

Retrieving Surface Snowfall With the GPM Microwave Imager: A New Module for the SLALOM Algorithm

Jean-François Rysman, Giulia Panegrossi, Paolo Sanò, Anna Cinzia Marra, Stefano Dietrich, Lisa Milani, Mark S. Kulie, Daniele Casella, Andrea Camplani, Chantal Claud, et al.

▶ **To cite this version:**

Jean-François Rysman, Giulia Panegrossi, Paolo Sanò, Anna Cinzia Marra, Stefano Dietrich, et al.. Retrieving Surface Snowfall With the GPM Microwave Imager: A New Module for the SLALOM Algorithm. Geophysical Research Letters, 2019, 46, pp.13,593-13,601. 10.1029/2019GL084576 . insu-03727009

HAL Id: insu-03727009
https://insu.hal.science/insu-03727009

Submitted on 28 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Check for updates

AGU100 ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL084576

Retrieving Surface Snowfall With the GPM Microwave Imager: A New Module for the SLALOM Algorithm

Jean-François Rysman[1,2], Giulia Panegrossi[1], Paolo Sanò[1], Anna Cinzia Marra[1], Stefano Dietrich[1], Lisa Milani[3,4], Mark S. Kulie[5], Daniele Casella[1], Andrea Camplani[1,6], Chantal Claud[2], and Léo Edel[2]

[1]Institute of Atmospheric Sciences and Climate-National Research Council, Rome, Italy, [2]Laboratoire de Météorologie Dynamique/IPSL, Centre National de la Recherche Scientifique, École Polytechnique, Université Paris-Saclay, Palaiseau, France, [3]NASA Goddard Space Flight Center, Greenbelt, MD, USA, [4]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA, [5]NOAA/NESDIS/STAR/Advanced Satellite Products Branch, Madison, WI, USA, [6]Geodesy and Geomatics Division-DICEA, Sapienza University of Rome, Rome, Italy

Abstract In this study, we present a new module for the Snow retrievaL ALgorithm fOr gMi (SLALOM) that retrieves surface snowfall rate using Global Precipitation Measurement (GPM) Microwave Imager measurements together with humidity and temperature vertical profiles. This module, named Surface Snowfall Rate Module, is tuned using colocated surface snowfall observations of the Cloud Profiling Radar onboard CloudSat. Using this new module, the SLALOM algorithm is able to predict surface snowfall rate with a relative bias of −13%, a root-mean-square error of 0.08 mm/hr, and a correlation coefficient of 0.7. Surface Snowfall Rate Module is then used to retrieve snowfall rate for three case studies and to provide a unique, 70°S to 70°N high-resolution distribution of average surface snowfall rate from 2014 to 2017. This new product will be useful for surface precipitation analyses, global water budget estimation, and climatological analyses.

Plain Language Summary A comprehensive monitoring of Earth surface snowfall can only be reached using spaceborne instruments. However, estimating surface snowfall from space is challenging as the signal measured by space sensors is only indirectly related to surface snowfall rate. In this study, we developed a new algorithm that is able to estimate surface snowfall rate using observations from the Global Precipitation Measurement Microwave Imager together with humidity and temperature vertical profiles. We used this algorithm to provide a unique, 70°S to 70°N high-resolution distribution of average surface snowfall rate from 2014 to 2017. This new product will be useful for surface precipitation analyses, global water budget estimation and climatological analyses.

Key Points:
- We developed a surface snowfall rate retrieval algorithm for the Global Precipitation Measurement Microwave Imager
- Surface snowfall rate is predicted with a relative bias of −13% and a correlation coefficient of 0.7 compare to CloudSat Cloud Profiling Radar observations
- A unique, 70°S to 70°N high-resolution distribution of average surface snowfall rate from 2014 to 2017 is provided

Correspondence to:
J.-F. Rysman,
jeanfrancois.rysman@artov.isac.cnr.it

Citation:
Rysman, J.-F., Panegrossi, G., Sanò, P., Marra, A. C., Dietrich, S., Milani, L., et al. (2019). Retrieving surface snowfall with the gpm microwave imager: a new module for the slalom algorithm. *Geophysical Research Letters*, *46*, 13,593–13,601. https://doi.org/10.1029/2019GL084576

Received 16 JUL 2019
Accepted 9 OCT 2019
Accepted article online 17 OCT 2019
Published online 29 NOV 2019

©2019. American Geophysical Union. All Rights Reserved.

1. Introduction

Although most spaceborne passive microwave sensors equipped with high-frequency channels (90–190 GHz) are highly sensitive to the radiation scattering by ice hydrometeors (e.g., Eriksson et al., 2015; Kulie et al., 2010; Skofronick-Jackson & Johnson, 2011), using them to retrieve surface snowfall rate (SSR) remains nowadays a particularly challenging task (Levizzani et al., 2011; Skofronick-Jackson et al., 2017). This complexity is due to several factors that include the high diversity of size and shape of snowflakes implying complex radiative signatures (e.g., Eriksson et al., 2015; Kulie et al., 2010; Kuo et al., 2016; Skofronick-Jackson & Johnson, 2011), vertical variability of snowfall within clouds, and difficulties to distinguish the signature of snowfall from nonprecipitating cloud ice or from surface background impaired by the low and variable emissivity of snow or ice-covered surfaces (e.g., Prigent et al., 2006; Takbiri et al., 2019). Overcoming these difficulties would improve greatly the ability to quantitatively characterize the Earth's water cycle and energy budget at the global scale. For this purpose, several attempts to develop snowfall detection and retrieval algorithms have been conducted in the past years (among others, Kongoli et al., 2003, Advanced Microwave Sounding Unit A/B, [AMSU-A/B]; Noh et al., 2009, AMSU-B; Liu & Seo, 2013, AMSU-B/Microwave Humidity Sounder; You et al., 2015, Special Sensor Microwave Imager and Sounder; Kummerow et al., 2015, GMI; and Tang et al., 2018, GMI). However, difficulties previously mentioned remain and further developments are needed in order to mitigate uncertainties about global snowfall estimates.

