EVALUATING SPATIO-TEMPORAL VARIABILITY OF GRIDDED PRECIPITATION DATASETS AND THEIR EFFECT ON RIVER BASIN RESPONSES

Hamed HAFIZI

Dissertation

Department of Civil Engineering

Programme in Hydraulics

Supervisor: Assoc. Prof. Dr. Ali Arda ŞORMAN

Eskişehir

Eskişehir Technical University

Institute of Graduate Programs

April 2023

This thesis has been supported by the project 20DRP214, which is approved by the SRP Committee.

FINAL APPROVAL FOR THESIS

This thesis titled EVALUATING SPATIO-TEMPORAL VARIABILITY OF GRIDDED PRECIPITATION DATASETS AND THEIR EFFECT ON RIVER BASIN RESPONSES has been prepared and submitted by Hamed HAFIZI in partial fulfillment of the requirements in "Eskişehir Technical University Directive on Graduate Education and Examination" for the Degree of PhD in Civil Engineering Department has been examined and approved on 06/04/2023.

Committee Members		**Title, Name and Surname**	**Signature**
Member	:	Prof. Dr. Uğur AVDAN	
Member	:	Prof. Dr. İsmail YÜCEL	
Member	:	Assoc. Prof. Dr. Gökçen UYSAL	
Member	:	Assoc. Prof. Dr. Gordana KAPLAN	
Member	:	Assoc. Prof. Dr. M. Tuğrul YILMAZ	

Prof. Dr. Murat TANIŞLI

Director of the Institute of Graduate Programs

06/04/2023

SUPERVISOR APPROVAL

PhD student Hamed HAFIZI, whom I supervise, has completed this thesis titled EVALUATING SPATIO-TEMPORAL VARIABILITY OF GRIDDED PRECIPITATION DATASETS AND THEIR EFFECT ON RIVER BASIN RESPONSES. According to my inspections, the work is scientifically and ethically appropriate for the student to the thesis defense exam.

Supervisor

Assoc. Prof. Dr. Ali Arda ŞORMAN

ABSTRACT

EVALUATING SPATIO-TEMPORAL VARIABILITY OF GRIDDED PRECIPITATION DATASETS AND THEIR EFFECT ON RIVER BASIN RESPONSES

Hamed HAFIZI

Department of Civil Engineering

Programme in Hydraulics

Eskişehir Technical University, Institute of Graduate Programs, April 2023

Supervisor: Assoc. Prof. Dr. Ali Arda ŞORMAN

This study presents a detailed investigation into the spatio-temporal consistency and hydrological utility of 18 Gridded Precipitation Datasets (GPDs) (one ground-based, seven satellite-based, two reanalysis, one reanalysis and satellite, and seven multi-source) over varying topography and climate conditions of Turkey.

Considering the regional performance of GPDs by direct comparison with 757 ground stations for the entire study period (2015–2019), the consistency increases from daily to monthly time step, where daily KGE varies from 0.18 to 0.52 and monthly KGE ranges between 0.35 and 0.73. Regardless of any regional classification and time step, the tendency of GPD detection ability is high for no-precipitation events and the detection strength decreases as the precipitation intensity increases. Moreover, comparing the streamflow reproducibility of selected datasets over three snow-dominated basins, namely Karasu, Çukurkışla, and Kayabaşı, many GPDs are able to reproduce streamflow with higher accuracy when their optimum model parameter set is utilized.

Overall, multi-source GPDs generally perform better both in direct comparison over various classifications and hydrological utility with different model parameterization scenarios. Furthermore, some reanalysis datasets also present relatively high performance compared to other satellite-based, reanalysis or combined GPDs.

Keywords: Gridded Precipitation Datasets, Validation, Hydrological modeling, Varying climate-topography, Turkey.

ÖZET

HÜCRESEL YAĞIŞ VERİSETLERİNİN MEKANSAL-ZAMANSAL DEĞİŞKENLİĞİ VE BUNLARIN AKARSU HAVZALARINDAKİ UYGULANABİLİRLİĞİNİN İNCELENMESİ

Hamed HAFIZI

İnşaat Mühendisliği Anabilim Dalı

Hidrolik Bilim Dalı

Eskişehir Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Nisan 2023

Danışman: Doç. Dr. Ali Arda ŞORMAN

Çalışma, değişken topoğrafya ve iklim koşullarının hakim olduğu Türkiye'de, 18 hücresel yağış verisetinin (HYV) (bir yer, yedi uydu, iki yeniden analiz, bir uydu ve yeniden analiz, yedi çok kaynak temelli) detaylı şekilde mekansal-zamansal tutarlılığını ve hidrolojik açıdan uygulananbilirliğini içermektedir.

Çalışma süresince (2015–2019) 757 yer gözlem istasyonu değerlendirildiğinde HYV lerin bölgesel performansı günlük karşılaştırmalardan aylığa geçildikçe artmakta olduğu gözlenirken, günlük KGE başarı oranı 0.18 ile 0.52, aylık KGE başarı oranı ise 0.35 ile 0.73 arasında değişmektedir. Bölgesel ve zamansal sınıflandırma haricinde, HYV lerin yağış tespit etme becerisi yağış olmayan olaylar için yüksek iken yağış şiddeti arttıkça bu kabiliyetin azaldığı görülmüştür. HYV lerin akım benzeşim değerlendirmesi üç kar baskın çalışma havzası (Karasu, Çukurkışla, Kayabaşı) için karşılaştırıldığında, birçok HYV nin kendi model parametreleri kullanıldığında daha başarılı sonuçlar ürettiği söylenebilmektedir.

Sonuç olarak, çok kaynaklı ürünlerin genel itibariyle hem direkt yer gözlem kıyasıyla hem de hidrolojik model uygulamasında daha başarılı performanslarıyla öne çıkarken, bazı yeniden analiz ürünleri de yüksek performansıyla dikkat çekmekte ve diğer uydu bazlı verisetleri sonlarda yer almaktadır.

Anahtar Sözcükler: Hücresel Yağış Verisetleri, Doğrulama, Hidrolojik Modelleme, Değişken İklim-Topoğrafya, Türkiye.

ACKNOWLEDGEMENTS

This thesis would never have become a reality without the full cooperation of many individuals. I wish to extend my praise and thanks to all of them.

First and foremost, it is a great pleasure to express my gratitude and sense of thanks to my kind research supervisor, Assoc. Prof. Dr. Ali Arda Şorman for the practical guide and continuous support of my Ph.D. study, as well as his patience, motivation, and immense knowledge. His scholarly advice and scientific guidance carried me through all the stages of writing this thesis. I could not have imagined having a better advisor and mentor through my research and Ph.D. studies.

I also owe deep, sincere gratitude to the rest of my thesis committee members: Prof. Dr. Uğur Avdan, Prof. Dr. İsmail Yücel, Assoc. Prof. Dr. Gökçen Uysal, Assoc. Prof. Dr. Gordana Kaplan, and Assoc. Prof. Dr. M. Tuğrul Yılmaz for their prompt inspiration, valuable suggestions, and insightful comments. A special thanks to Prof. Dr. Aynur Şensoy Şorman for providing beneficial courses and taking an interest in my research.

I would like to acknowledge and give my special thanks to the Presidency for Turks Abroad and Related Communities (YTB) of Turkey and the Ministry of Higher Education of Afghanistan for awarding me the scholarship with funding to cover my stay in Turkey and support my Ph.D.

Finally, I am extremely grateful to all my family members: my wife and my daughter, for their unlimited love, care, sacrifices, understanding, and continuous support, without whom I would not have been able to accomplish this research. My warmest sincere thanks to my parents for their spiritual support and appreciation throughout my life. I would like to thank my in-laws for their multilateral support and encouragement along my Ph.D. journey.

I also place my appreciation on record for all who, directly or indirectly, have contributed to this research.

Hamed HAFIZI

06/04/2023

STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with "scientific plagiarism detection program" used by Eskişehir Technical University, and that "it does not have any plagiarism" whatsoever. I also declare that, if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.

Hamed HAFIZI

CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

GLOSSARY OF SYMBOLS AND ABBREVIATIONS

AEG	:	Aegean
ASCAT	:	Advanced SCATterometer
BLS	:	Black Sea
CAN	:	Central Anatolia
CCI	:	Climate Change Initiative
CCS	:	Cloud Classification System
CDR	:	Climate Data Record
CFS	:	Coupled Forecast System
CFSR	:	Climate Forecast System Reanalysis
CHELSA	:	Climatologies at high resolution for the Earth's land surface
CHIRP	:	Climate Hazards group InfraRed Precipitation
CHIRPS	:	Climate Hazards group InfraRed Precipitation with Stations
CHPClim	:	Climate Hazards group Precipitation climatology
CHRS	:	Center for Hydrometeorology and Remote Sensing
CMORPH	:	Climate Prediction Center morphing technique
CPC	:	Climate Prediction Center
CRU TS	:	Climatic Research Unit Time-Series
EAN	:	Eastern Anatolia
ECMWF	:	European Centre for Medium-Range Weather Forecasts
EOS	:	Earth Observing System
ERA5	:	ECMWF ReanAlysis fifth generation
EUMETSAT	:	European Organisation for the Exploitation of Meteorological Satellites
FGGE	:	First Global Atmospheric Research Program Global Experiment
G	:	Ground
GDAS	:	Global Data Assimilation System
GDM	:	General Directorate of Meteorology
GEFS	:	Global Ensemble Forecast System
GEOS5	:	Goddard Earth Observing System Model, version 5
GHCN-D	:	Global Historical Climatology Network-Daily
GPCC	:	Global Precipitation Climatology Centre
GPCP	:	Global Precipitation Climatology Project
GPD	:	Gridded Precipitation Datasets
GPM	:	Global Precipitation Measurement
GSI	:	Grid-point Statistical Interpolation
GSMap	:	Global Satellite Mapping of Precipitation
GSOD	:	Global Summary Of the Day
GTS	:	Global Telecommunication System
HBV	:	Hydrologiska Byråns Vattenbalansavdelning
HKS	:	Hanssen–Kuiper Skill
IFS	:	Integrated Forecasting System

