Comparative analysis of snowfall accumulation and gauge undercatch correction factors from diverse data sets: in situ, satellite, and reanalysis

Milad Panahi[1], Ali Behrangi[1,2]

[1]University of Arizona, Department of Hydrology and Atmospheric Sciences
[2]University of Arizona, Department of Geosciences

Submitted to*: Asia-Pacific Journal of Atmospheric Sciences*

Corresponding author:

Ali Behrangi,
University of Arizona, Department of Hydrology and Atmospheric Sciences
Mail:1133 E. James E Rogers Way, Harshbarger Rm. 226b, Tucson, AZ, 85721-0011.
E-mail: Behrangi@email.arizona.edu

Key words: snowfall, SWE, gauge undercatch, satellite, reanalysis.

Abstract

Despite its importance for hydrology and water resources, accurate estimation of snowfall rate over snow-covered regions has remained a major observational challenge from both in-situ and remote sensing instruments. Snowfall accumulation can be measured by either accumulating snowfall estimates or measuring snowpack properties such as Snow Water Equivalent (SWE) and mass. By focusing on snowfall over snow accumulation period and using case studies and long-term average (2003 to 2015) over CONUS, this study compares snowfall accumulation from gauge stations (using GPCC and PRISM products), satellite products (GPCP and the suite of IMERG products), and reanalysis (ERA-interim, ERA5, and MERRA-2). Changes in SWE based on the recent UA-SWE product together with mass change observation from GRACE were used for assessment of precipitation products. We also investigated two popular gauge undercatch correction factors (CFs) used to mitigate precipitation undercatch in GPCC and GPCP. The results show that snow accumulation from most of the products is bounded by GPCC with and without correction, highlighting the critical importance of selecting proper CFs for gauge-undercatch correction. The CF based on Legates and Willmott method was found to be more consistent with the SWE-based analysis than CF based on the Fuchs method. Reanalysis show very similar spatial pattern among themselves, but represent large variation in simulating snow accumulation, with ERA-interim showing the least accumulation and MERRA-2 showing the highest accumulation and closest to the snow accumulation suggested by SWE.

1. Introduction

Accurate quantification of snowfall and its accumulation is critical for a wide range of applications including flood forecasting and water resources management. Snow acts as a large

natural reservoir and is an important source of freshwater during the warm season for many regions that rely on wintertime snowfall.

Precipitation, including snowfall, has commonly been measured by direct sampling through in-situ gauges that come in a variety of types and specifications. Snowfall measurement using gauges is particularly susceptible to undercatch issues as wind and turbulence can blow snow away from the mouth of a precipitation gauge (Rasmussen et al, 2012). While correction factors are used, they can be as large as a factor of three for snowfall (Legates and Willmott, 1990) and their uncertainty may exceed 100% (Yang et al., 2005; Fuchs et al., 2001). Furthermore, gauges provide point measurements with limited spatial representation as ground stations are often sparse or non-existent in cold and remote regions.

Remote sensing of precipitation provide areal estimate and can cover remote regions, but accurate estimation of snowfall from radar or satellites is challenging (Behrangi et al. 2014a, 2016a; Maggioni et al. 2017; Chen et al. 2016). Satellite-based precipitation estimates have been affected by insufficient sensitivity of sensors to light rain and snowfall (e.g., Behrangi et al. 2017; Smalley 2014), inaccurate knowledge of surface emissivities over frozen land (Ferraro et al. 2013), problems in distinguishing light rain and snow from cloud (Berg et al. 2006; Lebsock and L'Ecuyer 2011), insufficient understanding of precipitation microphysics (Bartsotas et al. 2018; Christensen et al. 2016), and difficulties in retrieving precipitation rate enhanced by orographic effect (Shige et al. 2013). Ground radar estimates can also contain large errors due to imperfect empirical relationships between radar reflectivity and precipitation rate, missing light snowfall (Chen et al. 2016), and mountain blockage. Furthermore, ground radars are available over few countries and most regions of the world are not covered by ground radars.

Bias adjustment of satellite products using in situ (i.e., rain gauge) measurements is a common practice (Adler et al. 2003, Huffman et al. 2007) and often results in major improvement of satellite or radar products for hydrologic applications (Behrangi et al. 2011, 2014b). However, for snowfall measurement, in situ data also have shortcomings (e.g., undercatch in gauges, sparseness of gauges over cold regions, and point versus areal measurement). Therefore, researchers have tried to also investigate other alternative methods.

Snowfall accumulation could be measured by either accumulating snowfall estimates or measuring snowpack properties (e.g., height, water equivalent, or mass). In theory, the two types of measurement should yield consistent results and differences should be minimum in an ideal case. For example, storage change observation from the Gravity Recovery and Climate Experiment (GRACE) has been used to quantify large scale bias in popular precipitation products over Tibetan plateau (Behrangi et al. 2017) and high latitudes (Behrangi et al. 2018a). Another alternative is using estimates of Snow Water Equivalent (SWE). In a proof of concept study, Tian et al. (2014) used SWE from AMSR-E Daily Global Snow Water Equivalent data (Kelly 2009; Tedesco and Narvekar 2010) and showed that uncertainty in determining snow accumulation using SWE, albeit large, is still lower than direct precipitation measurements over. The study was performed over the northern hemisphere and during snow accumulation period (e.g., winter-time), when snow ablation is minimum. In a more recent study, Behrangi et al. (2018b) used fairly accurate high resolution (i.e., 100 m) SWE estimates from the Airborne Snow Observatory (ASO) over two mountainous basins in California to estimate monthly snowfall accumulation. By comparing SWE accumulations from ASO, they evaluated a suite of in-situ based and remote sensing precipitation products and concluded that satellite and radar products largely underestimate snowfall accumulation compared to ASO (up to 50%), but can

capture monthly spatial patterns as indicated by overall correlation coefficients greater than 0.8 with ASO snow accumulation. However, ASO observation has been limited to few basins in California and Colorado and surveys are often performed during the melting season, limiting its application for assessment of snowfall accumulation across different regions and years.

In the present study we perform a comparative analysis of snow accumulation among several gridded precipitation products from satellites, reanalysis, and in situ measurements, guided by the recently developed UA-SWE product (Broxton et al. 2016). The analyses are then used to assess two popular gauge undercatch correction factors. In our analyses we also use GRACE observation to add additional insights on snow accumulation and quality of the UA-SWE product.

2. Method and dataset

We conduct our comparative analysis for snow accumulation period, defined as a period in which daily maximum surface air temperature stays below 1ºC, continuously for at least 30 days, and snow accumulation based on UA-SWE exceeds 6 mm. A case study is defined if at least 15 connected grids at 0.5°×0.5° resolution pass the snow accumulation criteria. These criteria minimize snow melting and allow comparing snowfall accumulation from UA-SWE and various precipitation products. Consequently, the study region is often located in North or Northeast CONUS and incudes periods that often coincide with winter months. UA-SWE and precipitation rates from various products were mapped onto a common 0.5°×0.5° grid within the identified regions prior to comparison.

The use of UA-SWE for snowfall accumulation study follows the mass conservation principle dictating that any change in one component of the water balance must be compensated for by the

same amount collectively in the other components (e.g. Dingman 2008). In other words, SWE accumulation can be used to estimate precipitation accumulation when sublimation is accounted for and there is no snowmelt and runoff. According to the temperature criteria used to define snow accumulation period, no snowmelt or runoff is expected. Similar concept has also been used for precipitation analysis using GRACE Total Water Storage Anomaly (TWSA) observations (e.g., Behrangi et al. 2017). Therefore, in this study we also use GRACE data to investigate the consistency between GRACE and UA-SWE observations. We used GRACE data for case studies where and when possible, because the reported timing for GRACE may not exactly match the accumulation period and due to the coarse resolution of GRACE the use of GRACE requires a fairly large region of connected grids. We used the following data sets:

2.1. In situ data

2.1.1. PRISM

Gridded observations of daily precipitation rate and maximum air temperature are obtained from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset (Daly et al. 1994; 2002) developed by the PRISM Climate Group at Oregon State University. PRISM data cover the CONUS and are high quality meteorological station observations interpolated to a 4-km grid using a human expert and statistical knowledge-based system that utilizes a digital elevation model (http://www.prism.oregonstate.edu). PRISM incorporates data from about 13,000 stations for precipitation and nearly 10,000 stations for maximum and minimum temperature across the CONUS. The PRISM products use a weighted regression scheme to account for complex climate regimes associated with orography, rain shadows, temperature inversions, slope aspect, and coastal proximity, and other factors. However, PRISM does not

155 account for gauge undercatch issue (personal communication with C. Daly in 2019) and thus
156 PRISM may largely underestimate snowfall rate. PRISM data are available at daily and monthly
157 time scales. We used daily 4 km PRISM data sets that are available since 1981.
158
159 2.1.2. GPCC
160 The Global Precipitation Climatology Centre (GPCC) product established in 1989 under World
161 Meteorological Organization (WMO). GPCC produces station-based gridded precipitation data
162 that covers global land (except Antarctica) using station data from several networks (Schneider
163 et al. 2014). GPCC uses interpolation of sparse data to produce a gridded map (Schneider et al.
164 2017). GPCC offers various products including: (1) the Full data daily (V 2018), currently
165 available at 1°×1° resolution from 1982 through 2016, incorporating all stations with the highest
166 quality-control process, but is not available in real-time and does not correct for gauge under-
167 catch problem; and (2) the latest Monitoring product (V6) that uses records from fewer gauging
168 stations and is available from 1982 to 3 months before present (Schneider et al. 2018). The V6
169 product also includes an empirical gauge correction factor (CF). The CF is based on a method
170 described in Fuchs et al. (2001), hereafter referred to as CF-F, that uses a dynamic correction
171 model (Rubel and Hantel 1999) and employs various synoptic observations such as air
172 temperature, relative humidity, precipitation intensity and wind speed. In case CF-Fs cannot be
173 derived (e.g., because of missing synoptic observations), Legates and Willmott climatology
174 (Legates and Willmott 1990) can be used as base pattern (see Fuchs et al. 2001 for further
175 details). The Legates and Willmott correction method is a monthly spatially varying climatology
176 of undercatch estimate using the best guess of gauge types around the world as of the early
177 1980's. Legates and Willmott (1990) generated a global climatology of monthly precipitation by

