1.1 INTEGRATED MULTI-SATELLITE EVALUATION FOR THE GLOBAL PRECIPITATION MEASUREMENT: IMPACT OF PRECIPITATION TYPES ON SPACEBORNE PRECIPITATION ESTIMATION

Pierre-Emmanuel Kirstetter[1], Walter A. Petersen[2] , David Wolff[3], and Jonathan J. Gourley[4]

Abstract Integrated multi-sensor assessment is proposed as a novel approach to advance satellite precipitation validation in order to provide users and algorithm developers with an assessment adequately coping with the varying performances of merged satellite precipitation estimates. Gridded precipitation rates retrieved from space sensors with quasi-global coverage feed numerous applications ranging from water budget studies to forecasting natural hazards cuased by extreme events. Characterizing the error structure of satellite precipitation products is recognized as a major issue for the usefulness of these estimates. The Global Precipitation Measurement (GPM) mission aims at unifying precipitation measurements from a constellation of low-earth orbiting (LEO) sensors with various capabilities to detect, classify and quantify precipitation. They are used in combination with geostationary observations to provide gridded precipitation accumulations. The GPM Core Observatory satellite serves as a calibration reference for consistent precipitation retrieval algorithms across the constellation. The propagation of QPE uncertainty from LEO active/passive microwave (PMW) precipitation estimates to gridded QPE is addressed in this study, by focusing on the impact of precipitation typology on QPE from the Level-2 GPM Core Observatory Dual-frequency Precipitation Radar (DPR) to the Microwave Imager (GMI) to Level-3 IMERG precipitation over the Conterminous U.S. A high-resolution surface precipitation used as a consistent reference across scales is derived from the ground radar-based Multi-Radar/Multi-Sensor. While the error structure of the DPR, GMI and subsequent IMERG is complex because of the interaction of various error factors, systematic biases related to precipitation typology are consistently quantified across products. These biases display similar features across Level-2 and Level-3, highligthing the need to better resolve precipitation typology from space and the room for improvement in global-scale precipitation estimates. The integrated analysis and framework proposed herein applies more generally to precipitation estimates from sensors and error sources affecting low-earth orbiting satellites and derived gridded products.

[1] Pierre-Emmanuel Kirstetter,
Advanced Radar Research Center and School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, Oklahoma, and NOAA/National Severe Storms Laboratory,
Norman, Oklahoma, U.S.
e-mail: pierre.kirstetter@noaa.gov
[2] Walter A. Petersen
NASA Marshall Space Flight Center,
Wallops, Virginia, U.S.
e-mail: david.b.wolff@nasa.gov
[3] David Wolff
NASA Goddard Space Flight Center,
Huntsville, Alabama, U.S.
e-mail: walt.petersen@nasa.gov
[4] Jonathan J. Gourley
NOAA/National Severe Storms Laboratory,
Norman, Oklahoma, U.S.
e-mail: jj.gourley@noaa.gov

1.1.1 Introduction

Precipitation is key to the Earth hydrologic and energy fluxes through its occurrence, type, and quantity. A host of scientific questions impacting society are concerned with the distribution of precipitation characteristics, from extreme events such as droughts and hurricanes to how the availability of fresh water evolves under a non-stationary climate. Precipitation physical processes take place over a range of spatial and temporal scales and drive its highly variable intermittency, intensity, areal extent, and duration. This variability poses challenges for observations specifically from remote sensing for hydrologic, meteorological, and climatic applications. Their quasi-global coverage promotes the use of satellite-based quantitative precipitation estimates (QPE) for such purposes. However, converting satellite measurements into QPE poses challenges, as it depends on the spatial heterogeneity of the precipitation fields, the indirect nature of the measurement, the sensor resolution and sensitivity, and the retrieval algorithm. Hence satellite-based precipitation estimates suffer from poorly characterized and quantified sources of uncertainty, which currently limit their assimilation into hydrologic and atmospheric models (Bauer et al., 2011; Stephens and Kummerow 2007; Weng et al., 2007). To improve the satellite estimates and maximize their usability, their uncertainty must be evaluated in terms of precipitation characteristics including intermittency, distribution of types (e.g. stratiform, convective, snow, hail) and rates, as highlighted by the International Precipitation Working Group (IPWG; see http://www.isac.cnr.it/~ipwg/) (Turk *et al.* 2008; Yang *et al.* 2006; Zeweldi and Gebremichael 2009; Sapiano and Arkin 2009; Wolff and Fisher 2009; Grimes and Diop 2003; Lebel *et al.* 2009). Comprehensive characterization of the satellite precipitation error structure relies on ground-validation research to ensure proper accuracy of spaceborne QPE missions (Petersen and Schwaller 2008) like the former Tropical Rainfall Measurement Mission (TRMM; http://trmm.gsfc.nasa.gov) and the current Global Precipitation Measurement Mission (GPM; http://gpm.gsfc.nasa.gov).

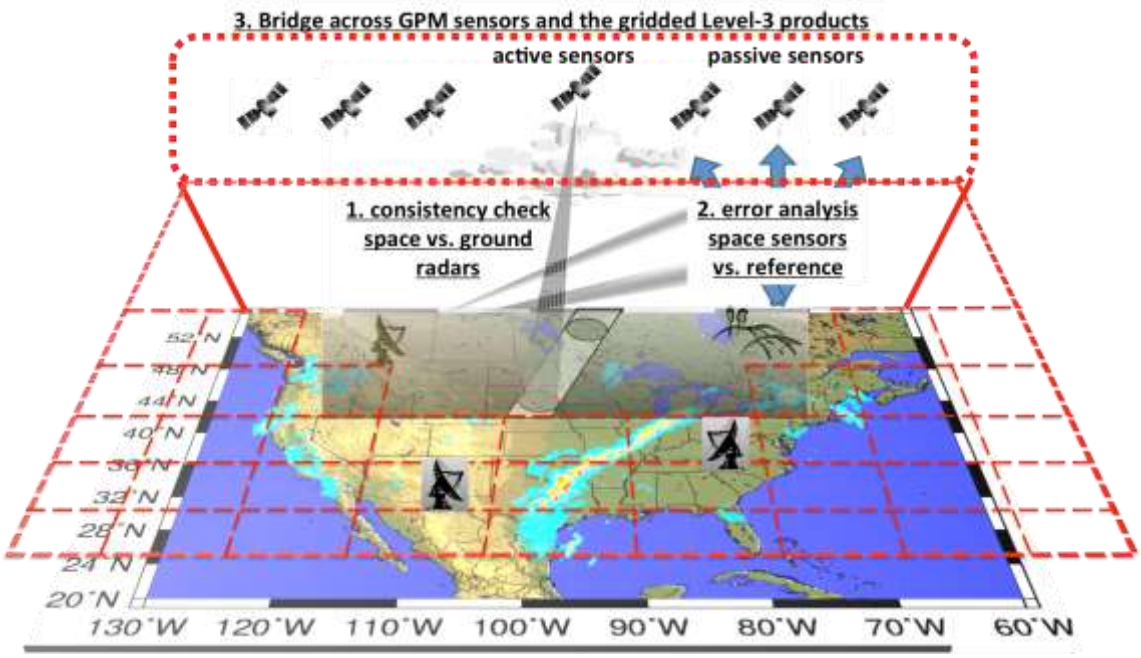

Figure 1.1.1. Research framework and overview flowchart to bridge from the Global Precipitation Measurement mission core satellite to the combined gridded IMERG product using ground radar-

based Multi-Radar/Multi-Sensor QPE. An example of Multi-Radar/Multi-Sensor instantaneous precipitation rates at 0725 UTC on 11 April 2011 is shown.

Current high-resolution and gridded satellite-based QPE have been developed by combining active and passive microwave (PMW) and infrared (IR) sensors into multi-sensor precipitation retrievals. PMW-based estimates from Low Earth Orbiting (LEO) satellites provide higher accuracy but limited spatial and temporal resolution. They are used to calibrate IR-based estimates from geosynchronous Earth-orbiting (GEO) satellite platforms with comparatively higher spatio-temporal resolution and lower latency (Joyce *et al.*, 2004; Huffman *et al.*, 2007; Ushio *et al.* 2006, Ebert 2007). As of today, our understanding on how uncertainties originate from LEO platforms and propagate through such combined products is critically lacking, and there is a need for comparison across product levels (Fig. 1.1.1). While single satellite precipitation products are typically evaluated independently in the literature, a consistent end-to-end error quantification tracking the uncertainty from Level-2 through Level-3 products is needed to fully understand the errors across scales and assess the room for improvement in global precipitation estimation.
During the last couple of decades, individual satellite-based precipitation products have been evaluated over various regions (Carr et al., 2015; Derin et al., 2016; Derin & Yilmaz, 2014; Dinku et al., 2010; Golian et al., 2015; Grecu & Anagnostou, 2001), precipitation regimes (Kirstetter et al., 2012; Kirstetter et al., 2013a,b, 2014, 2015; Smalley et al., 2017), seasons (Tang et al., 2014), scales (Tan et al., 2017), precipitation intensities (Kirstetter et al., 2012; Panegrossi, 2016; Wolff & Fisher, 2009), and for applications in flood prediction (Hossain & Anagnostou, 2004; Vergara et al., 2014) because their effective use necessitates the characterization of their uncertainties at fine spatio-temporal scales (e.g. Anagnostou, 2004; Huffman et al., 2007; Joyce et al., 2004; Munchak & Skofronick-Jackson, 2013; Tian et al., 2007; Sorooshian et al., 2011; Kirstetter et al. 2018). However, the extensive body of literature on satellite precipitation validation generally provides limited insight in their error characteristics for several reasons. Common assessment practices typically use a surface-based precipitation reference and bulk comparison metrics (e.g. correlation, bias) to assess performances over a given spatial and temporal domain. First, an objective assessment requires an independent precipitation reference often derived from ground-based sensors, usually gauges. Independence may not be verified when gauges already incorporated in the satellite product are re-used to assess its accuracy. Second, metrics such as correlation, bias, contingency statistics are often applied without necessarily checking the relevance of such criteria. For example, the definition of bias may be ambiguous as it can be defined as an additive satellite QPE-to-reference difference or a multiplicative difference (ratio), sometimes based on conditional (positive) precipitation values. The linear correlation is generally insufficient to describe the non-linear and heteroscedastic dependence structure between the satellite QPE and the reference. Third, the satellite product is often assumed to be consistent and display homogeneous performances over the spatial and temporal domain of comparison. Bulk metrics (correlation, bias, contingency, etc.) are computed over samples actually gathering a variety of precipitation characteristics (intermittency, typology, rates) for which the satellite algorithm (or combination of algorithms for merged products) is likely to behave differently. More generally the comparison is always performed with precipitation estimates ambiguously derived from the satellite sensor observation through the retrieval algorithm and associated assumptions. Individual PMW/IR retrievals are underconstrained by nature and sensitive to unobserved atmospheric parameters (Stephens and Kummerow, 2007). The combined products inherit the varying PMW/IR performances and create additional uncertainties with temporal/spatial resampling. Hence bulk error metrics depict averaged space/time properties while the errors tend to be non-stationary and sensitive to parameters not accounted for in the assessment formulation. Fourth, the representativeness of any overall satellite QPE assessment or error model is confined to the time and space domain over which it is performed. It tends to be specific to the satellite instrument (e.g. resolution), the retrieval algorithm (and associated version), the space-time-scale and the accuracy of