IMERGHH	:	Integrated Multi-satellitE Retrievals for GPM Half Hourly
IMERGHHE	:	IMERGHH Early run
IMERGHHF	:	IMERGHH Final run
IMERGHHL	:	IMERGHH Late run
IR	:	Infrared
ISD	:	Integrated Surface Database
JAXA	:	Japan Aerospace Exploration Agency
JRA-55	:	Japanese 55-year Reanalysis
KGE	:	Kling-Gupta Efficiency
MAR	:	Marmara
MED	:	Mediterranean
MERRA-2	:	Modern-Era Retrospective Analysis for Research and Applications 2
MSWEP	:	Multi-Source Weighted-Ensemble Precipitation
MSWX	:	Multi-Source Weather
MW	:	MicroWave
NCEP	:	National Center for Environmental Prediction
NRT	:	Near-Real-Time
NSE	:	Nash-Sutcliffe Efficiency
PDF	:	Probability Density Function
PDIR	:	PERSIANN Dynamic Infrared-Rain Rate
PERSIANN	:	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
PET	:	Potential Evapotrnspiration
PMW	:	Passive Microwave
R	:	Reanalysis
RT	:	Real-Time
S	:	Satellite
SAN	:	Southeastern Anatolia
SM	:	Soil Moisture
SM2RAIN	:	Soil Moisture to Rain
TIR CCD	:	Thermal infrared cold cloud duration
TMPA	:	TRMM Multi-satellite Precipitation Analysis
TRMM	:	Tropical Rainfall Measuring Mission
TUW	:	Technical University of Wien
VIS	:	Visiable
VR	:	Variabiliy Ratio
WMO	:	World Meteorological Organization

1. INTRODUCTION

1.1. Precipitation: A key factor subject to uncertainty

Global water demand has been increasing at a rate of 1 percent per year over recent decades as a function of increasing water pollution, population growth, and economic development, and nearly 6 billion (66%) of the world's future population will experience water stress by 2050 (World Water Assessment Programme (United Nations), 2018). Moreover, the availability of water resources is directly related to the redistribution of seasonal precipitation and precipitation patterns under climate variability (Mengistu et al., 2021; Saeed et al., 2018). Hence, comprehensive insight on precipitation variation over time and space is one of the fundamental aspects of water resources management (Talchabhadel, Aryal, Kawaike, Yamanoi, & Nakagawa, 2021).

Measuring precipitation by installing a rain gauge above the ground level is a well-known technique that provides direct and physical precipitation observations at the point scale (Sevruk, 1987; T. Zhao & Yatagai, 2014). However, systematic errors such as wind-induced, evaporation, scarcity, and infrequent distribution of rain gauges, mainly over complex terrain, sometimes cause unreliable capture of precipitation's spatial variability (Sevruk et al., 2009; Z. Yu et al., 2019). Nevertheless, rain gauge data is universally accepted as one of the references for validating precipitation estimates from other sources (Tapiador et al., 2012). Similarly, ground radars can provide high spatio-temporal resolution precipitation estimates over relatively larger and remote areas, and the concept of using returned echo power from radar for precipitation measurement dates back to the mid-1940s (Tapiador et al., 2012; Yan & Bárdossy, 2019). Nonetheless, the interaction between radar with topography and atmosphere, such as the vertical motion of air, precipitation drift, beam blockage, ground clutter, the range effects, radar signal attenuation, and anomalous propagation, introduces uncertainty into the result (Nikahd et al., 2016). In recent decades, the rapid development of remote sensing technology and satellite-based precipitation retrieval algorithms provide many satellite-based precipitation datasets with high spatio-temporal resolution, larger and continuous coverage (Awange et al., 2016). Precipitation retrieval algorithms in general have relied mainly on data from visible (VIS), infrared (IR), and active and passive microwave (MW) satellite sensors. Furthermore, precipitation estimation using satellite sensors information can be categorized into the following three groups; (1) using visible and infrared sensors information, (2) utilizing microwave sensor data, and (3) combining visible, infrared, and microwave radiation (Boushaki et al., 2009). Overall, the geosynchronous-orbiting IR measurements have a higher spatial and temporal resolution, while the indirect relationship between cloud top temperature and precipitation rate results in a higher bias in the final product. Compared with IR, the polar-orbiting MW measurements use more direct information on the vertical distribution of hydrometeors, but their low spatial and temporal resolution limits their reliability (Joyce et al., 2004). Besides, precipitation from numerical weather prediction models could be another alternative approach that provides precipitation estimates and forecasts (Xu et al., 2022).

In recent years, massive efforts have been made to evaluate Gridded Precipitation Datasets (GPDs) and report the presence of bias and errors (Ghaedamini et al., 2021; J. Hsu et al., 2021; Karmouda et al., 2022; Lei et al., 2021; Temimi et al., 2022; L. Zhang et al., 2021). Hence, the accuracy assessment of GPDs is essential for both applications and algorithm improvement (Lei et al., 2022).

1.2. Gridded Precipitation Datasets (GPDs)

From a global perspective, a large number of GPDs with varying spatio-temporal resolution and coverage are available (Q. Sun et al., 2018). On the basis of the input and technique used to retrieve the precipitation amounts, three groups of GPDs can be classified: (1) taking advantage of information retrieved from gauge networks, including Global Precipitation Climatology Centre (GPCC) (Iizumi et al., 2017) and Climate Prediction Center (CPC) unified version 1 (Xie et al., 2007), (2) utilizing Infrared (IR) and/or Passive Microwave (PMW) satellite sensors information namely Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Sorooshian et al., 2000), and Integrated Multi-satellitE Retrievals for GPM (IMERG) late run version 6 (G. J. Huffman et al., 2020), (3) using reanalysis data from numerical weather prediction models—for instance, the National Center for Environmental Prediction (NCEP) climate forecast system reanalysis (CFSR) (Saha et al., 2010), and European Centre for Medium-range Weather Forecasts ReAnalysis Interim (ERA-Interim) (Dee et al., 2011). It is essential to keep in mind that most GPDs include data from all three inputs (multi-source) to get the best level of accuracy. As proof, Climate Hazards group InfraRed Precipitation with Stations (CHIRPS) (Funk et al., 2015) and Multi-Source Weighted-Ensemble Precipitation (MSWEP) (Beck et al., 2017) use satellite, reanalysis, and ground information in their retrospective algorithms.

It is important to note that GPDs can be classified as either gauge uncorrected, which their temporal variations depend fully on satellite and/or reanalysis data, or gauge corrected, which their temporal variations rely partially on ground observations. To illustrate, the Climate Hazards group InfraRed Precipitation (CHIRP) (Funk et al., 2015), PERSIANN Dynamic Infrared-Rain Rate (PDIR-Now) (P. Nguyen et al., 2020), and the fifth generation of European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA5) (Hersbach et al., 2020) use only satellite and/or reanalysis data where some other GPDs, namely the Multi-Source Weather version 100 (MSWXv100) (Beck et al., 2022), and Modern-Era Retrospective Analysis for Research and Applications 2 (MERRA-2) (Gelaro et al., 2017) use ground observations on the top of satellite and/or reanalysis data. Moreover, the spatial and temporal resolution/coverage of GPDs draw end-users' attention correspondingly. Suppose that Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Cloud Classification System (PERSIANN-CCS) (Hong et al., 2004), have the highest spatial resolution (0.04°) and provides hourly precipitation data, which is available from 2003 to near real-time with 60° N/S spatial coverage, while CPCv1 provides relatively lower spatial resolution (0.50°) and worldwide coverage, displaying daily precipitation from 1997 to the present. Furthermore, the timing (latency) of the dataset's public release is critical for many hydro-climatological research, like flood forecasts and early warnings. For instance, the Integrated Multi-satellitE Retrievals for GPM (IMERG) Half Hourly early run version 6 (IMERGHHEv06) (G. Huffman et al., 2015) is released after 4 hours of real-time, but its ultimate research product (IMERGHHFv06) (G. J. Huffman et al., 2020) is published after 3.5 months. Finally, it's more important to have an insight into gauge correction temporal and spatial resolutions. Generally, GPDs are corrected by gauge observation with daily, five-days, and monthly time steps having different spatial resolutions. To give an idea, the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) 3B42 version 7 (TMPA-3B42v7) and IMERGHHFv06 use monthly Global Precipitation Climatology Centre (GPCC) data with a 1° spatial resolution (G. J. Huffman et al., 2010, 2020) and PERSIANN Climate Data Record (PERSIANN-CDR) datasets employ monthly Global Precipitation Climatology Project (GPCP) dataset with a 2.5° spatial resolution (Ashouri et al., 2015). In the same way, CHIRPSv2.0 includes 5-days precipitation estimates from the Climate Hazards group Precipitation climatology (CHPclim) dataset and daily precipitation data from other national meteorological agencies and private streams (Funk et al., 2015). MSWEPv2.8 includes WorldClim 2 dataset having 1 km spatial resolution and utilizing monthly GPCC, Global Historical Climatology Network-Daily (GHCN-D), Summary of the Day (GSOD), and other gauge observations (Beck et al., 2017).

As already stated, GPDs originated from distinct sources vary in spatial and temporal resolution/coverage, and a specific gauge correction approach is applied to each corrected GPD individually. Therefore, it is essential that the accuracy of GPDs should be investigated with in situ observations before utilizing them for a particular application.

1.2.1. Evaluation of GPDs over time and space

Regardless of GPD source(s), spatio-temporal resolution/coverage, time lag (latency), and retrospective algorithms, the reliability of GPDs is more conclusive when a potentially larger number of GPDs are collectively compared by using in situ observations as a benchmark over a specific time and space. In recent decades, extensive research has been conducted on the inter-comparison and validation of GPDs through including gauge observations as a reference (Andermann et al., 2011; Ayoub et al., 2020; Centella-Artola et al., 2020; Darand & Khandu, 2020; de Leeuw et al., 2015; Dinku et al., 2010; Ghulami et al., 2017; Hosseini-Moghari et al., 2018; Joyce et al., 2004; Liu, Shangguan, et al., 2019; Prakash, 2019; Salih et al., 2022; Satgé et al., 2020; Xu et al., 2022; X. Yuan et al., 2021; L. Zhang et al., 2022; T. Zhao & Yatagai, 2014). However, most of these accuracy assessments have been completed before the release of new and updated versions of GPDs (e.g., MSWEPv2.8 and/or MSWXv100, released in 2021 and 2022). Certain number of them (Ayoub et al., 2020; Dee et al., 2011; Hosseini-Moghari et al., 2018; Liu, Shangguan, et al., 2019; L. Zhang et al., 2022) have utilized a limited number of GPDs for validation. Some cases (Ayoub et al., 2020; Darand & Khandu, 2020; Ghulami et al., 2017) have considered less number of rain gauge stations for GPDs performance validation. In a few cases (Dinku et al., 2010; Liu, Shangguan, et al., 2019; X. Yuan et al., 2021), only rain/no-rain conditions were considered for GPD detectability analysis, which has not given detailed detectability information through different precipitation intensities. In the same way, several authors reported the consistency of GPDs over varying topography and/or climate conditions (Beck, Pan, et al., 2019; Chiaravalloti et al., 2018; Derin et al., 2019; El Kenawy et al., 2015; G. Li et al., 2021; Lu et al., 2021; Milewski et al., 2015; Moazami et al., 2016; Sharifi et al., 2019;

Sunilkumar et al., 2019; L. Yu et al., 2020; Zambrano-Bigiarini et al., 2017). Nevertheless, certain number of these validation assessments focused on limited number of GPDs. Few cases either considered terrain complexity (i.e., the spatial and temporal variation of GPDs from flat or low-land to highly elevated regions are ignored) or climate variability individually, and the recently released GPDs have not been taken into account. Based on the aforementioned context, there is a well-known gap in the literature that might be exploited for a larger number and recently released GPDs spatio-temporal consistency analysis.