In this context, we propose a new approach to estimate surface snowfall from spaceborne instrumentation using the Global Precipitation Measurement (GPM) Microwave Imager (GMI; Hou et al., 2014; Skofronick-Jackson et al., 2017). A first step toward this objective has been accomplished with the development of the Snow retrievaL ALgorithm fOr gMi (SLALOM) by Rysman et al. (2018). This algorithm is able to detect surface snowfall and the presence of supercooled cloud droplets, and to retrieve the associated snow water path (SWP). Despite some limitations in specific environmental conditions (very low humidity and very weak snowfall), SLALOM has proven to retrieve SWP accurately compared to the CloudSat Cloud Profiling Radar (CPR) globally and in specific environmental conditions. In addition, since GMI allows for much higher spatiotemporal sampling than CPR, SLALOM is able to provide a more comprehensive SWP climatology. The new SSR Module (SSR-Mod) presented in this letter extends the SLALOM algorithm to retrieve SSR using the CloudSat CPR 2C-SNOW-PROFILE (2CSP) product as a reference. Thus, SSR-Mod aims at retrieving global snowfall rates by inferring at best the CPR estimates, available on the 1.7-km-wide CPR swath, on the full 880-km-wide GMI swath. The data set and the SSR-Mod are presented in the second and third sections of this paper. The fourth section is devoted to the results and the evaluation of SSR-Mod, while a discussion of the results and the conclusion are presented in the last section.

2. Data Set

The data set used is similar to the one described in Rysman et al. (2018) except that it extends now from March 2014 to October 2017. Here we provide a brief summary of its main characteristics. This data set, adapted from the 2B-CSATGPM product (Turk, 2016), contains 821,842 GMI and CloudSat coincident observations spread over 1,010 days together with ancillary environmental variables. These coincidences are mainly found around 60°S to 60°N (see Figures 1 in Panegrossi et al., 2017, and Rysman et al., 2018) and are associated with surface snowfall occurrence in 9% of the cases. Since CPR probes the atmosphere at a higher spatial resolution than GMI, CPR pixels are averaged over the closest GMI pixel to provide a one-to-one correspondence. Specifically, the data set contains GMI brightness temperatures (GMI L1C-R), SWP and SSR (from the CPR 2CSP product; Wood et al., 2014), supercooled droplets occurrence (from the CloudSat/Calipso DARDAR product; Delanoë & Hogan, 2010), land-sea flag (2B-GEOPROF product), sea ice concentration (from the AMSR2 ASI product; Spreen et al., 2008), 2-m temperature (ECMWF-AUX product), total precipitable water (TPW; ECMWF-AUX product), and the four first principal components of the vertical profiles of temperature, specific humidity, and relative humidity (ERA-Interim reanalysis at $0.75° \times 0.75°$).

3. Method

The SLALOM algorithm initially consisted of three modules for snowfall detection (SD-Mod), supercooled droplets detection (SCDD-Mod), and SWP retrieval (SWP-Mod). In this work, we developed a fourth module to retrieve SSR (SSR-Mod).

In order to obtain the most reliable estimates of snowfall rate, we evaluated the ability of seven widely used statistical approaches: support vector machine (Boser et al., 1992; Cortes & Vapnik, 1995), gradient boosting (Chen & Guestrin, 2016), random forest (Breiman, 2001), linear regression, single-layer neural network (Ripley & Hjort, 1996), k-nearest neighbor (Altman, 1992), and regression tree (Breiman, 2017). We trained these algorithms on a sample of the data set corresponding to 80% of randomly chosen days (training data set) and evaluated their performances on the remaining 20% (testing data set). We tuned and fitted every model on the training data set to provide an estimate of SSR using the same input variables as those selected for the SLALOM SWP-Mod, that is, T2m, TPW, the four first principal components of the vertical profiles of temperature, specific humidity, and relative humidity, together with brightness temperatures of the 13 GMI radiometer channels (10.65 to 183.31 ± 7 GHz). Results showed that gradient boosting has the highest correlation coefficient and the lowest root-mean-square error (RMSE) among all tested algorithms compared to the reference and has thus been implemented in the SSR-Mod. It must be noted that SSR-Mod retrieves SSR only where SD-Mod detects snow but that SSR retrieval accuracy does not depend on the other module outputs.

4. Results

SSR-Mod evaluated on the testing data set shows a correlation coefficient of 0.7, a bias of −13%, and a RMSE of 0.08 mm/hr.

