Can We Estimate the Uncertainty Level of Satellite Long-Term Precipitation Records?

Veljko Petkovic,[a,b] Paula J. Brown,[b] Wesley Berg,[b] David L. Randel[b], Spencer R. Jones[b] and Christian D. Kummerow[b]

[a] *Earth System Science Interdisciplinary Center/Cooperative Institute for Climate and Satellites, University of Maryland, College Park, Maryland*

[b] *Colorado State University, Fort Collins, CO*

Corresponding author: Veljko Petkovic, veljko.petkovic@umd.edu

1

ABSTRACT

Several decades of continuous improvements in satellite precipitation algorithms have resulted in fairly accurate level-2 precipitation products for local-scale applications. Numerous studies have been carried out to quantify random and systematic errors at individual validation sites and regional networks. Understanding uncertainties at larger scales, however, has remained a challenge. Temporal changes in precipitation regional biases, regime morphology, sampling, and observation-vector information content, all play important roles in defining the accuracy of satellite rainfall retrievals. This study considers these contributors to offer a quantitative estimate of uncertainty in recently-produced global precipitation climate data records. Generated from inter-calibrated observations collected by a constellation of Passive Microwave (PMW) radiometers over the course of 30 years, this data record relies on Global Precipitation Measurement (GPM) mission enterprise PMW precipitation retrieval to offer a long-term global monthly precipitation estimates with corresponding uncertainty at $5°$ scales. To address changes in the information content across different constellation members the study develops synthetic datasets from GPM Microwave Imager sensor, while sampling- and morphology-related uncertainties are quantified using GPM's Dual-frequency Precipitation Radar (DPR). Special attention is given to separating precipitation into self-similar states that appear to be consistent across environmental conditions. Results show that the variability of bias patterns can be explained by the relative occurrence of different precipitation states across the regions and used to calculate product's uncertainty. It is found that at $5°$ spatial scale monthly mean precipitation uncertainties in Tropics can exceed 10%.

1. Introduction

Long term, homogeneous climate data records of precipitation are needed to both establish baselines against which future change can be detected (Xie and Arkin, 1997, Adler et al., 2003), as well as for validation of climate models (Tapiador et al. 2017). In Hydrology, long term time series are also essential to establish recurrence frameworks (Knapp et al. 2011). With a changing climate, and the possibility that rainfall statistics may be changing, this becomes particularly important.

Current observational precipitation products may be categorized as climatological products that aim to construct a homogeneous time series at the expense of using all available data (i.e., Climate Data Records - CDRs). The gauge based Global Precipitation Climatology Centre - GPCC (Schneider et al. 2000, 2017) has climatological products spanning over 70 years, but is limited to land, and uses only the subset of rain gauges with a continuous record for the entire time series (e.g., Beck et al. 2005). This reduces the available global gauge network from approximately 100,000 gauges to a mere 10,000 gauges for the 1951-2020 climatology (Schneider et al. 2008). Likewise, satellite products such as Global Precipitation Climatology Project – GPCP (Adler et al. 2003) represent a climatology in that a single passive microwave sensor is used to intercalibrate available IR data for the complete timeseries. This differs from products such as CMORPH (Xie et al. 2019), GSMaP (Okamoto et al. 2005), IMERG (Huffman et al. 2020) and others, that aim to produce precipitation with the highest space-time resolution available, but at the expense of consistency in the long-term time series.

A number of efforts have been undertaken to establish the accuracy of the products listed above by comparing them to each other as well as to the global model reanalysis. It was through one of these efforts (Adler et al. 2012), that GPCP estimated uncertainties of approximately 10% in the global mean precipitation. However, a requisite step in such product intercomparisons, including in the cited work, is the need to eliminate any products that do not converge with the ensemble mean to within a specified value. Masunaga et al. (2019) reported on such an intercomparison of widely used satellite merged products. While they found that mean values were relatively consistent among products, the extremes were very different. In that study, the gauge data had larger disagreements in the extreme than did the different satellite products. This shows that even

3

when means agree, it may not be for the same reasons, and rain gauges have their own set of issues – related largely to the need to interpolate across large distances.

Uncertainties are of course central to any climate time series as all of the products have limitations. Yet, no estimates of uncertainty currently exist beyond the product intercomparisons. This is likely due to the difficulty in quantifying uncertainties from first principles. In this study, the authors use the GPROF algorithm (Kummerow et al. 2015) to try to build an error model for a time series consisting of a number of different satellites – each with different channels, spatial resolution, equator crossing times, and potentially unresolved calibration issues. In principle, the GPROF algorithm can incorporate a diverse range of sensors in a fully parametric way that is essential to translate error characteristics from one sensor to another. The GPM program (Skofronick-Jackson et al. 2017; Hou et al. 2014), has already devoted significant resources to correct for calibration differences between spaceborne radiometers as part of the XCAL intercalibration effort (Berg et al. 2016).

As with any uncertainty, discussion immediately turns to random and systematic errors. Unfortunately, these prove inadequate for the current study. Random errors, which will be labeled as algorithm errors in the subsequent discussion, reduce as the number of satellite-retrieved pixels (i.e., observations) increase. As true random errors, they decrease at $1/\sqrt{n}$ rate, where n is the number of independent pixels/observations. As with any satellite product, n is very large when considering long-term products (e.g., monthly global estimates) and random errors do not contribute significantly to the overall uncertainty. Systematic errors, on the other hand, do not decrease with the number of samples. However, they should be constant for each satellite and easily determined from ground validation studies. Unfortunately, this is not the case, as time and space dependent errors from sources such as time-dependent calibration changes, diurnal sampling, the variability in convective organization, and the sensor information content all affect the satellite product. These error sources have been considered in this study as they are known to strongly link to potential biases, although the authors acknowledge that there may be additional sources of uncertainty. Chapter 3 describes the sensors, algorithms, and time series that constitute a CDR before describing our treatment of the Random, Sampling, Convective fraction, Diurnal cycle, and Information content errors. Chapter 4 lists the results, followed by Discussion and Conclusions in Chapter 5.

## 2.	Data collection and methods

To investigate satellite PMW precipitation CDR uncertainties, this study uses satellite and reanalysis datasets corresponding to the 34-year period – 1987 to 2020. The four data products considered are a) PMW observations from the intercalibrated Level 1C Tb dataset (Berg et al. 2018), b) precipitation estimates from NASA's GPM PMW precipitation retrieval – GPROF (Kummerow et al. 2015), c) precipitation estimates from GPM DPR-combined product – GPM_2BCMB (Olson 2017) and d) the most recent ECMWF reanalysis – ERA5 (C3S, 2017). The four products are interconnected and highly correlated across temporal/spatial scales (Watters et al 2021). PMW satellite data is limited to conical scanning microwave imagers only (see the full list of acronyms in Appendix A).

### a.	Data collection

#### 1)	Intercalibrated Level 1C Data

The Level 1C data containing intercalibrated microwave brightness temperatures relies on observations from 14 conical-scanning microwave imagers launched between 1987 and 2014 (Berg et al. 2018). Development of this data record involved quality control, corrections for cross-track biases, view angle and geolocation errors, emissive reflector issues, solar and lunar intrusions into the warm load, antenna pattern spillover effects, and intercalibration of the following PMW imagers: GMI, TMI, SSM/I (F08, F10, F11, F13, F14, F15), SSMIS (F16, F17, F18, F19), AMSR-

E, and AMSR2. The constellation members timeline is shown in Fig. 1, with sensor-specific channels listed in Table 1.
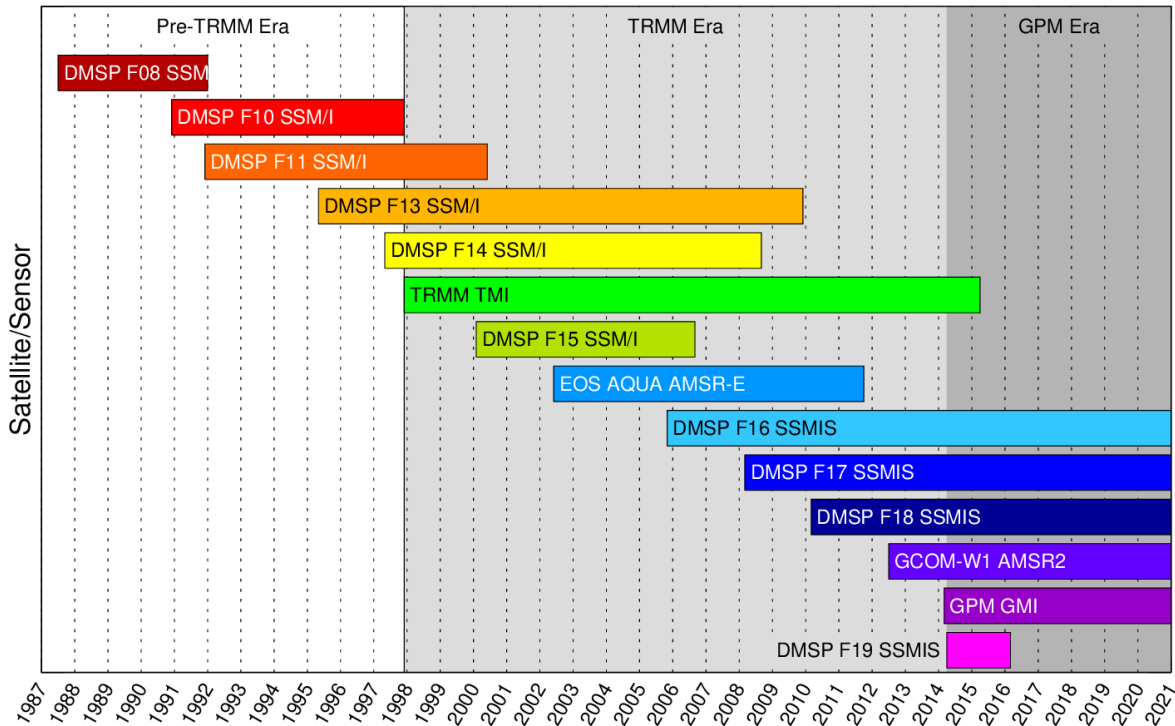

Fig. 1 Timeline of available window channel microwave imagers. The study uses satellite climate precipitation record created using a total of 14 PMW sensors covering pre-TRMM, TRMM and GPM era (1987-2020).