1.2.2. Hydrological Utility of GPDs

Hydrologic models are utilized to solve many practical problems related to water, including flood forecasting, early warning, droughts, water resources management, and development (H. H. Nguyen et al., 2019; Singh, 2018). Moreover, the accuracy of hydrological modeling for streamflow simulation relies on accurate meteorological forcing, especially precipitation, which is typically limited for many parts of the globe (Talchabhadel, Aryal, Kawaike, Yamanoi, Nakagawa, et al., 2021). On the other hand, many efforts have been made to employ GPDs as an alternative source for precipitation-runoff modeling and investigate their interactions with basin climatology and topography (Behrangi et al., 2011; Bitew & Gebremichael, 2011; Gebregiorgis et al., 2012; Gunathilake et al., 2020; S. Jiang et al., 2017; N. Li et al., 2017; X. H. Li et al., 2012; Omonge et al., 2022; Saddique et al., 2022; Satgé et al., 2021; F. Su et al., 2008; J. Su et al., 2021; R. Sun et al., 2016; Z. Wang et al., 2017; Yong et al., 2010; F. Yuan et al., 2018; Y. Zhang et al., 2019; C. Zhao et al., 2020; Zhu et al., 2018). Furthermore, for the hydrological utility, GPDs should be imposed through hydrological models as a meteorological forcing, where the well-known approaches are focused on the hydrologic model parameters calibration and validation by observed precipitation, then the observed precipitation is replaced by each GPD, keeping the optimal parameters and/or the model parameters calibration and validation by each GPD individually, though GPDs have their optimal set of parameters (E. Ahmed et al., 2020; Gunathilake et al., 2020; L. Jiang & Bauer-Gottwein, 2019; S. Jiang et al., 2012; Stisen & Sandholt, 2010; J. Su et al., 2021; R. Sun et al., 2016; Xue et al., 2013). However, the majority of these assessments were completed prior to the release of the most recent GPDs. Some authors (N. Li et al., 2017; X. H. Li et al., 2012; F. Su et al., 2008; Z. Wang et al., 2017; Y. Zhang et al., 2019) have utilized a limited number of GPDs and GPDs from the same platforms for validation. Some of them (Omonge et al., 2022; Satgé et al., 2021; F. Yuan et al., 2018) have not included the updated version of GPDs. Several authors (Behrangi et al., 2011; Gunathilake et al., 2020; S. Jiang et al., 2017; J. Su et al., 2021; C. Zhao et al., 2020; Zhu et al., 2018) have only considered satellite-based precipitation, whereas reanalysis datasets are not considered. Few cases used the second model parameter calibration approach or evaluated GPDs over the validation time period when the model parameters are calibrated by in situ observations. As a result, the aforementioned gap in literature regarding GPDs hydrological utilities might be considered for the recently released and larger number of GPDs in precipitation-runoff simulation analyses. Finally, it is important to note that the performance and consistency of GPDs evaluated in one region may not be applicable in another, and individual validation is required in order to address GPD reliability.

1.2.3. GPDs evaluation over Turkey

Turkey, with its diverse landscape and climate conditions, could be used as a case study to investigate GPDs spatio-temporal consistency across varying topography and climate conditions, as well as their hydrological utilities over highly elevated domains. Several studies have evaluated the spatio-temporal and hydrological utility of GPDs over Turkey and related basins (Aksu & Akgül, 2020; Amjad et al., 2020; Bıyık et al., 2009; Derin et al., 2016, 2019; Derin & Yilmaz, 2014; Irvem & Ozbuldu, 2019; Jaber, 2020; Karakoc & Patil, 2016; Saber & Yilmaz, 2018; Uysal, 2022; Uysal & Şorman, 2021; Yucel et al., 2015). Though, some limitations of these studies include a shorter temporal window (i.e., 1-2 days), a smaller number of gauge stations for validation, the evaluation of a few GPDs over limited regions, GPDs selected from the same platform, focusing only on regions with complex topography, and ignoring the hydrological utility of GPDs. To generalize GPDs reliability across varying topography and climate conditions of Turkey, a detailed investigation involving a larger number of recently released GPDs and the use of larger number of ground stations as a reference is required.

1.3. Objectives of the study

This study aims to evaluate the spatio-temporal consistency and hydrological utility of 18 recently released GPDs (CPCv1, MSWXv100, MSWEPv2.8, CHIRPSv2.0, CHIRPv2.0, ERA5, MERRA-2, CFSR, IMERGHHFv06, IMERGHHLv06, IMERGHHEv06, TMPA-3B42v7, TMPA-3B42RTv7, PERSIANN-CDR, PERSIANN-CCS, PERSIANN, PDIR-Now, SM2RAIN-ASCAT) over Turkey, which is characterized by varying topography and climate conditions. Thus, two main objectives are followed in this research:

1. Evaluating the spatio-temporal consistency of 18 GPDs in a daily and monthly time steps considering the seasonal variability of precipitation over the varying topography and climate conditions of Turkey by employing gauge station data as a benchmark (Meteorological Evaluation of GPDs);
2. Evaluating the potential and limitations of selected GPDs as a meteorological forcing for streamflow prediction under various model parameterization scenarios (Hydrological Evaluation of GPDs).

1.4. Innovation of the Study

To begin, at the regional scale, according to the author's knowledge, this would be the first time that 18 GPDs from various sources (satellite, reanalysis, and ground) are being evaluated over Turkey. Furthermore, at the global level, there are some studies that include a large number of GPDs and validate them considering the test region uniformly, while others take into account the topography and climate variability effects on a few GPDs performance, where some GPDs, such as MSWEPv2.8, MSWXv100, and PDIR-Now, are relatively new GPDs and they are not validated widely yet. Hence, in this study, the spatio-temporal consistency of larger GPDs is evaluated, and the validation is performed over variable topography and climate conditions rather than only taking the region into account uniformly. Furthermore, in order to have a better understanding of GPDs' interactions with the temporal distribution of precipitation, this study evaluates the seasonal performance of selected GPDs through discretized topography and climate conditions. Lastly, the most significant part of this study is the evaluation of 18 GPDs' potential for streamflow simulation under different model parameterization scenarios to obtain insight into the merit of these datasets.

1.5. Dissemination of the results

The peer-reviewed research articles and conference papers/abstract which present a variant of the study could be found at the end of this thesis.

2. METHODOLOGY

2.1. Study area

2.1.1. Topography-climate variability and hydro systems

Turkey is selected as a case study area, located between latitudes 36°N–42°N and longitudes 26°E–45°E at the crossroads of Europe, the Caucasus, the Middle East, and the eastern Mediterranean, with a total surface area of around 785350 km^2 (Figure 2.1a). The climate of Turkey is mainly driven by very sharp and parallel mountain ranges near the Black Sea and the Mediterranean Sea. Overall, the inner Anatolian plateau experiences limited precipitation, hot summers, and relatively cold winters, whereas coastal areas receive significant precipitation and have milder climates. For instance, the regions located on the Aegean and Mediterranean coasts receive annual precipitation ranging from 580 to 1300 mm. In the same way, the Black Sea region experiences a significant amount of precipitation, especially in Rize, which receives 2200 mm of precipitation per year and is the only region of the country that receives precipitation throughout the year. Due to mountain blockage, the central Anatolia region experiences the least annual precipitation of about 300 mm (Sensoy, 2004; Turkes, 1996). Moreover, the elevation rises from sea level to 5137 m from west to east, and the mean elevation of Turkey is around 1140 m. Considering the mountain ranges and the country's long borders (i.e., a total coastal length of 10765 km) with the Black Sea in the north, the Mediterranean Sea in the south and the Aegean Sea in the west, the climate in Turkey can be classified as semi-arid, the Mediterranean, continental, and sub-tropical climate predominate in various parts of the country (Derin & Yilmaz, 2014; Selek & Aksu, 2020). Furthermore, Turkey is primarily divided into 25 major river basins (i.e., Meriç, Marmara, Susurluk, North East Mediterranean, Gediz, Küçük Menderes, Büyük Menderes, West Mediterranean, Antalya, Burdur, Akarçay, Sakarya, West Black Sea, Yeşilırmak, Kızılırmak, Konya, East Mediterranean, Seyhan, Ceyhan, Asi, Euphrates-Tigris, East Black Sea, Çoruh, Aras, and Lake Van) where most rivers originate in Turkey and discharge into the Black Sea, Mediterranean Sea, Caspian Sea, Persian Gulf, and inland lakes within the country. The Euphrates-Tigris basin provides 28.4 percent of Turkey's water potential and is the largest basin in terms of both water potential and surface area (Frenken, 2009). However, a large number of water structures (i.e., 1000 small dam reservoirs and 293 large dams) have been built within mentioned basins and it is only possible to have a streamflow simulation in the headwaters of basins in Turkey.

The diverse climate and variable topography of Turkey provide many opportunities to evaluate the spatio-temporal consistency and hydrological utility of recently released GPDs.