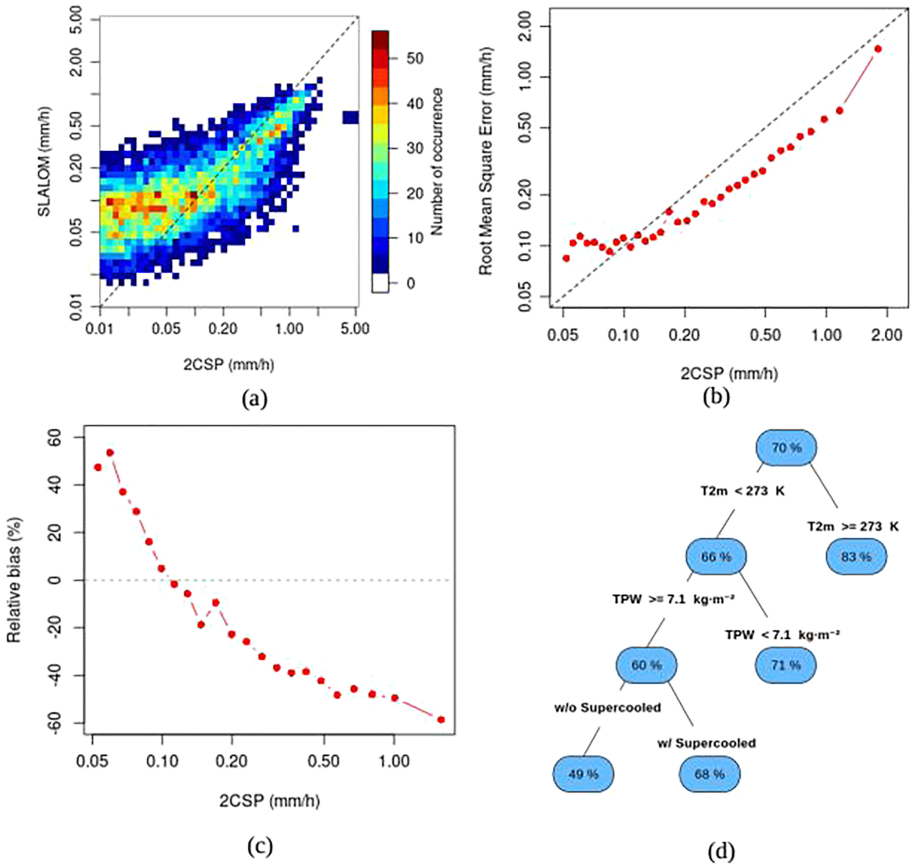


Figure 1. Two-dimensional histogram of surface snowfall of SLALOM versus surface snowfall of 2CSP (a). Relative bias (%; b) and root-mean-square error (mm/hr; c) as a function of 2CSP surface snowfall rate (mm/hr). Regression tree (d) with absolute value of the normalized difference percentage between predicted and observed surface snowfall rate indicated in the boxes, and partitioning variables and thresholds indicated on the branches. SLALOM = Snow retrievaL ALgorithm fOr gMi; 2CSP = 2C-SNOW-PROFILE; TPW = total precipitable water.

In Figure 1, we investigated the factors explaining sources of SSR-Mod uncertainties on snowfall rate retrieval. Figure 1a indicates that for very weak CPR SSR (below 0.05 mm/hr), SLALOM is not able to retrieve accurately SSR and predicts values between 0.02 and 0.2 mm/hr whatever the CPR SSR. On the contrary, above 0.05 mm/hr SLALOM and 2CSP are well correlated even if an uncertainty in the retrieval remains. Figure 1b shows that the relative bias ranges between 47% at 0.05 mm/hr (and higher for lower snowfall rate, not shown) and −59% for SSR higher than 1 mm/hr. This result implies that SSR-Mod tends to underestimate the heaviest snowfall events compared to 2CSP. We also plotted the RMSE as a function of SSR (Figure 1c). When RMSE is above (below) the dashed line, it means that RMSE is higher (lower) than the average SSR. It shows that SSR-Mod accuracy improves continuously from 0.05 to 1.2 mm/hr. Yet, in the last bin (corresponding to 1.8 mm/hr on average), RMSE increases significantly up to 1.5 mm/hr (i.e., retrieval is less accurate). The difficulty for SSR-Mod to retrieve surface snowfall in this last bin, that is, with the heaviest snowfall events of the data set, is probably related to the very low percentage of cases with SSR > 1.5 mm/hr in the coincidence database (i.e., only 314 cases over 56,380 snowfall cases, see Figure 6 in Casella et al., 2017, for the full probability distribution function for the 2014–2015 period).

In order to understand how the environmental conditions affect SSR-Mod skill, we built a regression tree of the normalized difference percentage between predictions and observations of SSR (Figure 1d). A normalized difference percentage equal to 0 indicates an optimal performance while increasing values mean lower overall performance. The regression tree clusters hierarchically this normalized difference percentage choosing, at each step, the best splitting variable (and associated threshold) among T2m, TPW, surface type (land, open sea, sea ice, and coast), or the presence of supercooled droplets. It is thus a very effective tool to