the reference, with limited applicability for other precipitation regimes, regions, products, etc. As a result, the actual benefit of these analyses to satellite precipitation users and developers is limited.
Integrated multi-sensor assessment is necessary to track the origin of uncertainties and their propagation throught various Level-2 active, passive then Level-3 merged satellite precipitation estimates. More informative assessment and information to algorithm improvement can be gained by stratifying (conditioning) the assessment according to relevant factors driving the state of the satellite estimation error. Hence targeting the most significant factors is essential to characterize uncertainties in satellite QPE and lead to a generalization of their assessment.
Generic issues have been identified for both LEO active and passive and subsequent gridded satellite-based precipitation retrieval algorithms and motivate ongoing and future research such as variability of the precipitation inside the resolution volumes (Non-Uniform Beam Filling effects - NUBF), precipitation typology and phase. In this chapter we focus on the impact of precipitation typology. While sensors measurements from ground-based radars, spaceborne radar and PWM and IR sensors differ in terms of frequency, polarization state, beam geometry, and incidence angle, they are physically consistent through the identification of hydrometeors and estimation of particle size distribution (PSD). The precipitation retrieval errors caused by the variability in the cloud vertical structure and the observed signal (e.g. reflectivity for radars, brightness temperatures for PMW and IR sensors)-to-precipitation relationship are related due to their dependence on the underlying precipitation microphysics. Precipitation type is a fundamental characteristic that drives the interpretation of the observed signal for precipitation estimation through the particle size distribution (PSD), hydrometeors properties and their evolution through precipitation microphysical processes. The satellite algorithms' capabilities to classify precipitation systems (e.g. Grams et al., 2016) therefore present great potential for the generalization of the uncertainty characteristics, specifically at coarser scales. Within the Global Precipitation Measurement (GPM) mission (https://www.nasa.gov/mission_pages/GPM/main/; Hou et al., 2014; Skofronick-Jackson et al., 2017), precipitation type is currently used only to constrain the space-borne radar precipitation retrievals. Since subsequent passive sensor retrievals and combined precipitation products do not include this fundamental characteristic, there is an interest in documenting how the absence of constraints on precipitation type impacts these precipitation products.
Targeting significant uncertainty factors for a quantitative and detailed characterization necessitate working at the primary satellite QPE scale for LEO (Level 2) through gridded precipitation products (Level-3). This task is often impaired by the difficulty of obtaining a consistent reference precipitation commensurate with the various scales of such products. To the best of our knowledge, no satellite assessment has been designed at the fine space–time resolutions of these primary satellite QPE scales. We propose an original framework to tackle these challenges. The problem is addressed by comparing the satellite QPE overall accuracy with respect to an external, independent reference precipitation product adapted to each type of space-based sensors and product. In order to match the resolution of various precipitation estaimtes, the primary resolution of the reference needs to finer than any satellite footprint or resolution. The reference has also to perform better than the space sensors regarding precipitation detection to ensure proper evaluation, and a correct reference precipitation classification (type) is required to target the physical factors contributing to erroneous satellite precipitation retrievals. A high-resolution surface rainfall product is used as a consistent reference across scales for robust comparison and evaluation over the CONUS (Figure 1.1.1). It is derived from the ground radar-based Multi-Radar/Multi-Sensor designed at the NOAA National Severe Storms Laboratory and the University of Oklahoma. Joint precipitation observations with the NEXRAD ground-based radar network and from space sensors provide unique opportunities for comparison of QPE as estimated from various sensors. Measurements from NEXRAD and GPM are physically consistent through the identification of hydrometeors and estimation of particle size distribution. We build on previous research performed on the ground-based MRMS data at unprecedented high resolution and accuracy and the high-resolution spaceborne precipitation measuring techniques to analyze the impact and propagation of uncertainty related to precipitation typology from the GPM Dual-

frequency Phased Array Radar (DPR) to passive sensors such as GPM Microwave Imager (GMI) to the gridded Integrated Multi-Satellite Retrievals for GPM (IMERG).
The aim of this study is to propose a novel approach for satellite QPE assessment by addressing the propagation of uncertainties related to precipitation type from LEO platforms through combined precipitation products. End-to-end integrated error quantification tracking the uncertainty from Level-2 through Level-3 precipitation products is particularly relevant for understanding the origin and impact of uncertainty. It addresses the critical problem that precipitation algorithm developers and users need more than just an overall assessment to adequately cope with the varying performances of precipitation products. To the best of our knowledge this approach offers, for the first time, uncertainty assessment across scales and products. It advances practices in the evaluation of remote sensing precipitation estimates by consistently analyzing precipitation estimation across levels. The focus on precipitation types is particularly relevant in the context of the current GPM mission and assesses the potential benefit of introducing precipitation type constrains to the PMW Level-2 QPE and subsequent Level-3 precipitation. Working at the primary satellite QPE scale benefits from the explicit integration of factors directly impacting the uncertainty. This framework is tested over the conterminous United States (CONUS) covered by Weather Surveillance Radar – 1988 Doppler (WSR-88D) radar data (Figure 1.1.1). This study uses a period of two years (from 2014 to 2016) of satellite and ground-based precipitation observations to obtain representative samples characterizing precipitation for various conditions of climatology, occurrence, type and rate.
The DPR, GMI, IMERG data and steps required to refine the MRMS ground-based precipitation to arrive at a consistent reference precipitation across scales used for comparisons are presented in section 2. Section 3 assesses the impact of precipitation typology on DPR, GMI and IMERG precipitation retrievals. The chapter is closed with concluding remarks in section 4.

1.1.2 Spaceborne and ground-based precipitation datasets

Following the Tropical Rainfall Measuring Mission (TRMM), the Global Precipitation Measurement mission led by NASA and JAXA aims at providing a comprehensive description of precipitation at the global scale (Hou et al. 2014). It consists of a core satellite in non-sun-synchronous orbit to serve as a physics observatory to gain insights into precipitation systems and as a calibration reference to unify precipitation estimates from a constellation of research and operational satellites involving passive microwave (PMW) sensors. The core satellite carries the Ku/Ka-band Dual-Frequency Precipitation Radar (DPR) measuring reflectivity profiles and the GPM microwave radiometer (GMI) measuring brightness temperatures (T_b). The overlapping Ku- and Ka-band measurements provide estimates of the shape and size of hydrometeors and higher sensitivity for detection of light rain and snow than the single frequency radar capabilities of the TRMM Precipitation Radar (PR; Ku-band). The GMI extends the range of frequencies on the TRMM Microwave Imager to provide brightness temperatures covering the range of frequencies on GPM constellation radiometers. Precipitation estimates from the constellation sensors enable the creation of a quasi-global-scale and gridded merged precipitation product named Integrated Multi-Satellite Retrievals for GPM (IMERG).
For each precipitation product (DPR, GMI, or IMERG), the satellite rainfall estimate $R(A,T)$ is compared with a reference rainfall $R_{ref}(A,T)$ over a spatial domain A (satellite footprint or pixel for DPR and GMI or IMERG, respectively) over the time period T (snapshot or 30-min timescale for DPR and GMI or IMERG, respectively). The reference rainfall $R_{ref}(A,T)$ is a proxy of the true (and unknown) area-averaged rainfall rate over the same area A and time period T. While we do not know the truth at ground we need to correctly assess the reference's uncertainties for a reliable quantitative comparison of precipitation products.

