1.2. Validating the intrinsic uncertainty: Implications for hydrologic applications

Pierre-Emmanuel Kirstetter[1]

[1]University of Oklahoma, Norman, OK, USA

Reliable quantitative information on the spatial distribution of precipitation is essential for hydrologic and climatic applications that range from real-time hydrologic hazards forecasting (e.g. floods, droughts, landslides), to water resources and urban drainage management and agriculture, to diagnosing hydroclimate patterns and trends, to evaluating regional and global atmospheric model simulations. Physical processes associated with these applications cover multiple scales, from minutes to decades and from metres to the synoptic scale. The critical importance of accurate water flux estimates for applications explains the large body of verification analyses focusing on precipitation estimates, in terms of occurrence, average and extremes. An abundance of independent validation has been carried out directly on Level-3 products using gauges and sometimes ground radar data from various over-land locations. Very few are implemented at the relevant scales to address the intrinsic uncertainty of precipitation products. Without relevant information on key uncertainty features, applications making use of satellite Level-3 precipitation products are impacted both in terms of outcomes and physical realism.

1.2.1. Benchmarks for satellite precipitation: Sensors

Accurately measuring rainfall has been a challenge for the research community predominantly because of its high variability in space and time. There are primarily three major types of techniques of precipitation measurement: (1) surface-based rain gauge, (2) weather radar and (3) space-based meteorological satellites.

Among precipitation sensors, only the rain gauge directly measures precipitation rates or time accumulations. Rain gauges collect rainfall directly in a small orifice and measure the water depth, weight or volume. Rain gauges provide quite reliable point measurements of precipitation and records frequently span more than 100 years. These are therefore the best source for long-term studies of precipitation extremes and trends. The global distribution of gauges is heterogeneous, with higher densities in more populated regions and lower densities in rural and remote areas. Critically, the number of gauges available also depends on their temporal sampling resolution, with stations sampling at finer scales being rarer. Gauges are routinely used to represent areas of 100 to 3000 m^2 from measurements taken over a few square centimeters. However, their measurements are affected by uncertainties (for example, wind undercatch, evaporation, snow) and lack areal representation, which becomes particularly problematic for intense rainfall with high spatial variability (for example, Zawadzki, 1975). The spatial representativeness of each gauge measurement depends on the autocorrelation distance of precipitation (for example, Delahaye et al., 2015). While the autocorrelation increases with time integration, it varies greatly with precipitation regime and is typically short for extreme events (for example, Lebel et al., 1987). Interpolation of rain gauge observations is mandatory to obtain spatial information. When it comes to comparing gauges with other area-averaged precipitation estimates such as from radars or satellites, the spatial variability of rainfall at small scales and the large resolution difference (as much as nine orders of magnitude in area) may cause large differences in the statistical sampling properties of the extremely variable rainfall process (for example, Habib et al., 2004). The added statistical noise when comparing the two measurements is especially significant for short accumulation periods (1 hour or less; Ciach and Krajewski, 1999).

DOI: 10.13021/gewex.precip.1.2

Precipitation is associated with specific generating processes, such as convection, orographic enhancement in complex terrain or warm rain processes. Measuring variations in the drop size distribution and the vertical structure of events is essential for understanding precipitation processes but cannot be captured by rain gauges. Remote sensing is the only way to explicitly observe the spatial distribution of precipitation. However, complex interactions between the spatiotemporal variability of precipitation processes, sensor resolution, sensitivity, calibration and the indirect nature of precipitation retrievals introduce complications (Section 1.1). In the last decades, weather radar systems have become a valuable tool to fill multiple observational gaps in time, surface 2D and 3D. As active sensors, ground-based radars provide range-resolved information on precipitation that is not available from most satellite sensors. Radar systems reveal precipitation characteristics, including intermittency, types (for example, stratiform, convective, snow and hail) and rates, with better resolution and accuracy than gauges and satellites, respectively. Through real-time and high-resolution volume scanning, weather radars offer more comprehensive information on the horizontal and vertical structure of rainfall. Radar networks upgraded with dual-polarization technology give additional insights into precipitation microphysics specifically on the size, shape, orientation and phase of hydrometeors. Ground-based weather radar data are now widely used by national weather services for quantitative precipitation estimation (QPE) at fine scales (for example, 1 km/5 min). Radar QPE is subject to specific uncertainties (that is, sensor calibration, attenuation depending on the radar frequency, ground clutter and beam blocking, variation of reflectivity with height, conversion from radar moments to precipitation rate, etc.; for example, Delrieu et al*.,* 2009; Villarini and Krajewski*,* 2010; Berne and Krajewski*,* 2013). The characterization of these uncertainties has motivated studies for several decades. Radar–rain gauge merging approaches combining the fine spatio-temporal resolution of radar and the local accuracy of gauges have been proposed for QPE (for example, Delrieu et al., 2014) and are applied operationally (for example, Zhang et al*.,* 2016), while novel approaches are being developed to integrate uncertainty as part of the quantitative estimation process (for example, Kirstetter et al., 2015; Neuper and Ehret*,* 2019).