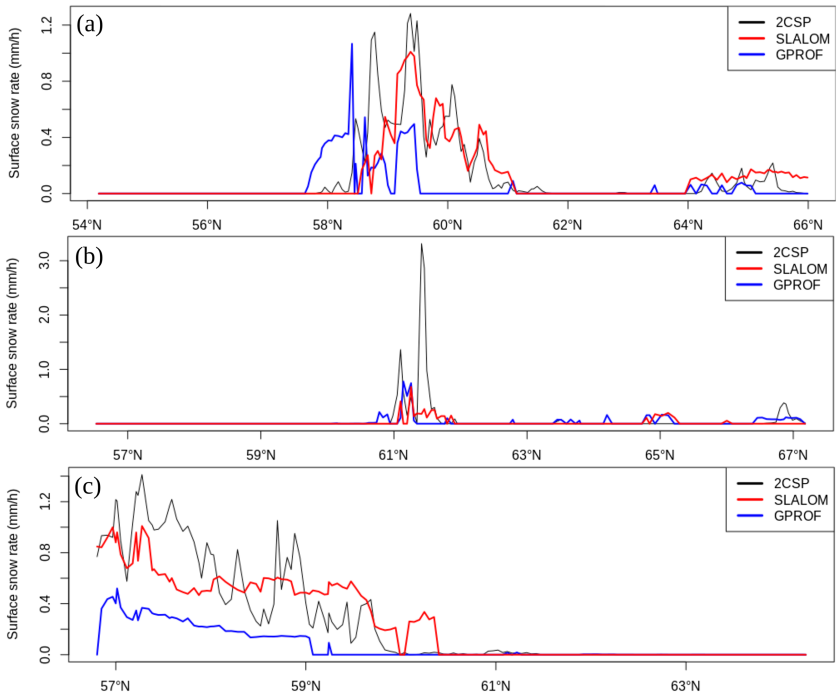


Figure 2. Surface snowfall rate from Cloud Profiling Radar-CloudSat (black), SLALOM (red), and GPROF (blue) for Snowfall event on (a) 30 April 2014, (b) 14 December 2014, and (c) 27 March 2014. SLALOM = Snow retrievaL ALgorithm fOr gMi; 2CSP = 2C-SNOW-PROFILE.

understand how snowfall rate retrievals are affected by environmental conditions (see also Rysman et al., 2018, for details about regression tree interpretation). The regression tree reveals first that predictions are less accurate when T2m is greater than 273 K. This is not surprising as for these temperatures mixed precipitation can occur and, while 2CSP product does not retrieve snowfall rate if the liquid fraction exceeds 15 %, GMI channels can still be affected by snowfall scattering in this situation and thus SSR-Mod can retrieve SSR. In addition, the quality of the retrieval also decreases when TPW is low (<7.1 kg/m^2) because in dry conditions snowfall detection can be problematic due to background surface contamination. Finally, the presence of supercooled droplets also impacts the retrieval negatively because the emission signal from the cloud liquid water tends to mask the ice scattering signal associated with snowfall (see Panegrossi et al., 2017). It is interesting to note that the obtained regression tree is quite similar to the one obtained for the SWP module (Figure 6a in Rysman et al., 2018).

5. Application of SSR-Mod

In the following, the new SSR-Mod is used to predict SSR on three case studies for which coincident GMI and CloudSat observations are available (Figure 2). Note that these case studies have been thoroughly analyzed by Panegrossi et al. (2017) and Rysman et al. (2018) and correspond to complex and insightful meteorological and environmental situations. We also included here the SSR provided by the NASA GMI product (GPROF V05, Kummerow et al., 2015) for a qualitative comparison. However it must be noted that comparing SLALOM (and 2CSP) with GPROF remains a difficult task as both products are based on different definitions. For instance, while 2CSP does not retrieve snowfall when liquid fraction is above 15%, GPROF V05 product retrieves snowfall as soon as the probability of having solid precipitation (liquid vs. solid classification methodology is detailed in Sims & Liu, 2015) is greater than 50%.

The first case study corresponds to a frontal system that occurred on 30 April 2014 in Eastern Siberia. CPR shows a main snowy region from 58°N to 61°N that corresponds to the frontal system and a weak snowfall region around 65°N (Figure 2a). Both regions variability is well captured by SLALOM. However, both maxima of snowfall higher than 1 mm/hr are underestimated by SLALOM, which is not surprising considering