178 removing systematic errors caused by wind, wetting on the interior walls of the gauge, and
179 evaporation from the gauge observations. The Global Precipitation Climatology Project (GPCP)
180 (Huffman et al. 2009; Adler et al. 2018) uses CFs from Legates and Willmott (1990) to correct
181 GPCC before combining it with satellite data.
182
183 2.1.3. UA-SWE
184 UA-SWE uses in situ data from the National Resources Conservation Service (NRCS) Snowpack
185 Telemetry (SNOTEL; Serreze et al. 1999) network and the National Weather Service's (NWS)
186 Cooperative Observer Program (COOP) network in the United States. The COOP snow depths
187 are converted to SWE using a relationship between the SNOTEL SWE and snow depth data at
188 the time of maximum snow accumulation. As discussed in Broxton et al. (2016a and 2016b),
189 SWE measurement at each station is divided by the amount of accumulated snowfall (from the
190 time that no snow was on the ground to the time of the SWE estimate) measured at the station
191 minus estimate of accumulated ablation (for the same period), which is given by a simple
192 temperature index snow model. These ratios are then interpolated to the PRISM 2.5-arc-min (4
193 km) grid and multiplied by net snow accumulation (i.e., accumulated snowfall minus the
194 estimate of accumulated ablation for each PRISM grid). Evaluation of UA-SWE using the
195 Airborne Snow Observatory (ASO) (Painter et al., 2016) and comparison with the Snow Data
196 Assimilation System (SNODAS; Barrett 2003) have shown remarkable skill over the CONUS
197 (Braxton et al. 2016b; Dawson et al. 2018). The UA-SWE product is available at 4km x 4km
198 spatial resolution and daily time scale since October 1981.
199
200 2.2. Satellite data

2.2.1 GPCP

GPCP is a widely-used community-based analysis of global precipitation under the auspices of the World Climate Research Program (WCRP) and the Global Energy and Water Exchanges (GEWEX) project. GPCP is a combined satellite-gauge product. Satellite data are used over land and ocean and obtained from the Special Sensor Microwave Imager (SSM/I), Special Sensor Microwave Imager/Sounder (SSMIS), geostationary imagers and polar orbiting infrared sounders [i.e., the Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) and Atmospheric Infrared Sounder (AIRS)]. GPCP uses GPCC full product as its in situ component, but prior to combination with satellite data, GPCC estimates are adjusted for gauge undercatch using the CFs suggested by Legates and Willmott (1990), hereafter we refer to them as CF-LW. Gauge measurements are averaged with satellite estimates using inverse estimated error variance as the weights (Huffman et al. 1997). Several upgrades have been made to the GPCP since its initial release (Huffman et al. 1997; Adler et al. 2003; Adler et al. 2018). We used the latest version of daily GPCP climate data record (Version 1.3) (Huffman et al. 2001; Adler et al. 2018), providing globally complete precipitation estimates at a spatial resolution of 1°×1° from October 1996 to few month before present. While it uses slightly different data sets and methods than that used in the GPCP Monthly analysis, the GPCP Daily sums up to the GPCP Monthly (V2.3) (Adler et al. 2018).

2.2.2 IMERG

The Integrated Multi-satellite Retrievals for GPM (IMERG) is the unified U.S. algorithm and its latest version (V6) provides half-hourly 0.1°×0.1° global map of precipitation since 2000. The precipitation estimates from the various precipitation-relevant satellite passive microwave

(PMW) sensors comprising the GPM constellation are computed using the 2017 version of the Goddard Profiling Algorithm (GPROF2017), intercalibrated to the GPM combined instrument product, gridded, and combined into half-hourly 0.1°×0.1° fields. Detailed description and derivation of the IMERG products are provided in Huffman et al. (2018) and in brief includes four main steps: 1) Precipitation estimates from the GPM constellation radiometers are gridded, intercalibrated to 2BCMB, and combined into half-hourly 0.1°x0.1° fields (variable name: "HQprecipitaiton"), 2) Maps of half-hourly IR precipitation rate ("IRprecipitation") are calculated using an IR-based precipitation retrieval method (PERSIANN-CCS; Hong et al. 2004), 3) By utilizing the Climate Prediction Center Morphing-Kalman Filter (CMORPH-KF) Lagrangian time interpolation scheme (e.g., Joyce et al. 2011), MW and IR estimates are used to create half-hourly estimates ("precipitationUncal") , and 4) the multi-satellite half-hour estimates are adjusted so that they sum to a monthly satellite-gauge combination ("precipitationCal") that employs GPCC monthly gauge analysis after being corrected by CF-LW.

This paper uses "precipitationCal", "precipitationUncal", "IRprecipitation", "HQprecipitaiton" fields from the Level 3 IMERG V6 "Final" half-hourly 0.1°×0.1° product (GPM_3IMERGHH.06), referred to by IMER-F-Cal, IMERG-F-Uncal, IMERG-IR, and IMERG-HQ, respectively.

2.2.3 GRACE

GRACE was a joint mission of NASA and the German Aerospace Center. Twin satellites provided measurements of Earth's gravity field anomalies from March 2002 to October 2017, the end of its mission. The continuation of the mission enables using the GRACE follow-on mission, with its twin satellites launched in May 2018. TWSA data are obtained from GRACE mass

concentration (mascon) data product (RL06M V1), representing the 3°×3° equal-area caps, downloaded from ftp://podaac-ftp.jpl.nasa.gov/allData/tellus/L3/mascon/RL06/JPL/CRI/netcdf/). This version includes improvements in both background geophysical models and orbital parameterization and reduces uncertainty and bias due to leakage effects. The coarse resolution GRACE mascon solution is downscaled to 0.5°×0.5° grid by multiplying it by a scaling factor grid available in https://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/.

2.3 Reanalysis products

Considering the lack of reliable observational data sets over cold regions, reanalysis precipitation products provide viable insights on precipitation estimates (Palerme et al. 2014; Behrangi et al. 2016b). Reanalyses products are based on the assimilation of various in situ and remote sensing observations into an atmospheric general circulation model and are among the most widely used datasets in Earth science. We used three popular reanalysis products: The recent Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2) (Reichle et al. 2017, Gelaro et al. 2017), ERA-Interim (Dee et al., 2011; Simmons et al., 2010), and ERA5 (C3S 2017). MERRA-2 was introduced to replace and extend the original MERRA dataset, which ended in February 2016. The MERRA-2 dataset is available at subdaily 0.5°×~0.67° resolution. It includes recent upgrades to the atmospheric assimilation that enable the use of newer satellite observations that could not be assimilated in the original MERRA system. Additionally, MERRA-2 benefits from the latest advances in the Goddard Earth Observing System, version 5 (GEOS-5). In this study we used daily data obtained from Goddard Earth Sciences Data and Information Services Center (GES DISC). ERA-Interim is the European Centre for Medium-Range Weather Forecasts global atmospheric reanalysis. ERA-interim relies on a 4D-VAR

system which uses observations within 12 h windows to initialize forecast simulations. ERA-interim treated several problems that were identified in ERA-40 by improving methods, in particular the hydrological cycle, the temporal consistency of the analyzed fields, and the quality of the stratospheric circulation. This study uses half-daily precipitation obtained from http://apps.ecmwf.int/datasets/ at 0.5°×0.5° spatial resolution. ERA Interim is being phased out, and the last date to be made available in ERA Interim will be 31 August 2019. ERA5 data released as the substitution for migration form ERA-interim. ERA5 is available from 1979 to 2-3 months before the present, claiming significantly improvements with spatial and temporal resolution and global balance of precipitation and evaporation. ERA5 was produced using 4D-Var data assimilation. The ERA5 dataset contains one (31 km) high resolution realisation (HRES) and a reduced resolution ten member ensemble (EDA). Generally, the data are available at a sub daily and monthly frequency and consist of analyses and short-term (18 hour) forecasts, initialized twice daily from analyses at 06 and 18 UTC. Most analyzed parameters are also available from the forecasts. Here we used 0.25°×0.25° hourly precipitation data from ERA5 provided by Copernicus Climate Change Service (C3S; 2017).

3. Results

Figure 1 shows seasonal maps of CF-LW and CF-F, calculated by averaging 13 years (2003-2015) of monthly CFs, over CONUS. CFs are largest in winter and smallest (around one) in summer, mainly because gauge undercatch is larger when it snows than when it rains. Snow particles have a much larger drag force than rain drops, so wind and turbulence can be more effective to blow the snow away from the mouth of a precipitation gauge (Rasmussen et al, 2012). However, CF-F is much larger than CF-LW over most of the CONUS (i.e., mountains and

293 cold region) in winter and western half of CONUS in other seasons. In winter, CF-F values
294 exceed 2.5 and differ from CF-LW by more than 40% over a wide region in the northern and
295 central U.S (Fig 1.c). This means that depending on the CF used, in situ estimates of winter
296 precipitation might significantly differ from each other, an important issue that often is not
297 considered.
298 It is important to assess the role of CFs in the context of other precipitation inter-comparisons as
299 most of the precipitation products face challenges in quantifying winter precipitation (especially
300 snowfall) and they tend to utilize in situ data in their final satellite-gauge product. Furthermore,
301 the popular GPCP and GPCC monitoring products use different CFs (i.e., GPCP uses CF-LW
302 and GPCC monitoring uses CF-F) to correct for gauge undercatch.
303 As discussed in the methodology section, we focus on snow accumulation period, in which no
304 melting is expected. This period represents the largest CFs and largest differences between CF-F
305 and CF-LW (Fig. 1; the first row). Focusing on snow-accumulation-period also enables to bring
306 together insights from independent snow accumulation estimates from UA-SWE and GRACE
307 gravity measurements.
308 Figure 2 shows maps of snow accumulation for a case study, between Dec 30 2010 and Feb 5
309 2011. GPCC (Fig 2b) produces the lowest accumulation across the entire region. The
310 precipitation products include PRISM, GPCC 2018, GPCC-F (i.e., GPCC corrected by CF-F),
311 GPCC-LW (i.e., GPCC corrected by CF-LW), GPCP daily V1.3, IMERG fields (Cal, Uncal, IR,
312 and HQ), ERA-Interim, ERA5, and MERRA-2. It can be seen that the overall pattern of PRISM
313 (Fig. 2a) is similar to GPCC (Fig. 2b), but it has several grids with elevated snow accumulation
314 compared to GPCC. The observed increase in these grids is likely because PRISM uses
315 additional in situ data, has higher resolution than GPCC, and accounts for several factors (e.g.,