The last decade has witnessed the growing use of satellite-based observations for seamless observation of precipitation over land and oceans, with quasi-global coverage that is not available with radar or gauge networks (for example, Skofronick-Jackson et al*.,* 2017). As shown in Section 1.1, most instantaneous-level spaceborne precipitation observations are performed with passive MW sensors, providing more indirect observations of surface rainfall amounts than radars. Many multi-sensor precipitation retrievals combine IR and passive MW data to produce near-real time estimates at high spatial and temporal resolution (for example, 30 min, 0.1°). A description and an intercomparison of current global precipitation datasets from stations and satellites can be found in Sun et al. (2017).

Sensor limitations discussed in this section are listed in Section 2.3.1. In order to overcome these limitations, it is crucial to recognize that no single sensor combines accuracy, resolution and representativeness over relevant spatial and temporal scales, which are essential characteristics for applications making use of precipitation inputs. Achieving these characteristics requires the expert combination of observations that maximize each sensor’s advantages while minimizing its weaknesses. Ultimately, such a combination does not produce perfects estimates with no uncertainty, but estimates with uncertainties that are deemed sufficiently low.

1.2.2. Benchmarks for satellite precipitation: Requirements for quantifying intrinsic uncertainty

Because precipitation displays variability at all scales, the satellite's intrinsic uncertainty can only be assessed at the primary scale of the precipitation retrievals. Preserving the product's intrinsic characteristics precludes any scale alteration such as interpolation, averaging, smoothing or oversampling, which affects key characteristics such as the retrieved rainfall amount, the rainy area and the distribution of precipitation rates. The true precipitation averaged over the spatial domains and time intervals corresponding to the primary scale of satellite precipitation retrievals is unknown. A reference precipitation used as a proxy for the true precipitation and as a benchmark should spatially and temporally match the satellite retrieval domain and display acceptable levels of accuracy.

Ground sensors constitute a natural choice to create a benchmark because their measurements are more directly sensitive to surface precipitation than satellite sensors. A trustworthy surface reference rainfall dataset should combine the complementary qualities of ground-based sensors, specifically the local accuracy of gauges and the spatial and temporal resolution provided by radars. An example of a satellite precipitation benchmark is the Ground Validation Multi-Radar/Multi-Sensor (GV-MRMS; Kirstetter et al., 2018b) that is derived from the MRMS system (Zhang et al., 2016). MRMS incorporates observations from all polarimetric Weather Surveillance Radar, 1988, Doppler (WSR-88D) radars and from gauge networks in the conterminous U.S. and creates a seamless 3D radar mosaic. Automatic quality controls and corrections procedures mitigate radar uncertainties (section 1.2.1) and generate high-resolution mosaicked radar-based surface precipitation products at a 0.01° horizontal resolution and 2 minute update cycles. Dual-polarization improves the radar data quality and enables the identification of hydrometeors where the ground-radar estimates are the most reliable. The radar-based data are integrated with atmospheric environmental data and rain gauge observations to generate a suite of severe weather and quantitative precipitation estimation (QPE) products. A surface reference precipitation framework is derived from MRMS to support the GPM mission for ground validation and intercompare satellite sensors (Kirstetter et al., 2014; Petersen et al., 2020) and to validate Level-3 precipitation products (Gebregiorgis et al., 2017; Tan et al., 2017). It applies conservative adjustments, quality controls and quantity controls on MRMS products to refine the most trustworthy radar-gauge precipitation estimates towards specific satellite purposes and needs. This processing is designed to maximize accuracy, minimize uncertainties and standardize the GV-MRMS precipitation reference products across the Continental United States (CONUS).

Thanks to their resolution, which is higher than any satellite precipitation product, GV-MRMS data are designed to be pixel-matched in both time and space, and to build statistics for comparing reference precipitation intensities to Level-2 and Level-3 satellite-based estimates. Note that no reference perfectly matches the true precipitation; however, eliminating systematic error sources and non-robust reference values is necessary to improve confidence in the reference precipitation. The reference data covers a broad range of land surface types (mountains, coasts, plains) and precipitation regimes and captures a variety of situations to document representative features of satellite intrinsic uncertainty.

An extended characterization of the reference precipitation should include additional key precipitation properties such as typology. The typology of rainfall can be assessed within the satellite sensor's field of view (FOV) (for Level-2 products) or pixel grid (for Level-3 products) through precipitation properties such as the Convective Percent Index (CPI). CPI quantifies the volume contribution of convective rainfall to the reference precipitation (Kirstetter et al., 2020).

The CPI is expressed in percent between 0% (purely stratiform rainfall within the FOV or pixel) to 100% (purely convective rainfall). CPI values between 0% and 100% indicate mixed precipitation types (Figure 1.2.1).