Rigorous quality control has resulted in removal and/or flagging of some instruments observations (details available in Berg et al. 2017). SSM/I-F13 is used as a reference for consistency over the SSM/I-era, while GPM GMI observations are considered to be an absolute calibration reference for the entire Level 1C (i.e., Tbs) dataset. The Level 1C product used in this study is available from the NASA Precipitation Processing System (https://pmm.nasa.gov/data-access/downloads/gpm).

SSM/I		19.35 v/h	22.23 v	37.0 v/h	85.5 v/h		
SSMIS		19.35 v/h	22.23 v	37.0 v/h	91.65 v/h	150.0 h	183+/- 1,3,7 h
TMI	10.65 v/h	19.35 v/h	21.3 v	37.0 v/h	85.5 v/h		
AMSR2, AMSR-E	10.65 v/h	18.7 v/h	23.8 v/h	36.5 v/h	89.0 v/h		
GMI	10.65 v/h	18.7 v/h	23.8 v	36.64 v/h	89.0 v/h	166 v/h	183 +/- 3,7 h

Table 1 PMW instruments and frequencies [GHz] used to generate the long-term precipitation CDR

6

2) PMW precipitation dataset – GPROF algorithm

Developed in the mid-1990s at Goddard Space Flight Center, the Goddard PROFiling (GPROF) precipitation algorithm (Kummerow and Giglio, 1994) has been serving as the operational PMW precipitation retrieval at NASA Precipitation Processing System (PPS) for over three decades. Its fully parametric scheme ensures consistency across a constellation of cross-track (e.g., Kidd et al. 2016) and conical scanning sensors, including the above-mentioned radiometers. An operational version of the algorithm (Version 5 – V05; operational at NASA PPS during 2017-2021 timeframe) is used in this study for production of precipitation estimates over the period of available brightness temperature (i.e., Level 1C) data record. While the evolution of the retrieval is documented in Kummerow et al. (2015), the most important algorithm properties relevant to this study are outlined below.

GPROF is a Bayesian scheme relying on an *a priori* information to establish the relationship between hydrometeor profiles and observed PMW radiances. This *a priori* knowledge is offered through a database of coupled DPR-combined algorithm precipitation profiles (i.e., a state vector) and corresponding computed Tbs (sub-section 2.a.1) following Kummerow et al. (2011). To constrain the problem, the algorithm subsets the *a priori* database using ancillary information on the observed large-scale conditions, namely TPW, surface type, and 2-meter temperature (Berg et al. 2016). In this process, surface type is defined using SSM/I observed emissivity climatology (Aires et al. 2011) updated daily by NOAA's AutoSnow product (Romanov et al. 2000), while the 2-meter temperature and TPW come from reanalysis datasets (e.g., ERA-Interim, Dee et al. 2011 or ERA5, C3S 2017). Once the *a priori* information is identified, the retrieval uses a Bayesian scheme to calculate a weighted mean, where each database element (i.e., DPR-combined precipitation rate) is assigned a weight proportional to its respective probability given by Eq. (1):

$$r_r = \frac{\sum_i r_i w_i}{\sum_i w_i}, \quad where \quad w_i = exp\left\{-0.5 \left[T_b - T_{b_f}(r_i)\right]^T S^{-1} \left[T_b - T_{b_f}(r_i)\right]\right\} \tag{1}$$

where, i is an element of the *a priori* database, S is the T_b error covariance (to account for both forward model and instrument errors), r_r is retrieved rain rate, r_i is database-element rain rate, T_b is the satellite FOV observed-, while $T_{b\_f}(r_i)$ is r_i-associated brightness temperature. Applicable to any PMW sensor, Eq (1) provides consistency of GPROF output and ensures preservation of the global precipitation rate distribution through its *a priori* reference (i.e., DPR-combined product).

7

3) Calibration and Reference Datasets

Launched in February 2014, the GPM mission Core Observatory includes a 13-channel conical-scanning GPM Microwave Imager (GMI) and a Dual-frequency Precipitation Radar. The imager is designed to 1) serve as a calibration standard for the entire GPM PMW constellation, and 2) in synergy with DPR, offer a link between the passive- and active-microwave signatures of the atmospheric column under all weather conditions. The DPR itself, being the most advanced satellite precipitation radar to date, provides invaluable information on vertical profile of global precipitation systems. Data collected over a 5-year period (i.e., March 2014 to February 2019) are used in the present study. The two products of particular interest are DPR-combined level-2 precipitation estimates (GPM_2BCMB; Olson 2017) and GMI Level 1C PMW brightness temperatures (Berg 2016).

4) ECMWF Reanalysis Data – ERA5

ECMWF reanalysis data (ERA) is used at its native spatial and temporal resolution (0.28125º, hourly) in this study to provide: 1) information on environmental conditions and 2) the necessary elements for constructing idealized satellite sampling at various scales. Parameters of particular interest include: 2-meter Temperature (2mT), Sea-Surface Temperature (SST) and precipitation rate.

b. Methodology – Defining and Assigning Uncertainties

Using the intercalibrated brightness temperature (i.e., Level 1C) data (sub-section 2.a.1) as input, GPROF delivers a reasonably consistent and unbiased long-term data record of global precipitation. With a goal to provide a dataset suitable for climate applications, monthly means of precipitation rate are generated at 5° global grids (S65° – N65°). The spatio-temporal scale of the product and the intercalibration of the Level 1C input greatly mitigate random errors of the timeseries. To parse and estimate the remaining uncertainty, we turn our attention to the error sources that contribute to systematic biases in the monthly mean precipitation estimates. These include, although are not limited to, calibration, diurnal sampling, information content, and convective organization. Choosing one of many options, uncertainty (u) of a measurement is described using the standard deviation (i.e., variance, Var(X)) around the mean of the timeseries

8

(i.e., precipitation data record). The variance is calculated as the average of the squared differences from the mean:

$$Var(X) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 \ , \qquad \mu = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad (2)$$

where μ is the mean and x is an individual element of the group with the sample size n. Standard deviation, describing how far from the mean a group of numbers is, is given by the square root of the variance. After estimating each single-source uncertainty (u_j), the assumption of independence between the individual contributors is applied allowing for the combined uncertainty ($U(X)$) to be determined using the following formula:

$$U(X) = \sqrt{\sum_{j=1}^{m} u_j^2} \qquad (3)$$

Below we describe the methodology we used to estimate each of the m individual uncertainties.

1) Uncertainty of Instantaneous Rainfall Estimates - Random Error Uncertainty

Random uncertainty, hereafter referred to as *Bayesian uncertainty*, is estimated by calculating gridded-product standard deviations using instantaneous rate errors and DPR-combined product as a reference. For this purpose, the GPROF algorithm is tasked against the same data that comprises its *a priori* database. This ensures that the bias, or non-representativeness, is not a potential source of error, and that all the errors are truly random. The full sequence of steps consists of the following: *i)* retrieve precipitation against the *a priori* database, *ii)* create monthly global grids of the output and reference data, *iii)* compute the differences across all grid boxes, sampled by surface type and rainfall rate. To preserve precipitation retrieval output dependance on surface type, GPROF surface type maps are resampled to match the output resolution (e.g., 5° x 5°) keeping only 4 main surface type grid boxes: Ocean, Land, Snow and Mixed (a mix of Land and Ocean). To address uncertainty's dependency on the rainfall rate itself, five balanced (i.e., equally populated) rain rate bins are identified: (0, 0.3, 0.9, 1.7, 3.6, inf) mm day^{-1}. Finally, the result provides Bayesian uncertainty as a function of surface type, rain rate and month. The method is repeated for each PMW sensor in the data.

9

2) Sampling-induced Uncertainty

Sampling uncertainty (i.e., observing frequency uncertainty) is defined as a deviation from the true value caused by limited sampling of precipitating systems within a grid box during a time-interval of interest. This uncertainty is driven by the satellite orbit properties and the total number of observations (i.e., samples) available during the sampling period. To estimate sampling uncertainty, modeled precipitation fields are employed in simulating a wide range of different sampling scenarios over a domain of choice (e.g., 5° monthly global grid). The simulations consider an ideal scenario, where a region of interest is sampled continuously, as a reference which is then compared against other sub-sampled cases to estimate uncertainty as a function of month and sampling frequency.

To ensure intra-annual variability does not play an important role, a 5-year GPM-era (2014 - 2018) period of ERA5 precipitation fields, at 1-hour 0.28125° resolution, is used to calculate the sampling uncertainty. The high-resolution hourly data is first gridded into 5° x 5° hourly global grids, for four different surface types (ocean, land, snow and mixed), and split into monthly timeseries with up to 740 individual hourly samples (24 hours for 30 days). These timeseries are then used to derive 1) adjustment factor for any bias caused by a satellite drift, and 2) uncertainty of caused by suboptimal sampling. In the first step, monthly means corresponding to different sampling local times are recorded to be applied to any satellite products that suffer from significant drift in Equatorial crossing time (e.g., DMSP sensor series). In the second step, the timeseries are randomly sampled to simulate sub-sampled monthly records. The process is repeated until each sampling frequency (1 to 740) of the monthly intervals is simulated 100 times. Next, the simulated timeseries are used to calculate monthly means at the 5° x 5° grid. The resulting means are assessed against the reference, providing a distribution (sample size 100) of the differences for each sampling frequency. Finally, using three-month centered data, standard deviations of these distributions define the sampling uncertainty of monthly precipitation product.