topography, etc.) in its high resolution interpolation scheme (Section 2.1.1). As a result, PRISM exceeds GPCC by about 7 mm snow accumulation at the end of accumulation period (Fig. 3). GPCC-F (Fig. 2c) and GPCC-LW (Fig. 2d) show large increase in snow accumulation compared to GPCC, but GPCC-F shows a much larger increase than GPCC-LW, as CF-F has significantly larger values than CF-LW (Fig. 1). While one may expect more similarities between GPCP (Fig. 2e), IMERG-F-Cal (Fig. 2f) and GPCC-LW (e.g., because both GPCP and IMERG-F-Cal use GPCC-LW for satellite-gauge analyses), regional differences can be observed due to contribution of satellites with different resolutions, differences between calendar month and study period, and slight differences in combination approach in GPCP and IMERG-F-Cal. IMERG-F-Uncal (Fig. 2g) and IMERG-IR (Fig. 2h) are similar. This is because IMERG-F-Uncal uses IMERG-IR precipitation estimates over snow and ice surface where current passive microwave precipitation estimate (here IMERG-HQ; Fig 2i) is believed to be low quality (based on personal communication with G. Huffman in April 2019). Among the reanalysis products, ERA-Interim (Fig. 2j) shows the smallest, MERRA-2 (Fig. 2l) shows the largest, and ERA5 (Fig. 2k) shows medium snow accumulations that often fall between the two. The three reanalysis products show a fairly comparable spatial pattern, but grid-level differences can also be seen. Figures 2m and 2n show the UA-SWE and UA-SWE with no ablation (hereafter UA-SWE-NA). The overall magnitude of UA-SWE is comparable to GPCC-LW, but for comparison with snowfall accumulation snow ablation needs to be added to UA-SWE. By adding the snow ablation, the UA-SWE-NA (representing UA-SWE with no ablation subtracted) suggests slightly larger values than GPCC-LW, but not as high as GPCC-F. In other words, UA-SWE-NA favors CF-LW over CF-F.

Figure 3 shows area-averaged daily cumulative snowfall from several products for the case study shown in Figure 2. UA-SWE and UA-SWE-NA are shown in blue and dark blue colors, respectively. UA-SWE-NA suggests that GPCC largely underestimates and GPCC-F largely overestimates snow accumulation, but GPCC-LW, GPCP, and IMERG-F-Cal produce snow accumulations that are comparable to UA-SWE-NA over the calendar month period (Jan 1 to Jan 31 2010). Note that GPCP and IMERG-F-Cal use GPCP-LW to perform bias adjustment at the end of calendar month, so comparison between Dec 30 2009 and Feb 5 2010 does not warranty similarity among the bias adjusted products. In fact, considerable differences appear among these product for the Dec 30 2009 to Feb 5 2010 accumulation period, mainly because the products take different paths in their daily accumulation. For example, during the first few days of accumulation, GPCP followed by IMERG-F-Cal show a slow rise, deviating from GPCC, PRISM, reanalysis and UA-SWE products. IMERG-HQ significantly underestimates snow accumulation (produces about 10% of snow accumulation suggested by UA-SWE-NA and GPCC-L) with accumulation pattern that does not follow other products. IMERG-IR produces about 50% of the accumulation suggested by UA-SWE and GPCC-LW, but suggest a rapid increase around Jan 20 2010 that is not observed in SWE, GPCC, PRISM and reanalysis products. Consequently, this increase is also seen in IMERG-F-Cal and IMERG-F-Uncal as they are based on IMERG-IR over frozen land. While PRISM has larger values than GPCC, it underestimates snow accumulation compared to both GPCC-LW and GPCC-F, partly because PRISM does not utilize any correction for gauge undercatch. Reanalysis products tend to show smaller snow accumulation than GPCC-LW and GPCC-F, among them MERRA-2 shows the highest accumulation and is closest to GPCC-LW and UA-SWE-NA. ERA-Interim is close to PRISM that does not account for undercatch correction, so it likely underestimates. ERA5 falls

between ERA-interim and MERRA-2 and is close to GPCP. While PRISM and reanalysis products show different total snow accumulation, they agree fairly well in capturing the daily accumulation pattern.

GRACE reported an observation centered on Dec 14 2010, representing the entire month of December, and another one centered on Feb 18 2011, representing the month of February (the dates are shown with black arrows in Fig. 5). The change in TWS between the two dates is 41.9 mm, only 5% more than the corresponding change in SWE (i.e., 39.9 mm), suggesting a great consistency between the observed changes in TWS and SWE.

Similarly, Figures 4 and 5 display another case study from Dec 30 2008 to Jan 31 2009. GPCC-LW shows similar snow accumulation map (Fig. 4), similar daily accumulation pattern (Fig. 5), and an almost identical total snow accumulation (Fig. 5) compared to UA-SWE-NA. On the other hand, GPCC-F produces snow accumulations that exceed all other products. IMERG-HQ produces the lowest snow accumulation, but performs better than previous case study, producing about 30% of total snow accumulation from UA-SWE-NA and GPCC-LW. IMERG-HQ also captures the timing of relatively rapid increase in snow accumulations starting around Jan 7, Jan 17, and Jan 27 2009 (Fig. 5), but with smaller rates than other products. However, Figure 4 shows that IMERG-HQ is unable to produce a spatial map similar to that obtained from gauge-based or reanalysis products. Interestingly, IMERG-IR produces total snow accumulation that is very close to that estimated by GPCP, IMERG-F-Cal and UA-SWE-NA (Fig. 5). Note that it shows different daily accumulation pattern (Fig. 5) and spatial distribution map (Fig. 4) compared to the UA-SWE-NA and reanalysis products. The similarity of IMERG-F-Uncal to IMERG-IR is not surprising as over snow surfaces, IMERG-F-Uncal uses IMERG-IR. GPCC

385 provides the second lowest total snow accumulation among all of the products (Fig. 5). PRISM
386 exceeds GPCC by total snow accumulation of about 8 mm on Jan 31 2009, but the two products
387 show large similarities in daily snow accumulation pattern (Fig. 5) and spatial map (Fig. 4),
388 except in east of the Great Lakes. Similar to the previous case study, MERRA-2 and ERA-
389 Interim produce larger and smaller snow accumulation than ERA5, respectively, but all of them
390 show fairly similar spatial maps and daily accumulation pattern.
391
392 GRACE reported an observation centered on Jan 16, representing the entire month of January,
393 and another one centered on Feb 15, representing the entire month of February (the dates are
394 shown with black arrows in Fig. 5). The change in TWS between the two dates is 4.86 mm, only
395 slightly (0.27 mm) more than the corresponding change in SWE (i.e., 4.59 mm), suggesting a
396 great consistency between the observed changes in TWS and SWE.
397
398 By averaging all of the snow accumulation cases across the study years (2003-2015), Figures 6
399 and 7 show maps of two distinct regions over the North and Northeast United States. We plotted
400 these two figures separately to improve the visualizations as snowfall accumulation over the
401 Northeast U.S. is considerably larger than snow accumulation west of the Great Lakes. The
402 results are generally consistent with the case studies. In particular, GPCC with no correction
403 (followed by PRISM) provides the least snow accumulation among all of the studied products,
404 GPCC-F overestimates snow accumulation compared to all other products, GPCC-LW has better
405 overall agreement with UA-SWE-NA (although not perfect) than GPCC-F, ERA-Interim and
406 MERRA-2 produce the lowest and highest snow accumulation respectively among reanalysis
407 and MERRA-2 is more consistent with UA-SWE-NA and GPCC-LW in terms of overall

magnitude. IMERG-F-Cal, however, shows smaller snow accumulation than GPCP and GPCC-L, especially over northeast (Fig. 7). This could be partly because GPCP and IMERG-F-Cal are bias adjusted at calendar-month with GPCC-L, but the snow accumulation periods studied here often do not fully match calendar month. GPCP and IMERG-F-Cal use different satellite products with different resolutions that may also contribute to the observed differences. Figures 6 and 7 also show that IMERG-F-IR and IMERG-HQ by themselves are not robust in capturing the magnitude and pattern of snow accumulation.

The overall cross-correlation and bias ratio among the products are summarized in Figures 8a and 8b. The correlation coefficient (CC) map (Fig 8a) suggests that the products are generally well correlated with each other (i.e., CC > 0.8 for all pairs except those that involve one of the IMERG-IR, IMERG-HQ or IMERG-F-Uncal products). UA-SWE and UA-SWE-NA shows slightly better correlation with GPCC-LW than GPCC-F, although both correlations are high (exceeding 0.9). Figure 8a also shows that reanalysis products are almost identical in capturing spatial patterns among themselves (i.e., CC>0.98; see the box with bright color in the middle of the Fig 8a), but they can have large differences in capturing snow accumulation amount (Fig. 8b), Bias ratios in Figure 8b is calculated by dividing the total snow accumulation from the products listed in the Y-axis by the total snow accumulation from the products listed in the X-axis. The value 1 means that the two products are almost identical in terms of capturing the total snow accumulation. From the UA-SWE-NA perspective, GPCC-F shows the largest and IMERG-HQ followed by GPCC and IMERG-IR show the lowest snow accumulation values among all of the products. MERRA-2 and GPCC-LW, followed by GPCP and IMERG-F, show the closest snow accumulation amount to UA-SWE-NA.

Figure 9 summarizes the relative skill of the studied products, each shown by a different letter on the diagram, using Taylor diagram (Taylor 2001) and from the UA-SWE-NA perspective.

The distance between each product and UA-SWE-NA, here used as comparison reference and represented by letter "A" on the X axis, is a measure of similarity between the products and the reference. For each product, three statistics are shown: (1) the Pearson correlation coefficient, a measure of similarity in pattern between the products and reference fields (blue contours representing the azimuthal angle), (2) the centered RMS error in the products is proportional to the distance from point "A" and is (shown in green contours); and (3) the standard deviation of the products (shown in black contours) that is proportional to the radial distance from the origin. The diagram shows, for example, that UA-SWE, GPCC-LW, MERRA-2, and ERA-5 are the closest to UA-SWE-NA, while IMERG-HQ, IMERG-IR, IMERG-F-Uncal are the least similar products to UA-SWE-NA. Therefore, from UA-SWE-NA point of view, CF-LW is preferred over CF-F to correct GPCC.