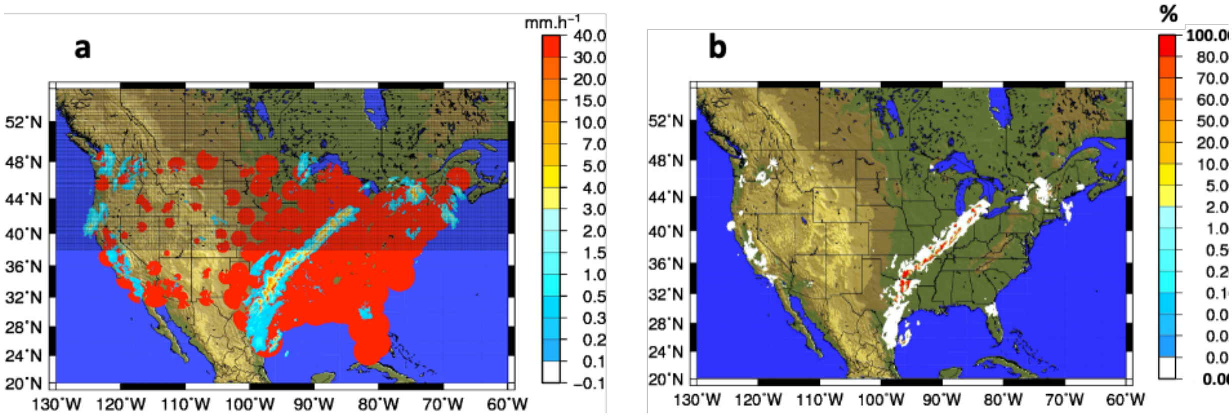

Figure 1.2.1. (a) Map of CONUS area with GV-MRMS instantaneous rain rates at 0725 UTC on 11 April 2011. The red area shows the good quality radar coverage; (b) the convective percent index (CPI).

1.2.3. Non-homogeneity of gridded satellite precipitation products: Implication for their hydrologic assessment

It is essential to recognize that gridded Level-3 satellite precipitation products are not homogeneous because of the dynamical interplay between a variety of error sources described in previous sections. Precipitation characteristics such as convection are a challenge for satellite retrievals, although convective precipitation is a strong driver of extremes. They condition systematic biases at all levels (for example, see Figure 1.2.2). The estimation error varies also depending on which sensor is weighted more in the retrieval. For example, estimates originating from IR display different error patterns from passive MW (for example, see Figure 1.2.2d). Consistency and homogeneity are properties that any Level-3 satellite precipitation product is designed to achieve, but these properties are often overlooked in assessment exercises. It follows that a gap remains with our ability to consistently merge precipitation estimates into gridded products and assess the procedure.

An endemic limitation in the extensive body of literature on satellite precipitation validation and error modeling is that the satellite product is implicitly assumed to be consistent and homogeneous over the spatial and temporal domain of comparison. This is rarely the case because comparison samples gather a variety of precipitation characteristics (for example, intermittency, typology, rates) for which the satellite algorithm (or combination of algorithms for Level-3 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 passive MW/IR retrievals are underconstrained by nature and sensitive to unobserved atmospheric parameters (Stephens and Kummerow, 2007). The combined products inherit the varying passive MW/IR performances and create additional uncertainties with temporal/spatial resampling.