3) Diurnal Cycle Uncertainty

Accounting for the variability in sampling frequency across the globe, as described above, provides an estimate of uncertainty induced by smoothing of the precipitation diurnal cycle. However, regardless of how frequent the sampling is, the variability of the diurnal cycle itself also contributes to the overall uncertainty. To assess its contribution to the overall uncertainty, the same

10

GPM-era period of native-resolution hourly ERA5 data is employed once again. This time the hourly data is gridded into monthly 5° x 5° grid boxes for each hour of a month, keeping track of local standard time. For each month in the year, three-month centered means (e.g., DJF for January) are used to generate three products: *i)* monthly mean hourly rates, *ii)* monthly mean daily rates, and *iii)* monthly mean rates. The first product is stored into a 24-element vector, each element holding the mean hourly rate (h_{mean}) for each hour of the day. The second product, the mean daily rate (d_{mean}), holds a single value corresponding to a mean daily precipitation rate calculated for the given month. The last, third product, is a simple monthly mean precipitation rate (m_{mean}). Using these three parameters, the diurnal precipitation anomaly ratio ($h\_anomaly_{ratio}$) is calculated at each hour of the day according to the following equation:

$$h\_anomaly_{ratio} = \frac{h_{mean} - d_{mean}}{m_{mean}} \quad (4)$$

This ratio describes the amplitude of the diurnal cycle of the precipitation field and provides information on how much (on average) a single satellite estimate (for a given local hour) deviates from the true daily mean. This information can be used to remove the mean diurnal cycle bias from the dataset. After removing the mean diurnal cycle bias, however, a residual uncertainty associated with variability in the diurnal cycle remains. This is because the diurnal cycle varies both regionally as well as seasonally or over time. These variations are affected by a large number of factors. While our approach assumes that to first order most of the diurnal cycle variability relates to surface type and precipitation regime (e.g., the west coasts' stratiform systems, MJO, Amazon/Congo convection), this is an assumption that cannot capture the full range of variations in the diurnal cycle due to limited data, errors in the ERA5 reanalysis, etc. In our analysis, the solution of Eq. (4) is used to create diurnal cycle data at each 5° grid box, over the 5-year period from the ERA5 data. With each grid box being characterized by 60 (12 months x 5 years) diurnal cycle sinusoid curves, a k-means clustering method is used to identify grid boxes with similar diurnal precipitation variability regimes. After iterating through the results for a range between 2 to 10 clusters, five clusters were identified as those producing distinctive diurnal patterns (shown in Fig. 2 for the month of August). These five clusters were named based on the local time and amplitude of the anomaly's peak: High Amplitude Early Peak (HAEP), Low Amplitude Late Peak (LALP), High Amplitude Late Peak (HALP), Low Amplitude Early Peak (LAEP) and Medium

11

Amplitude Late Peak (MALP) diurnal types. Fig. *3* offers an insight to the clusters' spatial distribution implying their links to general precipitation regimes.

With each 5° grid box being characterized by one of the five clusters, diurnal cycle uncertainty is defined as the standard deviation of the precipitation ratio anomaly. This uncertainty is calculated as a function of a local hour and cluster using all global three-month centered diurnal cycle sinusoids.
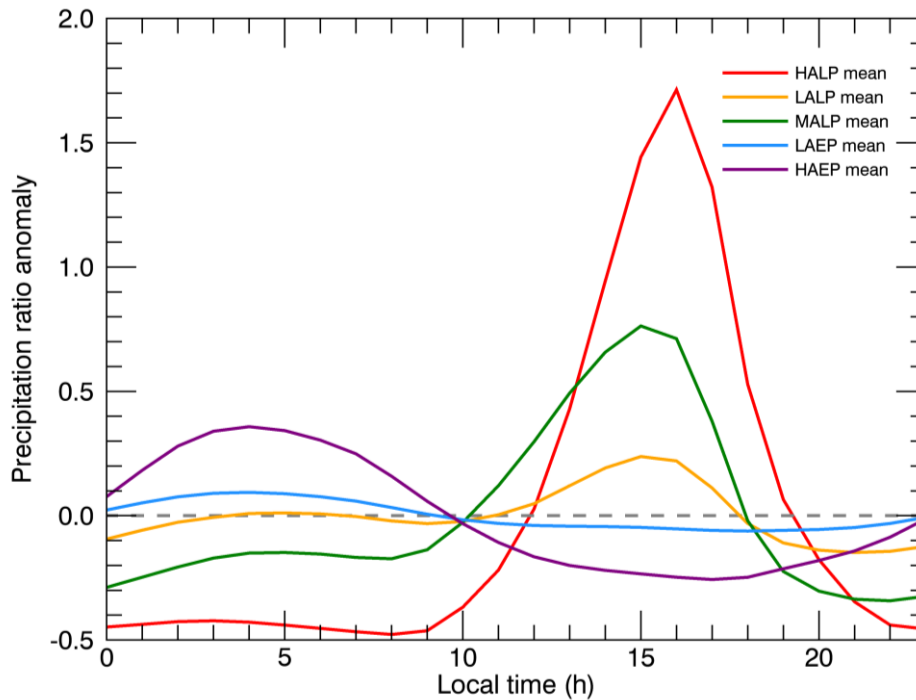

Fig. 2 Monthly averages of precipitation diurnal anomaly (for month of August) characterized by five distinct clusters: High Amplitude Early Peak (HAEP - red), Low Amplitude Late Peak (LALP - yellow), High Amplitude Late Peak (HALP - green), Low Amplitude Early Peak (LAEP - blue) and Medium Amplitude Late Peak (MALP - magenta). Data source: ERA5 hourly product, precipitation fields re-gridded into 5° monthly fields for the period of 5 years (2014-2019).
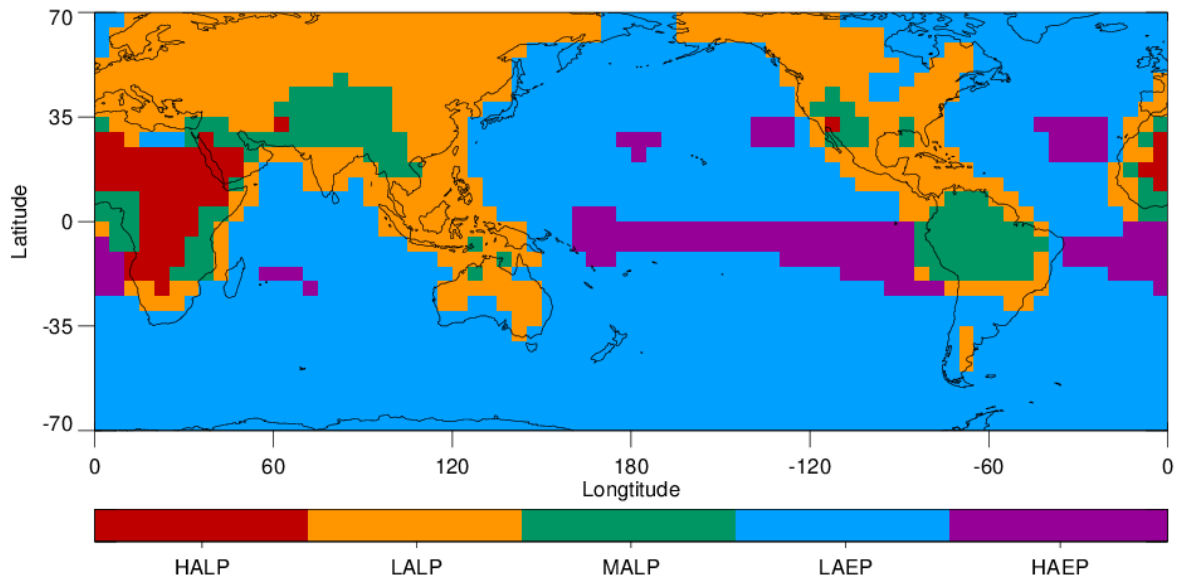
Fig. 3 Map of the 5 clusters shown in Fig. 2 (month of August). High Amplitude Early Peak (HAEP), Low Amplitude Late Peak (LALP), High Amplitude Late Peak (HALP), Low Amplitude Early Peak (LAEP) and Medium Amplitude Late Peak (MALP).

4) Information Content Uncertainty

In remote sensing theory, the available information content defines how strongly (or loosely) the inverse problem can be constrained. PMW observations used for generating precipitation estimates in this study come from sensors with different channels and capabilities resulting in varying information content related to precipitation. In particular, the footprint sizes and the number of observed frequencies and polarizations (i.e., channels) a sensor uses to collect the radiometric signature of the underlying atmospheric column greatly affects the accuracy of the retrieval. Increased resolution and a greater number of observed frequencies typically deliver significantly higher information content, resulting in a more accurate precipitation estimate. To estimate the information content contribution to the uncertainties of the precipitation data record, GMI is once again considered as a reference standard and used to simulate observations from the other sensors, including their channel configuration and sampling geometry. These simulated observations are then used to retrieve precipitation, which is assessed against the precipitation estimates obtained using the reference standard (i.e., GMI). Since the simulated Tb for each sensor is based on the GMI sampling, and thus covers the same spatio-temporal domain as GMI over the 5-year of GPM-era (2014 – 2019), the resulting differences between the simulated products and

13

the GMI-based estimates should come exclusively from the differences in the information content between GMI and the simulated sensor (e.g., SSM/I). The GMI level-1 product (2014 – 2019) is used to simulate a 7-channel SSM/I, a 9-channel AMSR/TMI, and an 11-channel SSMIS (see Table 1) synthetic level-1 data. As the GMI orbit covers only 66°N to 66°S, all simulated datasets in this analysis correspond to same latitudinal region, except for the TMI, which is restricted to the TRMM domain (40°N to 40°S). The FOV resampling process uses the Bachus-Gilbert (1970) approach to (de)convolve GMI Tbs to the resolution and sampling of the corresponding sensor. For each simulated FOV, sensor-specific (de)convolution coefficients are applied to a 11 x 11 pixels patch of GMI-observed Tbs (Fig. 4 left panel). Applying an optimized gain function to account for sensors' channel specifications including the sampling geometry of the sensor, Tbs are computed for the sensor to be considered using the closest available GMI frequency and polarization. The scan-geometry of GMI is thus converted into a pseudo-swath of a simulated sensor that is slightly narrower than the original GMI ground track. An example of GMI-SSM/I re-sampled geometry is given in Fig. 4 (right panel).