1.1.2.1 Dual-frequency Phased Array Radar

The DPR version 5 product used in this work is described in Iguchi *et al.* (2009, 2010). It provides 3-D reflectivity and 2-D precipitation rate fields at ground. The product classifies snowfall and rain into stratiform and convective (Awaka *et al.*, 2007). The DPR algorithm relies on a hybrid attenuation correction method that combines the surface reference technique and Hitschfeld-Bordan method (Iguchi *et al.* 2000; Meneghini *et al.* 2000, 2004; Takahashi *et al.* 2006). It uses models to describe the hydrometeor particle size distributions (PSD) depending on the precipitation type, which are adjusted to match the observed dual-frequency Path Integrated Attenuation. The DPR is a well-calibrated and very stable radar. The scan geometry and sampling rate of the DPR lead to footprints spaced approximately 5 km cross- and along-track, over 245-km and 120-km wide swaths at Ku-band and Ka-band respectively, centered within the 885-km-wide GMI swath. DPR observations provide a more direct measurement of the precipitation rates than GMI. The minimum theoretical detectable precipitation rate by the DPR is fixed by its sensitivity and is better than 18 dBZ or 0.5 $mm.h^{-1}$, at Ku-band and around 12 dBZ at Ka-band or about 0.2 $mm.h^{-1}$. Through the combined DPR and GMI-based retrievals, the DPR may be regarded as a "calibrator" of the PMW precipitation estimates, while the passive sensors, already a component of several polar-orbiting observatories (e.g., SSMI, SSMI-S, AMSR-2), provide more extensive sampling of precipitation events over the globe. Uncertainty sources affecting the DPR-based precipitation estimates propagate into PMW estimates. DPR observations provide a more direct measurement of the precipitation rates than PMW sensors. However similarly to the TRMM-PR primary errors in rainfall retrievals are mainly attributed to attenuation correction of the radar signal and conversion from reflectivity-to-precipitation intensity. Both involve incorrect physical assumptions related to snowfall and convective versus stratiform rainfall classification and assumed particle size distribution. Contamination by surface backscatter and Non-uniform beam filling (NUBF) are other challenges in correctly interpreting the radar signal into precipitation (Wolff and Fisher, 2008; Iguchi *et al.* 2009). Retrieval of precipitation rate requires knowledge of precipitation type, on which the PSD depends (Battan 1973; Awaka et al. 2007), and which has profound impacts on the accuracy of the quantitative retrievals. It drives the vertical model of microphysics used to correct for the attenuation of the DPR signal, to estimate the vertical profile of reflectivity, and the rainfall rate at ground (Iguchi et al. 2009). Satellite precipitation classification relies partly on subjective analysis based on interpretation of reflectivity spatial variability. While the DPR algorithm classifies rain into three categories: convective, stratiform and others, there is actually a mixture of processes and types within the footprint (Kirstetter et al. 2015). Classification capabilities and their impact have not been evaluated extensively.

Our dataset covers 30 months (June 2014 - October 2016) of satellite overpasses over the CONUS. The variable precipRateESurface (estimated surface rain) from the matched scan GPM-DPR/Ka-Ku (hereafter DPR) was extracted from the DPR files as the DPR surface QPE. The use of gridded MRMS data for reference provided a large sample size totaling 1,012,151 non-zero DPR-reference pairs including 798,155 pixels classified as stratiform by the DPR and 196,801 convective pixels. Along with the GPM-DPR/Ka-Ku, other products used in this study are the GPM-DPR/Ku and GPM-DPR/Ka. A TRMM-PR dataset is also used for comparison, covering the period from March to October 2011 (Kirstetter et al. 2015).

1.1.2.2 GPM Microwave Imager

The GPM Microwave Imager (GMI) measures brightness temperatures T_b at thirteen microwave frequencies ranging from 10 GHz to 183 GHz, which are used by the Goddard Profiling Algorithm (GPROF) to estimate surface rain rates and vertical hydrometeor profiles (Kummerow *et al.* 2017). The GPROF uses a Bayesian approach and an a priori database populated with combined DPR and GMI-based precipitation. Note the DPR precipitation types are inputs for the combined DPR and GMI-based precipitation retrievals hence impact the GPROF a priori database.

Sources of systematic errors in GPROF are related to the a priori database and ancillary information used to subset the a priori database. Over land the hydrometeor information is entangled with highly

variable surface characteristics in T_b. The main information used for the retrieval comes from the scattering channels because of the radiometrically warm land surface and variable surface emissivity. As the T_b at these frequencies are mostly sensitive to the scattering processes in the higher regions of the cloud (Wilheit *et al.* 2003), the available information for the retrieval is not directly related to surface precipitation (the correlation between ice aloft and surface rainfall is variable). Hence, there are particular needs for a detailed assessment of GPROF performances especially regarding the quantitative retrieval (Gopalan *et al.*, 2010). Currently GPROF does not condition precipitation retrievals by precipitation types (convective/stratiform), although recent works indicate that atmospheric stability and precipitating system structure impact its performance (Petkovic et al. 2017, 2018; Hendersen et al. 2017).
The GPM constellation includes a number of satellites with GMI-like radiometers or microwave sounding instruments, including the DMSP F19 and F20 (U.S.; imager), GCOM-W1 (JAXA; imager), Megha–Tropiques (CNES/ ISRO; sounder), MetOp (EUMETSAT; sounder) and NOAA 19 (NOAA; sounder) satellites. The GPM core satellite sensors provide self-consistent radiometric observations across the constellation. The DPR-calibrated GMI is, in turn, used as the calibrator for other passive microwave sensors as GPROF is consistently applied on the GPM constellation PMW sensors that collectively enable the creation of quasi-global-scale combined precipitation products. Thus, DPR and GMI have fundamental impacts on satellite-based precipitation estimates from other low Earth-orbiting passive microwave measurements and a number of satellite-based, high-resolution precipitation products.
The GPROF-GMI version 5 product is used here. The datasets for GMI covers the same period as for DPR (from June 2014 to October 2016) with a sample size totaling 3,782,453 non-zero GPROF-GMI-reference pairs.

1.1.2.3 Integrated Multi-satellitE Retrievals for GPM

The radiometer-based precipitation estimates impact the merged precipitation product IMERG by creating a uniformly gridded precipitation product at the global scale. IMERG blends complementary satellite-based precipitation estimates, i.e. from IR and PMW sensors (Huffman et al., 2014). IR data have a more indirect relation with surface precipitation occurrence and rates than PMW observations (Kirstetter et al. 2018). IMERG take advantage of the complementary more accurate GPROF retrievals from all constellation sensors and more frequent IR-based Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System algorithm (PERSIANN-CCS; Hong et al., 2004) to produce near-real time estimates at high spatial and temporal resolution (0.1°; 0.5h) with quasi-global coverage. To increase their coverage PMW estimates are morphed with the Climate Prediction Center (CPC) Morphing-Kalman Filter using the motion of clouds from infrared imagery (Joyce et al., 2004) before being merged with PERSIANN-CCS precipitation. IMERG has three separate products called Early with 4h latency used e.g. for flood and landslide monitoring, Late with 12h latency for drought monitoring and agricultural applications, and Final with 2 months latency for research applications. The IMERG Early and Late products are adjusted to climatological sets of coefficients and IMERG final product to gauge observations.
Whereas IMERG provides an approach to high-resolution precipitation estimation, uncertainty increases with resolution due to the combined effect of higher variability of precipitation and more indirect information provided by higher-resolution IR relative to PMW remote sensing. While fine-resolution quantitative precipitation estimates (QPEs) are needed for many applications, they come at the expense of degraded accuracy. Satellite estimates are currently limited in capturing extreme precipitation events often related to convective precipitation, which impedes short-term warning systems based on satellite precipitation.
IMERG version V05 Late is employed in this study. The datasets for IMERG covers the same period (from June 2014 to October 2016) with a sample size totaling 11,796,935 non-zero IMERG-reference

pairs. The information about the data used in this study is archived in the following fields in the IMERG files: *precipCal* for the merged products, *HQprecipitation* for the MW component, and *IRprecipitation* for the IR component (Huffman et al., 2014).

1.1.2.4 Ground-based reference precipitation

Over the U.S., Kirstetter et al. (2012, 2014) set up a framework to compare the GPM constellation sensors using the ground-based Multi-Radar/Multi-Sensor (MRMS) system (Zhang, 2016). It provides an independent reference for space-based precipitation products regarding resolution, accuracy and sample size, and a consistent database in time and space up to 55° in latitude including various geographical (plains, mountains) and meteorological (subtropical to midlatitudes) conditions for robust comparison (Fig. 1.1.1). To perform end-to-end uncertainty characterization the ground reference derived from MRMS is designed to bridge the scales of multiples precipitation measurements from local gauges to regional ground-based radar coverage to satellite DPR and GMI field of view (FOV) and swaths to the resolution of IMERG gridded products. It combines the fine spatial and temporal depiction of precipitation variability by ground-based radars with the quantitative accuracy of gauges and match this information at the scale of the satellite precipitation estimates.

The NOAA/NSSL and University of Oklahoma (OU) MRMS system (https://mrms.nssl.noaa.gov/) incorporates data from all S-band polarimetric WSR-88D radars and automated rain gauge networks in the CONUS (Zhang *et al.* 2016). The WSR-88D radars' sensitivity (about -5dBZ at a distance of 50 km) allows provides high performance for rainfall detection at least at confined ranges from the radar (e.g. < 75 km) compared to satellite sensors. At S-band the radar signal is relatively unattenuated. Dual-polarization improves the radar data quality and enables targeting specific microphysical situations where the ground-radar estimates are the most reliable (Ryzhkov et al., 2005, 2014; Chandrasekar et al., 2008). The MRMS system generates and high-resolution 3-D reflectivity mosaic grids and a suite of severe weather and QPE products at a 0.01° horizontal resolution and 2-minute update cycle: rain types, precipitation phase, rain rate, freezing level height, etc. The geographical extension and time period covered by MRMS compiles a broad sample of precipitation situations (stratiform, convective, orographically enhanced precipitation, etc.) from a stable and uniform observation system (Kirstetter et al. 2012, 2014). At hourly time step, MRMS adjusts radar estimates with automated rain gauge networks using a spatially variable bias multiplicative factor. A radar quality index (RQI) is produced to represent the radar QPE uncertainty associated with reflectivity changes with height and near the melting layer (Zhang *et al.* 2011).