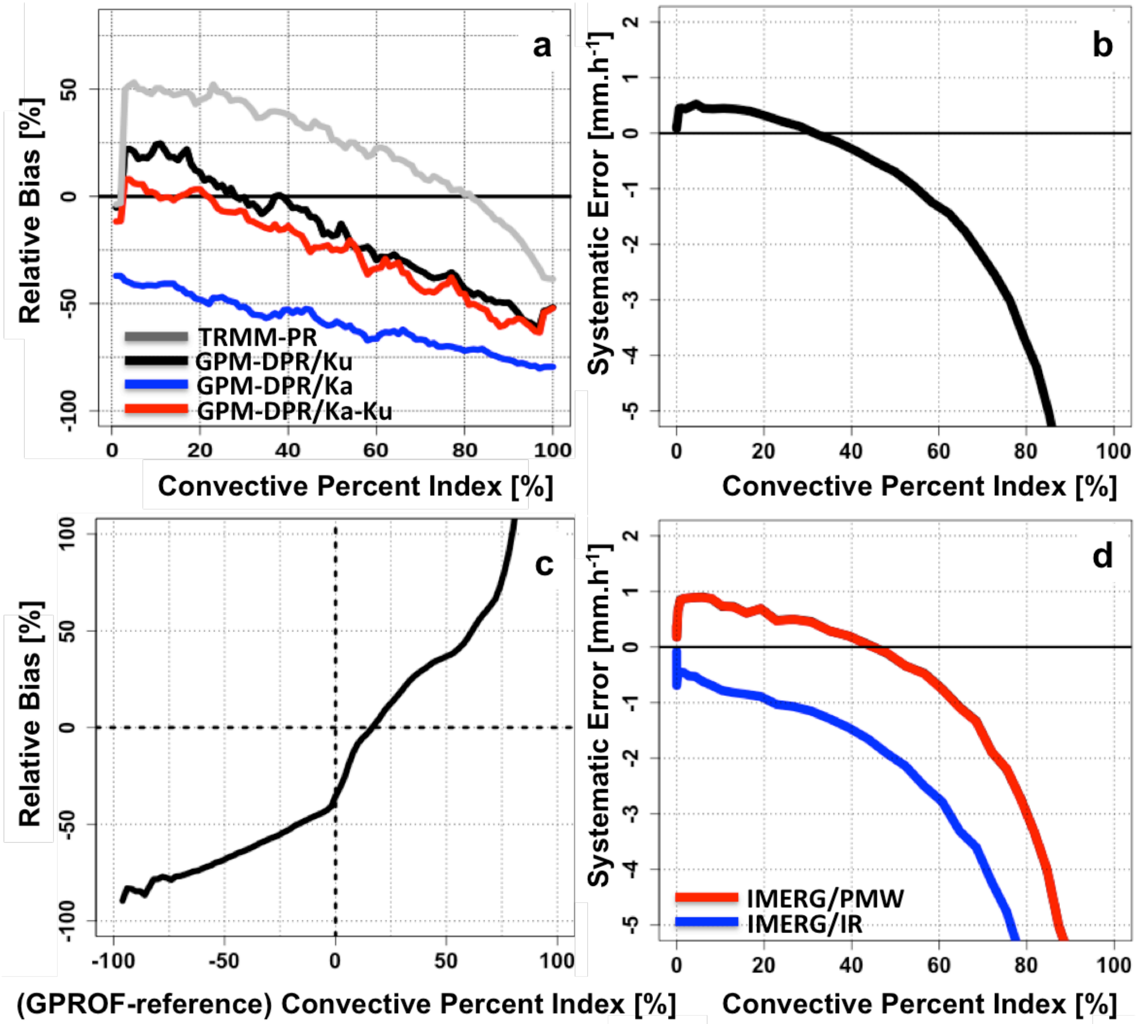

Figure 1.2.2. Performances of space-based QPE as functions of Convective Percent Index with respect to GV-MRMS: (a) space-based radars' relative bias for TRMM-PR (grey), DPR/Ku (black), DPR/Ka (blue) and DPR/Ka-Ku (red); (b) GPROF-GMI systematic error; (c) GPROF-GMI relative bias as a function of CPI difference with GV-MRMS; and (d) IMERG systematic error for the passive MW (red) and IR (blue) components. Convective and stratiform situations correspond to CPI=100% and CPI=0%, respectively. Comparison data include 2M+ matched ground-satellite pairs from June 2014 to September 2016 (from Kirstetter et al., 2020).

Common assessment typically uses bulk comparison metrics (for example, probability of detection, correlation, bias) that depict averaged space/time properties while the errors tend to be non-stationary and sensitive to parameters not accounted for in the assessment formulation. These metrics are sometimes applied without necessarily checking their relevance (for example, the linear correlation is generally insufficient to describe the non-linear and heteroscedastic dependence structure between a satellite precipitation estimate and the precipitation reference). Hence validation practice generally provides limited insight in the complex error characteristics of satellite precipitation estimates.

In addition, 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 (for example, resolution), the retrieval algorithm, the space-time-scale and the accuracy of the reference, and has limited applicability for other precipitation regimes, regions, products, etc. The actual benefit of these analyses to satellite precipitation users and developers is limited. There is a need to formulate the goals of validation and error modeling, and to design appropriate comparison practices.

Integrated assessment of the intrinsic uncertainty across multiple sensors and products is necessary to track the origin of errors and their propagation through various Level-2 active to passive MW to Level-3 merged satellite precipitation estimates. Targeting the most significant factors driving the state of the satellite estimation error (for example, precipitation types) is essential to characterize uncertainties in satellite QPE and lead to a generalization of their assessment (Kirstetter et al., 2020; Shige et al., 2013; Taniguchi et al., 2013; Yamamoto et al., 2017). Figure 1.2.2 illustrates the propagation of uncertainty that arises from precipitation types in the form of systematic biases from spaceborne radars (for example, GPM DPR) through MW precipitation estimates (GPROF-GMI) to the IMERG Level-3 merged product.

1.2.4. Impact of satellite precipitation intrinsic uncertainty on hydrologic applications

Hydrologic applications of satellite data include agriculture, freshwater availability and natural disasters monitoring (for example, floods, droughts, landslides; Serrat-Capdevila et al., 2014). Each application is characterized by specific spatial and temporal requirements that can vary significantly. For example, global hydrological modeling to assess the occurrence of flood events is uniquely enabled with the coverage provided by Level-3 precipitation products (such as the Global Flood Monitoring System, http://flood.umd.edu/; Wu et al., 2014). Anticipating flood events enables the assessment of associated risks and optimized decision making, specifically in developing countries (Kirschbaum et al., 2017). The detection of floods and inundations is critical for hazard response by agencies such as the United Nations World Food Program and the International Federation of Red Cross and Red Crescent Societies (Gray, 2015). On the other side of the precipitation spectrum, precipitation deficits are monitored with satellites as drivers of drought and food and water security (for example, the Famine Early Warning Systems Network, FEWS NET; www.fews.net). Water resources applications are reservoir operations that use precipitation products at monthly time scales (for example, Yang et al., 2017). However, applications of Level-3 products have not been demonstrated yet in contexts involving hydrologic processes over short scales (for example, a few kilometres, hourly), such as flash-flood monitoring, because the Level-3 resolution or latency are limiting factors.

Hydrologic applications often require an understanding of the error structure in the satellite precipitation products. Errors in hydrologic simulations result from a complex interaction between the forcing uncertainty (that is, precipitation), the model structure and approximations, the estimation of model parameters, and observations (for example, gauged streamflow). Hence precipitation errors and uncertainty sources have the potential to affect hydrologic applications (Maggioni and Massari, 2018). For example, simulations using Level-3 products for predicting streamflow and runoff are greatly impacted by their performance in terms of precipitation detection and quantification. Common issues involve systematic bias in the precipitation estimates, since hydrologic simulation highly depends on basin-scale water budget assumptions that directly impact the streamflow simulation (for example, Thiemig et al., 2012). Systematic biases arising at the satellite Level-2 and propagating to the Level-3 products are conditioned on a number of factors, some of which are independent of surface hydrology, such as precipitation physics, climatologies, sensors and algorithms (see Figure 1.2.1).