The pseudo-swath Tbs are delivered at the same channel frequencies as those of GMI but reflect their sensor's spatial resolution and sampling. Therefore, prior to running the retrieval on the simulated-sensor pseudo-Tbs, the GPROF *a priori* database must be adjusted for any systematic differences between the two sets of Tbs (pseudo vs. GMI-observed ones). To estimate the differences due to the information-content, monthly 5° averaged precipitation estimates for each simulated-sensor are compared to those from GMI. The differences are recorded as a function of year, month, and location (i.e., grid box). For each month in the year, using the three-month centered windows (e.g., DJF for January), the information content uncertainty is defined as standard deviation of the corresponding differences.
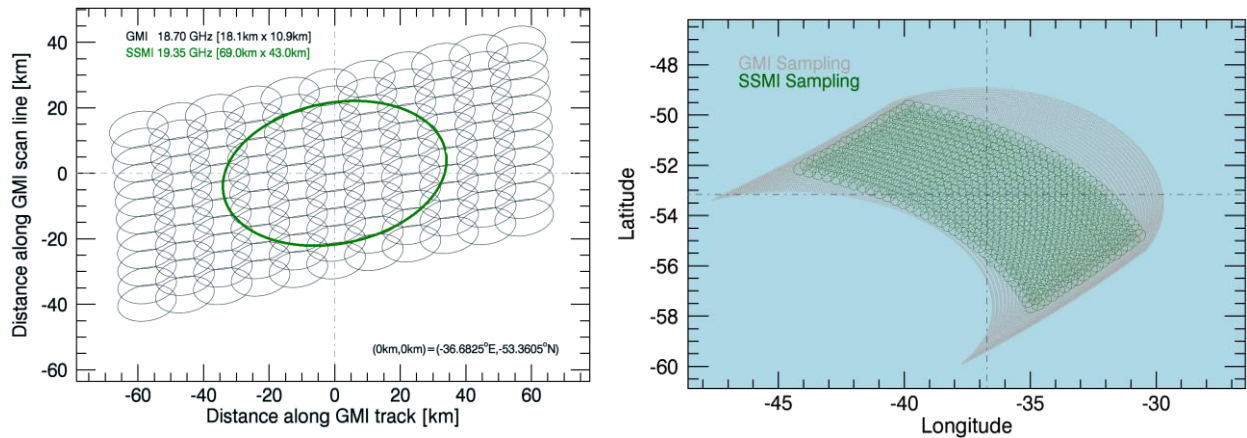
Fig. 4 Spatial distribution of GMI native and SSM/I-simulated scanning pattern. Each SSM/I FOV is derived using a centered 11x11 pixel patch of GMI observations. The non-overlapping portion of the GMI swath was used only for the (de)convolving purposes, but excluded from GPROF runs comparisons when determining information-content variability.

5) Convective Fraction Uncertainty

Another important contribution to the overall uncertainty comes from the precipitation retrieval itself. While the GPROF algorithm (sub-section 2.a.2) is a state-of-the-art enterprise retrieval, it is important to consider the impact of errors in the retrieval and their propagation into spatially- and temporally-averaged estimates (Elsaesser and Kummerow 2015). For the applications considered in this study, many of the retrieval errors make a negligible contribution to the time-averaged large-scale precipitation products. However, the ability of the retrieval to distinguish between radiometrically similar scenes characterized by different precipitation rates, is one that requires attention even when considering 5° monthly products. Due to the equal treatment of distinct hydrometeor profiles in the GPROF algorithm with similar Tbs, an error introduced to the instantaneous precipitation rates becomes a function of precipitation type (Petkovic et al 2017). Given that there is no direct information on precipitation type, a proxy for convective fraction over a grid box is identified based on surface temperature (Fig. 5). This proxy is thus used to estimate the effect of the variability in precipitation systems morphology on the total uncertainty of the retrieval's output.
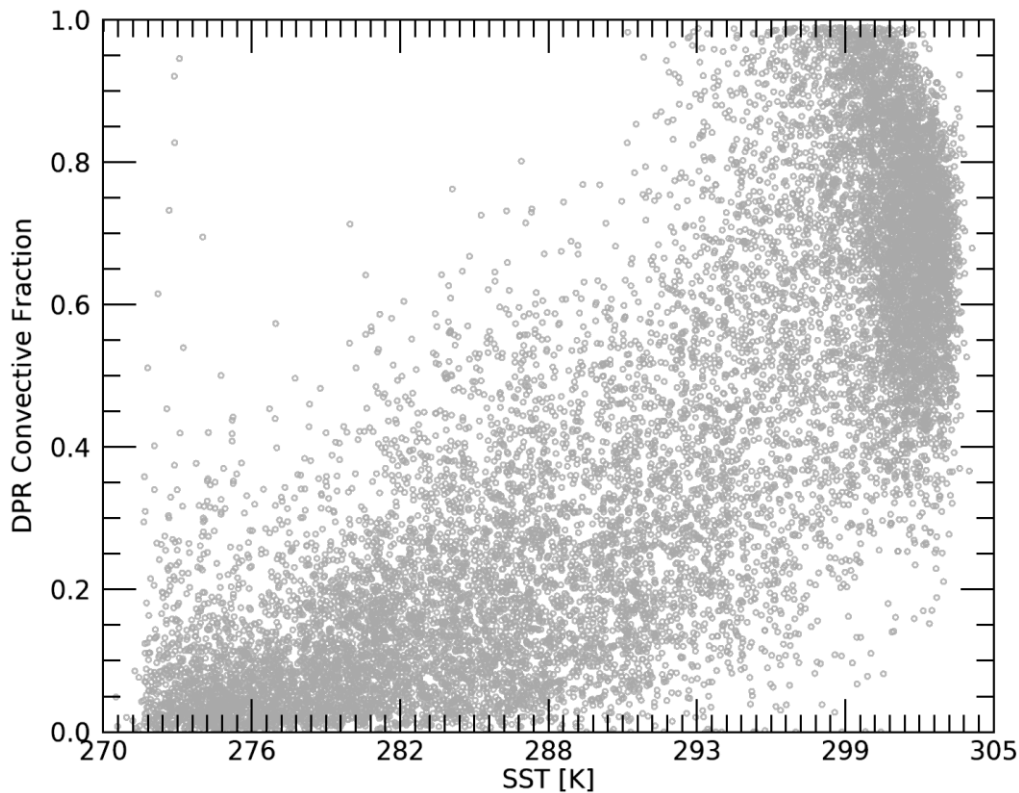
15

Fig. 5 Convective fraction as a function of Sea Surface Temperature. The relationship uses monthly 5° gridded products of DPR global ocean observations and collocated ERA5 Sea Surface Temperatures (SST). Each 5° grid box has at least 50,000 DPR samples.

Using estimates from the combined-DPR product over the same five-year GPM-era period, the relationship between GPROF precipitation biases and the convective ratio proxy (i.e., 2m Temperature from ERA5) is established for each month of the year (Fig. 6) for four different surface types and five rainfall rate bins (as defined in sub-section 2.b.1). Uncertainty induced by varying the convective fraction is estimated as a standard deviation of precipitation bias over the three-month centered periods for each month of the year.
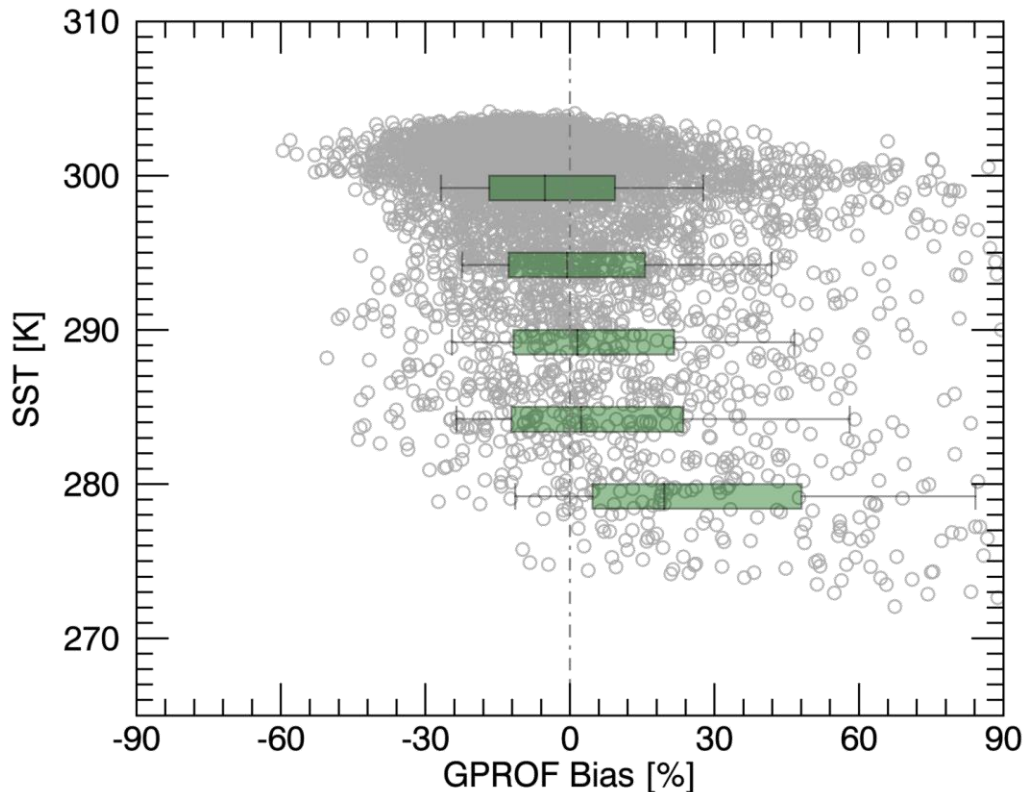
16

Fig. 6. GPROF bias as a function of Sea Surface Temperature – proxy for convective fraction. The relationship of monthly 5° gridded product generated from GPROF output for global oceans and collocated ERA5 SSTs. Box and whisker bars denote 10th, 25th, 50th, 75th and 90th percentile of the bias values (by count), for every 5 K SST.

