The Global Precipitation Measurement Mission: Performance evaluation across multiple spatio-temporal scales

Thesis submitted for the degree of
Doctor of Philosophy

by
Rajani Kumar Pradhan

Department of Water Resources and Environmental Modeling
Faculty of Environmental Sciences
Czech University of Life Sciences Prague

April 2024

Supervised by
doc. Ing. Mgr. Ioannis Markonis, Ph.D.

DECLARATION

I hereby declare that the thesis is independently written by me, and all the sources are quoted properly according to the methodological instruction provided to write a doctoral thesis. Furthermore, this thesis has not been submitted or accepted for any other degrees.

Place: Prague

Date: **Signature of the Candidate**

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Yannis Markonis, for supervising this Ph.D. thesis. It is an absolute honour and pleasure for me to be mentored by Yannis. He is not only an excellent researcher and mentor but also a great human being. Without his continuous encouragement and support, it would not have been possible to finish this work.

I would also like to express my sincere gratitude to Prof. Martin Hanel and Dr. Peter Máca for their important input and constant support throughout the years. Additionally, I would like to express my sincere gratitude to Dr. Francesco Marra for his invaluable guidance, support, and mentorship during my internship at the University of Padova (Department of Geosciences). Moreover, I would also like to acknowledge Prof. Vincenzo Levizzani and his group from CNR-ISAC Bologna, Italy, for their time, providing crucial insights, and hosting me a few times during the internship. Additionally, I would like to thank all of the group members and colleagues, such as Johanna, Riya, Ujjwal, Shailendra, Mijael, Akif, Vishal, Akbar, etc., for their valuable discussion and insights. Special thanks to Alex for providing technical support and consistently recovering the server whenever I encountered issues and each time I crashed it. Moreover, it has never been an easy journey without the crucial support from the administrative department, and hence, I would like to express my sincere thanks to all administrative people, and especially to Helena Michálková, Petr Bašta and Petra Kadlecová, among others.

Finally, I would like to thank my parents and friends for their constant support throughout the Ph.D. journey. Especially, I would like to thank my Father, Niranjan, my mother, Pankaj, and my brother, Prabhat for their continuous support and love over the years.

Abstract

Precipitation is probably one of the most important water cycle variables to measure. Considering its high spatio-temporal variability, satellite-based techniques emerge as the most reliable method for accurate measurements with near-global coverage. However, being the indirect measurement method, satellite estimates are subject to many assumptions and thus suffer from several limitations. Therefore, it is essential to thoroughly evaluate satellite estimates and comprehend their uncertainties prior to their utilization in scientific endeavors.

This thesis aims to comprehensively evaluate the GPM IMERG precipitation estimates at the global level. Initially, it assesses the IMERG product's performance based on published studies between 2016 – 2019. This quantitative review lays down the foundation for the two other important research objectives. Firstly, the evaluation of IMERG precipitation estimates across the tropical oceans. Secondly, the assessment of IMERG alongside other satellites and reanalysis estimates at sub-daily scales, focusing on their ability to depict the diurnal cycle at the global level.

The results reveal that IMERG effectively captures the spatial-temporal variation in precipitation across diverse geographic and climatic conditions, often outperforming TRMM products. However, IMERG exhibits certain limitations, particularly during heavy and extreme precipitation events. Despite its potential for detecting and tracking extreme events, IMERG underestimates the high precipitation intensities, especially over the core of typhoons and hurricanes. Its performance shows a relative improvement during summer compared to winter and is more effective in areas characterized by high-intensity and humid conditions than in drier regions with lower precipitation. In hydrological applications, it has shown comparatively better performance than TRMM, though its performance varies with the region, model, calibration datasets, etc.

In terms of IMERG performance over the ocean, it represents well the precipitation spatial patterns across the tropical oceans, similar to the buoys. However, IMERG tends to overestimate the precipitation in the high precipitation regions, such as over the Indian and West Pacific Oceans. IMERG detection ability is best over these high-precipitation regions compared to the low-precipitation regions (i.e., East Pacific and Atlantic regions). The false alarms and positive hit bias are the major components that contribute to the total error.

Finally, IMERG, along with other satellite estimates (GSMaP, CMORPH, PERSIANN), and the

reanalysis product (ERA5), agree on producing major global diurnal features with an afternoon peak over land and an early morning peak over the ocean. However, discrepancies do exist among the estimates in terms of the exact time of peak, amount, frequency, and intensities. In particular, ERA5 suffers from overestimation (underestimation) of precipitation frequency (intensity) throughout the globe. Among the satellite precipitation estimates, IMERG and CMORPH are in close agreement in most instances, while GSMaP and PERSIANN agree in a few instances.

The thesis contributes to a comprehensive understanding of the uncertainties and inaccuracies inherent in IMERG across a variety of geographical and climatic conditions at the global level. This not only provides valuable insights for algorithm developers to identify strengths and limitations but also empowers end-users and the scientific community to select datasets accordingly. As a result, this enhanced understanding improves the quality of outcomes and facilitates more informed decision-making processes. Given the widespread application of IMERG precipitation data in various fields such as weather forecasting, cyclone tracking, flash flood and landslide forecasting, elucidating the uncertainties associated with IMERG datasets holds paramount importance. Ultimately, this information leads to improved overall results and decision-making, thereby mitigating the societal impact of adverse events.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Importance of precipitation

Precipitation, a vital element of the water cycle, plays a pivotal role in maintaining Earth's energy balance and regulating the global climate system (Kidd and Huffman, 2011). As the availability of freshwater to humans is a frontier in the coming decades, precipitation directly impacts societal needs. In particular, precipitation, being an integral part of the global climate system, directly influences climate change dynamics, runoff patterns, groundwater levels, crop and food security, water management practices, and other critical factors that profoundly impact our daily lives.

In addition, the occurrence of extreme precipitation events, such as floods and landslides resulting from intense precipitation, carries profound implications for human society, ecological systems, and hydrological processes (Hou et al., 2014; Tapiador et al., 2019). To enhance our ability to predict, manage, or devise mitigation strategies for such events, the availability of long-term and accurate quality observation is of utmost importance. Moreover, precise precipitation estimation is very sensitive and holds particular importance in the studies of weather and climate forecasts, as well as in the field of extreme events and agricultural forecasts. Therefore, ensuring accurate precipitation estimation at both regional and global scales is imperative for scientific understanding and effective decision-making.

1.2 Precipitation measurement and its challenges

Precipitation is probably one of the most commonly measured climatic variables, though it varies highly in space and time. Its complex nature and its interaction with other variables in the system across various spatiotemporal scales make it one of the most challenging variables to accurately measure. Traditionally, precipitation has been measured using rain gauges, and is considered one of the most trusted sources of primary precipitation data (Sun et al., 2018). To date, rain gauges are considered the ground truth for evaluation of other sources of precipitation measurements.

Despite being considered the most trusted and accurate source of precipitation data, rain gauge measurements suffer from several limitations (Kidd et al., 2017; Tapiador et al., 2011). Rain gauges measure precipitation at point scales and represent small areal precipitation depending

on the homogeneity of the region (Tapiador et al., 2017). Consequently, the sparse and uneven distribution of rain gauges limits the capture of the spatial variability of precipitation, especially over mountains and oceans. Furthermore, for the largest part of the globe, most rain gauges only provide daily measurements, not fulfilling the demand for high spatiotemporal resolution precipitation data, leading to a need for downscaling methods to transform precipitation from coarse to fine scales (Papalexiou et al., 2018). Additionally, rain gauges are susceptible to the effects of wind, splashing, evapotranspiration, and other factors. However, their unequal distribution, intermittent coverage across most continents, and point-based measurements are the primary hampers to using rain gauges for global precipitation measurement.

To overcome these issues, various indirect measures of precipitation have been developed over the years. Such as remote sensing-based estimates, which include both space-based satellite estimates and ground radar estimates. Ground-based radar measurements are characterised by better accuracy and precision, having the additional advantage of capturing the three-dimensional structure of precipitation in real-time. Yet radars are expensive to maintain and install and require skilled experts, which hampers their deployment at a global scale and still, station data are needed to calibrate the precipitation fields (Sun et al., 2018). In addition, the beam blockage by the continental feature is another major drawback of radar-based measurements (Lang et al., 2009). Therefore, deploying radar systems globally to achieve comprehensive precipitation measurements is highly impractical. This challenge becomes particularly pronounced in data-scarce regions like mountains, oceans, and deserts, where the maintenance of radar infrastructure is nearly unattainable.

1.3 Satellite precipitation

Unlike ground-based remote sensing precipitation, satellite-based estimates present themselves as an alternative and viable solution for achieving global precipitation measurements. During the last three decades, satellite data have become a promising source of precipitation observations at the global scale (Levizzani and Cattani, 2019). They provide continuous measurement of precipitation in both space and time with quasi-global coverage (Derin and Yilmaz, 2014), making them

especially important over data-scarce regions (Kidd and Levizzani, 2011). Moreover, satellite-based estimation of precipitation is the only way to achieve near-real-time global precipitation measurement with finer spatiotemporal resolution. In broad terms, satellite-based precipitation estimation techniques can be categorised into three main methods; (i) the visible (VIS) and infrared (IR) based, (ii) the passive microwave (PMW) based, and (iii) the merged VIS/IR and PMW based (Kidd and Levizzani, 2011).

As its name implies, the visible (VIS) and infrared (IR) based method uses VIS/IR images obtained from geostationary satellites and estimates precipitation based on cloud-top temperature. In particular, it relies on the assumption that a colder cloud top means larger vertical development of the cloud and, therefore, is associated with more precipitation (Tapiador et al., 2012). However, the cloud-top temperature and surface precipitation relationship is not remains consistent across all conditions and varies with the precipitating cloud types. For instance, the relationship is weak for the shallow and warm clouds, whereas it is more robust for cold and tall clouds (So and Shin, 2018). Considering the high spatial-temporal resolution of the IR sensors onboard the geostationary satellites, this method provides important insights into the continuous monitoring and tracking of hazardous storms and evolving weather systems. Nevertheless, uncertainties in the indirect relationship between the cloud top temperature and precipitation rate impede its precipitation estimates (Maggioni et al., 2016; Tapiador et al., 2017).

On the other hand, the PMW-based method links microwave scattering/emission and rain/ice particles and thus provides more direct precipitation estimates, as opposed to the IR-based, which is solely based on cloud-top temperature (Sun et al., 2018). In particular, it mainly exploits the emission and scattering properties of precipitation hydrometeors and frozen hydrometeors. While emission by precipitation hydrometeors results in an increase in total radiation, the scattering by precipitation ice particles leads to a decrease in the total radiation (Kidd and Levizzani, 2011). Unlike over the homogeneous ocean and large water bodies, the emission of microwave radiation from heterogeneous land is more complex. Consequently, the retrieval of precipitation from PMW estimates often consists of separate processes over land and ocean. For instance, due to the low emissivity of the ocean, the retrieval of precipitation over oceans is mainly based on emission at lower microwave frequencies (< 20 GHz). In contrast, scattering by ice particles at

higher frequencies ($>$ 35 GHz) is predominantly used over land, taking into account the higher emissivity of land.

Even though, compared to the IR-based methods, PMW estimates provide better estimates of precipitation, they also suffer from various issues. A major drawback is that, due to some technical reasons, all the microwave sensors are primarily limited to the low-orbit satellites, and hence can only measure a very small fraction of the earth precipitation at a time. This leads to poor sampling in time, which yields significant gaps in the precipitation estimation (Hong et al., 2019). Another major challenge arises in the estimation of snow over land. It always poses a difficult challenge for the PMW estimates to distinguish snow over the land from falling snow as precipitation from the atmosphere.

The complementary properties of the PMW and VIS/IR technologies led to the development of the VIS/IR and PMW merged data products. Even though IR estimates are less skillful, the sparse PMW estimates cannot adequately cover all times, so IR estimates are used to help cover the periods which are not well-represented by the PMW sensors (directly or by morphing). Some of these merged data products include the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network (PERSIANN) (Sorooshian et al., 2000; Hsu et al., 1997), Climate Prediction Center Morphing Technique (CMORPH) (Joyce et al., 2004), Tropical Rainfall Measuring Mission (TRMM) (Huffman et al., 2007), Integrated Multi-Satellite Retrievals for GPM (IMERG) (Huffman et al., 2015), and Global Satellite Mapping of Precipitation (GSMaP) (Kubota et al., 2020). Among them, the TRMM mission, which was designed to estimate precipitation over the tropics and subtropics, provided significant information on rainfall and its associated properties (Huffman et al., 2007). The Multi-Satellite Precipitation Analysis (TMPA), the most important data product of TRMM, has been one of the widely used precipitation datasets for a range of applications (Hamada et al., 2014; Li et al., 2009; Siddique-E-Akbor et al., 2014; Rozante and Cavalcanti, 2008). The continuation of the TMPA is the latest IMERG precipitation, another major centerpiece of satellite precipitation over the past few years.

1.4 GPM Mission

Considering the importance of precipitation measurement to the global water and energy cycle, the Global Precipitation Measurement Mission (GPM) stands as a collaborative initiative between the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA). This mission is primarily aimed at delivering precise precipitation measurements at fine spatio-temporal resolutions on a global scale to enhance our understanding of the global water cycle. The GPM mission is the continuation of its predecessor, i.e., TRMM, which also started in 1997 based on similar objectives of accurate precipitation measurement, however limited to the tropics. Following the success of TRMM, on February 27, 2014, NASA and JAXA jointly launched the Global Precipitation Measurement (GPM) Core Observatory (GPM CO) satellite (Hou et al., 2013; Liu, 2016). The GPM CO satellite served as a calibration reference for the various microwave satellite estimates onboard the international spacecraft (Hou et al., 2014). Similar to the TRMM, the GPM mission is based on the concept of integrating precipitation from a constellation of satellites. This involves utilising and collaborating with various satellite precipitation providers from national and international partners, and merging these various estimates based on their respective strengths (Figure 1.1). It is also worth mentioning that except for the GPM CO satellite, NASA has no control over the orbit, channels, data policy, and overall operation of the constellation satellites. Therefore, the success of the GPM mission is very much dependent on the international collaboration and cooperation of the partner satellites.

Compared to its predecessor i.e., TRMM, GPM not only improved its orbital inclination, extending TRMM's coverage from 35° N-S to 65° N-S, but it also incorporated several advanced sensors. This includes the active microwave Dual-Frequency Precipitation Radar (DPR) and the passive microwave GPM Microwave Imager (GMI). In fact, the GPM CO carried the first multi-frequency space-borne precipitation radar (i.,e, DPR) and, overall, the second precipitation radar in space after the TRMM's Precipitation Radar (PR). The DPR, encompassed with the Ku-band (35.5 GHz) and Ka-band (13.6 GHz) channels, enables GPM to measure the three-dimensional structure of precipitation from space (Skofronick-Jackson et al., 2017). Similarly, the GMI, a 13-channel conical scanning microwave radiometer, enables the CO spacecraft to serve as a pre-

cipitation standard for other constellation members. This assists the GPM mission in achieving near-global coverage and frequent revisit time requirements. Moreover, these advanced sensors onboard the GPM mission have substantially improved the accuracy of quantitative precipitation estimation, particularly enhancing sensitivity to a broader range of hydrometeors, including light and solid precipitation (Hou et al., 2014).

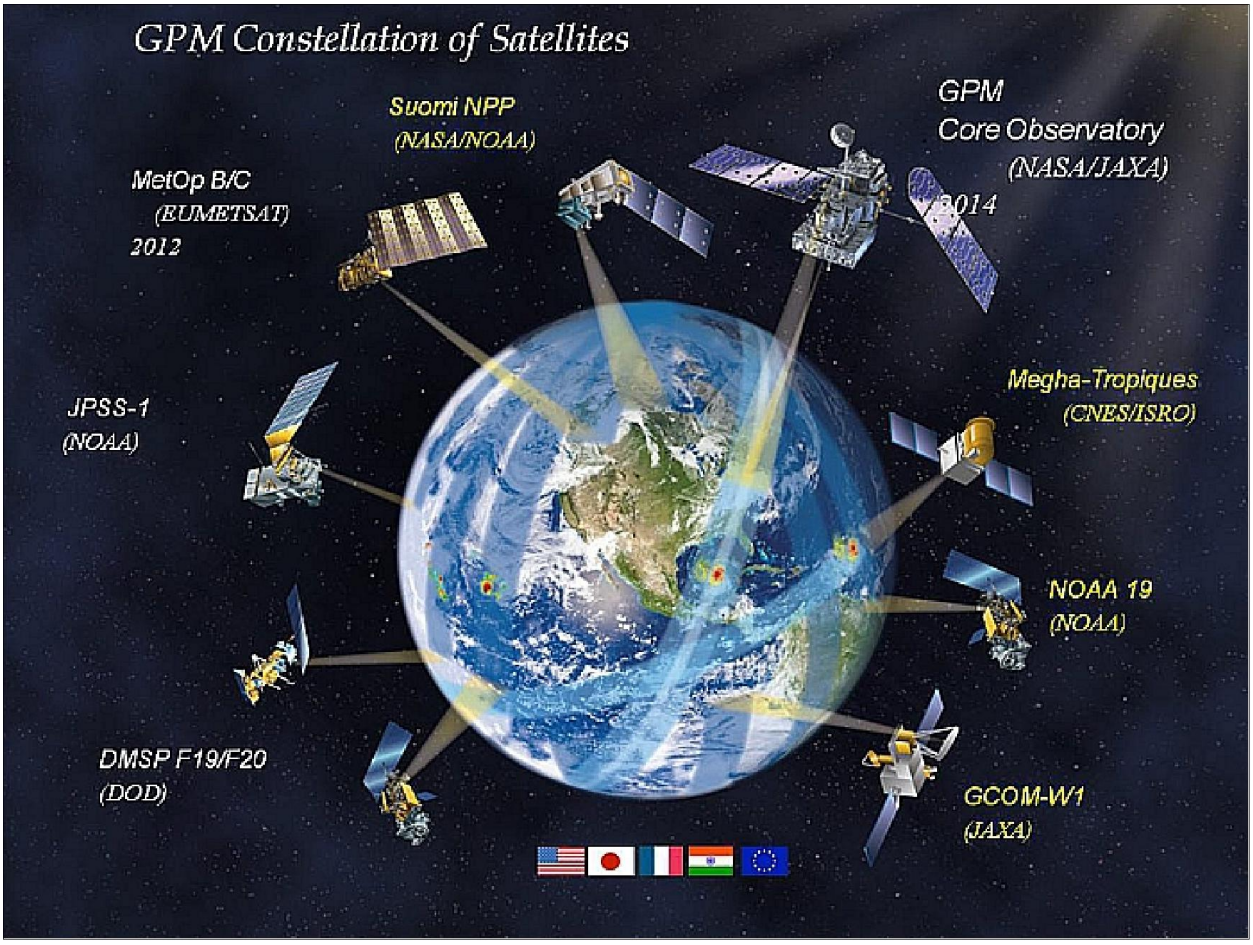


Figure 1.1: GPM mission architecture with a constellation of national and international satellites (Image credit: NASA)

1.5 IMERG precipitation

On the processing side, the Integrated Multi-Satellite Retrievals for GPM (IMERG) algorithm incorporates, merges, and inter-calibrates various IR, microwave (MW), from the GPM constellation satellites and gauge observations to provide precipitation estimates at relatively high spatial

($0.1^\circ \times 0.1^\circ$) and temporal resolution (30 min) (Huffman et al., 2015). At present, IMERG stands out as one of the leading sources of global precipitation, offering exceptional accuracy at very fine spatial resolution across the globe (Huffman et al., 2015).

The IMERG algorithms consist of five main steps: 1) Estimating precipitation rates from passive microwave (PMW) sensors, 2) Cross and inter-calibration of PMW retrieved precipitation rates with the GPM Core DPR, 3) Correction of IR precipitation rates using the calibrated PMW estimates, 4) Combing the IR and PMW estimates, and 5) Monthly scale bias correction using the GPCC products over the land.

The estimation of precipitation from all the PMW satellite constellations is mainly performed using the Goddard Profiling Algorithm V05 (GPROF) (Kummerow et al., 2015; Randel et al., 2020). However, the only exception for this is the Sounder for Probing Vertical Profiles of Humidity (SAPHIR) data, for which the Precipitation Retrieval and Profilling Scheme (PRPS) (Kidd, 2019) is applied. Additionally, the conical scanner estimates take precedence over cross-scale estimates when multiple precipitation estimates are accessible for the same grid. Furthermore, preference is given to estimates that are closest to the half-hour grid (Tan et al., 2019c). The microwave precipitation estimates are then converted to $0.1^\circ \times 0.1^\circ$ grid resolution using the nearest-neighbour interpolation technique. The inter-calibration of the gridded PMW estimates are then performed using the Ku-band Combined Precipitation Radar and Radiometer (CORRA) (Olson et al., 2016) product. The CORRA product is already seasonally adjusted with the Global Precipitation Climatology Project V2.3 (GPCP) (Adler et al., 2003, 2018) product to compensate the biases over the known region with CORRA deficiencies over land and ocean. Since the partner satellites have various orbital characteristics and owing to the Ku-band narrow swath width, this results in very sparse coincident sampling with the partner satellites. Therefore, partner satellite calibration directly with the CORRA is not feasible. Instead, it inter-calibrates the PMW estimates in two steps. First, the PMW estimates are calibrated with the GMI and TMI imagers for the respective GPM/TRMM era. Second, the GMI and TMI imagers are calibrated with the CORRA estimates using probability matching. The intercalibration also performed for the GMI/TMI calibrated partner sensors as well.

The gridded PMW estimates are then gap-filled through the Morphing schemes (Joyce et al.,

2004). This scheme employs the quasi-Lagrangian interpolation using the estimated motion vector from ancillary data to propagate precipitation. Until the V05, Geo-IR cloud images were used to derive the motion vectors. However, it was recently replaced with the total column water vapour (TQV), derived from Modern-Era Retrospective Analysis for Research and Application (MERRA2) (Gelaro et al., 2017) (in IMERG V06 Final) and Goddard Earth Observing System Model (GEOS) Forward Processing (PF) (for Early and Late runs). These changes are based on improved precipitation demonstrations, as noted by Tan et al. (2019c). Furthermore, the estimation of motion vectors from the total column water vapour instead of IR images also allowed IMERG to extend precipitation propagation to the entire globe, which was initially limited to 60°N-S. In addition, the morphed precipitation is complemented via the Kalman Filter approach (Joyce and Xie, 2011), in which the input precipitation is computed using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS) (Hong et al., 2004; Nguyen et al., 2018). The PERSIANN algorithm uses the PMW-calibrated IR brightness temperature to derive precipitation using the neural network. Finally, the precipitation estimate is bias corrected against the Global Precipitation Climatology Centre (GPCC) (Schneider et al., 2014) based monthly estimates, which prodcuces the IMERG-Final precipitation estimates.

Furthermore, to accommodate various user requirements, IMERG provides three types of datasets: the IMERG-Early run (IMERG-E), IMERG-Late run (IMERG-L), and IMERG-Final run (IMERG-F). IMERG-E and IMERG-L, being near-real-time products, are available with a latency of 4 hours and 14 hours, respectively, and can serve as a potential data source for flood forecasting and real-time disaster management (Huffman et al., 2020). IMERG-F is available 3.5 months after observation, and it is mainly aimed for research purposes. Unlike IMERG-E and IMERG-L, IMERG-F incorporates the GPCC monthly gauge analysis. More detailed information about IMERG precipitation products can be found in Tan et al. (2019c). Since IMERG's release in early 2015, a substantial number of studies have used and recommended it for various applications, such as streamflow simulation (Tang et al., 2016b), flood forecasting (Wang et al., 2017b), and analysis of extreme events (Huang et al., 2019). Recently, IMERG version 06 (V06) extended its temporal coverage to the TRMM era (2000 - 2015) and now provides 20-year-long datasets from 2000 to present (Huffman, 2020). The high-quality precipitation estimates and long-term cover-

age of IMERG is expected to provide insights into various hydro-meteorological processes and climatological studies in the future.

1.6 Importance of uncertainty assessment in satellite estimates

Also, all these satellite-based precipitation estimates have enormous advantages compared to the point-based gauge/station estimates, but they also come with substantial challenges. Especially since all these estimates are indirect measures of precipitation, their accuracy and, therefore, their reliability is a major concern. Since precipitation is a highly complex and major driver of the hydrological processes, uncertainties in the input datasets would have unimaginable consequences. Therefore, the evaluation and validation of this estimation prior to their application in related domains has been a major research domain over the past few decades. Especially over the last two decades, there has been incredible growth in the development and availability of various precipitation estimates; it is a major challenge to choose the proper estimates. In particular, since the GPM mission is relatively new and the IMERG algorithm has undergone several changes (V03 to V07 in the last eight years) over the years with continuous improvement, it is of paramount importance to evaluate how well these retrievals perform under different environmental and geographical conditions across the globe. This will give meaningful insights to both the algorithm developers and the end users to know the strengths and weaknesses of the product and, thus, its applicability.

Chapter 2

Objectives of dissertation

2.1 Objectives of dissertation

The thesis seeks to provide valuable insights into the characteristics and behaviour of GPM IMERG precipitation across a range of geographic and climatic conditions at the global level, contributing to the advancement of scientific understanding in this field. The key research objectives addressed in the thesis include: i) Quantitatively assessing the current status of IMERG precipitation estimates on a global scale, ii) Evaluating the performance of IMERG precipitation products over the tropical oceans, and iii) Examining the diurnal variability of IMERG along with various satellite and reanalysis estimates at the global level.

The following research questions (RQ's) will be addressed specifically:

- RQ-1. What is the current status of IMERG precipitation estimates across the globe? What are the main strengths and weaknesses of IMERG, and how does it change with season, region, etc.?
- RQ-2. How effective is IMERG in the representation of tropical ocean precipitation, and how does its bias change with spatial regions and precipitation intensities?
- RQ-3. How do different precipitation estimates capture the diurnal variability of global precipitation, and how is this variability influenced by distinct geographical features, such as land, ocean, etc.?

2.2 Approach and outline

To address the first research question (RQ1), a quantitative review of GPM IMERG evaluation and validation studies across the globe published between 2016 - 2019 is carried out (Pradhan et al., 2022). Upon the collection of studies, a detailed database is prepared to record various details of the performed studies such as their global distribution, number of studies over the years, the spatial and temporal resolution at which the studies were carried out, precipitation types (rain or snow), etc. Furthermore, through a comprehensive review and recorded statistical matrices

(Bias, RMSE, MAE, POD, FAR, CSI, etc), IMERG's strengths and weaknesses are identified. Moreover, IMERG performance in terms of the region, season, intensity, and potential in extremes and hydrological application are detailed.

To deal with the second question (RQ2), IMERG estimates are evaluated against the buoy observations over the tropical ocean provided by the Global Tropical Moored Buoy Array (GTMBA) (Pradhan and Markonis, 2023). All three IMERG runs, i.e., IMERG-E, IMERG-L, and IMERG-F, are evaluated against the buoys estimates for the period 2001 – 2020. In addition, IMERG estimates are also evaluated against buoys with different precipitation intensities to identify how the IMERG precipitation estimates perform under different precipitation intensities. Moreover, the resulting errors are further decomposed to address the contribution of each error type to the total error.

The third research question (RQ3) deals with characterising the satellite and reanalysis estimates in terms of their diurnal variation across the globe. Along with IMERG, other satellite estimates such as GSMaP, CMORPH, PERSIANN, and reanalysis estimates, i.e., ERA5, are considered for the analysis. The diurnal characterisation among the estimates is carried out using the mean precipitation amount, frequency and intensity. First, the 20 years (2001 - 2020) mean of all the precipitation properties (amount, frequency and intensity) are compared in terms of their spatial variation and different latitudinal zones. Then, the diurnal variations among the estimates are evaluated.

The thesis begins with a theoretical background of the GPM IMERG precipitation. Chapter 3 deals with a quantitative review of the GPM IMERG precipitation, its current status, and its overall performance across a range of climates, geographical conditions, and scales at the global level. Following the research gap identified in Chapter 3, Chapter 4 evaluates the performance of IMERG precipitation estimates at the tropical ocean against the buoy's observations. Chapter 5 represents the sub-daily evaluation of various satellite and reanalysis products at the global level. In particular, it compares the sub-daily precipitation estimates and their diurnal structure at the global level. Finally, Chapter 6 summarises the major findings of the thesis, along with highlighting both limitations and future research recommendations.

Chapter 3

Current status of IMERG precipitation

3.1 Introduction

A considerable number of studies have evaluated the performance of IMERG precipitation products at various temporal and spatial scales (e.g. Navarro et al., 2019; Shawky et al., 2019; Watters and Battaglia, 2019a; Palomino-Ángel et al., 2019; Prakash et al., 2018b; Tan et al., 2016; Manz et al., 2017). Nonetheless, most of them focus on local/regional domains (e.g., bounded by national boundaries) addressing specific climatic or topographic conditions across the globe. The few studies that investigate the global performance of IMERG present contradictory results (Wang et al., 2018; Liu, 2016; Derin et al., 2019). In this context, the main objective of this chapter is to review the state-of-the-art of the IMERG precipitation products and summarize the results of the recent efforts to evaluate the IMERG products in a quantitative manner across the globe. We aim to identify the strengths and weaknesses of IMERG products, providing information to the user community and product developers to further improve IMERG algorithms in future versions. The chapter is organized as follows: the second section describes the methodology employed herein, briefly discusses data collection, database preparation, and its analysis. The third section presents the results in terms of validation design and geographical distribution of the publications. The fourth section discusses the post-2019 developments, strengths, and weaknesses of IMERG products across the globe. Finally, the last section summarizes the findings, and reports the conclusions.

3.2 Approach

To analyze the global performance of GPM IMERG products, we performed an exhaustive literature review using the Google Scholar and Scopus databases. We used the keywords “GPM" and “IMERG", focusing on the period between 2016 to 2019. We limited the used articles to ones that had the evaluation of the performance of IMERG products within their scope. To assure the quality of the scientific articles used in this analysis, we focused only on articles published in Q1 and Q2 journals according to Scopus (first and second quantile of journals according to their ranking in the hydrology, climate or remote sensing fields). These selection crite-

ria resulted in a 101 articles database with information regarding the performance of IMERG precipitation products across the globe. This information includes a unique identification (id) code, study area, country, continent, surface category, precipitation type, IMERG product, record length, temporal and spatial resolution, validation method, validation data, and statistical metrics (The database is available at `https://ars.els-cdn.com/content/image/1-s2.0-S0034425721004740-mmc1.xlsx`).

The ID code was generated as the first three letters of the first author's name followed by a two-digit number representing the year of publication. The study area varies widely from small river basins to global-scale analyses; therefore, the country and continent were also reported whenever the study area was not global. The surface category considered two major groups, namely land and ocean. Precipitation type was registered as rainfall or snowfall. IMERG products, as described in the previous section, could be IMERG-E, IMERG-L, or IMERG-F, and their version ranged between IMERG V03 and IMERG V06. The record length or length of the evaluation period was reported in months. The temporal resolution ranged from sub-hourly to yearly, and the spatial resolution was expressed in degrees. Validation methods considered the characteristics of the IMERG product (gridded) and the validation data (gridded or point) being compared, i.e., grid vs. grid or grid vs. point. Validation data reports the source of reference data used (gauge-, radar-, satellite-, or model-based). Statistical metrics were classified into volumetric and categorical indices. The volumetric indices mainly include the correlation coefficient (COR), root mean square error (RMSE), and bias. The key metrics in the categorical indices include the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI). POD, also known as hit rate, represents the detection capability of the satellite; measures the proportion of the events detected by the satellite to the total number of precipitation events. FAR denotes the fraction of events detected by satellite that is not real or not detected by gauge, while CSI, a function of POD and FAR, represents a balanced score. Furthermore, where available, the database recorded additional metadata like best/worst performances, limitations, year of publication, and the journal of publication.

We followed a two-step evaluation approach: Firstly, the experimental design of the studies was evaluated based on the generated database. Information such as latitude, longitude, and

the study area (both country and continents) were used to analyze the geographical distribution of the studies. Then, the publication years and their corresponding counts were employed for the investigation of the chronological evolution. In addition, the information regarding the spatiotemporal resolutions, validation period length, reference data types, statistical metrics, etc., were used for the analysis of additional aspects of the validation design. Secondly, the performance of IMERG was summarised and categorized in terms of the continents, hydrological applications (i.e., streamflow simulation), and extreme events. The statistical metrics such as bias, COR, RMSE, NSE, POD, and FAR were employed to support the results.

Herein, we must note that the analysis is not complete. The search criteria were restricted to the “IMERG or GPM" keywords on Google Scholar search engine, and studies that do not have those keywords in their title or abstract were excluded from the database. Only studies whose scope was solely IMERG validation were gathered. In addition, data collection is limited to studies published in Q1 and Q2 journals. There may be more studies in Q3 and Q4 category’s journals that were excluded from this analysis. Another limitation of the analysis is that we did not consider studies published outside journals, such as conferences, book chapters, and reports. Finally, the analysis focuses on articles published between 2016 and 2019 (there may be more evaluation studies published in the later 2020). It is worth noting that this analysis is aimed explicitly at GPM’s IMERG products only, so it does not imply to the entire GPM mission (IMERG is one of the products of GPM, and GPM provides other data sets as well).

3.3 Characteristics of IMERG studies

From the geographical distribution of the 101 studies (Figure 3.1) Asia not only has the majority of studies (66) but is also the continent with most studies covering different countries. The Americas hold a total of 21 studies, 12 in North America and 9 in South America, while there were 7 studies in Europe and 2 in Africa. Note that until 2019 no available studies had assessed IMERG performance over Australia. Additionally, 5 studies validated IMERG products at the global level (not shown in Figure 3.1). On a per-country basis, IMERG performance was evaluated over a total of 34 different countries. There is an unequal distribution where China, the United States, and

India represent around 55% of the studies. Yet, China alone accounts for 40% of the total. Countries like Brazil, Iran, Pakistan, Japan, Myanmar, Malaysia, and the Netherlands have at least two validation studies associated with each, whereas the rest of the countries are featured with a single study. Given the number of studies, IMERG validation is spatially well-distributed over the Asian continent.

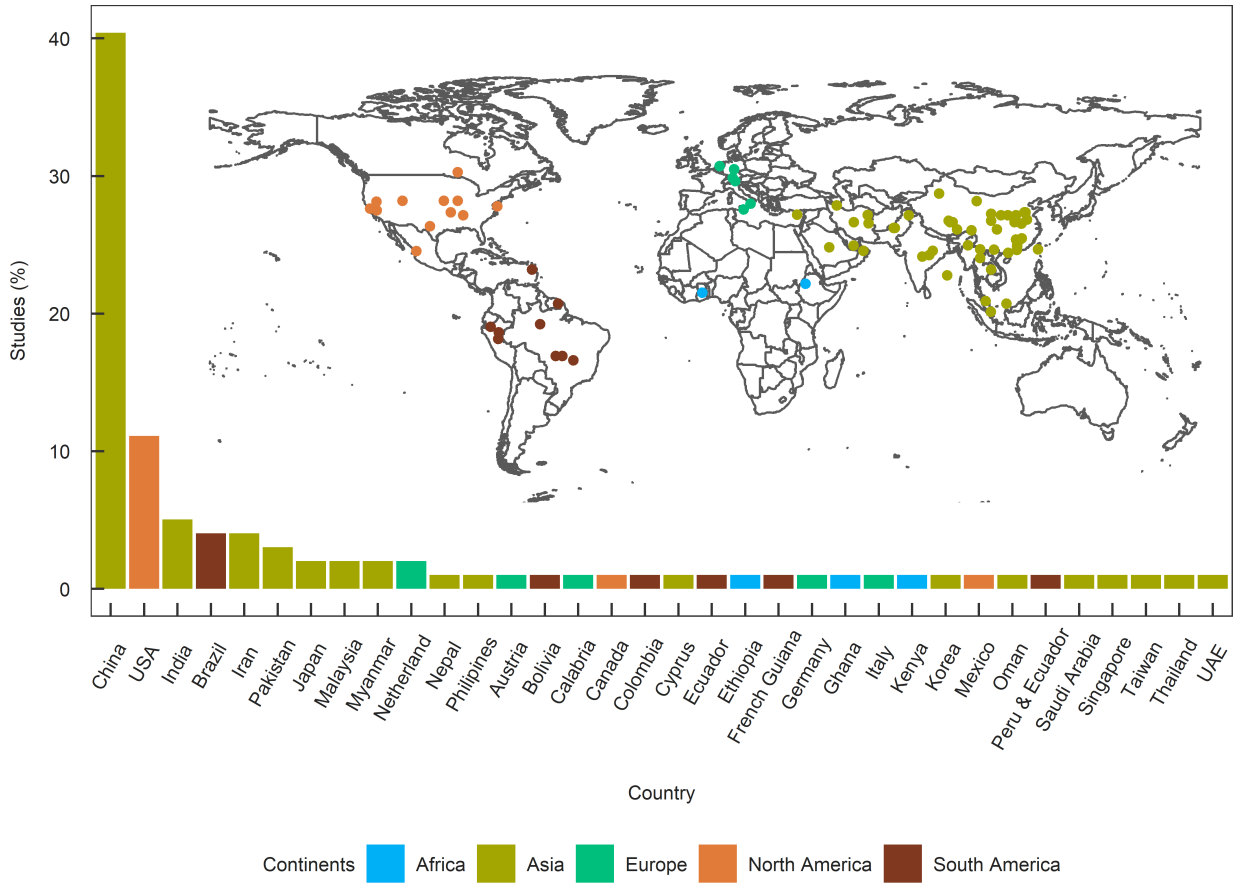


Figure 3.1: Geographical distribution of IMERG validation studies across the globe. The points do not represent the extent of the study domain but rather the mean latitude and longitude of the domain.

The chronology of the studies indicates a growing research interest in the topic (Figure 3.2). A total of 14 studies were published in 2016, 21 in 2017, 30 in 2018, and 36 in 2019. Africa has the lowest number of studies, it showed no increment with time, and no studies were published in 2018 and 2019. All other continents show an increase in the number of publications per year. Asia exhibited the highest growth in the number of publications per year. Starting with less than 10 studies in 2016, it has reached 28 in 2019, representing approximately three-quarters of all studies

published that year. In Europe, the number of studies remains similar each year, except for 2018, in which the number of studies doubled. North America reported an increasing trend of studies from 2016 to 2018, while a decrease in 2019. Finally, in South America, there is fluctuation in the number of studies with the years without any trend. For the global studies, 2016 and 2018 have the same number of publications, but there is a significant increase in 2019.

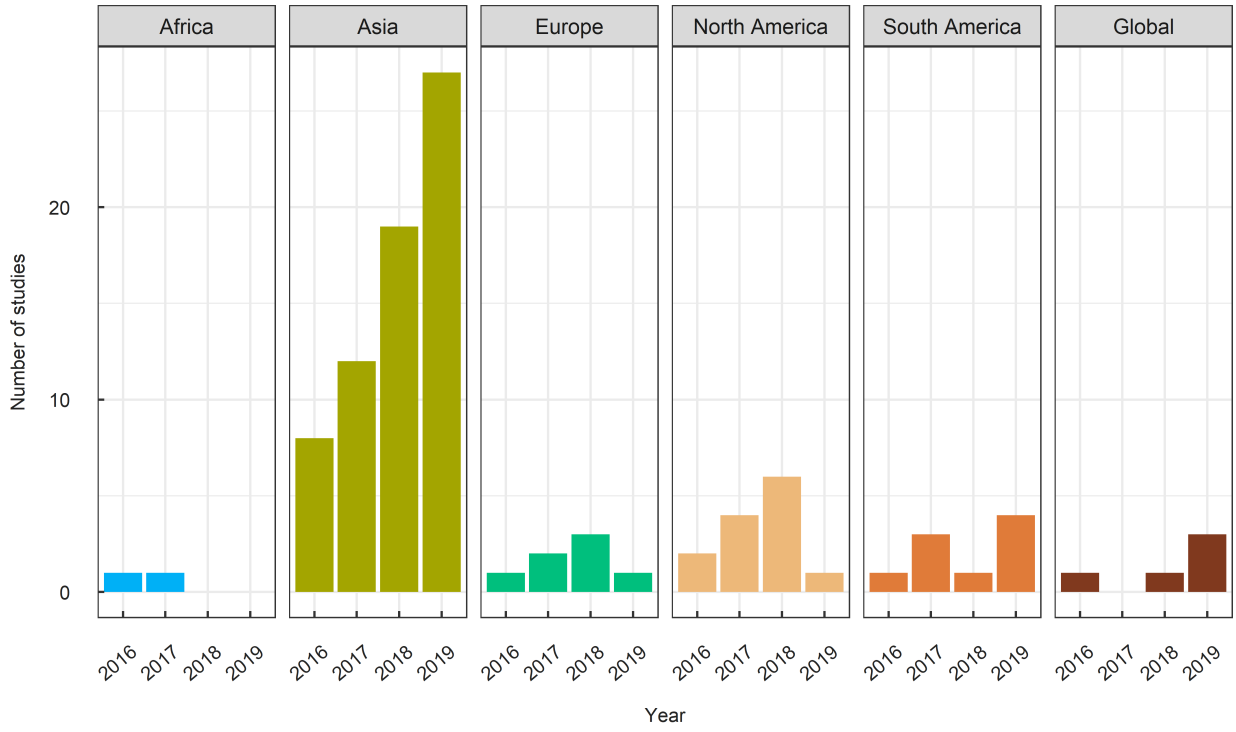


Figure 3.2: Number of IMERG validation studies published between 2016 and 2019

The temporal resolution at which the validation is performed can significantly affect the results. Therefore, we identified the different temporal resolutions of IMERG products used in the aforementioned studies (Figure 3.3). Those using products at daily resolution account for 35%, followed by monthly (22%), sub-daily (21%), annual (13%), and seasonal (9%) resolutions. When combined, the daily, monthly, annual and seasonal scales account for 80% of the total studies, whereas the rest (20%) are sub-daily scales. One possible explanation for this may be the availability of observational data sets. For example, this may be that the availability of observational data with high temporal resolution (sub-daily scale) at a regional scale is scarce. Another possible explanation could be that two precipitation data sets typically have a better agreement when they are upscaled to a coarser resolution in space and time. Therefore, evaluating IMERG at higher

temporal scales (i.e., sub-daily) is a bit challenging (Tan et al., 2016). Continent-wise, Asia dominates in most temporal resolutions beyond the sub-daily scale. Studies over Europe are evenly distributed across the different temporal resolutions. In general, the least number of studies were at sub-daily resolutions (attributed to the lack of corresponding reference data), indicating that IMERG's raw resolution (30 min) has still not been adequately evaluated.