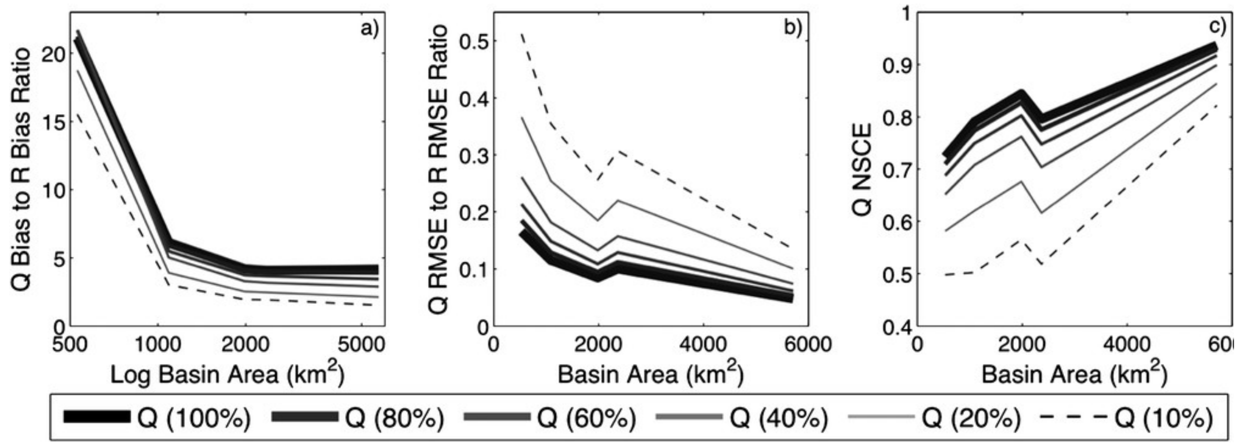

Figure 1.2.3. An example of propagation of resolution-induced rainfall error to streamflow simulations, using a distributed hydrologic model on the Tarboro watershed (North Carolina, U.S.) over the period 2002–2009. Metrics of hydrologic model performance as functions of basin area and streamflow threshold are presented: (a) ratio of streamflow (Q) relative bias to rainfall (R) relative bias, (b) ratio of streamflow (Q) relative RMSE to rainfall (R) relative RMSE and (c) Nash–Sutcliffe coefficient of efficiency (NSCE) of streamflow simulations. The values between parentheses indicate the probability of occurrence of the corresponding threshold. From Vergara et al., 2014

One strategy dealing with the uncertainty is to mitigate it by (1) debiasing for reconciliation with higher-resolution hydrologic models, (2) averaging/filtering/smoothing to obtain coarser-resolution products (typically 1-day 1-degree; see Chapter 2). Applications running at the monthly or seasonal time scales are less affected by biases because Level-3 satellite products increasingly benefit from gauge-based adjustments at coarser scales. However, many applications require spatial resolutions finer than 25 km and temporal resolutions less than 3 hours (Kirschbaum et al., 2017). In many cases, the uncertainty is transferred into the applications. Bias correction techniques can reduce streamflow errors (for example, Serrat-Capdevila et al., 2014); for example, by applying climate-scale bias corrections (for example, Beck et al., 2017). However, the multi-factor and nonstationary nature of satellite-based precipitation biases (that are not well understood yet; see Chapter 1.1) hinders the effectiveness of correction techniques. Another option is to compensate the forcing biases with hydrologic model calibration (for example, Xue et al., 2013; Nikolopoulos et al., 2013). It is made possible because the observed hydrologic response (discharge at the basin outlet) results from unobserved and integrated contributions of surface and subsurface processes. This endemic lack of observational hydrologic constraints leaves a considerable range of options to adjust hydrologic model parameters to reproduce the observed behavior (Beven, 2001), sometimes at the expense of physical realism. Model recalibration has been applied across watersheds with various geomorphologies and climatologies around the world to cope with satellite precipitation biases and improve streamflow prediction. This transfer of uncertainties from the satellite precipitation estimates to modeled hydrologic processes estimates hinders the broad application of hydrologic modeling, especially at sub-basin scales.

Uncertainty also arise due to the resolution of current satellite-based rainfall products and impacts applications of hydrologic modeling and forecasting systems. Resolution modifies the spatial structure of rainfall fields, and its interplay with basin area conditions the propagation of biases in distributed hydrologic models (for example, Vergara et al., 2014; Figure 1.2.3). The effects of precipitation resolution can be accounted for during the calibration of hydrologic models. The systematic analysis of the complex and combined effects arising from satellite uncertainties and resolution, basin geomorphologic characteristics and hydrologic modeling approaches remains a great challenge for the hydrologic application of satellite precipitation estimates. For more reliable flood simulations, additional physical constraints can be brought by observations, such as using soil moisture as a fingerprint of past rain occurrence (for example, Crow et al., 2011; Ciabatta et al., 2015).