4. Concluding remarks

Quantitative knowledge on snowfall accumulation is critical for water resources management and hydrologic applications such as flood forecasting and drought analysis. However, accurate estimation of snowfall rate over snow-covered regions has remained a major observational challenge from both in-situ and remote sensing instruments. In situ observation are mainly composed of gauges that are not necessarily designed for snowfall measurement, thus they have to use large post-corrections to mitigate the effect of snowfall undercatch. There is a large uncertainty in correction factors (CFs) used for compensating the gauge undercatch, for example,

454 the two popular correction factors (i.e., CF-F and CF-LW) can be different by more than 50%
455 (Fig. 1). Satellite products also face problems for snowfall estimation over snow covered area as
456 they have to overcome several obstacles such as uncertain emissivity over land, weak sensitivity
457 of instrument to snow particles, and other challenges discussed in Section 1.
458 Here, by focusing on snowfall over snow accumulation period, we compared snowfall
459 accumulation from rain gauges (through GPCC and PRISM), satellite products (GPCP and the
460 suite of IMERG products), and reanalysis (ERA-interim, ERA5, and MERRA-2) with changes in
461 SWE based on the recent observational SWE product (UA-SWE) and after removing ablation
462 terms from that (UA-SWE-NA). The analysis were performed using case studies as well as long-
463 term average over 13 years (from 2003-2015). GRACE observation of total water storage
464 anomaly was also utilized to verify UA-SWE where and when possible.
465
466 The results suggests that the use of CF-F, for correction of snowfall undercatch in GPCC, results
467 in significant overestimation compared to UA-SWE-NA and all of the other products studied
468 here. UA-SWE-NA, however, supports CF-LW and is also reasonably consistent with MERRA-
469 2, and ERA-5, based on Tylor diagram analysis. Nonetheless, averaged over all case studies, the
470 difference between the GPCC and GPCC-F snowfall accumulation was found to encompass
471 snow accumulation from all of the other products, excluding IMERG-HQ. This highlights the
472 critical importance of selecting proper CFs for gauge-undercatch correction. While this study
473 favored CF-LW over CF-F for snowfall undercatch correction over CONUS, previous studies
474 over Eurasia and Siberia showed an opposite conclusion based on mass change analysis and
475 water vapor convergence (Behrangi et al. 2014c, 2018a). Different countries use different

476 gauges and the CF's performance might be gauge-type, regional or regime dependent. This calls
477 for future investigation to further check the performance of these correction methods.
478
479 PRISM shows higher snow accumulation than GPCC but often lower values than GPCC-LW and
480 GPCC-F, perhaps because PRISM does not utilize gauge undercatch correction factors.
481 Reanalysis show very similar spatial pattern among themselves, but represent large variation in
482 simulating snow accumulation, with ERA-interim showing the least accumulation and MERRA-
483 2 showing the highest accumulation and closest to UA-SWE-NA. For the cases studied, a great
484 consistency between GRACE observation of total water storage and SWE-UA was observed.
485
486 This study provides an example for how a diverse set of observations (e.g., in situ, satellite
487 precipitation, SWE, and mass change) can be used to provide recommendation on gauge
488 correction schemes and understanding the uncertainties exist in quantifying snowfall
489 accumulation over frozen land. It is hoped that with further improvements in remote sensing of
490 precipitation (e.g., under efforts supported by GPM), in situ observations (e.g., more appropriate
491 gauges and undercatch correction schemes), SWE estimates (from both in situ and satellites), and
492 mass change products (e.g., efforts supported under GRACE-FO), further convergence among
493 the products can be achieved.
494
495 **Acknowledgments**
496 The GPCP daily V1.3 was obtained from National Centers for Environmental Information
497 (NESDIS), https://www.ncei.noaa.gov/data/global-precipitation-climatology-project-gpcp
498 daily/access/, IMERG and MERRA-2 data were obtained Goddard Earth Sciences Data and

Information Services Center (GES DISC). ERA-Interim data was accessed from the ECMWF website, www.ecmwf.int/en/forecasts/datasets/archivedatasets/reanalysis-datasets/era-interim. ERA5 data provided by Copernicus Climate Change Service (C3S). PRISM was obtained from the Climate Group at Oregon State University (http://www.prism.oregonstate.edu), GPCC data were obtained from Deutscher Wetterdienst (DWD) (https://www.dwd.de/EN/ourservices/gpcc/gpcc.html). UA-SWE product was accessed through Dr. Patrick Broxton of U. of Arizona, also available from the NSIDC: https://nsidc.org/data/nsidc-0719/versions/1. Financial support was also made available from NASA GRACE and GRACE-FO (NNH15ZDA001N–GRACE) and NASA MEaSUREs (NNH17ZDA001N-MEASURES) awards.

REFERENCES

Adler, R., Sapiano, M., Huffman, G., Wang, J.-J., Gu, G., Bolvin, D., Chiu, L., Schneider, U., Becker, A., Nelkin, E., Xie, P., Ferraro, R., Shin, D.-B.: The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. Atmosphere **9**(4), 138 (2018)

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., Nelkin, E.: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). Journal of Hydrometeorology **4**(6), 1147-1167 (2003).

Bartsotas, N.S., Anagnostou, E.N., Nikolopoulos, E.I., Kallos, G.: Investigating Satellite Precipitation Uncertainty Over Complex Terrain. Journal of Geophysical Research: Atmospheres **123**(10), 5346-5359 (2018). doi:doi:10.1029/2017JD027559

Barrett, A. P.: National Operational Hydrologic Remote Sensing Center Snow Data Assimilation System (SNODAS) products at NSIDC. NSIDC Special Rep. 11, 19 pp (2003). [Available online at http://nsidc.org/pubs/documents/special/nsidc_special_ report_11.pdf.]

Behrangi, A., Gardner, A., Reager, J.T., Fisher, J.B., Yang, D., Huffman, G.J., Adler, R.F.: Using GRACE to Estimate Snowfall Accumulation and Assess Gauge Undercatch Corrections in High Latitudes. Journal of Climate **31**(21), 8689-8704 (2018a). doi:10.1175/jcli-d-18-0163.1

Behrangi, A., Bormann, K.J., Painter, T. H.: Using the Airborne Snow Observatory to Assess Remotely Sensed Snowfall Products in the California Sierra Nevada. Water Resources Research **54**(10), 7331-7346 (2018b). doi:10.1029/2018wr023108.

Behrangi, A., Gardner, A.S., Reager, J.T., Fisher, J.B.: Using GRACE to constrain precipitation amount over cold mountainous basins. Geophysical Research Letters **44**(1), 219-227 (2017). doi:10.1002/2016gl071832.

Behrangi, A., Guan, B., Neiman, P.J., Schreier, M., Lambrigtsen, B.: On the Quantification of Atmospheric Rivers Precipitation from Space: Composite Assessments and Case Studies over the Eastern North Pacific Ocean and the Western United States. Journal of Hydrometeorology **17**(1), 369-382 (2016a). doi:10.1175/jhm-d-15-0061.1.

Behrangi, A., Christensen, M., Richardson, M., Lebsock, M., Stephens, G., Huffman, G.J., Bolvin, D., Adler, R.F., Gardner, A., Lambrigtsen, B., Fetzer, E.: Status of high-latitude precipitation estimates from observations and reanalyses. Journal of Geophysical Research: Atmospheres **121**(9), 4468-4486 (2016b). doi:10.1002/2015jd024546.

543 Behrangi, A., Tian, Y., Lambrigtsen, B.H., Stephens, G. L.: What does CloudSat reveal about
544 global land precipitation detection by other spaceborne sensors? Water Resources Research
545 **50**(6), 4893-4905 (2014a). doi:10.1002/2013wr014566.
546 Behrangi, A., Andreadis, K., Fisher, J.B., Turk, F.J., Granger, S., Painter, T., Das, N.: Satellite-
547 Based Precipitation Estimation and Its Application for Streamflow Prediction over
548 Mountainous Western U.S. Basins. Journal of Applied Meteorology and Climatology **53**(12),
549 2823-2842 (2014b). doi:10.1175/jamc-d-14-0056.1.
550 Behrangi, A., Wong, S., Mallick, K., Fisher, J.B.: On the net surface water exchange rate
551 estimated from remote-sensing observation and reanalysis. International Journal of Remote
552 Sensing **35**(6), 2170-2185 (2014c). doi:10.1080/01431161.2014.889866.
553 Behrangi, A., Khakbaz, B., Jaw, T.C., AghaKouchak, A., Hsu, K., Sorooshian, S.: Hydrologic
554 evaluation of satellite precipitation products over a mid-size basin. Journal of Hydrology
555 **397**(3-4), 225-237 (2011). doi:10.1016/j.jhydrol.2010.11.043.
556 Broxton, P.D., Dawson, N., Zeng, X.: Linking snowfall and snow accumulation to generate
557 spatial maps of SWE and snow depth. Earth and Space Science **3**(6), 246-256 (2016b).
558 doi:10.1002/2016EA000174.
559 Broxton, P.D., Zeng, X., Dawson, N.: Why Do Global Reanalyses and Land Data Assimilation
560 Products Underestimate Snow Water Equivalent? Journal of Hydrometeorology **17**(11), 2743-
561 2761 (2016a). doi:10.1175/jhm-d-16-0056.
562 Berg, W., L'Ecuyer, T., Kummerow, C.: Rainfall Climate Regimes: The Relationship of
563 Regional TRMM Rainfall Biases to the Environment. Journal of Applied Meteorology and
564 Climatology **45**(3), 434-454 (2006). doi:10.1175/jam2331.1.