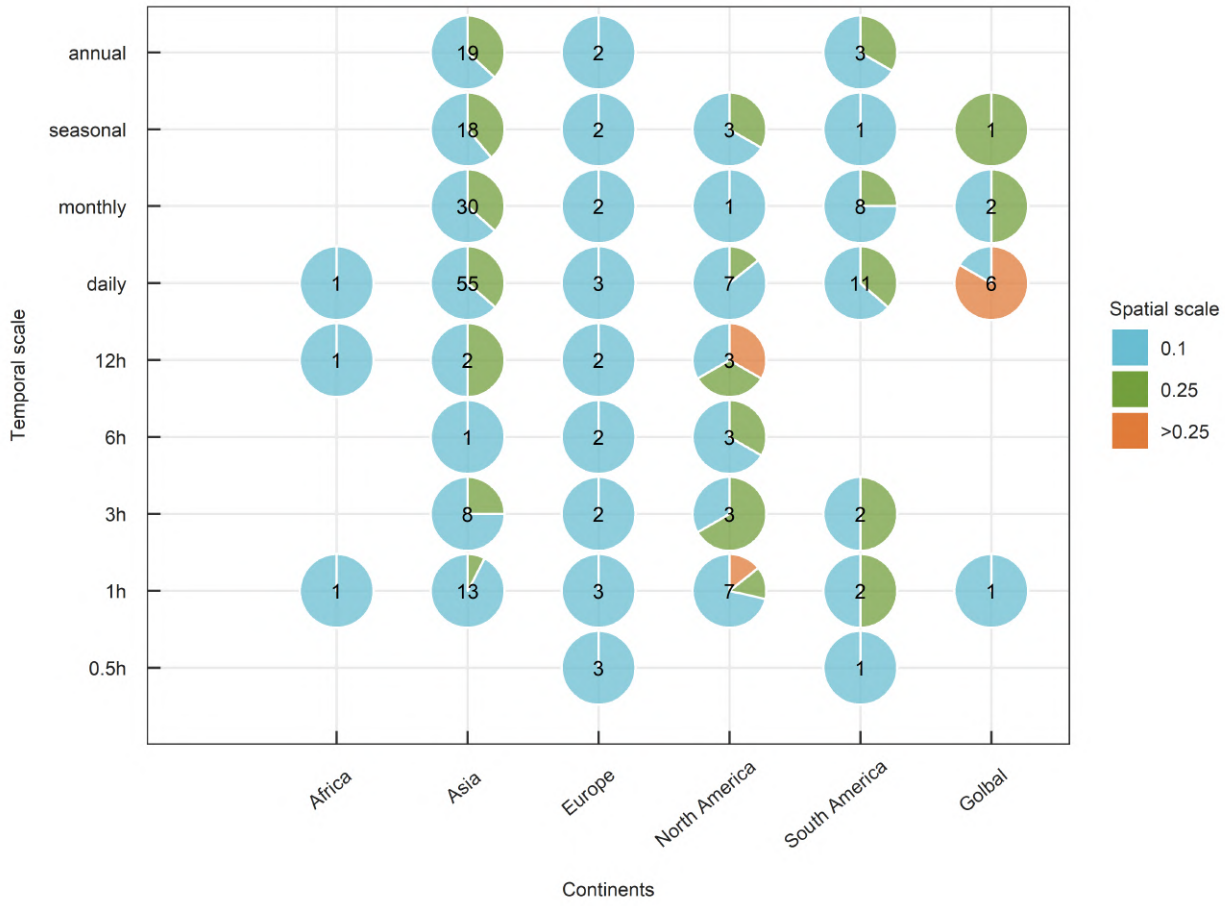
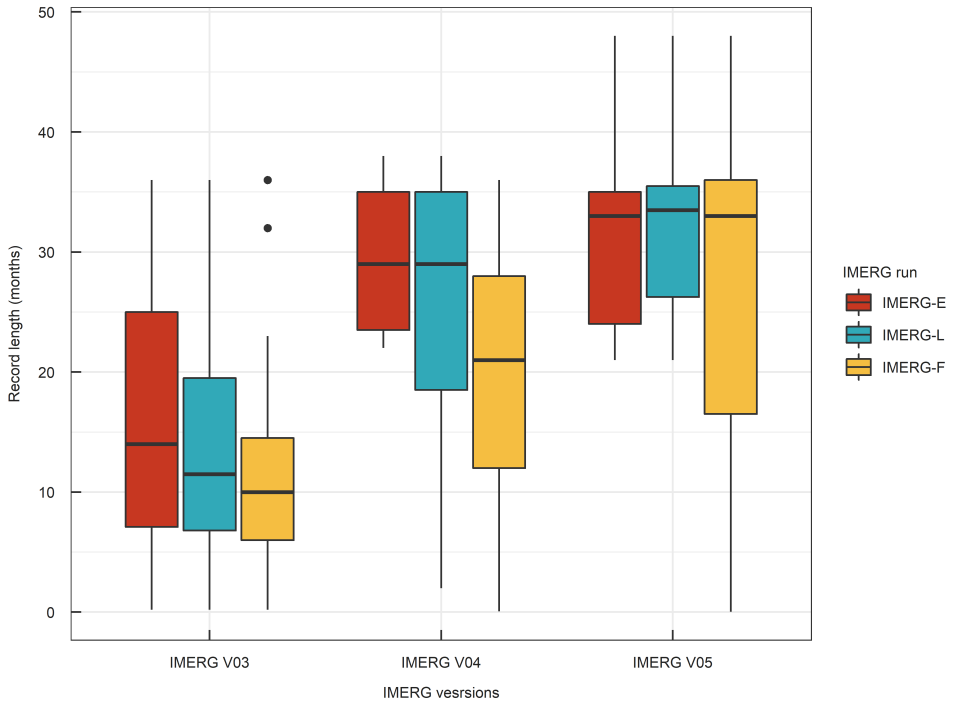


Figure 3.3: Spatial and temporal scales of IMERG validation studies per continent. The numbers inside the circles do not represent the actual number of studies, because a single study can have evaluated in multiple temporal scale, and has the possibility of counted more than once.

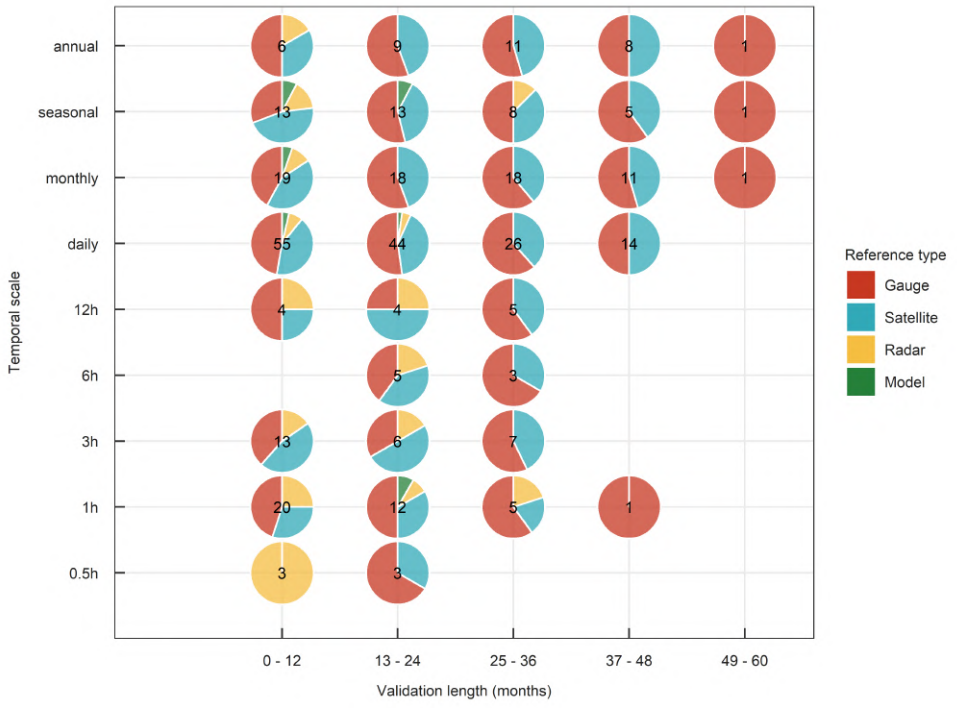
In terms of spatial scale, validation of IMERG data has been performed at $0.1° \times 0.1°$, $0.25° \times 0.25°$, $0.5° \times 0.5°$, $1° \times 1°$, $2.5° \times 2.5°$, or $3° \times 3°$ resolutions (Figure 3.3). The majority of validation studies were at $0.1° \times 0.1°$ and $0.25° \times 0.25°$ resolution. This could be a consequence of IMERG products' nominal resolution of $0.1° \times 0.1°$, and the fact that most gridded reference data sets used are typically available at $0.25° \times 0.25°$ (e.g., TRMM TMPA). We note that studies

that evaluated IMERG products using gauges (point vs. pixel-based method) were also considered as evaluated at $0.1^\circ \times 0.1^\circ$ resolutions; this also contributes to the higher number of studies on this resolution. Despite Asia being the subject of most of the studies, none of them evaluated IMERG products at resolutions coarser than $0.25^\circ \times 0.25^\circ$; the same holds for Africa, Europe, and South America. In contrast, North America (Tan et al., 2017) and at the global scale (Khan and Maggioni, 2019); despite a smaller number of studies, conducted IMERG evaluation studies at different spatial resolutions. Therefore, more validation studies at multiple spatiotemporal resolutions are needed to better understand and achieve a more in-depth analysis of the IMERG data set properties over different scales.

Generally, the validation period length of the studies is increasing with recent IMERG versions (Figure 3.4a). As expected, IMERG V03 has the shortest validation period (median around 12 months), whereas IMERG V05 has the most extended (median around 33 months). In terms of IMERG runs, IMERG-E and IMERG-L have similar validation lengths within the corresponding IMERG versions, unlike the IMERG-F, which has a relatively shorter period length. The main reason for this is that IMERG-F is available at 3.5 months of latency. Although IMERG V06 is available contemporary to the TRMM era (June 2000 - present), up to 2019, no study has evaluated this period. Furthermore, the length of the validation period was also reported on a monthly scale and classified into five ranges, each a multiple of 12-month duration (Figure 3.4b). Considering the short record of IMERG up through V05 (i.e., available from early 2014 onward), most of the studies' record length falls under the shortest range (0 – 12 months). Based on the database created, it is evident that the number of studies is inversely related to the validation period. Around 35% of the studies have a validation period length between 0 – 12 months, 31% between 13 – 24 months, 22% between 25 – 36 months, 11% between 37 – 48 months, and 1% between 49 – 60 months. It is interesting that studies using longer validation periods were often associated with coarser temporal resolution (daily and longer), whereas studies with shorter validation periods were associated with sub-daily temporal resolution. In terms of reference data, radar- and model-based data sets are mainly used for short evaluation periods, whereas gauge- and satellite-based data sets are used for long evaluation periods.



(a)



(b)

Figure 3.4: (a) Validation length by IMERG versions and runs, (b) Temporal scale versus validation length per reference type of IMERG validation studies. The numbers inside the circles do not represent the actual number of studies, because a single study can have evaluated in multiple references, and has the possibility of counted more than once.

The types of reference data sets, and thus their accuracy, play a significant role in the evaluation results. Generally, ground (gauge- and/or radar-based) data sets are preferred as the source of reference data to assess the accuracy of GPM precipitation. However, the low density of ground stations globally forces the scientific community to rely on different sources for evaluation, namely satellite, model, reanalysis, and merged products (Figure 3.4b). The most common satellite-based products for IMERG comparison are TMPA, GSMaP, and CMORPH. As GPM is the immediate successor mission of TRMM, most studies compare IMERG using TMPA (i.e., IMERG vs. TMPA) and compare their individual performance versus gauge precipitation (i.e., IMERG vs. TMPA vs. gauge). Only a couple of reference data sets came from different sources other than satellites, namely ERA-Interim (reanalysis) and Weather Research and Forecasting (WRF) (model) when ground observations were unavailable. In addition, radar precipitation data as a reference to evaluate the IMERG is very limited in number as well. The expensive installation and maintenance cost of radars could be the main reason for such fewer studies. Radar evaluation of IMERG is mainly dominated by developed countries/continents like the United States of America (USA) and Europe. The Multi-Radar/Multi-Sensor (MRMS) is the dominant radar product in the USA. Furthermore, all the IMERG evaluation studies against radar assessed the IMERG products on a sub-daily scale.

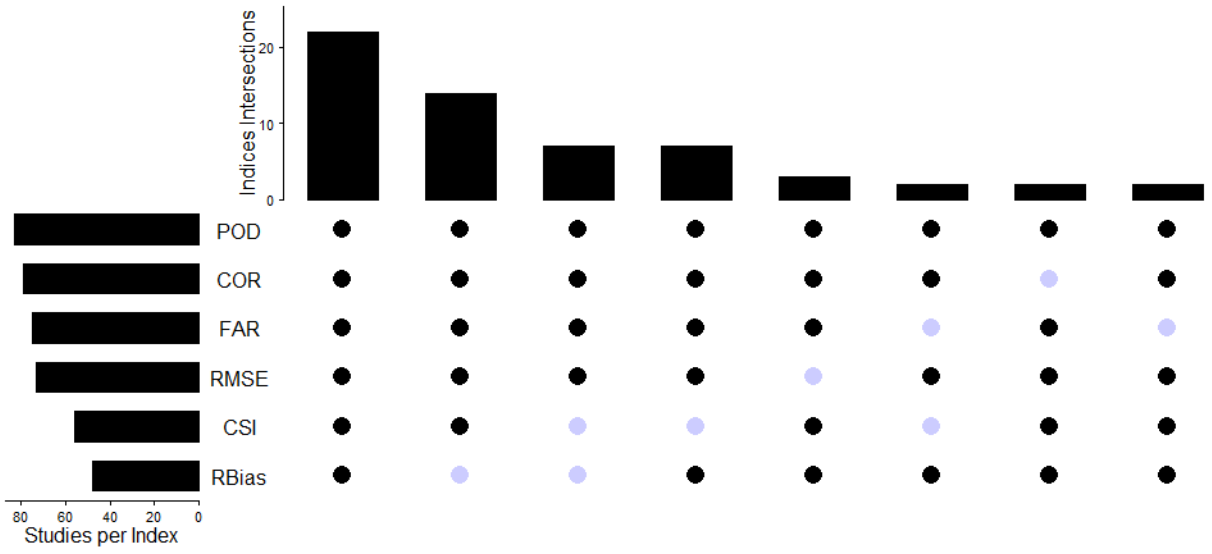


Figure 3.5: Statistical metrics used for evaluation of IMERG products (POD = Probability of Detection, COR = Coefficient of Correlation, FAR = False Alarm Ratio, RMSE = Root Mean Square Error, CSI = Critical Success Index, RBias = Relative Bias).

Generally, various statistical metrics were often used to validate satellite precipitation with the reference data sets. The statistical measures used for the IMERG validation can be categorized into two main types; i) Volumetric and ii) Categorical metrics.

The UpSet plot (Figure 3.5) shows the number of studies that employed different statistical metrics combinations. The UpSet plot is a standard format to depict the intersection of sets when the number sets are more than three or four, and it was developed by Lex and Gehlenborg (2014). Each bar represents a unique combination of the metrics, and underneath the table shows their combination types. The empty cells (light grey) indicate the particular metric is not part of the intersection, whereas the filled cells (black) indicate it participates in the intersection. From the left to right direction, the number of studies is decreasing. POD, COR, FAR, RMSE, CSI, and RBias are the most frequently used metrics combination (n = 22), followed by POD, COR, FAR, RMSE, CSI (n = 14), and POD, COR, FAR, RMSE (n = 7). Other combinations appear in a very low number of studies, mostly fewer than 3. In addition, the small barplot on the left side represents the unconditional (without combination) metrics used for the studies. From the top, POD is the most reported metrics used in 82 studies, followed by COR (78), FAR (74), and RMSE (72). On the other hand, CSI and RBias fall between $40 - 60$.

3.4 IMERG's performance by continents

3.4.1 Asia

Asia is characterized by diverse climate patterns and a variety of regional topography. Based on the studies assessing IMERG in China, IMERG captures the overall spatiotemporal behavior of precipitation over the country. However, there are substantial differences in local climatic conditions, which can affect IMERG-F V05 performance (Chen et al., 2018). Geographically, IMERG-F tends to be more accurate in the lower latitudes than mid/high latitudes of China (Chen and Li, 2016). IMERG-E, -L V05 showed more accurate estimates of high-intensity precipitation over wet/humid regions compared to low-intensity precipitation over dry regions (Wu et al., 2018). This is also supported by other studies as well (e.g., Fang et al. 2019; Jiang and Bauer-Gottwein

2019; Wei et al. 2018; Asong et al. 2017). With regards to the diurnal variation, IMERG-F V06 performs poorly between 06:00 and 10:00 UTC (Xu et al., 2019b). Furthermore, on average, IMERG-F products at daily and sub-daily time steps did not perform as well as in the monthly time scale (Xu et al., 2019b,a; Chen et al., 2018; Wang et al., 2019b). When assessing IMERG-F at the seasonal scale, its performance is worse during winter (Chen and Li, 2016). However, it is clear that IMERG-F V03 performance improved compared to TMPA, still leaving substantial room for further improvement over China (Tang et al., 2016a).

3.4.1.1 Eastern China

In eastern China (humid/semi-humid climate), although IMERG-F V04 has good agreement when measuring light precipitation ($<$8 mm/day), it tends to overestimate high precipitation rates ($>$64 mm/day), and underestimate precipitation rates between 8 – 64 mm/day. Furthermore, IMERG-F V04 showed an overestimation of up to 17.9% for the 99th percentile of precipitation on wet days (RR99P) and 11.5% relative bias for the R20TOT index (total precipitation sum of daily precipitation that is more than 20 mm) (Ning et al., 2017). The intensity of precipitation at which IMERG over-/underestimates varies in space according to the topography. For instance, over the Huang-Huai-Hai Plain (eastern coastal region of China), the range of IMERG-F V05 overestimation of precipitation rates lies between 2 – 50 mm/day, and there is an underestimation of heavier precipitation rates ($>$50 mm/day) (Xu et al., 2019a). On the other hand, over the Huaihe river basin, IMERG-F V05 overestimates precipitation between 0.5 – 25 mm/day and underestimates it above and below that range (Chen et al., 2018).

3.4.1.2 Southern China

In southern China, IMERG-F tends to overestimate the light rainfall, and underestimate the high rainfall. Compared to TMPA, however, the underestimation of light rainfall is lower, and IMERG better captures the probability density function (PDF) and the inter-annual precipitation variability, especially over the lower Mekong river basin (Wang et al., 2017a). In addition, IMERG-F V03, V04 and V05 consistently overestimate precipitation over mountains and underestimates

it in coastal regions of Guangdong Province. The positive hit bias and false detection of moderate to heavy precipitation events are attributed to the above-described overestimation (Wang et al., 2019b). IMERG's performance over high elevation (mountains) areas during dry seasons need further improvement. Nonetheless, IMERG estimates denote an improvement over TMPA estimates in terms of light/heavy precipitation detection and hit bias (Wang et al., 2019b).

3.4.1.3 Northwestern China

In the northwest region (arid climate), IMERG-F V05 did not show significant improvement compared to TMPA. IMERG underestimated precipitation at low altitudes and overestimated it at high elevations. On the seasonal scale, IMERG performs better in summer than in winter. Additionally, IMERG seems to suffer from poor detection capability of light rainfall, i.e., 0 − 2 mm/day. However, it better performs in moderate (>5 − 10 mm/day) and heavy precipitation events (>25 mm/day) (Wang et al., 2019c). Over the Tianshan mountain, IMERG-F V06 did not show significant improvements compared to the IMERG-F V05 (Anjum et al., 2018). IMERG products perform better in the eastern region compared to the western (eg., Boertala Valley, Yili Valley, and West Tianshan). Overall, IMERG products are reliable enough to be used in precipitation trend analysis over the Tianshan mountain, but caution should be taken for the western regions (Anjum et al., 2018). Compared to TMPA, IMERG-E, -F V05 products have significant discrepancies over high latitudes and thus can be considered less reliable for the Tianshan mountain (Yang et al., 2019). Overall, IMERG has almost similar performance as TMPA and is significantly affected by the northwestern region's topography and aridity.

3.4.1.4 Tibetan Plateau

IMERG-F V03 and V06 validation studies over the Tibetan Plateau (Ma et al., 2018; Xu et al., 2017; Lu and Yong, 2018) show an overestimation of total precipitation in the southwest regions and an underestimation in the northeast regions. Moreover, a high correlation was observed in the northeast and southeast regions, while a low correlation was reported in the southern regions of the Tibetan Plateau (COR <0.40). Most likely, the complex topography of the Himalayan moun-

tains can be the reason for the low correlation (Ma et al., 2018). Additional findings confirm that the accuracy of IMERG-F V05 decreases as elevation increases, which indicates the direct effect of elevation on IMERG products (Wang et al., 2019b). Furthermore, the detection of light precipitation is particularly affected at elevations above 4500 m (Xu et al., 2017). IMERG-F V05 tends to overestimate trace or light precipitation (0 – 1 mm/day) and underestimates highly intense precipitation (>50 mm/day) (Wang et al., 2019b). Overestimation could be attributed to the evaporation of light precipitation in the atmosphere before reaching the surface (Wang et al., 2019b). In terms of IMERG versions, IMERG-F V04 did not show significant improvements to its predecessors (i.e., IMERG-F V03). Compared to the IMERG-F V03, IMERG-F V04 showed significant underestimation of daily precipitation's annual average with a relative bias of -60.91% over the Tibetan plateau, which becomes more profound in the winter (-72.33%) (Zhao et al., 2018). Although IMERG-F V03 outperformed IMERG-F V04 over the Tibetan Plateau, yet both products underestimate winter precipitation with relative bias of -6.47% and -70.62% respectively (Wei et al., 2018). Despite the fact that IMERG-F V06 captures the average distribution of total precipitation in space, detection of light rainfall, winter snowfall, and detection of precipitation at high elevations remain major challenges over the Tibetan Plateau (Lu and Yong, 2018).

3.4.1.5 India and Pakistan

Over India, IMERG-F V03 showed a noticeable improvement over GSMaP and TMPA, capturing southwest monsoon mean rainfall and its variability (Prakash et al., 2015). Nevertheless, there was a higher total negative bias and hit bias in IMERG-F V03 and V04 in mountainous regions such as the Himalaya foothills and Western Ghats and underestimation of northeastern orographic precipitation (Prakash et al., 2018b, 2015, 2018a). A large fraction of the FAR and hit bias over the south peninsula can be attributed to the Western Ghats (leeward side or rain shadow region). In contrast, IMERG-E performs better over plains and coastal regions (Singh et al., 2019), significantly improving TMPA systematic error dependency with topography. In addition, there were improvements in rainfall estimates of varying intensities across different topographies over most river basins, except northwest semi-arid basins (Beria et al., 2017). IMERG-F V03 showed reasonable improvement over TMPA capturing heavy precipitation events during the summer monsoon

season, especially over Himalaya and northwest India (Prakash et al., 2015). IMERG-F appears to overestimate precipitation in high elevation zones in Pakistan and slightly underestimate it in semi-arid regions. Additionally, it tended to overestimate pre-monsoon and monsoon precipitation but underestimated post-monsoon and winter precipitation (Rahman et al., 2018). Despite overestimation of light precipitation (0 – 1 mm/day) and underestimation of moderate (1 – 20 mm/day) to heavy rainfall (>20 mm/day) over the north hill highlands of Pakistan, IMERG-F V04 represents the spatial variation of precipitation better than TMPA (Anjum et al., 2018). Overall, in southeast Asia, the uncertainties about orographic precipitation remain a considerable challenge.

3.4.1.6 Eastern Asia

In Eastern Asian countries, the performance of IMERG varies with space. For instance, in Japan and Korea, IMERG-F's V03 average POD (0.69) for convective rainfall over mountains and coastal regions during pre-monsoon and monsoon season is 8% better than TMPA (POD = 0.61) (Kim et al., 2017). In addition, it outperforms TMPA in both pre- and post-monsoon precipitation, as well as in terms of spatial precipitation patterns. However, contrary to the expectations, TMPA outperformed IMERG-F V04 precipitation estimates at daily and monthly scales both for total and heavy precipitation over Myanmar (Yuan et al., 2017). A poor detection and estimation skill of IMERG-F V04 is found both in light and heavy precipitation with a significant underestimation of total precipitation in the Chindwin river basin of Myanmar (Yuan et al., 2017). Similarly, over the same region, 3B42RT shown the best estimates followed by IMERG-F V05, whereas the near-real-time products (IMERG-E V05 and IMERG-L V05) have the lowest quality (Yuan et al., 2019). Furthermore, despite IMERG-F's V04 better detection of daily precipitation, overall, it did not show significant improvement compared to TMPA over Singapore (Tan and Duan, 2017). In Taiwan, IMERG-F V05 can reproduce different precipitation characteristics like the seasonal variation and temporal bimodal peak of annual precipitation. By validating IMERG with gauge data, it appears again that spatial discrepancies and underestimation bias are higher over mountain regions than plains. Furthermore, in terms of seasonality, IMERG shows poor performance in winter (Huang et al., 2018). Despite underestimating heavy precipitation, IMERG-F V05 agrees with the APHRODITE data product in Japan, the Philippines, and Nepal (Sunilkumar et al., 2019).

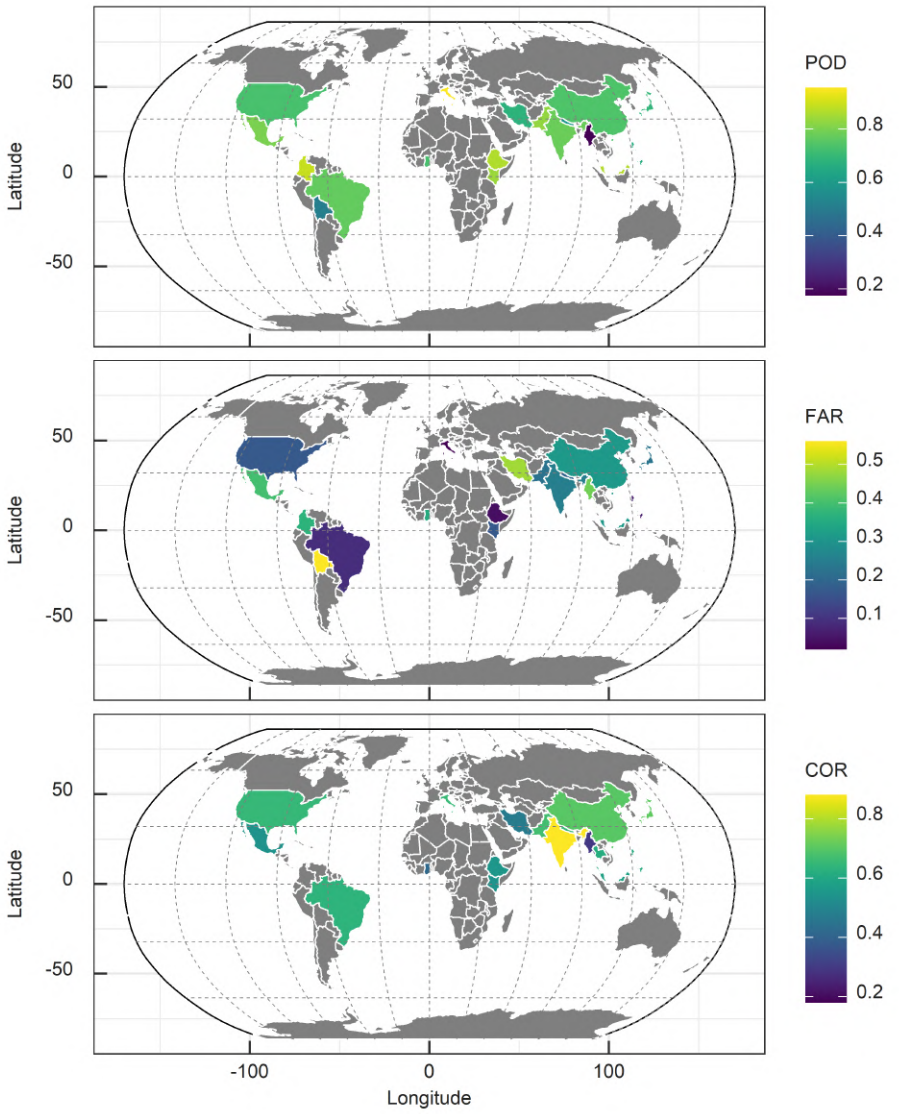


Figure 3.6: Spatial distribution of POD, FAR, and correlation values of IMERG evaluation studies across the countries at daily scale. The matrices were calculated based on the daily scale, and median values were considered in case of more than one study available for the same country (e.g., China, Iran, India, USA etc.).

3.4.1.7 Western Asia

In western Asia, IMERG-F V03 had acceptable performance compared to TMPA and European Centre for Medium Range Weather Forecasts (ECMWF). Accordingly, it was shown that it could be used as a substitute for ground observations in regions lacking observational precipitation over Iran (Sharifi et al., 2016; Khodadoust Siuki et al., 2017; Maghsood et al., 2020). In the study by

Mahmoud et al. (2019), IMERG-F V03 showed good agreement with ground data in the southern, middle, and northern parts of the UAE. However, in the eastern and northeastern parts of the UAE, characterized by mountainous topography and coastal areas, there were errors in detection and estimation. IMERG-F outperformed IMERG-E and IMERG-L in terms of POD, bias, Mean Absolute Error (MAE), and RMSE. However, against expectations, IMERG-E outperformed IMERG-F in terms of correlation. In Oman, it was observed that as precipitation intensity increases, so does IMERG-E, -L, -F V04 underestimation. The Mean Difference (MD) was -3.11 for 2.5-10 mm/day, -12.30 for 10-50 mm/day, and -50.74 for $>$ 50 mm/day intensity classes (Shawky et al., 2019). In Saudi Arabia, IMERG-F V05 also outperformed IMERG-E and IMERG-L, and its uncertainty was higher in the southern and northern parts of the country (Mahmoud et al., 2018).

3.4.2 North America

North America has the second-highest number of IMERG validation studies. It was shown that compared to the Multi-Radar/Multi-Sensor (MRMS) precipitation data set, IMERG-F V03 overestimates drizzle (light rainfall) and underestimates heavy rainfall over CONUS (Contiguous United States) (Tan et al., 2016). Furthermore, compared to TMPA, IMERG-L V03 improved the missed rain bias, and false hits over the same region (Gebregiorgis et al., 2018). Tan et al. (2017) evaluated the IMERG-F V03 product against the MRMS as a function of spatial-temporal scale over the southern United States. They found an enhancement of performance with the increase in the spatial-temporal scale, both capturing the rain occurrence and its estimation. Over the central United States, IMERG-E and -L V05 hourly products show close agreement and higher correlation with NCEP products when the temperature (i.e., hourly) exceeds 280 K (Zhang et al., 2018). IMERG-F V05 performs noticeably well in representing the spatial variability of storms, despite some errors in high-intensity precipitation regions (storm core) (Omranian et al., 2018). In terms of the diurnal and semi-diurnal cycle, IMERG-F V04 agrees with the reference precipitation. However, it overestimates the normalized amplitude over the central US and underestimates it in the western and eastern US mountainous regions (Kirstetter et al., 2018). It also shows substantial differences in the peak of diurnal precipitation for convective and stratiform precipitation of mesoscale convective systems over the Great Plains (Kirstetter et al., 2018). Despite the overesti-

mation of heavy rain events and low performance in mountainous regions, IMERG-F V03 is satisfactory reproducing the spatial distribution and precipitation amount over Canada. In addition, its performance is relatively better in the continental semiarid region than in the humid regions (Asong et al., 2017). In Mexico, IMERG-F V03 underestimates heavy precipitation at daily and hourly scales, but it reduces the error over high-elevation terrains (Mayor et al., 2017). In terms of snowfall, IMERG-F V04 underestimates precipitation compared to SNOTEL with a relative bias between -71% and -82% over the western mountain regions. Furthermore, the discrepancy between IMERG-F V04 and SNOTEL observations increases as daily temperature increases from -14°C and approaches 0°C. Concerning precipitation intensity, the IMERG products have better performance between 0 – 5 mm/day but show significant underestimation at $>$10 mm/day (Wen et al., 2016). Similarly, Sadeghi et al. (2019) reported the IMERG-F's V04 underestimation of snow accumulation, although it detects the snowfall events comparatively better than the MRMS. It may be concluded that IMERG's current snowfall estimation performance is unreliable for hydrological and climatological applications.

3.4.3 Europe

Europe is the second to last continent with the least published IMERG validation studies (6%). The first study validated IMERG-F V03 over the Netherlands using one year of the data set developed by Gaona et al. (2016). It concluded that IMERG could reproduce the spatiotemporal distribution of precipitation over the nation despite a very small (2%) underestimation across all resolutions (i.e., 0.5h, daily, monthly, and annually). Furthermore, IMERG-F V03 has a small relative bias on the 30 min (-1.51%) and daily (-1.49%) scale, highlighting IMERG's potential in hydrological applications (Gaona et al., 2017). When comparing IMERG-F V05 to radar-based precipitation data, IMERG showed a significant overestimation of precipitation, especially during winter over Germany with low correlation ($<$0.4), POD (0.38), CSI (0.28), and high FAR (0.48). Furthermore, it showed difficulties in reproducing spatial variability across Germany's diverse topography (Ramsauer et al., 2018). The performance of IMERG-E V04, IMERG-F V05, and IMERG-F V06 was influenced by complex terrain and had problems capturing precipitation over mountainous regions (Chiaravalloti et al., 2018; Ramsauer et al., 2018; Navarro et al., 2019). Additionally, IMERG-F V06

performed better during summer than in the winter (Navarro et al., 2019) and on the monthly scale than in the daily and sub-daily ones (Gaona et al., 2016; Ramsauer et al., 2018).

3.4.4 South America

In South America, IMERG-F V06 effectively represents the spatial pattern of precipitation and shows reasonably better performance than TMPA throughout Brazil (Rozante et al., 2018). In the northeast coast of Brazil, which is characterized by warm rain events, IMERG-F V05 showed significant errors and underestimated daily precipitation. The above could be attributed to the inability of GPM sensors to detect orographically forced warm-rain processes (Gadelha et al., 2019). Large biases appear in the North and Central-west regions, associated with the sparse density of gauges. In the Central Plateau of Brazil, IMERG-F V05 reproduced annual and monthly precipitation better than daily precipitation. Furthermore, IMERG-F V05 exhibited strong seasonal variability as numerous errors, and estimation difficulties occur with low and sparse dry season precipitation (Salles et al., 2019). For diurnal precipitation, IMERG-F V03 overestimated the frequency of heavy precipitation over the Negro, Solimões, and Amazon rivers and underestimated dry season precipitation compared to S-band weather radar measurements. The above was attributed to IMERG's difficulties in detecting isolated convective cells and the poor calibration over water surfaces (Oliveira et al., 2016). In the high Andes, IMERG-F V03 efficiently captured rainfall intensity but showed substantial discrepancies with gauge-based observations along the dry Peruvian coastline (Manz et al., 2017).

3.4.5 Africa

Africa has the least IMERG validation studies, with only two works so far (2019). The lack of access to reliable observational data sets could be the reason for such a small number of validation studies. These analyses agree that the performance of IMERG-F V03 and V04 varies with the season, climate, and topography (Dezfuli et al., 2017; Sahlu et al., 2016). For instance, IMERG performs better over Eastern and humid regions of Africa than Southern Sahel, and the discrepancies between IMERG-F V04 and TMPA were higher over mountainous regions (Dezfuli et al., 2017).

In addition, despite a slightly superior performance of IMERG-F V03 over the Blue Nile basin, its overall performance was similar to CMORPH (Sahlu et al., 2016). Another finding suggests that IMERG-F V03 detection capability decreases with increasing precipitation, emphasizing the need to improve IMERG's detection capability of heavy precipitation (Sahlu et al., 2016).

3.4.6 Spatial distribution of POD, FAR and COR

The spatial distribution of POD, FAR, and correlation values reported among the reviewed studies are shown in Figure 3.6. These values are based on a comparison of IMERG daily precipitation with the corresponding gauge observation on a daily scale. Only the correlation, FAR, and POD values of the most recent version and IMERG final run (IMERG-F) were considered when a single study evaluated successive versions (IMERG V03, IMERG V04, etc.) and various IMERG runs (IMERG-E, -L, and -F) in the same study. More or less, the majority of the studies reported good detection skills (POD $>$ 0.6), fewer false alarms (FAR $<$ 0.5), and reasonable agreement with reference data sets (COR $>$ 0.5). Myanmar is the only country that reported very poor values for both detection (POD = 0.17, FAR = 0.45) and estimation (COR = 0.29) of precipitation. The above is attributed to IMERG's poor performance in n detecting and estimating light and heavy precipitation along the Chindwin river basin of Myanmar (Yuan et al., 2017). However, the study region is characterized by very sparse observation stations, and the evaluation results are based on only four gauge stations.

In terms of continents, IMERG has good detection skills over Africa with POD ranging from 0.73 to 0.84, and FAR up to 0.35, but moderate correlation ranging from 0.42 to 0.54. Similarly, North America (POD = 0.73 – 0.8, FAR = 0.17) and South America (POD = 0.51 – 0.89, FAR = 0.12 – 0.37) revealed a good detection skill, except Bolivia which reports the highest FAR of 0.56. In addition, both the continents reported a reasonable correlation (0.54 – 0.65) of IMERG with the reference data sets. High spatial variation in both POD and correlation is observed in Asia, varying from the poor results over Myanmar to high values over India. In addition, the East Asian countries such as Thailand, the Philippines, Nepal, and Malaysia have correlation values ranging from 0.5 - 0.69, whereas Japan, India, and China have COR $>$ 0.7. In western Asian

countries, Iran and Pakistan present a medium correlation (0.47 and 0.67, respectively). On the other hand, Pakistan and Malaysia report slightly better detection skills (> 0.8) than India (0.78) and China (0.74). Detection skills over Japan, Iran, and Nepal vary from 0.59 to 0.69. In terms of false detection, Iran reported a FAR of 0.51, which could be attributed to the prevailing arid condition, and evaporation of light precipitation before reaching the ground's surface. Moreover, based on the POD, FAR, and COR values among the studies, it can be concluded that IMERG shows its good detection skills throughout various climatic and topographic conditions and has better agreement with the reference observations in overall precipitation estimation.

3.5 IMERG performance by run (IMERG-E, -L, and -F) types

Generally, it is believed that the IMERG-F, which has gauge correction at the monthly scale, has superior performance compared to both -E and -L run products as shown in China (Guo et al., 2016; Tang et al., 2016a), East Asia (Kim et al., 2017), Austria (temperate climate) (Sungmin et al., 2017), Saudi Arabia (Mahmoud et al., 2018), Pakistan (Anjum et al., 2018), and Italy (Chiaravalloti et al., 2018). However, few studies contradicted this and reported either no significant improvement of IMERG-F run or even outperformance by IMERG early run at least in some aspect (i.e., POD, correlation, extreme events). For instance, Maghsood et al. (2020) reported no significant IMERG-F improvement in POD and FAR over Iran at the daily scale, but it does at the monthly scale. In the same study, they report that IMERG-E and -L products are more suitable for extreme precipitation. Moreover, Shawky et al. (2019) found no significant improvements of IMERG-F over IMERG-E in the arid environment of Oman. Mahmoud et al. (2019) reported that IMERG-E outperformed the IMERG-F products in terms of correlation over the UAE. However, in terms of error and bias, IMERG-F outperformed both the early and late products (on average MAE and RMSE decreased by 10% and 11%, and bias from 1.1% to 0.4%). Tan and Santo (2018) observed similar results over Malaysia. This counter-intuitive behavior could be attributed to the sparse gauge availability and consequently the GPCC calibration of IMERG-F over those regions. Moreover, it should be noted that the IMERG-E and IMERG-L runs serve for near-real-time applications (e.g., flood, drought, and crop forecasting), whereas the IMERG-F is best intended for research purposes.

3.6 IMERG performance by versions (IMERG V03, IMERG V04, IMERG V05, and IMERG V06)

The IMERG algorithm and its version are episodically updated, yet only a few studies compared successive IMERG versions (Figure 3.7). Xu et al. (2019c) compared the IMERG V04 and IMERG V05 over mainland China and found that IMERG V05 estimates precipitation better except for false precipitation. Wang et al. (2019a) found that IMERG V05 possesses significant enhancements in precipitation estimation compared to the IMERG V03 and V04 over the Guangdong Province, China. Similarly, Satgé et al. (2018) reported IMERG V05's expected improvement compared to its predecessors (i.e., IMERG V03 and V04) throughout Pakistan except for the extreme arid region where the IMERG V04 had the best performance. In addition, Wang et al. (2018) compared the IMERG V03, V04, and V05 at the global level. They found that IMERG V05 significantly improved over the previous IMERG V03 and V04, and the improvements are mainly observed in the estimation of mean oceanic precipitation.