6) Calibration Uncertainty

Although the Level 1C Tbs are intercalibrated for consistency between sensors (sub-section 2.a.1), residual calibration errors likely remain. A number of issues can impact the calibration and thus the quality of the observed Tb. These include, but are not limited to, permanent or intermittent loss of a channel, an increase in the channel noise or NEDT, changes in the orbit and the local observing time, deviations in the satellite orientation, and changes in antenna and feedhorn properties. To provide an estimate on how these effects may translate to the level-2 products used in this study, a set of synthetic experiments is performed. The GPROF retrieval is run multiple times using one month of SSMIS simulated observations. With each run, a varying amount of noise and/or bias is introduced to the Bayesian error covariance **S** (see Eq. (1)) and the level-1 input Tbs, respectively. The retrieval outputs are then used to provide global mean monthly

17

precipitation rates as function of the Tb bias and sensors sensitivity, providing an upper limit on possible calibration uncertainty contributions.

c. Methodology - Combining the Uncertainties

With the five main contributors to the uncertainty defined as described above, and without accounting for any effects arising from calibration uncertainty causes (sub-section 2.b.6), the total uncertainty of the CDR product is calculated using Eq. (3) but accounting for the number of observations within each grid box (i.e., sample size). The total uncertainty is computed for each 5° grid box in units of mm h^{-1} based on the month of the year, surface type, rain rate, diurnal cycle mode, 2-m temperature and the number of samples during the month. Table 2 lists the individual components of the uncertainty along with the dimensions and the properties used for assigning the uncertainty information to the CDR.

Uncertainty	Dimensions	Dependency
Bayesian/random**	[12, 5, 4]	Month, Rain Rate, Sfc. Type
Sampling freq.	[12, 4, 720]	Month, Sfc. Type, Num. of hourly samples in a month
Diurnal cycle	[12, 5]	Month, Cluster ID
Info. Content**	[12]	Month
Conv. fraction	[12, 11, 5, 4]	Month, 2-m temperature, Rain Rate, Sfc. Type

Table 2 Uncertainty contributors and their dimensions. TPW and 2m-Temperature bins are of 1 mm and 5 K widths, respectively. (**uncertainty calculated for each PMW constellation member separately).

3.   Results (per uncertainty contributor and combined)

The uncertainties for each of the elements listed in Table 2 are presented individually before the merged result is shown.

a. Uncertainty from Random Retrieval Errors

The pixel-level random uncertainty contribution to any monthly-scale product, resulting from the Bayesian estimation described above, is expected to be small. Table 3. lists mean pixel-level Bayesian uncertainties over 5° monthly grids for SSMIS sensor during month of August as a function of Surface Type and Rain Rate. They are taken directly from the retrieval output that reports these errors based on the brightness temperature fit of various rain profiles and the observed Tb. It is notable that uncertainties stay within 0.4 mm h^{-1} for rain rates characteristic to 5° monthly grids (e.g., up to 1 mm h^{-1}). The independent nature of pixel-level random uncertainties and a

18

typically large sample size of observations available to 5° monthly products will yield (see Eq. (3)) a total uncertainty from Random Retrieval Errors to be by far the smallest among all contributors considered in this study. Bayesian uncertainty for other sensors (not shown here) are found to be similar to those of the SSMIS.

Rain rate [mm h^{-1}]	Ocean	Land	Snow	Mixed
(0.00 – 0.01]	0.007	0.007	0.008	0.012
(0.01 – 0.10]	0.044	0.034	0.034	0.055
(0.10 – 1.00]	0.182	0.231	0.286	0.437
(1.00 – 10.0]	0.759	1.349	1.681	1.934
(10.0 – inf]	5.821	13.43	16.806	20.012

Table 3 Mean Bayesian (i.e., random) pixel-level uncertainty for the month of August as a function of the surface type and rainfall rate; 5° monthly scale.

b. Uncertainty Induced by Varying Sampling Frequency

Variability in the sampling frequency of precipitating systems across the globe is driven by the orbit inclination of the individual satellites and the total number of orbiting sensors at a given time. The timeline of sampling frequency over the 30-year period is presented at the end of this section (sub-section 3.g). Fig. *7* shows the sampling-induced uncertainty as a function of the number of hourly samples during the month of August. While an increase in sampling frequency leads to the expected reduction in uncertainty, differences in the amplitude of the sampling uncertainty for different surface types depicts the variability in precipitation rate distribution at hourly-to-monthly scales (e.g., afternoon convection over ocean vs. random, system-specific events over land and costal mixed surface type).
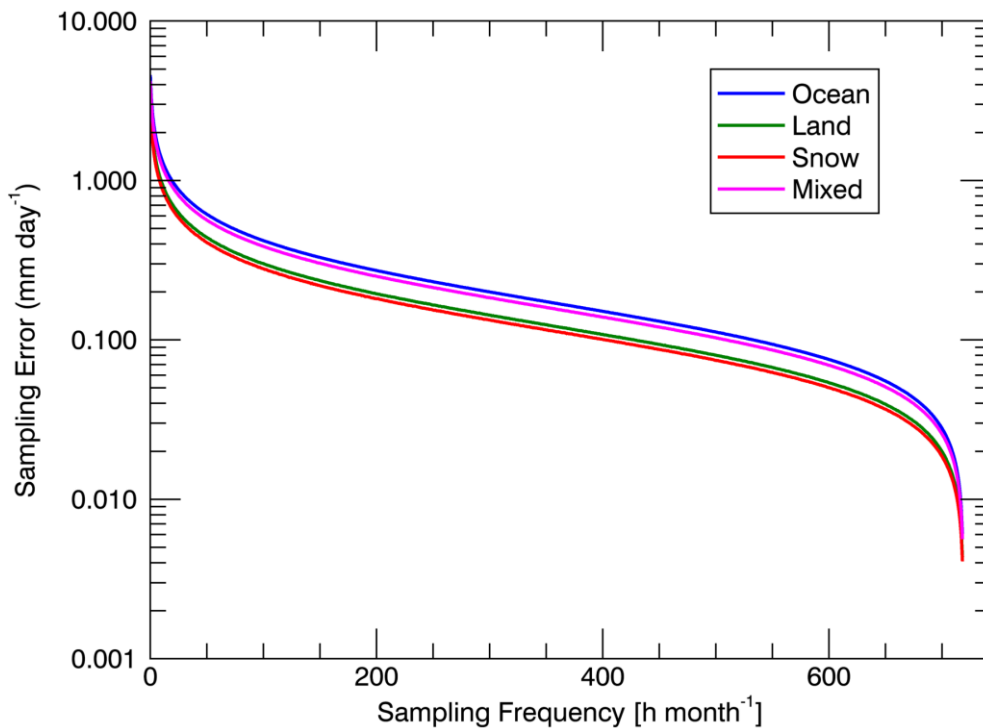
19

Fig. 7 Sampling uncertainty for month of August as a function of the surface type and sampling frequency; 5° gridded product.

c.   Uncertainty Induced by Variations of Diurnal Cycle

For large-scale climate applications, it is important to account for the fact that sun-synchronous sensors do not capture diurnal cycle of precipitation. In this study, monthly precipitation rate estimate is corrected based on the ratio of the mean daily precipitation rate to the mean estimated rate at the local time(s) specific to a given satellites overpass schedule. While this mean correction removes the overall diurnal bias, the variability of the diurnal cycle regionally and over monthly scales both during and prior to the GPM-era remains and can propagate into the spatially and temporally averaged level-3 product. An example of the residual diurnal uncertainty for month of August is shown in Fig. 8. Not surprisingly, when comparing the regimes (i.e., surface types/clusters), the amplitude of diurnal cycle uncertainty is strongly correlated to the total amount of precipitation within a given regime. The example shown in Fig. 8 may be related to an MJO event in the Indian Ocean that has different diurnal characteristics than the background precipitation.
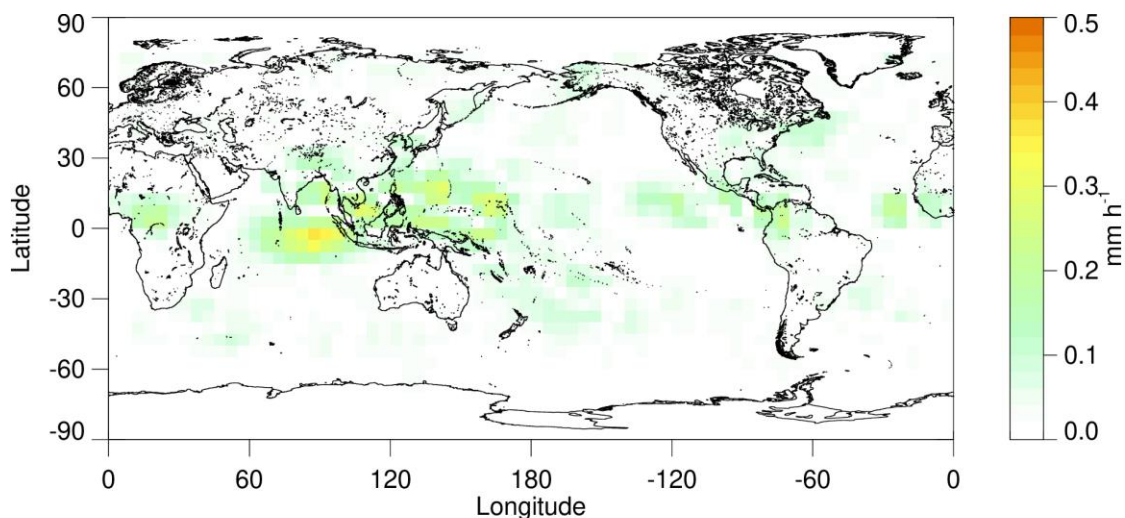
20

Fig. 8 Diurnal cycle variability induced uncertainty for month of August as a function of surface type; 5°
gridded product.