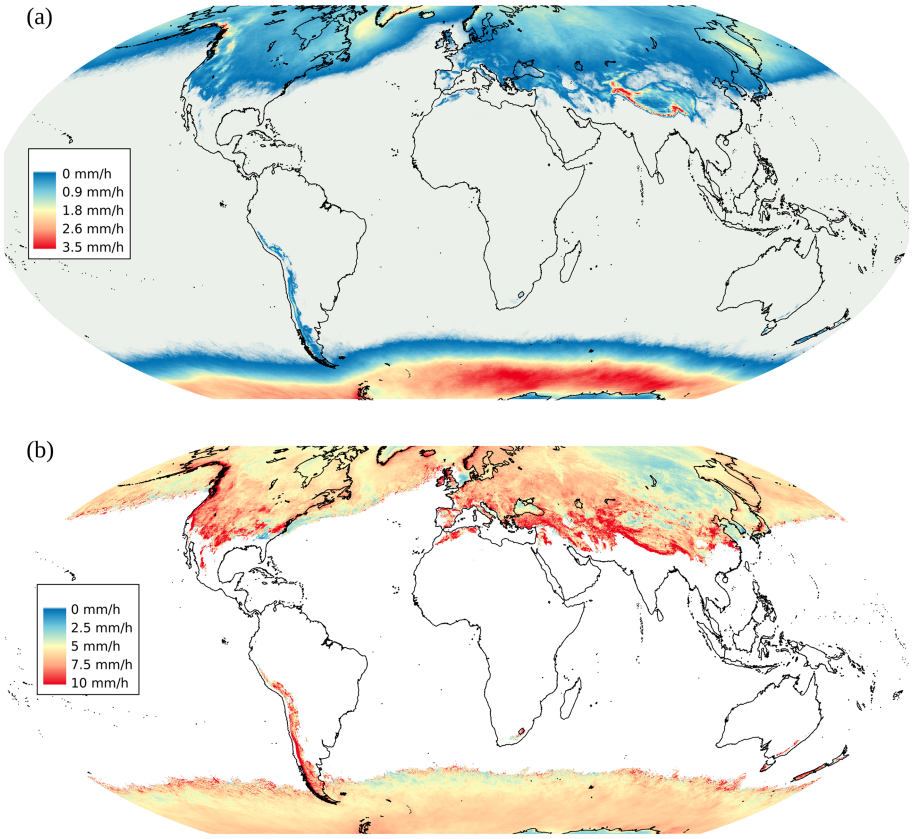


Figure 3. Unconditioned (a) and conditioned (i.e., for SSR > 0 mm/hr; b) average SSR (mm/hr), as estimated by the Snow retrievaL ALgorithm fOr gMi SSR module between May 2014 and May 2017. SSR = surface snowfall rate.

that a negative relative bias occurs for heavy snowfall events (Figure 1a). GPROF, on the other hand, is able to capture the precipitation around 58°N (likely mixed-phase, as evidenced also in Panegrossi et al., 2017), where it evidences two peaks of SSR (up to 1 and 0.4 mm/hr) undetected by CPR, while it misses almost completely the snowfall north of 60°N.

The second case study occurred on 14 December 2014 in Alaska and is associated with orographic precipitation (Figure 2b). While CloudSat shows an impressive peak of snowfall rate (>3 mm/hr) around 61.5°N, in spite of the relatively low SWP around 0.8 kg/m^2 (see Figure 8 of Rysman et al., 2018), neither SLALOM nor GPROF is able to reproduce it, but both algorithms find slight snowfall rate (around 0.5 mm/hr) in this region. As already mentioned in Rysman et al. (2018), this could be due to the low TPW (<5 kg/m^2) for this case, and to the contamination by the background surface of the weak scattering signal associated with this orographic snowfall event (as evidenced in Panegrossi et al., 2017).

The last case study took place over the Labrador Sea on 27 March 2014 and is associated with a synoptic snowfall event (Figure 2c). CloudSat shows a snowfall rate that decreases with intermittent peaks from 1.4 mm/hr at 57°N to 0 mm/hr at 61°N. SLALOM succeeds in retrieving the same snowfall pattern but again tends to underestimate snowfall rates higher than 1 mm/hr. GPROF captures well the trend but significantly underestimates the snowfall rate values and misses the precipitation north of 59°N. It is worth noticing that for all the three case studies shown here the SLALOM SWP retrieval shown in Rysman et al. (2018) showed better agreement with the 2CSP retrieval than the SSR. This confirms that GMI brightness temperatures respond mostly to the 3-D distribution of ice (or liquid) hydrometeors within the cloud, while the SSR retrieval is less directly associated with brightness temperature (see You & Liu, 2012).

In this last section, we analyzed the average SSR between 68°S and 68°N from May 2014 to May 2017. For each GMI orbit for this period, we applied SSR-Mod to retrieve the SSR. Then, we computed the unconditioned and conditioned (i.e., for SSR > 0 mm/hr) average SSR for each 0.1° × 0.1° pixel of a regular grid.

Note that in this analysis a "0" SSR can be anything (e.g., clear sky and rain) except SSR > 0 mm/hr. The two maps presenting the unconditioned and conditioned average surface snowfall intensity are shown in Figure 3 (respectively in a and b pannels). Figure 3a shows high average SSR values in the Southern Ocean region reaching 3.5 mm/hr in its eastern part. In the Northern Hemisphere, coastal regions of North America and Greenland show local maxima higher than 2 mm/hr. Labrador Sea, western Siberia, and Okhotsk Sea (along the Kamchatka Peninsula) also presents high average snowfall rate. Finally, heavy snowfall rates are found in most of mountainous regions (the Himalayas with local maxima above 3.5 mm/hr, the Andes, the Alps, and the Rocky Mountains). Those patterns are globally similar to those found in Behrangi et al. (2016), Kulie et al. (2016), Mateling (2016), and Skofronick-Jackson et al. (2019). Figure 3b highlights a rather smooth pattern of snowfall intensity between 5 and 7.5 mm/hr in the Southern Ocean region. The situation is more contrasted in the Northern Hemisphere where values reaching 10 mm/hr are found over mountainous ranges and for low latitudes. Surprisingly, the Europe-Asia continent is divided in two areas with very distinct behaviors: Over Europe and western Siberia the average snowfall rate intensity is mainly between 5 and 7.5 mm/hr with high variability, while over eastern Siberia the average intensity is between 2.5 and 5 mm/hr. Average surface snowfall intensity in North America decreases northward and away from the coastal regions. Finally, it is interesting to notice that some regions (South Africa and Australia) where average surface snowfall is extremely low as shown in Figure 3a can experience some rare but intense snowfall events.

6. Conclusions

In this letter, we presented a new module (SSR-Mod) of the SLALOM algorithm for retrieving SSR using GMI observations. The SSR-Mod module is trained and tuned using coincident observations between GMI and the CloudSat CPR. SSR-Mod also takes into account ancillary environmental variables, such as humidity and temperature. Our evaluation focused on the accuracy of SLALOM SSR retrievals with respect to CPR-derived SSR, while SLALOM snowfall detection capabilities have been evaluated in Rysman et al. (2018). SLALOM is able to retrieve SSR with a relative bias of −13%, a correlation coefficient of 0.7 and a RMSE of 0.08 mm/hr compared to CloudSat observations. Surface snowfall retrieval is more accurate in cold and/or moist environmental conditions and when supercooled droplets does not occur. In addition, the highest snowfall rates tend to be underestimated by SLALOM. SSR-Mod is used in this letter to provide high-resolution nearly global maps of unconditioned and conditioned (i.e., for SSR > 0 mm/hr) average SSR.