565 Chen, S., Hong, Y., Kulie, M., Behrangi, A., Stepanian, P.M., Cao, Q., You, Y., Zhang, J., Hu,
566 J., Zhang, X.: Comparison of snowfall estimates from the NASA CloudSat Cloud Profiling
567 Radar and NOAA/NSSL Multi-Radar Multi-Sensor System. Journal of Hydrology **541**, 862-
568 872 (2016). doi:10.1016/j.jhydrol.2016.07.047.
569 Christensen, M.W., Behrangi, A., L'ecuyer, T.S., Wood, N.B., Lebsock, M.D., Stephens, G.L.:
570 Arctic Observation and Reanalysis Integrated System: A New Data Product for Validation
571 and Climate Study. Bulletin of the American Meteorological Society **97**(6), 907-916 (2016).
572 doi:10.1175/bams-d-14-00273.1
573 Copernicus Climate Change Service (C3S): ERA5: Fifth generation of ECMWF atmospheric
574 reanalyses of the global climate . Copernicus Climate Change Service Climate Data Store
575 (CDS), date of access March 2019 (2017).
576 Daly, C., Neilson, R.P., Phillips, D.L.: A Statistical-Topographic Model for Mapping
577 Climatological Precipitation over Mountainous Terrain. Journal of Applied Meteorology
578 **33**(2), 140-158 (1994).
579 Daly, C.; Gibson, W.P.; Taylor, G.H.; Johnson, G.L.; Pasteris, P. A knowledge-based approach
580 to the statistical mapping of climate. Clim. Res., 22, 99–113 (2002).
581 Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J.,
582 Pasteris, P.P.: Physiographically sensitive mapping of climatological temperature and
583 precipitation across the conterminous United States. International Journal of Climatology
584 **28**(15), 2031-2064 (2008). doi:doi:10.1002/joc.1688
585 Dawson, N., Broxton, P., Zeng, X.: Evaluation of Remotely Sensed Snow Water Equivalent and
586 Snow Cover Extent over the Contiguous United States. Journal of Hydrometeorology **19**(11),
587 1777-1791 (2018). doi:10.1175/jhm-d-18-0007.1

588 Dee D. P., Uppala S. M., Simmons A. J., Berrisford P., Poli P., Kobayashi S., et al. : The ERA-
589 Interim reanalysis: configuration and performance of the data assimilation system. Quarterly
590 Journal of the Royal Meteorological Society **137**(656), 553-597 (2011).
591 doi:doi:10.1002/qj.828
592 Dingman, S. L., Physical Hydrology. 2nd ed. Long Grove, IL: Waveland Press (2008).
593 Ferraro, R.R., Peters-Lidard, C.D., Hernandez, C., Turk, F.J., et al. : An Evaluation of
594 Microwave Land Surface Emissivities Over the Continental United States to Benefit GPM-
595 Era Precipitation Algorithms. IEEE Transactions on Geoscience and Remote Sensing **51**(1),
596 378-398 (2013). doi:10.1109/tgrs.2012.2199121
597 Fuchs, T., Rapp, J., Rubel, F., Rudolf, B.: Correction of synoptic precipitation observations due
598 to systematic measuring errors with special regard to precipitation phases. . Phys. Chem.
599 Earth, **B29**, 689–693 (2001)
600 Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, et al..: The Modern-Era
601 Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Journal of
602 Climate **30**(14), 5419-5454 (2017). doi:10.1175/JCLI-D-16-0758.1
603 Huffman, G. J., D.T. Bolvin, D. Braithwaite, K. Hsu et al., Integrated Multi-satellitE Retrievals
604 for GPM (IMERG) Technical Documentation (V5.2), pp 35 (2018).
605 Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman,
606 K.P., Stocker, E.F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global,
607 Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of
608 Hydrometeorology **8**(1), 38-55 (2007). doi:10.1175/jhm560.1
609 Huffman, G.J., Adler, R.F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab,
610 A., Rudolf, B., Schneider, U.: The Global Precipitation Climatology Project (GPCP)

611 Combined Precipitation Dataset. Bulletin of the American Meteorological Society **78**(1), 5-20
612 (1997)
613 Huffman, G.J., Adler, R.F., Morrissey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B.,
614 Susskind, J.: Global Precipitation at One-Degree Daily Resolution from Multisatellite
615 Observations. Journal of Hydrometeorology **2**(1), 36-50 (2001)
616 Joyce, R.J., Xie, P.: Kalman Filter–Based CMORPH. Journal of Hydrometeorology **12**(6), 1547-
617 1563 (2011). doi:10.1175/jhm-d-11-022.1
618 Kelly, R.: The AMSR-E Snow Depth Algorithm: Description and Initial Results. Journal of The
619 Remote Sensing Society of Japan **29**(1), 307-317 (2009). doi:10.11440/rssj.29.307
620 Lebsock Matthew D., L'Ecuyer Tristan S.: The retrieval of warm rain from CloudSat. Journal of
621 Geophysical Research: Atmospheres **116**(D20) (2011). doi:doi:10.1029/2011JD016076
622 Legates, D., Willmott, C.J.: Mean Seasonal and Spatial Variability in Gauge-Corrected, Global
623 Precipitation. Internat. J. Climatol. **10**, 111-127 (1990)
624 Maggioni, V., Nikolopoulos, E.I., Anagnostou, E.N., Borga, M.: Modeling Satellite Precipitation
625 Errors Over Mountainous Terrain: The Influence of Gauge Density, Seasonality, and
626 Temporal Resolution. IEEE Transactions on Geoscience and Remote Sensing **55**(7), 4130-
627 4140 (2017). doi:10.1109/TGRS.2017.2688998
628 Painter, T.H., Berisford, D.F., Boardman, J.W., Bormann, K.J., et al.: The Airborne Snow
629 Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling
630 for mapping snow water equivalent and snow albedo. Remote Sensing of Environment **184**,
631 139-152 (2016).

632 Palerme, C., Kay, J.E., Genthon, C., L'Ecuyer, T., Wood, N.B., Claud, C.: How much snow falls
633 on the Antarctic ice sheet? The Cryosphere Discussions **8**(1), 1279-1304 (2014).
634 doi:10.5194/tcd-8-1279-2014
635 Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A.P., Black, J.,
636 Thériault, J.M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., Gutmann, E.:
637 How Well Are We Measuring Snow: The NOAA/FAA/NCAR Winter Precipitation Test Bed.
638 Bulletin of the American Meteorological Society **93**(6), 811-829 (2012). doi:10.1175/bams-d-
639 11-00052.1
640 Schneider, Udo; Becker, Andreas; Finger, Peter; Meyer-Christoffer, Anja; Ziese, Markus: GPCC
641 Monitoring Product: Near Real-Time Monthly Land-Surface Precipitation from Rain-Gauges
642 based on SYNOP and CLIMAT data (2018). DOI: 10.5676/DWD_GPCC/MP_M_V6_100;
643 http://dx.doi.org/10.5676/DWD_GPCC/MP_M_V6_100
644 Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M., Becker, A.:
645 Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation
646 Climatology from the Global Precipitation Climatology Centre (GPCC). Atmosphere **8**(12)
647 (2017). doi:10.3390/atmos8030052.
648 Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., Rudolf, B.: GPCC's new
649 land surface precipitation climatology based on quality-controlled in situ data and its role in
650 quantifying the global water cycle. Theoretical and Applied Climatology **115**(1-2), 15-40
651 (2014).
652 Serreze, M.C., Clark, M.P., Armstrong, R.L., McGinnis, D.A., Pulwarty, R.S.: Characteristics of
653 the western United States snowpack from snowpack telemetry (SNOTEL) data. Water
654 Resources Research **35**(7), 2145-2160 (1999). doi:10.1029/1999WR900090

655 Simmons, A.J., Willett, K.M., Jones, P.D., Thorne, P.W., Dee, D.P.: Low-frequency variations in
656 surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and
657 monthly gridded observational data sets. Journal of Geophysical Research **115**(D1) (2010).
658 doi:10.1029/2009jd012442
659 Tian, Y., Peters-Lidard, C.D., Harrison, K.W., Prigent, C., Norouzi, H., Aires, F., Boukabara, S.-
660 A., Furuzawa, F.A., Masunaga, H.: Quantifying Uncertainties in Land-Surface Microwave
661 Emissivity Retrievals. IEEE Transactions on Geoscience and Remote Sensing **52**(2), 829-840
662 (2014). doi:10.1109/tgrs.2013.2244214
663 Reichle, R.H., Draper, C.S., Liu, Q., Girotto, M., Mahanama, S.P.P., Koster, R.D., De Lannoy,
664 G.J.M.: Assessment of MERRA-2 Land Surface Hydrology Estimates. Journal of Climate
665 **30**(8), 2937-2960 (2017). doi:10.1175/JCLI-D-16-0720.1
666 Rubel, F., Hantel, M.: BALTEX 1/6-degree daily precipitation climatology 1996–1998.
667 Meteorology and Atmospheric Physics **77**(1), 155-166 (2001). doi:10.1007/s007030170024
668 Shige, S., Kida, S., Ashiwake, H., Kubota, T., Aonashi, K.: Improvement of TMI Rain Retrievals
669 in Mountainous Areas. Journal of Applied Meteorology and Climatology **52**(1), 242-254
670 (2013). doi:10.1175/jamc-d-12-074.1
671 Smalley, M., L'Ecuyer, T., Lebsock, M., Haynes, J.: A Comparison of Precipitation Occurrence
672 from the NCEP Stage IV QPE Product and theCloudSatCloud Profiling Radar. Journal of
673 Hydrometeorology **15**(1), 444-458 (2014). doi:10.1175/jhm-d-13-048.1
674 Tedesco, M., Narvekar, P.S.: Assessment of the NASA AMSR-E SWE Product. IEEE Journal of
675 Selected Topics in Applied Earth Observations and Remote Sensing **3**(1), 141-159 (2010).