d. *Uncertainty Induced by Variations in Sensor Information Content*

As described previously, the information content uncertainty is estimated using GMI as an absolute reference. Fig. *9* compares the information content uncertainty for each of the constellation sensors relative to GMI. Improvements in spatial resolution and the addition of new channels for more current sensors leads to a decrease in the information content uncertainty for the more recent observations. With similar channels and spatial resolution to GMI, estimates from TMI and AMSR2 exhibit the smallest amount of information content uncertainties relative to GMI. On the other hand, the limited channel availability and low spatial resolution of the SSM/I sensors, made even worse by the loss of the 85 GHz channels on DMSP F08, results in the highest information content errors. Because sensors with lower information content will tend to more closely follow the retrieval's database mean, regional biases are not only possible, but highly likely as the precipitation differs from its statistical average.
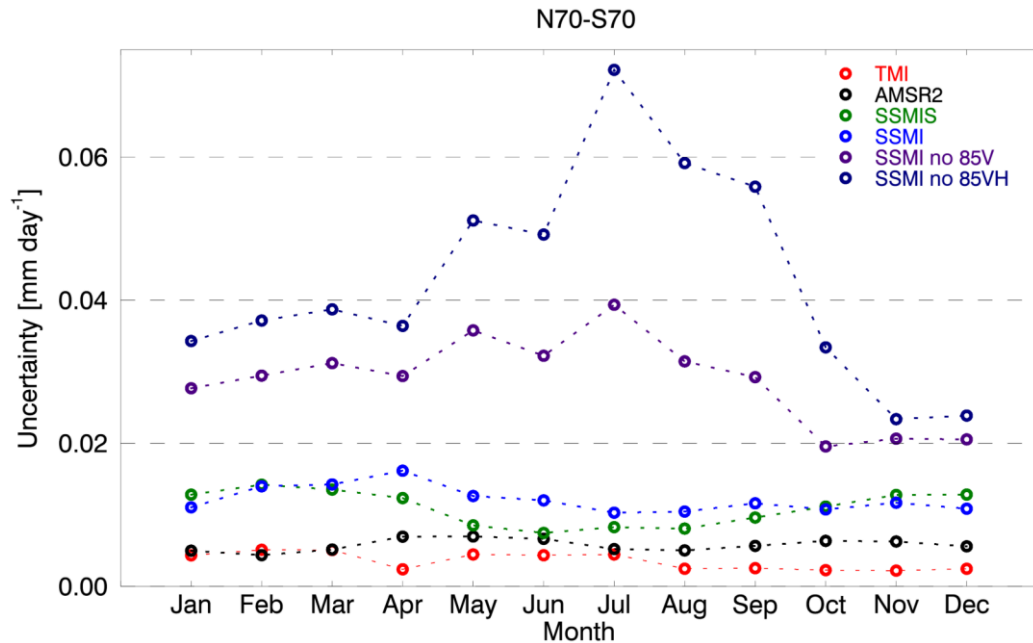
21

Fig. 9 Information content uncertainty for TMI, SSM/I, SSMIS and AMSR2 as a function of month.

e. *Uncertainty Induced by Variations of Convective Fraction of Precipitating Systems*

Another consequence of using limited information content to retrieve precipitation rate, are uncertainties related to the cloud system morphology. Unable to distinguish between precipitation types, retrieval uses a same Tb-to-rain-rate relationship for all storm profiles with similar Tb even when those storms have significantly different surface precipitation rates. Consequently, seasonal, sub-seasonal and regional biases are introduced to instantaneous level-2 precipitation estimates and then propagated to the spatially- and temporally-averaged level-3 products. Using GPM GMI and DPR observations during the GPM-era, the contribution to the total uncertainty due to variations in storm morphology is estimated to be between 0.1–0.4 mm day^{-1}. Table 4 shows the convective fraction uncertainty as a function of month. The large variation in the amplitude of the uncertainty estimates shown here is related to the prevalence of strong convection which is typically associated with high precipitation rate regimes in the global monthly means.

22

Rain rate	Temperature [K]							
[mm day⁻¹]	265	270	275	280	285	290	295	300
(0.00 – 0.31]	-	0.17	0.20	0.18	0.08	0.09	0.12	0.15
(0.31 – 0.87]	0.19	0.29	0.26	0.26	0.26	0.19	0.25	0.32
(0.87 – 1.74]	0.30	0.40	0.36	0.41	0.44	0.35	0.40	0.37
(1.74 – 3.59]	0.42	0.60	0.55	0.62	0.70	0.59	0.66	0.52
(3.59 – inf]	-	-	0.92	0.99	1.12	1.05	1.36	1.26

Table 4 Uncertainty [mm day⁻¹] of long-term 5° girded product induced by variations of Convective Fraction, for combined surface types in month of August, as a function of rain rate and temperature.

f. Estimating Uncertainty Induced by non-optimal Calibration

As explained above, increased sensor noise and/or errors in the calibration can also lead to increased uncertainties in the precipitation estimates. To provide an estimate on the scale of such effects, Table 5 shows what one could expect, in terms of the uncertainty introduced by all unaccounted calibration effects to the brightness temperature Level 1C data used in this study.

Description	Global mean [mm day⁻¹]	Global mean bias [mm day⁻¹]	Ocean mean [mm day⁻¹]	Land mean [mm day⁻¹]
Baseline	2.7382	/	2.9400	2.2464
Low ch. +1K bias	2.8839	0.1457	3.1765	2.1802
Low ch. 1K noise	2.7292	-0.0090	2.9310	2.2351
Low ch. +3K bias	3.2429	0.5047	3.7443	2.0624
Low ch. 3K noise	2.7016	-0.0366	2.8986	2.2132
High ch. +1K bias	2.6933	-0.0449	2.9371	2.1077
High ch. 1K noise	2.7299	-0.0083	2.9394	2.2229
High ch. +3K bias	2.6408	-0.0974	2.9389	1.9355
High ch. 3K noise	2.7082	-0.0300	2.9367	2.1640
All ch. +1K bias	2.8395	0.1013	3.1738	2.0432
All ch. 1K noise	2.7206	-0.0176	2.9304	2.2103
All ch. +3K bias	3.1528	0.4146	3.7469	1.7720
All ch. 3K noise	2.6672	-0.071	2.8947	2.1147

Table 5 Changes in the mean daily global precipitation rate of the SSMIS sensor as a consequence of added noise and bias to the level-1 product (i.e., GPROF input). Values correspond to one-month global means (June 2017). Low channels – less or equal to 37 GHz; High channels – greater or equal to 91 GHz.

g. The Overall Uncertainty of long-term Precipitation CDR

The overall uncertainty is estimated by combining the contributors listed in Table 2. The uncorrelated nature of their origin makes Eq. (3) suitable for calculating the total uncertainty of the precipitation data record. Fig. *10* presents global distribution of individual and combined uncertainties for the month of August 2017 at 5° resolution. With five available conical-scanning radiometers during this month (SSMIS-F17, -F18, -F19, AMSR2 and GMI) the overall uncertainty is dominantly driven by convective fraction variability, followed by contributions from sampling, diurnal, information content, and random errors. The same hierarchy of the contributors' importance is seen throughout the entire data record.
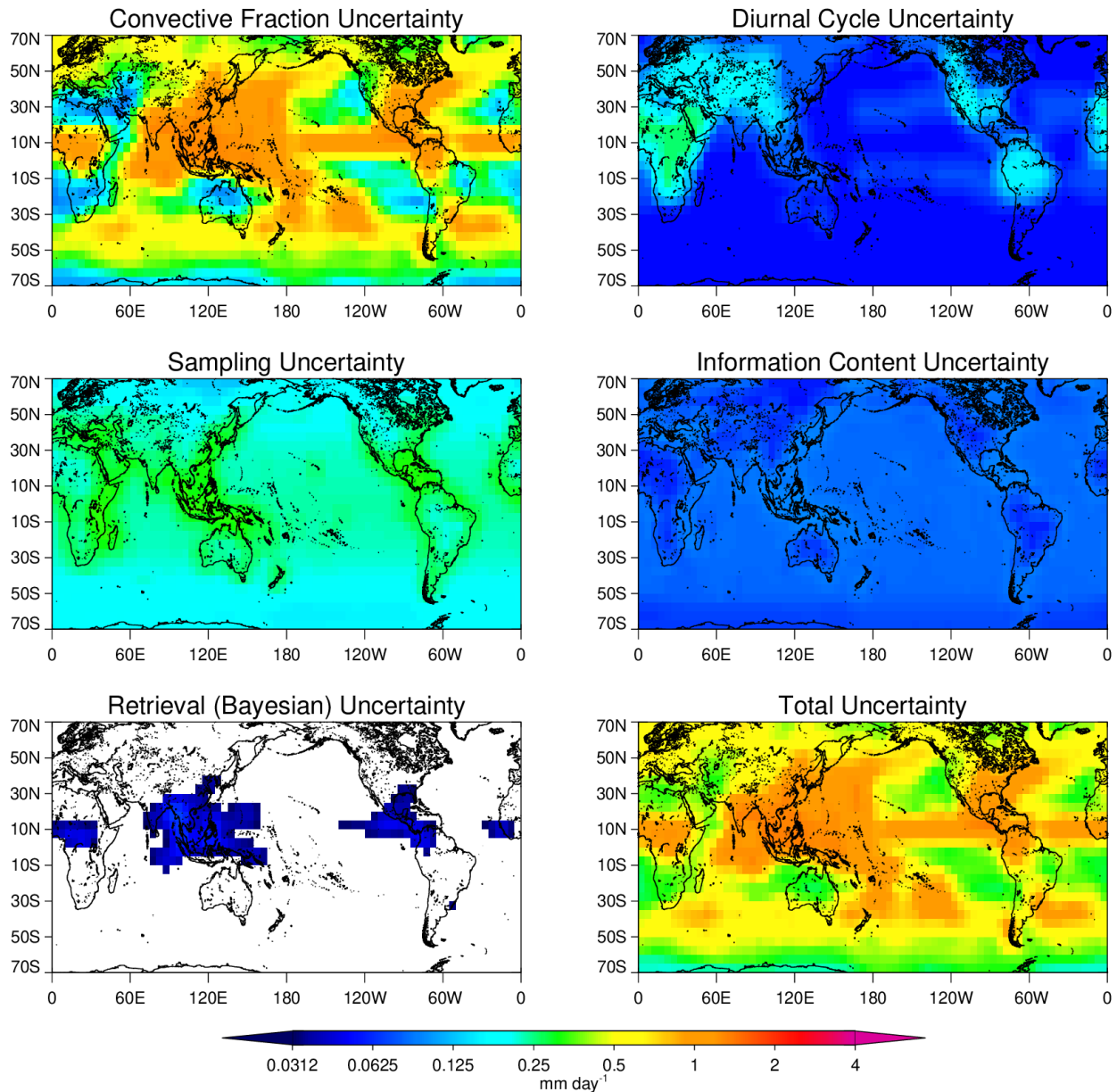
24

Fig. 10 Monthly mean precipitation data record uncertainties for August 2017 at 5° grid. Available sensors: GMI, SSMIS F17, F18, F19 and AMSR2.