A significant contribution of the SSR-Mod module comes from the possibility to infer accurately the CPR SSR on the full GMI swath (880 km wide vs. 1.7 km for CPR). Indeed, while CPR is nowadays the most accurate instrument to retrieve surface snowfall at the global scale, its very narrow swath of 1.7 km implies a revisit time of two weeks for a 100 × 100-km grid pixel. This does not allow to provide a comprehensive climatology characterization of snowfall. Inferring CPR SSR on the full GMI swath allows to obtain a much higher spatial and temporal coverage of the Earth (between −68°S and 68°N) with a revisit time of about 1 day for 10 × 10-km grid pixels. In addition, SSR-Mod could be adapted to other available passive microwave radiometers, which could improve global surface snowfall characterization, including in the polar regions, which are not covered by GMI.

The ability of SSR-Mod to retrieve SSR with a good accuracy can be explained by several characteristics. First using all high-frequency channels (with dual-polarization at 89 and 166 GHz) help to deal with the problem of the vertical variability and microphysical variability of snowfall. Indeed, depending of their frequencies and of their polarizations, high-frequency channels are more or less sensitive to the different layers of a snow cloud and to the size or shape of snowflakes (e.g., Gong & Wu, 2017; Panegrossi et al., 2017). In addition, the use of low frequency channels provides information on the background surface at the time of the overpass without using climatology or ancillary information on the surface type. Another advantage of SSR-Mod comes from the fact that it uses environmental variables to adjust its retrieval strategy. For instance, a dryer atmosphere can increase the surface contamination for channels that probe closer to the surface. In these conditions, SSR-Mod will tend to favor channels that probe at higher altitudes to retrieve SSR.

However, as highlighted by Rysman et al. (2018), SLALOM algorithm and SSR-Mod present several limitations. The main one lies in the training data set that takes 2C-SNOW-PROFILE (2CSP) products as a reference, which is itself affected by several source of uncertainties that must be recalled here. First, 2CSP does

not provide measurements close to the surface to avoid ground clutter contamination, which is a potential limitation for providing accurate surface snowfall estimates. In addition, several authors (Cao et al., 2014; Chen et al., 2016; Norin et al., 2015) showed that, while 2CSP snowfall estimates agree favorably with those from ground-based radar products, it tends to overestimate light snowfall events and to underestimate moderate-heavy snowfall events. Another issue comes from the fact that 2CSP does not include melting snow or mixed-phase precipitation. Indeed, the 2CSP snowfall rate is provided only either when surface snowfall is probable or certain, or when the liquid fraction is lower than 15% (see also Casella et al., 2017). Therefore, CPR pixels where 2CSP SSR equals zero, considered as no-snowfall cases, might contain mixed-phase precipitation (as illustrated in the 30 April 2014 case presented in Figure 2a). However, the high-frequency GMI channels are sensitive to melting snowfall, even if the liquid fraction is higher than 15%. GMI instruments also struggle to detect surface snowfall in dry atmospheric conditions because the brightness temperatures measured in the high-frequency channels, sensitive to the snowfall scattering signal in most cases, are contaminated by the background surface (emission and scattering) signal. As a matter of fact, a recent study by Takbiri et al. (2019) showed the different GMI channel response to the presence of dry (and cold) or wet (warmer) snow on the surface. Melting snowfall events where solid and liquid water droplets/crystals coexist are also problematic cases. Indeed, it is unclear if, for these cases, high-frequency brightness temperatures are mainly impacted by solid or liquid phases or by a combination of both. Finally, supercooled droplets often found on top of ice clouds at high latitudes also impact the measured signal (e.g., Panegrossi et al., 2017) and are detected in a dedicated SLALOM module.

In the future, as more and more coincident observations will be available the algorithm training database will expand. This will help improving heavy snowfall events retrieval as they are currently underrepresented in the data set. It will be also interesting to explore the possibility to combine GMI and CPR coincident observations with DPR measurements, following Adhikari et al. (2018) approach, to improve heavy snowfall events characterization. Finally, as no consensus exists regarding global snowfall estimates, it will be insightful to compare the SSR climatology provided by SLALOM with those from other products.

Tools

This project has been conducted using the R software (version 3.5; RCoreTeam, 2013) with the following packages: rhdf5 (Fischer, 2015), ncdf4 (Pierce, 2012), rpart (Therneau et al., 2015), data.table (Dowle et al., 2014), randomForest (Liaw & Wiener, 2002), hydroGOF (Zambrano-Bigiarini, 2014), tree (B. Ripley, 2005), nabor (Elseberg et al., 2012), snow (Tierney et al., 2008), snowfall (Knaus, 2010), mlr (Bischl et al., 2016), mlrMBO (Bischl et al., 2017), xgboost (T. Chen et al., 2018), kknn (Schliep & Hechenbichler, 2016), svm (Meyer et al., 2017), and nnet (Venables & Ripley, 2002).