One should note that it is not possible to "validate" satellite precipitation estimates in a strict sense because independent precipitation estimates with no uncertainty do not exist. Yet trustworthy values of the MRMS rainfall estimates within the satellite pixel are needed to evaluate the satellite estimates. A reference matched to satellite sampling properties is derived for Level-2 and Level-3 precipitation. Postprocessing complements procedures already in place within the MRMS system to further refines and standardize the reference dataset. Several quality control steps are applied in order to use ground radar data only where confidence in skill is very high and artifacts (beam blockage due to mountain, range effects, etc.) are essentially minimized and deemed negligible. Blending techniques build on the locally reliable rain gauge measurement and the space-time resolution of the radar (Kirstetter *et al.* 2012, 2013a). A conservative approach is followed by (i) filtering out instances when the radar and gauge have significant quantitative disagreement (i.e. radar- rain gauge ratios outside of the range [0.1-10]) and (ii) by retaining only the best measurement conditions (i.e., no beam blockage and radar beam below the melting layer) using the RQI product as described in Kirstetter *et al.* (2012, 2013a). These data quality controls standardize the reference product and filter out the less trustworthy MRMS estimates, e.g. radar measurements at far range and in the Intermountain West. These improvements may not screen out all possible errors in ground-based radar estimates. Kirstetter *et al.* (2012) showed the increased consistency between TRMM-PR and the MRMS-based reference

following sequential MRMS data quality control steps including bias correction using rain gauges and filtering using the RQI product. This finding highlights the importance of matching the scales and refining the accuracy of the reference dataset as much as possible before reaching meaningful conclusions about the satellite sensor's accuracy.

Because of the highly variable precipitation processes, the resolution of the ground reference should match the satellite products resolution to reduce noise when comparing satellite retrievals to $R_{ref}(A,T)$. The spatial resolution of MRMS is finer than any satellite sensor, allowing the resolution of the reference to be specifically adapted to each product using spatial and temporal sampling techniques (Kirstetter *et al.* 2012). To determine the reference rainfall $R_{ref}(A,T)$ over the sensor pixel A and time period T, a block-MRMS precipitation pixel matches each sensor/product pixel. For Level-2 products the MRMS products closest in time to the GPM core satellite local overpass schedule time are used. The time difference between the MRMS data and satellite data (max 1 min.) could add random noise (but not bias) in the comparison, although other factors like the difference in resolution may have more significant impacts. For Level-3 products the MRMS products are aggregated to match the grid spacing at 0.1° and the 30-min temporal accumulation. All of the MRMS pixels (rainy and non-rainy) found within a footprint sensor or pixels are located to compute unconditional mean rain rates for the MRMS at the sensor pixel scale. When more than 25% of MRMS pixels have missing values, the data pair is discarded from the comparison. Matched DPR/GMI/IMERG and $R_{ref}(A,T)$ estimates only exist at locations where both the satellite and ground radars have taken actual observations. The satellite products remain untouched, hence preserving their characteristics: the total rainfall amount, the total rainy area, the convective/stratiform contribution and the PDF shapes, and may therefore be compared to the reference at once. An extended description of the reference precipitation regarding types of rainfall within the satellite sensor's FOV or pixel grid is assessed through a Convective Percent Index quantifying the volume contribution of convective rainfall to $R_{ref}(A,T)$. The CPI is expressed in percent between 0% (purely stratiform rainfall within the FOV or pixel) to 100% (purely convective rainfall). CPI values betweem 0% and 100% indicate mixed precipitation types. Given its spatial extension and native resolution, the MRMS products provide large samples of matched ground-satellite comparison pairs (Kirstetter et al. 2012; 2013a,b; 2014; 2015a,b; 2018). This allows judicious selections of data pairs to monitor the comparison quality. All significant rain fields observed coincidentally by DPR and GMI overpasses and IMERG and the NEXRAD radar network from June 2014 to October 2016 are collected.

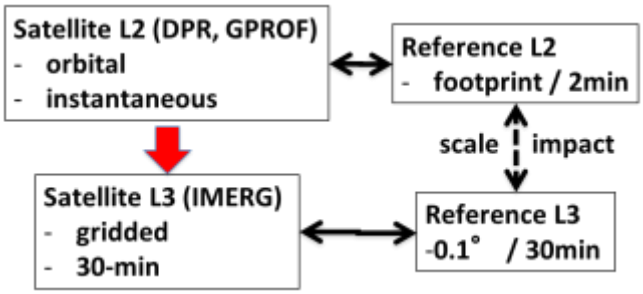

Figure 1.1.2. Comparison framework across Level-2 and Level-3 satellite QPE to bridge from the GPM core satellite to IMERG.

Given the highly variable precipitation processes, a consistent reference derived from MRMS is instrumental to bridge in-depth characterization across Level 2 constellation sensors from the GPM core to the constellation to IMERG (Level 3), directly assessing the influence of precipitation types.

1.1.3 Impact of precipitation typology on satellite-based active, passive and merged precipitation estimation

Precipitation error characterization is conditioned on precipitation types by taking full advantage of the MRMS insights into precipitation across scales (from sub-satellite Level-2 FOV to IMERG grid). The state dependency of the error is diagnosed by stratifying the error according to the CPI across sensors and products.

1.1.3.1 DPR QPE and precipitation typology

The DPR algorithm uses different reflectivity-to-precipitation rate relationships in the convective/stratiform profiling components. The CPI continuous classification provides a fine assessment of the DPR categorical classification and its impact on rate retrievals. Figure 1.1.3 shows the reference and DPR rainfall rate distributions as functions of the convective contribution CPI. All coincident and collocated DPR values are considered and sorted according to the reference sample. Figure 1.1.3 shows a shift toward higher rainfall rates as CPI increases, as we would expect. Despite these consistencies, we note rain rate distributions indicating higher rainfall rates for the reference compared to those of DPR (i.e. Figure 1.1.3(a) compared to 1.1.3(b)). The dynamic ranges of rain rate distribution are greater for the reference than for the DPR. Such differences, which will undoubtedly result in some bias, could be related to the DPR reflectivity-to-precipitation relationships. The difference is larger for convective contribution > 80%, e.g. for CPI = 90% the conditional reference median rate is 15 mm.h^{-1} while the DPR median rate is 7 mm.h^{-1}. Apparently, the DPR profiling algorithm lacks sufficient dynamics to deal with extreme rainfall amounts.

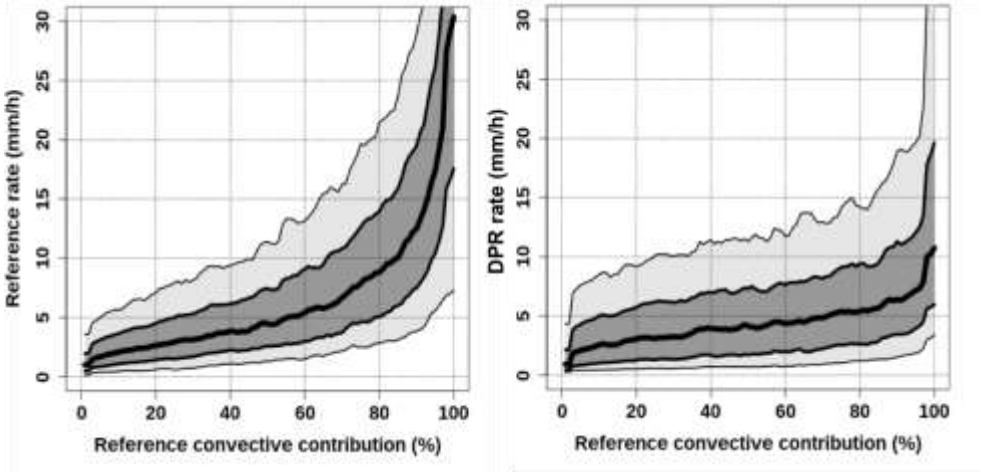

Figure 1.1.3. Reference (left) and DPR (right) rainfall rate distributions (mm.h^{-1}) as functions of the CPI (%). The thick black line represents the median (50% quantile), the dark grey-shaded region represents the area between the 25 and 75% quantiles, the light grey-shaded region represents the area between the 10 and 90% quantiles.

The precipitation type impacts the discrepancies of the DPR relative to the reference. In order to provide some insight into the influence of precipitation typing on the DPR error, the departures of DPR estimates from the reference values are analyzed as functions of the convective contribution on a point-to-point basis. The residuals are defined as the difference between the reference rainfall (R_{ref}) and the satellite estimates (R): $\epsilon = \left(R - R_{ref}\right)$. Only pairs for which R_{ref} and R are both nonzero are considered in the calculations, so as to remove any discrepancies related to detectability. Figure 1.1.4 shows the residuals as a function of the convective contribution CPI. All coincident and collocated PR values are considered and sorted according to the reference sample.

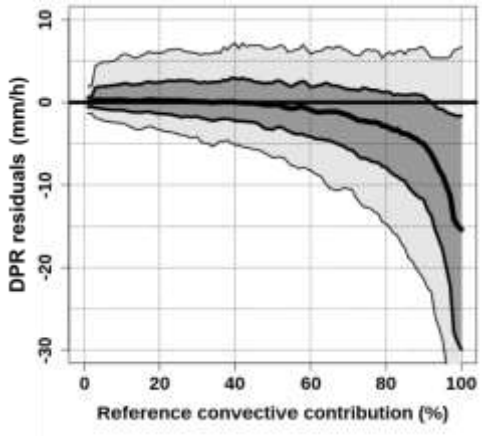

Figure 1.1.4. DPR residual distributions (mm.h^{-1}) as functions of the CPI (%). The thick black line represents the median (50% quantile), the dark grey-shaded region represents the area between the 25 and 75% quantiles, the light grey-shaded region represents the area between the 10 and 90% quantiles.