To depict the evolution of the uncertainty during the three decades of passive microwave observations, and the effect the introduction of additional sensors has to the precipitation product quality, Fig. 11 presents a timeline of several relevant parameters. Notably, the higher sampling yields lower uncertainties, while a decline in sensor performance, such as that of SSM/I in late 1987, strongly decreases the reliability of precipitation data record. The loss of the 85 GHz channel

on the SSM/I instrument in the 1980's, in combination with low sampling, led to the highest uncertainty in the record. However, the change in the precipitation rate anomaly related to this loss of SSM/I channels significantly exceeds the estimated uncertainty. As the data record evolves, the introduction of TMI in the 1990's, followed by AMSR-E, AMSR2 and GMI in the 2002–2014 period, mitigates the uncertainty sources related to relatively low information content of SSM/I and SSMIS instrument series.

Low but notable trends in global precipitation rate appear to be well outside the uncertainty bars (the blue shade around the mean anomaly shown in the middle panel). However, as discussed in the next section, this is not a sufficient condition to qualify the data record as a global precipitation trend reference.
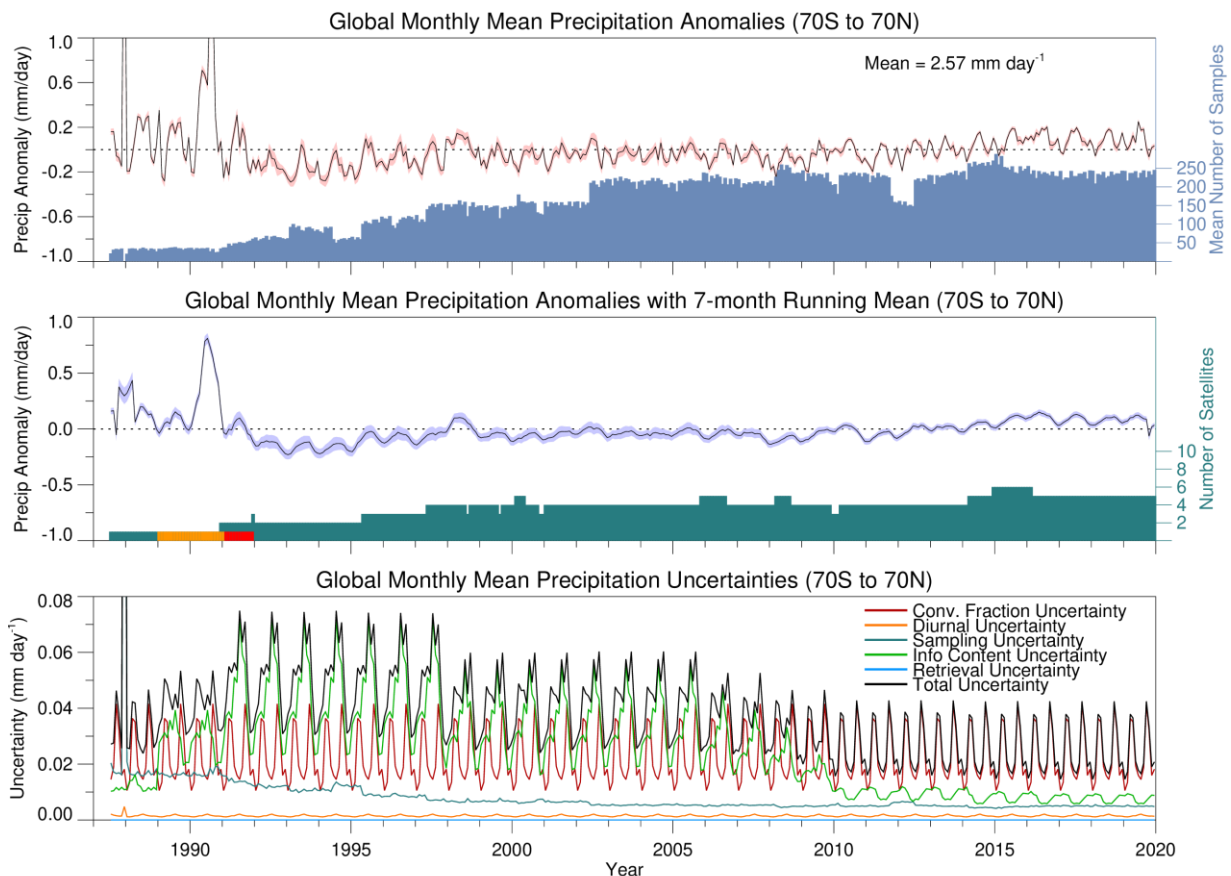

Fig. 11 Timeline of the 34-year global monthly mean precipitation data record (1987-2020) at 5° grid. (Top panel) Precipitation monthly anomalies (solid line), its uncertainty (red shade) and mean number of samples per grid box (blue). (Middle panel) 7-month running mean global monthly precipitation anomalies, the corresponding uncertainty (blue share) and the number of passive microwave instruments contributing to the data record (green). (Bottom panel) Individual contributors and total uncertainty of the CDR precipitation product.

26

4.   Conclusions and discussions

The results presented here consider the main contributors to the uncertainty in a long-term global precipitation record obtained from passive microwave conical-scanning satellite sensors from the GPROF algorithms. While the choices made in identifying and estimating these uncertainties are relatively easy to explain and deliver, combining them into a single uncertainty proved challenging. The challenge is twofold. The first has to do with sensor calibration. Lasting 2 to 5 times their design-defined lifetime, satellite radiometers, just as any other space-born sensors, are exposed to the environment that inevitably affects their performance. Relying on pre-launch calibration and testing, with no physical access to the instrument, leads to potential errors that are not simple to quantify. Simulations of added noise and bias (see sub-section 3.f) suggest that realistic effects of those changes exceed the identified trends in large-scale global precipitation rate. These time-dependent calibration-induced effects are likely responsible for sensors' occasional but large departures in the mean global precipitation rate. As documented in Fig. 12, the sensors from the SSM/I (F08/F10) and SSMIS (F16/F18) series are most prone to this problem. The second issue is related to the assumptions in the algorithm. The assumptions in the GPROF algorithm that were known to impact convective and stratiform precipitation estimates led to the use of convective organization as a predictor of uncertainty during periods where no radars are available to quantify the degree of convective organization. However, it is not only possible, but even likely, that the algorithm produces biased results, although perhaps with smaller magnitudes, in response to other large-scale cloud and precipitation properties. We contend that significant work is still needed to not only validate precipitation products, but to learn how to predict validation results based on large scale environmental conditions that led to the precipitation.

Work remains to be done on both calibration methods and predicting algorithm uncertainties. One important result of this study, however, is that the uncertainties are not simply random or systematic, but have a strong dependence on time and space scales that couple to the source of error itself. The full analysis, requiring an examination of which errors persist and which errors become small due to natural variability, must be undertaken for each space and time resolution under consideration. It is not simply enough to assign an uncertainty to a satellite pixel. Instead, a full elaboration of errors requires each satellite pixel to provide the source of the uncertainty, as

27

well as the time and space correlation of those uncertainties. For now, it may be necessary to provide these uncertainties in a post processing step that considers the particular time-space resolution of the product – be it instantaneous, 5° monthly, or global monthly time series.
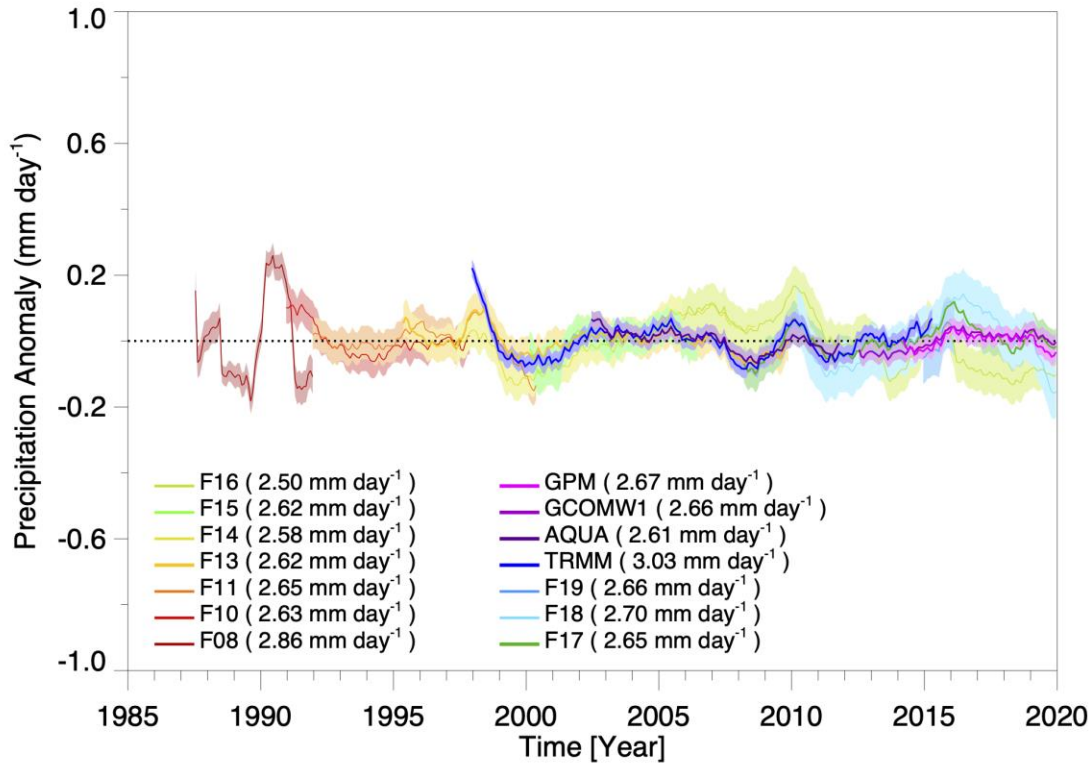

Fig. 12 Global mean precipitation anomalies (70S to 70N) with uncertainties by satellite. Time series is smoothed using a 7-month running mean filter. Shaded color denotes uncertainty of each satellite product.

Acknowledgments.

Data Availability Statement.