676 Yang, D., Kane, D., Zhang, Z., Legates, D., Goodison, B.: Bias corrections of long-term (1973-
677 2004) daily precipitation data over the northern regions. Geophysical Research Letters **32**(19),
678 (2005). doi:10.1029/2005gl024057.
679 Watkins, M.M., Wiese, D.N., Yuan, D.-N., Boening, C., Landerer, F.W.: Improved methods for
680 observing Earth's time variable mass distribution with GRACE using spherical cap mascons.
681 Journal of Geophysical Research: Solid Earth **120**(4), 2648-2671 (2015).
682 doi:10.1002/2014jb011547.
683 Wiese, D.N., Landerer, F.W., Watkins, M.M.: Quantifying and reducing leakage errors in the
684 JPL RL05M GRACE mascon solution. Water Resources Research **52**(9), 7490-7502 (2016).
685 doi:10.1002/2016wr019344.
686

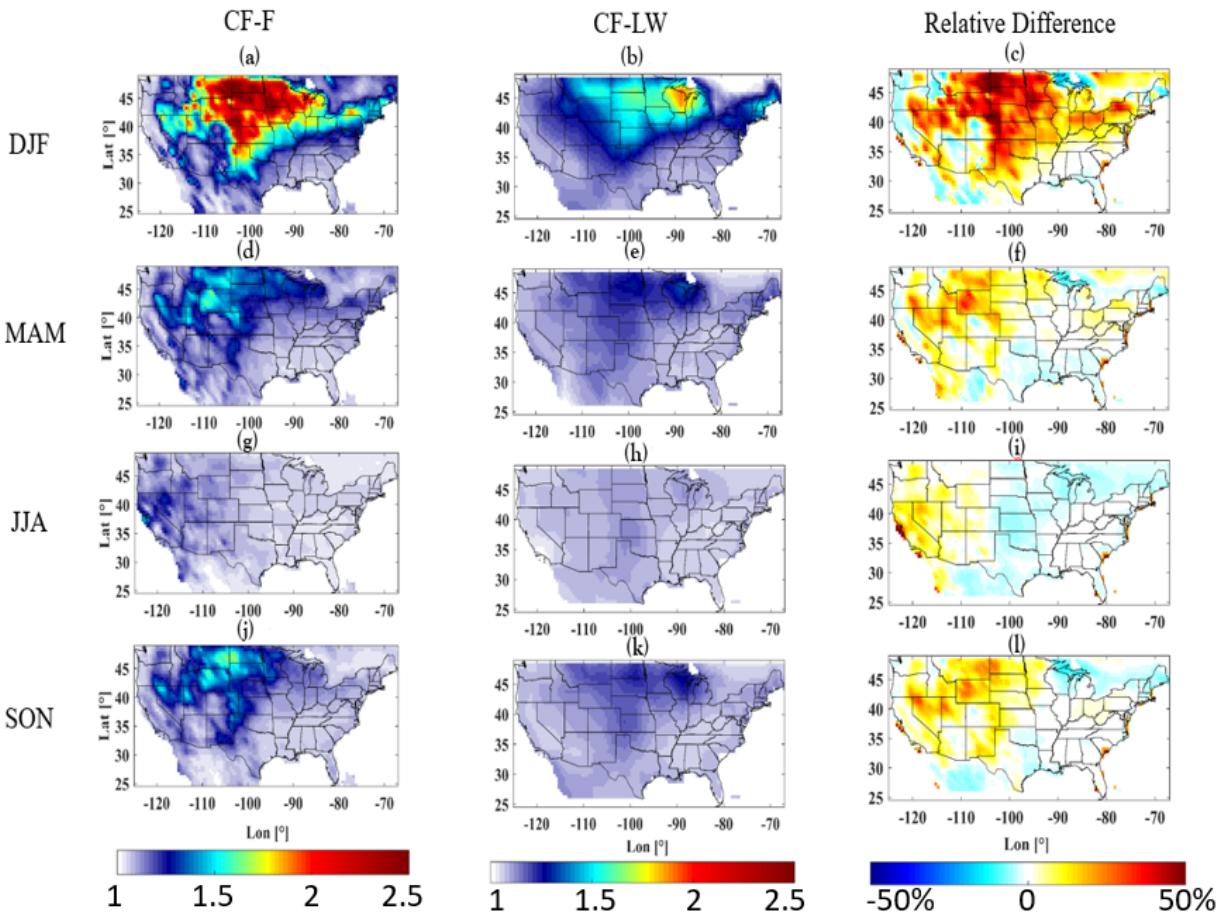


Figure 1. Seasonal maps of CF-F (left column) and CF-LW (middle column), and relative differences between the two (right-column) calculated by averaging over the studied years (2003-2015) of monthly CFs over CONUS. From top to bottom rows represent winter, spring, summer, and fall, respectively.

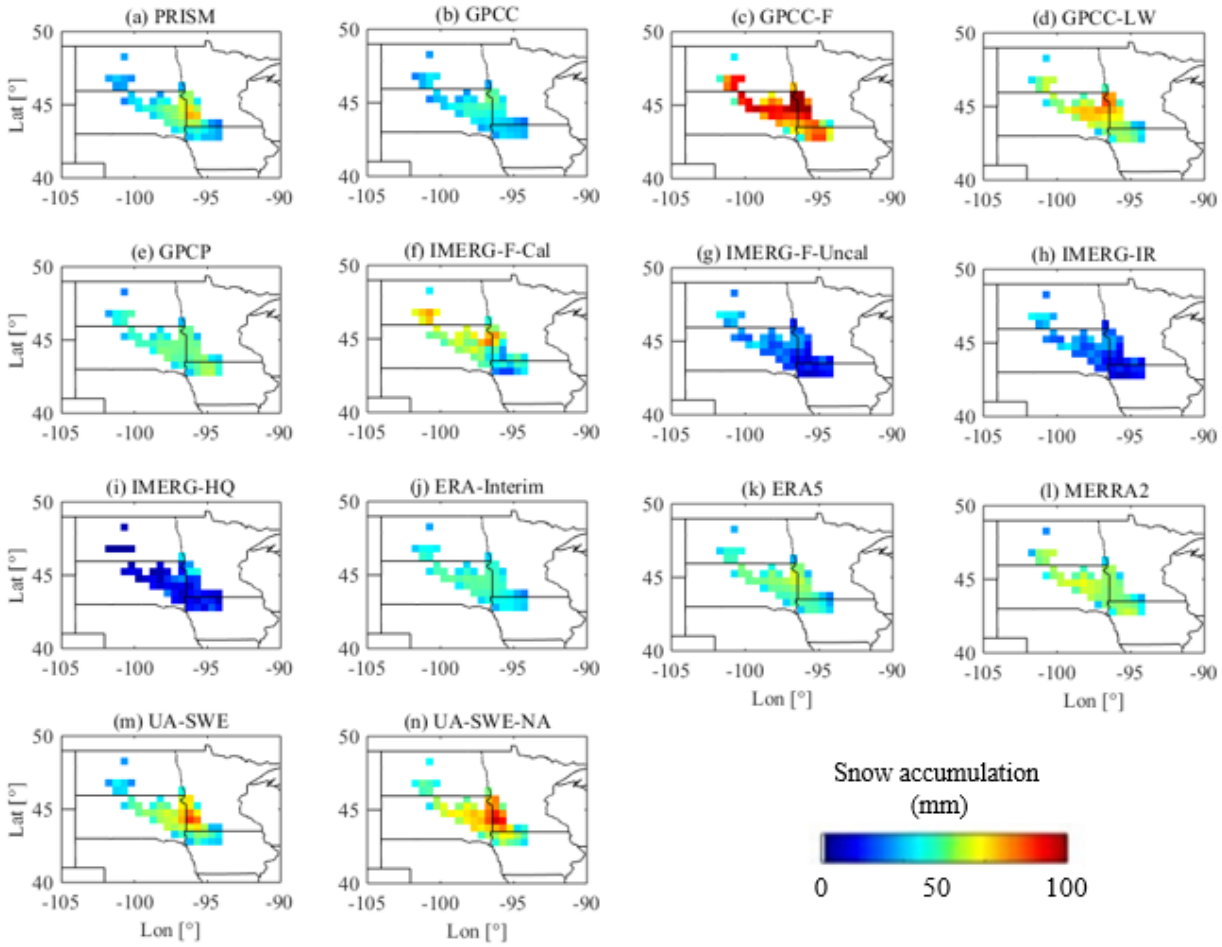


Figure 2. Maps of snow accumulation over a snow accumulation case study (from Dec 30 2010 to Feb 5 2011), corresponding to the cumulative daily snow plot shown in Figure 5.

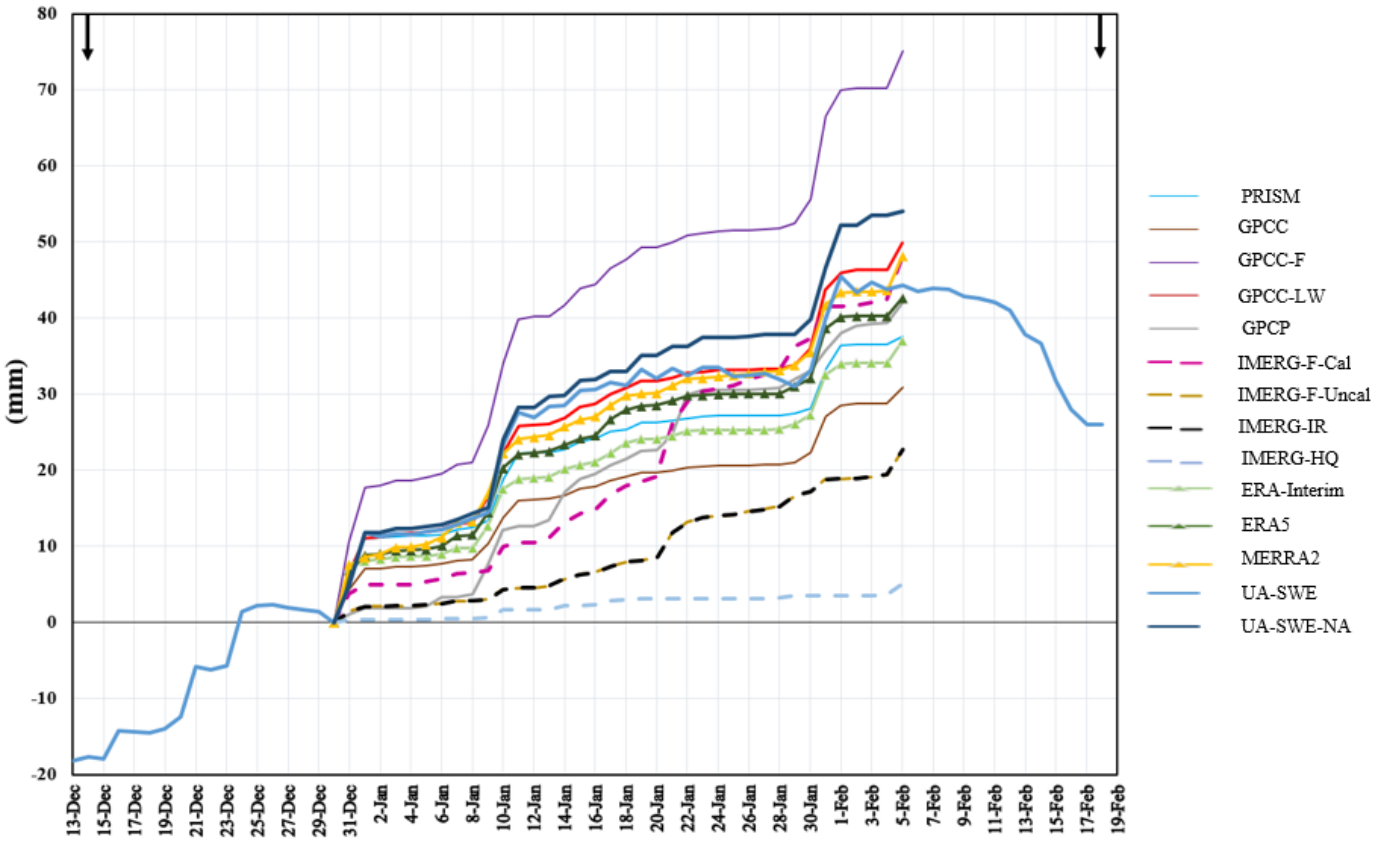


Figure 3. Cumulative daily snowfall plot using several products for a snow accumulation period (Dec 30. 2010 to Feb 5 2011) over a region west of the Great Lakes shown in Figure 2. GRACE reported an observation on Dec 14 2010 and another one on Feb 18 2011, shown with black arrows. The change in TWS between the two dates was only 5% more than the corresponding change in SWE, suggesting a great consistency between the observed changes in TWS and SWE.