Acknowledgments
The data sets used in this study have been collected from the NASA PPS website ftp://arthurhou.pps.eosdis.nasa.gov (GPM products, 2B-CSATGPM product), from the CloudSat website ftp.cloudsat.cira.colostate.edu (CPR products), from http://www.icare.univ-lille1.fr/projects/dardar (DARDAR product), and from https://seaice.uni-bremen.de/sea-ice-concentration/ (AMSR2 sea ice data set). The PMM Research Program and EUMETSAT are warmly acknowledged for supporting the H SAF and GPM collaboration through the approval of the no-cost H SAF-GPM proposal, "H SAF and GPM: Precipitation algorithm development and validation activity." The authors want to sincerely express their gratitude to Joe Turk (NASA JPL), for developing and sharing the 2B-CSATGPM data set used in this study, and to Norm Wood, for the valuable discussions and suggestions on the use of the CloudSat 2C-SNOW-profile product. This work is supported by the EUMETSAT H SAF [CDOP-2 and CDOP-3]. NASA support for Mark Kulie (NASA Grant NNX16AE21G) and Lisa Milani (NASA GPM Mission) is also gratefully acknowledged. We would also like to thank the two referees for their valuable comments and suggestions.

References

Adhikari, A., Liu, C., & Kulie, M. S. (2018). Global distribution of snow precipitation features and their properties from 3 years of GPM observations. *Journal of Climate*, *31*(10), 3731–3754. https://doi.org/10.1175/JCLI-D-17-0012.1

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, *46*(3), 175–185. https://doi.org/10.1080/00031305.1992.10475879

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

Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., et al. (2016). mlr: Machine learning in R. *Journal of Machine Learning Research*, *17*(170), 1–5.

Bischl, B., Richter, J., Bossek, J., Horn, D., Thomas, J., & Lang, M. (2017). mlrMBO: A modular framework for model-based optimization of expensive black-box functions.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory – COLT'92* (pp. 144–152). Pittsburgh, Pennsylvania, United States: ACM Press. https://doi.org/10.1145/130385.130401

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L. (2017). *Classification and regression trees*. Abingdon, UK: Routledge.

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

Casella, D., Panegrossi, G., Sano, P., Marra, A. C., Dietrich, S., Johnson, B. T., & Kulie, M. S. (2017). Evaluation of the GPM-DPR snowfall detection capability: Comparison with CloudSat-CPR. *Atmospheric Research*, *197*, 64–75. https://doi.org/10.1016/j.atmosres.2017.06.018

Chen, S., Hong, Y., Kulie, M., Behrangi, A., Stepanian, P. M., Cao, Q., et al. (2016). Comparison of snowfall estimates from the NASA CloudSat Cloud Profiling Radar and NOAA/NSSL Multi-Radar Multi-Sensor System. *Journal of Hydrology*, *541*, 862–872. https://doi.org/10.1016/j.jhydrol.2016.07.047

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining – KDD'16* (pp. 785–794). San Francisco, CA: ACM Press. https://doi.org/10.1145/2939672.2939785

Chen, T., He, T., Benesty, M., Khotilovich, V., & Tang, Y. (2018). xgboost: Extreme gradient boosting.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Delanoë, J., & Hogan, R. J. (2010). Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *Journal of Geophysical Research*, *115*, D00H29. https://doi.org/10.1029/2009JD012346

Dowle, M., Short, T., Lianoglou, S., Saporta, R., Srinivasan, A., & Antonyan, E. (2014). data. table: Extension of data. frame. *R package version*, *1*(4).

Elseberg, J., Magnenat, S., Siegwart, R., & Nüchter, A. (2012). Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration. *Journal of Software Engineering for Robotics*, *3*(1), 2–12.

Eriksson, P., Jamali, M., Mendrok, J., & Buehler, S. A. (2015). On the microwave optical properties of randomly oriented ice hydrometeors. *Atmospheric Measurement Techniques*, *8*(5), 1913–1933. https://doi.org/10.5194/amt-8-1913-2015

Fischer, B. (2015). rhdf5-HDF5 interface for R. *R Package Version*, *2*.

Gong, J., & Wu, D. L. (2017). Microphysical properties of frozen particles inferred from Global Precipitation Measurement (GPM) Microwave Imager (GMI) polarimetric measurements. *Atmospheric Chemistry and Physics*, *17*(4), 2741–2757. https://doi.org/10.5194/acp-17-2741-2017

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

Knaus, J. (2010). Snowfall: Easier cluster computing (based on snow). *R Package Version*, *1*.

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

Kulie, M. S., Bennartz, R., Greenwald, T. J., Chen, Y., & Weng, F. (2010). Uncertainties in Microwave Properties of Frozen Precipitation Implications for Remote Sensing and Data Assimilation. *Journal of the Atmospheric Sciences*, *67*(11), 3471–3487. https://doi.org/10.1175/2010JAS3520.1

Kulie, M. S., Milani, L., Wood, N. B., Tushaus, S. A., Bennartz, R., & L'Ecuyer, T. S. (2016). A Shallow Cumuliform Snowfall Census Using Spaceborne Radar. *Journal of Hydrometeorology*, *17*(4), 1261–1279. https://doi.org/10.1175/JHM-D-15-0123.1

Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N.-Y., Ferraro, R., Joseph Munchak, S., & Petkovic, V. (2015). The evolution of the Goddard profiling algorithm to a fully parametric scheme. *Journal of Atmospheric and Oceanic Technology*, *32*(12), 2265–2280.