The conditional PDFs of residuals ϵ present a high conditional shift from the 0 line and a large conditional spread. The spread increases with CPI, indicating larger uncertainties in quantifying the reference convective rainfall. The panel shows also a tendency to overestimate rain rates for low CPI values (CPI < 30%, the conditional median of residuals is positive), a shift toward lower rainfall rates as CPI increases, and rain rates underestimation for CPI values higher than 50% (negative median of residuals). As an example, the DPR model overestimates at CPI = 20% (mainly stratiform) rain rates with an occurrence of 55% and underestimates at CPI = 100% (convective) rain rates with an occurrence of 80%.

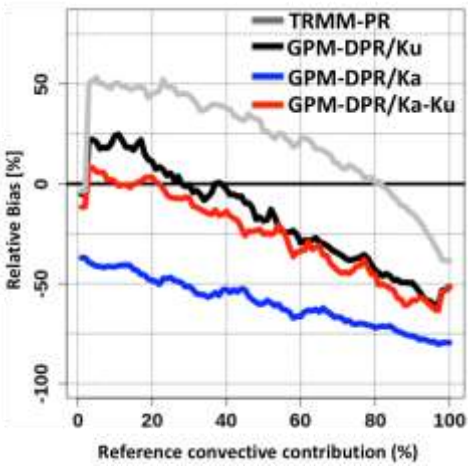

Figure 1.1.5. Conditional bias of spaceborne radars relative to the reference as a function of the CPI (%) for the GPM-DPR/Ku (black), GPM-DPR/Ka (blue), GPM-DPR/Ka-Ku (red) and TRMM-PR (grey).

Because of the asymmetric density of residuals and for a better representativity and intercomparison across products, we consider the conditional mean relative bias of the residuals ($MRE = 100\frac{\overline{R_{ref}}-\overline{R}}{\overline{R_{ref}}}$ in percent) to compare the systematic error components. Figure 1.1.5 shows the conditional biases as a function of the CPI for GPM-DPR/Ku, GPM-DPR/Ka, GPM-DPR/Ka-Ku and TRMM-PR. The conditional biases are distinct according to the satellite-based radar product but present also similar features. Except for GPM-DPR/Ka, all biases are slightly negative and within 10% for CPI=0%

(stratiform precipitation). These biases display an abrupt shift toward higher values (e.g. +50% for TRMM-PR, +23% for GPM-DPR/Ku and +11% for GPM-DPR/Ka-Ku) for CPI shifting from 0% to 3%. This might be due to the satellite profiling algorithms interpreting this mixed microphysics within the FOV with the stratiform parameterization. For CPI < 20% the GPM-DPR/Ka-Ku bias covers a narrower range (from -10% to +11%) than TRMM-PR and GPM-DPR/Ku. It probably illustrates the benefits of using two frequencies to interpret the vertical structure of PSD and hydrometeor properties and estimate surface precipitation rates. All biases decrease with CPI and reach negative values for CPI>80%. Apparently, all profiling algorithms lack sufficient dynamics to deal with extreme rainfall rates. In addition to the reflectivity-to-precipitation rate relationships, the consistent underestimation at high rainfall rates associated with high CPI values could also be related to the attenuation correction algorithm, or even total loss of the radar signal, specifically for the GPM-DPR/Ka as the signal at this frequency is more attenuated than at Ku-band. The TRMM-PR bias presents a shift towards higher values compared to GPM-DPR and present relatively limited biases for high CPI values (biases within 50% for CPI>70%), which is consistent with the design and application of the TRMM-PR profiling algorithm on intertropical precipitation. Note that the PR eatimates surface rain rates over the southern US up to a latitude of 37°N, which is dominated by deep convection during the warm season. The GPM-DPR algorithms display relatively limited biases for light CPI values (biases within 25% for CPI<50%) consistent with their application to mid-latitude precipitation. GPM-DPR/Ka shows systematic underestimation (biases ranging from -77% to -35%) possibly related to the strong attenuation of the signal at Ka-band.
To summary the categorical classification of precipitation types in the space-borne radar algorithms does not handle mixed types and microphysics within the FOV and appear to lack of sufficient dynamics to deal with extreme rainfall amounts.

1.1.3.2 GMI QPE and precipitation typology

Currently GPROF does not condition precipitation retrievals by precipitation types, although recent works indicate that atmospheric stability and precipitating system structure impact its performance (Petkovic et al. 2017, 2018; Hendersen et al. 2017). However GPROF-GMI quantifies the convective contribution in precipitation rate retrievals (CPI_{GPROF}). CPI provides a reliable assessment of the output convective contribution CPI_{GPROF} inside a passive sensor FOV. We focus here on cases when both GPROF-GMI and the reference precipitation are greater than zero over the CONUS dataset, so that precipitation detection is not a factor.
Figure 1.1.6 shows the cumulative distributions of the reference and GPROF-GMI convective contributions. The reference stratiform precipitation (CPI=0%) dominates at 81% the reference precipitation typology in terms of the fraction of type occurrence, while the fraction of convective/stratiform mix (0% < CPI < 80%) is 15%, and the fraction of primarily convective precipitation (CPI > 80%) is only 4%. The GPROF-GMI cumulative distribution is significantly different as stratiform precipitation (CPI=0%) is not frequently retrieved (<1%), the fraction of convective/stratiform mix is significantly higher than the reference at 96%, and the fraction of primarily convective precipitation (CPI > 80%) is 4%. While the Bayesian methodology in the GPROF-GMI yields estimates that are well constrained as designed from the a-priori database, this approach apparently doesn't accommodate precipitation types.

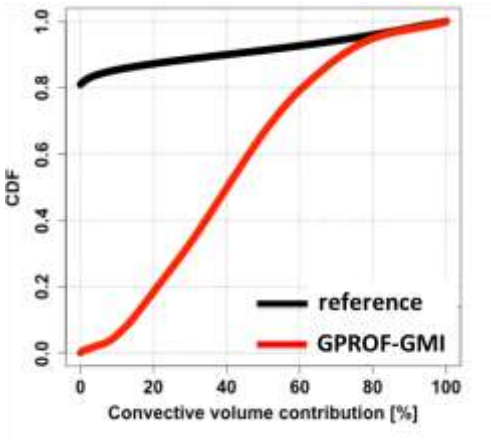

Figure 1.1.6. Reference and GPROF-GMI cumulative distribution of convective contribution.

A conditional error analysis is applied to target biases over different precipitation type conditions. Figure 1.1.7 illustrates GPROF-GMI residuals as a function of the convective contribution CPI. All coincident and collocated GPROF-GMI values are considered and sorted according to the reference sample.

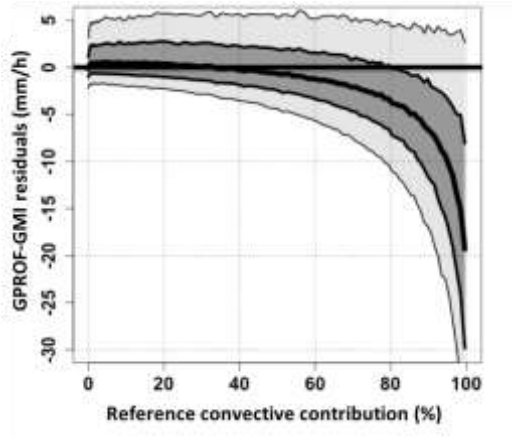

Figure 1.1.7. GPROF-GMI residual distributions ($mm.h^{-1}$) as functions of the CPI (%). The thick black line represents the median (50% quantile), the dark grey-shaded region represents the area between the 25 and 75% quantiles, the light grey-shaded region represents the area between the 10 and 90% quantiles.

The systematic error is extracted as the conditional median of residuals for a better representativity given the asymmetric density of residuals. The random error is assessed with the interquantile (90% - 10%) value. Both systematic and random errors are displayed in Figure 1.1.8. Similarly to the DPR (Fig. 1.1.4), the conditional PDFs of residuals ϵ present a high conditional shift from the 0 line and a large conditional spread. The spread increases with CPI, indicating larger uncertainties in quantifying the reference convective rainfall. It is confirmed with the random error increasing with CPI (Fig. 1.1.8). As for DPR for stratiform situations (CPI = 0%), the systematic error is low, and the random error minimal. Systematic error displays an abrupt shift toward higher values for CPI from 0% to 3%. There is a systematic overestimation of rain rates for low CPI values (CPI < 25%, the conditional median of residuals is positive in Fig. 1.1.8), a shift toward lower rainfall rates as CPI increases, and rain rates underestimation for CPI values higher than 50% (negative median of residuals, see Fig. 1.1.8). As an example at CPI = 20% (mainly stratiform) the GPROF-GMI retrieval overestimates rain rates with an occurrence of 55%, and at CPI = 100% (convective) it underestimates rain rates with an occurrence of 85%. The similarities between the GPROF-GMI and DPR residuals as functions of CPI

(Figures 1.1.4 and 1.1.7) may be caused by the propagation of precipitation type related errors from the DPR profiling algorithm to the GPROF-GMI retrieval database. In particular DPR interprets mixed microphysics within the FOV with a specific parameterization type (e.g. stratiform) and lacks sufficient dynamics to deal with extreme rainfall amounts.
To investigate the potential improvement in QPE if GPROF-GMI correctly estimates the convective contribution, the CPI difference $\epsilon_{CPI} = CPI_{GPROF} - CPI$ is computed. A positive (negative) ϵ_{CPI} indicates that the GPROF-GMI retrieval is more convective (stratiform) than the reference. Figure 1.7.8 shows the GPROF surface precipitation mean relative bias and correlation as functions of ϵ_{CPI}.