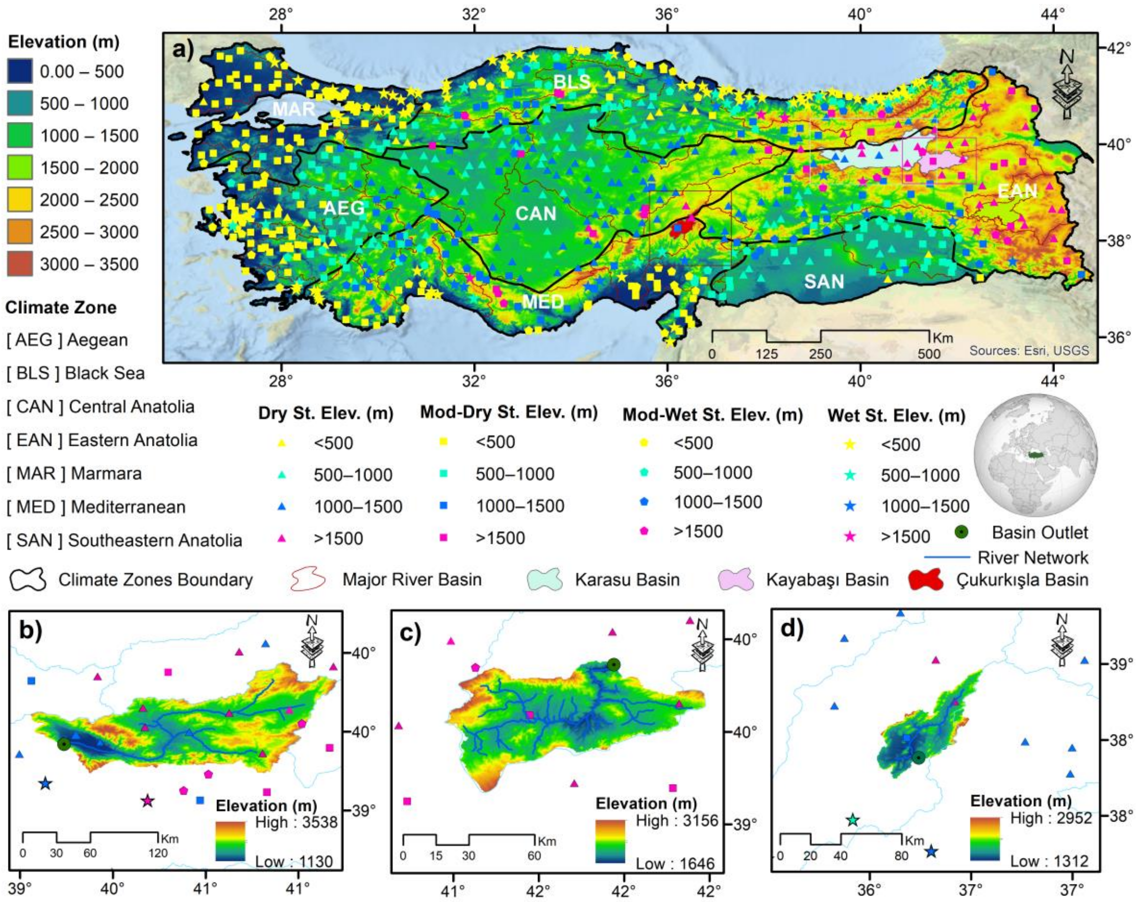

Figure 2.1. *a) Elevation map, geographical location of stations, and division of study area into seven climate zones, four wetness classes, and four elevation ranges. b) Geographical location of Karasu River Basin. c) Geographical location of Kayabaşı River Basin. d) Geographical location of Çukurkışla River Basin.*

2.1.2. On the selection of catchments for hydrological utility of GPDs

As it was mentioned earlier (section 2.1.1), most river basins in Turkey are drained out to the surrounding seas and neighboring countries, including Iran, Iraq, and Armenia. Despite the large spatial extent of these river basins, the cascade water structure systems along these rivers disrupt the precipitation-runoff modeling. Moreover, some basins (i.e., the Aegean catchments, namely Büyük and Küçük Meanders, and Gediz) have been almost completely transformed into regulated land. Hence, it is only possible to have a streamflow simulation in the headwaters and highly elevated parts of major basins on the outside of any hydraulic structure interruption. Subsequently, this study evaluates the hydrological utility of selected GPDs in the following three catchments, namely Karasu, Kayabaşı, and Çukurkışla river basins and having different drainage areas.

2.1.2.1. The upper Euphrates (Karasu) basin

The Karasu basin, which is located in eastern part of Turkey between latitudes 39°23′ to 40°25′N and longitudes 38°58′ to 41°39′E (Figure 2.1b). It has an area of around 10250 km^2 and elevations ranging from 1130 to 3538 m, with a hypsometric mean of 1972 m. Moreover, the basin is known as one of the major tributaries of the Euphrates River, which is recognized as the longest transboundary river in southwest Asia and the largest river basin (127300 km^2) with a 17% total water potential in Turkey. The basin is mainly covered by agricultural area (31.5%), bare land (27.5%), grass (35.0%), forest (3.5%), urban areas (1.84%), and wetland (0.66%). Karasu experiences precipitation in both rain and snow forms and snowmelt contributes to the flow (i.e., the snowmelt contribution is about 60–70% of the annual runoff volume) once the temperatures rise in the spring and early summer seasons. The discharge from Karasu basin is controlled by Kemah (E21A019) stream gauging station. The basin is one of the pilot catchments in Turkey which carries many national and international scientific studies (Akyurek et al., 2010; Montero et al., 2016; Parajka et al., 2019; Şensoy et al., 2023; A. A. Şorman et al., 2019; A. Ü. Şorman et al., 2007; A. U. Sorman & Beser, 2013; Tekeli et al., 2005; Uysal et al., 2023).

2.1.2.2. The upper Aras (Kayabaşı) basin

Kayabaşı basin is in the headwaters of Aras River (i.e., Aras basin, with a total area of 27548 km^2, originates from mountains exceeding around 3000 m above mean sea level in the west. It then joins the Arpaçay, a major tributary draining the northern basin, forming the Turkish-Armenian boundary, before entering Iran and discharging to the Caspian Sea in Azerbaijan.), situated in the eastern part of Turkey, within the latitudes of 39°18' to 39°52'N and longitudes of 41°10' to 41°55'E (Figure 2.1c). It has a drainage area of around 2730 km^2, and the D24A096 stream gauging station controls its outlet. The basin surface area is primarily covered by agricultural land (15.0%), bare land (30.8%), and grass (50.4%). The upper Aras basin, Kayabaşı is a relatively mountainous catchment where its elevation varies from 1646 m to 3156 m, with a hypsometric mean of 2218 m.

Precipitation usually occurs in the form of rain and snow, where snow plays a significant role in the hydrological cycle of the basin. Kayabaşı is an important area for hydrologic model application (Akbulut et al., 2022; A. A. Sorman et al., 2020).

2.1.2.3. The upper Seyhan (Çukurkışla) basin

The Çukurkışla basin is located in southern part of Turkey, between latitudes 38°01′ to 38°04′N and longitudes 36°02′ to 36°40′E (Figure 2.1d). It is one of the major tributaries of Seyhan River (i.e., the Seyhan River, with a 560 km long and 20450 km^2 basin area, called Sarus River in ancient times, originates in the Tahtalı Mountains. It flows through the Taurid Mountains, passes through the Çukurova coastal plain, and finally discharges to the Mediterranean Sea) along with Zamantı, Çakıt, and Körkün Rivers (Akbulut et al., 2022). It has a drainage area of around 1510 km^2, with elevations ranging from 1312 m to 2952 m, and has a hypsometric mean of 1710 m, and its outlet is controlled by Çukurkışla stream gauging station (E18A024). The basin is covered by wetland (0.10%), urban areas (0.33%), forest (5.09%), grass (23.86%), agricultural area (34.96%), and bare land (35.66%). The topographic complexity, high altitude, and snow-dominant characteristics of these three basins are advantageous in validating GPDs and operating hydrologic models for areas under such difficult conditions. The properties of selected basins are summarized in Table 2.1.

Table 2.1. *Properties of selected basins*

No.	Basin	Stream gauging station No.	Elevation (m)			Drainage area (km^2)
			Minimum	Maximum	Hypsometric mean	
1.	Karasu	E21A019	1130	3538	1972	10250
2.	Kayabaşı	D24A096	1646	3156	2218	2730
3.	Çukurkışla	E18A024	1312	2952	1710	1510

2.2. Datasets

In this context, 18 GPDs from various sources (i.e., one ground-based, seven satellite-based, two reanalysis, one reanalysis and satellite, and seven multi-sources) are being evaluated. Each GDP has its own set of characteristics and could be differentiated based on the information level utilized for precipitation estimates and the output features like latency (lag time) and spatio-temporal resolution/coverage (Table 2.2). Furthermore, a large number of ground stations (757 in total) and stream gauging stations (3 in total) have been selected carefully after quality control, providing the opportunity to evaluate the consistency and hydrological utility of GPDs. The properties of ground stations (which are accepted as reference data in this study), stream gauging stations, and selected GPDs are described in detail below:

2.2.1. Observed precipitation and streamflow data

It is widely accepted that gauge-based precipitation data could be used as a reference for GPDs validation. Similarly, employing streamflow data for GPDs hydro-validation is the most common procedure in the literature. In this study, daily precipitation data from 1870 stations distributed all over Turkey were provided by the General Directorate of Meteorology (GDM). Similarly, daily quality-controlled streamflow data at the outlet of selected basins were obtained from the General Directorate of Hydraulic Works (GDHW) of Turkey. Moreover, dependent stations (i.e., stations whose data are shared and evaluated with worldwide research centers for bias correction of gauge corrected GPDs), missing values, outliers (which are not recorded by GDM), continuous zero value for extended periods (during rainy days), minimum cumulative value, data entry repetition, were among the anomalies found in the collected data. As a result, the daily precipitation data for the period of five water years (October 2014–September 2019) was subjected to certain quality control procedures to eliminate outliers or any observable discrepancies in the raw meteorological data. The details of quality control and data cleaning will be discussed in section 2.3.1.

2.2.2. CPCv1 dataset

The Climate Prediction Center (CPC) unified gauge-based daily precipitation dataset was developed by the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC). The project's main objective is to create a unified gridded precipitation with reliable quantity and enhanced quality by merging all information sources existing at CPC and by taking advantage of the optimal interpolation (OI) objective analysis technique. Gauge-based information from over 30,000 weather stations is collected, and the quality control is performed by considering historical records and information from surrounding stations, radar, and satellite observation. The data from the quality-controlled station records have been interpolated to produce a ground-based gridded precipitation dataset with a 0.5-degree spatial resolution from 1979 to the present. This dataset has two components: (a) the retrospective version, which utilizes 30K stations and spans 1979-2005; and (b) the real-time version, which uses 17K stations and spans 2006-present. The real-time version will be reprocessed in the future to be consistent with the retrospective analysis (Xie et al., 2007). The data is available at (http-1).