On the other hand, Anjum et al. (2019) revealed no significant improvement of IMERG V06 over IMERG V05 in the Tianshan Mountains, China. Likewise, Derin et al. (2019) reported that when capturing light or heavy precipitation IMERG V06 failed to outperform IMERG V05 over the mountain regions. In addition, IMERG V04 did not shown paramount enhancement compared to its predecessor IMERG V03 (Satgé et al., 2018). The above claim is supported by studies in China (Zhao et al., 2018), and the Tibetean Plateau and Weihe River Basin (Wei et al., 2018). Despite successive IMERG versions aiming towards more accurate estimation of precipitation, its performance varies with the precipitation type and topographical features.

3.7 IMERG's performance over ocean

Out of the 101 peer review articles in our database, only three studies evaluated the performance of IMERG over the ocean, making it difficult to draw any definitive conclusions. The lack of standard observational data sets could be the possible reason for such a low number of studies

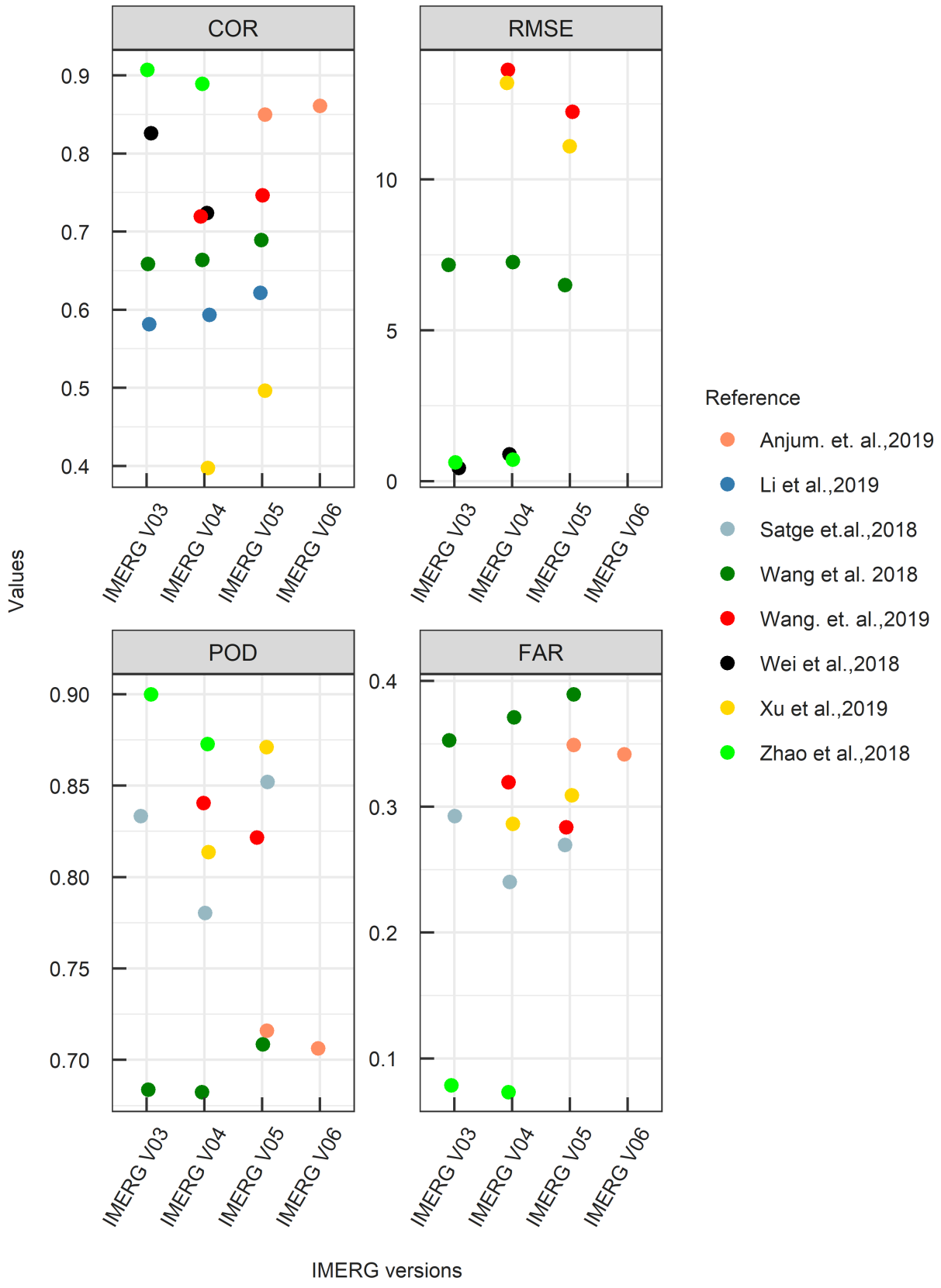


Figure 3.7: Performance of IMERG successive versions (IMERG V03, V04, V05 and V06) in term of COR, RMSE, POD, and FAR values across the countries at daily scale.

over the ocean. Khan and Maggioni (2019) assessed the performance of IMERG-E, -L, -F V05 daily products over the ocean using Ocean-Rain and satellite-based radar products as validation targets. IMERG was found to underestimate Ocean-Rain measurements significantly. However, 80% of the time, IMERG products detect rain. Prakash et al. (2018a) assessed the performance of IMERG-F V04 over the north Indian ocean using hourly moored buoy observations. The authors reported that IMERG performed better over the Arabian sea than over the Bay of Bengal. Despite low errors and good detection capability, IMERG suffered from a high FAR and overestimated rainfall (especially for light and extreme precipitation). Considering IMERG's resolution (0.1° x 0.1°), it is more probable that small-scale showers can occur at any part of the IMERG grid, but not over the exact buoys locations, which are sparser than the gauges over land. Thus, leading to the apparent overestimation of precipitation. Besides, the evaporation of light precipitation in the atmosphere could be another possible reason. Wang et al. (2018) reported that mean oceanic precipitation of IMERG-F V04 and IMERG-F V05 shows significant improvement over IMERG-F V03, and their estimates are close to the ones by Multi-Source Weighted-Ensemble Precipitation (MSWEP) (Beck et al., 2017) and Global Precipitation Climatology Project (GPCP) (Adler et al., 2003). Even though observational data sets are scarce over the ocean, other global products such as GPCP and MSWEP could be used as a reference for validation, as shown above.

3.8 IMERG's performance in representing extremes

The high spatiotemporal resolution, homogeneous global coverage, and near-real-time availability of satellite-based precipitation data are essential to understand extreme events better. Various recent studies are evaluating IMERG's performance for multiple extreme event applications (Table 3.1). Omranian et al. (2018) evaluated the capability of IMERG-F V05 to reproduce the precipitation of hurricane Harvey in Texas. They found that IMERG captured the storm with POD $>$0.82 and FAR $<$0.2, and precipitation spatial variability with 62% accuracy. Nevertheless, several aspects appear to need improvements, such as underestimation over the coastal region, overestimation in the high-intensity region, and discrepancies between observational data and IMERG precipitation, especially near the storm's center. Huang et al. (2019) analyzed the performance of IMERG-E and -F V05 products to capture six major typhoons over southern China

during 2016 and 2017. They found that IMERG captured precipitation spatial variability and areal hourly precipitation. Furl et al. (2018) also revealed a significant underestimation in capturing extreme storms. Interestingly, they observed that the IMERG-E and -L V03 products have lesser underestimation than the IMERG-F V03 products, which can be attributed to the -F run's gauge correction. Wen et al. (2018) evaluated the performance of IMERG-L V04 in terms of Atmospheric Rivers over the western USA. Despite good detection of heavy precipitation events, IMERG significantly underestimates (-40%) the total precipitation volume. Additionally, Fang et al. (2019) found good agreement between IMERG-F V05 and gauge observations regarding extreme precipitation spatial patterns throughout China. IMERG's performance was more consistent over the southeast (humid) than over the northwest (arid) regions. In the same study, an underestimation of extreme precipitation events was reported, implying that although IMERG reproduces the spatial precipitation pattern and volume better than TMPA, limitations in detecting extreme events remain. Zhang et al. (2019) evaluated the performance of IMERG-F V05 products in capturing a 60-year return period extreme precipitation storm over southern China. The authors reported that IMERG products significantly underestimated the event. However, IMERG's performance was reported to vary at different intensities; i.e., IMERG performs poorly when rainfall intensity is above 17 mm/hr and best when the intensity is below 5 mm/hr.

Table 3.1: List of the studies evaluated the IMERG products in extreme events.

Location	Events	Data	Period	COR	Bias	RMSE	Reference
USA	Hurricane	IMERG-F	Aug-2017	0.61	– –	– –	(Omranian et al., 2018)
China	Typhoon	IMERG-E	2016-2017	0.61	8.38	44.97	(Huang et al., 2019)
		IMERG-F		0.57	13.50	47.50	
China	>90th percentile	IMERG-F	2014-2017	0.63	-22.82	23.52	(Fang et al., 2019)
China	Storm	IMERG-F	May-2017	0.70	-58.77	9.70	(Zhang et al., 2019)
Global	>90th percentile	IMERG-E	2014-2017	– –	-3.18	– –	(Mazzoglio et al., 2019)
		IMERG-L		– –	-3.1	– –	

3.9 IMERG's performance in hydrological applications

A large number of studies have evaluated the performance of IMERG products in terms of simulating streamflow (Table 3.2). However, as the studies used different hydrological models, calibration and validation methods, basin locations, and climatic conditions, the direct comparison of the re-

sults cannot be conclusive. Nevertheless, we compiled the Nash–Sutcliffe coefficient of efficiency (NSE) and statistical bias indices from these studies (Figure 3.8) to summarise their results. Our main objective here is to highlight the IMERG performance and its different runs (IMERG-E, -L, and -F) in each study separately, rather than inter-comparing different studies results with each other.

Table 3.2: List of the studies evaluated the hydrological applications of IMERG products. (CREST = Coupled Routing and Excess Storage, VIC = Variable Infiltration Capacity, XAJ = Xinanjiang, MGB-IPH = Large Basins Model and Institute of Hydraulic Research, SWAT = Soil and Water Assessment Tool, GXAJ = Grid-based Xinanjiang hydrological model.)

Basin	Location	Model	Data	Period	NSE	Bias	Reference
Ganjiang	China	CREST	IMERG-F	May2014-Sep2014	0.77	−14.09	(Tang et al., 2016b)
Beijiang	China	VIC	IMERG-E	Apr2015-Dec2015	0.34	28.48	(Wang et al., 2017b)
			IMERG-L		0.28	28.5	
			IMERG-F		0.74	10	
Chindwin	Myanmar	XAJ	IMERG-F	Apr2014-Dec2014	0.65 to 0.72	−23.5 to −28.7	(Yuan et al., 2017)
Mahanadi	India	VIC	IMERG-F	Apr2014-Dec2014	0.64	41.4	(Beria et al., 2017)
Amazon	Peru-Ecurdor	MGB-IPH	IMERG-F	Mar2014-June2015	−24.21 to −0.9	– –	(Zubieta et al., 2017)
Mekong	China	XAJ	IMERG-F	May-Oct2015-May-Oct2016	0.53	– –	(He et al., 2017)
Ganjiang	China	CREST	IMERG-F	June2014-Sep2014	0.7	−12.6	(Li et al., 2017)
Yellow	China	VIC	IMERG-F	Jan2015-Dec2015	0.62	−7.2	(Lu and Yong, 2018)
Mishui	China	XAJ	IMERG-E	Apr2014-Dec2015	0.73	−19.52	(Jiang et al., 2018)
			IMERG-L		0.71	−25.23	
			IMERG-F		0.81	−6.53	
Kelantan	Malaysia	SWAT	IMERG-E	Mar2014-Dec2016	0.7	−27.6	(Tan et al., 2018)
			IMERG-L		0.66	−36.3	
			IMERG-F		0.71	−5.3	
Chindwin	Myanmar	GXAJ	IMERG-F	Mar2014-Dec2016	0.84	−18.5	(Yuan et al., 2019)
Huaihe	China	VIC	IMERG-E	Apr2014-Dec2015	0.18	−39.91	(Su et al., 2019)
			IMERG-L		0.16	−43.95	
			IMERG-F		0.64	−16.51	
Nanliu	China	XAJ	IMERG-F	Mar2014-Dec2016	0.28	−7.83	(Liang et al., 2019)
			IMERG-E		0.29	−59.49	
Mun-chi	Thailand	VIC	IMERG-F	Apr2014-Mar2017	−0.98	– –	(Li et al., 2019)

Li et al. (2017) found more promising results when the hydrological model parameters were calibrated by gauge and RQPE (Radar-corrected Quantitative Precipitation Estimation) compared to the IMERG-F V04 data set. Furthermore, when the model was calibrated using raw IMERG data, the results were quite problematic over the Ganjiang river basin. In another study, Wang et al. (2017b) evaluated both near real-time and post real-time V03 products using the VIC (Variable Infiltration Capacity) model over the Beijiang River Basin. They found promising results for IMERG-F, while near real-time products showed poor performance (NSE < 0.35). Nonetheless, both products showed reliable flood forecasting results and thus could be considered useful for

such applications. The promising performance of IMERG-F V03 products was reported over the Mekong River Basin, suggesting its use for similar mountainous basins He et al. (2017). Yuan et al. (2017) observed the propagation of IMERG-F V04 error through the Xinanjiang model, which led to significant underestimation of streamflow over the Chindwin river basin in Myanmar; in this context, TMPA showed better results. The evaluation of IMERG-E V05 and IMERG-F V05 products over the Nanliu River Basin in Tropical Humid Southern China showed that IMERG had poor performance estimating streamflow at a daily scale with NSE of $<$0.4 Liang et al. (2019). In the Upper Huaihe River Basin, IMERG-F V06 had better performance in flood simulations than IMERG-E V06 and IMERG-L V06 Su et al. (2019). Nevertheless, due to the significant underestimation of runoff (-16.51%), the authors noted that IMERG-F V05 products should be used cautiously. Yuan et al. (2019) reported significant improvements of IMERG-F V05 performance after model specific input calibration with NSE increasing from 0.66 to 0.84, and the relative bias decreasing from -32.3 to -18.5, enabling the replacement of TMPA in hydrological applications. Tan et al. (2018) found that IMERG-F V05 (NSE = 0.71 and relative bias of -5.3%) outperformed IMERG-E V05 (NSE = 0.70 and relative bias of -27.6%) and IMERG-L (NSE = 0.66 and relative bias of -36.3%) over the Kelantan river basin in Malaysia. Lu and Yong (2018) also reported IMERG-F V06's potential to estimate streamflow on a daily scale over the Yellow river basin of the Tibetan Plateau.

The number of studies validating only IMERG-F products was significantly higher than those including IMERG-E and IMERG-L. In general, IMERG-F outperformed both IMERG-E and IMERG-L products across different basins and hydrological models in terms of NSE, with values ranging between 0.60 – 0.78. This is considered highly acceptable for hydrological simulation (Moriasi et al., 2007). On the other hand, the IMERG-E and IMERG-L products show significant uncertainties between different basins and hydrological models. For example, most studies report NSE values below 0.35 in China, with only a few exceeding 0.65. Relative bias metrics follow a similar behavior to NSE, where IMERG-F outperforms the other products. IMERG-E's streamflow relative bias ranges between -59.4% and 28.4%, while IMERG-L is between -43.9% and 28%, and IMERG-F is between -23.5% and 41.4% (Figure 3.8). Succeeding IMERG versions should address these issues since the near-real-time application needs such as flood, landslide, and crop forecast-

ing is only possible for -E and -L runs, whereas the -F runs are mainly for research purposes.

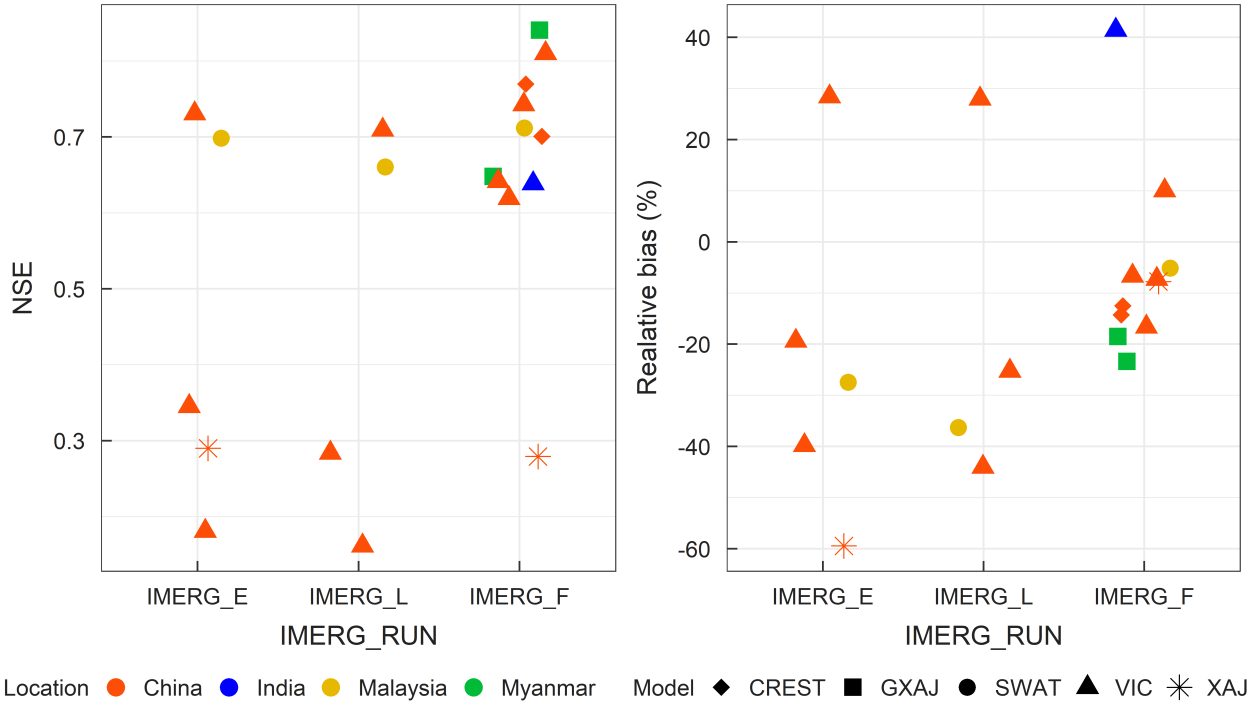


Figure 3.8: IMERG performance in hydrological performances (NSE and relative bias). CREST = Coupled Routing and Excess Storage, VIC = Variable Infiltration Capacity, XAJ = Xinanjiang, SWAT = Soil and Water Assessment Tool, GXAJ = Grid-based Xinanjiang hydrological model

3.10 Recent developments

Many studies continue to assess the IMERG products and push the boundaries (e.g., spatial coverage and validation length) observed up to 2019. For instance, Islam et al. (2020) assessed the IMERG-F V06 products over Australia (no study had evaluated IMERG over Australia before 2019) on daily, monthly, and annual scales. They found IMERG performed well, despite some discrepancies over regions with high precipitation. It was also revealed that IMERG's best performance takes place during winter in terms of seasonal scale. Similarly, Tang et al. (2020) evaluated the IMERG-F V06 products between 2000 - 2018 (the first study assessing the IMERG for the TRMM period) over China and revealed that IMERG had improved its quality over time, attributed to the increase in the number of passive microwave samples. They also reported that the performance of IMERG deteriorates when it comes to snowfall, and thus further improvement in cold climates

is needed.

Most of the studies showed that IMERG robustly represents the spatio-temporal patterns of precipitation (Sharma et al., 2020; Yu et al., 2020; Zhang et al., 2020b; Peng et al., 2020; Yu et al., 2020; Hamza et al., 2020). In addition, studies have given insights into the influence of rainfall type and topography on IMERG's performance (Yang et al., 2020a; Navarro et al., 2020). For instance, Zhang et al. (2020b) revealed that IMERG-F V06 has a high correlation with gauge observations at low elevations and low at high elevations over the Tianshan Mountains, China. Yu et al. (2020) reported a decrease in correlation and increase in RMSE with the elevation over China. In contrast, Zhang et al. (2020a) reported IMERG-E V06 performance decrease from the high-altitude regions to the low-altitude regions in terms of correlation and CSI at the daily scale over the Huang plain. Liu et al. (2020) evaluated the IMERG-F V05 product over the Bali island and concluded that IMERG best detects precipitation events at different altitudes, despite some overestimation in the high altitude. Zhou et al. (2020) evaluated the IMERG-E, -L, and -F V05 products over mainland China and found more consistent performance over southern China than northern China, and more accurate at lower latitude and elevation compared to higher latitudes and elevation. The sparse number of rain gauges and light precipitation observed at the higher latitude could be the possible reason and is in line with Navarro et al. (2020), Retalis et al. (2020), Cui et al. (2020). Besides, the process and mechanism of precipitation formation over the high elevation region could be another possible reason. Moreover, the monthly gauge adjustment to final data sets reduces the elevation and other surface sensitivity uncertainties (Sui et al., 2020).

The difficulties of IMERG-F V06 estimating light rainfall were observed over UAE (Alsumaiti et al., 2020), western Pakistan (Hamza et al., 2020), and Huang plain, China (Peng et al., 2020). In addition, Abebe et al. (2020) observed a significant decline in the detection skill of IMERG-F V06 as the intensity of precipitation increases. Furthermore, Shi et al. (2020) reported that IMERG-E and -L V06 products underestimate light rain events, whereas they overestimate the moderate to heavy precipitation events concerning precipitation detection. Freitas et al. (2020) evaluated the IMERG-F V06 product for sub-daily scale and concluded that IMERG-F has to improve its rainfall intensity and duration estimates over Brazil. Yang et al. (2020b) and Li et al. (2020a) further stress that IMERG-F V06 and V05 have to improve their sub-daily scale performances. Moreover,

Afonso et al. (2020) reported IMERG-F V06 better represent the diurnal cycle over the region characterized by deep convective cloud (warm rain process over land) compared to the shallow convection or low-level circulation and recommended further investigation in these perspectives.

In terms of hydrological evaluation, IMERG performance varies with the regions. For instance, Song et al. (2020) evaluated the IMERG-F V06 product over the Quing river basin China and concluded that IMERG has a satisfactory performance simulating the daily streamflow over the humid tropical climate. Similar results were reported for IMERG-F V06 over the Chenab River, Pakistan (Ahmed et al., 2020). Saouabe et al. (2020) evaluated the IMERG-E V06 product in terms of flood modeling in a semi-arid region of Morocco and concluded that IMERG-E has satisfactory performance in the simulation of flood events and can be applied for flood modeling in this climate in the absence of ground observations. In contrast, Mo et al. (2020) reported the unsatisfactory performance of IMERG-E V06 in both estimating precipitation and simulating the runoff over the Xiajia River basin, China, unless the Geographic Difference Analysis (GDA) method is used to correct IMERG data sets. Le et al. (2020) revealed that IMERG-E and -F V06 products outperformed other satellite products, and comparatively, IMERG-F showed better performance than the early product (IMERG-E).

For extreme events, IMERG-F V06 has the potential to capture the storm track and its spatial variation. However, IMERG has difficulties capturing the storm core and underestimates both high precipitation (>90 mm/hr) and accumulated precipitation (Li et al., 2020b). Getirana et al. (2020) evaluated IMERG-E and -F V06 products in terms of monitoring natural disaster/extreme precipitation events over Brazil and revealed the superior performance of IMERG-F compared to IMERG-E. Furthermore, they reported that both products have considerable skills in detecting extreme events (despite a slight underestimation of rain rate) and have the potential application in disaster detection. Li et al. (2020a) and Chen et al. (2020) also recommended the IMERG V05 and V06 near-real-time products for flood forecasting and early warning system.

3.11 Weaknesses and strengths

The latest versions of IMERG products show significant improvement over the TRMM data set. However, some discrepancies remain when compared to ground measurements. Most studies revealed that estimation of light precipitation should be improved (Lu and Yong, 2018; Anjum et al., 2018; Huang et al., 2019). Another common issue is the substantial overestimation or underestimation of precipitation over mountainous regions and its poor performance over complex topographies (Sharifi et al., 2016; Dezfuli et al., 2017; Asong et al., 2017; Kim et al., 2017; Sungmin and Kirstetter, 2018; Huang et al., 2018; Anjum et al., 2018). In addition, IMERG products show substantial bias in dry climates (Tang et al., 2016a; Su et al., 2018; Fang et al., 2019) and over ocean (Liu, 2016; Prakash et al., 2018a). Finally, when it comes to the seasonality, there are discrepancies in winter precipitation (Chen and Li, 2016; Chen et al., 2019; Lee et al., 2019). There is a general agreement that the IMERG algorithm needs further improvements in the aforementioned areas (Tang et al., 2016a; Su et al., 2018; Wang et al., 2018; Prakash et al., 2018a).

On the other hand, IMERG products perform robustly in various cases. For instance, their performance is equally good with the ground observations in the estimation and detection of regional precipitation patterns and their spatial averages (Gaona et al., 2016; Rozante et al., 2018; Palomino-Ángel et al., 2019). In addition, IMERG has a higher detection capability of snowfall (Sadeghi et al., 2019) and light precipitation (Wang et al., 2017a, 2019a; Yang et al., 2019), compared to other satellite products (e.g., TMPA, CMORPH). Furthermore, IMERG has the potential to detect and estimate hurricane precipitation, indicates its significant applicability for estimation of precipitation during the extremes, and thus can be used for impact modeling studies (Omranian et al., 2018). All these factors reveal promising potential and a wide range of future applications.

3.12 Conclusion

This chapter compiled the reported performance of IMERG products across different climatic conditions and geographic locations throughout the globe. Along with IMERG performance representing precipitation, we also investigated the performance of IMERG regarding extreme pre-

cipitation events and hydrological application.

In terms of the spatial and temporal distribution of IMERG evaluation, Asia, and China in specific, are dominant in number of studies followed by North and South America, while Africa and Europe recorded the least number. Regarding the spatial and temporal resolutions, $0.1^\circ \times 0.1^\circ$, $0.25^\circ \times 0.25^\circ$ with daily, monthly and annual scales are the most evaluated resolutions of IMERG products. In addition, 12 – 24 months is the validation period length used in most studies. Studies with longer validation periods were often associated with coarser temporal resolution (daily and longer), whereas studies with shorter validation periods were associated with sub-daily temporal resolution. As expected, the studies at coarser time scales surpass the finer ones, highlighting the need for more research in sub-daily resolutions. Surprisingly, also very few studies exist that investigate how IMERG products perform as they move from daily to monthly scale. Understanding how biases propagate across the time scale-continuum is crucial for the proper validation and application of any data set Markonis et al. (2021).

IMERG showed better performance compared to the TMPA estimates in the representation of spatio-temporal variability of precipitation across the climatic and geographic conditions. However, IMERG showed significant over/underestimation in different precipitation intensities that varies with region and climatic conditions. When it comes to climate regimes, IMERG tends to more consistent precipitation estimates over humid regions (wet and high intensity precipitation) compared to semi-arid and arid regions (dry and low intensity precipitation), which is especially true for China. Also, IMERG still has difficulties in estimating precipitation over complex terrains and mountainous regions. Orographic precipitation associated with high mountains is the major cause of the poor performance over such conditions. Most of the studies found that the accuracy of IMERG increased significantly with temporal aggregation, i.e., monthly results were reasonably better than daily ones, and annual results were better than monthly ones. IMERG performance is also affected by seasonal variation. Usually, IMERG performs poorly in winter compared to the summer season, which is attributed to the inability to detect light rainfall (more common in winter). Another reason could be that during winter the solid phase of precipitation (snowfall) is more common.

IMERG captures well the spatiotemporal patterns and variability of extreme precipitation.

Nonetheless, IMERG has issues when measuring over the center of the typhoon or hurricanes. In addition, it has some limitations estimating high-intensity precipitation and shows significant underestimation across various geographical locations and climatic conditions. From the perspective of the hydrological application, although most of the studies conclude that there is potential for the application of IMERG products in simulating the streamflow their performances differ depending on the hydrological model used, calibration methods, and basin types. Despite IMERG showing better detection capacity, its performance over the ocean shows substantial over/underestimation of total precipitation. However, the reliability of observational data over the ocean is questionable, and it is too early to draw any conclusions.

Overall, the performance of IMERG varies with climatic conditions, geographical locations, seasons, precipitation types, and intensities. More studies throughout the globe, especially in the regions that are under-represented, are needed for a better evaluation of IMERG's performance. It remains to be seen how the extension of the dataset back to 2000 has influenced the regional and overall performance of IMERG. This will allow also for an investigation of the climatic properties of precipitation, which will be valuable for a better quantification of the global water cycle Vargas Godoy et al. (2021). What is most promising, though, is that each new version improves the previous one, in almost every validation metric examined in this analysis. Thus, despite its limitations, IMERG remains one of the most robust alternatives to ground observational records.

Chapter 4

Assessment of IMERG precipitation estimates over the tropical oceans

4.1 Introduction

Covering 71% of Earth's surface and receiving about 78% of global total precipitation, the ocean plays an important role in the Earth's climate system and hydrological cycle (Trenberth et al., 2007; Vargas Godoy et al., 2021). The tropical oceans which receives a major proportion of total oceanic precipitation have significant effects on the global radiation budget. Therefore, understanding the amount, rate, and distribution of precipitation on the tropical oceans not only assists in the accurate estimation of the global water cycle and energy fluxes, but also enhance our understating about the processes of the global water cycle over the ocean. Thus, the precise quantification of tropical oceanic precipitation is one of the utmost research interests.

One of the reasons for poor understanding and very few numbers of studies on the tropical ocean is a lack of observational network over oceans. At present time, bridging this important information gap satellite estimations have been providing an important and promising source of precipitation for data-scarce regions like mountains, and so over oceans. Since the launch of Tropical Rainfall Measuring Mission (TRMM) core instrument in 1998, the TRMM Multi-Satellite Precipitation Analysis (TMPA) has been one of the most widely used satellite precipitation products over the tropics (Huffman et al., 2007). Following its success, the Global Precipitation Measurement mission (GPM) launched its core satellite in the early 2014 (Hou et al., 2013; Liu, 2016). Thereafter, several studies have evaluated the performance of GPM products, either by validating them using reference gauge observations or by comparing them with TRMM products across a range of climates (Pradhan et al., 2022).

Despite the scarcity of reference datasets, there have been efforts to evaluate the satellite precipitation over the ocean either by using radar observations, island gauges, or buoy observations. A substantial number of studies performed quantitative analysis of satellite precipitation over the tropical oceans using the gauges datasets from buoys. Among them, Bowman (2005) compared the TRMM precipitation using 25 buoys over the Pacific Ocean, and reported the validation challenges in point-area averaging between the satellite and gauges. Sahany et al. (2010) analysed the TMPA with buoy observation for the estimation of the diurnal cycle over the Indian Ocean and found overall a good agreement between these two products. In addition, Sapiano and Arkin

(2009) evaluated the TMPA over the Pacific Ocean using buoys observations and revealed an underestimation by TMPA. Prakash et al. (2011) compared the TRMM Microwave Imager (TMI) with the available buoys observations at the Indian, Pacific and Atlantic oceans. They found reasonable agreement in precipitation rate between the TMI and buoys observation over the Atlantic Ocean, followed by the Pacific and the Indian Ocean.

In recent times, there have been limited attempts to assess the performance of IMERG products over the ocean. For instance, Prakash et al. (2018a) first evaluated the IMERG-F V04 products over the north Indian Ocean against the buoys observation from March 2014 to December 2015. They noticed substantial positive bias and false alarms in IMERG estimates despite its better improvement compared to TRMM. Kucera and Klepp (2018) evaluated the IMERG V03 products over the ocean and revealed an underestimation of IMERG compared to the OceanRain. Similarly, IMERG V05 has been evaluated using the OceanRain and the Dual-frequency Precipitation Radar (3DPRD) as references and revealed an overall underestimation in IMERG precipitation compared to the OceanRain; despite an accurate detection of precipitation events (Khan and Maggioni, 2019). In addition, IMERG comparison with TMPA and buoys observation over the tropical oceans revealed that IMERG performs best when representing the mean precipitation at the Pacific and Indian oceans, except for the high-precipitation regions of Atlantic (Wu and Wang, 2019). Evaluation results of IMERG-F V06 monthly estimates with the PACRAIN atoll daily observations over the ocean also showed an overall overestimation by IMERG (Bolvin et al., 2021). The study also revealed that IMERG tends to underestimate light precipitation while overestimating intense precipitation rates.

It should also be noted that most of the aforementioned studies are based on the TMPA dataset. Thus, so far, very little is known regarding the GPM IMERG products' performance. In particular, the most recent IMERG version, V06, which has not yet been evaluated at a daily scale. In this context, the main objective of this chapter is to comprehensively investigate the performance of the IMERG V06B precipitation estimates using observation buoys across the tropical oceans. This will bring out important information for the user community, the GPM ground validation group, and the algorithm developers, and thus improve its applicability over the remote ocean regions.

This chapter is organized as follows. After this introduction, the second section briefly de-

scribes the datasets used and the methodological approach employed in this analysis. The third section represents the results, including spatial and temporal evaluation of IMERG. The fourth section discusses the significant findings. Finally, the fifth and last section summarise and conclude the findings.

4.2 Data and Methods

4.2.1 IMERG

IMERG is a gridded precipitation product and algorithm that merges a variety of satellite precipitation estimations from the GPM constellation (Huffman et al., 2015). Although IMERG V06 is extended beyond 60°N - 60°S (i.e., until 90°N - 90°S), it is partially outside of that latitude band (i.e., only limited to non-ice covered surfaces). In this study, the IMERG V06 Early, Late and Final run daily products are evaluated over the period of 2001 - 2020 (Table 5.1). An overview of the IMERG algorithm is presented in Chapter 1.5.

4.2.2 Buoys

The in-situ data from the Global Tropical Moored Buoy Array (GTMBA) program was used to evaluate the IMERG products over the tropical ocean. This program includes the Tropical Atmosphere Ocean/Triangle Trans-Ocean Buoy Network (TAO/TRITON) in the Pacific (McPhaden et al., 1998), the Prediction and Research Moored Array in the Tropical Atlantic (PIRATA) (Bourlès et al., 2008), and the Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction (RAMA) in the Indian Ocean (Mcphaden et al., 2009). The buoys are equipped with a Y.M young R3125 rain gauge and measure precipitation around 3 meters above the sea surface. The relevant datasets of each (i.e., TAO, PITRA and RAMA) buoys can be accessed from the Global Moored Tropical website (`https://www.pmel.noaa.gov/gtmba/`) or OceanSites website (`http://www.oceansites.org/data/index.html`) (Table 5.1). An important feature of the buoys observation is unlike to the overland rain gauges, it excludes from

the island orographic and other thermodynamic effects (Bowman et al., 2005; Prakash et al., 2013; Mcphaden et al., 2009). The buoys observations provides an independent precipitation source, as IMERG didn't used them for the gauge correction over ocean. Moreover, no additional wind correction was applied to the buoys observation.

For the Indian ocean the 15 RAMA buoys were selected with starting from 2004 to 2020. Among them, 7 are situated in the northern hemisphere whereas the remaining are on the southern hemisphere. For the Atlantic ocean 25 buoys were selected and and the data ranges from 2001 - 2020. For the Pacific Ocean 40 buoys are selected and the the data ranges from 2001 - 2020, of which 22 is over the western Pacific and 18 over the eastern Pacific. The detailed spatial distribution of the buoys over the tropical ocean is depicted in the Figure 4.1.

Table 4.1: Summary of the datasets used in this study.

Dataset name	**Spatial scale**	**Temporal scale**	**Record length**	**Reference**
Buoys	point	hourly	2001 - 2020	McPhaden et al. (1998); Bourlès et al. (2008)
IMERG-E, -L, -F	$0.1^\circ \times 0.1^\circ$	0.5h	2001 - 2020	Huffman et al. (2015)

4.2.3 Methodology

This study is focused on the tropical oceans, covering an extent of 25° N - S in both the hemispheres (Figure 4.1). The selection of the study area is driven by two main reasons; i) the significance of the region in global water and energy balance, and ii) the availability of the in-situ observational datasets (i.e., the buoys). In order to compare the grided satellite estimates with point-based buoys measurement, the point-pixel based approach was employed (Xie et al., 2022; Chen et al., 2018). Following the extraction of IMERG pixels ($0.1^\circ \times 0.1^\circ$) overlying the buoy stations, a pairwise time series was prepared for the evaluation. Despite the fact that buoy datasets are available from 1998 onward, the datasets from 2001 to 2020 were selected for the analysis to match the IMERG period. Due to the lack of continuous time series of the buoy datasets, only the days present both in the buoys and the IMERG were considered for the analysis. Consequently, the sample size for each ocean was different, ranging from 5809 days over the Indian Ocean to 7519 over the Pacific Ocean. In addition, for the basin scale comparison, the mean areal precipitation of all the IMERG pixels and buoys falling inside the particular basin is taken into account.

To account for spatial heterogeneity in precipitation rate and to better understand the regional influence on IMERG, the Pacific Ocean was divided into East (longitude < 0) and West Pacific (longitude > 0). The buoy precipitation rate was converted from intensity rates (mm/hr) to daily precipitation rates (mm/day), to match the IMERG format, and the evaluation methodology was carried out on a daily timescale.

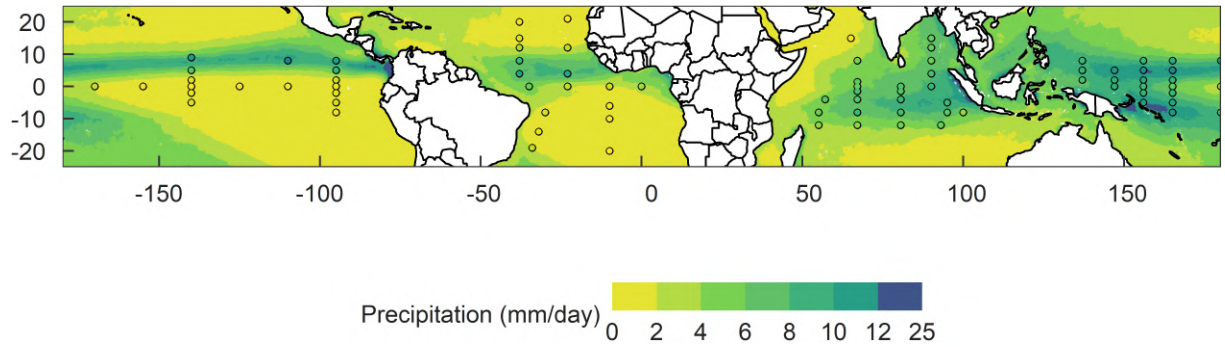


Figure 4.1: Spatial distribution of mean daily precipitation of buoys and IMERG (IMERG-F) for the period of 2001 - 2020 over the tropical oceans.

The standard continuous and categorical metrics were used to quantitatively evaluate the IMERG products against the buoys observations. The main continuous metrics include Bias (Eq. 1), Root Mean Square Error (RMSE; Eq. 2), and Mean Absolute Error (MAE; Eq. 3). The Pearson correlation coefficient (COR; Eq. 4) was used to estimate the cross-correlation between the two datasets.

$$Bias = \frac{\sum(S_i - G_i)}{N} \tag{4.1}$$

$$MAE = \frac{\sum|S_i - G_i|}{N} \tag{4.2}$$

$$RMSE = \sqrt{\frac{\sum(S_i - G_i)^2}{N}} \tag{4.3}$$

$$COR = \frac{(\sum(S_i - \bar{S})(G_i - \bar{G}))^2}{(\sum(S_i - \bar{S})^2 \sum(G_i - \bar{G})^2)} \tag{4.4}$$

Where S_i and G_i represent the IMERG and buoys precipitation, $\bar{S}$ and $\bar{G}$ their mean precipitation correspondingly.

Table 4.2: Contingency table of satellite and buoys precipitation data.

	Buoys $\geq$ Threshold	Buoys $<$ Threshold
Satellite $\geq$ Threshold	Hits (H)	False alarms (F)
Satellite $<$ Threshold	Misses (M)	Correct negative (C)

Unlike the continuous skill scores which were used for precipitation volume estimation, the detection skill scores represent the IMERG's ability to detect precipitation events. A 2×2 contingency table was constructed with four variables including hit (H), miss (M), false alarm (F), and correct negative (CN) for the IMERG-buoys pairs (Table 4.2). Based on these scores, the three most commonly used categorical metrics, namely the Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI), were calculated (Eq. 5-7).

$$POD = \frac{H}{H + M} \tag{4.5}$$

$$FAR = \frac{F}{H + F} \tag{4.6}$$

$$CSI = \frac{H}{H + M + F} \tag{4.7}$$

POD represents the total number of correctly detected precipitation events to the total events. FAR includes the total number of times IMERG detects a false event to the total events, whereas the CSI reflects a more balanced score between the POD and FAR. In this analysis, a rain/no-rain threshold of 0.1 mm/day was used to compute the metrics (values, < 0.1 are very low and may be considered noise). This threshold was selected to account for light precipitation events, which make up a significant portion of the total precipitation events. For instance, IMERG-F has 12%, 35%, 32%, and 10% of precipitation events between 0.1 to 1 mm/day for Indian, Atlantic, East Pacific, and West Pacific, respectively. Although the detection threshold is subjective (Behrangi

et al., 2012), previous studies have also used a threshold of 0.1 mm/day for estimating the detection scores (Tian and Peters-Lidard, 2007; Wu et al., 2018). In addition to the main threshold, categorical metrics were also examined with varying thresholds ranging from 0.1 mm/day to 1.0 mm/day. The results demonstrate that while the choice of threshold has a slight effect on the categorical metrics, the overall pattern remains consistent (Figure A1).