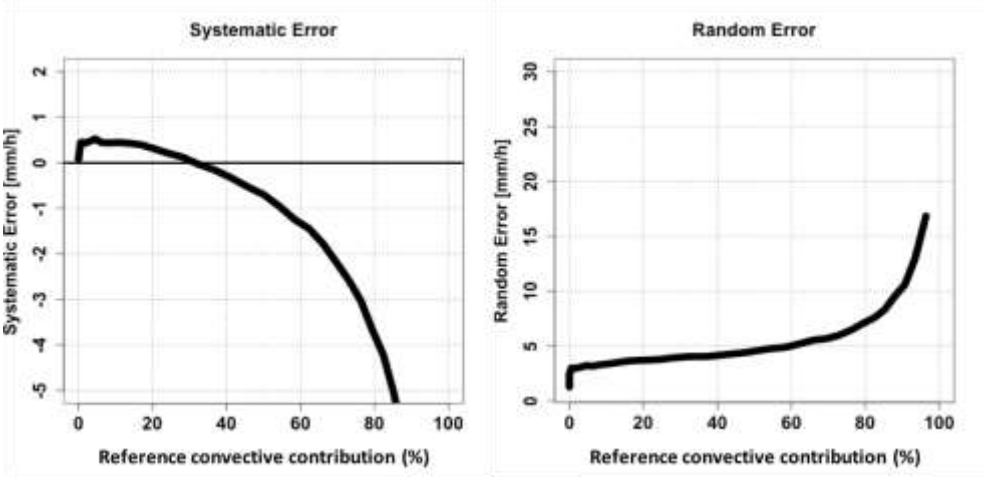

Figure 1.1.8. GPROF-GMI (left) systematic part –conditional median - and (right) random part – interquantile 10%-90%- of error.

The precipitation rate correlation between GPROF-GMI and the reference displays a strong dependence with ϵ_{CPI}, taking on lower values when GPROF-GMI and the reference precipitation types are significantly different. The correlation is above 0.65 when ϵ_{CPI} ranges from -1% to +11% and maximizes at 0.71 when ϵ_{CPI} is +8%. The mean relative bias of GPROF-GMI increases with ϵ_{CPI} and shows underestimation (overestimation) when GPROF is more stratiform (convective) than the reference, as one can expect.

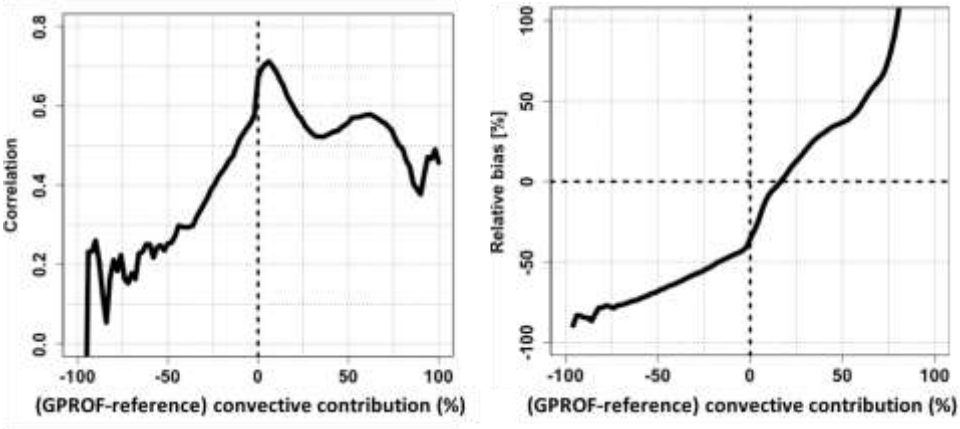

Figure 1.1.9. GPROF-GMI (left) correlation and (right) relative bias as a function of convective contribution difference between GPROF-GMI and the reference.

Considering that GPROF-GMI retrieves much less stratiform precipitation than the reference (Fig. 1.1.6), it can partly explain the positive precipitation rate bias for reference CPI < 30% (Fig. 1.1.8).

The bias is 0 for the ϵ_{CPI} = +8%, which is also where the correlation maximizes. Improving both the bias and correlation is noteworthy. Ciach et al. (2000) show that postprocessing optimization of a precipitation product relative to a reference can be done by improving the bias (systematic error) or the mean square error (random error), but not both. Matching the precipitation types output from GPROF-GMI shows improvements in both bias and correlation, which demonstrate the potential benefit in conditioning precipitation rate retrievals by precipitation types in the algorithm.

1.1.3.3 IMERG and precipitation typology

To investigate the impact of precipitation types on IMERG, Figure 1.1.10 displays the residuals of IMERG estimates relative to the reference as functions of the convective contribution on a pixel-to-pixel basis. The residuals are consistently defined as the difference between the IMERG estimates (R) the reference rainfall (R_{ref}): $\epsilon = \left(R - R_{ref}\right)$ at the resolution [0.1°; 30-min]. Only pairs for which R_{ref} and R are both nonzero are considered in the calculations, so as to remove any discrepancies related to detectability. All coincident and collocated IMERG values are considered and sorted according to the reference sample.

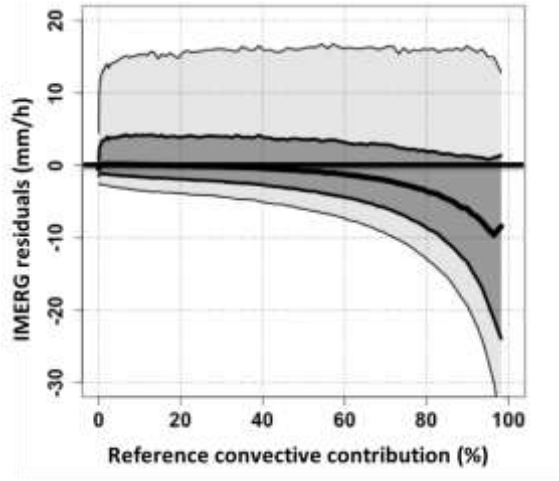

Figure 1.1.10. IMERG residual distributions (mm.h^{-1}) as functions of the CPI (%). The thick black line represents the median (50% quantile), the dark grey-shaded region represents the area between the 25 and 75% quantiles, the light grey-shaded region represents the area between the 10 and 90% quantiles.

Consistent with DPR and GPROF-GMI, IMERG conditional PDFs of residuals ϵ present a high conditional shift from the 0 line and a large conditional spread. The spread increases with CPI, indicating larger uncertainties in quantifying the reference convective rainfall at the IMERG scale. The panel shows a shift toward lower rainfall rates as CPI increases, and rain rates underestimation for CPI > 50% (negative median of residuals). It illustrates the propagation of the GPROF biases with respect to precipitation types into IMERG.
Further insight is provided in Figure 1.1.11 showing the systematic error (median of residuals) and random error (interquantile 90%-10%) of the PMW and IR components of IMERG. The conditional biases are distinct according to the IMERG components. Both PMW and IR systematic and random errors are minimized for stratiform situations (CPI = 0%). Over the range of CPI values, systematic and random PMW errors show similar behavior to GROF-GMI (Fig. 1.1.8), especially for CPI values from 0% to a few percent. Over the same CPI range, while PMW systematic errors display a positive shift the IR systematic errors display the opposite behavior. The PMW systematic biases present a shift towards higher values (+1.5 mm.h-1) compared to IR biases for CPI < 50%. They are both decreasing functions of the convective contribution CPI, with the PMW component overestimating

at CPI < 45% while the IR component underestimates over the whole CPI range. Both components lack sufficient dynamics to deal with extreme rainfall amounts. The random discrepancies increase consistently with CPI for both components, and it is worth noting that the PMW random error is greater (~ +5 mm.h-1) than the IR component.

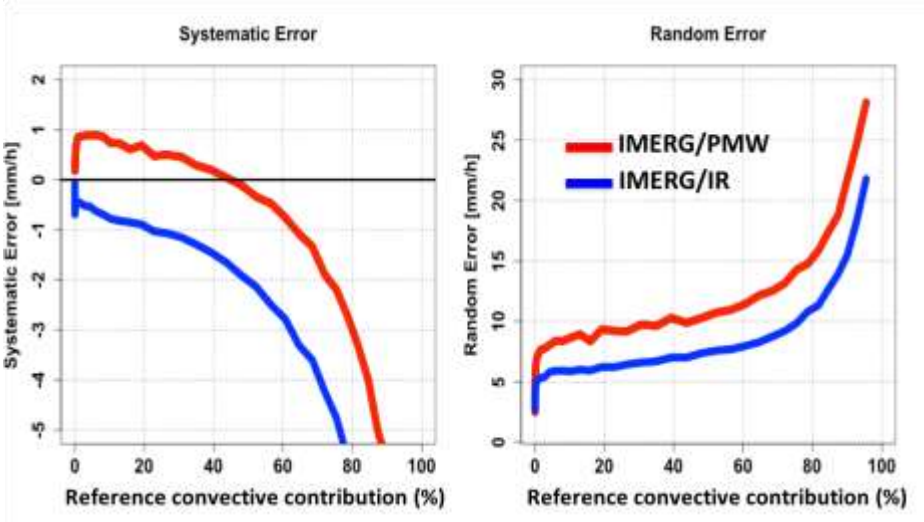

Figure 1.1.11. IMERG (left) systematic part –conditional median - and (right) random part – interquantile 10% - 90% - of error for the PMW component (red) and IR component (blue).

This analysis confirms that satellite estimates are currently limited in capturing extreme precipitation events often related to convective precipitation.