2.2.3. MSWXv100 dataset

The Multi-Source Weather (MSWX) version 100 is an operational global near-surface gridded meteorological dataset with high spatial (0.1°) and temporal (3-hourly) resolutions and a short time lag (3 hours of latency). Moreover, MSWXv100 provides medium-range (up to ten days) and long-range (up to seven months) bias-corrected forecast ensembles. It takes advantage of different datasets such as CHELSA, CRU TS, ERA5, FLUXNET, GDAS, GEFS, GHCN-D, GSOD, and ISD. The dataset contains data records for ten meteorological variables (precipitation, air temperature, daily minimum and maximum air temperature, surface pressure, relative and specific humidity, wind speed, and downward shortwave and longwave radiation), and its historical records are available from January 1979 to the present (Beck et al., 2022). MSWXv100 provides a wide range of applications, including hydrological modeling, water resources management, flood forecasting, drought monitoring, and disease tracking at the global/regional scale. The MSWXv100 data is available at (http-2).

2.2.4. MSWEPv2.8 dataset

The Multi-Source Weighted-Ensemble Precipitation, version 2.8 (MSWEPv2.8), is a fully global historical precipitation dataset (1979–present) with 3-hourly temporal and 0.1° spatial resolution. MSWEPv2.8 utilizes spatial information provided by gauges, including WorldClim, Global Historical Climatology Network-Daily (GHCN-D), Global Summary of the Day (GSOD), and Global Precipitation Climatology Centre (GPCC). Similarly, it includes satellite-based information, including Climate Prediction Center morphing technique (CMORPH), Gridded Satellite (GridSat), Global Satellite Mapping of Precipitation (GSMaP), and Tropical Rainfall Measuring Mission (TRMM) Multi satellite Precipitation Analysis (TMPA-3B42RT). It also includes reanalysis of precipitation, namely, European Centre for Medium-Range Weather Forecasts

(ECMWF) interim reanalysis (ERA-Interim) and Japanese 55-year Reanalysis (JRA-55). MSWEPv2.8 provides unique opportunities to explore spatio-temporal variations in precipitation, enhance our understanding of hydrological processes and their parameterization, and improve the performance of hydrological models. (Beck, Wood, et al., 2019). The MSWEPv2.8 data is available at (http-3).

2.2.5. CHIRP(S)v2.0 datasets

The Climate Hazards group Infrared Precipitation Version 2.0 (CHIRPv2.0) displays precipitation at daily, 5-days, and monthly temporal resolutions with a spatial resolution of 0.05°, quasi-global coverage (50 N/S), and a complete data record from 1981 to the present. CHIRPv2.0 primarily relies on the thermal infrared cold cloud duration (TIR-CCD) to produce precipitation estimates. At first, a fixed temperature (235°K) threshold for precipitation and no-precipitation events was considered. Subsequently, local linear regression relationships were derived from both 5-days TMPA-3B42v7 (0.25°) and 5-days CCD data for each month. Later on, the regression parameters were resampled into a 0.05° × 0.05° grid, and then a 5-days precipitation dataset was produced. The 5-days precipitation estimates for each grid were divided by that grid cell's long-term (1981–2013) mean values and multiplied with the corresponding CHPclim value to generate the final 5-days CHIRPv2.0 dataset. The primary computing time step for the CHIRPv2.0 is the pentad. All other time steps are either aggregated (dekadal and monthly) or disaggregated (daily). d-days CHIRPv2.0 values are disaggregated into daily precipitation estimates based on daily Coupled Forecast System version 2 (CFSv2) fields rescaled to 0.05° resolution. At each pixel, the CHIRPv2.0 pentad total is redistributed in proportion to the daily values of the CFSv2. CHIRPv2.0 is available at (http-4).

The Climate Hazards group Infrared Precipitation with stations version 2.0 (CHIRPSv2.0) dataset with a spatial resolution of 0.05° was initially designed to support the United States Agency for International Development Famine Early Warning System Network (FEWS NET) for drought monitoring at large scales. CHIRPSv2.0 employs several ground-based gauge networks, including Global Historical Climate Network (GHCN) daily, CHEN monthly, Global Summary of the Day (GSOD), World Meteorological Organization's Global Telecommunication System (GTS), Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL), and several public data streams. CHIRPSv2.0 utilizes the modified inverse distance weighting algorithm for station blending, and it is available with a lag of 1–3 weeks following the generation of CHIRPv2.0 dataset (Funk et al., 2015). CHIRPSv2.0 is available at the Climate Hazards Group's official website (http-5).

2.2.6. MERRA-2 dataset

The Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2) is a reanalysis-based dataset derived from NASA's Goddard Earth Observing System Model, version 5 (GEOS5) and its data assimilation system. MERRA-2 replaces the original MERRA dataset created by NASA utilizing a fixed assimilation system, the GEOS5, to incorporate NASA's Earth Observing System (EOS) observations into a consistent representation of mass and energy flows from the earth's surface to the upper atmosphere (Rienecker et al., 2011). In preparing MERRA-2 dataset, its atmospheric assimilation system was updated to assimilate newer satellite observations. A newer GEOS5 system and a Grid-point Statistical Interpolation (GSI) analysis scheme were also employed (Bosilovich et al., 2015). Moreover, the simulated precipitation is bias-corrected, utilizing observations from satellites and gauges. As a consequence, MERRA-2 precipitation dataset has a lower bias and is more highly correlated to reference data than its earlier version. The dataset has a temporal resolution of 1 hour and spatial resolution of around 0.5° (Gelaro et al., 2017; Reichle, Liu, et al., 2017). Furthermore, MERRA-2 precipitation was corrected within the coupled atmosphere-land modeling system, allowing the near-surface air temperature and humidity to respond to the improved precipitation data (Reichle, Draper, et al., 2017). The MERRA-2 datasets are can be downloaded from their website (http-6).

2.2.7. ERA5 dataset

The European Centre for Medium Range Weather Forecasts (ECMWF) reanalysis fifth generation (ERA5) is the latest reanalysis dataset, which provides a variety of atmospheric and land-surface variables and follows the predecessors, including the First Global Atmospheric Research Program Global Experiment (FGGE), ERA-15, ERA-40, and ERA-Interim. ERA5 employs the 4-Dimensional Variational (4D-var) data assimilation technique in the Integrated Forecasting System (IFS) version Cy41r2 to assimilate large amounts of ground-based observations, remote sensing data, and atmospheric sounding data and delivers reasonable spatio-temporal variability of precipitation at a large scale and especially demonstrates high potential in areas where dense ground data have been assimilated. The ERA5 presents precipitation with a temporal resolution of 1 hour and a spatial resolution of 0.25° and includes detailed meteorological reanalysis data from 1950 onwards globally (Hersbach et al., 2020). The ERA5 data can be downloaded from the following website (http-7).

2.2.8. CFSR dataset

The Climate Forecast System Reanalysis (CFSR) dataset was developed by the United States (U.S.) National Centers for Environmental Prediction (NCEP). CFSR is a global, high-resolution, coupled atmosphere–ocean–land surface–sea ice system that is designed to deliver the most accurate estimate of the coupled domains' states. The horizontal native spatial resolution of CFSR global atmosphere data is approximately 38 km, interpolated to a 0.5° latitude and 0.5° longitude grid. It was originally provided with a temporal resolution of 6 hours, but its hourly and daily (derived from 6-hourly) data are also available (Saha et al., 2010). CFSR data have a temporal coverage from 1979 to the present and can be downloaded from the following website (http-8).

2.2.9. SM2RAIN-ASCAT dataset

SM2RAIN-ASCAT is a quasi-global (only over land) precipitation dataset utilizing Advanced SCATterometer (ASCAT) satellite soil moisture data through the SM2RAIN algorithm, and it is the successor of SM2RAIN-Climate Change Initiative (SM2RAIN-CCI). The soil moisture data through a depth of 2 cm of soil is obtained from the ASCAT radar system operating in the C-band (5.255 GHz) on board three Meteorological Operational (MetOp) satellites, launched each with a 6-year lag (i.e., 2006, 2012, and 2018), as part of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Polar System programme. SM2RAIN-ASCAT presents precipitation with a temporal resolution of 1 day and a spatial resolution of 12.5 km, which is available since 2007 (Brocca et al., 2019). The dataset is can be downloaded from the following website (http-9).

2.2.10. TMPA datasets

The Tropical Rainfall Measuring Mission (TRMM) was launched by the National Aeronautics and Space Administration (NASA) in cooperation with the Japan Aerospace

Exploration Agency (JAXA) on November 1997 (X. Wang et al., 2019). The TRMM Multisatellite Precipitation Analysis (TMPA) uses microwave and infrared information from multiple satellites to deliver the best precipitation estimates. It has spatial coverage of 50°N/S and a spatial resolution of 0.25° × 0.25°. TMPA is available both after (TMPA-3B42v7) and in real-time (TMPA-3B42RTv7) and only post-research product incorporates gauge data at present. For both post and real-time products, the daily (accumulated from 3-hourly) data is available from 1998 to the present (G. J. Huffman et al., 2007). The data for both TMPA-3B42v7 and TMPA-3B42RTv7 were downloaded from the TRMM website (http-10).

2.2.11. GPMIMERG datasets

The Global Precipitation Measurement (GPM) core observatory was successfully deployed by NASA and JAXA in February 2014. GPM mission is the successor of TRMM, provides the new generation of precipitation datasets at a spatial resolution of 0.1° × 0.1° with a 30 min temporal resolution. Furthermore, GPM has a larger spatial coverage (60°N/S) compared to TRMM (50°N/S). The recently released Integrated Multi-satellitE Retrievals for GPM Half Hourly (IMERGHH) V06 precipitation dataset retrospectively treated the TRMM precipitation by employing a consistent GPM-TRMM calibration algorithm, resulting in a long series GPM precipitation dataset with a global coverage since June 2000. The IMERGHHv06 consists of two types of near-real-time datasets (IMERGHH Early run and IMERGHH Late run with a latency of ~4 h and ~14 h) and one post-real-time dataset (IMERGHH Final run with a latency of ~3.5 months). The late run (IMERGHHLv06) and final run (IMERGHHFv06) include more passive microwave (PMW) estimates than the early run (IMERGHHEv06), and both the late run and the final run employ forward-and backward-morphing, while the early run only morphs forward. Lastly, the final run uses the monthly GPCC dataset to correct the systematic bias (G. Huffman et al., 2015; X. Wang et al., 2019). Three types of IMERGHHv06 datasets were downloaded from the GPM website (http-11).