Kuo, K.-S., Olson, W. S., Johnson, B. T., Grecu, M., Tian, L., Clune, T. L., et al. (2016). The microwave radiative properties of falling snow derived from nonspherical ice particle models. Part I: An extensive database of simulated pristine crystals and aggregate particles, and their scattering properties. *Journal of Applied Meteorology and Climatology*, *55*(3), 691–708. https://doi.org/10.1175/JAMC-D-15-0130.1

Levizzani, V., Laviola, S., & Cattani, E. (2011). Detection and measurement of snowfall from space. *Remote Sensing*, *3*(1), 145–166. https://doi.org/10.3390/rs3010145

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

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

Mateling, M. E. (2016). Global snowfall: A combined reanalysis and spaceborne remote sensing perspective (Master Thesis). University of Wisconsin–Madison.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2017). e1071: Misc functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien.

Noh, Y.-J., Liu, G., Jones, A. S., & Haar, T. H. V. (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

Norin, L., Devasthale, A., L'Ecuyer, T. S., Wood, N. B., & Smalley, M. (2015). Intercomparison of snowfall estimates derived from the CloudSat Cloud Profiling Radar and the ground-based weather radar network over Sweden. *Atmospheric Measurement Techniques*, *8*(12), 5009–5021. https://doi.org/10.5194/amt-8-5009-2015

Panegrossi, G., Rysman, J.-F., Casella, D., Marra, A. C., Sano, P., & Kulie, M. S. (2017). CloudSat-based assessment of GPM microwave imager snowfall observation capabilities. *Remote Sensing*, *9*(12), 1263. https://doi.org/10.3390/rs9121263

Pierce, D. (2012). ncdf4: Interface to Unidata netCDF (version 4 or earlier) format data files. *R Package, URL Http://CRAN. R-Project. Org/Package = Ncdf4*.

Prigent, C., Aires, F., & Rossow, W. B. (2006). Land surface microwave emissivities over the globe for a decade. *Bulletin of the American Meteorological Society*, *87*(11), 1573–1584. https://doi.org/10.1175/BAMS-87-11-1573

RCoreTeam. (2013). R: A language and environment for statistical computing.

Ripley, B. (2005). Tree: Classification and regression trees. *R Package Version*, *1*.

Ripley, B. D., & Hjort, N. (1996). *Pattern recognition and neural networks*. Cambridge: Cambridge University Press.

Rysman, J.-F., Panegrossi, G., Sanò, P., Marra, A., Dietrich, S., Milani, L., & Kulie, M. (2018). SLALOM: An all-surface snow water path retrieval algorithm for the GPM Microwave Imager. *Remote Sensing*, *10*(8), 1278. https://doi.org/10.3390/rs10081278

Schliep, K., & Hechenbichler, K. (2016). kknn: Weighted k-Nearest Neighbors.

Sims, E. M., & Liu, G. (2015). A parameterization of the probability of snow–rain transition. *Journal of Hydrometeorology*, *16*(4), 1466–1477.

Skofronick-Jackson, G., & Johnson, B. T. (2011). Surface and atmospheric contributions to passive microwave brightness temperatures for falling snow events. *Journal of Geophysical Research*, *116*, D02213. https://doi.org/10.1029/2010JD014438

Skofronick-Jackson, G., Kulie, M., Milani, L., Munchak, S. J., Wood, N. B., & Levizzani, V. (2019). Satellite estimation of falling snow: A Global Precipitation Measurement (GPM) Core Observatory Perspective. *Journal of Applied Meteorology and Climatology*, *58*(7), 1429–1448. https://doi.org/10.1175/JAMC-D-18-0124.1

Skofronick-Jackson, G., Munchak, S. J., Ringerud, S., Petersen, W., & Lott, B. (2017). Falling Snow Estimates from the Global Precipitation Measurement (gpm) Mission. In *2017 Ieee International Geoscience and Remote Sensing Symposium (igarss)* (pp. 2724–2727). New York: Ieee.

Spreen, G., Kaleschke, L., & Heygster, G. (2008). Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal of Geophysical Research*, *113*, C02S03. https://doi.org/10.1029/2005JC003384

Takbiri, Z., Ebtehaj, A., Foufoula-Georgiou, E., Kirstetter, P.-E., & Turk, F. J. (2019). A prognostic nested k-nearest approach for microwave precipitation phase detection over snow cover. *Journal of Hydrometeorology*, *20*(2), 251–274. https://doi.org/10.1175/JHM-D-18-0021.1

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

Therneau, T., Atkinson, B., & Ripley, B. (2015). rpart: Recursive partitioning and regression trees. *R package version*, *4*, 1–10.

Tierney, L., Rossini, A. J., Li, N., & Sevcikova, H. (2008). Snow: Simple network of workstations. *R Package Version 0.3-3, URL Http://CRAN. R-Project. Org/Package = Snow*.

Turk, J. (2016). CloudSat-GPM coincidence dataset (version 1C). NASA Technical Report.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S (Fourth)*. New York: Springer.

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

You, Y., & Liu, G. (2012). The relationship between surface rainrate and water paths and its implications to satellite rain rate retrieval: RELATION BETWEEN RAINRATE AND WATER PATH. *Journal of Geophysical Research*, *117*, D13207. https://doi.org/10.1029/2012JD017662

You, Y., Wang, N.-Y., & Ferraro, R. (2015). A prototype precipitation retrieval algorithm over land using passive microwave observations stratified by surface condition and precipitation vertical structure. *Journal of Geophysical Research: Atmospheres*, *120*, 5295–5315. https://doi.org/10.1002/2014JD022534

Zambrano-Bigiarini, M. (2014). hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. *R Package Version*, *0*, 3–8.