To further assess the precipitation volume error, an error decomposition method was applied (Wang et al., 2022; Tian et al., 2009). This method divides the total bias (e) into three independent constituents; hit bias (h), missed bias (m), and false alarm bias (f). The hit bias which is generated from the hit events can be either positive (overestimate the volume from the detected events) or negative (underestimate the volume from the detected events). On the other hand, the false alarm bias which comes from the IMERG's falsely identified events when there is no precipitation detection in buoys, obviously can only lead to positive bias. Similarly, the missed bias can only have negative bias since it arises from the precipitation events that are not detected by the IMERG. It can also be noted that the total bias may be smaller than the individual biases since they may cancel each other out, leading to a smaller total error (Wang et al., 2022). Therefore, by decomposing the total error into individual components, we can gain important insight into the main contributors to the total error. This relationship between the individual components and the total error can be represented as follows;

$$e = h - m + f \tag{4.8}$$

4.3 Result and Discussion

4.3.1 Mean daily precipitation maps

The basic visual inspection of satellite and buoy precipitation is one of the simplest verification methods between the two datasets. The mean daily precipitation for IMERG-F and buoys from 2001 to 2020 is shown in Figure 4.1. Even though a perfect match between IMERG-F's grided areal precipitation estimates and buoy point measurements is not accomplished, IMERG-F accurately

represents the spatial pattern of buoy precipitation across tropical oceans. This is in line with other studies (Pradhan et al., 2022), which found that IMERG-F provided a better representation of the spatio-temporal pattern of precipitation than the other satellite estimates across a variety of climatic and regional conditions. IMERG-F estimates are in well agreement with the buoys in the high-precipitation regions, such as the Inter Tropical Convergence Zone (ITCZ), over the Pacific and the Indian oceans. Additionally, it also represents well the low-precipitation areas of the East Pacific and Atlantic Oceans. However, IMERG-F substantially overestimates precipitation with the bias varying with location. The mean daily precipitation for buoys ranges from 0 – 11 mm/day, whereas the corresponding range for IMERG-F is nearly twice as large (0 – 25 mm/day). Given that we are comparing the entire tropical oceanic precipitation with very few point measurements, such differences are expected. Moreover, since the buoy point measurements are very sparser in spatial distribution compared to IMERG's complete coverage, it is very likely that buoys could have missed some precipitation from high-precipitation regions due to their absence over those areas.

4.4 Point-pixel evaluation

The statistical performance of various IMERG products in comparison to buoy precipitation data has been evaluated using several metrics such as Bias, RMSE, COR, POD, FAR, and CSI. The daily-scale estimates for each buoy-IMERG pair have been computed separately for different oceanic regions (Figure 4.2 and Table 4.3). All the IMERG runs tend to overestimate the buoy precipitation, (except IMERG-F over the East Pacific), though the magnitude varies across oceanic basin. In terms of bias, IMERG-F shows the best performance over the East Pacific with a mean bias of -0.07 mm/day (-3%), which is much lower than the mean bias observed at the Indian (1.70 mm/day and 38%), Atlantic (0.33 mm/day and 16%), and West Pacific (2.37 mm/day and 36%) oceans. The high precipitation of the Indian Ocean (mean 4.43 mm/day) and West Pacific (mean 6.67 mm/day) could be one possible reason for the overestimation. On the other hand, while the Indian and West Pacific exhibit the highest RMSE values (Figure 4.2b), the Atlantic (395%) and East Pacific (351%) have the highest RMSE values in terms of relative errors (Table 4.3). Despite IMERG's lower bias over the East Pacific, the higher variability in terms of RMSE indicates the possibility that

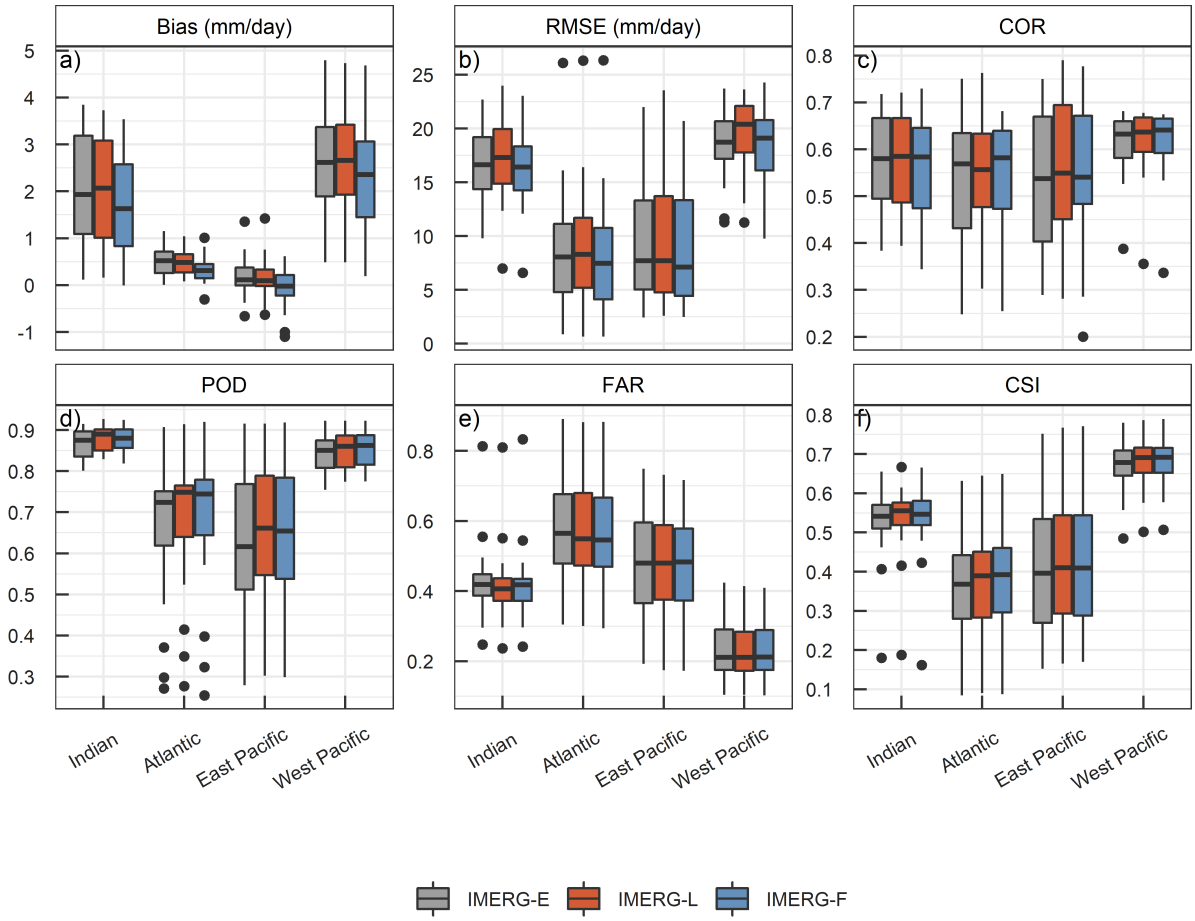


Figure 4.2: Evaluation metrics of daily IMERG precipitation over the tropical oceans for the period of 2001 - 2020; a) Bias (mm/day), b) Root Mean Square Error (mm/day), c) Correlation Coefficient, d) Probability of Detection, e) False Alarm Ratio, and f) Critical Success Index.

the positive and negative biases may offset each other, resulting in a low total bias. In addition, a similar pattern is also observed for COR, with the highest values over the West Pacific (0.61) and Indian (0.56%), followed by the East Pacific (0.55) and Atlantic (0.54) oceans (Figure 4.2c). In terms of IMERG runs, IMERG-F products have slightly better performance than IMERG -E and -L products. This is because, while none of the IMERG runs has any gauge correction over the ocean, there is a slight difference among the runs. The early run only has forward propagation, whereas the late and final runs have both backward and forward propagation. Moreover, the only difference between the late and final over the ocean is the period of calibration.

Interestingly, the analysis of precipitation events detection reveals that the categorical scores are quite opposite to the volumetric scores. All the IMERG runs show good detection capability

over the West Pacific (0.86) and Indian oceans (0.88), where the volumetric scores are worse. In fact, both the Indian and West Pacific oceans have similar POD values with a median above 0.85 (Figure 4.2d). However, the distribution of values for the Indian Ocean is slightly narrower compared to the West Pacific, indicating IMERG's high detection capability over all the buoys of the Indian Ocean, whereas some spatial heterogeneity for the West Pacific (values vary from 0.70 to 0.92). Similar to POD, IMERG has shown the best FAR over the West Pacific with the lowest median FAR (0.21), followed by the Indian (0.41), East Pacific (0.48), and Atlantic (0.54) oceans (Figure 4.2e). This could possibly suggest that the IMERG's high detection capability over the Atlantic (i.e., POD $>$ 0.6) comes at the expense of a high false alarm ratio. Furthermore, IMERG shows high variability of POD and FAR among the buoys for the Atlantic and East Pacific oceans compared to the Indian and the West Pacific oceans. Similar results were also shown for CSI with the highest values over the West Pacific (0.69), followed by the Indian (0.54), East Pacific (0.40), and Atlantic (0.39) oceans (Figure 4.2f).

Table 4.3: Volumetric scores (mm/day) for all the IMERG runs in comparison with Buoys at daily scale.

Ocean	IMERG run	Mean (Buoys)	Bias	RMSE	COR
	IMERG-F		1.70 (38%)	16.11 (364%)	0.56
Indian	IMERG-L	4.43	2.05 (46%)	17.24 (389%)	0.57
	IMERG-E		2.04 (46%)	16.76 (379%)	0.57
	IMERG-F		0.33 (16%)	8.38 (395%)	0.54
Atlantic	IMERG-L	2.12	0.49 (23%)	9.06 (427%)	0.54
	IMERG-E		0.50 (23%)	8.72 (411%)	0.53
	IMERG-F		-0.07 (-3%)	9.08 (351%)	0.55
East Pacific	IMERG-L	2.59	0.17 (7%)	9.98 (385%)	0.56
	IMERG-E		0.20 (8%)	10.13 (391%)	0.54
	IMERG-F		2.37 (36%)	18.52 (278%)	0.61
West Pacific	IMERG-L	6.67	2.70 (41%)	19.49 (292%)	0.61
	IMERG-E		2.64 (40%)	18.35 (275%)	0.61

4.4.1 Spatial distribution of categorical scores

In order to better understand the regional performance of IMERG detection skills, we further investigate the spatial distribution of categorical metrics, including the POD, FAR, and CSI in different regions (Figure 4.3). Since the performance difference among the three IMERG data products is negligible, we only present the detection scores of IMERG Final (IMERG-F) products. In terms of POD spatial distribution, IMERG-F exhibited relatively low variability over the West Pacific (0.7 - 0.9) and Indian (0.8 - 0.9) oceans. However, IMERG-F has demonstrated high variability in its detection capability over the East Pacific and Atlantic oceans, with POD values varying between 0.3 and 0.9. For the East Pacific, in the low-precipitation zone along and south of the equator between 95°W and 120°W the POD scores are considerably low (< 0.5). In contrast, buoys located along the high-precipitation band north of the equator have better performance, with POD > 0.7. Similar to the East Pacific, POD scores over the Atlantic Ocean follow the high/low-precipitation regions, with the highest detection scores in the west Atlantic along the equatorial ITCZ band, while the moderate detection scores (0.5 – 0.7) throughout the rest of the low-precipitation regions. However, IMERG-F's lowest (< 0.5) values are reported in the southernmost and northernmost low-precipitation regions along the 10°W and 22°W meridians, respectively.

The majority of IMERG-F grid cells exhibit comparatively better FAR performance over the Indian and West Pacific regions, with most values varying between 0.3 to 0.5. Among them an outlier is the northernmost Arabian Sea, a low-precipitation region with a FAR around 0.8, that stands out from the rest of the Indian Ocean (FAR < 0.3). Over the East Pacific, most of the IMERG-F grid cells have very high FAR with values reaching up to 0.71. Especially, in the low-precipitation region along and south of the equator, FAR scores range between 0.5 and 0.7. In contrast, to the north of the equator, along the high-precipitation ITCZ band, IMERG-F demonstrates good performance with FAR values < 0.5. Over the Atlantic Ocean, FAR shows high variability, varying from 0.29 to 0.89. Although no particular spatial pattern exists for the distribution of FAR, similarly to the POD, the low FAR values are observed in the high-precipitation regions at the west Atlantic, whereas the worst performance (> 0.8) is reported over low-precipitation regions. The

analysis of the regional patterns further confirms that IMERG-F has better detection ability and fewer false alarms in the high-precipitation regions than in the low-precipitation regions.

The CSI results are quite similar to those of POD. IMERG-F exhibits notable variability in its performance over the East Pacific and the Atlantic Ocean, with scores varying between 0.1 and 0.7. In fact, in the low-precipitation zone along and south of the equator in the East Pacific, not even any single IMERG-F pixel has CSI above 0.5, which indicates a generally lower level of detection skill. This also indicates that the low POD and excessive FAR in this region reduce the CSI. In contrast, over the northern high-precipitation regions between 5°N to 10°N latitude, IMERG-F has moderate CSI (0.5 – 0.7). Similar to the POD, the high precipitation regions at the west Atlantic show good detection (> 0.5), whereas the worst (< 0.3) is observed in low-precipitation regions. On average, the northern Atlantic has shown comparatively better CSI than the southern Atlantic. This pattern could be attributed to the greater extent of the low-precipitation region over the southern Atlantic relative to the North Atlantic.

4.4.2 Spatial distribution of volumetric scores

The spatial distribution of bias varies from -1.10 to 4.68 mm/day, indicating an overall and more pronounced overestimation of precipitation than underestimation throughout the tropical oceans (Figure 4.4). The best performance of IMERG-F is observed over the Atlantic and East Pacific oceans, with bias varying between -1.10 to 1.01 mm/day. Despite the similar bias range between the Atlantic and East Pacific oceans, most of the buoys (17 out of 18) over the Atlantic show positive bias, indicating an overall overestimation by IMERG-F. However, in the case of the East Pacific, it tends to be the other way around, i.e., most of the buoys (11 out of 18) show negative bias, revealing an overall underestimation by IMERG-F. Except for one buoy (Bias = 1.01) that lies around the northern Atlantic (high-precipitation ITCZ bands) between 2° - 8° N latitude, IMERG-F shows consistently low bias with values < 0.8 mm/day. In contrast, bias values in the Indian Ocean display more variability, ranging up to 4 mm/day, with comparatively less bias over the Arabian Sea than the Bay of Bengal, in line with previous findings using OMNI buoys (Prakash et al., 2018a). This difference may be explained by the difference in precipitation distribution

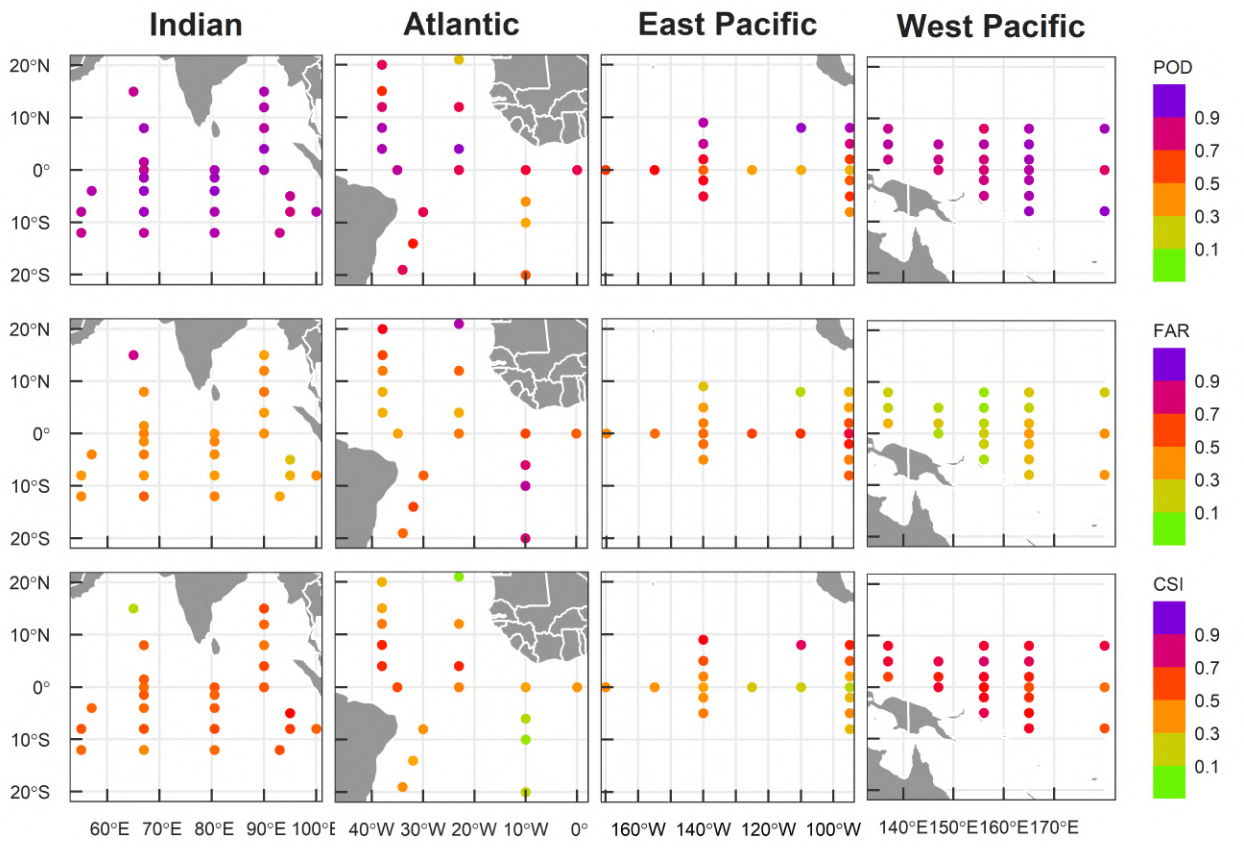


Figure 4.3: Spatial distribution of precipitation detection metrics across the tropical oceans at daily scale.

across the basins, with the Bay of Bengal generally receiving more precipitation than the Arabian Sea. On the other hand, IMERG-F exhibits the highest overestimation over the West Pacific, with most of the grid cells having a bias between 2 - 4 mm/day. The buoys, in particular, located between the 130°E to 160°E longitude have a bias $>$ 2 mm/day.

Unlike the bias, the spatial distribution of RMSE values across the tropical oceans exhibits considerable variability, especially over the Atlantic and East Pacific oceans (Figure 4.4). IMERG-F dataset shows comparatively better RMSE performance over the Atlantic and East Pacific than in the Indian and West Pacific. The worst performance is yielded over the West Pacific with most of the IMERG-F grid cells reporting RMSE values $>$ 20 mm/day. Likewise, IMERG-F also exhibits a moderately high RMSE over the Indian Ocean, with most locations having RMSE values $>$ 10 mm/day. On the other hand, in the East Pacific, IMERG-F exhibits lower RMSE ($<$ 10 mm/day) over the low-precipitation region along and south of the equator. In contrast, in the

high-precipitation region north of the equator, RMSE values range from 5 to 25 mm/day. Unlike the other oceanic regions, IMERG-F presents the highest variability over the Atlantic with RMSE varying from 0.63 to 26.34 mm/day. In fact, most of the IMERG-F grid cells in the southern Atlantic (low-precipitation region) have comparatively lower RMSE ($<$ 10 mm/day) than their corresponding northern counterparts. Especially along the high-precipitation region between 0° - 12° N latitude, IMERG-F has the highest RMSE with values greater than 15 mm/day. Whereas along the low-precipitation regions between the 15° - 20° N/S latitude in both hemispheres have the lowest RMSE ($<$ 5 mm/day). Nonetheless, when it comes to relative RMSE values (i.e., standardised by their means), in particular, over the Atlantic and East Pacific, the regions with low-precipitation exhibit the highest values compared to the high-precipitation regions (Figure A2).

On the other hand, the spatial distribution of correlation between the IMERG-F and buoys estimates is quite diverse. Furthermore, unlike bias and RMSE, it does not exhibit any particular spatial pattern (Figure 4.4). The correlation values range from 0.2 to 9.78. Not surprisingly, the East Pacific exhibits the greatest variability, with values ranging from $<$ 0.3 to $>$ 0.7. On the other hand, the Atlantic Ocean has values between 0.3 to 0.7. IMERG-F pixels along the 30°W - 40°W longitude regions, in particular, have a low correlation with values $<$ 0.5. For the Indian Ocean, the majority of the IMERG-F pixels (11) have a correlation of $>$ 0.6, with most (8) located south of the equator. Correlation values with $<$ 0.5 are predominantly observed in the IMERG-F pixels in the regions between 1°S - 10°N. In contrast to the other oceanic basins, IMERG-F pixels across the West Pacific Ocean exhibit strong correlate with the buoys precipitation, with most having values between 0.6 and 0.7. These results are consistent with previous findings on buoy and CMORPH correlation values, which have been reported to range between 0.5 and 0.7 (Xie et al., 2017).

One of the main reasons for the high variability of IMERG-F's performance over the Atlantic and East Pacific compared to the Indian and West Pacific could be attributed to the spatial distribution of precipitation patterns in these regions. The West Pacific and Indian Ocean exhibit a more extensive high-precipitation zone, resulting in buoys being located within homogeneous precipitation regions (Figure 4.1). In contrast, the high-precipitation zone of the East Pacific and Atlantic oceans is narrower and limited to the ITCZ region, leaving out more regions on either

side of the equator with very low precipitation. Substantially, a greater number of buoys are located over the low-precipitation regions than in the high-precipitation ITCZ regions, leading to higher variability in IMERG-F's performance in these regions.

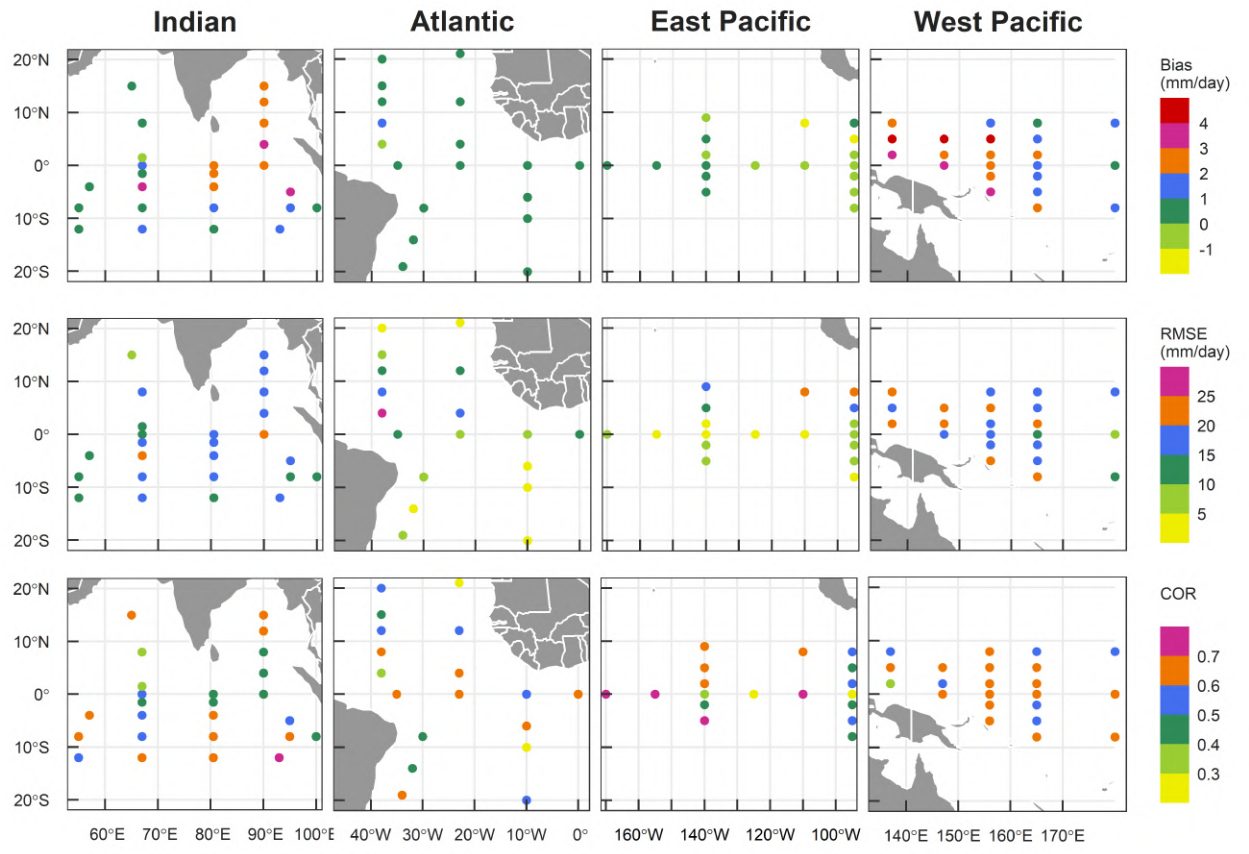


Figure 4.4: Spatial distribution of volumetric metrics across the tropical oceans at a daily scale.

4.5 Basin scale evaluation

In this section, the spatial mean precipitation of IMERG-F grid cells for each oceanic region and their corresponding buoys are compared for each oceanic basin (Figure 4.5). IMERG-F has better agreement with buoy precipitation over the Atlantic and East Pacific oceans than over the Indian and West Pacific oceans. For the Indian Ocean, IMERG-F tends to substantially overestimate buoy precipitation, with high-density peaks above the 1:1 line. The notable overestimation of the higher precipitation magnitudes could have caused the comparatively high RMSE and MAE. For the Atlantic and East Pacific, IMERG-F and buoy precipitation are in good agreement and are

comparatively less scattered than in the Indian and West Pacific oceans. Especially for higher magnitudes, IMERG-F shows comparatively less overestimation than in the Indian and West Pacific. On the other hand, the higher density precipitation peaks along the 1:1 line imply a better correlation for the West Pacific. Similar to the Indian Ocean, the high RMSE and MAE of IMERG-F over the West Pacific could be attributed to the relatively high mean precipitation intensities.

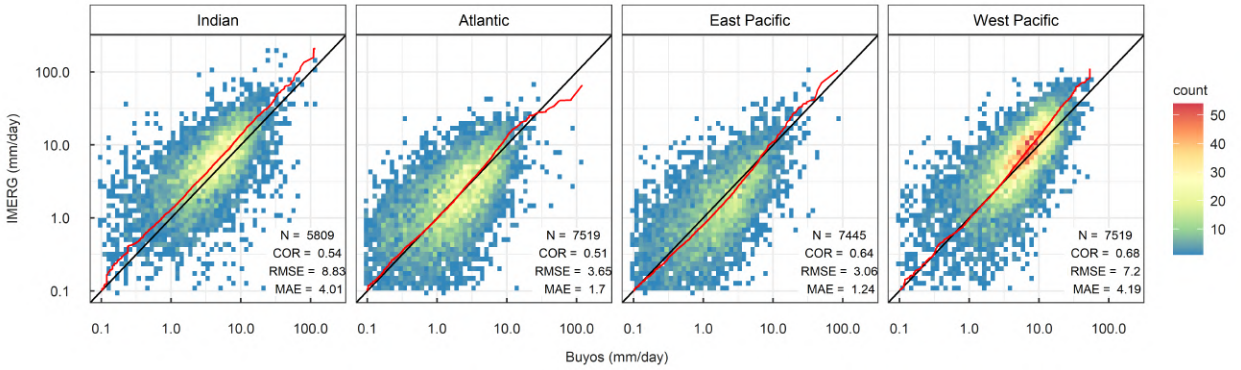


Figure 4.5: Scatter density plot between IMERG-F and buoys daily precipitation (mm/day) over the tropical oceans (the red line denotes the quantile-quantile match-ups).

Furthermore, it can also be noticed how the evaluation metrics significantly improved for all the oceanic basins compared to the IMERG-F-buoys individual point-pixel errors. For instance, the RMSE values for the Indian, Atlantic, East Pacific, and West Pacific are reduced from 16.1, 8.3, 9.08, and 18.52 from the point-pixel scale to 8.83, 3.65, 3.06, and 7.2 for the basin scale, respectively. Considering the smoothing effect of spatial aggregation, it is generally expected that the agreement between the regional mean precipitation is better than the individual IMERG-F's point-pixel agreement (Behrangi and Wen, 2017). This is also in line with other findings (Tan and Santo, 2018; Wang et al., 2017b), which also confirmed the positive effect of scale on satellite errors. Nonetheless, the pattern of IMERG-F performance remains similar to the previous results, i.e., IMERG-F's best performance is over the Atlantic and East Pacific, followed by the Indian and West Pacific oceans.

4.5.1 Frequency

In order to better understand the IMERG-F performance in terms of representing the frequency of different precipitation rates, the total precipitation is divided into various intensities (Fig-

ure 4.6). Overall, IMERG-F underestimates the occurrence frequency of no-rain days compared to the buoys throughout the tropical oceans, although the extent of underestimation varies among the regions. The Atlantic exhibits the largest underestimation (10%), followed by the East Pacific (5.2%), Indian (5%), and West Pacific (3%) oceans. Similar results were also previously reported for IMERG-F by Prakash et al. (2018a) over the Indian Ocean, though the reference data was different (i.e., OMNI buoys). Even though the rain/no-rain threshold of 0.1 mm/day could probably influence the results, the primary reason could be attributed to the point-pixel comparison (Behrangi et al., 2012). For instance, as IMERG-F pixels represent around 11×11 km over the equator, it is highly probable that the light rainfall could occur anywhere in the grid and probably be missed by the buoys if it does not occur precisely over the buoy point location. Consequently, buoys can record a higher number of non-rainy days than the IMERG-F.

Furthermore, except in the Indian Ocean, IMERG-F tends to overestimate the light precipitation (0.1 – 2 mm/day) throughout the tropical oceans. Especially, in the Atlantic and East Pacific, the overestimation of light precipitation is more pronounced than in the West Pacific Ocean. This is due to the frequency distribution over the Atlantic and East Pacific oceans which is right-skewed, indicating the dominance of light precipitation. On the other hand, IMERG-F's overestimation of the heavy precipitation events ($>$ 10 mm/day) is more pronounced over the Indian and west Pacific oceans than the Atlantic and east Pacific ones. IMERG-F's overestimation of heavy precipitation was also reported over the Indian (Prakash et al., 2018a) and Pacific oceans (Bolvin et al., 2021). Again, this is due to the higher fraction of heavy precipitation over the Indian and West Pacific oceans. Moreover, this may be the one probable reason for IMERG-F's overall overestimation of precipitation over the Indian and West Pacific oceans. Finally, IMERG-F tends to slightly underestimate moderate precipitation (3 – 5 mm/day) in all basins, except for the Atlantic, where it agrees with the buoys.

4.5.2 Error decomposition

To further understand the IMERG estimation errors, the total bias is divided into hit bias, missed bias, and false alarm bias (Figure 4.7). It can also be noted that unless a specific IMERG run is

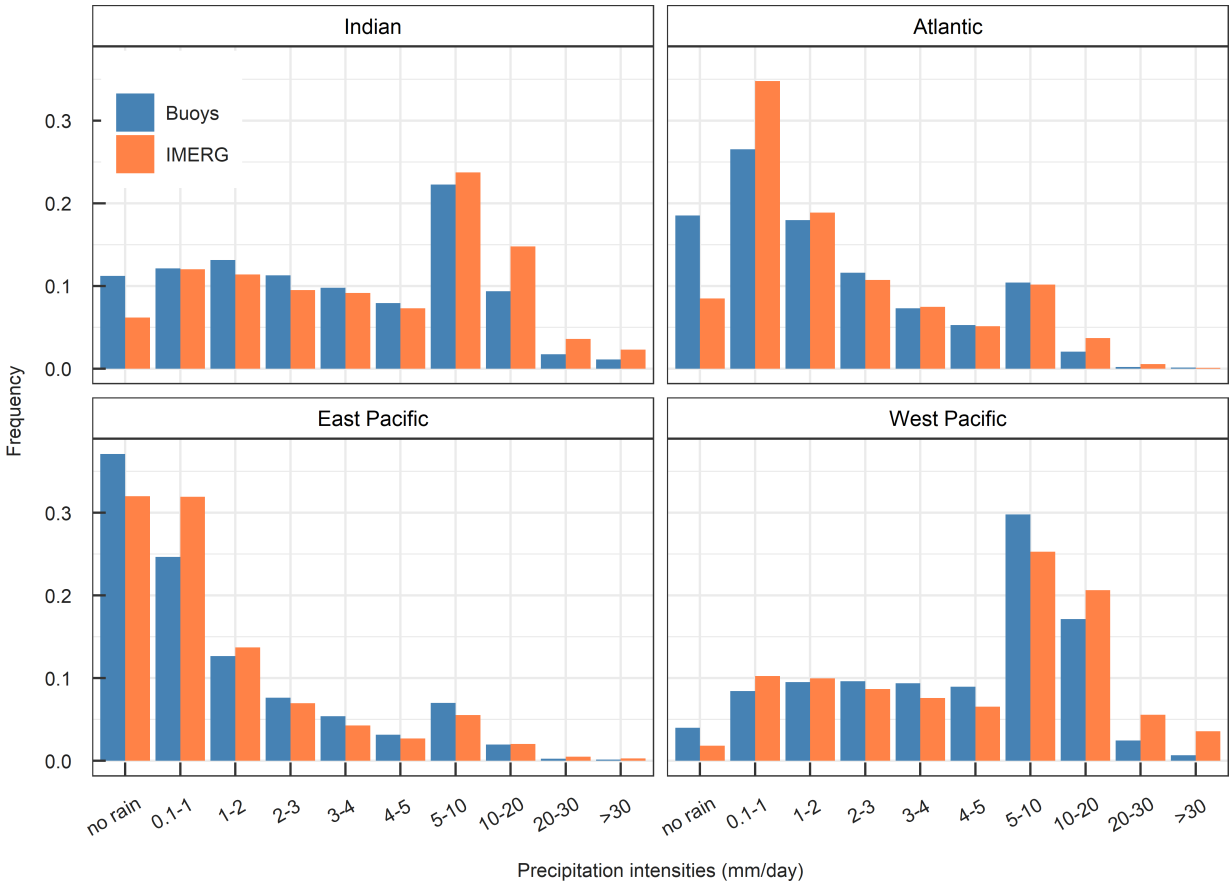


Figure 4.6: IMERG-F capability in reproducing the occurrence frequency of different precipitation intensities across the tropical oceans.

mentioned, the results reported here refer to all three IMERG runs. Overall, the positive hit bias and false alarm bias are the leading contributors to the total error. Across all oceanic basins, IMERG exhibits a total positive hit bias, with a maximum over the Indian Ocean (IMERG-F = 36%), followed by the West Pacific (IMERG-F = 32%), and the Atlantic (IMERG-F = 10%) oceans. However, the East Pacific is an exception, as the IMERG-F shows a slight negative hit bias (-2%). This indicates that, except for the East Pacific, IMERG predominately overestimates precipitation magnitude from correctly detected events. In terms of error contribution, similar to the previous results, the Indian and West Pacific have shown a more or less similar error pattern, and so have the Atlantic and East Pacific. Even though the Indian and West Pacific have a similar total bias ($>$ 30%), a significant difference exists in the contributions from individual errors; the false alarms and missed biases over the Indian Ocean are almost twice that of the West Pacific. Nonetheless, the cancellation among the positive hit bias, negative missed bias, and positive false alarm bias lead to a total positive bias similar to the West Pacific.

The error components show a similar pattern of contribution over the Atlantic Ocean, with hit bias and false alarm bias being the major contributors. In contrast, the East Pacific exhibits a different behaviour for IMERG-F. Unlike the other oceanic basins, where the hit bias is the major contributor to the total error, the East Pacific region shows that false alarm bias (7%) and missed bias (-2%) are the major contributors of IMERG-F's total bias. For IMERG-E and IMERG-L, hit bias contributes relatively more than other biases, consistent with the pattern observed in other oceanic basins. Moreover, here, all the IMERG runs have higher false alarms and missed bias compared to the Indian and West Pacific oceans. Since the accurate detection of light precipitation has remained a challenging task for the satellite estimates (Li et al., 2021b), the dominance of light precipitation over the East Pacific and Atlantic oceans can be attributed to the higher false alarms and missed bias over these regions. Moreover, the final IMERG runs outperform the early and late runs, while the differences between the latter two are negligible.

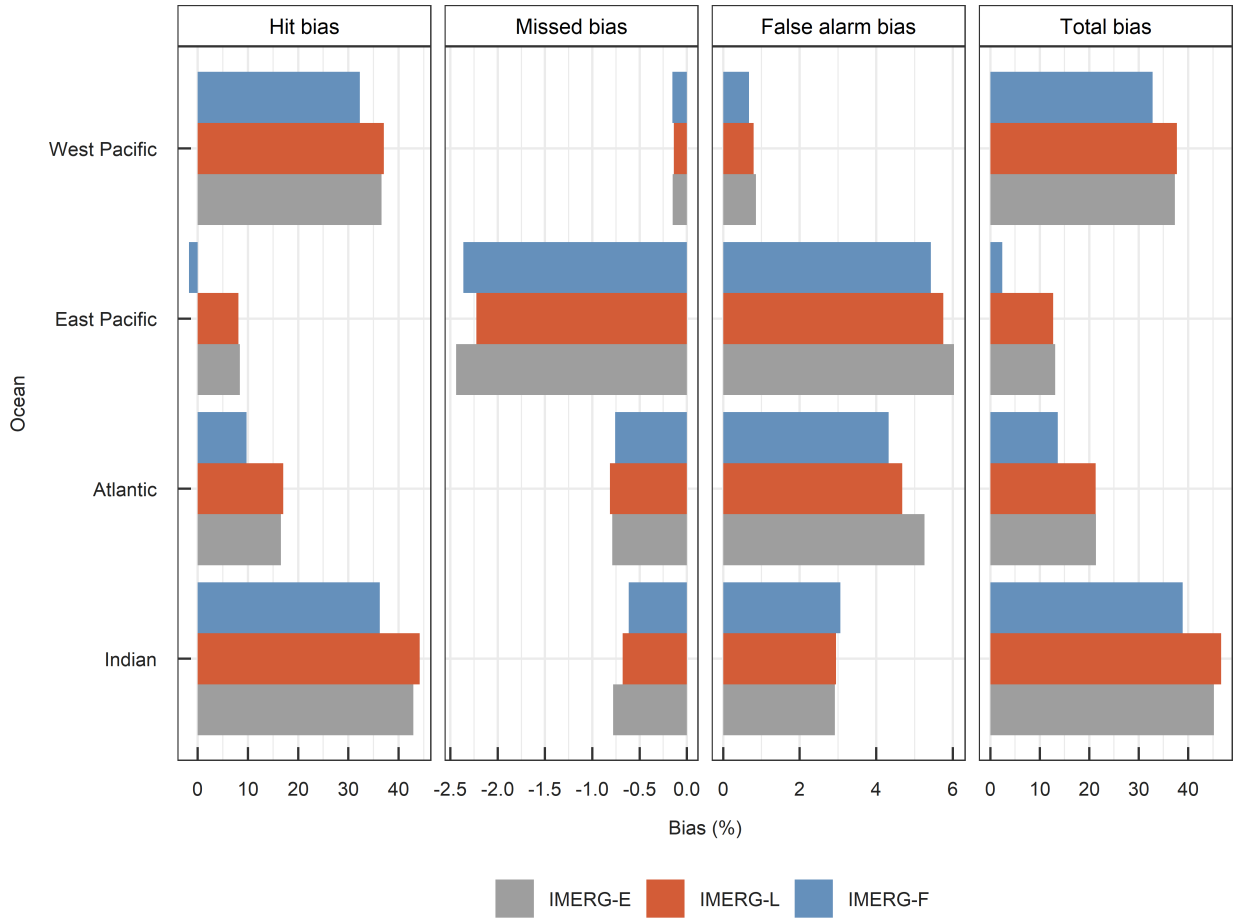


Figure 4.7: Decomposition of the IMERG total bias into hit bias, missed bias, and false alarm bias (shown in percentage relative to their corresponding buoys precipitation).

4.5.3 IMERG performances at extremes

The performance of IMERG-F in detecting different rainfall intensities is also assessed against the buoys throughout the tropical oceans (Figure 4.8). In terms of POD, IMERG-F performs reasonably well in detecting precipitation events until the 25 percentile (values > 0.8). However, the detection capability of IMERG-F decreases gradually with increasing percentiles. Notably, a clear distinction can be observed between the Indian and West Pacific, and Atlantic and East Pacific oceans. This indicates IMERG-F's relatively better precipitation detection over the West Pacific and Indian oceans than in the Atlantic and East Pacific. Additionally, POD scores falls below 0.5 for the Atlantic and East Pacific when the precipitation reaches the 75 percentile. For the Indian Ocean, the POD scores remain above 0.5 beyond the 95 percentile, and until the 99 percentile

for the West Pacific. These findings suggest that IMERG-F performs better at detecting extreme precipitation events over the West Pacific and Indian oceans compared to the East Pacific and Atlantic.