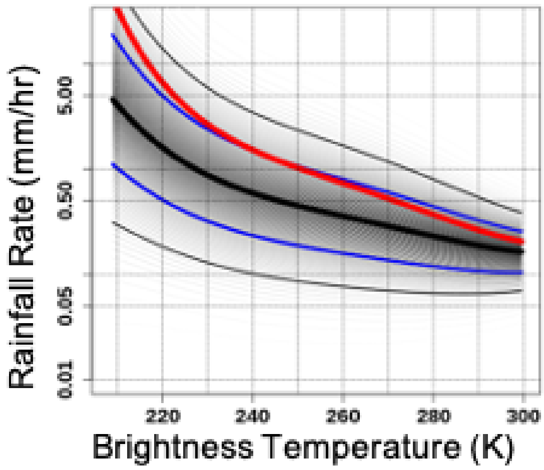

Figure 1.2.4. Precipitation rate distributions conditioned on the IR brightness temperature. The thick black line represents the median (50% quantile), the blue curves represent the 25 and 75% quantiles, the thin black lines represent the 10 and 90% quantiles. The red curve represents the expected value. The intrinsic bias in Figure 1.1.12 is mitigated by the probabilistic approach, while the intrinsic uncertainty is represented by the spread of the conditional precipitation rate distribution (adapted from Kirstetter et al., 2018a).

Another strategy dealing with satellite-based precipitation is to explicitly integrate uncertainty into the precipitation estimation process (see Chapter 3.1). Most precipitation products are deterministic and represent a single "best guess" realization of precipitation but are blind to their intrinsic uncertainty. Recent probabilistic precipitation estimates are developed to explicitly represent uncertainty (Kirstetter et al., 2015, 2018a; Wright et al., 2017), as shown in Figure 1.2.4.

1.2.5. Summary on intrinsic uncertainty and implications for hydrologic applications

Understanding hydrometeorological processes and applications requires more than just one deterministic “best estimate” to adequately cope with the intermittent, highly-skewed distribution that characterizes precipitation. The intrinsic uncertainty structure of satellite-based quantitative precipitation estimates is still largely unknown at the spatiotemporal scales near the sensor measurement scale. Advancing the use of uncertainty as an integral part of QPE in the relationship between sensor measurements and the corresponding “true” precipitation has the potential to provide a framework for diagnosing intrinsic uncertainty when instruments sample raining scenes or processes challenging QPE algorithms’ assumptions. It provides the basis for multisensor merging and precipitation assimilation, hydrometeorological hazard mitigation, decision making and hydrological modeling. Hydrologic applications are not generally configured to directly ingest probabilistic estimates of precipitation, but current research explores this avenue (for example, Hartke et al., 2020).

Numerous validation studies have been performed on satellite precipitation products. Chapter 1.3 summarizes what has been done to date by the IPWG validation subgroup. Limited progress has been made on quantifying the intrinsic uncertainty and its impact on hydrologic applications. This is because few studies are carried out at the primary precipitation retrieval scale. This endeavor requires the expert use of other precipitation sensors such as radar-gauge combinations.

1.2.6. Recommendations

Recommendation 1.2.1: Encourage more satellite precipitation comparisons at the actual satellite retrieval scale to study the intrinsic uncertainty.

The homogeneity of satellite precipitation is often overlooked in the evaluations while it remains an endemic challenge in the generation of such products and their applications. The dynamic interplay between precipitation characteristics, sensors and satellite algorithms is critical to study in order to make progress. There is a need to formulate the goals of validation and error modeling and to design appropriate comparison practices.

Recommendation 1.2.2: The non-homogeneity of satellite estimates at the retrieval scale needs more attention in order to improve products and their applications.

Currently hydrological applications are greatly impacted by the satellite precipitation intrinsic uncertainty. A better understanding of this uncertainty is critical to make progress and mitigate precipitation forcing errors and to avoid error propagation in other modeling components of the water cycle. Explicitly accounting for uncertainty in precipitation products is a promising way to explore coordination with hydrologic applications.

Recommendation 1.2.3: Further means to explicitly represent uncertainty in precipitation products and their hydrologic applications (beyond root mean squared additive error) should be explored.

1.2.7. References

Beck, H.E., A.I.J.M. van Dijk, V. Levizzani, J. Schellekens, D.G. Miralles, B. Martens and A. de Roo, 2017: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21, 589–615, https://doi.org/10.5194/hess-21-589-2017.

Berne, A., and W.F. Krajewski, 2013: Radar for hydrology: Unfulfilled promise or unrecognized potential? *Advances in Water Resources*, 51, 357–366, https://doi.org/10.1016/j.advwatres.2012.05.005.

Beven, K., 2001: How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5, 1–12, https://doi.org/10.5194/hess-5-1-2001.

Ciabatta, L., L. Brocca, C. Massari, T. Moramarco, S. Puca, A. Rinollo, S. Gabellani and W. Wagner, 2015: Integration of Satellite Soil Moisture and Rainfall Observations over the Italian Territory. *Journal of Hydrometeorology*, 16, 1341–1355, https://doi.org/10.1175/JHM-D-14-0108.1.

Ciach, G.J., and W.F. Krajewski, 1999: On the estimation of radar rainfall error variance. *Advances in Water Resources*, 22, 585–95, https://doi.org/10.1016/S0309-1708(98)00043-8.

Crow, W.T., M.J. van den Berg, G.J. Huffman and T. Pellarin, 2011: Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART). *Water Resources Research*, 47, W08521, doi:10.1029/2011WR010576.

