{cr} - S \) relation is applied to the equivalent radar reflectivity factor observations, and error limits of the snowfall rate are

<table>
<thead>
<tr>
<th>Table 5. FAR, POD, and CSI for all overpasses.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMI V04A; PoP &gt; 9</td>
</tr>
<tr>
<td>FAR</td>
</tr>
<tr>
<td>POD</td>
</tr>
<tr>
<td>CSI</td>
</tr>
</tbody>
</table>
FIG. 14. The (a) CSI, (b) FAR, and (c) POD as a function of cloud echo top.

FIG. 15. As in Fig. 14, but as a function of total water vapor index product.
determined from the 5-min instantaneous values of prefactor $a_{zs}$ and exponent $b_{zs}$ of power-law relation. The upper and lower error limit of the $a_{zs}$ is evaluated from a CDF at percentiles of 25% and 75% when the exponent $b_{zs}$ is kept as a mean constant value. The radar-estimated snowfall rates have been compared with the observations of the operational gauges in the 30–70-km distance from the radar. Both total LWE accumulation and averaged 10-min snowfall rate are studied, and the accuracy of the radar-based estimates is good even though, for example, snow drifting or radar beam broadening is not considered in the comparison. The radar-based snowfall rate has instant high values, which are not observable in the gauge measurements as these are averaged over the 10-min period; thus, when comparing a longer LWE accumulation period calculated by summing the snowfall rates together, this is one of the reasons the radar-estimated rate seems to overestimate accumulation relative to gauges. Given a strong dependence of prefactor $a_{zs}$ on the changes in PSD parameters and prefactor of $m(D)$ relation (von Lerber et al. 2017), the determined error limits of the $Ze-S$ relation are wide. Limits can be utilized, for example, in the probabilistic nowcasting of winter precipitation.

In the events where only the edge of the precipitating system passed the Hyytiälä measurement station, the event-specific factors are not applicable. In these cases, a long-term average relation can be applied; it is defined for 24 snowfall cases during two consecutive winters 2014–15 with error limits. It must be noted that this relation seems to be characteristic of a geographical region; for example, it was observed (Tiira et al. 2016; Moisseev et al. 2017) that ensemble mean snow density as a function of median volume diameter during the
winters of 2014 and 2015 in Finland was higher in comparison with similar relations diagnosed for Colorado snow (Brandes et al. 2007).

When utilizing the event-specific $Z_c-S$ relation obtained in this work, the radar-estimated snowfall rate can be applied as a direct ground validation for space-based snowfall products. In this study, this is demonstrated with the GPM GMI surface snowfall rate retrieved with GPROF algorithm versions V04A and V05A. Altogether, 26 cases are studied where snowfall was occurring at the time of GPM overpass. By calculating the ratio of the GMI-observed snowfall rate relative to the radar-estimated rate, a clear bias is demonstrated. Both V04A and V05A GMI GPROF products underestimate the detected snowfall rate. The median ratio value for version V04A is 0.16 and for version V05A is 0.32. Based on skill scores presented herein, the snowfall detection of GPROF closely follows radar echo top height, peaking for heights exceeding 3 km. The implication is that events associated with echo tops < 3 km will be less well detected or missed altogether. Generally, GPROF seems to detect the snowfall well; POD is high for the whole dataset, the values being 0.90 (V04A) and 0.84 (V05A).

Acknowledgments. Instrumentation and efforts used in this study were supported by NASA Global Precipitation Measurement mission ground validation and Precipitation Measurement Missions science programs. We are grateful for the valuable work of the scientists and technicians of NASA Global Precipitation Measurement mission ground validation program and personnel at Hyytiala station and University of Helsinki, especially Matt Wingo, Larry F. Bliven, Matti Leskinen, and Janne Levula. We acknowledge Markus Peura, FMI, for providing cloud echo top data. Author von Lerber was funded by a grant of the Vilho, Yrjö and Kalle Väisälä Foundation and by SESAR Joint Undertaking Horizon 2020 Grant Agreement 699221 (PNOWWA). The research of author Moisseev was supported by Academy of Finland (Grant 305175) and the Academy of Finland Finnish Center of Excellence program (Grant 3073314); he also acknowledges the funding received by ERA-PLANET, transnational project iCUPE (Grant Agreement 689443), funded under the EU Horizon 2020 Framework Programme. Research funding for author Petersen is provided by the NASA Precipitation Measurement Missions science program.

REFERENCES


Hudak, D., P. Rodriguez, and N. Donaldson, 2008: Validation of the CloudSat precipitation occurrence algorithm using the


——, and E.-K. Seo, 2013: Detecting snowfall over land by satellite high-frequency microwave observations: The lack of scatter-