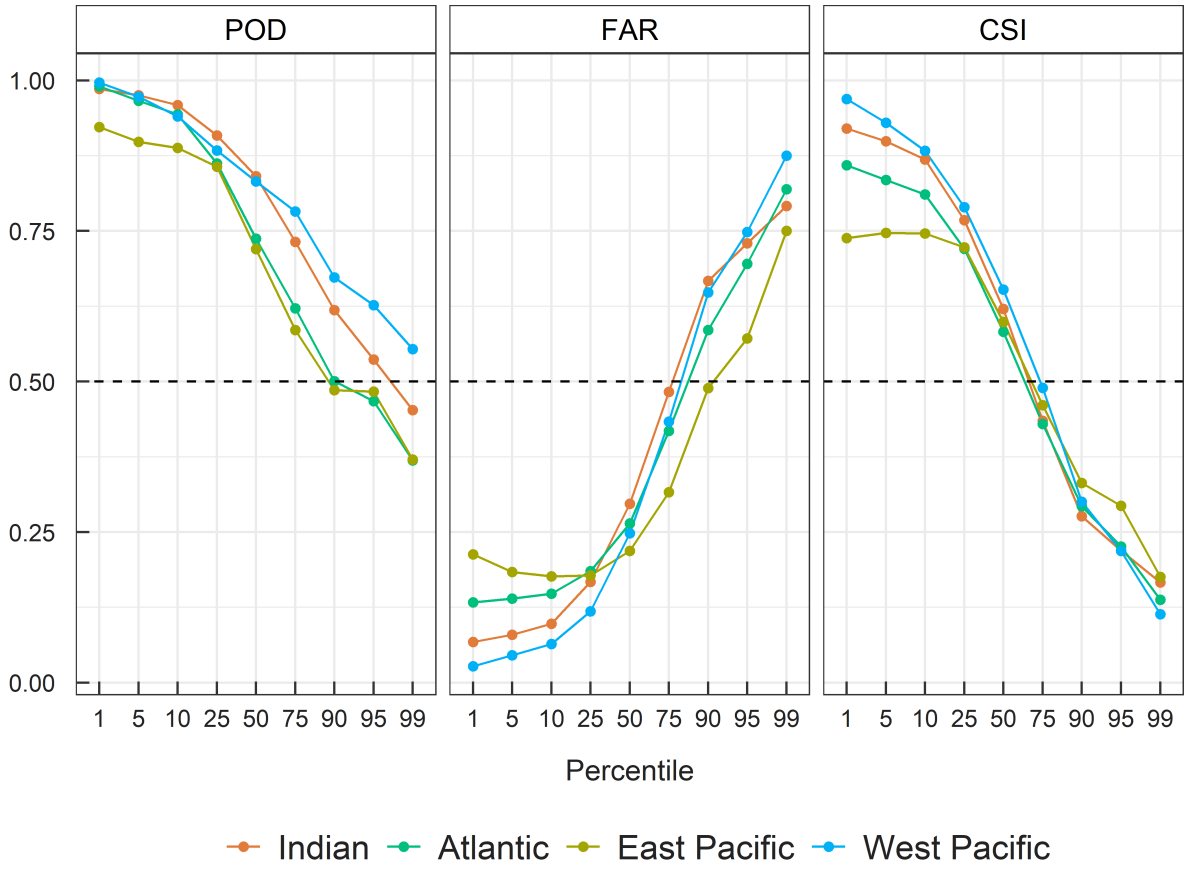


Figure 4.8: Detection capability of IMERG-F at different rainfall thresholds over the tropical oceans.

In the case of FAR, precipitation between 1 to 10 percentiles exhibits the highest variability among the oceanic basins, with the East Pacific showing the highest FAR, followed by the Atlantic, West Pacific, and the Indian Ocean. However, above the 25 percentile, IMERG-F follows a similar pattern for all the oceanic basins; a steady increase in FAR with percentile, though the increase for the East Pacific is somewhat lower. Subsequently, IMERG-F shows higher false alarms for precipitation percentiles exceeding 75, with FAR values greater than 0.5. This indicates that the capability of IMERG-F to detect extreme precipitation comes at the cost of higher false alarms, especially for the Indian and West Pacific oceans. In contrast, for the East Pacific, FAR remains below 0.5 until the 90 percentile, indicating that IMERG-F exhibits fewer false alarms for extreme

precipitation events in this region than in other oceanic basins.

The CSI scores are also very similar to the POD, at least until the 25 percentile, with the West Pacific performing the best, followed by the Indian, Atlantic, and East Pacific. In addition, it can also be noticed that for the 1 to 10 percentile, East Pacific has a lower CSI with values around 0.75. However, starting from the 25 percentile, all oceanic basins have a similar pattern; gradual decline in CSI with percentiles. The values also fall below 0.5 when the precipitation reaches the 75 percentile. Moreover, the CSI values are below 0.25 for all the oceanic basins for the 95 to 99 percentile. Nonetheless, the East Pacific displays a slightly better overall CSI score than other basins despite having the lowest POD for high percentiles. This could be due to the IMERG-F's lower false alarms for the higher percentiles over the East Pacific than the other basins.

4.6 Discussion

Overall, IMERG-F captures the spatial properties of oceanic precipitation well. Still, some substantial biases appear. Our results indicate that the biases associated with hit and false alarm events are the major contributors to the total error over the tropical oceans. Notably, the hit bias is the leading cause of the total positive bias observed over high-precipitation regions, such as India and the West Pacific. However, over the East Pacific, and especially for the IMERG-F, false alarm bias (-5%) is slightly higher than the hit bias (-2%). This is due to the excessive overestimation of the light precipitation (0.1 - 2 mm/day) events over the Atlantic and East Pacific oceans (Figure 4.6). Therefore, despite the higher false alarms, IMERG-F yet better estimates the total precipitation over the Atlantic and East Pacific oceans. On the other hand, the overestimation of extreme precipitation frequencies is the main reason of significant overestimation observed over the high-precipitation regions.

IMERG's issues with overestimation (underestimation) of light (heavy) precipitation, are not uncommon and have also been observed by other studies (Ehsani et al., 2022; Li et al., 2022; Tan et al., 2021). In particular, the overestimation of light precipitation has been reported over coastal regions (Derin et al., 2022) and land in general (Kazamias et al., 2022; Su et al., 2018), especially in arid and semi-arid climates where light precipitation often evaporates before reaching the sur-

face. IMERG tends to over-detect the occurrence frequency of such light precipitation events in these regions (Guo et al., 2016; Tang et al., 2016a). Additionally, despite the point-pixel measurement mismatches, the occurrence of virga (i.e., precipitation that evaporates before reaching the surface), may contribute to IMERG's overestimation of light precipitation over the tropical oceans (Prakash et al., 2018a). Additionally, buoy underestimation due to wind under-catch error, which is more prevalent in light precipitation events, may also be a contributing factor.

One probable reason for the observed biases in precipitation estimates could be the lack of appropriate gauge density to accurately represent the small-scale precipitation. For instance, sometimes, the extreme/light events are very small in horizontal scale and cannot be captured by the sparse gauge networks, leading to an underestimation of the actual precipitation by the gauges. Another potential explanation could be due to tropical cyclone-related extreme events. For instance, it can also be possible that the high wind speed associated with tropical cyclones further exacerbated the under-catching issue of oceanic buoys. As a result, when compared to the buoys, IMERG may appear to overestimate precipitation rates and total precipitation. This is further confirmed by Prakash et al. (2018a) who found that IMERG consistently reported higher precipitation during three tropical cyclones over the Indian Ocean.

Furthermore, a decreasing trend in IMERG detection scores was also found with an increasing precipitation threshold (Figure 4.8). This can be partially attributed to the fact that when the threshold increases, the total number of precipitation events (H+M+F) decreases, and is compensated by an increase in the number of non-precipitation events or correct negatives (Table A1). Consequently, the number of hits decreases compared to the number of misses and leads to an overall decreasing POD. Similarly, in the case of FAR, the higher number of false alarms than the number of hits leads to an overall higher FAR. The lower detection scores and higher FAR with increasing precipitation threshold can also be due to the inability of the satellite to detect the exact precipitation threshold. In this case, although IMERG may not detect events within the given threshold, it did detect events with a little more or less from the given thresholds, resulting in lower detection scores but contributing to the total amount. These results are in line with Manz et al. (2017)(V05) and Wu et al. (2019) though the study regions are over the tropical Andes and Yangtze River Basin in China correspondingly. Similar results were also reported by Rojas et al.

(2021) over the mountain regions of south-central Chile, and by Retalis et al. (2018) (V04A) over Cyprus.

An overall overestimation of IMERG V06 monthly estimates was also reported by Bolvin et al. (2021) over the Pacific Ocean. In contrast, Khan and Maggioni (2019) found an overall underestimation by IMERG V05. However, they evaluated IMERG against the OceanRAIN, and the results were applied to the entire global ocean instead of specific regions. Similarly, IMERG V06 underestimation of precipitation rate against the radar precipitation was also reported over the Kwajalein region of the Central Pacific (Wang et al., 2022). In addition, our findings show good agreement with the overestimation of the IMERG-final product over the Pacific Ocean, but disagree with the underestimation reported for the Atlantic and Indian oceans by Wu and Wang (2019). The authors state that the IMERG-E product aligns well with buoys, although their study was limited to three years (April 2014–2017) and focused on IMERG V05.

However, it can also be noted that most of the above-mentioned IMERG's underestimations over the ocean were for the previous versions (i.e., earlier than IMERG V06). Compared with the previous versions, the IMERG V06 has introduced several changes in the inter-calibration, Kalman filter, and morphing system, etc. (Tan et al., 2019a,c). Although all of these changes are intended to improve the IMERG V06 even further, the possibility of causing instead an overall overestimation over the tropical oceans cannot be ignored. In fact, Wang et al. (2022) reported an increase in FAR in the IMERG V06 compared to the IMERG V05 over the central Pacific Ocean. This remains an open question as a comparison with the previous IMERG versions is not part of this analysis which focuses on the evaluation only of the last IMERG version. However, the IMERG V06 overestimation is most likely due to the implementation of a new microwave satellite inter-calibration scheme, as well as the discovery of swath-dependent precipitation biases in satellite microwave estimates. It is expected that these issues will be addressed in the upcoming IMERG V07. Moreover, the Scheme for Histogram Adjustment with Ranked Precipitation Estimates in the Neighborhood (SHARPEN) is formulated to preserve precipitation rates (Tan et al., 2021), which in general is smoothed by the averaging of precipitation fields by the Kalman filter (Rajagopal et al., 2021). Ground validation has shown that the SHARPEN scheme improves performance and increases detection skills, suggesting that its implementation in the IMERG V07

will further reduce false alarms.

4.7 Conclusion

This chapter is focused on the tropical oceans and aims to evaluate the IMERG precipitation against the in-situ buoys precipitation. In this perspective, we quantitatively evaluated the IMERG-E, -L, and -F precipitation products with the buoys observation for a common overlapping period of 2001 – 2020 on a daily scale. This evaluation was carried out through both point-pixel and basin-scale approaches. The main conclusions of the analysis are summarised as follows;

- Overall, IMERG represents well the spatial pattern of mean daily buoys precipitation throughout the tropical oceans (Figure 4.1). However, at the same time, it has shown significant overestimation in total precipitation estimation with its magnitude varying with precipitation regimes.

- IMERG is better in detecting the occurrence of precipitation over the high-precipitation regions, such as the West Pacific and the Indian Ocean, than the low-precipitation regions of the Atlantic and East Pacific (Figure 4.2).

- The opposite is true in terms of volumetric scores (Figure 4.2). IMERG estimates precipitation over the Atlantic and East Pacific better than the Indian and West Pacific. This suggests that despite IMERG's ability to correctly detect precipitation events over the West Pacific and the Indian Ocean, it tends to significantly overestimate precipitation amounts over these two regions. Moreover, IMERG shows a higher estimation error over the high-precipitation regions than the low-precipitation regions, irrespectively of the oceanic regions (Figure 4.4). However, over the Atlantic and East Pacific, relative errors exhibit an opposite trend; high-precipitation regions tend to have lower relative errors compared to low-precipitation regions (Figure A2).

- IMERG shows excessive overestimation of the occurrence frequency of light precipitation events (0.1 – 1 mm/day) over the Atlantic and East Pacific, and heavy precipitation events

over the Indian and West Pacific (>10 mm/day onwards) (Figure 4.6). This excessive overestimation of heavy precipitation events is the primary reason for IMERG's overall overestimation over the Indian and West Pacific oceans.

- With regard to the detection of precipitation events, IMERG's detection capability tends to deteriorate with increasing precipitation rates (Figure 4.8). This is strongly evident above the 75 percentile. Moreover, although IMERG exhibits good detection capability for extreme precipitation rates, it also comes with the expense of higher false alarms.

- The error decomposition reveals that IMERG's positive hit bias and false alarms are the major contributors to IMERG's overall overestimation throughout the tropical oceans (Figure 4.7). This is especially true for the Indian and West Pacific oceans.

- We found very slight differences among the different IMERG runs in terms of their performance, with IMERG-F outperforming -E and -L.

Our results reveal the challenge that it is extremely difficult to conclude whether the reported biases are due to buoy measurement error or IMERG retrieval deficiency. However, it provides a clear overview of the uncertainties encountered and their structural properties. This can offer valuable insight to remote sensing communities for further research inquiries and methodological improvements. Most importantly, it highlights the need for multi-source observation networks over the oceans that will provide a variety of independent sources for evaluation. The combination of buoys, satellite, radar, and ship-based data in a homogeneous network could help us further constrain the uncertainties presented here and get closer to the true estimates of oceanic precipitation globally. In other words, it will provide a more accurate representation of oceanic precipitation and will help to address the uncertainties and limitations associated with any single dataset.

Chapter 5

Diurnal characteristics of satellite precipitation estimates

5.1 Introduction

The diurnal cycle of precipitation plays an important role in global precipitation variability (Tan et al., 2019b). The sud-daily variation in precipitation is of high importance for various applications, including water resource management, agricultural, and disaster management, etc. In particular, extreme precipitation events that occur in a short duration are more prone to cause devastation, such as flash floods, landslides, soil erosion, etc (Fowler et al., 2021). Sub-daily estimates evidently play a crucial role in accurately forecasting such extremes.

In addition, it is also well known that uncertainties among the precipitation estimates are often higher at the sub-daily scale compared to their corresponding daily, monthly and annual scales. Climate models are also known to have substantial biases, primarily rooted mainly in the deficiency in representing these small features occurring on sub-daily scale (Fiedler et al., 2020; Trenberth et al., 2017). Furthermore, the sub-daily scale variation, although on a longer scale driven by solar radiation, is highly regional (Yang and Smith, 2006). Therefore, understanding the sub-daily scale precipitation and its diurnal variation at the global level provides a more comprehensive understanding of the overall precipitation properties of the region.

Given their fine spatial-temporal resolution, near-global coverage, and access to remote areas, the high-resolution multi-satellite estimates such as the Tropical Rainfall Measuring Mission (TRMM) (Yang and Smith, 2006), and Integrated Multi-satellitE Retrievals for GPM (IMERG) (Huffman et al., 2015) have extensively been used in the estimation of the diurnal variability at both regional (Tan et al., 2019b; O and Kirstetter, 2018; Dezfuli et al., 2017) and global (Watters et al., 2021; Watters and Battaglia, 2019b) scales. Recently, Tan et al. (2019b), for instance, evaluated the IMERG versions in terms of diurnal variation over the different regions of the globe. They find that IMERG V06 has shown a maturing ability to capture the diurnal variability than V05. In addition, it has also been reported that IMERG well represents the regional diurnal variability in a variety of climatic regions, such as over Africa (Dezfuli et al., 2017), Brazil (Afonso et al., 2020), China (Li et al., 2018) and the contiguous United States (CONUS) (O and Kirstetter, 2018). Furthermore, O and Kirstetter (2018) revealed the potential of IMERG as an alternative to the ground measures, even at a sub-daily scale over the CONUS. However, they also highlighted

the fact that there are some region-specific biases to be considered. At the global scale, IMERG has also shown promising results in capturing the key features of the diurnal cycle (Watters and Battaglia, 2019b). Moreover, IMERG has also been used as a reference to evaluate the ability of reanalysis and model datasets to represent the diurnal variability (Watters et al., 2021).

A substantial number of regional studies have inter-compared and evaluated various satellite products in terms of sub-daily scale at different regions of the world (e.g., Sapiano and Arkin 2009; Pfeifroth et al. 2016; Janowiak et al. 2005). For instance, Afonso et al. (2020), evaluated the diurnal cycle of satellite estimates (IMERG, GSMaP, and CMORPH) over South American Brazil. They reported that all the products provided a better representation of the diurnal precipitation cycle in regions with deep clouds generated by local thermal heating compared to regions dominated by shallow clouds. GSMaP products effectively capture the diurnal cycle of precipitation over the Indonesian Maritime Continent as noted by (Ramadhan et al., 2023), though significant differences emerge in regions with more than one peak. Shawky et al. (2019) examined the sub-daily (3h, 6h, and 18 h) GSMaP and IMERG precipitation estimates in the arid region of Oman, and found that GSMaP outperformed IMERG. However, issues were identified with light and extreme precipitation events. Recently, the GSMaP products have also been evaluated at an hourly scale by Lv et al. (2024) over mainland China, and reported the significant improvement of the gauge corrected versions compared to the near-real-time products. Furthermore, Haile et al. (2013) assessed CMORPH and TRMM datasets against the gauge observations in terms of representing the diurnal cycle of precipitation occurrence and rate over the Nile River basin. They observed the over-detection of precipitation occurrence over the lake surfaces, including water bodies and islands, while an underestimation occurred over the mountaintops. They also reported a superior performance of CMORPH in the diurnal cycle of precipitation rate than TRMM, especially over the Lake Tana basin. Similarly, Zhang et al. (2021) revealed the comparatively better performance of CMORPH followed by TRMM and PERSIANN over the Three Gorges Reservoir area in China at 6h and 12h scales. All three estimates are in close agreement with the observations in terms of the diurnal cycle, though PERSIANN exhibits some biases.

In addition, the reanalysis datasets, such as the ECMWF Reanalysis v5 (ERA5) (Hersbach et al., 2020), also have the capability of capturing the diurnal cycle given its long-term and consistent

high-resolution estimates at the global level. Beyond precipitation, it also has the advantage of providing other variables, such as temperature, humidity, water vapour, etc., at multiple levels. Therefore, the application of reanalysis datasets in studying the diurnal precipitation has extensively increased over the years (Chen et al., 2014; Jiang et al., 2021; Qin et al., 2021). Furthermore, several attempts have been made to evaluate the performance of ERA5 at various spatial-temporal scales (Nogueira, 2020; Beck et al., 2019). For instance, Beck et al. (2019) revealed that ERA5 has shown better performance in precipitation estimation than IMERG in complex regions; however, the opposite is true in regions characterised with short-lived convective systems. On the other hand, Sharifi et al. (2019) reported that IMERG outperforms the ERA5 at the complex terrain on the daily and monthly scale over Austria. Studies have also attempted to evaluate the ERA5 at a sub-daily scale (Kumar et al., 2021; Hong et al., 2021; Tang et al., 2020) and most of them find ERA5 has difficulties in estimation of diurnal precipitation when compared to IMERG.

Despite the increasing number of sub-daily and diurnal precipitation studies at the regional scale, such studies are rarely available at the global level. Moreover, to our knowledge, there have not been many studies that evaluated the various satellite estimates in terms of their diurnal variation and their performance at sub-daily scales over the global level (Dai et al., 2007). Therefore, concerning the research gap, here we aim to compare five state-of-the-art precipitation estimates such as IMERG, Global Satellite Mapping of Precipitation (GSMaP), Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN), Climate Prediction Center Morphing (CMORPH), and ERA5 estimates at their sub-hourly scales. This enables us to examine the region-specific strength and limitation of the precipitation estimates at the sub-daily scales. Moreover, the analysis paid particular attention in the comparison and quantification of the differences in the diurnal cycle of precipitation among the various estimates.

The chapter is organized as follows. Section 2 introduces the five datasets used in the analysis as well as the methodological approaches employed. Then, in Section 3, we present the findings of the analysis, starting with the spatial mean precipitation across the globe and their zonal distribution, and following with the diurnal cycle and its variation across the globe and among the datasets. In Section 4, we focus on the underlying mechanisms responsible for the observed results. Finally, we conclude this chapter, highlighting what we have learned from this

intercomparison, in Section 5.

5.2 Data and Methodology

Satellite estimates

The Global Precipitation Measurement (GPM) mission is a constellation of international satellites that aims to provide high-quality precipitation with quasi-global coverage. IMERG is a unique algorithm that merges and inter-calibrates precipitation estimates from a range of sources, such as Passive Microwave (PMW), Infrared (IR) and gauges in order to produce $0.1^\circ \times 0.1^\circ$ and 30 min precipitation products (Huffman et al., 2015). A substantial number of studies have validated the IMERG performance in a range of climatic conditions, and it performed extremely well in a wide range of applications (Pradhan et al., 2022). An overview of the IMERG algorithm is presented in Chapter 1.5, and more details regarding the IMERG precipitation estimation algorithms and their technical details can be found in Huffman et al. (2015). In this analysis, the IMERG V06B Final Run Half Hourly product is used.

GSMaP is a gridded multi-satellite precipitation estimates product developed jointly by the Japan Aerospace Exploration Agency (JAXA) and Japan Science and Technology Agency (JSTA) (Mega et al., 2018). GSMaP merges precipitation estimates from a range of several low earth orbit passive microwaves and geostationary IR precipitation sensors. Like IMERG, GSMaP also provides precipitation in near-real-time, as well as gauge-corrected final products. Nonetheless, unlike IMERG, GSMaP uses the Climate Prediction Centre (CPC) unified global daily gauge precipitation for gauge correction. Moreover, validation studies have reported consistence performance of GSMaP with the observation (Zhou et al., 2020; Lu and Yong, 2018), and relatively better than IMERG at least in a few cases (Li et al., 2021a; Ning et al., 2017; Salles et al., 2019). In the present analysis, we have used the GSMaP gauge corrected V08 product from 2001 – 2020, available at hourly and 0.1° resolutions.

On the other hand, the CMORPH products use low earth orbiter satellite passive microwave estimates in order to produce high-quality global precipitation estimates (Joyce et al., 2004; Joyce

and Xie, 2011). Considering the low-sampling nature of microwave estimates, it uses the geostationary Infrared (IR) image-derived information to propagate the precipitation system (i.,e CPC Morphing technique) in the absence of microwave estimates. CMORPH uses the CPC daily precipitation estimates over land and GPCP pentads merged analysis over ocean for bias correction (Xie et al., 2017). Even though CMORPH is available with very high spatial resolution i.,e, 0.07277 degrees lat/lon, the CMORPH bias-corrected V.1 product with a spatial resolution of $0.25^\circ \times 0.25^\circ$ and hourly temporal resolution is used in the current analysis (`https://www.ncei.noaa.gov/data/cmorph-high-resolution-global-precipitation-estimates/access/hourly/`).

Unlike the CMORPH, PERSIANN is mainly based on the geostationary infrared cloud images, providing precipitation at $0.25^\circ \times 0.25^\circ$ spatial resolution and hourly temporal scales (Hsu et al., 1997; Sorooshian et al., 2000). As its name suggests, PERSIANN uses artificial neural networks to estimate precipitation based on the cloud top temperature from the geostationary satellite-derived infrared images (Nguyen et al., 2019). It can also be noted that PERSIANN also uses the LEO satellites PMW estimates to continuously adapt the parameters of the model. The PERSIANN product used in the current analysis is obtained from the CHRS Data Portal at `https://chrsdata.eng.uci.edu`.

Reanalysis estimates-ERA5

ERA5 is the latest fifth-generation global atmospheric reanalysis product developed by the European Center for Medium-Range Weather Forecast (ECMWF) using the 4D-Var data assimilation techniques in cycle 41r2 (Hersbach et al., 2020; Jiang et al., 2021). Recently, ERA5 replaced its predecessor, i.e., the ERA-Interim reanalysis product. Compared to ERA-Interim, ERA5 has been updated with a more advanced data assimilation system and physical model, and more importantly, the spatial resolution is reduced to 31 km. In addition, ERA5 datasets are now available at an hourly scale and have extended to 1950. We have chosen ERA5 over the other global reanalysis products, because it has been reported to show better agreement with the observations of the ground stations (Keller and Wahl, 2021; McClean et al., 2021). Here, we have used the hourly

ERA5 reanalysis data with a spatial resolution of $0.25^\circ \times 0.25^\circ$.

Table 5.1: Summary of the datasets used in this analysis.

Dataset name	Spatial scale	Temporal scale	Record length	Reference
IMERG	$0.1^\circ \times 0.1^\circ$	0.5h	2000 – present	Huffman et al. (2018)
GSMaP	$0.1^\circ \times 0.1^\circ$	hourly	2000 – present	Mega et al. (2018)
ERA5	$0.25^\circ \times 0.25^\circ$	hourly	1950 – present	Hersbach et al. (2020)
PERSIANN	$0.25^\circ \times 0.25^\circ$	hourly	2000 – present	Hsu et al. (1997)
CMORPH	$0.25^\circ \times 0.25^\circ$	hourly	1998 – present	Joyce et al. (2004)

Methodology

The methodological approach includes an inter-comparison of sub-daily scale precipitation derived from multiple sources of precipitation datasets (Table 5.1). The multi-source precipitation datasets from the state-of-the-art satellite and reanalysis product at their original resolution (i.e., 30 minutes or hourly) are collected covering the global land and ocean between 60°N – -60°S. Considering the different temporal coverage of each dataset, a common overlapping period from 2001 – 2020 is selected as the analyzing period. Moreover, given the different spatial and temporal resolutions of the datasets, to have a consistent and fair analysis, all the estimates are converted into a common spatial and temporal resolution of $0.25^\circ \times 0.25^\circ$ and hourly scale.

The sub-daily scale evaluation among the estimates is based on the diurnal cycle. According to Watters et al. (2021), the first and second-order harmonics are often not efficient in capturing the diurnal variability. Therefore, in this analysis, we are not fitting any harmonic function or empirical orthogonal function to estimate diurnal parameters. Instead, the diurnal variability of global precipitation is estimated using three variables, namely, the mean precipitation amount, frequency, and intensity (Marzuki et al., 2021). The mean precipitation amount is estimated by accumulating all the hourly precipitation divided by the total available hours for each grid. For frequency, the total number of precipitating hours (precipitation $>$ 0.1 mm/hr) is divided by the total available hours. Finally, the intensity is estimated with the total precipitation divided by the precipitating hours (precipitation $>$ 0.1 mm/hr). As mentioned above, all these metrics are

estimated for each grid.

The mean precipitation for each latitude (ϕ) and longitude (λ) at the Universal Coordinated Time (UTC) time hour (t_{UTC}) is estimated by using the following equation (Eq. 5.1)

$$P(\phi, \lambda, t_{UTC}) = \frac{\sum_{i=1}^{N} Pi}{N} \tag{5.1}$$

Here, P_i represents the ith precipitation estimate of the study period, and N represents the total number of precipitation estimates (including no precipitation events).

The UTC hour of each precipitation dataset is converted to Local Solar Time (LST) by using the following equation (Eq. 5.2)

$$t_{LST}(h) = t_{UTC}(h) + \frac{\lambda(^{\circ})}{15(^{\circ}h^{-1})} \tag{5.2}$$

Furthermore, in order to examine how the diurnal shapes change at each grid level, the K-means clustering algorithm is employed. The K-means clustering method clusters similarly behaving pixels into a single cluster. To determine the optimal number of clusters, the process is iterated from K = 1 to K = 10, ultimately selecting k = 4 as the appropriate number of clusters based on distinct diurnal patterns. Subsequently, four clusters are extracted from each dataset to depict global diurnal precipitation variability. Finally, each cluster is named according to its peak hour of local solar time.

5.3 Results

5.3.1 Spatial distribution of mean hourly precipitation properties

In order to investigate the differences and similarities among precipitation products, first, we examined the spatial distribution of mean hourly precipitation amount, frequency, and intensities at 0.25° resolutions for the period of 2001 – 2020. The distribution of hourly mean precipitation

exhibits a consistent spatial pattern across the globe (Figure 5.1). In particular, visually, all of them depict similar spatial patterns characterized by high precipitation across the Intertropical Convergence Zone (ITCZ) and South Pacific Convergence Zone (SPCZ) belt, and low precipitation over the dry regions in the subtropical high and across the Sahara regions of African continents. As most of the precipitation datasets are somewhat dependent on each other, it is not surprising to have similar spatial mean precipitation among the datasets. However, small regional differences in the dry regions (e.g., southern Pacific Ocean, Southern Atlantic Ocean, and southern Indian Ocean near Australia) can be observed between the PERSIANN and ERA5. Unlike other data products, PERSIANN exhibits more widespread dry regions, while the opposite is true for ERA5. Considering the known issues of infrared (IR)-based estimates in accurately detecting precipitation generated by warm clouds (Behrangi et al., 2012), the underestimation of precipitation by PERSIANN over the tropical oceans, which are known for warm rainfall, could be attributed to this limitation.

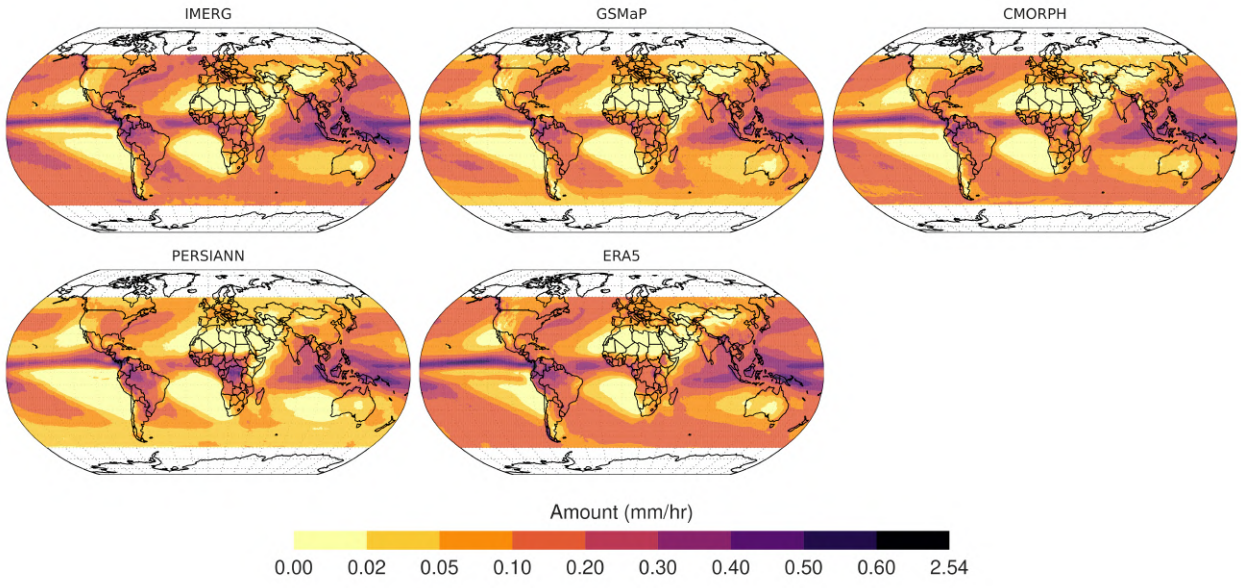


Figure 5.1: Spatial distribution of global mean (2001 – 2020) hourly precipitation amount (mm/hr).

In terms of the spatial distribution of hourly precipitation frequency, there is a more or less a similar pattern among the estimates (Figure 5.2), and it resembles those of mean precipitation amounts (Figure 5.1). Nonetheless, ERA5 appears markedly different from the remote sensing data products, showing substantially high frequencies across the globe. In particular, across the

tropical belts and, more precisely, over the ocean, ERA5 showed substantially higher precipitation frequency (40 – 90%) than the rest of the estimates. Similar to the mean precipitation amount, the frequency of hourly precipitation appears to be quite low in PERSIANN, particularly over dry regions such as the subtropical high and the Sahara regions of the African continent. Additionally, compared to the IMERG and GSMaP, CMORPH also exhibits relatively lower frequencies across the globe.

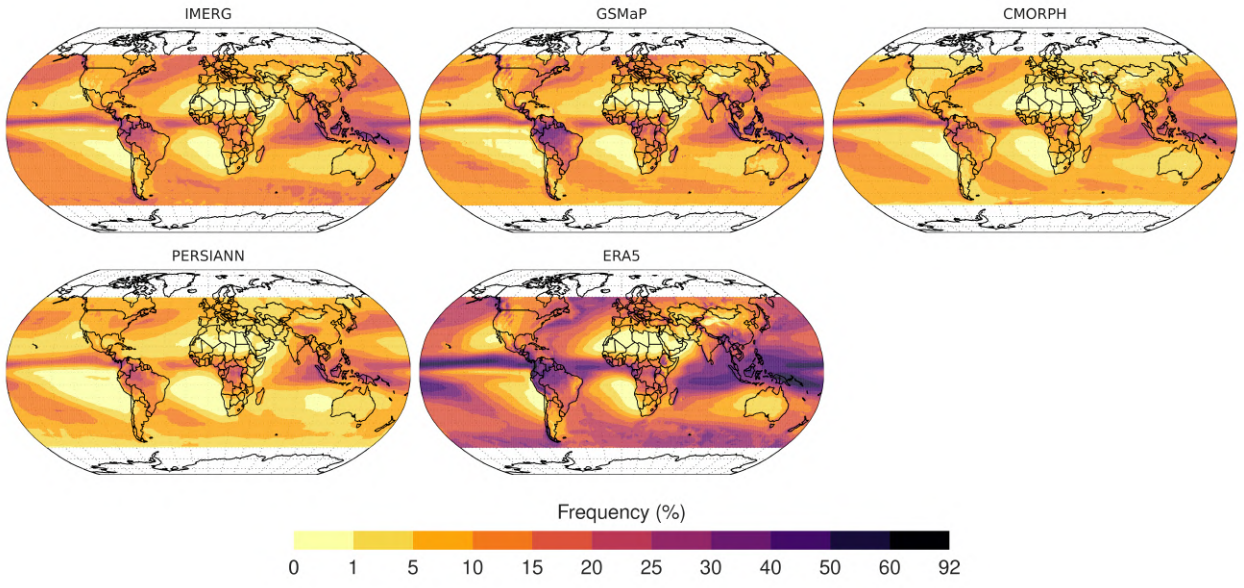


Figure 5.2: Same as Figure 5.1, but for precipitation frequency (%).

Unlike the mean precipitation amount and frequency, the spatial pattern of mean precipitation intensity is not very homogeneous among the datasets (Figure 5.3). As expected, the mean intensity of precipitation seems comparatively very low in the ERA5 estimates compared to the rest of the datasets. It further confirms that models and reanalysis products exhibit high-frequency, low-intensity issue, a concern that has been extensively reported over the years (e.g., Watters et al. 2021; Qin et al. 2021). In terms of satellite estimates, there is a relatively good agreement between the IMERG and CMORPH throughout the globe, compared to the GSMaP and PERSIANN. In fact, surprisingly the mean intensity of precipitation for GSMaP also appears to be relatively low, especially over global land. Moreover, it can also be noted that CMORPH exhibits relatively higher intensities, despite having low frequency. This could be attributed to the relatively lower fraction of light precipitation events in the CMORPH.

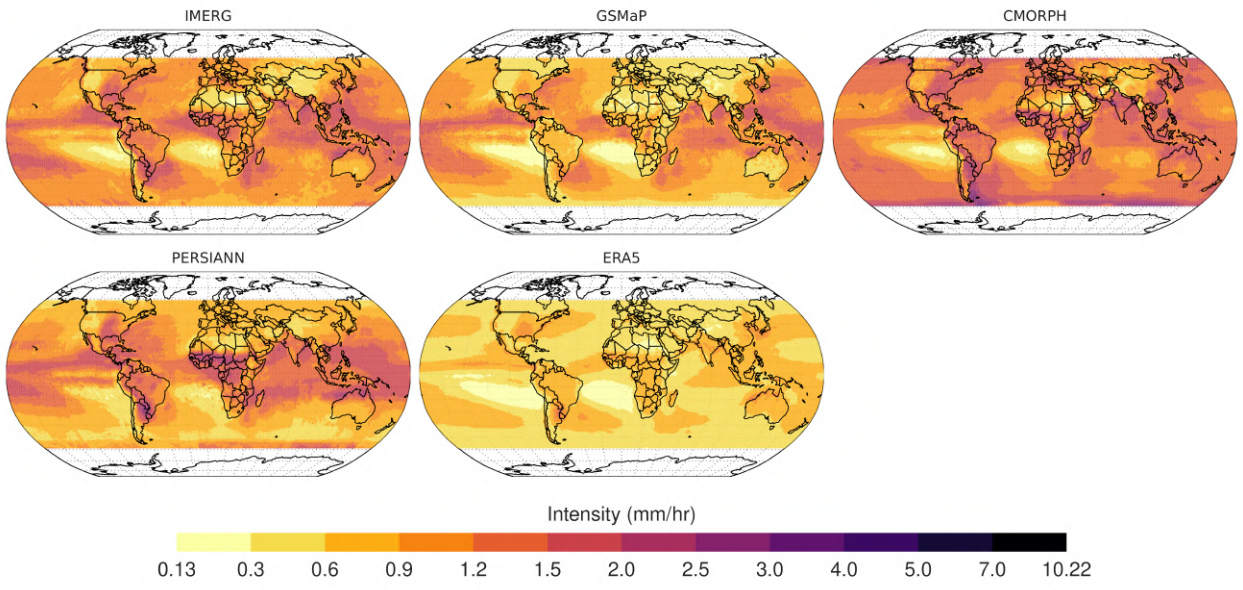


Figure 5.3: Same as Figure 5.2, but for precipitation intensity (mm/hr).

To further explore and compare the various estimates, the latitudinal average of the mean precipitation amount, frequency, and intensities are also examined (Figure 5.4a, 5.4b and, 5.4c). In terms of hourly mean latitudinal precipitation, all the products agree and exhibit a similar pattern with a peak at the ITCZ followed by a minimum in sub-tropical belts and so on (Figure 5.4a). At the ITCZ, and between 0°N to 10°N, in particular, all datasets accurately depict the peak. However, ERA5 (0.3 mm/hr) and IMERG (0.27 mm/hr) show close agreement, which distinguishes them from the other products that have values greater than 0.25 mm/hr. Moreover, all the datasets have similar values from 35°N to -20°S, and after that, the uncertainties among the estimates increase with the latitude. Furthermore, compared to the northern hemisphere, the variability among the datasets is notably higher in the southern hemisphere (-20°S – -60°S). Additionally, while ERA5 exhibits the highest mean precipitation amounts with a peak at the ITCZ, IMERG surpasses ERA5 in extra-tropical belts (-20°S to 20°N) in the southern hemisphere, followed by CMORPH. Conversely, in the northern hemisphere, ERA5 retains the highest mean precipitation amounts. GSMaP aligns closely with CMORPH and PERSIANN, in contrast to IMERG. In the southern hemisphere, GSMaP has shown a sharp decline from the -30°S latitudes until the -60°S.

In terms of frequency, ERA5 exhibits a spatial pattern similar to other datasets throughout the latitudes (Figure 5.4b). However, its frequency estimation is often significantly higher than that

of other datasets across all latitude zones. At the ITCZ belt, where the peak frequency occurs, ERA5 has a frequency reaching up to 40%, which is almost double that of the rest of the products ($< 20\%$). Furthermore, among the latitudinal zones, $20^\circ - 30^\circ$ is the only region in both the hemispheres, where the difference between the ERA5 and the rest of the datasets is relatively minimal. When it comes to the remote sensing estimates, IMERG and GSMaP have a very close agreement throughout the latitudinal zones. However, from -40°S onward, the difference between IMERG and GSMaP keeps increasing with the latitude. GSMaP exhibits a sharp decline in a manner similar to the mean precipitation amount (Figure 5.4a). Although the PERSIANN and CMORPH are in close agreement throughout the latitudinal zones, PERSIANN remained lowest among the estimates, particularly in the southern hemisphere.

As a consequence of the high frequency, the intensity of ERA5 remains the minimum among the datasets throughout the latitudinal zones (Figure 5.4c). In fact, the highest uncertainties among the estimates are observed in terms of representing the precipitation intensity. Again, compared to the northern hemisphere, the discrepancies are highest over the southern hemisphere, with the highest occurring towards the higher latitudes. PERSIANN shows the highest intensity over the ITCZ belts (-20°S to 20°N), with values up to 1.5 mm/hr, followed by CMORPH, IMERG, and GSMaP. However, from the extra-tropical regions, especially from 20°N/S on-wards, CMORPH shows the highest intensity with an increasing trend with the latitudes, which is quite the opposite of the rest of the products. Despite having low frequencies, both PERSIANN and CMORPH show high intensity, probably due to the missing light precipitations.

Overall, all the datasets effectively capture the spatial variability and distribution of mean precipitation, frequency, and intensity. However, the zonal plots reveal significant uncertainties among them, especially towards the high latitudes. In fact, the challenge of precipitation retrievals toward high latitudes has been reported in previous studies as well (Protat et al., 2019; Grecu et al., 2016; Skofronick-Jackson et al., 2017). ERA5 exhibits the highest frequency and lowest intensity, while CMORPH depicts the highest intensity. The remaining datasets fall in between, contributing to the observed variations.

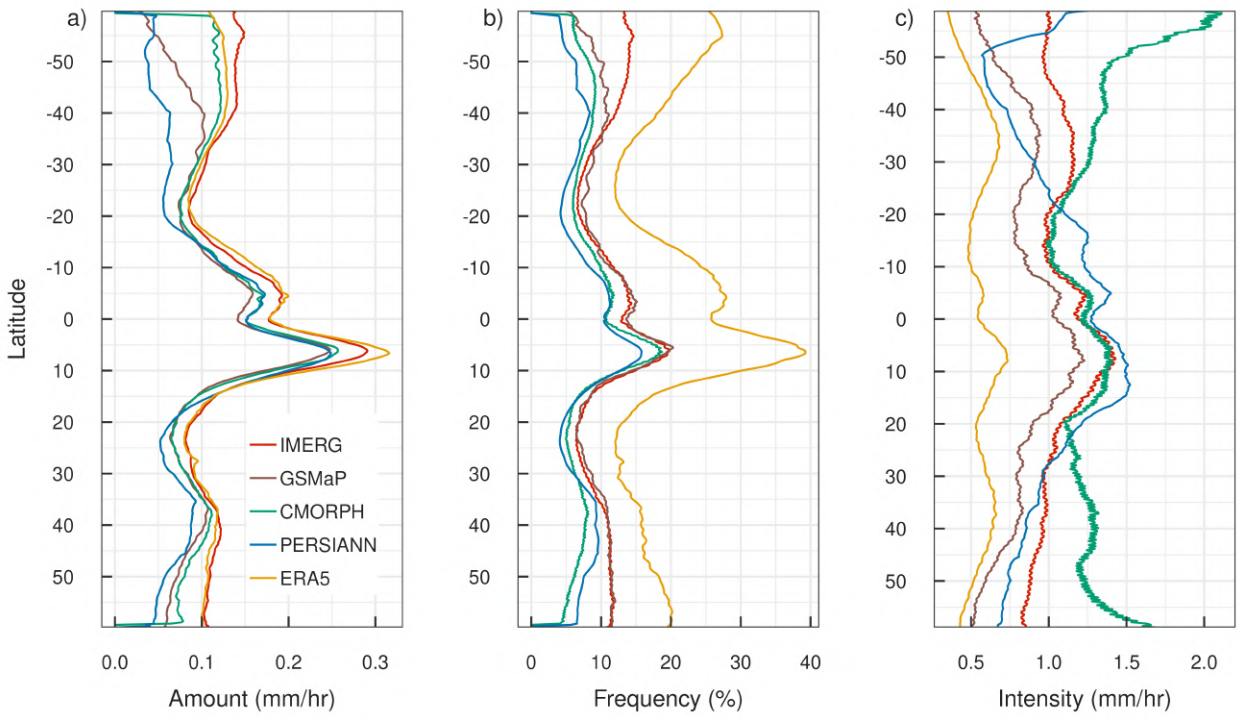


Figure 5.4: Latitudinal average of mean hourly precipitation a) amount (mm/hr), b) frequency (%), and c) intensity (mm/hr).