Delahaye, F., P.-E. Kirstetter, V. Dubreuil, L.A.T. Machado, D.A. Vila and R. Clark III, 2015: A consistent gauge database for daily rainfall analysis over the legal Brazilian Amazon. *Journal of Hydrology*, 527, 292–304, https://doi.org/10.1016/j.jhydrol.2015.04.012.

Delrieu, G., B. Boudevillain, J. Nicol, B. Chapon, P.-E. Kirstetter, H. Andrieu and D. Faure, 2009: Bollène-2002 Experiment: Radar Quantitative Precipitation Estimation in the Cévennes–Vivarais Region, France. *Journal of Applied Meteorology and Climatology*, 48, 1422–1447, https://doi.org/10.1175/2008JAMC1987.1.

Delrieu, G., A. Wijbrans, B. Boudevillain, D. Faure, L. Bonnifait and P.-E. Kirstetter, 2014: Geostatistical radar-raingauge merging: A novel method for the quantification of the rain estimation accuracy. *Advances in Water Resources*, 71, 110–24, https://doi.org/10.1016/j.advwatres.2014.06.005.

Gebregiorgis, A., P.-E. Kirstetter, Y.E. Hong, N.J. Carr, J.J. Gourley, W. Petersen and Y. Zheng, 2017: Understanding multi-sensor satellite precipitation error structure in level-3 TRMM products. *Journal of Hydrometeorology*, 18, 285–306, https://doi.org/10.1175/JHM-D-15-0207.1.

Gray, E., 2015: Satellite-based flood monitoring central to relief agencies' disaster response. NASA, July, www.nasa.gov/feature/goddard/satellite-based-flood-monitoring-central-to-relief-agencies-disaster-response.

Habib, E., G.J. Ciach and W.F. Krajewski, 2004: A method for filtering out raingauge representativeness errors from the verification distributions of radar and raingauge rainfall. *Advances in Water Resources*, 27, 967–980, https://doi.org/10.1016/j.advwatres.2004.08.003.

Hartke, S.H., D.B. Wright, D.B. Kirschbaum, T.A. Stanley and Z. Li, 2020: Incorporation of Satellite Precipitation Uncertainty in a Landslide Hazard Nowcasting System. *Journal of Hydrometeorology*, 21, 1741–1759, https://doi.org/10.1175/JHM-D-19-0295.1.

Kirschbaum, D.B., G.J. Huffman, R.F. Adler, S. Braun, K. Garrett, E. Jones, A. McNally, G. Skofronick-Jackson, E. Stocker, H. Wu and B.F. Zaitchik, 2017: NASA's Remotely Sensed Precipitation: A Reservoir for Applications Users. *Bulletin of the American Meteorological Society*, 98, 1169–1184, https://doi.org/10.1175/BAMS-D-15-00296.1.

Kirstetter, P.E., Y. Hong, J.J. Gourley, Q. Cao, M. Schwaller and W. Petersen, 2014: Research framework to bridge from the global precipitation measurement mission core satellite to the constellation sensors using ground-radar-based national Mosaic QPE. In *Remote Sensing of the Terrestrial Water Cycle* (V. Lakshmi, D. Alsdorf, M. Anderson, S. Biancamaria, M. Cosh, J. Entin, G. Huffman, W. Kustas, P. van Oevelen, T. Painter, J. Parajka, M. Rodell and C. Rüdiger, eds.). Hoboken, New Jersey, Wiley, https://doi.org/10.1002/9781118872086.

Kirstetter, P.E., J.J. Gourley, Y. Hong, J. Zhang, S. Moazamigoodarzi, C. Langston and A. Arthur, 2015: Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resources Research*, 51, 1422–1442, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2014WR015672.

Kirstetter, P.-E., N. Karbalaee, K. Hsu and Y. Hong, 2018a: Probabilistic precipitation rate estimates with space-based infrared sensors. *Quarterly Journal of the Royal Meteorological Society*, 144 (Suppl. 1): 191–205, https://doi.org/10.1002/qj.3243.

Kirstetter, P.-E., W. Petersen and J.J. Gourley, 2018b: GPM Ground Validation Multi-Radar/Multi-Sensor (MRMS) Precipitation Reanalysis for Satellite Validation Product. Dataset available online from the NASA Global Hydrology Resource Center DAAC, Huntsville, Alabama, U.S.A., DOI: http://dx.doi.org/10.5067/GPMGV/MRMS/DATA101.

Kirstetter, P.-E., W.A. Petersen, C.D. Kummerow and D.B. Wolff, 2020: Integrated multi-satellite evaluation for the Global Precipitation Measurement mission: Impact of precipitation types on spaceborne precipitation estimation. In: *Satellite Precipitation Measurement*, Chapter 31 (V. Levizzani, C. Kidd, D.B. Kirschbaum, C. Kummerow, K. Nakamura, F.J. Turk, eds.), Advances in Global Change Research, 69, Springer-Nature, Cham, 583–608, https://doi.org/10.1007/978-3-030-35798-6_7.

Lebel, T., G. Bastin, C. Obled and J.D. Creutin, 1987: On the accuracy of areal rainfall estimation: A case study. *Water Resources Research*, 23, 2123–2134, https://doi.org/10.1029/WR023i011p02123.

Maggioni, V., and C. Massari, 2018: On the performance of satellite precipitation products in riverine flood modeling: A review. *Journal of Hydrology*, 558, 214–224, https://doi.org/10.1016/j.jhydrol.2018.01.039.