2.2.12. PERSIANN datasets

Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) utilizes longwave Infrared (10.2-11.2 μm) information from geostationary satellites and the model parameters are updated by passive microwave

(PMW) sensors data from low-orbit satellites to deliver the best global (60°N/S) precipitation estimates (K. L. Hsu et al., 1997). Unlike PERSIANN, which employs the traditional constant temperature threshold method, PRESIAN Cloud Classification System (PERSIANN-CCS) includes a variable temperature threshold cloud segmentation algorithm that enables the categorization of cloud-patch features based on texture, dynamic evolution, geometry, and cloud top height from satellite images (Hong et al., 2004). The PERSIANN Climate Data Record (PERSIANN-CDR) take advantages of PERSIANN algorithm by using Gridded Satellite (GridSat-B1) infrared (IR) data. The algorithm first calibrates by using NCEP stage IV hourly precipitation data then the calibrated model is run with GridSat-B1 infrared (IR) data model then the GPCPv2.2 monthly data is utilized for the systematic bias correction (Ashouri et al., 2015). Finally, PERSIANN Dynamic Infrared Rain Rate near real-time (PDIR-Now) utilizing PERSIANN-CCS algorithm uses a higher temperature threshold to capture warm precipitation and includes improvements to the method of cloud segmentation that result from the use of monthly cloud-type sets to expand the cloud classification system (P. Nguyen et al., 2020). Information on the spatio-temporal coverage/resolution of PERSIANN datasets could be found in Table 2.2. The datasets were downloaded from CHRS portal (http-12).

2.3. Selection of temporal window for the study

Table 2.2 summarizes the overall properties of ground data and 18 selected GPDs, as well as their temporal coverage, where each GPD has its own set of temporal coverage. On the other hand, a large number of stations have been operating over Turkey since 2014, which are being used as a reference for spatio-temporal consistency analysis of GPDs in this study. As a result, the gauge data was compiled between October 2014 and September 2019 (i.e., the validation study of GPDs started in 2019). Finally, gauge-based precipitation time length (five water years) is applied as a reference for the temporal selection of GPDs.

2.4. Pre-processing of data

2.4.1. Quality control of observed data

As mentioned earlier (section 2.2.1), observed daily precipitation data from 1870 stations was initially provided by the General Directorate of Meteorology (GDM) of

Table 2.2. *Properties of selected 18 GPDs. Abbreviations in the data source(s) column represent; G, gauge; S, satellite; R, reanalysis; and A, analysis. In the spatial coverage column, "global" indicates fully global coverage including ocean areas. Abbreviation in the temporal coverage represent; NRT, Near Real-Time.*

Dataset	**Source (s)**	**Resolution**		**Coverage**		**Reference**
		Spatial	**Temporal**	**Spatial**	**Temporal**	
Observed Data	Ground	-	Daily	-	2015–2019	GDM, Turkey
CPCv1	G	0.5°	Daily	Global	1979–NRT[2]	(Xie et al., 2007)
MSWXv100	S, R, G	0.1°	3-hourly	Global	1979–NRT[1]	(Beck et al., 2022)
MSWEPv2.8	S, R, G	0.1°	3-hourly	Global	1979–NRT[3]	(Beck, Wood, et al., 2019)
CHIRPSv2.0	S, R, G	0.05°	Daily	50°N/S[L]	1981–NRT[2]	(Funk et al., 2015)
CHIRPv2.0	S, R	0.05°	Daily	50°N/S[L]	1981–NRT[2]	(Funk et al., 2015)
MERRA-2	S, R, G	0.5°	Hourly	Global	1980–NRT[3]	(Reichle, Liu, et al., 2017)
ERA5	R	0.25°	Hourly	50°N/S	1979–NRT[3]	(Hersbach et al., 2020)
CFSR	R	0.5°	Hourly	Global	1979–NRT[2]	(Saha et al., 2010)
SM2RAIN-ASCAT	S	~0.1°	Daily	Global	2007–NRT[3]	(Brocca et al., 2019)
TMPA-3B42v7	S, G	0.25°	3-hourly	50°N/S	2000–NRT[3]	(G. J. Huffman et al., 2007)
TMPA-3B42RTv7	S	0.25°	3-hourly	50°N/S	1998–NRT[2]	(G. J. Huffman et al., 2010)
IMERGHHFv06	S, G	0.1°	30 min	60°N/S	2014–NRT[3]	(G. J. Huffman et al., 2020)
IMERGHHEv06	S	0.1°	30 min	60°N/S	2014–NRT[1]	(G. J. Huffman et al., 2020)
IMERGHHLv06	S	0.1°	30 min	60°N/S	2014–NRT[1]	(G. J. Huffman et al., 2020)
PERSIANN-CDR	S, G	0.25°	Daily	60°N/S	1983–NRT[3]	(Ashouri et al., 2015)
PERSIANN-CCS	S	0.04°	Hourly	60°N/S	2003–NRT[1]	(Hong et al., 2004)
PERSIANN	S	0.25°	Hourly	60°N/S	2000–NRT[2]	(K. L. Hsu et al., 1997)
PDIR-Now	S	0.04°	Hourly	60°N/S	2000–NRT[1]	(P. Nguyen et al., 2020)

[1] Available until the present with a delay of several hours. [2] Available until the present with a delay of several days. [3] Available until the present with a delay of several months. [L] Indicates that the coverage is limited to the Land.

Turkey. Precipitation data from various time periods (for example, 2010–2019 and 2014–2019) was trimmed to the required five water years (October 2014–September 2019). Furthermore, a certain quality control approach is imposed on the data to remove outliers or any other observable discrepancies, which is explained in the following paragraphs.

To eliminate outliers from daily precipitation data, the data is subjected to a maximum precipitation threshold filter of 490 mm/day (i.e., the threshold was applied based on GDM report on extreme daily precipitation over Turkey). Moreover, the stations were checked for continuous missing values by setting a threshold filter of three months (i.e., 90 days of continuous missing data) within the study period, and stations with consecutive missing values greater than the assigned threshold were removed. Later on, the stations which could not pass threshold filters either through minimum annual cumulative precipitation (150 mm/year in 2017 and 200 mm/year for the rest of the study period) or the intra-annual cumulative precipitation (i.e., 100 mm of cumulative precipitation between October and March of each year during the study period) are excluded. Following that, the stations whose data were shared internationally and evaluated by global research centers for bias correction of gauge corrected GPDs were also removed. It is worth mentioning that the threshold filters on the raw daily precipitation data were applied based on information collected from GDM of Turkey.

Finally, after the deployment of all quality control steps and a visual inspection, the daily precipitation data for five consecutive water years (October 2014–September 2019) from 757 stations which is distributed over all Turkey were selected to evaluate the spatio-temporal consistency and hydrological utility of GPDs.

It's important to keep in mind that the streamflow data at the outlet of three basins was provided by the General Directorate of Hydraulic Works (GDHW) of Turkey for the study period. The temperature data which was provided by GDM and discharge data are already quality controlled and will be employed for hydrologic model parameters calibration by introducing GPDs and observed precipitation as meteorological forcing and evaluating GPDs simulate streamflow.

2.4.2. Meteorological stations classification

The diverse landscape and variable climate of the study area (section 2.1.1) provides the opportunity for a detailed evaluation of GPDs' performance through varying topography and climate conditions. As a result, rather than assuming the region uniformly (i.e., the long-term annual mean precipitation standard deviation reaches 285.9 mm and the standard deviation of elevation between coastal areas and high-altitude regions in the eastern part of the country exceeds by more than 5%), which limits evaluation of GPDs' generalization, the area is discretized based on its topographic elevation and climate condition. Hence, the stations that passed the quality control steps (757 in total) are classified based on terrain elevation, climate zone, and wetness to investigate GPDs reliability over varying topography (elevation ranges) and climate conditions (climate zone, and wetness classes).

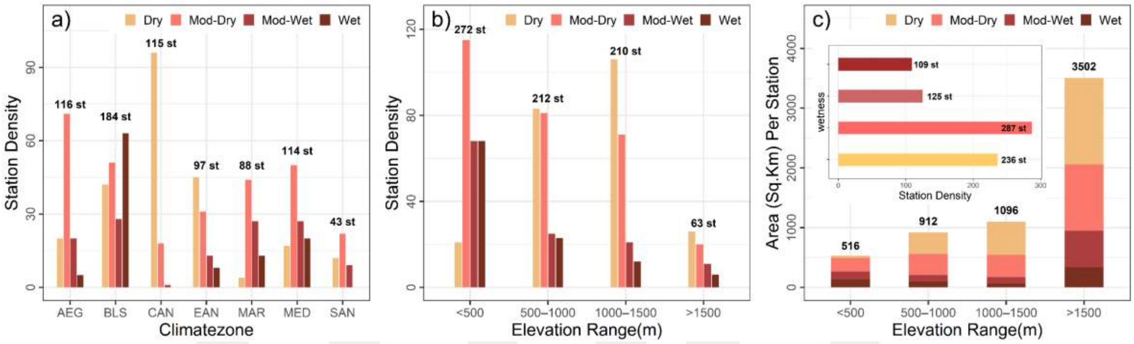

Figure 2.2. *a) Stations' classification and density over seven climate zones ordered by wetness classes; b) stations' density over four elevation classes considering four wetness classes over each elevation range; and c) elevation-area of ground stations and stations.*

Firstly, the stations are grouped into seven zones (Figure 2.2a) based on their location in various climate zones (i.e., Aegean, Black Sea, Eastern Anatolia, Central Anatolia, Marmara, Mediterranean, and Southeastern Anatolia), which are provided by GDM of Turkey.

Secondly, for the wetness-based station classification, stations are mandatory divided into four categories (Figure 2.2): dry, moderately-dry, moderately-wet, and wet, which are determined by assigning the mandatory boundary of mean monthly precipitation (mm) of ground data (40, 40–60, 60–80, and > 80). Finally, based on stations' altitude from mean sea level, they are classified by specifying randomly selected threshold intervals into four distinct elevation ranges, including less than 500m, 500–1000m, 1000–1500m, and more than1500m (Figure 2.2b). Figure 2.1a maps the spatial distribution of 757 meteorological stations over different elevation ranges, climate zones, and wetness classes. Furthermore, the inter-relations and density of stations over each classification are summarized in Figure 2.2.