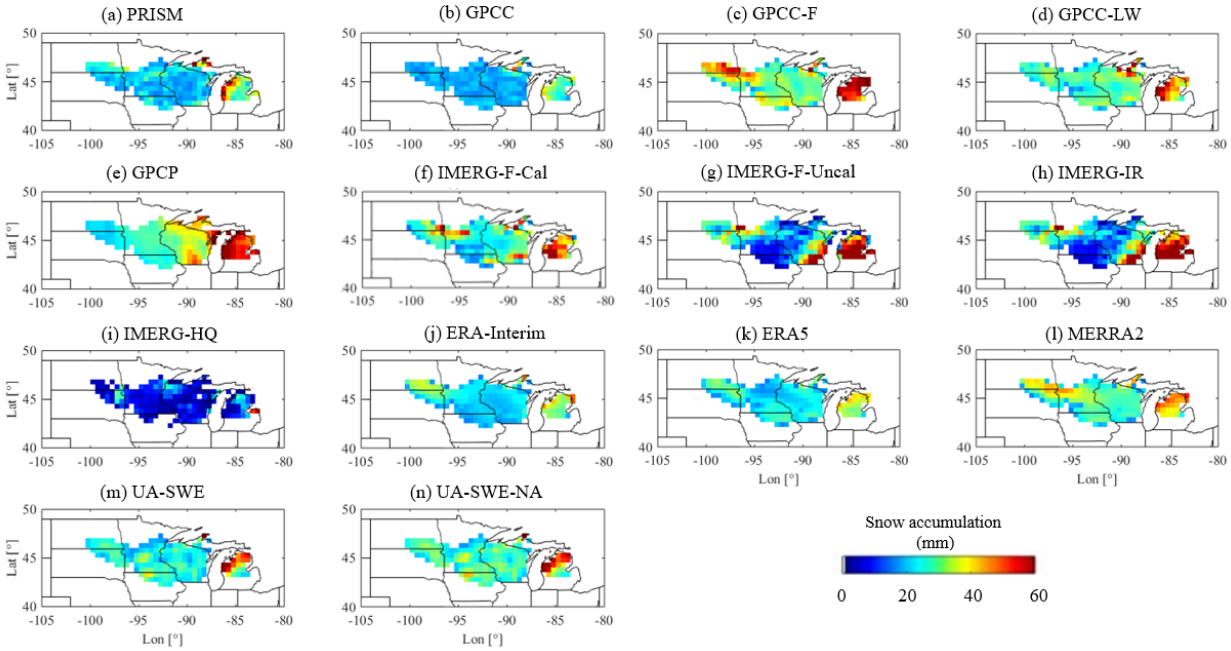
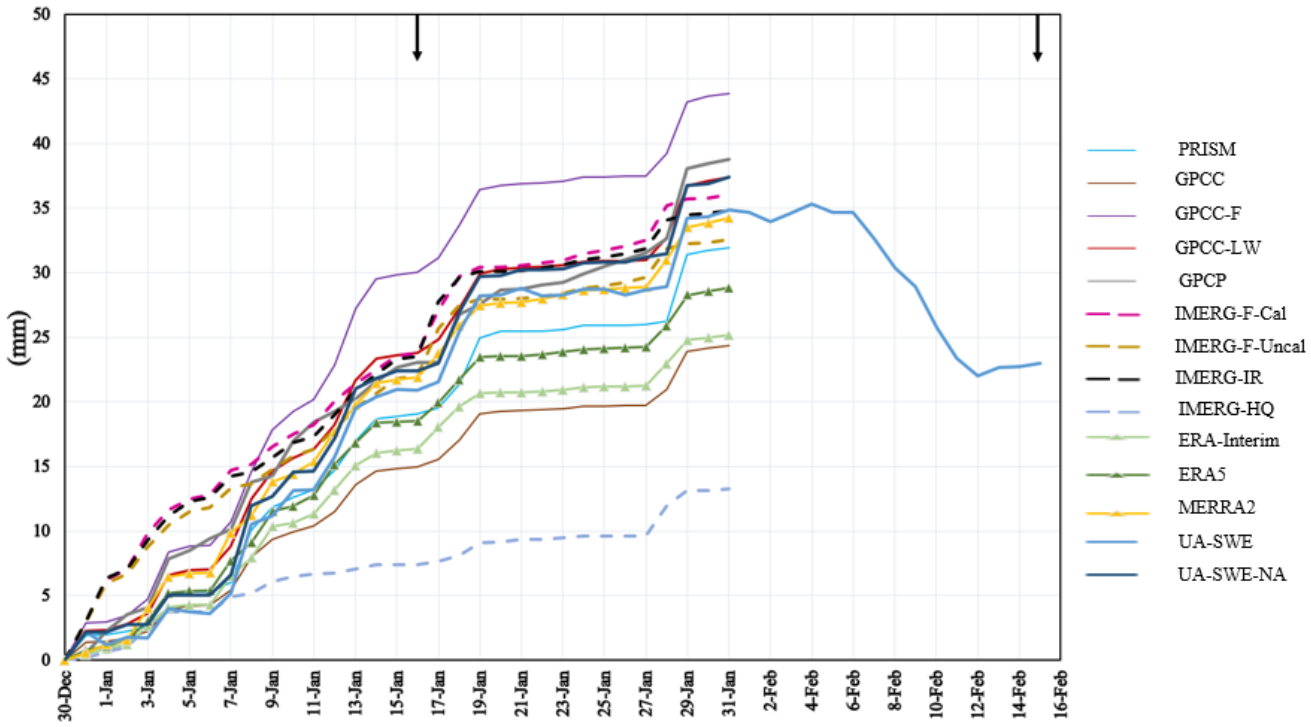


Figure 4. Maps of snow accumulation over a snow accumulation case study (from Dec 30. 2008 to Jan 31 2009), corresponding to the cumulative daily snow plot shown in Figure 3.

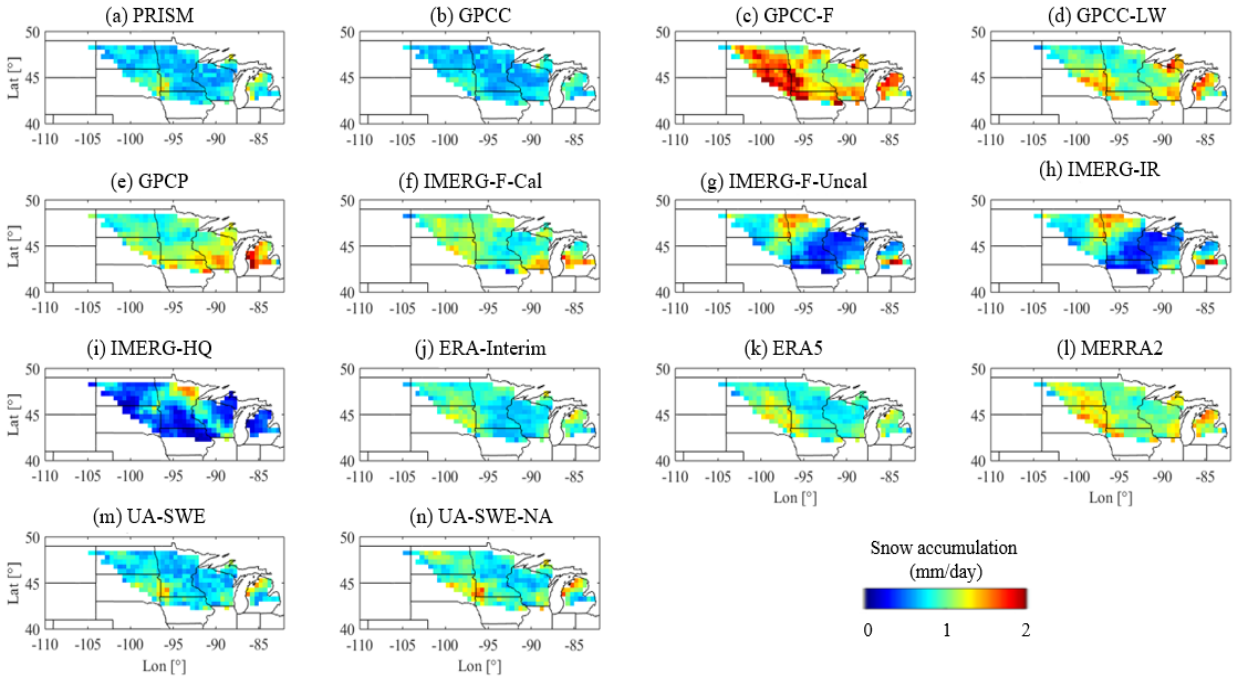


733

734 Figure 5. Cumulative daily snowfall plot using several products for a snow accumulation period
735 (Dec 30. 2008 to Jan 31 2009) over a region west of the Great Lakes shown in Figure 2. GRACE
736 reported an observation on Jan 17 2009 and another one on Feb 16th. The change in TWS
737 between the two dates was only slightly (0.27 mm) more than the corresponding change in SWE,
738 suggesting a great consistency between the observed changes in TWS and SWE.