5.3.2 Diurnal variation

5.3.2.1 Diurnal mean

The diurnal variation of mean precipitation amount among the datasets is quite similar in shape (Figure 5.5a). In other words, all the products agreed well in terms of producing specific features of the diurnal variation over the globe: an afternoon peak over the land and an early morning peak over the ocean. In addition, a bimodal peak with peaks in the early morning (from the ocean) and afternoon (from land) can be observed at the global level. These diurnal results are consistent with previous studies (Dai et al., 2007; Watters et al., 2021). However, significant differences exist among the estimates as well. To start with, at the global level, IMERG and ERA5 look quite close to each other, whereas the CMORPH is in between, and GSMaP and PERSIANN are at the bottom. GSMaP behaves quite differently from the rest of the products, exhibiting multiple peaks throughout the day. Over the ocean, the behaviour of the products is also quite similar to the global level, except for one difference: the early morning peak. This also indicates that the ocean

diurnal cycle dominates at the global level, which is expected as the ocean receives the lion's share of global precipitation compared to land.

Over the land, all the products well reproduce the afternoon peak, a common feature of the diurnal cycle and has been consistent with other studies over the years. ERA5 shows the peak a little earlier, around 15 LST over land, compared to the other estimates which are mostly between 16 to 18 LST. The earlier peak from the ERA5 reanalysis and other model-generated precipitation is, in fact, not uncommon (Hayden et al., 2023). Additionally, ERA5 also shows the highest peak with a mean precipitation of around 0.18 mm/hr, followed by PERSIANN (0.15 mm/hr), IMERG (0.14 mm/hr), CMORPH (0.11 mm/hr) and GSMaP (<0.1 mm/hr). While ERA5 shows the peak and diurnal cycle slightly earlier than other datasets, the uncertainty among them is greater between 11 to 18 LST, whereas it is minimal during other times, such as early morning and night.

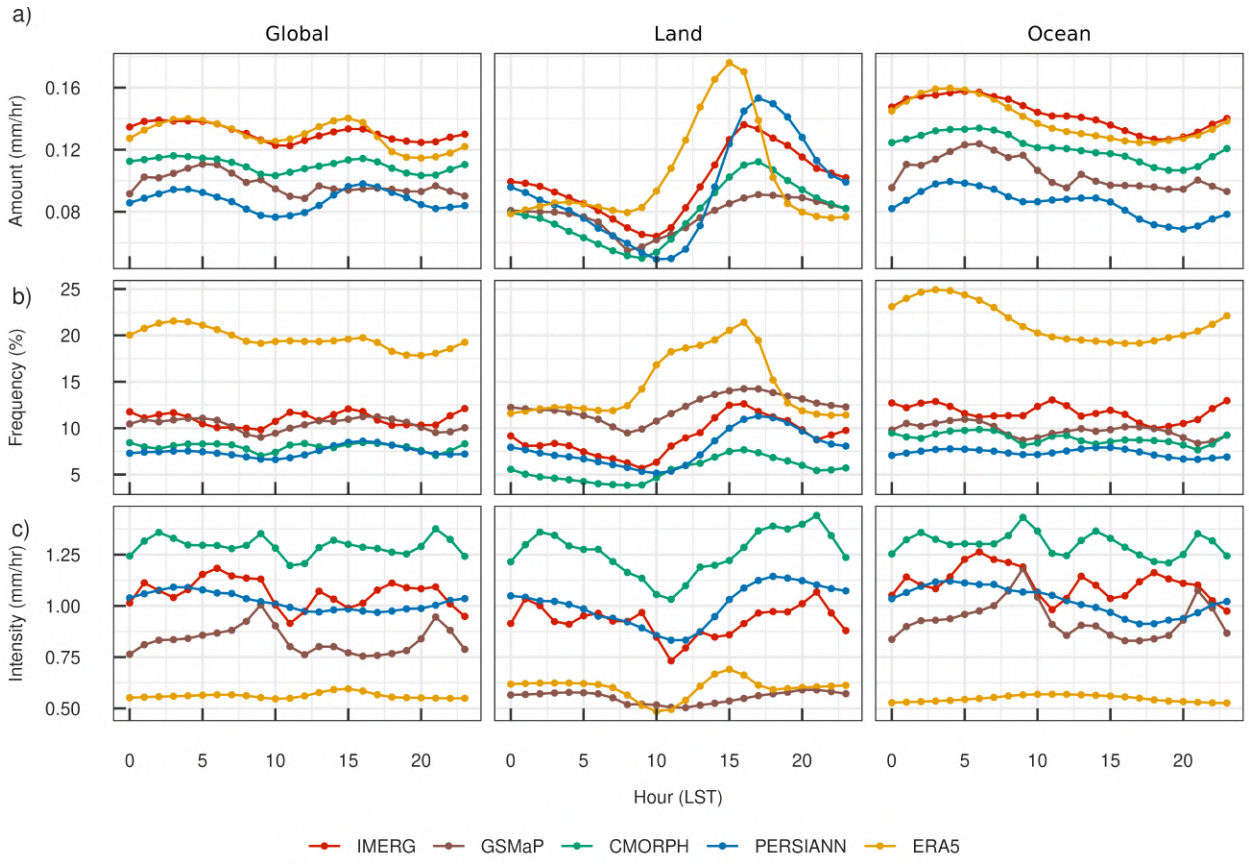


Figure 5.5: Diurnal variation of precipitation a) amount (mm/hr), b) frequency (%), and c) intensity (mm/hr), of each dataset for 2001 – 2020.

5.3.2.2 Diurnal frequency

The diurnal variation in precipitation frequency across the datasets looks quite different compared to the mean precipitation amount (Figure 5.5b). As expected, ERA5 displays considerable deviation from the other datasets, irrespective of the region (i.e., globe, land, ocean). Despite the significant overestimation, ERA5 exhibits a peak during the afternoon around 16 LST, aligning with the peak observed in other satellite precipitation datasets over land. At the global level, and over the ocean, the variation in diurnal frequency is not particularly pronounced across the satellite datasets. However, ERA5 shows a distinct peak in frequency early in the morning (between 03 to 04 LST), with the diurnal variation being more predominant over the ocean. All the other estimates show little variation (a flatter shape) with multiple peaks throughout the 24-hour period. On the contrary, over land, the peak of diurnal frequency is between 15 LST to 16 LST. Regarding amplitude, while all datasets show similar peaks in the afternoon, substantial disparities among them are apparent. ERA5 exhibits relatively high discrepancies (exceeding 20%) during the daytime from late morning (08 LST) until late evening (19 LST). In terms of the satellite products, their agreement varies with the region. At a global level, IMERG and GSMaP appear similar, as do CMORPH and PERSIANN. Over land, notably, GSMaP is comparatively high, followed by IMERG, PERSIANN, and CMORPH.

5.3.2.3 Diurnal intensity

Unlike the diurnal mean precipitation amount and frequency, the diurnal intensity is not so pronounced, and thus, it does not exhibit a clear diurnal variation or pattern (Figure 5.5c). Peaks are not distinctly evident globally and over the ocean. However, over land, a bimodal peak can be seen, with an early morning peak between 00 – 05 LST, and a late afternoon peak at 15 – 21 LST.

In terms of different precipitation products, it is evident that CMORPH exhibits the highest intensity (exceeding 1.25 mm/hr) regardless of whether it is over land, ocean, and globally throughout the 24-hour period. Following, IMERG shows the highest intensity over global and oceanic levels, whereas it is PERSIANN over land. As expected, ERA5 exhibits the minimum intensity throughout the hours, with a very weak diurnal variation over the ocean and globe

compared to land. Despite having a similar pattern, GSMaP does not follow IMERG and has the lowest precipitation intensity over land, even lower than ERA5. GSMaP consistently has the lowest intensity after ERA5, whether over global, land or ocean regions. It also has the largest discrepancy with the other datasets, particularly over land. This behavior of GSMaP over land is notable, considering both IMERG and GSMaP use a similar constellation of satellite estimates. Nevertheless, it should also be noted that both datasets use different gauge corrections over land. IMERG uses Global Precipitation Climatology Centre (GPCC) corrections on a monthly scale, while GSMaP uses CPC corrections on a daily scale, and this could be the probable reason for such observational differences between the datasets.

5.3.3 Peak hour of diurnal mean precipitation amount, frequency, and intensity

To further investigate the timing of maximum precipitation properties (i.e., amount, frequency, and intensity) and their variations across different climates and topographic regions, the peak hours are also examined (Figure 5.6, 5.7, and 5.8). The peak hour denotes the hour at which the maximum precipitation properties occur at each grid. Regarding the peak hour of precipitation amount, the continental/land regions are mostly dominated by the evening peak hours (15 – 18 LST), compared to the early morning peak hours over the ocean (02 – 06 LST) (Figure 5.6). However, there are regions, particularly over land, where a slightly inhomogeneous distribution of peak precipitation hours is observed. This inhomogeneous distribution is mainly seen in dry regions such as Africa, Australia, and the Middle East. Elevation appears to have a significant impact on the timing of peak precipitation hours. In high-elevation mountain regions such as the Himalayas and Andes, peak hours tend to occur in the early morning, in contrast to the early afternoon peak hours observed in the surrounding land. This phenomenon is likely due to orographic precipitation.

The oceanic regions are mainly characterized by midnight to early morning peak hours (00 – 06 LST). In fact, the high precipitation regions are mainly dominated by early morning peak hours, approximately between 03 – 07 LST, while the dry regions, such as the Atlantic and Pacific warm

pools, are a bit earlier, between 01 to 03 LST. It is even earlier, around 22 to 01 LST, towards the high latitudes (pole wards). Especially over the southern hemisphere, which is clearly observed in IMERG and CMORPH but not so clearly by GSMaP.

In the coastal regions near the land, precipitation peaks in-between, i.,e 06 – 12 LST, and as it progresses towards the land, the peak hours keep increasing and reach the typical late-afternoon/early evening peaks (15 – 18 LST). The exact opposite pattern follows towards the ocean with precipitation peak in the early morning (03 – 06 LST), although regional differences exist among the estimates (e.g., over the Southern Ocean). Similar results were also observed in previous studies, such as by Hayden et al. (2023) and Bai and Schumacher (2022) over the maritime continents.

In terms of agreement among different datasets, most of them portray similar spatial patterns, although some noticeable differences exist among them. In particular, ERA5 differs slightly from the rest of the datasets, with an early peak hour over both the land and oceanic regions. PERSIANN is quite different from the other estimates, which is expected as it is solely derived from comparatively less accurate IR sensors. IMERG and CMORPH show a high degree of similarity, while GSMaP depicts noticeable differences. In particular, IMERG and CMORPH agree well, not only in producing peak hours in the early morning hours over the ocean and in the late afternoon over land, but they are also consistent in depicting the small regional differences. For instance, both IMERG and CMOPRH agree on the occurrence of midnight peak hours in the high-latitudes Southern Ocean and various dry regions, such as the sub-tropical region of the southern Atlantic and Pacific, as well as over the Africa and Middle East regions. This agreement could be due to the fact that both IMERG and CMORPH have relied on microwave estimates, and, more importantly, the CMORPH algorithm is a crucial component of IMERG.

Over land, similar to the peak hour of mean precipitation amount, precipitation frequency peaks during the afternoon between 14 – 18 LST (Figure 5.7). However, over the ocean, depending on the region, it varies with time and datasets. The discrepancies among the estimates are larger over the ocean than over land. For instance, as observed by IMERG, sub-tropical dry regions such as the Pacific and Atlantic warm pool zones show nighttime frequency peaks around 02 – 03 LST, while in the high precipitation ITCZ belts, peaks occur during the day (11 – 12 LST). In the high

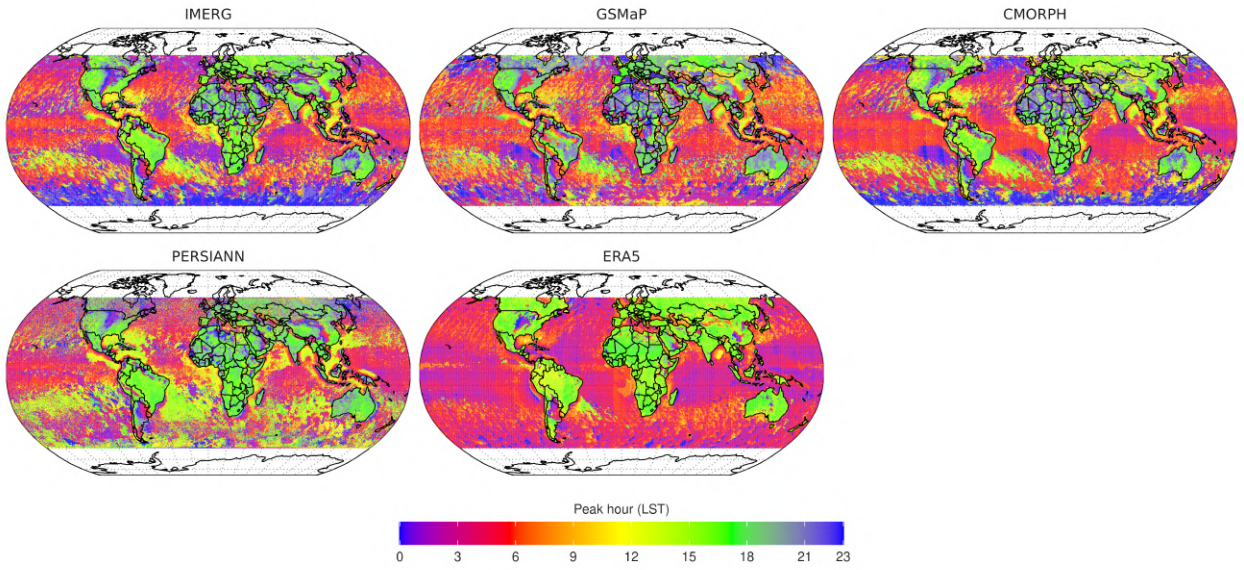


Figure 5.6: Peak hour of mean precipitation for each dataset during 2001 – 2020.

latitude zones (40° – 60°), the northern hemisphere experiences midnight peaks (22 – 01 LST), whereas the southern hemisphere peaks during midday (10 – 12 LST). All these differences in peak hour over the ocean by IMERG are not consistent across other datasets. However, IMERG and CMORPH agree over the high latitude northern oceans (> 40°N) with nocturnal peaks during 21 – 02 LST. Nonetheless, discrepancies arise on the southern oceans, as IMERG shows late-morning/midday (09 – 12 LST), while CMORPH is nocturnal with few midday peaks. In fact, GSMaP also depicts a similar pattern; however, it is observed in the early morning (03 – 06 LST), and the nocturnal peaks are restricted to the high latitudes (> 50°N).

In terms of different datasets, similar to the peak hour of mean precipitation amount, all the remote sensing estimates are consistent in reproducing the nocturnal peaks (21 – 01 LST) over the Great Plains in the United States, southern Brazil, central and northern Africa, eastern China, and parts of Australia. These region-specific features, however, are not produced by ERA5, and overall, it does not show the spatial variability in peak hour frequency. Instead, ERA5 exhibits a more uniform pattern with peak frequency hours predominantly varying between 00 to 06 LST, showing little distinction whether they occur in the polar oceans or in the tropical regions.

Unlike the peak hour of mean precipitation amount and frequency, the peak hour of intensity exhibits significant heterogeneity over both land and ocean (Figure 5.8). The difference in

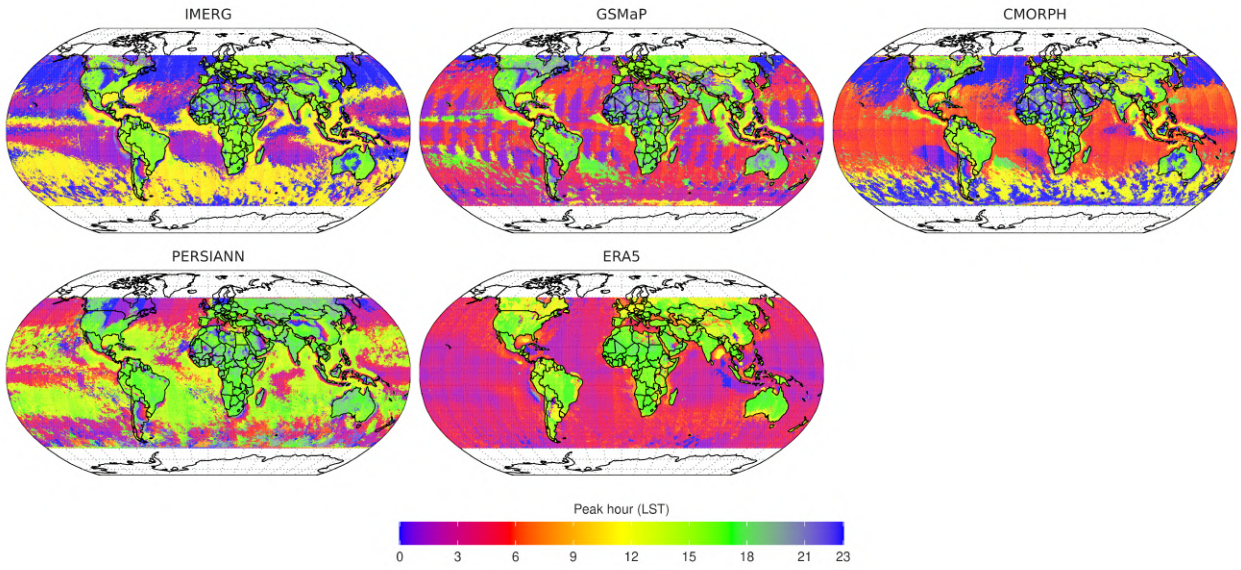


Figure 5.7: Peak hour of precipitation frequency for each dataset during 2001 – 2020.

peak hours of intensity between land and ocean is not very pronounced. Over land, the peak hour of intensity occurs either during the late night/early morning (02 – 06 LST) or late afternoon between 15 – 18 LST, depending on the region and the dataset. This observation highlights that although the mean and frequency of peak hours predominantly occur during the late afternoon over land, high-intensity precipitation mainly occurs during late nights or early mornings, though some regions also exhibit late afternoon peaks as well. Despite all the satellite estimates being consistent in depicting the regional pattern, slight differences exist among them. When compared to IMERG, CMORPH exhibits a similar pattern, whereas GSMaP exhibits a slight delay. In PERSIANN and ERA5, the peak hours are further delayed, and hence the majority of the land regions depict the peak hours during the late afternoon between 15 – 18 LST. Especially in regions such as northern South America (Brazil), southern Africa, and Canada and Russia in the northern hemisphere, ERA5 depicts afternoon peaks. In contrast, satellite estimates depict a mix of peak hours ranging from late night/early morning to a few afternoon peak hours. Unlike the peak hour of mean and frequency, ERA5 is consistent with satellite estimates, producing regional differences in peak hour intensity in regions such as the Great Plains of the USA, the southern region of South America (Peru), central regions of Africa, and parts of Australia.

Over the ocean, all satellite estimates indicate an early to late morning peak between 03 –

09 LST. In ERA5, however, the peak extends from late morning to early afternoon, covering the period from 09 – 15 LST. IMERG exhibits an early morning peak between 03 – 06 LST throughout most of the tropical oceans and the majority of the northern hemisphere. Towards the poles and in most of the southern hemisphere, it shows late evening peaks between 06 – 08 LST. Despite a similar pattern between IMERG and GSMaP, GSMaP exhibits the peak hour slightly later, occurring between 07 – 09 LST and 19 – 21 LST. CMORPH, on the other hand, shows peak hours falling between the range of IMERG and CMORPH. Regarding PERSIANN, it exhibits peak hours even earlier than the other satellite estimates, with the majority of the ocean showing peaks between 03 – 05 LST. Among the datasets, ERA5 shows the greatest deviation, with most oceanic regions exhibiting peaks in the late morning to early afternoon (09 – 15 LST). The only exception is in the Indian Ocean and western Pacific regions, where ERA5 shows early morning peaks between 06 – 09 LST, which is consistent with the other datasets. These results are not surprising, given the significant differences among the datasets, as illustrated in Figure 5.5c.

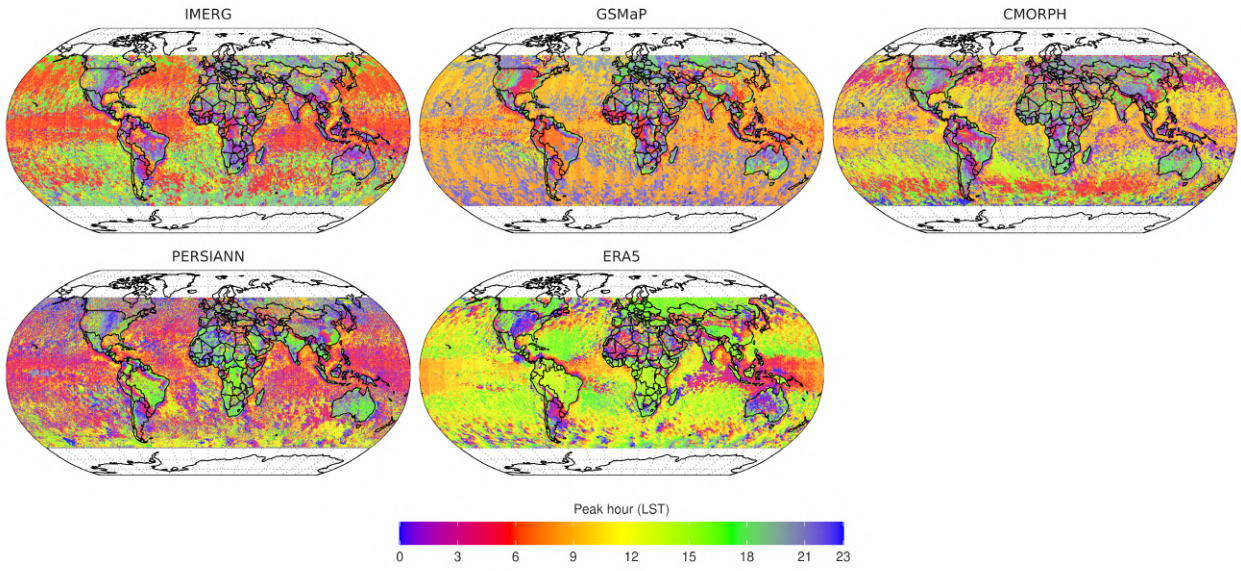


Figure 5.8: Peak hour of precipitation intensity for each dataset during 2001 – 2020.

5.3.4 Spatial distribution of diurnal characteristics

Although it is observed how diurnal precipitation varies between datasets in terms of land, ocean and globe, it does not provide information on how these shapes change at each grid level. To

accomplish this, the K-means clustering algorithm is employed. (Figure 5.9). The clusters are named according to its peak hour of local solar time: afternoon peak (red), early morning peak (grey), late morning peak (green), midnight peak (yellow), and early afternoon peak (blue).

Unsurprisingly, both IMERG and CMORPH produce similar clusters, while GSMaP and PERSIANN fall in between. However, ERA5 appears distinct. All datasets share an afternoon peak hour cluster, with a maximum between 15 – 17 LST and a minimum between 09 – 10 LST, although there are noticeable differences in magnitude. In contrast, the early morning peak, at 08 – 10 LST, is another cluster that can be seen in all the products. The early afternoon peak, observed over the GSMaP, PERSIANN, and ERA5 with peaks between 14 – 15 LST, has quite different amplitudes among the estimates. ERA5 shows the highest amplitude (> 0.20 mm/hr), whereas it remains $<$ 0.10 mm/hr in GSMaP and PERSIANN. The peak in the late morning, which closely resembles the early morning peaks, is delayed by one hour and is present in all estimates except for PERSIANN. Conversely, the peak at midnight has been observed in all estimates except for ERA5 and GSMaP, which peak between 23 – 02 LST. In addition, GSMaP and ERA5 show notable differences from the other datasets. In particular, ERA5 has only two different types of clusters: morning peak hour (brown) and afternoon peak hour (red). On the other hand, the other two clusters are very similar in shape to the red and brown clusters, the only difference being a slight delay.

The spatial distribution of the clusters reveals that afternoon peaks are more frequent over land, while early and late morning peaks occur over the ocean, which is consistent across all datasets (Figure 5.10). The midnight peak is observed in high latitudes, the oceans of the Southern Hemisphere, as well as over some land regions (i.e. the Great Plains of the USA, northern Africa, the Middle East, parts of north and eastern China, and Australia). Such regional discrepancies are captured and remain consistent in IMERG, CMORPH, and PERSIANN. In contrast, neither GSMaP nor ERA5 exhibits the midnight peaks. Instead, it is slightly delayed and these regions are occupied by the early and late morning peaks in GSMaP and ERA5. Conversely, the early and late morning peaks in ERA5 are not as pronounced as in GSMaP, especially in Africa. Over the ocean, IMERG and CMOPRH show more or less a similar pattern, with early morning peaks in most regions, late morning peaks in coastal regions, and nocturnal peaks towards high latitudes. For PERSIANN, the nocturnal peaks are more pronounced than for its counterparts IMERG and

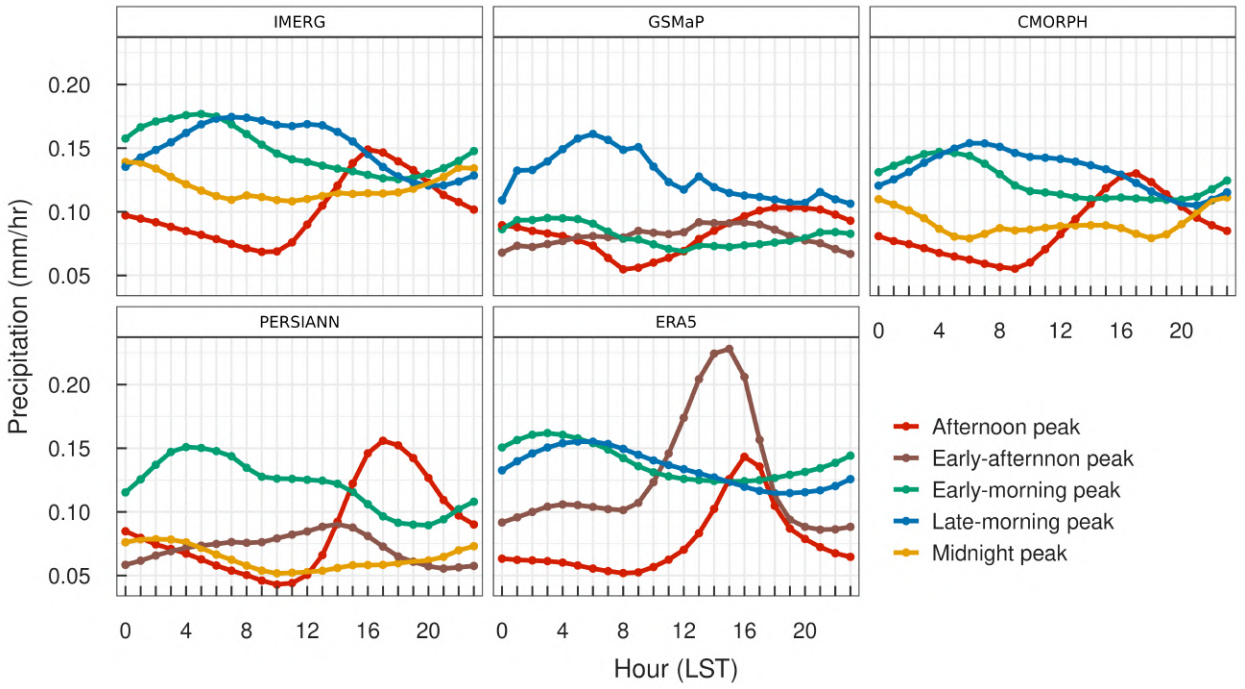


Figure 5.9: K-means clustering (k = 4) generated distinct cluster types depicting the shapes of diurnal variation of mean hourly precipitation for the entire globe (60°NS) for each dataset.

CMORPH. As far as GSMaP and ERA5 are concerned, both do not show the diversity observed in the remote sensing datasets, especially in ERA5.

5.4 Discussion

Our findings demonstrate that all the datasets agree in producing the broad spatial pattern and represent the major global precipitation features (e.g., high precipitation ITCZ, SPCZ, and low precipitation dry regions, etc.) across the globe. Nonetheless, regional discrepancies do exist among them. In particular, the regional disagreements can be observed through the latitudinal zonal precipitation. Additionally, the precipitation estimates among the products have shown some uncertainties in the dry regions such as the Sahara region, northern Africa, Asia, and the dry region of the Atlantic and Pacific Oceans. Consistent with other studies (Sun et al., 2018; Cattani et al., 2016; Dinku et al., 2011), our findings highlight higher inconsistencies among precipitation estimates for drier regions compared to humid regions. The lack of efficient ground observations over these dry and sparsely populated regions could contribute to the large uncer-

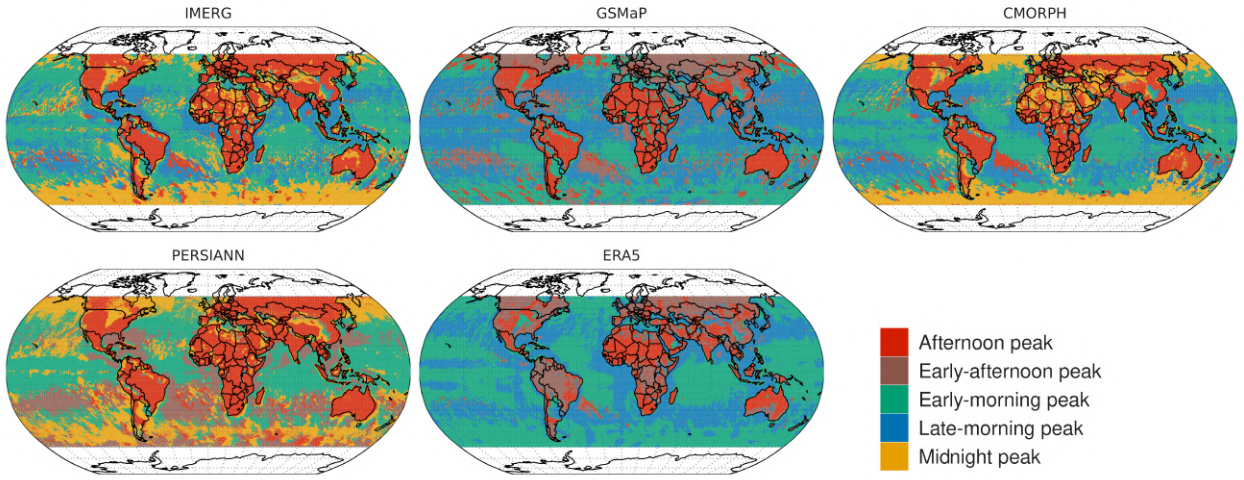


Figure 5.10: Spatial distribution of each cluster type in Figure 5.9

tainties in precipitation over these regions.

Another important aspect is the high uncertainty among the datasets over the southern hemisphere, especially within the latitudinal range of 30°S – 60°S. Even though similar concerns have previously been reported over the years, in particular over the Southern Ocean (Montoya Duque et al., 2023; Siems et al., 2022; Watters and Battaglia, 2021; Behrangi and Song, 2020), the exact reason for such behaviour is yet to be known. A fundamental challenge in this perspective is the lack of long-term, high-quality ground truth over the Southern Ocean, a reason that can be primarily attributed to the substantial difference among the estimates (Siems et al., 2022). Furthermore, as reported by Behrangi and Song (2020), the Southern Ocean exhibits the highest precipitation frequency (40%) in terms of zonal averages, and most of the precipitation occurs in the form of light precipitation. As the accurate detection of light precipitation is a persistent problem among the satellite and reanalysis datasets, this could be another probable reason for such discrepancies among them. Additionally, due to fundamental disparities in landmass distribution between the hemispheres, the Southern Ocean experiences distinct influences from atmospheric and oceanic circulation. This results in the formation of unique cloud and precipitation systems, contributing to variations in the intensity and frequency of precipitation when compared to the northern hemisphere (Siems et al., 2022). Moreover, the large inconsistencies among the products over

these regions mean there is a need for further studies exploring the physical mechanism behind such behaviours.

As previously stated, ERA5, being a reanalysis product, exhibits a slight deviation in behaviour compared to other satellite estimates, featuring high frequency and low intensity. To further confirm that ERA5's overestimation of frequency is mainly contributed by the light precipitation events, we have analyzed and repeated a similar analysis with some additional thresholds (0.2 mm/hr and 0.5 mm/hr) as well (Figure B1). The results illustrate that as the precipitation threshold increases, the overestimation of the ERA5 frequency gradually decreases and the variation between datasets decreases, confirming the challenge of ERA5 in estimating light precipitation events. Indeed, the 'drizzle problem,' characterized by excessively frequent and insufficient precipitation, is a prevalent pattern consistently observed in model precipitation estimates (Dai and Trenberth, 2004). It is mainly attributed to poor representation of convection and model characterizations (Watters et al., 2021). In particular, for ERA5, these high-frequency and low-intensity behaviours have already been reported across the regions, i.e, Eastern China (Qin et al., 2021), Tibetan Plateau (Hu and Yuan, 2021; Chen et al., 2023), eastern Himalaya (Kumar et al., 2021), Alpine Basins over Italy (Shrestha et al., 2023), Africa (Terblanche et al., 2022), southern ocean (Montoya Duque et al., 2023), etc., among others. Most of the above studies further reported that although ERA5 overestimates the frequency due to its low intensity, it is in better agreement with the total amount of precipitation. In fact, Montoya Duque et al. (2023) also found that despite the high-frequency and low-intensity issues, ERA5 has a better estimate of total precipitation than IMERG over the southern oceanic region. As the ERA5 model forecasts precipitation, it overestimates light precipitation, and to compensate for the amount, it underestimates the precipitation intensities (Qin et al., 2021).

Similarly, the early peak hours of the diurnal cycle by ERA5 are also consistently reported by previous studies (Hayden et al., 2023; Chen et al., 2023; Watters et al., 2021). The early peak bias in ERA5 could be mainly attributed to the ineffective representation of the convective parameterization schemes (Chen et al., 2023). In particular, the premature convection in the tropics could be the probable reason (Watters et al., 2021). Moreover, the challenge of estimating precipitation over the ocean is also a persistent problem in ERA5 (Montoya Duque et al., 2023).

Furthermore, similar to the previous studies, our results also report some unique diurnal regional features across the globe. For example, one such instance is the eastward shift of the diurnal peak over the central United States (Tan et al., 2019b). All the estimates depict the nocturnal peaks across the Great Plains, a unique feature of the US diurnal cycle. Nevertheless, unlike the IMERG, GSMaP, CMORPH and, to some extent, PERSIANN, the changes in peak hours by ERA5 exhibit minimal spatial variation (Figure 5.6). Another example can be seen in the Amazon region, where ERA5 does not exhibit spatial variation in the peak hour of precipitation amount, but instead shows a widespread noon peak hour between 11 – 12 LST (Figure 5.6). In contrast, the remaining satellite estimates show substantial spatial variation in the peak hour across the Amazon basin, which is especially more pronounced in GSMaP. These results are consistent with those observed by Hayden et al. (2023); Tai et al. (2021) for ERA5 over the Amazon region. They further attributed that in ERA5, the precipitation diurnal variability was mainly driven by solar radiation rather than the small-scale regional processes.

Moreover, it should be noted that the findings of this analysis do not indicate that IMERG is the best representation of the diurnal variation of precipitation. Although studies have shown that IMERG can be considered as a reference for the evaluation of other diurnal precipitation estimates globally (O and Kirstetter, 2018; Watters et al., 2021), it is not error-free and still has regional issues (O and Kirstetter, 2018). Nonetheless, the results of O and Kirstetter (2018) are based on IMERG V04, whereas the latest IMERG version V06 has shown significant improvements over the earlier ones, i.,e IMERG V06 has a very short lag (average +0.59) over the eastern United States (Tan et al., 2019b).

5.5 Conclusions

This chapter compared four state-of-the-art satellite and one reanalysis precipitation data products at the sub-daily scale. In particular, it compared the diurnal variation of precipitation estimates in terms of mean precipitation amount, frequency, and intensity at the global level. The main findings of the analysis are as follows:

- Overall, all estimates provide a good representation of the spatial pattern of mean hourly

precipitation characteristics across the globe. They are more consistent in producing spatial patterns regarding mean hourly precipitation amount and frequency than intensity.

- The discrepancies among the datasets are most pronounced at high latitudes (30°N/S to 60°N/S) compared to the tropical regions (0°N/S to 30°N/S). The agreement among the estimates is higher in the Northern Hemisphere than in the Southern Hemisphere. In particular, the discrepancies are larger at the latitudes between 35°S to 60°S, which could be mainly due to the prevalence of oceans in the Southern Hemisphere.

- ERA5 significantly overestimates the precipitation frequency, and is characterised by very low intensity compared to the rest of the precipitation products. GSMaP also depicts very low precipitation frequency, which is more pronounced over land than the ocean.

- All datasets effectively capture the major diurnal features: an afternoon peak over land and an early morning peak over the ocean. In terms of inter-product comparison, ERA5 detects the peak slightly earlier, around 15 LST over land, compared to the other datasets, which peak at 16 LST. Moreover, ERA5 tends to overestimate the amount of precipitation compared to IMERG and PERSIANN estimates, while CMORPH and GSMaP consistently show lower values.

- In terms of diurnal frequency, the ERA5 precipitation frequency is significantly higher than the rest of the estimates, regardless of whether it is over land, over the ocean, or at the global level. However, compared to the land, the high frequencies seem way more dominant over the ocean.

- Different from precipitation mean and frequency, precipitation intensity exhibits a weaker diurnal cycle. ERA5 displays the lowest precipitation intensity among the estimates, while CMORPH exhibits the highest. Surprisingly, GSMaP also shows the lowest intensity among the datasets, even lower than ERA5 over land.

- All the estimates have smaller discrepancies in the peak hour of mean precipitation amount than in the peak hour of frequency and intensity. The highest discrepancies among the datasets of peak hours are observed mainly over the Southern Ocean.

- The K-means clustering results also depict that all the estimates are consistent in reproducing the early-morning peak over the ocean and the afternoon peak over land. Moreover, the IMERG and CMORPH estimates exhibit a high degree of agreement in terms of diurnal shapes, producing similar patterns. However, the remaining products show variations in their diurnal shapes.

Chapter 6

Conclusions

6.1 Summary of thesis

The thesis aims to contribute towards a better understanding of the GPM IMERG precipitation in a range of geographic and climatic conditions at the global level. Accordingly, the thesis performs a quantitative review of IMERG precipitation evaluation studies between 2016 – 2019 (Chapter 3). This analysis identifies the geographical distribution of IMERG precipitation evaluation/validation studies across the globe, the spatial-temporal scale, the method of analysis (point-pixel or pixel-pixels), whether it is over land or ocean, rain or snow, etc. Additionally, it identifies IMERG's performance over extreme events and its hydrological application. Furthermore, based on the review datasets, this analysis also identifies the major research gaps that remain, and recommends future studies that could grasp opportunities to provide further insights into IMERG evaluation/validation.

Building upon the research gaps identified in Chapter 3, the thesis undertakes the challenging task of evaluating IMERG precipitation datasets over the global ocean (Chapter 4). Given the limited knowledge of IMERG precipitation over the global ocean, all the IMERG products including Early, Late, and Final runs are compared against the buoy datasets across the tropical oceans. It further decomposes the IMERG total errors to identify the main source of error. Moreover, it also evaluates the IMERG products in terms of different precipitation intensities to provide further insights into how IMERG performs under different precipitation intensities across the tropical oceans.

Another important research question identified in Chapter 3 is the evaluation of IMERG sub-daily performance at the global level. This task is undertaken, and a comprehensive assessment is performed along with a range of satellite (IMERG, GSMaP, CMORPH, PERSIANN) and reanalysis (ERA5) products (Chapter 5). All the datasets considered are the major global precipitation available at the time of the analysis, with fine spatial-temporal resolutions (hourly and $0.25°$) covering both land and ocean. The intercomparison is performed in terms of sub-daily precipitation and their diurnal variation at the global level. It considered three main properties; the mean precipitation amount, the frequency, and the intensity. Considering the role of reanalysis and satellite precipitation estimates, this kind of analysis provides great insights to future algorithm

developers and end users.

6.2 Key findings

The key research questions (RQ's), as well as the dissertation's major findings, are revisited below.

RQ-1. What is the current status of IMERG precipitation estimates across the globe? What are the main strengths and weaknesses of IMERG, and how does it change with season, region, etc?

- IMERG outperforms its predecessor, TMPA, and better depicts the spatial and temporal pattern of mean precipitation across the geographic and climatic conditions throughout the globe.
- Its performance is influenced by the region, season, and intensity. IMERG has better performance in humid and high-intensity precipitation regions than in dry regions with low-intensity precipitation. Its performance in the summer is comparatively better than in the winter seasons. The challenge of IMERG in the detection of light precipitation and solid precipitation in winter, indicating potential for further improvement in this aspect.
- IMERG still has issues in accurate estimation of precipitation over the complex terrains and mountain regions.
- It has better captured the extreme precipitation events and their track; however, it underestimates the high precipitation intensities, especially over the core of the typhoon and hurricanes.
- Regarding IMERG's hydrological application, it has demonstrated potential for various hydrological applications. However, its performance highly varies with the basin type, the hydrological model used, validation data, calibration, etc.