Neuper, M., and U. Ehret, 2019: Quantitative precipitation estimation with weather radar using a data- and information-based approach. *Hydrology and Earth System Sciences*, 23, 3711–3733, https://doi.org/10.5194/hess-23-3711-2019.

Nikolopoulos, E.I., E.N. Anagnostou and M. Borga, 2013: Using High-Resolution Satellite Rainfall Products to Simulate a Major Flash Flood Event in Northern Italy. *Journal of Hydrometeorology*, 14, 171–185, https://doi.org/10.1175/JHM-D-12-09.1.

Petersen, W.A., P.-E. Kirstetter, J. Wang, D.B. Wolff and A. Tokay, 2020: The GPM Ground Validation Program. In: *Satellite Precipitation Measurement* (V. Levizzani, C. Kidd., D.B. Kirschbaum, C.D. Kummerow, K. Nakamura, F.J. Turk, eds.). Springer Nature, Cham, Advances in Global Change Research, 69, 471–502, https://doi.org/10.1007/978-3-030-35798-6_2.

Serrat-Capdevila, A., J.B. Valdes and E.Z. Stakhiv, 2014: Water Management Applications for Satellite Precipitation Products: Synthesis and Recommendations. *Journal of the American Water Resources Association*, 50(2): 509–525, DOI: 10.1111/jawr.12140.

Shige, S., S. Kida, H. Ashiwake, T. Kubota and K. Aonashi, 2013: Improvement of TMI rain retrievals in mountainous areas. *Journal of Applied Meteorology and Climatology*, 52, 242–254, https://doi.org/10.1175/JAMC-D-12-074.1.

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 Thomas Wilheit, 2017: The Global Precipitation Measurement (GPM) mission for science and society. *Bulletin of the American Meteorological Society*, 98, 1679–95, https://doi.org/10.1175/BAMS-D-15-00306.1.

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, https://doi.org/10.1175/2006JAS2375.1.

Sun, Q., C. Miao, Q. Duan, H. Ashouri, S. Sorooshian and K. Hsu, 2017: A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56, 79–107, https://doi.org/10.1002/2017RG000574.

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

Taniguchi, A., S. Shige, M.K. Yamamoto, T. Mega, S. Kida, T. Kubota, M. Kachi, T. Ushio and K. Aonashi, 2013: Improvement of high-resolution satellite rainfall product for Typhoon Morakot (2009) over Taiwan. *Journal of Hydrometeorology*, 14, 1859–1871, https://doi.org/10.1175/JHM-D-13-047.1.

Thiemig, V., R. Rojas, M. Zambrano-Bigiarini, V. Levizzani and A. De Roo, 2012: Validation of Satellite-Based Precipitation Products over Sparsely Gauged African River Basins. *Journal of Hydrometeorology*, 13, 1760–1783, https://doi.org/10.1175/JHM-D-12-032.1.

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

Villarini, G., and W.F. Krajewski, 2010: Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surveys in Geophysics*, 31, 107–19, https://doi.org/10.1007/s10712-009-9079-x.

Wright, D.B., D.B. Kirschbaum and S. Yatheendradas, 2017: Satellite Precipitation Characterization, Error Modeling, and Error Correction Using Censored Shifted Gamma Distributions. *Journal of Hydrometeorology*, 18, 2801–2815, https://doi.org/10.1175/JHM-D-17-0060.1.

Wu, H., R.F. Adler, Y. Tian, G.J. Huffman, H. Li and J. Wang, 2014: Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model. *Water Resources Research*, 50, 2693–2717, doi: https://doi.org/10.1002/2013WR014710.

Xue, X., Y. Hong, A.S. Limaye, J.J. Gourley, G.J. Huffman, S.I. Khan, C. Dorji and S. Chen, 2013: Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *Journal of Hydrology*, 499, 91–99, https://doi.org/10.1016/j.jhydrol.2013.06.042.

Yamamoto, M.K., S. Shige, C.-K. Yu and L.-W. Cheng, 2017: Further improvement of the heavy orographic rainfall retrievals in the GSMaP algorithm for microwave radiometers. *Journal of Applied Meteorology and Climatology*, 56, 2607–2619, https://doi.org/10.1175/JAMC-D-16-0332.1.

Yang, N., K. Zhang, Y. Hong, Q. Zhao, Q. Huang, Y. Xu, X. Xue and S. Chen, 2017: Evaluation of the TRMM multisatellite precipitation analysis and its applicability in supporting reservoir operation and water resources management in Hanjiang basin, China. *Journal of Hydrology*, 549, https://doi.org/10.1016/j.jhydrol.2017.04.006.

Zawadzki, I., 1975: On radar-raingauge comparisons. *Journal of Applied Meteorology and Climatology*, 14, 1430–1436, https://doi.org/10.1175/1520-0450(1975)014<1430:ORRC>2.0.CO;2.

Zhang J., K. Howard, C. Langston, B. Kaney, Y. Qi , L. Tang, H. Grams, Y. Wang, S. Cocks, S. Martinaitis, A. Arthur, K. Cooper, J. Brogden and D. Kitzmiller, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: initial operating capabilities. *Bulletin of the American Meteorological Society*, 97, 621–38, https://doi.org/10.1175/BAMS-D-14-00174.1.