1.1.4 Conclusion

Integrated multi-sensor assessment is proposed as a novel approach to advance satellite QPE validation, in order to provide precipitation algorithm developers and users with more than just an overall assessment and adequately cope with the varying performances of satellite precipitation estimates. End-to-end integrated error quantification tracking the origin and propagation of uncertainty from Level-2 active, passive then through Level-3 precipitation products is particularly relevant for understanding the origin and impact of uncertainty. Precipitation typology is addressed in this context as a relevant factor driving the state of the GPM satellite estimation error and leading to a generalization of the assessment.

Such assessment is performed at the primary satellite QPE scale across products for a quantitative and detailed characterization. It is tested over a multi-year and multi-scale data sample of TRMM-PR, DPR, GMI and IMERG satellite precipitation products and a high-quality precipitation reference derived from MRMS. The MRMS-based convective contribution was adapted at multiple resolutions to perform precipitation error characterization conditioned on precipitation types, by taking full advantage of the MRMS insights into precipitation across scales (from sub-satellite Level-2 FOV to IMERG grid). The state depency of the error is diagnosed by stratifying the error according to the CPI across sensors and products.

For DPR (and TRMM-PR), incorrect physical assumptions related to convective versus stratiform rainfall classification are confirmed to be a primary error impacting the conversion from reflectivity-to-rainfall intensity. The non-uniformity of precipitation types within the FOV is a driving error contributor. Both DPR stratiform and convective profiling algorithms seem to lack sufficient dynamics to deal with extreme rainfall amounts. For stratiform FOV-filling conditions DPR algorithms understimate precipitation rates relative to the reference, overestimates mixed stratiform dominated conditions and underestimates for convective situations. Uncertainty in estimating

precipitation rates increases in case of convective precipitation. Results from the error analysis presented herein provide insights into significant characteristics of DPR rainfall retrieval errors that need to be considered to improve the retrieval and when such data are used in applications.
Based on passive sensors observations and conditioned on DPR outputs, GPROF-GMI retrievals display similar error features as DPR with respect to precipitation types. It probably inherits the lack of sufficient dynamics to deal with extreme rainfall amounts observed with the active profiling algorithm. This may be amplified by the GPROF Bayesian approach used to retrieve precipitation in GPROF-GMI. Room for improvement is shown to reside in conditioning the GPROF retrievals by precipitation types, and this can be considered for future versions of the algorithm. GPROF retrievals systematic and random errors with respect to precipitation types propagate into the IMERG PMW component.

References

Anagnostou, E. N., 2004: Overview of overland satellite rainfall estimation for hydrometeorological applications. *Surv. Geophys.*, **25**, 511–537.

Awaka, J., T. Iguchi, and K. Okamoto, 2007: Rain Type Classification Algorithm. *Advances In Global Change Research*, **28**, 213-224.

Battan, L. J., 1973: Radar Observations of the Atmosphere. *University of Chicago Press*, 324 pp.

Bauer P., T. Auligné, W. Bell, A. Geer, V. Guidard, S. Heilliette, M. Kazumori, M.-J. Kim, E. H.-C. Liu, , A.P. McNally, B. Macpherson, K. Okamoto, R. Renshaw and L.P. Riishøjgaard, 2011: Satellite cloud and precipitation assimilation at operational NWP centres. *Q.J.R. Meteorol. Soc.*, **137**, 1934–1951. doi:10.1002/qj.905.

Carr, N., Kirstetter, P., Hong, Y., Gourley, J.J., Schwaller, M., Petersen,W.,Wang, N., Ferraro, R.R. and Xue, X., 2015: The influence of surface and precipitation characteristics on TRMMmicrowave imager rainfall retrieval uncertainty. *Journal of Hydrometeorology*, **16**, 1596–1614. https://doi.org/10.1175/JHMD-14-0194.1.

Chandrasekar, V., A. Hou, E. Smith, V.N. Bringi, S.A. Rutledge, E. Gorgucci, W.A. Petersen, G. Skofronick Jackson, 2008: Potential Role Of Dual- Polarization Radar In The Validation Of Satellite Precipitation Measurements: Rationale and Opportunities. *Bull. Am. Meteorol. Soc.*, **89** (2008), pp. 1127–1145

Derin, Y., E. Anagnostou, A. Berne, M. Borga, B. Boudevillain, W. Buytaert, C. Chang, G. Delrieu, Y. Hong, Y.C. Hsu, W. Lavado-Casimiro, B. Manz, S. Moges, E.I. Nikolopoulos, D. Sahlu, F. Salerno, J. Rodríguez-Sánchez, H.J. Vergara, and K.K. Yilmaz, 2016: Multiregional Satellite Precipitation Products Evaluation over Complex Terrain. *J. Hydrometeor.*, **17**, 1817–1836, https://doi.org/10.1175/JHM-D-15-0197.1

Derin, Y., and K. K. Yilmaz, 2014: Evaluation of multiple satellite-based precipitation products over complex topography. *J. Hydrometeor.*, **15**, 1498-1516

Dinku, T., S. J. Connor, and P. Ceccato, 2010: Comparison of CMORPH and TRMM-3B42 over mountainous regions of Africa and South America. Satellite Rainfall Applications for Surface Hydrology, M. Gebremichael and F. Hossain, Eds., Springer, 193–204.

Ebert, E., 2007: Methods for verifying satellite precipitation estimates. In: *Measuring Precipitation from space: EURAINSAT and the future*. Levizzani, V., P. Bauer, and F.J. Turk, eds., Springer, 345-356.

Golian, S., S. Moazami, P.E. Kirstetter, Y. Hong, 2015: Evaluating the Performance of Merged Multi-Satellite Precipitation Products Over a Complex Terrain. *Water Resour. Manage.*, **29**(13), 4885-4901. doi:10.1007/s11269-015-1096-6

Gopalan K., N.Y. Wang, R. Ferraro and Chuntao Liu, 2010. Status of the TRMM 2A12 Land Precipitation Algorithm. *Journal of Atmospheric and Oceanic Technology*, **27**:8, 1343-1354.

Grams, H.M., P. Kirstetter, and J.J. Gourley, 2016: Naïve Bayesian Precipitation Type Retrieval from Satellite Using a Cloud-Top and Ground-Radar Matched Climatology. *J. Hydrometeor.*, **17**, 2649–2665. doi.org/10.1175/JHM-D-16-0058.1

Grimes, D.I.F., and M. Diop, 2003: Satellite-based rainfall estimation for river flow forecasting in Africa. I: Rainfall estimates and hydrological forecasts. *Hydrolog. Sci. J.*, **48**, 567–584.

Grecu, M., and E. N. Anagnostou, 2001: Overland precipitation estimation from TRMM passive microwave observations. *J. Appl. Meteorol.*, **40**, 1367-1380.

Henderson, D.S., C.D. Kummerow, D.A. Marks, and W. Berg, 2017: A Regime-Based Evaluation of TRMM Oceanic Precipitation Biases. *J. Atmos. Oceanic Technol.*, **34**,2613–2635.

Hong, Y., K. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. *J. Appl. Meteor.*, **43**, 1834–1853. doi.org/10.1175/JAM2173.1

Hossain, F., and E. N. Anagnostou, 2004: Assessment of current passive-microwave and infrared satellite rainfall remote sensing for flood prediction. *J. Geophys. Res.*, **109**, D07102.

Hou, A. Y., R. K. Kakar, S. A. Neeck, A. Azarbarzin, C. D. Kummerow, M. Kojima, R. Oki, K. Nakamura, and T. Iguchi, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701-722. doi:10.1175/BAMS-D-13-00164.1.

Huffman, George J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeor*, **8**, 38–55.

Huffman, G., and Coauthors, 2014: last updated 2018: Algorithm Theoretical Basis Document (ATBD) Version 5.2 NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). NASA/GSFC, Greenbelt, MD, USA (http://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2.pdf)

Iguchi, T., T. Kozu, J. Kwiatkowski, R. Meneghini, J. Awaka, and K. Okamoto, 2009: Uncertainties in the rain profiling algorithm for the TRMM Precipitation Radar. *J. Meteor. Soc. Japan*, **87A**, 1–30.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto, 2000: Rain-profiling algorithm for the TRMM precipitation radar. *J. Appl. Meteor.*, **39**, 2038–2052.

Iguchi, T., S. Seto, R. Meneghini, N. Yoshida, J. Awaka, T. Kubota, 2010: Algorithm Theoretical Basis Document (ATBD). GPM/DPR Level-2 Algorithm Theoretical Basis Document (https://pmm.nasa.gov/sites/default/files/document_files/ATBD_GPM_DPR_n3_dec15.pdf).

Joyce, J. R., E. J. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487-503.