2.4.3. Distribution of meteorological stations

Figure 2.2 presents the distribution of ground stations over different classifications. The majority of stations are concentrated over moderately dry (Figure 2.2c) and low-elevation lands, primarily in the Aegean (AEG), Black Sea (BLS), Central Anatolia (CAN), and Mediterranean (MED) regions, where station density decreases as the elevation increases (Figure 2.2b). On the other hand, according to the World Meteorological Organization's minimum required station density standards for different regions (WMO, 2008), the average gauge density of 1 gauge is 900 km^2 over coastal areas and 1 gauge per 250 km^2 over mountainous regions. Unlike the WMO guidelines, ground stations are disseminated in the other way (Figure 2.2c). For instance, stations that are located in areas more than 1500 m above mean sea level each cover a 3502 km^2 area (i.e., the WMO standard suggests 250 km^2 per station over mountainous areas), whereas in the coastal and low land areas (areas with elevation of less than 500 m), ground stations are extensively dense (1 gauge per 516 km^2) and they are more numerous than the WMO standard (1 gauge per 900 km^2 area). Furthermore, insufficient ground-based stations directly affect areal-average precipitation estimation. As a result, we selected point-based precipitation evaluation of GPDs over the study area.

2.4.4. Selection of Point-Grid evaluation

The comparison of GPDs against ground-based data could be performed either grid-to-grid or point-to-grid. The ground-based gridded dataset is obtained from rain gauge records by interpolation techniques, and then all GPDs, including the ground-based gridded dataset, are linearly interpolated to a coarser spatial resolution (i.e., 0.1 degree) and give the opportunity to compare GPDs values at each grid box, respectively. For the point-to-grid comparison, the values of each grid box, which station is located within, will be extracted by linear interpolation, and then the extracted values will be compared with the ground observations (Chua et al., 2020).

However, the quality of ground-based gridded precipitation dataset highly depends on the ground station density (Ly et al., 2013). In our case, there are insufficient number of stations, especially over high-altitude regions. Hence, the second methodology (point-to-grid) is carefully employed for the analysis of GPDs in this study. It's worth mentioning that in this study, the R platform is used for all statistical work and data visualization.

2.5. Hydrologic model and parametrization scenarios

The TUW (Technical University of Wien) conceptual hydrologic model was developed by the Technical University of Vienna, tested over numerous Austrian catchments (Neri et al., 2020; Parajka et al., 2007). It is built on a structure similar to HBV (Hydrologiska Byråns Vattenbalansavdelning) (Bergstrom, 1995; Bergström & Lindström, 2015), operates on a daily time step. The TUW model is able to simulate runoff, snow, and soil moisture (Figure 2.3) with 15 parameters (Table 2.2), and the inputs are total daily precipitation (mm), mean daily air temperature (°C), and daily potential evapotranspiration (mm). Moreover, the snow routing is based on a simple degree-day concept, and it is governed by five parameters: two threshold temperature (T_R, and T_S) parameters that identify rain and snow, a melting temperature (T_M), a snow correction factor (SCF) and the degree-day factor (DDF). Furthermore, the soil moisture routing depicts soil moisture state changes and runoff generation, and includes three parameters: the field capacity (FC) which represent the maximum soil moisture storage, a parameter representing the soil moisture condition above which evapotranspiration is at its potential rate (LP), and a parameter β controlling the non-linear function of runoff generation. Finally, an upper and lower soil reservoirs and a triangular transfer function compose the runoff response and routing model which includes seven parameters. The sum of excess rainfall and snowmelt first enters the upper zone reservoir then leaves this reservoir through three paths: (a) outflow from the reservoir based on a fast storage coefficient (k_1); (b) percolation to the lower zone with a constant percolation rate (C_{PERC}) (c) if a threshold of the upper storage state (L_{UZ}) is exceeded, through an additional outlet based on a very fast storage coefficient (k_0). Water leaves the lower zone based on a slow storage coefficient (k_2). The outflows from both reservoirs are then routed by a triangular transfer function representing runoff routing in the streams, where the base of the transfer function (BQ), is estimated with the scaling of the outflow by the free scaling parameter (C_{ROUTE}) and maximum base at low flows parameter (B_{MAX}) (Ceola et al., 2015; Parajka et al., 2007).

In general, two types of precipitation-runoff scenarios are extensively employed for the hydrological utility of GPDs in the literature. Depending on the level of information

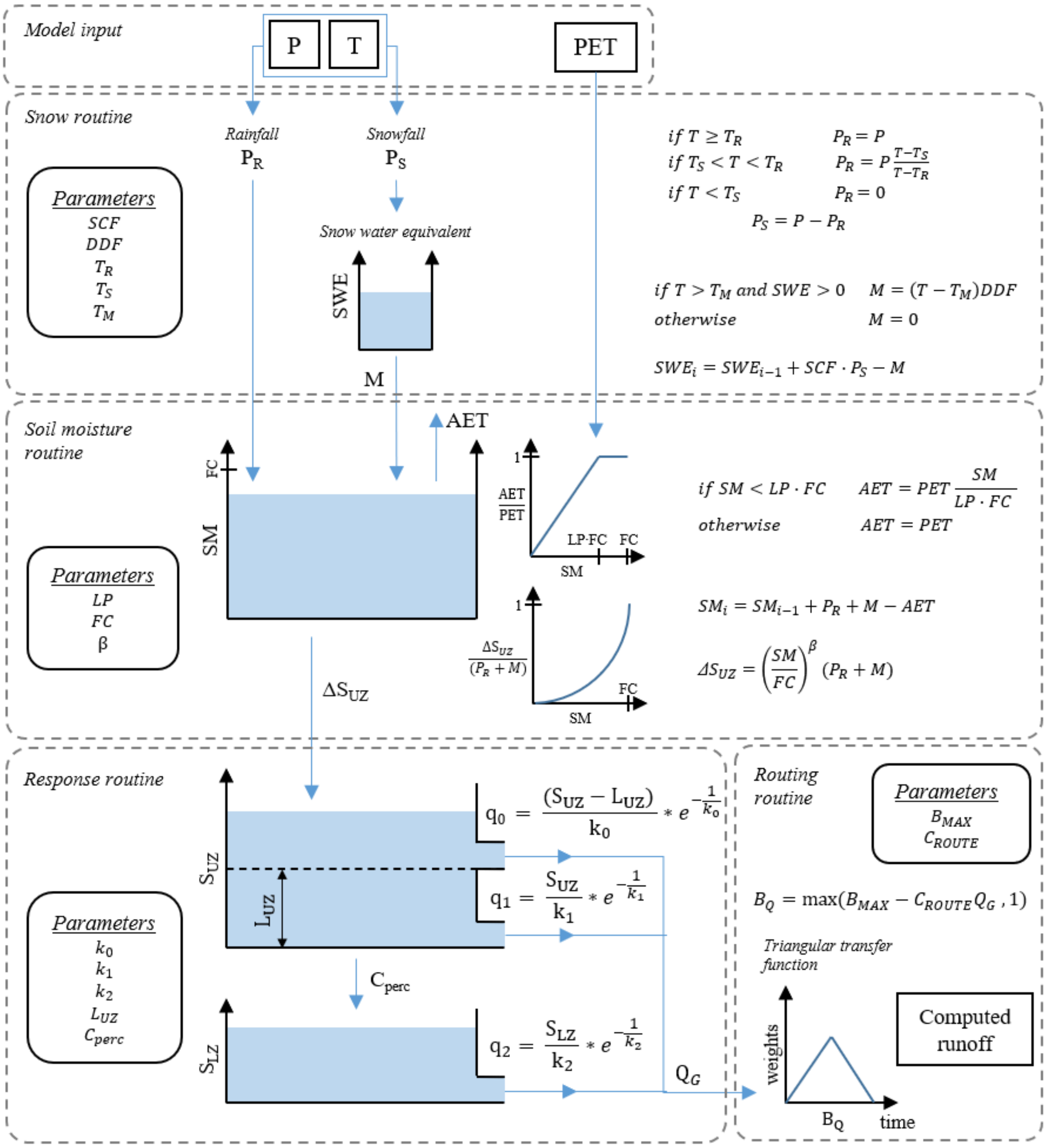

Figure 2.3. *The schematic lumped version of TUW model maps snow routing, soil moisture routing, and runoff routing (Neri et al., 2020).*

provided, (a) model parameters are calibrated/validated according to simulated and observed streamflow time series, employing observed precipitation data as meteorological forcing, and then, the observed precipitation is replaced by GPDs for validation (Scheme-1). This method of model calibration is more efficient for gauged basins. In the second approach, (b) model parameters are calibrated/validated with streamflow data, utilizing each GPD as a model input independently (Scheme-2). This method is suggested for ungauged basins where only observed streamflow and GPDs are available (Jiang et al., 2012; Xue et al., 2013). Hence, in Scheme-1, there will be a single optimal model parameters based on observed precipitation, which are also being applied to the GPDs. In contrast to Scheme-1, each GPD and observed precipitation have their own optimal parameters set in Scheme-2. As a consequence, in this study, both schemes are considered for the hydrologic response of the basin based on observed and selected GPDs inputs. Finally, the TUW model parameters are automatically calibrated by employing the hydroPSO R package, which incorporates the particle swarm global optimization algorithm (Zambrano-Bigiarini & Baez-Villanueva, 2020; Zambrano-Bigiarini & Rojas, 2013).

Table 2.3. *TUW model parameters and their ranges. Abbreviations in the Process column represent; S, Snow; SM, Soil Moisture; and R, Runoff.*

Parameter	Units	Range	Process
Snow correction factor [SCF]	-	0.9–1.5	S
Degree-day factor [DDF]	mm/°C /day	0.0–5.0	S
Temperature threshold above which precipitation is rain [T_R]	°C	1.0–3.0	S
Temperature threshold below which precipitation is snow [T_S]	°C	−3.0–1.0	S
Temperature threshold above which melt starts [T_M]	°C	−2.0–2.0	S
Parameter related to the limit for potential evaporation [LP]	-	0.0–1.0	SM
Field capacity [FC]	mm	0.0–600	SM
Non-linear parameter for runoff production [β]	-	0.0–20	SM
Constant percolation rate [C_{PERC}]	mm/day	0.0–8.0	R
Storage coefficient for very fast response [k_0]	day	0.0–2.0	R
Storage coefficient for fast response [K_1]	day	2.0–30	R
Storage coefficient for slow response [K_2]	day	30–250	R
Threshold storage state [L_{UZ}]	mm	1.0–100	R
Maximum base at low flows [B_{MAX}]	day	0.0–30	R
Free scaling parameter [C_{ROUTE}]	day^2/mm	0.0–50	R

2.6. Verification metrics

In this context, the direct comparison of GPDs with observed precipitation is performed by exploiting modified Kling-Gupta Efficiency (Gupta et al., 2009; Kling et al., 2012) at the daily and monthly time steps, considering precipitation seasonal variability over entire region and different classifications. Kling-Gupta Efficiency (KGE) is a relatively new objective function combining correlation (r), bias (β) and variability ratio (γ) components. KGE is used to show the temporal dynamics (measured by r) and volume distribution (measured by β and γ) of precipitation, which is important for hydrology and water resource management.