RQ-2. How effective is IMERG in the representation of tropical ocean precipitation, and how does its bias change with spatial regions and precipitation intensities?

- Across the tropical ocean, IMERG adeptly captures the average spatial patterns, distinguishing between regions of intense precipitation within the ITCZ belts and areas of lower precipitation in the subtropical oceans. However, the over/under estimation of precipitation observed across the tropical oceans with magnitude varies with the regions.
- IMERG overestimates precipitation in the high-precipitation regions such as the Indian and West Pacific oceans, mainly due to overestimation of highly intense precipitation.
- Its detection capability is higher in the high-precipitation regions compared to the low-precipitation regions (Atlantic and east Pacific). The dominance of light precipitation in low-precipitation regions is the main reason for such low detection capability and high-false alarms.
- The positive hit bias and false alarm bias are the dominant sources of IMERG's total error which leads to an overall overestimation of precipitation amount, especially over the Indian and Pacific oceans.
- No noticeable differences are found among the products in terms of different IMERG runs (i.e., IMERG-E, -L, and -F). This lack of discrepancy is expected due to the absence of gauge correction over the ocean. The only distinctions between the products are the sampling size and the integration of microwave precipitation estimates.

RQ-3. How do different precipitation estimates capture the diurnal variability of global precipitation, and how is this variability influenced by distinct geographical features, such as land, ocean, etc.?

- In terms of sub-hourly precipitation estimation, IMERG, GSMaP, CMORPH, PERSIANN, and ERA5 well represent the global precipitation features; high precipitation across the ITCZ and SPCZ, and low precipitation in the sub-tropical regions of northern Africa, the Atlantic, and Pacific Oceans.
- Regarding precipitation frequency and intensity, ERA5 has the highest precipitation occurrence and lowest intensity compared to the rest of the estimates. This is more obvious over the ocean than over land.

- All products agree on reproducing the afternoon peak precipitation hour over land and the early morning peak over the ocean. Nevertheless, there are regional variations among the estimates. Moreover, discrepancies in the amplitude persist among them, resulting in over- or underestimation.

- In terms of diurnal variation, ERA5 detects the peak hour for mean precipitation amount at 15 LST, which is earlier than other estimates that observe the peak at 16 LST over land. Furthermore, ERA5 fails to capture small-scale regional variations across the globe.

- The discrepancies among the datasets are most pronounced at high latitudes (30°N/S to 60°N/S) compared to the tropical regions (0°N/S to 30°N/S). The agreement among the estimates is higher in the Northern Hemisphere than in the Southern Hemisphere. In particular, the discrepancies are larger at the latitudes between 35°S to 60°S, which could be mainly due to the prevalence of oceans in the Southern Hemisphere.

6.3 Limitations and future research recommendations

Regarding RQ1, it is important to stress that all considered studies suffer from a common limitation: the very short validation period due to the lack of long-term IMERG records when the analyses were performed. Even though the evaluation methods used in most studies are reliable, there are often issues with their underlying assumptions. For instance, there are different approaches to comparing gauge measurements regarding their number, distribution, and density over the validation sites. Additionally, there is significant uncertainty associated with the point-to-area representation for gauge measurements (Dezfuli et al., 2017). The type of interpolation techniques applied to the point measurements might severely affect the evaluation outcome. It is worth mentioning that all the studies reviewed herein have employed a pixel-based approach which may result in problems like "double-penalty error" when comparing two data sets with high spatiotemporal resolution. This is one of the reasons limited studies are evaluating IMERG on sub-daily scales. Therefore, an object-based approach could be a solution to mitigate this effect. In addition, this approach could provide us with more information about the storm's characteristics like size, shape, translation speed, and direction, which cannot be accessed using a

pixel-based approach. Some efforts have been made recently like Ayat et al. (2021b), and Cui et al. (2020), but more studies are needed on this topic.

Another limitation associated with the studies examined appears when they use radar observations, reanalysis results, or other satellite products in the absence of rain gauges. In such cases, the reference used may not represent the actual precipitation of the region (assumption of ground truth), and is often associated with significant uncertainties. Even if the studies use gauge-based data as a reference for IMERG validation, gauges also present systematic and random errors. Therefore, we should keep in mind that it is hard to assess the "ground truth" in nature, and thus, we always make some assumptions. In addition, most often, the gauge-based data are not completely independent. For instance, IMERG-F products are adjusted with the GPCC monthly product, which itself uses gauge observations. Thus, the validation results would be biased if there are overlapping stations when compared with IMERG products. Furthermore, most studies seem to be restricted to local/regional climate; thus, the applicability of the results is limited to those particular climate conditions. Finally, most studies did not assess the performance at sub-daily resolutions but instead focused on daily, monthly, and/or annual resolutions.

Considering the identified limitations, gaps, and suggestions in the studies reviewed in this chapter, we can provide some recommendations that could potentially improve the performance of IMERG products across the globe: Since IMERG recently extended its temporal coverage back to the TRMM era, providing precipitation observations from early 2000-present, it now offers possibilities for longer evaluation of precipitation characteristics across the globe. For instance, the long-term data will help better understand the capacity of IMERG to represent changes in the annual, inter-annual, and seasonal precipitation at the regional scale. To fully exploit the advantages of IMERG products' spatiotemporal resolution (30 min and 0.1° x 0.1°) compared to TMPA (3 hourly and 0.25° x 0.25°), future studies should consider sub-daily scales and assess the diurnal and semi-diurnal precipitation in different regions. More studies over oceans and mountainous regions could help better understand the effectiveness of IMERG in such conditions.

Additionally, as very few studies have evaluated solid precipitation, future studies should focus on the detailed evaluation of snowfall. Since many studies revealed the poor performance of IMERG during the winter, these biases could be related. A more detailed evaluation of the hydro-

logical performance of IMERG will be another area of future research. For instance, understanding how the error propagation occurs from precipitation to runoff (Mei et al., 2017; Ehsan Bhuiyan et al., 2019) and their quantitative analysis will help the hydrological community better understand the performance of IMERG products in hydrological applications. Evaluating IMERG products at multiple scales simultaneously rather than constraining the analysis to a single spatial and temporal resolution could help us understand how the accuracy and errors vary with spatiotemporal aggregation. Additionally, it will help identify the effective resolution to be used for various hydro-meteorological purposes. Even though few studies have evaluated IMERG versions V03, V04, and V05, the results varied with the type of study. Thus, future studies could comprehensively evaluate and answer questions on the effectiveness of the current versions compared to previous ones.

Along with orographic precipitation and coastal areas evaluation, future studies could also consider evaluating the effect of rainfall intensity and gauge density (Maggioni et al., 2017) on the performance of IMERG products across the climatic regimes and geographical conditions across the globe. Even though the gauge-calibrated IMERG final run has advantages over the uncalibrated one, there is still room for improvement. For example, the scarce distribution of gauges over some regions could be solved by exploring the adjustment with other denser observational networks such as E-OBS (Ensembles-OBServation) over Europe in the future (Navarro et al., 2019). It will be beneficial to have more studies over Europe since many dense, well-maintained, observation networks exist. As the observational data sets are not free from uncertainty, many authors recommended considering more than one reference type to evaluate IMERG products. For example, the newly available blended data sets such as MSWEP (overland), OceanRain (over Ocean), other radar data sets, and reanalysis model data sets will help for better evaluation.

Uncertainties from different interpolation methods used when evaluating IMERG (grid) with gauge (point) date are not well represented. Thus, comparing different interpolation techniques and their effect on the IMERG evaluation could provide a more detailed error estimation. An additional consideration for future research is the evaluation of IMERG for different types of storms/environmental conditions (e.g., temperature profiles) and microphysical structures (derived from polarimetric radar and NWP) (Bartsotas et al., 2018). Another important topic for

future research is the effect of different sensors in the final merged products. Although a couple of studies like Tan et al. (2016), and Ayat et al. (2021a) have evaluated this effect in the IMERG final precipitation product, the need for further investigation remains.

Regarding IMERG performance evaluation over tropical oceans (Chapter 4), we could mention some additional limitations of our analysis, which pave the road to future research. To begin with, the buoys data used in the analysis were not applied any wind correction. Since the issue of undercatch in the buoy's measurement is well documented in past studies (Wu and Wang, 2019; Serra and McPhaden, 2003), it is assumed that the buoys would have slightly underestimated the actual precipitation. Even though no specific wind correction formulas are available for the R.M Young capacitance gauge mounted over open oceanic buoys, few studies have estimated wind correction for other gauges over land (Koschmieder, 1934; Yang et al., 1998). Although specific information on wind underestimation for the R.M Young capacitance gauge over open oceanic buoys is limited, these studies provide insights into the general effects of wind on gauge measurements. Further research and expert input may be necessary to address this specific issue.

Therefore, applying a wind correction factor to the buoys will reduce noise, which will further improve the IMERG evaluation by reducing the uncertainties associated with the validation results. In addition, it will help to identify the main source of error in the IMERG precipitation and their spatial distribution throughout precipitation regimes. Moreover, an appropriate wind correction would provide greater insight into IMERG error characterisation, especially for errors associated with light precipitation events. By considering other sources of independent reference precipitation, such as OceanRAIN (Klepp et al., 2018), Radar (e.g., Kwajalein Polarimetric S-band Weather Radar), and acoustic rain-gauges could be a means of further confirming the error characteristics of IMERG. This will provide a more accurate representation of the oceanic precipitation and will help to address the uncertainties and limitations associated with any single dataset.

In addition, another important limitation could stem from the point-pixel-based comparison approach (Tian et al., 2018). These discrepancies in the point-pixel approach could cause a higher false alarm ratio and over/underestimation of the light and extreme precipitation events. Furthermore, due to the lack of continuous time series in the buoys measurements, this analysis is limited

to a daily scale and has not explored the sub-daily scale analysis. Therefore, considering other sources of reference precipitation, such as OceanRAIN (Klepp et al., 2018) could be an addition to confirm the IMERG characteristics further. Furthermore, this would aid to evaluate the IMERG on a sub-daily scale and evaluating IMERG performance based on precipitation types, i.e., stratiform and convective. Another future direction could be error decomposition and investigating the main source of errors. Moreover, considering how the buoys location and high precipitation ITCZ affect IMERG performance, a seasonal evaluation of IMERG will provide further insights into this.

Regarding the sub-daily evaluation and diurnal cycle estimation of various satellite and reanalysis datasets (Chapter 5), given that the analysis is carried out at the $0.25^\circ \times 0.25^\circ$ resolution, some uncertainties could be associated with the re-gridding, especially for IMERG and GSMaP, which are available at the original resolution of $0.1^\circ \times 0.1^\circ$. Therefore, future studies could consider the IMERG and GSMaP products and evaluate their diurnal variation at their original resolutions. Future studies can also consider the duration of precipitation to gain more comprehensive insights into how the different precipitation durations have distinct diurnal variations and the mechanism behind each precipitation structure. In addition, the diurnal variation of different precipitation intensities will provide further insights.

One potential direction for future investigation involves examining the IMERG V07, which was not available during the analysis carried out, and identifying the effects of major changes brought about by the recent version compared to its predecessor (i.e., V06). In addition, considering the application-oriented importance of near-real-time satellite datasets, such as IMERG-Early, Late runs, and GSMaP near-real-time version (GSMaP-NRT), assessing their capability and identifying uncertainties in their representation of the diurnal cycle can offer additional perspectives, particularly for regions where these datasets could be potentially applied. Furthermore, a notable limitation of our current analysis is the strict filtering criteria applied, leading to the exclusion of various precipitation datasets available at very high spatial and temporal resolutions but limited to land (e.g., ERA5land, CHRIPS, etc.). Therefore, forthcoming research endeavours could incorporate these high-resolution products to provide more detailed insights into the uncertainties associated with estimating diurnal precipitation.

Another major limitation of the current analysis is that it did not consider any model data, despite their high demand and utility in long-term climate research, as well as their pivotal role in the current array of precipitation datasets. The utilisation of K-means clustering across the entire global domain may have overlooked certain small-scale and unique diurnal variations associated with distinct precipitation regimes. Future investigations could explore applying clustering techniques to individual continents or specific precipitation regimes, thereby capturing regional features more effectively. Additionally, alternative clustering or machine learning methods may provide additional details beyond those obtained from K-means clustering alone. Moreover, given the availability of nearly two decades of satellite precipitation datasets at a global level, assessing the trend and inter-annual variability of the diurnal cycle can offer further understanding into the changes in the diurnal cycle over time.

6.4 Epilogue

The thesis provides a comprehensive global examination of GPM IMERG precipitation products. Its outcomes enrich our scientific knowledge regarding the uncertainties of IMERG precipitation products and contribute to a deeper understanding of various perspectives of IMERG error characterisation. It offers a detailed yet global-level assessment of IMERG, facilitating algorithm developers in pinpointing critical limitations and areas necessitating enhancement. In addition, it will aid the end-users in selecting the appropriate precipitation dataset, consequently to overall better decision-making. The thesis outcomes will also assist the scientific community, including hydrologists and meteorologists, especially those engaged in global water and energy budget studies, in accurately discerning and attributing the contribution of satellite precipitation-related uncertainties to their final outcomes and claims. Moreover, the thesis also highlights the significant research gap, particularly concerning IMERG precipitation, thereby suggesting directions for future research. In doing so, it aids the scientific community by identifying areas where further investigation is needed to advance our understanding of IMERG data and its applications.

Accurate precipitation measurement and understanding its uncertainties play a pivotal role in weather and agricultural forecasting. The utilisation of IMERG precipitation products has re-

cently expanded into various domains, spanning from hurricane/tropical cyclones to predicting flash floods, landslides, and agricultural water needs. Our findings will enhance precise and accurate forecasting, aiding decision-making in near-real-time disaster management, such as flash flood, landslide prediction, etc. This, in turn, will enhance preparedness and mitigate societal and human losses. Moreover, the uncertainties, limitations, and strengths of IMERG, which is the most widely used remote sensing data product, will profoundly influence outcomes and consequently affect the overall well-being of human society.

Appendix A

Supplementary Data for Chapter-4

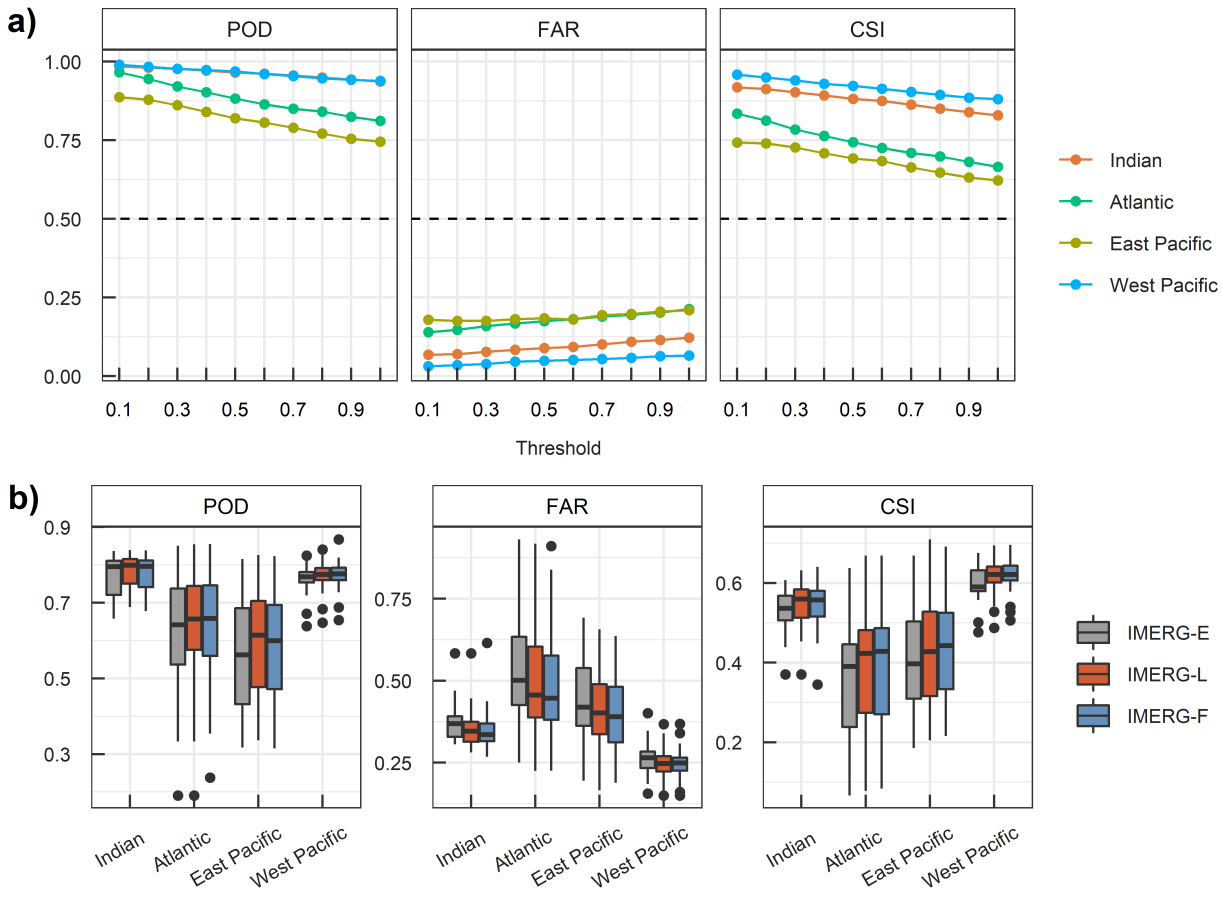


Figure A1: Categorical metrics for daily precipitation over the tropical oceans; a) as a function of different detection thresholds (IMERG-F), b) As Figure 4.2 but for a rain and no-rain threshold of 1 mm/day.

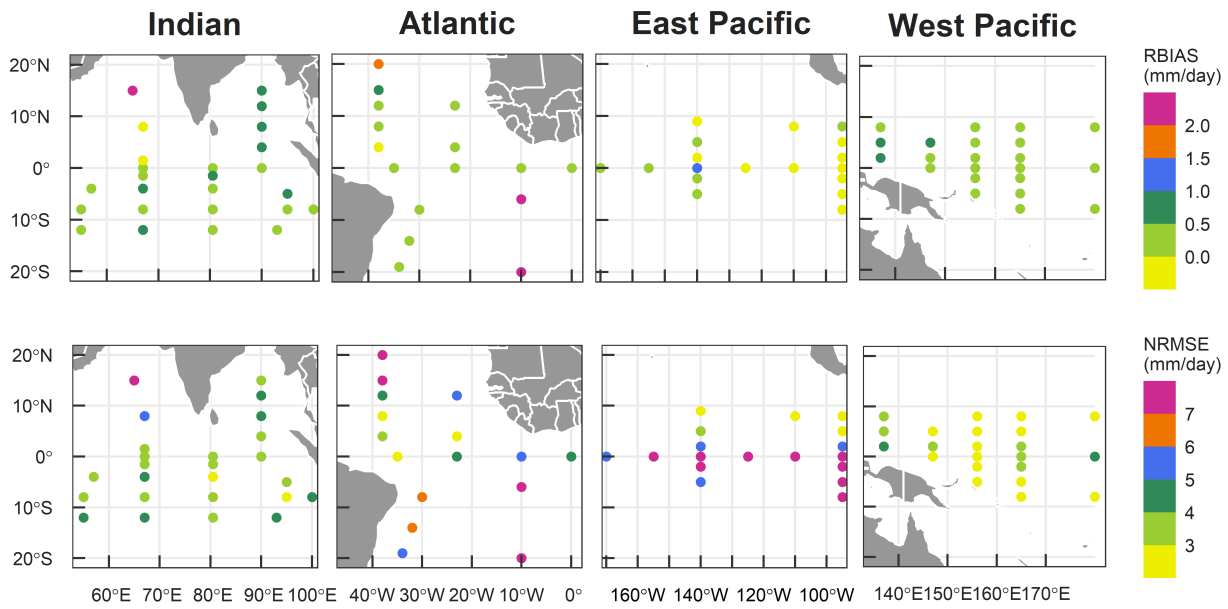


Figure A2: As Figure 4.4 but the values are normalised by the buoy's mean precipitation.

Table A1: Categorical metrics and their scores for different rainfall intensity percentiles.

Ocean	Threshold (Percentile)	POD	FAR	CSI	H	M	F	CN
Indian	1	0.99	0.07	0.92	5099	72	371	267
	25	0.91	0.17	0.77	3558	357	717	1177
	50	0.84	0.30	0.62	2195	415	928	2271
	99	0.45	0.79	0.17	24	29	91	5665
Atlantic	1	0.99	0.13	0.86	6317	62	974	166
	25	0.86	0.19	0.72	4164	667	947	1741
	50	0.74	0.26	0.58	2376	846	854	3443
	99	0.37	0.82	0.14	24	41	109	7345
East Pacific	1	0.92	0.21	0.74	4866	409	1317	853
	25	0.86	0.18	0.72	3424	572	741	2708
	50	0.72	0.22	0.60	1918	746	538	4243
	99	0.37	0.75	0.18	20	34	60	7331
West Pacific	1	1.00	0.03	0.97	7240	26	205	48
	25	0.88	0.12	0.79	4864	641	654	1360
	50	0.83	0.25	0.65	3054	616	1007	2842
	99	0.55	0.88	0.11	41	33	287	7158

Appendix B

Supplementary Data for Chapter-5

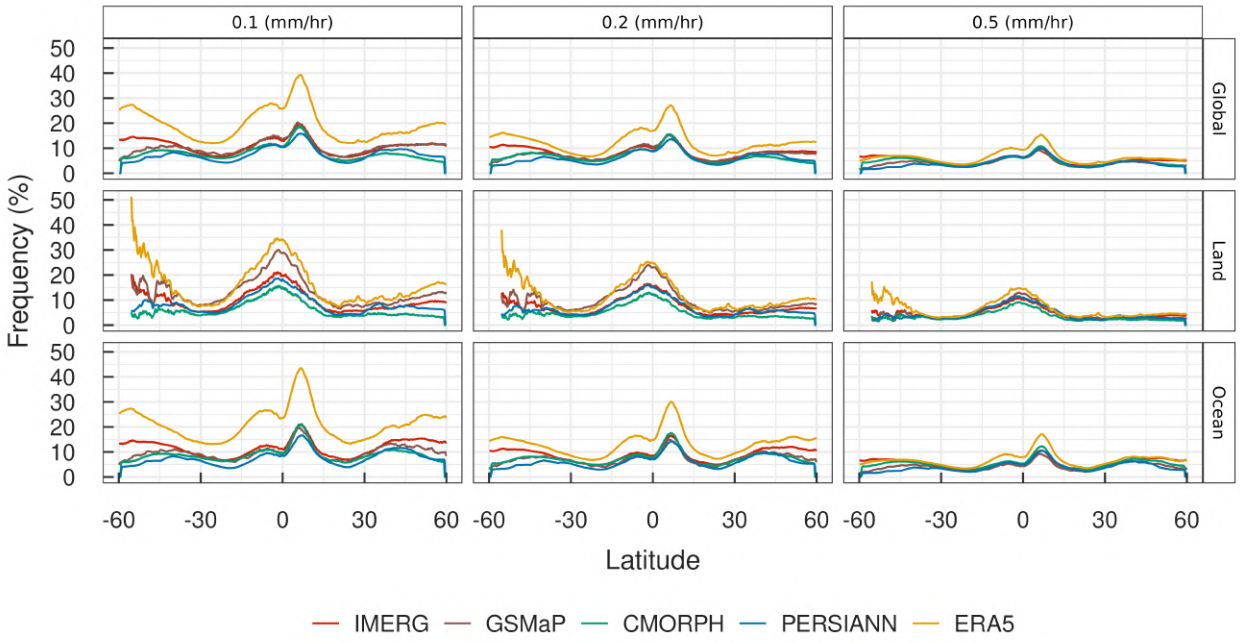


Figure B1: Latitudinal average of precipitation frequency among the estimates at different thresholds.

Bibliography

Abebe, S. A., Qin, T., Yan, D., Gelaw, E. B., Workneh, H. T., Kun, W., Shanshan, L., and Biqiong, D. (2020). Spatial and temporal evaluation of the latest high-resolution precipitation products over the upper blue nile river basin, ethiopia. *Water*, 12(11):3072.

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., et al. (2003). The version-2 global precipitation climatology project (gpcp) monthly precipitation analysis (1979–present). *Journal of hydrometeorology*, 4(6):1147–1167.

Adler, R. F., Sapiano, M. R., Huffman, G. J., Wang, J.-J., Gu, G., Bolvin, D., Chiu, L., Schneider, U., Becker, A., Nelkin, E., et al. (2018). The global precipitation climatology project (gpcp) monthly analysis (new version 2.3) and a review of 2017 global precipitation. *Atmosphere*, 9(4):138.

Afonso, J. M. d. S., Vila, D. A., Gan, M. A., Quispe, D. P., Barreto, N. d. J. d. C., Huamán Chinchay, J. H., and Palharini, R. S. A. (2020). Precipitation diurnal cycle assessment of satellite-based estimates over brazil. *Remote Sensing*, 12(14):2339.

Ahmed, E., Al Janabi, F., Zhang, J., Yang, W., Saddique, N., and Krebs, P. (2020). Hydrologic assessment of trmm and gpm-based precipitation products in transboundary river catchment (chenab river, pakistan). *Water*, 12(7):1902.

Alsumaiti, T. S., Hussein, K., Ghebreyesus, D. T., and Sharif, H. O. (2020). Performance of the cmorph and gpm imerg products over the united arab emirates. *Remote Sensing*, 12(9):1426.

Anjum, M. N., Ahmad, I., Ding, Y., Shangguan, D., Zaman, M., Ijaz, M. W., Sarwar, K., Han, H., and Yang, M. (2019). Assessment of imerg-v06 precipitation product over different hydro-climatic regimes in the tianshan mountains, north-western china. *Remote Sensing*, 11(19):2314.

Anjum, M. N., Ding, Y., Shangguan, D., Ahmad, I., Ijaz, M. W., Farid, H. U., Yagoub, Y. E., Zaman, M., and Adnan, M. (2018). Performance evaluation of latest integrated multi-satellite retrievals for global precipitation measurement (imerg) over the northern highlands of pakistan. *Atmospheric Research*, 205:134–146.

Asong, Z., Razavi, S., Wheater, H., and Wong, J. (2017). Evaluation of integrated multisatellite retrievals for gpm (imerg) over southern canada against ground precipitation observations: A preliminary assessment. *Journal of Hydrometeorology*, 18(4):1033–1050.

Ayat, H., Evans, J. P., and Behrangi, A. (2021a). How do different sensors impact imerg precipitation estimates during hurricane days? *Remote Sensing of Environment*, 259:112417.

Ayat, H., Evans, J. P., Sherwood, S., and Behrangi, A. (2021b). Are storm characteristics the same when viewed using merged surface radars or a merged satellite product? *Journal of Hydrometeorology*, 22(1):43–62.

Bai, H. and Schumacher, C. (2022). Topographic influences on diurnally driven mjo rainfall over the maritime continent. *Journal of Geophysical Research: Atmospheres*, 127(6):e2021JD035905.

Bartsotas, N., Anagnostou, E., Nikolopoulos, E., and Kallos, G. (2018). Investigating satellite precipitation uncertainty over complex terrain. *Journal of Geophysical Research: Atmospheres*, 123(10):5346–5359.

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., Van Dijk, A. I., Huffman, G. J., Adler, R. F., and Wood, E. F. (2019). Daily evaluation of 26 precipitation datasets using stage-iv gauge-radar data for the conus. *Hydrology and Earth System Sciences*, 23(1):207–224.

Beck, H. E., Van Dijk, A. I., Levizzani, V., Schellekens, J., Gonzalez Miralles, D., Martens, B., and De Roo, A. (2017). Mswep: 3-hourly 0.25 global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21(1):589–615.

Behrangi, A., Lebsock, M., Wong, S., and Lambrigtsen, B. (2012). On the quantification of oceanic rainfall using spaceborne sensors. *Journal of Geophysical Research: Atmospheres*, 117(D20).

Behrangi, A. and Song, Y. (2020). A new estimate for oceanic precipitation amount and distribution using complementary precipitation observations from space and comparison with gpcp. *Environmental Research Letters*, 15(12):124042.

Behrangi, A. and Wen, Y. (2017). On the spatial and temporal sampling errors of remotely sensed precipitation products. *Remote Sensing*, 9(11):1127.

Beria, H., Nanda, T., Bisht, D. S., and Chatterjee, C. (2017). Does the gpm mission improve the systematic error component in satellite rainfall estimates over trmm? an evaluation at a pan-india scale. *Hydrology and Earth System Sciences*, 21(12):6117–6134.

Bolvin, D. T., Huffman, G. J., Nelkin, E. J., and Tan, J. (2021). Comparison of monthly imerg precipitation estimates with pacrain atoll observations. *Journal of Hydrometeorology*, 22(7):1745–1753.

Bourlès, B., Lumpkin, R., McPhaden, M. J., Hernandez, F., Nobre, P., Campos, E., Yu, L., Planton, S., Busalacchi, A., Moura, A. D., et al. (2008). The pirata program: history, accomplishments, and future directions. *Bulletin of the American Meteorological Society*, 89(8):1111–1126.

Bowman, K. P. (2005). Comparison of trmm precipitation retrievals with rain gauge data from ocean buoys. *Journal of Climate*, 18(1):178–190.

Bowman, K. P., Collier, J. C., North, G. R., Wu, Q., Ha, E., and Hardin, J. (2005). Diurnal cycle of tropical precipitation in tropical rainfall measuring mission (trmm) satellite and ocean buoy rain gauge data. *Journal of Geophysical Research: Atmospheres*, 110(D21).

Cattani, E., Merino, A., and Levizzani, V. (2016). Evaluation of monthly satellite-derived precipitation products over east africa. *Journal of Hydrometeorology*, 17(10):2555–2573.

Chen, C., Chen, Q., Duan, Z., Zhang, J., Mo, K., Li, Z., and Tang, G. (2018). Multiscale comparative evaluation of the gpm imerg v5 and trmm 3b42 v7 precipitation products from 2015 to 2017 over a climate transition area of china. *Remote Sensing*, 10(6):944.

Chen, F. and Li, X. (2016). Evaluation of imerg and trmm 3b43 monthly precipitation products over mainland china. *Remote Sensing*, 8(6):472.

Chen, G., Iwasaki, T., Qin, H., and Sha, W. (2014). Evaluation of the warm-season diurnal variability over east asia in recent reanalyses jra-55, era-interim, ncep cfsr, and nasa merra. *Journal of climate*, 27(14):5517–5537.

Chen, J., Wang, Z., Wu, X., Chen, X., Lai, C., Zeng, Z., and Li, J. (2019). Accuracy evaluation of gpm multi-satellite precipitation products in the hydrological application over alpine and gorge regions with sparse rain gauge network. *Hydrology Research*, 50(6):1710–1729.

Chen, M., Nabih, S., Brauer, N. S., Gao, S., Gourley, J. J., Hong, Z., Kolar, R. L., and Hong, Y. (2020). Can remote sensing technologies capture the extreme precipitation event and its cascading hydrological response? a case study of hurricane harvey using ef5 modeling framework. *Remote Sensing*, 12(3):445.

Chen, T., Li, J., Zhang, Y., Chen, H., Li, P., and Che, H. (2023). Evaluation of hourly precipitation characteristics from a global reanalysis and variable-resolution global model over the tibetan plateau by using a satellite-gauge merged rainfall product. *Remote Sensing*, 15(4):1013.

Chiaravalloti, F., Brocca, L., Procopio, A., Massari, C., and Gabriele, S. (2018). Assessment of gpm and sm2rain-ascat rainfall products over complex terrain in southern italy. *Atmospheric Research*, 206:64–74.

Cui, W., Dong, X., Xi, B., Feng, Z., and Fan, J. (2020). Can the gpm imerg final product accurately represent mcss' precipitation characteristics over the central and eastern united states? *Journal of Hydrometeorology*, 21(1):39–57.

Dai, A., Lin, X., and Hsu, K.-L. (2007). The frequency, intensity, and diurnal cycle of precipitation in surface and satellite observations over low-and mid-latitudes. *Climate dynamics*, 29:727–744.

Dai, A. and Trenberth, K. E. (2004). The diurnal cycle and its depiction in the community climate system model. *Journal of Climate*, 17(5):930–951.

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C.-H., Chen, H., Delrieu, G., Hsu, Y. C., et al. (2019). Evaluation of gpm-era global satellite precipitation products over multiple complex terrain regions. *Remote Sensing*, 11(24):2936.

Derin, Y., Kirstetter, P.-E., Brauer, N., Gourley, J. J., and Wang, J. (2022). Evaluation of imerg satellite precipitation over the land-coast-ocean continuum–part ii: Quantification. *Journal of Hydrometeorology*.

Derin, Y. and Yilmaz, K. K. (2014). Evaluation of multiple satellite-based precipitation products over complex topography. *Journal of Hydrometeorology*, 15(4):1498–1516.

Dezfuli, A. K., Ichoku, C. M., Huffman, G. J., Mohr, K. I., Selker, J. S., van de Giesen, N., Hochreutener, R., and Annor, F. O. (2017). Validation of IMERG Precipitation in Africa. *Journal of Hydrometeorology*, 18(10):2817–2825.

Dinku, T., Ceccato, P., and Connor, S. J. (2011). Challenges of satellite rainfall estimation over mountainous and arid parts of east africa. *International journal of remote sensing*, 32(21):5965–5979.

Ehsan Bhuiyan, M. A., Nikolopoulos, E. I., Anagnostou, E. N., Polcher, J., Albergel, C., Dutra, E., Fink, G., Martínez-De La Torre, A., and Munier, S. (2019). Assessment of precipitation error propagation in multi-model global water resource reanalysis. *Hydrology and Earth System Sciences*, 23(4):1973–1994.

Ehsani, M. R., Heflin, S., Risanto, C. B., and Behrangi, A. (2022). How well do satellite and reanalysis precipitation products capture north american monsoon season in arizona and new mexico? *Weather and Climate Extremes*, page 100521.

Fang, J., Yang, W., Luan, Y., Du, J., Lin, A., and Zhao, L. (2019). Evaluation of the TRMM 3B42 and GPM IMERG products for extreme precipitation analysis over China. *Atmospheric Research*, 223:24–38.

Fiedler, S., Crueger, T., D'Agostino, R., Peters, K., Becker, T., Leutwyler, D., Paccini, L., Burdanowitz, J., Buehler, S. A., Cortes, A. U., et al. (2020). Simulated tropical precipitation assessed across three major phases of the coupled model intercomparison project (cmip). *Monthly Weather Review*, 148(9):3653–3680.

Fowler, H. J., Wasko, C., and Prein, A. F. (2021). Intensification of short-duration rainfall extremes and implications for flood risk: Current state of the art and future directions. *Philosophical Transactions of the Royal Society A*, 379(2195):20190541.

Freitas, E. d. S., Coelho, V. H. R., Xuan, Y., de CD Melo, D., Gadelha, A. N., Santos, E. A., Galvão, C. d. O., Ramos Filho, G. M., Barbosa, L. R., Huffman, G. J., et al. (2020). The performance of the imerg satellite-based product in identifying sub-daily rainfall events and their properties. *Journal of Hydrology*, 589:125128.

Furl, C., Ghebreyesus, D., and Sharif, H. O. (2018). Assessment of the performance of satellite-based precipitation products for flood events across diverse spatial scales using gssha modeling system. *Geosciences*, 8(6):191.

Gadelha, A. N., Coelho, V. H. R., Xavier, A. C., Barbosa, L. R., Melo, D. C., Xuan, Y., Huffman, G. J., Petersen, W. A., and Almeida, C. d. N. (2019). Grid box-level evaluation of imerg over brazil at various space and time scales. *Atmospheric Research*, 218:231–244.

Gaona, M. F. R., Overeem, A., Brasjen, A., Meirink, J. F., Leijnse, H., and Uijlenhoet, R. (2017). Evaluation of rainfall products derived from satellites and microwave links for the netherlands. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12):6849–6859.

Gaona, M. R., Overeem, A., Leijnse, H., and Uijlenhoet, R. (2016). First-year evaluation of gpm rainfall over the netherlands: Imerg day 1 final run (v03d). *Journal of Hydrometeorology*, 17(11):2799–2814.

Gebregiorgis, A. S., Kirstetter, P.-E., Hong, Y. E., Gourley, J. J., Huffman, G. J., Petersen, W. A., Xue, X., and Schwaller, M. R. (2018). To what extent is the day 1 gpm imerg satellite precipitation estimate improved as compared to trmm tmpa-rt? *Journal of Geophysical Research: Atmospheres*, 123(3):1694–1707.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of climate*, 30(14):5419–5454.

Getirana, A., Kirschbaum, D., Mandarino, F., Ottoni, M., Khan, S., and Arsenault, K. (2020). Potential of gpm imerg precipitation estimates to monitor natural disaster triggers in urban areas: The case of rio de janeiro, brazil. *Remote Sensing*, 12(24):4095.

Grecu, M., Olson, W. S., Munchak, S. J., Ringerud, S., Liao, L., Haddad, Z., Kelley, B. L., and McLaughlin, S. F. (2016). The gpm combined algorithm. *Journal of Atmospheric and Oceanic Technology*, 33(10):2225–2245.

Guo, H., Chen, S., Bao, A., Behrangi, A., Hong, Y., Ndayisaba, F., Hu, J., and Stepanian, P. M. (2016). Early assessment of integrated multi-satellite retrievals for global precipitation measurement over china. *Atmospheric Research*, 176:121–133.

Haile, A. T., Habib, E., Elsaadani, M., and Rientjes, T. (2013). Inter-comparison of satellite rainfall products for representing rainfall diurnal cycle over the nile basin. *International Journal of Applied Earth Observation and Geoinformation*, 21:230–240.

Hamada, A., Murayama, Y., and Takayabu, Y. N. (2014). Regional characteristics of extreme rainfall extracted from trmm pr measurements. *Journal of Climate*, 27(21):8151–8169.

Hamza, A., Anjum, M. N., Masud Cheema, M. J., Chen, X., Afzal, A., Azam, M., Kamran Shafi, M., and Gulakhmadov, A. (2020). Assessment of imerg-v06, trmm-3b42v7, sm2rain-ascat, and persiann-cdr precipitation products over the hindu kush mountains of pakistan, south asia. *Remote Sensing*, 12(23):3871.

Hayden, L. J., Tan, J., Bolvin, D. T., and Huffman, G. J. (2023). Variations in the diurnal cycle of precipitation and its changes with distance from shore over two contrasting regions as observed by imerg, era5, and spaceborne ku radar. *Journal of Hydrometeorology*.

He, Z., Yang, L., Tian, F., Ni, G., Hou, A., and Lu, H. (2017). Intercomparisons of rainfall estimates from trmm and gpm multisatellite products over the upper mekong river basin. *Journal of Hydrometeorology*, 18(2):413–430.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Hong, T., Li, H., and Chen, M. (2021). Comprehensive evaluations on the error characteristics of the state-of-the-art gridded precipitation products over jiangxi province in 2019. *Earth and Space Science*, 8(8):e2021EA001787.

Hong, Y., Hsu, K.-L., Sorooshian, S., and Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology*, 43(12):1834–1853.

Hong, Y., Tang, G., Ma, Y., Huang, Q., Han, Z., Zeng, Z., Yang, Y., Wang, C., and Guo, X. (2019). Remote sensing precipitation: Sensors, retrievals, validations, and applications. *Observation and Measurement; Li, X., Vereecken, H., Eds*, pages 1–23.

Hou, A., Kakar, R., and Neeck, S. (2013). Azarbarzin. *AA, Kummerow, CD, Kojima, M., Oki, R., Nakamura, K., and Iguchi, T.: The Global Precipitation Measurement (GPM) mission. Bull. Am. Meteorol. Soc*, 95:701–722.

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):701–722. Publisher: American Meteorological Society.

Hsu, K.-l., Gao, X., Sorooshian, S., and Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, 36(9):1176–1190.

Hu, X. and Yuan, W. (2021). Evaluation of era5 precipitation over the eastern periphery of the tibetan plateau from the perspective of regional rainfall events. *International Journal of Climatology*, 41(4):2625–2637.

Huang, C., Hu, J., Chen, S., Zhang, A., Liang, Z., Tong, X., Xiao, L., Min, C., and Zhang, Z. (2019). How well can imerg products capture typhoon extreme precipitation events over southern china? *Remote Sensing*, 11(1):70.

Huang, W.-R., Chang, Y.-H., and Liu, P.-Y. (2018). Assessment of imerg precipitation over taiwan at multiple timescales. *Atmospheric Research*, 214:239–249.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J., and Xie, P. (2018). Nasa global precipitation measurement (gpm) integrated multi-satellite retrievals for gpm (imerg). *Algorithm Theoretical Basis Document (ATBD) Version*, 4.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., and Yoo, S.-H. (2015). NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theoretical Basis Document (ATBD) Version*, 4:26.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Joyce, R. J., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., et al. (2020). Integrated multi-satellite retrievals for the global precipitation measurement (gpm) mission (imerg). In *Satellite Precipitation Measurement*, pages 343–353. Springer.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F. (2007). The trmm multisatellite precipitation analysis (tmpa): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1):38–55.

Huffman, G. J. B. (2020). Reaching for 20 Years with the IMERG Multi-Satellite Products. Library Catalog: NASA NTRS.

Islam, M. A., Yu, B., and Cartwright, N. (2020). Assessment and comparison of five satellite precipitation products in australia. *Journal of Hydrology*, 590:125474.

Janowiak, J. E., Kousky, V. E., and Joyce, R. J. (2005). Diurnal cycle of precipitation determined from the cmorph high spatial and temporal resolution global precipitation analyses. *Journal of Geophysical Research: Atmospheres*, 110(D23).