Kirstetter, P.E., Y. Hong, J.J. Gourley, Q. Cao, M. Schwaller, and W. Petersen, 2014: A research framework to bridge from the Global Precipitation Measurement mission core satellite to the constellation sensors using ground radar-based National Mosaic QPE. *In L. Venkataraman, Remote Sensing of the Terrestrial Water Cycle. AGU books Geophysical Monograph Series, Chapman monograph on remote sensing. John Wiley & Sons Inc. ISBN: 1118872037.*

Kirstetter, P.E., Y. Hong, J.J. Gourley, S. Chen, Z. Flamig, J. Zhang, K. Howard, W. Petersen, 2012: Toward a Framework for Systematic Error Modeling of Spaceborne Precipitation Radar with NOAA/NSSL Ground Radar-based National Mosaic QPE. *Journal of Hydrometeorology*, **13**(4), 1285-1300. doi:10.1175/JHM-D-11-0139.1

Kirstetter, P.E., Y. Hong, J.J. Gourley, M. Schwaller, W. Petersen, and Q. Cao, 2015: Impact of sub-pixel rainfall variability on spaceborne precipitation estimation: evaluating the TRMM 2A25 product. *Q.J.R. Meteorol. Soc.*, **141**, 953–966. doi:10.1002/qj.2416

Kirstetter, P.E., Y. Hong, J.J. Gourley, M. Schwaller, W. Petersen and J. Zhang, 2013a: Comparison of TRMM 2A25 Products Version 6 and Version 7 with NOAA/NSSL Ground Radar-based National Mosaic QPE. *Journal of Hydrometeorology*, 14(2), 661-669. doi:10.1175/JHM-D-12-030.1

Kirstetter P-E, Karbalaee N, Hsu K, Hong Y, 2018: Probabilistic precipitation rate estimates with space-based infrared sensors. *Q J R Meteorol Soc*; 1–15. https://doi.org/10.1002/qj.3243

Kirstetter, P.E., N. Viltard, and M. Gosset, 2013b: An error model for instantaneous satellite rainfall estimates: Evaluation of BRAIN-TMI over West Africa. *Quarterly Journal of the Royal Meteorological Society*, **139**, 894-911. doi:10.1002/qj.1964.

Kummerow, C. 2017: last updated 2017: GPROF2017 version 1. Algorithm Theoretical Basis Document (https://pmm.nasa.gov/sites/default/files/document_files/ATBD_GPM_GPROF_June1_2017.pdf).

Kummerow C, Berg W, Thomas-Stahle J, Masunaga H, 2006: Quantifying Global Uncertainties in a Simple Microwave Rainfall Algorithm. *J. Atmos. Oceanic Technol.*, **23** ,23–37.

Lebel, T., C. Cappelaere, S. Galle, N. Hanan, L. Kergoat, S. Levis, B. Vieux, L. Descroix, M. Gosset, and E. Mougin, 2009: AMMA-CATCH studies in the Sahelian region of West-Africa: an overview. *Journal of Hydrology*, **375**(1-2), 3-13.

Munchak., S. J., and G. Skofronick-Jackson, 2013: Evaluation of precipitation detection overvarious surfaces from passive microwave imagers and sounders. *Atmos. Res.*, **131**, 81-94.

Menegheni, R., T. Iguchi, T. Kozu, L. Liao, K. Okamoto, J.A. Jones, and J.R. Kwiatkowski, 2000: Use of the surface reference technique for path attenuation estimates from the TRMM Precipitation Radar. *J. Atmo. and Ocean. Tech.*, **40**, 2053–2070.

Menegheni, R., J.A. Jones, T. Iguchi, K. Okamoto, and J.R. Kwiatkowski, 2004: A hybrid surface reference technique and its application to the TRMM Precipitation Radar. *J. Atmo. and Ocean. Tech.*, **21**, 1645–1658.

Panegrossi, G., and Coauthors, 2016: Use of the GPM Constellation for Monitoring Heavy Precipitation Events Over the Mediterranean Region. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **9**, 2733-2753.

Petersen, W.A., and M.R. Schwaller, 2008: NASA GPM Ground Validation Science Implementation Plan. *NASA report*, http://pmm.nasa.gov/ sites/default/files/document_files/GPM_GVS_imp_plan_Jul08.pdf, 41 pp.

Petković, V. and C.D. Kummerow, 2017: Understanding the Sources of Satellite Passive Microwave Rainfall Retrieval Systematic Errors Over Land. *J. Appl. Meteor. Climatol.*, **56**, 597–614.

Petković, V., C.D. Kummerow, D.L. Randel, J.R. Pierce, and J.K. Kodros, 2018: Improving the Quality of Heavy Precipitation Estimates from Satellite Passive Microwave Rainfall Retrievals. *J. Hydrometeor.* **19**, 69–85.

Ryzhkov, A., M. Diederich, P. Zhang, and C. Simmer, 2014: Potential Utilization of Specific Attenuation for Rainfall Estimation, Mitigation of Partial Beam Blockage, and Radar Networking. *J. Atmos. Oceanic Technol.*, **31**, 599–619.

Ryzhkov, A.V., T.J. Schuur, D.W. Burguess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrnic´, 2005: The Joint Polarization Experiment: Polarimet- ric rainfall measurements and hydrometeor classification. *Bull. Amer. Meteor. Soc.*, **86**, 809–824.

Sapiano, M.R.P., and P.A. Arkin, 2009: An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data. *J. Hydrometeor.*, **10**, 149–166.

Skofronick-Jackson, G., W.A. Petersen, W. Berg, C. Kidd, E. F. Stocker, D.B. Kirschbaum, R. Kakar, S.A. Braun, G.J. Huffman, T. Iguchi, P.E. Kirstetter, C. Kummerow, R. Meneghini, R. Oki, W.S. Olson, Y.N. Takayabu, K. Furukawa, and T. Wilheit, 0: The Global Precipitation Measurement (GPM) Mission for Science and Society. *Bull. Amer. Meteor. Soc.*, **0**, https://doi.org/10.1175/BAMS-D-15-00306.1.

Smalley, M., P. Kirstetter, and T. L'Ecuyer, 2017: How Frequent is Precipitation over the Contiguous United States? Perspectives from Ground-Based and Spaceborne Radars. *J. Hydrometeor.*, **18**, 1657–1672, https://doi.org/10.1175/JHM-D-16-0242.1

Sorooshian, S., A. AghaKouchak, P. Arkin, J. Eylander, E. Foufoula-Georgiou, R. Harmon, J.M. Hendrickx, B. Imam, R. Kuligowski, B. Skahill, and G. Skofronick-Jackson, 2011: Advanced Concepts on Remote Sensing of Precipitation at Multiple Scales. *Bull. Amer. Meteor. Soc.*, **92**, 1353–1357.

Stephens, G.L., and C.D. Kummerow, 2007: The Remote Sensing of Clouds and Precipitation from Space: A Review. *Journal of the Atmospheric Sciences* **64**(11), 3742-3765.

Takahashi, N., H. Hanado and T. Iguchi, 2006: Estimation of path-integrated attenuation and its nonuniformity from TRMM/PR range profile data. *IEEE Transactions on Geoscience and Remote Sensing*, **44**(11), 3276-3283.

Tan, J., W.A. Petersen, P. Kirstetter, and Y. Tian, 2017: Performance of IMERG as a Function of Spatiotemporal Scale. *J. Hydrometeor.*, **18**, 307–319. doi.org/10.1175/JHM-D-16-0174.1

Tang, L., Y. Tian and X. Lin, 2014: Validation of precipitation retrievals over land from satellite based passive microwave sensors. *J. Geophys. Res. Atmos.*, **119**, 4546-4567.

Tian, Y., C. D. Peters-Lidard, B. J. Chaudhury, and M. Garcia, 2007: Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. *J. Hydrometeor*., **8**, 1165–1183.

Turk, F. J., P. Arkin, E. E. Ebert, and M. R. P. Sapiano, 2008: Evaluating high-resolution precipitation products. *Bull. Amer. Meteor. Soc.*, **89**, 1911-1916.

Ushio T, Okamoto K, Kubota T, Hashizume H, Shige S, Noda S, Iida Y, Aonashi K, Inoue T, Oki R, Kachi M, Takahashi N, Iguchi T., 2006: A combined microwave and infrared radiometer approach for a high resolution global precipitation mapping in the GSMAP project Japan, *3rd Int. Precipitation Working Group Workshop on precipitation measurements*, Melbourne, Australia.

Vergara, H., Y. Hong, J.J. Gourley, E.N. Anagnostou, V. Maggioni, D. Stampoulis, and P. Kirstetter, 2014: Effects of Resolution of Satellite-Based Rainfall Estimates on Hydrologic Modeling Skill at Different Scales. *J. Hydrometeor.*, **15**, 593–613, https://doi.org/10.1175/JHM-D-12-0113.1

Weng, F., T. Zhu, and B. Yan, 2007: Satellite data assimilation in numerical weather prediction models. Part II: Uses of rain affected radiances from microwave observations for hurricane vortex analysis. *J. Atmos. Sci.*, **64**, 3910–3925.

Wilheit, T., C.D. Kummerow, R. Ferraro, 2003. Rainfall algorithms for AMSR-E. *IEEE Trans. Geosci. Remote Sensing*, **41**, 204–214.

Wolff, D.B., and B.L. Fisher, 2008: Comparisons of instantaneous TRMM ground validation and satellite rain-rate estimates at different spatial scales. *J. Appl. Meteor. Climatol.*, **47**, 2215–2237.

Wolff, D.B., and B.L. Fisher, 2009: Assessing the Relative Performance of Microwave-Based Satellite Rain-Rate Retrievals Using TRMM Ground Validation Data. *Journal of Applied Meteorology and Climatology*, **48**(6), 1069-1099.

Yang, S., W.S. Olson, J.J. Wang, T.L. Bell, E.A. Smith, and C.D. Kummerow, 2006: Precipitation and Latent Heating Distributions from Satellite Passive Microwave Radiometry. Part II: Evaluation of Estimates Using Independent Data. *Journal of Applied Meteorology and Climatology*, **45**(5), 721-739.

Zeweldi, D.A., and M. Gebremichael, 2009: Sub-daily scale validation of satellite-based high-resolution rainfall products. *Atmospheric Research*, **92**(4), 427-433. doi:10.1016/j.atmosres.2009.01.001.

Zhang, J., Y. QI, K. Howard, C. Langston, and B. Kaney, 2011: Radar Quality Index (RQI) – A combined measure of beam blockage and VPR effects in a national network. *Proceedings, International Symposium on Weather Radar and Hydrology*, IAHS Publ. 351 (2011).

Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation 773 Estimation: Initial Operating Capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621-638.