Data used in this study can be freely accessed either through the regular public repositories, as cited in the text (e.g., ERA5: 10.24381/cds.bd0915c6; GPM-DPR product: https://storm.pps.eosdis.nasa.gov/storm/), or at Colorado State University ftp, as per user request

(e.g., Level 1C intercalibrated brightness temperature product and level 3 multi-sensor long-term precipitation record).

APPENDIX

AMSR-E	Advanced Microwave Scanning Radiometer - Earth Observing System sensor
AMSR2	Advanced Microwave Scanning Radiometer 2
CDR	Climate Data Records
CMORPH	Climate Prediction Center MORPHing
DMSP	Defense Meteorological Satellite Program
DPR	Dual-frequency Precipitation Radar
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF atmospheric reanalysis
FCDR	Fundamental Climate Data Record
FOV	Field Of View
GMI	GPM Microwave Imager
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GPM	Global Precipitation Measurement
GPROF	Goddard profiling algorithm
GSMaP	Global Satellite Mapping of Precipitation
HAEP	High Amplitude Early Peak
HALP	High Amplitude Late Peak
IMERG	Integrated Multi-satellitE Retrievals for *GPM*
IR	InfraRed
LAEP	Low Amplitude Early Peak
LALP	Low Amplitude Late Peak
MALP	Medium Amplitude Late Peak
MHS	Microwave Humidity Sounder
MJO	Madden–Julian Oscillation
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
PMW	Passive Microwave
PPS	Precipitation Processing System
SSM/I	Special Sensor Microwave/Imager
SSMIS	Special Sensor Microwave Imager/Sounder
SST	Sea Surface Temperature
Tb	Brightness temperature
TMI	TRMM Microwave Imager
TPW	Total Precipitable Water
TRMM	Tropical Rainfall Measuring Mission
XCAL	inter-satellite calibration process
2mT	2-meter Temperature

Table A1. List of acronyms

REFERENCES

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *J. Hydrometeor.*, **4**, 1147–1167, https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Adler, R.F., Wang, J.J., Gu, G. and Huffman, G.J., 2009. A ten-year tropical rainfall climatology based on a composite of TRMM products. *Journal of the Meteorological Society of Japan. Ser. II*, *87*, pp.281-293.

Adler, R. F., Gu, G., & Huffman, G. J. (2012). Estimating Climatological Bias Errors for the Global Precipitation Climatology Project (GPCP), *Journal of Applied Meteorology and Climatology*, *51*(1), 84-99

Aires, F., C. Prigent, F. Bernardo, C. Jiménez, R. Saunders, and P. Brunel, 2011: A Tool to Estimate Land–Surface Emissivities at Microwave frequencies (TELSEM) for use in numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **137**, 690–699, doi:10.1002/qj.803.

Backus, G., and F. Gilbert, 1970: Uniqueness in the inversion of inaccurate gross earth data. *Philos. Trans. Roy. Soc. London,* **266,** 123–192.

Beck, C., J. Grieser and B. Rudolf (2005): A New Monthly Precipitation Climatology for the Global Land Areas for the Period 1951 to 2000. DWD, Klimastatusbericht 2004, 181-190.

Berg, W., T. L'Ecuyer, and C. D. Kummerow, 2006: Rainfall cli- mate regimes: The relationship of regional TRMM rainfall biases to the environment. J. Appl. Meteor. Climatol., 45, 434–454, doi:10.1175/JAM2331.1.

Berg, W., 2016, GPM GMI_R Common Calibrated Brightness Temperatures Collocated L1C 1.5 hours 13 km V05, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: March 2019, 10.5067/GPM/GMI/R/1C/05

Berg, W., Bilanow, S., Chen, R., Datta, S., Draper, D., Ebrahimi, H., Farrar, S., Jones, W.L., Kroodsma, R., McKague, D. and Payne, V., 2016. Intercalibration of the GPM microwave radiometer constellation. *Journal of Atmospheric and Oceanic Technology*, *33*(12), pp.2639-2654.

Berg, W., 2017, July. Towards developing a long-term high-quality intercalibrated TRMM/GPM radiometer dataset. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 248-250). IEEE.

Berg, W., Kroodsma, R., Kummerow, C.D. and McKague, D.S., 2018. Fundamental Climate Data Records of Microwave Brightness Temperatures. *Remote Sensing*, *10*(8), p.1306.

Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), *Jan 2019*. https://cds.climate.copernicus.eu/cdsapp#!/home

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.

Elsaesser, G. S., & Kummerow, C. D. (2015). The Sensitivity of Rainfall Estimation to Error Assumptions in a Bayesian Passive Microwave Retrieval Algorithm, *Journal of Applied Meteorology and Climatology*, *54*(2), 408-422.

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K. and Iguchi, T., 2014. The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), pp.701-722.

Huffman, G.J. and coauthors, 2020: Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). In: Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J. (eds) Satellite Precipitation Measurement. Advances in Global Change Research, vol 67. Springer, Cham. https://doi.org/10.1007/978-3-030-24568-9_19

Knapp, K. R., Ansari, S., Bain, C. L., Bourassa, M. A., Dickinson, M. J., Funk, C., Helms, C. N., Hennon, C. C., Holmes, C. D., Huffman, G. J., Kossin, J. P., Lee, H., Loew, A., & Magnusdottir, G. (2011). Globally Gridded Satellite Observations for Climate Studies, *Bulletin of the American Meteorological Society*, *92*(7), 893-907.

Kummerow, C. D., and L. Giglio, 1994: A passive microwave technique for estimating rainfall and vertical structure information from space. Part I: Algorithm description. J. Appl. Meteor., **33**, 3– 18, doi:10.1175/1520-0450(1994)033,0003:APMTFE.2.0.CO;2.

Kummerow, C. D., S. Ringerud, J. Crook, D. Randel, and W. Berg, 2011: An observationally generated a priori database for microwave rainfall retrievals. J. Atmos. Oceanic Technol., **28**, 113–130, doi:10.1175/2010JTECHA1468.1.

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

Kunkee, D. B., G. Poe, D. Boucher, S. Swadley, Y. Hong, J. Wessel, and E. Uliana, 2008: Design and evaluation of the first Special Sensor Microwave Imager/Sounder. *IEEE Trans. Geosci. Remote Sens.*, **46**, 863–883, doi:10.1109/ TGRS.2008.917980.

Masunaga, H., Schröder, M., Furuzawa, F.A., Kummerow, C., Rustemeier, E. and Schneider, U., 2019. Inter-product biases in global precipitation extremes. *Environmental Research Letters*, *14*(12), p.125016.

Muth, C., W. A. Webb, W. Atwood, and P. Lee, 2005: Advanced technology microwave sounder on the National Polar- Orbiting Operational Environmental Satellite System. Proc. IEEE Int. Conf. on Geoscience and Remote Sensing Symp., Seoul, South Korea, Institute of Electrical and Electronics Engineers, 99–103, doi:10.1109/IGARSS.2005.1526113.

Olson, William, 2017, GPM DPR and GMI Combined Precipitation L2B 1.5 hours 5 km V05, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: March 2019, 10.5067/GPM/DPRGMI/CMB/2B/05

Okamoto, K., T. Iguchi, N. Takahashi, K. Iwanami and T. Ushio, 2005: The global satellite mapping of precipitation (GSMaP) project, 25th IGARSS Proceedings, pp. 3414-3416.

Romanov, P., G. Gutman, and I. Csiszar, 2000: Automated monitoring of snow cover over North America with multispectral satellite data. *J. Appl. Meteor.*, **39**, 1866–1880, doi:10.1175/ 1520-0450(2000)039,1866:AMOSCO.2.0.CO;2.

Schneider, U., Ziese, M., Meyer-Christoffer, A., Finger, P., Rustemeier, E., and Becker, A., 2000: The new portfolio of global precipitation data products of the Global Precipitation Climatology Centre suitable to assess and quantify the global water cycle and resources, Proc. IAHS, **374**, 29–34, https://doi.org/10.5194/piahs-374-29-2016, 2016.

Schneider, U., Fuchs, T., Meyer-Christoffer, A. and Rudolf, B., 2008. Global precipitation analysis products of the GPCC. *Global Precipitation Climatology Centre (GPCC), DWD, Internet Publikation*, *112*.

Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M. and Becker, A., 2017. Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere*, *8*(3), p.52.

Shimoda, H., 2005: GCOM missions. Proc. IEEE Int. Conf. on Geoscience and Remote Sensing Symp., Seoul, South Korea, Institute of Electrical and Electronics Engineers, 4201–4204, doi:10.1109/IGARSS.2005.1525844.

Skofronick-Jackson, G., Petersen, W.A., Berg, W., Kidd, C., Stocker, E.F., Kirschbaum, D.B., Kakar, R., Braun, S.A., Huffman, G.J., Iguchi, T. and Kirstetter, P.E., 2017. The Global Precipitation Measurement (GPM) mission for science and society. *Bulletin of the American Meteorological Society*, *98*(8), pp.1679-1695.

Tapiador, F.J., A. Navarro, V. Levizzani, E. García-Ortega, G.J. Huffman, C. Kidd, P.A. Kucera, C.D. Kummerow, H. Masunaga, W.A. Petersen, R. Roca, J.-L. Sánchez, W.-K. Tao, F.J. Turk, 2017: Global precipitation measurements for validating climate models, *Atmospheric Research*, *197*, 1-20, ISSN 0169-8095, https://doi.org/10.1016/j.atmosres.2017.06.021.

Xie and Arkin, 1997: Global Precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bulletin of the American Meteorological Society*, *78*, 2539-2558.

Xie, Pingping; Joyce, Robert; Wu, Shaorong; Yoo, S.-H.; Yarosh, Yelena; Sun, Fengying; Lin, Roger, NOAA CDR Program (2019): NOAA Climate Data Record (CDR) of CPC Morphing Technique (CMORPH) High Resolution Global Precipitation Estimates, Version 1. NOAA National Centers for Environmental Information. https://doi.org/10.25921/w9va-q159

Watters, D., A. Battaglia, and R. P. Allan, 2021: The Diurnal Cycle of Precipitation according to Multiple Decades of Global Satellite Observations, Three CMIP6 Models, and the ECMWF Reanalysis. *J. Climate*, *34*, 5063–5080, https://doi.org/10.1175/JCLI-D-20-0966.1.