Jiang, L. and Bauer-Gottwein, P. (2019). How do gpm imerg precipitation estimates perform as hydrological model forcing? evaluation for 300 catchments across mainland china. *Journal of Hydrology*, 572:486–500.

Jiang, Q., Li, W., Fan, Z., He, X., Sun, W., Chen, S., Wen, J., Gao, J., and Wang, J. (2021). Evaluation of the era5 reanalysis precipitation dataset over chinese mainland. *Journal of hydrology*, 595:125660.

Jiang, S., Ren, L., Xu, C.-Y., Yong, B., Yuan, F., Liu, Y., Yang, X., and Zeng, X. (2018). Statistical and hydrological evaluation of the latest integrated multi-satellite retrievals for gpm (imerg) over a midlatitude humid basin in south china. *Atmospheric Research*, 214:418–429.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P. (2004). Cmorph: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5(3):487–503.

Joyce, R. J. and Xie, P. (2011). Kalman filter–based cmorph. *Journal of Hydrometeorology*, 12(6):1547–1563.

Kazamias, A.-P., Sapountzis, M., and Lagouvardos, K. (2022). Evaluation of gpm-imerg rainfall estimates at multiple temporal and spatial scales over greece. *Atmospheric Research*, 269:106014.

Keller, J. D. and Wahl, S. (2021). Representation of climate in reanalyses: An intercomparison for europe and north america. *Journal of Climate*, 34(5):1667–1684.

Khan, S. and Maggioni, V. (2019). Assessment of level-3 gridded global precipitation mission (gpm) products over oceans. *Remote Sensing*, 11(3):255.

Khodadoust Siuki, S., Saghafian, B., and Moazami, S. (2017). Comprehensive evaluation of 3-hourly trmm and half-hourly gpm-imerg satellite precipitation products. *International Journal of Remote Sensing*, 38(2):558–571.

Kidd, C. (2019). Algorithm theoretical basis document (atbd) version 4.5: Nasa global precipitation measurement (gpm) integrated multi-satellite retrievals for gpm (imerg). *National Aeronautics and Space Administration (NASA): Washington, DC, USA*, 34.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B. (2017). So, how much of the Earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, 98(1):69–78.

Kidd, C. and Huffman, G. (2011). Global precipitation measurement. *Meteorological Applications*, 18(3):334–353. Publisher: Wiley Online Library.

Kidd, C. and Levizzani, V. (2011). Status of satellite precipitation retrievals. *Hydrology & Earth System Sciences*, 15(4).

Kim, K., Park, J., Baik, J., and Choi, M. (2017). Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmospheric Research*, 187:95–105.

Kirstetter, P.-E. et al. (2018). Evaluation of diurnal variation of gpm imerg-derived summer precipitation over the contiguous us using mrms data. *Quarterly Journal of the Royal Meteorological Society*.

Klepp, C., Michel, S., Protat, A., Burdanowitz, J., Albern, N., Kähnert, M., Dahl, A., Louf, V., Bakan, S., and Buehler, S. A. (2018). Oceanrain, a new in-situ shipboard global ocean surface-reference dataset of all water cycle components. *Scientific data*, 5:180122.

Koschmieder, H. (1934). Methods and results of definite rain measurements: Iii. danzig report (1). *Monthly Weather Review*, 62(1):5–7.

Kubota, T., Aonashi, K., Ushio, T., Shige, S., Takayabu, Y. N., Kachi, M., Arai, Y., Tashima, T., Masaki, T., Kawamoto, N., et al. (2020). Global satellite mapping of precipitation (gsmap) products in the gpm era. *Satellite precipitation measurement*, 1:355–373.

Kucera, P. and Klepp, C. (2018). Evaluation of high resolution imerg satellite precipitation over the global oceans using oceanrain. In *EGU General Assembly Conference Abstracts*, page 17673.

Kumar, M., Hodnebrog, Ø., Daloz, A. S., Sen, S., Badiger, S., and Krishnaswamy, J. (2021). Measuring precipitation in eastern himalaya: Ground validation of eleven satellite, model and gauge interpolated gridded products. *Journal of Hydrology*, 599:126252.

Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N.-Y., Ferraro, R., Munchak, S. J., and Petkovic, V. (2015). The evolution of the goddard profiling algorithm to a fully parametric scheme. *Journal of atmospheric and oceanic technology*, 32(12):2265–2280.

Lang, T. J., Nesbitt, S. W., and Carey, L. D. (2009). On the correction of partial beam blockage in polarimetric radar data. *Journal of Atmospheric and Oceanic Technology*, 26(5):943–957.

Le, M.-H., Lakshmi, V., Bolten, J., and Du Bui, D. (2020). Adequacy of satellite-derived precipitation estimate for hydrological modeling in vietnam basins. *Journal of Hydrology*, 586:124820.

Lee, J., Lee, E.-H., and Seol, K.-H. (2019). Validation of integrated multisatellite retrievals for gpm (imerg) by using gauge-based analysis products of daily precipitation over east asia. *Theoretical and Applied Climatology*, 137(3):2497–2512.

Levizzani, V. and Cattani, E. (2019). Satellite remote sensing of precipitation and the terrestrial water cycle in a changing climate. *Remote Sensing*, 11(19):2301.

Lex, A. and Gehlenborg, N. (2014). Points of view: Sets and intersections. *nature methods*, 11(8):779.

Li, L., Hong, Y., Wang, J., Adler, R. F., Policelli, F. S., Habib, S., Irwn, D., Korme, T., and Okello, L. (2009). Evaluation of the real-time trmm-based multi-satellite precipitation analysis for an operational flood prediction system in nzoia basin, lake victoria, africa. *Natural Hazards*, 50(1):109–123.

Li, N., Tang, G., Zhao, P., Hong, Y., Gou, Y., and Yang, K. (2017). Statistical assessment and hydrological utility of the latest multi-satellite precipitation analysis imerg in ganjiang river basin. *Atmospheric Research*, 183:212–223.

Li, R., Shi, J., Ji, D., Zhao, T., Plermkamon, V., Moukomla, S., Kuntiyawichai, K., and Kruasilp, J. (2019). Evaluation and hydrological application of trmm and gpm precipitation products in a tropical monsoon basin of thailand. *Water*, 11(4):818.

Li, R., Wang, K., and Qi, D. (2018). Validating the integrated multisatellite retrievals for global precipitation measurement in terms of diurnal variability with hourly gauge observations collected at 50,000 stations in china. *Journal of Geophysical Research: Atmospheres*, 123(18):10–423.

Li, X., Chen, S., Liang, Z., Huang, C., Li, Z., and Hu, B. (2021a). Performance assessment of gsmap and gpm imerg products during typhoon mangkhut. *Atmosphere*, 12(2):134.

Li, X., Chen, Y., Wang, H., and Zhang, Y. (2020a). Assessment of gpm imerg and radar quantitative precipitation estimation (qpe) products using dense rain gauge observations in the guangdong-hong kong-macao greater bay area, china. *Atmospheric Research*, 236:104834.

Li, X., Sungmin, O., Wang, N., Huang, Y., et al. (2021b). Evaluation of the gpm imerg v06 products for light rain over mainland china. *Atmospheric Research*, 253:105510.

Li, Z., Chen, M., Gao, S., Hong, Z., Tang, G., Wen, Y., Gourley, J. J., and Hong, Y. (2020b). Cross-examination of similarity, difference and deficiency of gauge, radar and satellite precipitation measuring uncertainties for extreme events using conventional metrics and multiplicative triple collocation. *Remote Sensing*, 12(8):1258.

Li, Z., Tang, G., Kirstetter, P., Gao, S., Li, J.-L., Wen, Y., and Hong, Y. (2022). Evaluation of gpm imerg and its constellations in extreme events over the conterminous united states. *Journal of Hydrology*, 606:127357.

Liang, Z., Chen, S., Hu, J., Huang, C., Zhang, A., Xiao, L., Zhang, Z., and Tong, X. (2019). Hydrologic evaluation of integrated multi-satellite retrievals for gpm over nanliu river basin in tropical humid southern china. *Water*, 11(5):932.

Liu, C.-Y., Aryastana, P., Liu, G.-R., and Huang, W.-R. (2020). Assessment of satellite precipitation product estimates over bali island. *Atmospheric Research*, 244:105032.

Liu, Z. (2016). Comparison of integrated multisatellite retrievals for GPM (IMERG) and TRMM multisatellite precipitation analysis (TMPA) monthly precipitation products: initial results. *Journal of Hydrometeorology*, 17(3):777–790.

Lu, D. and Yong, B. (2018). Evaluation and hydrological utility of the latest gpm imerg v5 and gsmap v7 precipitation products over the tibetan plateau. *Remote Sensing*, 10(12):2022.

Lv, X., Guo, H., Tian, Y., Meng, X., Bao, A., and De Maeyer, P. (2024). Evaluation of gsmap version 8 precipitation products on an hourly timescale over mainland china. *Remote Sensing*, 16(1):210.

Ma, Y., Yang, Y., Han, Z., Tang, G., Maguire, L., Chu, Z., and Hong, Y. (2018). Comprehensive evaluation of ensemble multi-satellite precipitation dataset using the dynamic bayesian model averaging scheme over the tibetan plateau. *Journal of Hydrology*, 556:634–644.

Maggioni, V., Meyers, P. C., and Robinson, M. D. (2016). A review of merged high-resolution satellite precipitation product accuracy during the tropical rainfall measuring mission (trmm) era. *Journal of Hydrometeorology*, 17(4):1101–1117.

Maggioni, V., Nikolopoulos, E. I., Anagnostou, E. N., and Borga, M. (2017). Modeling satellite precipitation errors over mountainous terrain: The influence of gauge density, seasonality, and temporal resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7):4130–4140.

Maghsood, F. F., Hashemi, H., Hosseini, S. H., and Berndtsson, R. (2020). Ground validation of gpm imerg precipitation products over iran. *Remote Sensing*, 12(1):48.

Mahmoud, M. T., Al-Zahrani, M. A., and Sharif, H. O. (2018). Assessment of global precipitation measurement satellite products over saudi arabia. *Journal of Hydrology*, 559:1–12.

Mahmoud, M. T., Hamouda, M. A., and Mohamed, M. M. (2019). Spatiotemporal evaluation of the gpm satellite precipitation products over the united arab emirates. *Atmospheric Research*, 219:200–212.

Manz, B., Páez-Bimos, S., Horna, N., Buytaert, W., Ochoa-Tocachi, B., Lavado-Casimiro, W., and Willems, B. (2017). Comparative ground validation of imerg and tmpa at variable spatiotemporal scales in the tropical andes. *Journal of Hydrometeorology*, 18(9):2469–2489.

Markonis, Y., Pappas, C., Hanel, M., and Papalexiou, S. M. (2021). A cross-scale framework for integrating multi-source data in earth system sciences. *Environmental Modelling & Software*, page 104997.

Marzuki, M., Suryanti, K., Yusnaini, H., Tangang, F., Muharsyah, R., Vonnisa, M., and Devianto, D. (2021). Diurnal variation of precipitation from the perspectives of precipitation amount, intensity and duration over sumatra from rain gauge observations. *International Journal of Climatology*, 41(8):4386–4397.

Mayor, Y. G., Tereshchenko, I., Fonseca-Hernández, M., Pantoja, D. A., and Montes, J. M. (2017). Evaluation of error in imerg precipitation estimates under different topographic conditions and temporal scales over mexico. *Remote Sensing*, 9(5):503.

Mazzoglio, P., Laio, F., Balbo, S., Boccardo, P., and Disabato, F. (2019). Improving an extreme rainfall detection system with gpm imerg data. *Remote Sensing*, 11(6):677.

McClean, F., Dawson, R., and Kilsby, C. (2021). Intercomparison of global reanalysis precipitation for flood risk modelling. *Hydrology and Earth System Sciences Discussions*, pages 1–13.

McPhaden, M. J., Busalacchi, A. J., Cheney, R., Donguy, J.-R., Gage, K. S., Halpern, D., Ji, M., Julian, P., Meyers, G., Mitchum, G. T., et al. (1998). The tropical ocean-global atmosphere observing system: A decade of progress. *Journal of Geophysical Research: Oceans*, 103(C7):14169–14240.

Mcphaden, M. J., Meyers, G., Ando, K., Masumoto, Y., Murty, V., Ravichandran, M., Syamsudin, F., Vialard, J., Yu, L., and Yu, W. (2009). Rama: the research moored array for african–asian–australian monsoon analysis and prediction. *Bulletin of the American Meteorological Society*, 90(4):459–480.

Mega, T., Ushio, T., Takahiro, M., Kubota, T., Kachi, M., and Oki, R. (2018). Gauge-adjusted global satellite mapping of precipitation. *IEEE Transactions on Geoscience and Remote Sensing*, 57(4):1928–1935.

Mei, Y., Anagnostou, E. N., Shen, X., and Nikolopoulos, E. I. (2017). Decomposing the satellite precipitation error propagation through the rainfall-runoff processes. *Advances in Water Resources*, 109:253–266.

Mo, C., Zhang, M., Ruan, Y., Qin, J., Wang, Y., Sun, G., and Xing, Z. (2020). Accuracy analysis of imerg satellite rainfall data and its application in long-term runoff simulation. *Water*, 12(8):2177.

Montoya Duque, E., Huang, Y., May, P., and Siems, S. (2023). An evaluation of imerg and era5 quantitative precipitation estimates over the southern ocean using shipborne observations. *Journal of Applied Meteorology and Climatology*, 62(11):1479–1495.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3):885–900.

Navarro, A., García-Ortega, E., Merino, A., Sánchez, J. L., Kummerow, C., and Tapiador, F. J. (2019). Assessment of imerg precipitation estimates over europe. *Remote Sensing*, 11(21):2470.

Navarro, A., García-Ortega, E., Merino, A., Sánchez, J. L., and Tapiador, F. J. (2020). Orographic biases in imerg precipitation estimates in the ebro river basin (spain): The effects of rain gauge density and altitude. *Atmospheric Research*, 244:105068.

Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., Ashouri, H., and Thorstensen, A. R. (2018). The persiann family of global satellite precipitation data: A review and evaluation of products. *Hydrology and Earth System Sciences*, 22(11):5801–5816.

Nguyen, P., Shearer, E. J., Tran, H., Ombadi, M., Hayatbini, N., Palacios, T., Huynh, P., Braithwaite, D., Updegraff, G., Hsu, K., et al. (2019). The chrs data portal, an easily accessible public repository for persiann global satellite precipitation data. *Scientific data*, 6(1):1–10.

Ning, S., Song, F., Udmale, P., Jin, J., Thapa, B. R., and Ishidaira, H. (2017). Error analysis and evaluation of the latest gsmap and imerg precipitation products over eastern china. *Advances in Meteorology*, 2017.

Nogueira, M. (2020). Inter-comparison of era-5, era-interim and gpcp rainfall over the last 40 years: Process-based analysis of systematic and random differences. *Journal of Hydrology*, 583:124632.

O, S. and Kirstetter, P.-E. (2018). Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. *Quarterly Journal of the Royal Meteorological Society*, 144(S1):270–281.

Oliveira, R., Maggioni, V., Vila, D., and Morales, C. (2016). Characteristics and diurnal cycle of gpm rainfall estimates over the central amazon region. *Remote Sensing*, 8(7):544.

Olson, W. S., Masunaga, H., and Gpm, C. R.-R. A. T. (2016). Gpm combined radar-radiometer precipitation algorithm theoretical basis document (version 4). *NASA: Washington, DC, USA*.

Omranian, E., Sharif, H. O., and Tavakoly, A. A. (2018). How well can global precipitation measurement (gpm) capture hurricanes? case study: Hurricane harvey. *Remote Sensing*, 10(7):1150.

Palomino-Ángel, S., Anaya-Acevedo, J. A., and Botero, B. A. (2019). Evaluation of 3b42v7 and imerg daily-precipitation products for a very high-precipitation region in northwestern south america. *Atmospheric Research*, 217:37–48.

Papalexiou, S. M., Markonis, Y., Lombardo, F., AghaKouchak, A., and Foufoula-Georgiou, E. (2018). Precise temporal disaggregation preserving marginals and correlations (dipmac) for stationary and nonstationary processes. *Water Resources Research*, 54(10):7435–7458.

Peng, F., Zhao, S., Chen, C., Cong, D., Wang, Y., and Ouyang, H. (2020). Evaluation and comparison of the precipitation detection ability of multiple satellite products in a typical agriculture area of china. *Atmospheric Research*, 236:104814.

Pfeifroth, U., Trentmann, J., Fink, A. H., and Ahrens, B. (2016). Evaluating satellite-based diurnal cycles of precipitation in the african tropics. *Journal of Applied Meteorology and Climatology*, 55(1):23–39.

Pradhan, R. K. and Markonis, Y. (2023). Performance evaluation of gpm imerg precipitation products over the tropical oceans using buoys. *Journal of Hydrometeorology*, 24(10):1755–1770.

Pradhan, R. K., Markonis, Y., Godoy, M. R. V., Villalba-Pradas, A., Andreadis, K. M., Nikolopoulos, E. I., Papalexiou, S. M., Rahim, A., Tapiador, F. J., and Hanel, M. (2022). Review of gpm imerg performance: A global perspective. *Remote Sensing of Environment*, 268:112754.

Prakash, S., Kumar, M. R., Mathew, S., and Venkatesan, R. (2018a). How accurate are satellite estimates of precipitation over the north indian ocean? *Theoretical and Applied Climatology*, 134(1-2):467–475.

Prakash, S., Mahesh, C., and Gairola, R. (2013). Comparison of trmm multi-satellite precipitation analysis (tmpa)-3b43 version 6 and 7 products with rain gauge data from ocean buoys. *Remote Sensing Letters*, 4(7):677–685.

Prakash, S., Mahesh, C., Gairola, R. M., Pokhrel, S., et al. (2011). Surface freshwater flux estimation using trmm measurements over the tropical oceans. *Atmospheric and Climate Sciences*, 1(04):225.

Prakash, S., Mitra, A. K., AghaKouchak, A., Liu, Z., Norouzi, H., and Pai, D. S. (2018b). A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *Journal of Hydrology*, 556:865–876.

Prakash, S., Mitra, A. K., Momin, I. M., Pai, D. S., Rajagopal, E. N., and Basu, S. (2015). Comparison of TMPA-3B42 versions 6 and 7 precipitation products with gauge-based data over India for the southwest monsoon period. *Journal of Hydrometeorology*, 16(1):346–362.

Protat, A., Klepp, C., Louf, V., Petersen, W. A., Alexander, S. P., Barros, A., Leinonen, J., and Mace, G. G. (2019). The latitudinal variability of oceanic rainfall properties and its implication for satellite retrievals: 1. drop size distribution properties. *Journal of Geophysical Research: Atmospheres*, 124(23):13291–13311.

Qin, S., Wang, K., Wu, G., and Ma, Z. (2021). Variability of hourly precipitation during the warm season over eastern china using gauge observations and era5. *Atmospheric Research*, 264:105872.

Rahman, K. U., Shang, S., Shahid, M., and Li, J. (2018). Developing an ensemble precipitation algorithm from satellite products and its topographical and seasonal evaluations over pakistan. *Remote Sensing*, 10(11):1835.

Rajagopal, M., Zipser, E., Huffman, G., Russell, J., and Tan, J. (2021). Comparisons of imerg version 06 precipitation at and between passive microwave overpasses in the tropics. *Journal of Hydrometeorology*, 22(8):2117–2130.

Ramadhan, R., Marzuki, M., Yusnaini, H., Muharsyah, R., Tangang, F., Vonnisa, M., and Harmadi, H. (2023). A preliminary assessment of the gsmap version 08 products over indonesian maritime continent against gauge data. *Remote Sensing*, 15(4):1115.

Ramsauer, T., Weiß, T., and Marzahn, P. (2018). Comparison of the gpm imerg final precipitation product to radolan weather radar data over the topographically and climatically diverse germany. *Remote Sensing*, 10(12):2029.

Randel, D. L., Kummerow, C. D., and Ringerud, S. (2020). The goddard profiling (gprof) precipitation retrieval algorithm. *Satellite Precipitation Measurement: Volume 1*, pages 141–152.

Retalis, A., Katsanos, D., Tymvios, F., and Michaelides, S. (2018). Validation of the first years of gpm operation over cyprus. *Remote Sensing*, 10(10):1520.

Retalis, A., Katsanos, D., Tymvios, F., and Michaelides, S. (2020). Comparison of gpm imerg and trmm 3b43 products over cyprus. *Remote Sensing*, 12(19):3212.

Rojas, Y., Minder, J. R., Campbell, L. S., Massmann, A., and Garreaud, R. (2021). Assessment of gpm imerg satellite precipitation estimation and its dependence on microphysical rain regimes over the mountains of south-central chile. *Atmospheric Research*, 253:105454.

Rozante, J. R. and Cavalcanti, I. F. A. (2008). Regional eta model experiments: Salljex and mcs development. *Journal of Geophysical Research: Atmospheres*, 113(D17).

Rozante, J. R., Vila, D. A., Barboza Chiquetto, J., Fernandes, A. d. A., and Souza Alvim, D. (2018). Evaluation of trmm/gpm blended daily products over brazil. *Remote Sensing*, 10(6):882.

Sadeghi, L., Saghafian, B., and Moazami, S. (2019). Evaluation of imerg and mrms remotely sensed snowfall products. *International Journal of Remote Sensing*, 40(11):4175–4192.

Sahany, S., Venugopal, V., and Nanjundiah, R. S. (2010). Diurnal-scale signatures of monsoon rainfall over the indian region from trmm satellite observations. *Journal of Geophysical Research: Atmospheres*, 115(D2).

Sahlu, D., Nikolopoulos, E. I., Moges, S. A., Anagnostou, E. N., and Hailu, D. (2016). First evaluation of the Day-1 IMERG over the upper Blue Nile basin. *Journal of Hydrometeorology*, 17(11):2875–2882.

Salles, L., Satgé, F., Roig, H., Almeida, T., Olivetti, D., and Ferreira, W. (2019). Seasonal effect on spatial and temporal consistency of the new gpm-based imerg-v5 and gsmap-v7 satellite precipitation estimates in brazil's central plateau region. *Water*, 11(4):668.

Saouabe, T., El Khalki, E. M., Saidi, M. E. M., Najmi, A., Hadri, A., Rachidi, S., Jadoud, M., Tramblay, Y., et al. (2020). Evaluation of the gpm-imerg precipitation product for flood modeling in a semi-arid mountainous basin in morocco. *Water*, 12(9):2516.

Sapiano, M. and Arkin, P. (2009). An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *Journal of Hydrometeorology*, 10(1):149–166.

Satgé, F., Hussain, Y., Bonnet, M.-P., Hussain, B. M., Martinez-Carvajal, H., Akhter, G., and Uagoda, R. (2018). Benefits of the successive gpm based satellite precipitation estimates imerg–v03,–v04,–v05 and gsmap–v06,–v07 over diverse geomorphic and meteorological regions of pakistan. *Remote Sensing*, 10(9):1373.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B. (2014). Gpcc's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, 115:15–40.

Serra, Y. L. and McPhaden, M. J. (2003). Multiple time-and space-scale comparisons of atlas buoy rain gauge measurements with trmm satellite precipitation measurements. *Journal of Applied Meteorology*, 42(8):1045–1059.

Sharifi, E., Eitzinger, J., and Dorigo, W. (2019). Performance of the state-of-the-art gridded precipitation products over mountainous terrain: A regional study over austria. *Remote Sensing*, 11(17):2018.

Sharifi, E., Steinacker, R., and Saghafian, B. (2016). Assessment of gpm-imerg and other precipitation products against gauge data under different topographic and climatic conditions in iran: Preliminary results. *Remote Sensing*, 8(2):135.

Sharma, S., Chen, Y., Zhou, X., Yang, K., Li, X., Niu, X., Hu, X., and Khadka, N. (2020). Evaluation of gpm-era satellite precipitation products on the southern slopes of the central himalayas against rain gauge data. *Remote Sensing*, 12(11):1836.

Shawky, M., Moussa, A., Hassan, Q. K., and El-Sheimy, N. (2019). Performance assessment of sub-daily and daily precipitation estimates derived from gpm and gsmap products over an arid environment. *Remote Sensing*, 11(23):2840.

Shi, J., Yuan, F., Shi, C., Zhao, C., Zhang, L., Ren, L., Zhu, Y., Jiang, S., and Liu, Y. (2020). Statistical evaluation of the latest gpm-era imerg and gsmap satellite precipitation products in the yellow river source region. *Water*, 12(4):1006.

Shrestha, S., Zaramella, M., Callegari, M., Greifeneder, F., and Borga, M. (2023). Scale dependence of errors in snow water equivalent simulations using era5 reanalysis over alpine basins.

Siddique-E-Akbor, A., Hossain, F., Sikder, S., Shum, C., Tseng, S., Yi, Y., Turk, F., and Limaye, A. (2014). Satellite precipitation data–driven hydrological modeling for water resources management in the ganges, brahmaputra, and meghna basins. *Earth Interactions*, 18(17):1–25.

Siems, S. T., Huang, Y., and Manton, M. J. (2022). Southern ocean precipitation: Toward a process-level understanding. *Wiley Interdisciplinary Reviews: Climate Change*, 13(6):e800.

Singh, A. K., Tripathi, J., Singh, K., Singh, V., and Sateesh, M. (2019). Comparison of different satellite-derived rainfall products with imd gridded data over indian meteorological subdivisions during indian summer monsoon (ism) 2016 at weekly temporal resolution. *Journal of Hydrology*, 575:1371–1379.

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., et al. (2017). The global precipitation measurement (gpm) mission for science and society. *Bulletin of the American Meteorological Society*, 98(8):1679–1695.

So, D. and Shin, D.-B. (2018). Classification of precipitating clouds using satellite infrared observations and its implications for rainfall estimation. *Quarterly Journal of the Royal Meteorological Society*, 144:133–144.

Song, Y., Zhang, J., Meng, X., Zhou, Y., Lai, Y., and Cao, Y. (2020). Comparison study of multiple precipitation forcing data on hydrological modeling and projection in the qujiang river basin. *Water*, 12(9):2626.

Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H. V., Imam, B., and Braithwaite, D. (2000). Evaluation of persiann system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, 81(9):2035–2046.

Su, J., Lü, H., Zhu, Y., Cui, Y., and Wang, X. (2019). Evaluating the hydrological utility of latest imerg products over the upper huaihe river basin, china. *Atmospheric Research*, 225:17–29.

Su, J., Lü, H., Zhu, Y., Wang, X., and Wei, G. (2018). Component analysis of errors in four gpm-based precipitation estimations over mainland china. *Remote Sensing*, 10(9):1420.

Sui, X., Li, Z., Ma, Z., Xu, J., Zhu, S., and Liu, H. (2020). Ground validation and error sources identification for gpm imerg product over the southeast coastal regions of china. *Remote Sensing*, 12(24):4154.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1):79–107.

Sungmin, O., Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J., and Petersen, W. A. (2017). Evaluation of gpm imerg early, late, and final rainfall estimates using wegenernet gauge data in southeastern austria. *Hydrology & Earth System Sciences*, 21(12).

Sungmin, O. and Kirstetter, P.-E. (2018). Evaluation of diurnal variation of gpm imerg-derived summer precipitation over the contiguous us using mrms data. *Quarterly Journal of the Royal Meteorological Society*, 144:270–281.

Sunilkumar, K., Yatagai, A., and Masuda, M. (2019). Preliminary evaluation of gpm-imerg rainfall estimates over three distinct climate zones with aphrodite. *Earth and Space Science*, 6(8):1321–1335.

Tai, S.-L., Feng, Z., Ma, P.-L., Schumacher, C., and Fast, J. D. (2021). Representations of precipitation diurnal cycle in the amazon as simulated by observationally constrained cloud-system resolving and global climate models. *Journal of Advances in Modeling Earth Systems*, 13(11):e2021MS002586.

Tan, J., Huffman, G. J., Bolvin, D. T., and Nelkin, E. J. (2019a). Diurnal cycle of imerg v06 precipitation. *Geophysical Research Letters*, 46(22):13584–13592.

Tan, J., Huffman, G. J., Bolvin, D. T., and Nelkin, E. J. (2019b). Diurnal Cycle of IMERG V06 Precipitation. *Geophysical Research Letters*, 46(22):13584–13592.

Tan, J., Huffman, G. J., Bolvin, D. T., and Nelkin, E. J. (2019c). Imerg v06: Changes to the morphing algorithm. *Journal of Atmospheric and Oceanic Technology*, 36(12):2471–2482.

Tan, J., Huffman, G. J., Bolvin, D. T., Nelkin, E. J., and Rajagopal, M. (2021). Sharpen: A scheme to restore the distribution of averaged precipitation fields. *Journal of Hydrometeorology*, 22(8):2105–2116.

Tan, J., Petersen, W. A., Kirstetter, P.-E., and Tian, Y. (2017). Performance of imerg as a function of spatiotemporal scale. *Journal of hydrometeorology*, 18(2):307–319.

Tan, J., Petersen, W. A., and Tokay, A. (2016). A novel approach to identify sources of errors in imerg for gpm ground validation. *Journal of Hydrometeorology*, 17(9):2477–2491.

Tan, M. L. and Duan, Z. (2017). Assessment of GPM and TRMM precipitation products over Singapore. *Remote Sensing*, 9(7):720.

Tan, M. L., Samat, N., Chan, N. W., and Roy, R. (2018). Hydro-meteorological assessment of three gpm satellite precipitation products in the kelantan river basin, malaysia. *Remote Sensing*, 10(7):1011.

Tan, M. L. and Santo, H. (2018). Comparison of gpm imerg, tmpa 3b42 and persiann-cdr satellite precipitation products over malaysia. *Atmospheric Research*, 202:63–76.

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., and Hong, Y. (2020). Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote Sensing of Environment*, 240:111697.

Tang, G., Ma, Y., Long, D., Zhong, L., and Hong, Y. (2016a). Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. *Journal of Hydrology*, 533:152–167.

Tang, G., Zeng, Z., Long, D., Guo, X., Yong, B., Zhang, W., and Hong, Y. (2016b). Statistical and hydrological comparisons between trmm and gpm level-3 products over a midlatitude basin: Is day-1 imerg a good successor for tmpa 3b42v7? *Journal of Hydrometeorology*, 17(1):121–137.

Tapiador, F., Navarro, A., Levizzani, V., García-Ortega, E., Huffman, G., Kidd, C., Kucera, P., Kummerow, C., Masunaga, H., Petersen, W., et al. (2017). Global precipitation measurements for validating climate models. *Atmospheric Research*, 197:1–20.

Tapiador, F. J., Angelis, C. F., Viltard, N., Cuartero, F., and De Castro, M. (2011). On the suitability of regional climate models for reconstructing climatologies. *Atmospheric Research*, 101(3):739–751.

Tapiador, F. J., Roca, R., Del Genio, A., Dewitte, B., Petersen, W., and Zhang, F. (2019). Is precipitation a good metric for model performance? *Bulletin of the American Meteorological Society*, 100(2):223–233.

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., et al. (2012). Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research*, 104:70–97.

Terblanche, D., Lynch, A., Chen, Z., and Sinclair, S. (2022). Era5-derived precipitation: Insights from historical rainfall networks in southern africa. *Journal of Applied Meteorology and Climatology*, 61(10):1473–1484.

Tian, F., Hou, S., Yang, L., Hu, H., and Hou, A. (2018). How does the evaluation of the gpm imerg rainfall product depend on gauge density and rainfall intensity? *Journal of Hydrometeorology*, 19(2):339–349.

Tian, Y. and Peters-Lidard, C. D. (2007). Systematic anomalies over inland water bodies in satellite-based precipitation estimates. *Geophysical Research Letters*, 34(14).

Tian, Y., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., Hsu, K.-l., Turk, F. J., Garcia, M., and Zeng, J. (2009). Component analysis of errors in satellite-based precipitation estimates. *Journal of Geophysical Research: Atmospheres*, 114(D24).

Trenberth, K. E., Smith, L., Qian, T., Dai, A., and Fasullo, J. (2007). Estimates of the global water budget and its annual cycle using observational and model data. *Journal of Hydrometeorology*, 8(4):758–769.

Trenberth, K. E., Zhang, Y., and Gehne, M. (2017). Intermittency in precipitation: Duration, frequency, intensity, and amounts using hourly data. *Journal of Hydrometeorology*, 18(5):1393–1412.

Vargas Godoy, M. R., Markonis, Y., Hanel, M., Kyselỳ, J., and Papalexiou, S. M. (2021). The global water cycle budget: A chronological review. *Surveys in Geophysics*, pages 1–33.

Wang, C., Tang, G., Han, Z., Guo, X., and Hong, Y. (2018). Global intercomparison and regional evaluation of GPM IMERG Version-03, Version-04 and its latest Version-05 precipitation products: Similarity, difference and improvements. *Journal of Hydrology*, 564:342–356.

Wang, D., Wang, X., Liu, L., Wang, D., Huang, H., and Pan, C. (2019a). Evaluation of tmpa 3b42v7, gpm imerg and cmpa precipitation estimates in guangdong province, china. *International Journal of Climatology*, 39(2):738–755.

Wang, J., Wolff, D. B., Tan, J., Marks, D. A., Pippitt, J. L., and Huffman, G. J. (2022). Validation of imerg oceanic precipitation over kwajalein. *Remote Sensing*, 14(15):3753.

Wang, S., Liu, J., Wang, J., Qiao, X., and Zhang, J. (2019b). Evaluation of gpm imerg v05b and trmm 3b42v7 precipitation products over high mountainous tributaries in lhasa with dense rain gauges. *Remote Sensing*, 11(18):2080.

Wang, W., Lu, H., Zhao, T., Jiang, L., and Shi, J. (2017a). Evaluation and comparison of daily rainfall from latest gpm and trmm products over the mekong river basin. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(6):2540–2549.

Wang, X., Ding, Y., Zhao, C., and Wang, J. (2019c). Similarities and improvements of gpm imerg upon trmm 3b42 precipitation product under complex topographic and climatic conditions over hexi region, northeastern tibetan plateau. *Atmospheric Research*, 218:347–363.

Wang, Z., Zhong, R., Lai, C., and Chen, J. (2017b). Evaluation of the gpm imerg satellite-based precipitation products and the hydrological utility. *Atmospheric Research*, 196:151–163.

Watters, D. and Battaglia, A. (2019a). The summertime diurnal cycle of precipitation derived from imerg. *Remote Sensing*, 11(15):1781.

Watters, D. and Battaglia, A. (2019b). The Summertime Diurnal Cycle of Precipitation Derived from IMERG. *Remote Sensing*, 11(15).

Watters, D. and Battaglia, A. (2021). The nasa-jaxa global precipitation measurement mission–part ii: New frontiers in precipitation science. *Weather*, 76(2):52–56.

Watters, D., Battaglia, A., and Allan, R. (2021). The Diurnal Cycle of Precipitation According to Multiple Decades of Global Satellite Observations, Three CMIP6 Models, and the ECMWF Reanalysis. *Journal of Climate*.

Wei, G., Lü, H., T Crow, W., Zhu, Y., Wang, J., and Su, J. (2018). Evaluation of satellite-based precipitation products from imerg v04a and v03d, cmorph and tmpa with gauged rainfall in three climatologic zones in china. *Remote Sensing*, 10(1):30.

Wen, Y., Behrangi, A., Chen, H., and Lambrigtsen, B. (2018). How well were the early 2017 california atmospheric river precipitation events captured by satellite products and ground-based radars? *Quarterly Journal of the Royal Meteorological Society*, 144:344–359.

Wen, Y., Behrangi, A., Lambrigtsen, B., and Kirstetter, P.-E. (2016). Evaluation and uncertainty estimation of the latest radar and satellite snowfall products using snotel measurements over mountainous regions in western united states. *Remote Sensing*, 8(11):904.

Wu, L., Xu, Y., and Wang, S. (2018). Comparison of tmpa-3b42rt legacy product and the equivalent imerg products over mainland china. *Remote Sensing*, 10(11):1778.

Wu, Q. and Wang, Y. (2019). Comparison of oceanic multisatellite precipitation data from tropical rainfall measurement mission and global precipitation measurement mission datasets with rain gauge data from ocean buoys. *Journal of Atmospheric and Oceanic Technology*, 36(5):903–920.

Wu, Y., Zhang, Z., Huang, Y., Jin, Q., Chen, X., and Chang, J. (2019). Evaluation of the gpm imerg v5 and trmm 3b42 v7 precipitation products in the yangtze river basin, china. *Water*, 11(7):1459.

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., and Lin, R. (2017). Reprocessed, bias-corrected cmorph global high-resolution precipitation estimates from 1998. *Journal of Hydrometeorology*, 18(6):1617–1641.

Xie, W., Yi, S., Leng, C., Xia, D., Li, M., Zhong, Z., and Ye, J. (2022). The evaluation of imerg and era5-land daily precipitation over china with considering the influence of gauge data bias. *Scientific Reports*, 12(1):1–21.

Xu, F., Guo, B., Ye, B., Ye, Q., Chen, H., Ju, X., Guo, J., and Wang, Z. (2019a). Systematical evaluation of gpm imerg and trmm 3b42v7 precipitation products in the huang-huai-hai plain, china. *Remote Sensing*, 11(6):697.

Xu, J., Ma, Z., Tang, G., Ji, Q., Min, X., Wan, W., and Shi, Z. (2019b). Quantitative evaluations and error source analysis of fengyun-2-based and gpm-based precipitation products over mainland china in summer, 2018. *Remote Sensing*, 11(24):2992.

Xu, R., Tian, F., Yang, L., Hu, H., Lu, H., and Hou, A. (2017). Ground validation of gpm imerg and trmm 3b42v7 rainfall products over southern tibetan plateau based on a high-density rain gauge network. *Journal of Geophysical Research: Atmospheres*, 122(2):910–924.

Xu, S., Shen, Y., and Niu, Z. (2019c). Evaluation of the imerg version 05b precipitation product and comparison with imerg version 04a over mainland china at hourly and daily scales. *Advances in Space Research*, 63(8):2387–2398.

Yang, D., Goodison, B. E., Metcalfe, J. R., Golubev, V. S., Bates, R., Pangburn, T., and Hanson, C. L. (1998). Accuracy of nws 8" standard nonrecording precipitation gauge: Results and application of wmo intercomparison. *Journal of atmospheric and oceanic technology*, 15(1):54–68.

Yang, M., Li, Z., Anjum, M. N., and Gao, Y. (2019). Performance evaluation of version 5 (v05) of integrated multi-satellite retrievals for global precipitation measurement (imerg) over the tianshan mountains of china. *Water*, 11(6):1139.

Yang, M., Liu, G., Chen, T., Chen, Y., and Xia, C. (2020a). Evaluation of gpm imerg precipitation products with the point rain gauge records over sichuan, china. *Atmospheric Research*, 246:105101.

Yang, S. and Smith, E. A. (2006). Mechanisms for diurnal variability of global tropical rainfall observed from trmm. *Journal of climate*, 19(20):5190–5226.

Yang, X., Lu, Y., Tan, M. L., Li, X., Wang, G., and He, R. (2020b). Nine-year systematic evaluation of the gpm and trmm precipitation products in the shuaishui river basin in east-central china. *Remote Sensing*, 12(6):1042.

Yu, C., Hu, D., Liu, M., Wang, S., and Di, Y. (2020). Spatio-temporal accuracy evaluation of three high-resolution satellite precipitation products in china area. *Atmospheric Research*, 241:104952.

Yuan, F., Zhang, L., Soe, K. M. W., Ren, L., Zhao, C., Zhu, Y., Jiang, S., and Liu, Y. (2019). Applications of trmm-and gpm-era multiple-satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in myanmar. *Remote Sensing*, 11(2):140.

Yuan, F., Zhang, L., Win, K. W. W., Ren, L., Zhao, C., Zhu, Y., Jiang, S., and Liu, Y. (2017). Assessment of gpm and trmm multi-satellite precipitation products in streamflow simulations in a data-sparse mountainous watershed in myanmar. *Remote Sensing*, 9(3):302.

Zhang, A., Xiao, L., Min, C., Chen, S., Kulie, M., Huang, C., and Liang, Z. (2019). Evaluation of latest gpm-era high-resolution satellite precipitation products during the may 2017 guangdong extreme rainfall event. *Atmospheric Research*, 216:76–85.

Zhang, J., Lin, L.-F., and Bras, R. L. (2018). Evaluation of the quality of precipitation products: A case study using wrf and imerg data over the central united states. *Journal of Hydrometeorology*, 19(12):2007–2020.

Zhang, L., Li, X., Cao, Y., Nan, Z., Wang, W., Ge, Y., Wang, P., and Yu, W. (2020a). Evaluation and integration of the top-down and bottom-up satellite precipitation products over mainland china. *Journal of Hydrology*, 581:124456.

Zhang, T., Yang, Y., Dong, Z., and Gui, S. (2021). A multiscale assessment of three satellite precipitation products (trmm, cmorph, and persiann) in the three gorges reservoir area in china. *Advances in Meteorology*, 2021:1–27.

Zhang, Y., Hanati, G., Danierhan, S., Liu, Q., and Xu, Z. (2020b). Evaluation and comparison of daily gpm/trmm precipitation products over the tianshan mountains in china. *Water*, 12(11):3088.

Zhao, H., Yang, S., You, S., Huang, Y., Wang, Q., and Zhou, Q. (2018). Comprehensive evaluation of two successive v3 and v4 imerg final run precipitation products over mainland china. *Remote Sensing*, 10(1):34.

Zhou, Z., Guo, B., Xing, W., Zhou, J., Xu, F., and Xu, Y. (2020). Comprehensive evaluation of latest gpm era imerg and gsmap precipitation products over mainland china. *Atmospheric Research*, 246:105132.

Zubieta, R., Getirana, A., Espinoza, J. C., Lavado-Casimiro, W., and Aragon, L. (2017). Hydrological modeling of the peruvian–ecuadorian amazon basin using gpm-imerg satellite-based precipitation dataset. *Hydrology and Earth System Sciences*, 21(7):3543–3555.