739

740



741

742 Figure 6: Average snow accumulation across all snow accumulation cases collected from all
743 studied years (2003-2015) over a region west of the Great Lakes.

744

745

746

747

748

749

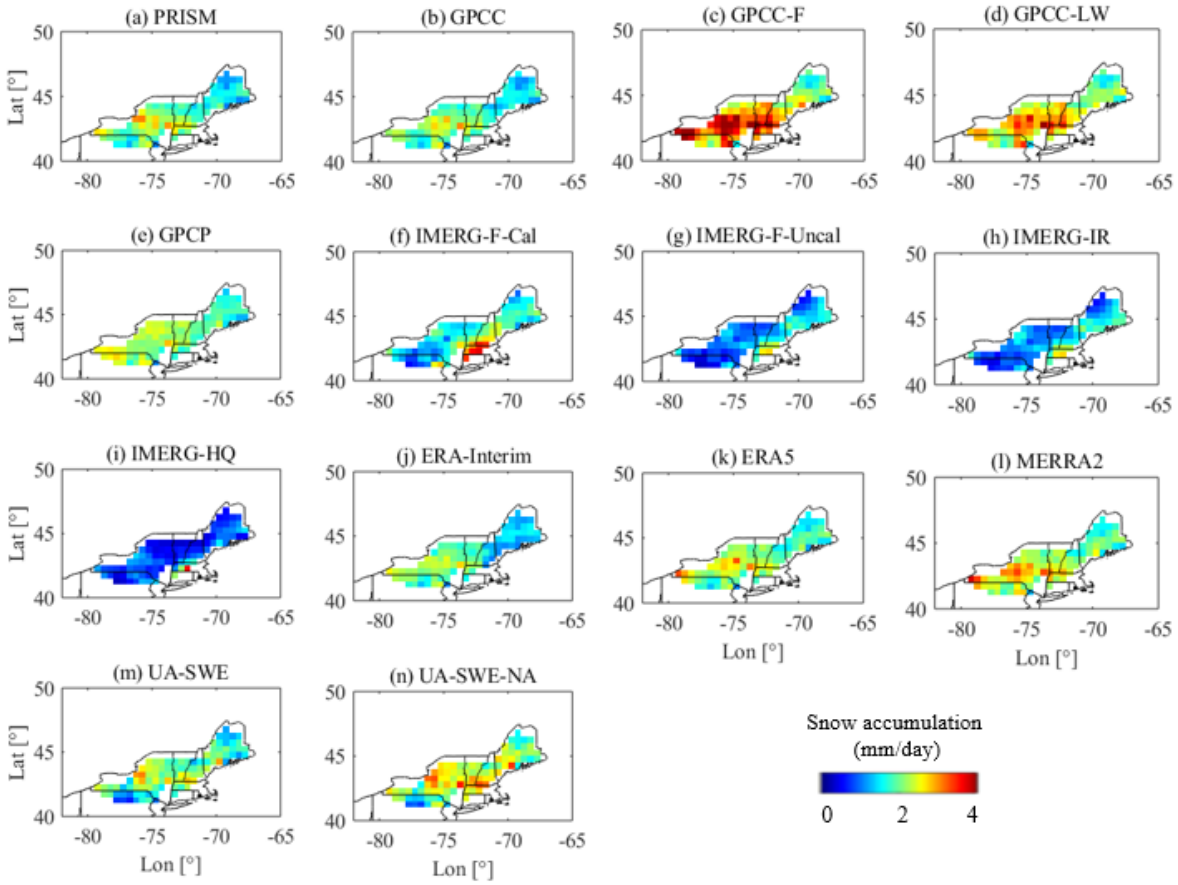


Figure 7: Similar to Figure 6, but over the Northeast United States.

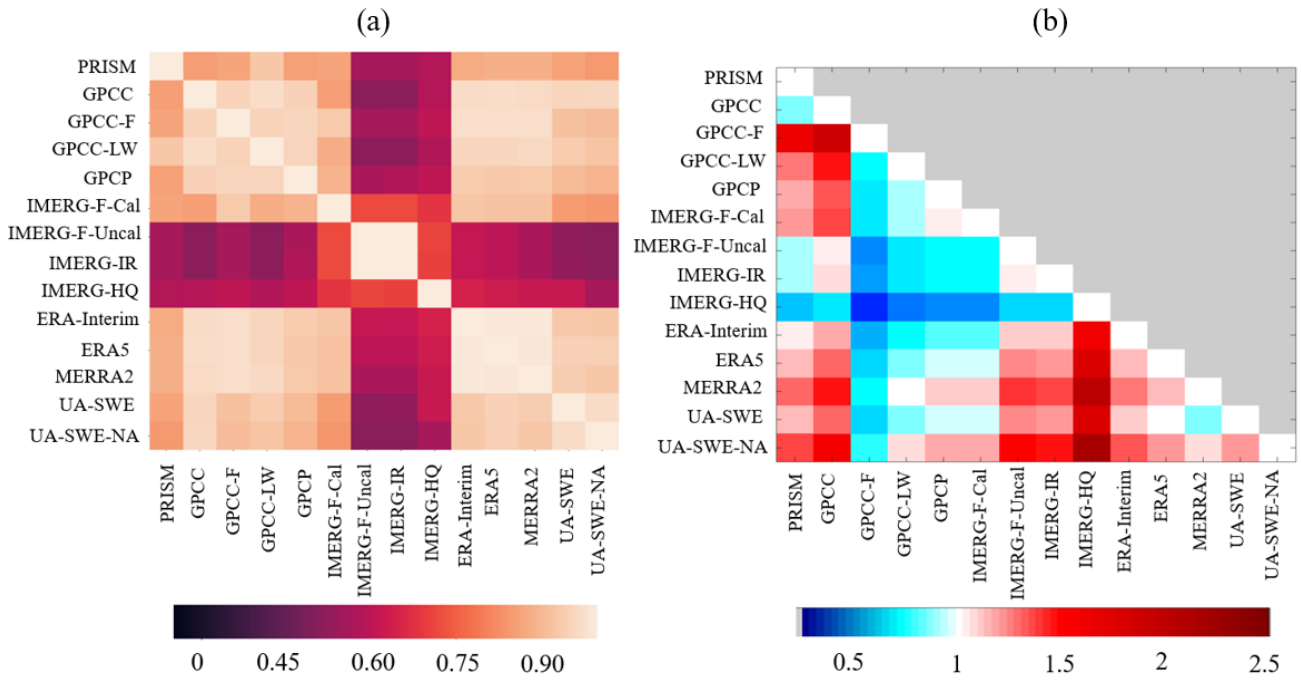


Figure 8, The overall cross-correlation and bias ratio among the products: (a) correlation coefficient, (b) bias ratio. Bias ratios in Figure 8b is calculated by dividing the total snow accumulation from the products listed at Y-axis by the total snow accumulation from the products listed at X-axis.

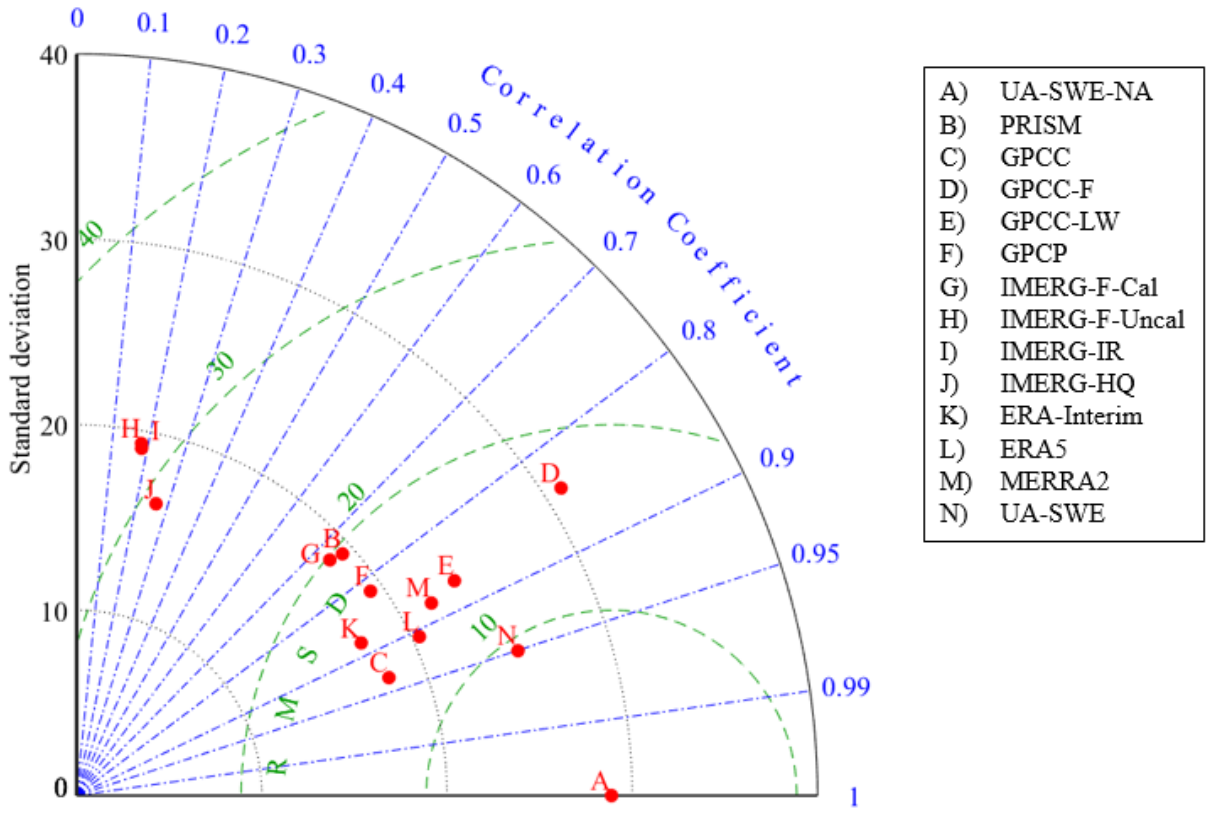


Figure 9: Taylor diagram summarizing the skill of the studied products relative to UA-SWE-NA.