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{2.1}$$

$$r = \frac{1}{n}\sum_{1}^{n}[(o_n - \mu_o) \times (s_n - \mu_s)] \times (\delta_o \times \delta_s)^{-1} \tag{2.2}$$

$$\beta = \mu_s \times (\mu_o)^{-1} \tag{2.3}$$

$$\gamma = (\delta_s \times \mu_o) \times (\delta_o \times \mu_s)^{-1} \tag{2.4}$$

Where, r represents the Pearson correlation coefficient, β is the ratio of mean observed and GPD precipitation, and γ is the ratio between coefficient of variation of observed and GPD precipitation. δ and μ show the standard deviation and mean of distribution, where o and s indicate the observed and estimated precipitation, and n shows sample size of observed or GPD, respectively.

In general, knowing the frequency of precipitation for different intensities is essential as having information about mean and spatio-temporal variation of precipitation (Sohn et al., 2013; Tian et al., 2007), where the same precipitation amount in the form of long-lasting light rain or a short-duration storm demonstrates different impacts on natural hazards, including floods and landslides (Z. Li et al., 2013; Y. Shen et al., 2010).

In this context, the Probability Density Function (PDF) is utilized to understand the frequency of precipitation for different intensities. Furthermore, the PDF is employed to measure the number of observations that are above or below a specific value in a data set. In other words, this indicates how frequently the GPD measurements are lower or higher than the precipitation from stations (Gebere et al., 2015). Hence, for the frequency and detectability analysis, daily precipitation from observed and GPDs are discretized and four possible outcomes are visible (Table 2.4).

Several authors divided precipitation events into rain/no-rain by applying a certain threshold (Blacutt et al., 2015; Ebert et al., 2007; Liu, Shangguan, et al., 2019; Sharannya et al., 2020), where limited GPDs detectability for precipitation with different intensities. Hence, daily precipitation data from observed and GPDs are discretized into the five thresholds, including no-precipitation (less than 1 mm/day), light precipitation (1–5 mm/day), moderate precipitation (5–20 mm/day), heavy precipitation (20–40 mm/day), and extreme precipitation (more than 40 mm/day) (Zambrano-Bigiarini et al., 2017). This classification is crucial for hydrological studies, where precipitation with different intensities may exhibit different hydrologic responses across the basin.

Table 2.4. *Contingency table used to define HKS.*

		Observed data		Total
		Precipitation	**No-Precipitation**	
GPD	**Precipitation**	Hit (H)	False (F)	GPD precipitation
	No-Precipitation	Miss (M)	Correct Negative (CN)	GPD No-Precipitation
Total		Observed precipitation	Observed No-Precipitation	Total

The Hanssen-Kuiper Skill (*HKS*) score, which is also known as true skill statistic or Pierce skill score, is employed to measure the detectability strength of GPD, distinguishing between occurrences and non-occurrences of a certain event.

$$HKS = \frac{(H \times CN) - (F \times M)}{(H + M)\,(F + CN)} \tag{2.5}$$

Where: H (Hit); when the observed precipitation is correctly detected, M (Miss); when the observed precipitation is not detected, F (False); when the precipitation is detected but not observed, and CN (Correct Negative); a no precipitation event is detected.

Finally, the Nash-Sutcliffe Efficiency (*NSE*) and KGE are used to evaluate the hydrological utility of GPDs for streamflow simulation.

$$NSE=1 - \frac{\sum_{i=1}^{n} (Q_i^s - Q_i^o)^2}{\sum_{i=1}^{n} (Q_i^o - \overline{Q_i^o})^2} \tag{2.6}$$

Where, n is the sample size of the observed or calculated streamflow. Q_i^o and Q_i^s present the observed and simulated streamflow, $\overline{Q_i^o}$ present the mean observed streamflow. KGE, r, β, γ, HKS, and NSE values all have their optimum at unity.

Figure 2.4 demonstrates the methodology flowchart for the selected GPDs spatio-temporal evaluation and their hydrological utility.

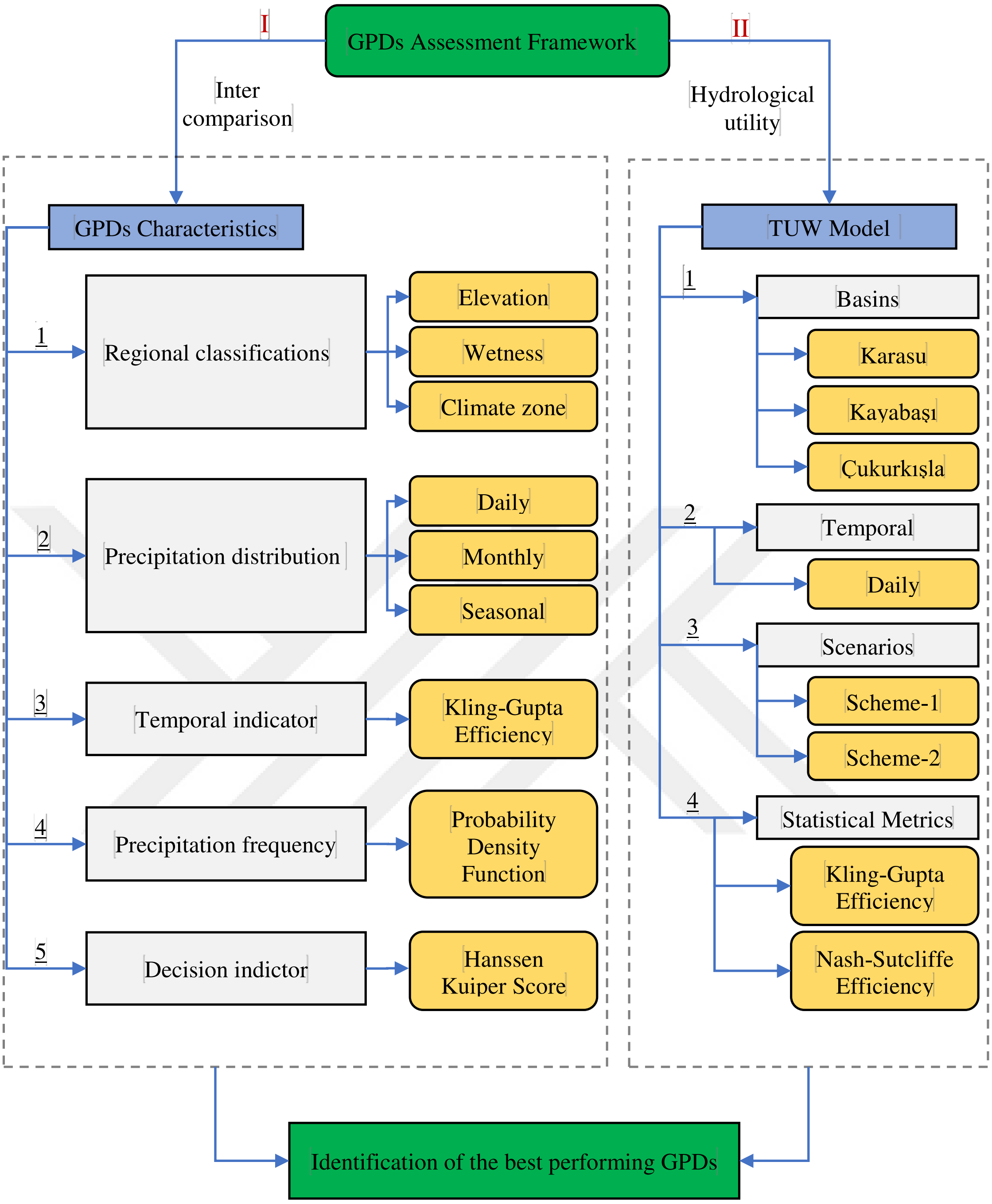

Figure 2.4. *Methodology flowchart for the direct performance evaluation of GPDs and their hydrological utility.*

3. RESULTS

3.1. Meteorological evaluation of GPDs

In this context, selected GPDs are being directly evaluated by observed precipitation data. Overall, this section of the study deals with a total of 26 263 358 daily observations obtained from 757 ground stations and information extracted from 18 GPDs at the station’s location for five hydrological years. Furthermore, the monthly precipitation was obtained from daily accumulated precipitation. It is important to mention that the terms "gauge-corrected" and "multi-source" are used interchangeably during the performance evaluation of GPDs.

3.1.1. Spatial and temporal evaluation of daily and monthly precipitation

The spatial distribution of mean daily and monthly precipitation from 18 GPDs is evaluated by direct comparison with observed precipitation. Figure 3.1 maps the mean daily and monthly precipitation at the station location, obtained from observed precipitation and selected GPDs for the entire period (2015–2019). Overall, the spatial distribution of observed precipitation demonstrates that the coastal areas receive higher precipitation amounts than the inland parts which is intense in the eastern region of the Black Sea. Likewise, as the distance increases from the coastal areas to the direction of the inner lands, the precipitation amount decreases, and the central regions experience the least amount of precipitation comparatively. From the result, CPCv1 displays mean daily and monthly distributed precipitation close to the reference (Obs). Moreover, among multi-source GPDs, MSWXv100, MSWEPv2.8, and CHIRPSv2.0 present mean daily and monthly precipitation closer to the observed, followed by MERRA-2, with slightly higher precipitation in the northern part of the Black Sea region. IMERGHHFv06 shows higher precipitation in the west and northwestern parts, where TMPA-3B42v7 maps lower precipitation in the northern regions and PERSIANN-CDR seems to map higher precipitation in western areas and lower precipitation in the eastern Black Sea regions. However, several gauge-corrected GPDs show higher precipitation in the central regions. Furthermore, among gauge uncorrected GPDs, CHIRPv2.0, which combines information from satellite and reanalysis, shows the same amount of mean daily and monthly precipitation as its corrected version (CHIRPSv2.0) and is close to observed precipitation, while ERA5 comes with the second-best spatial distribution of mean precipitation compared to gauge uncorrected GPDs, comparatively.

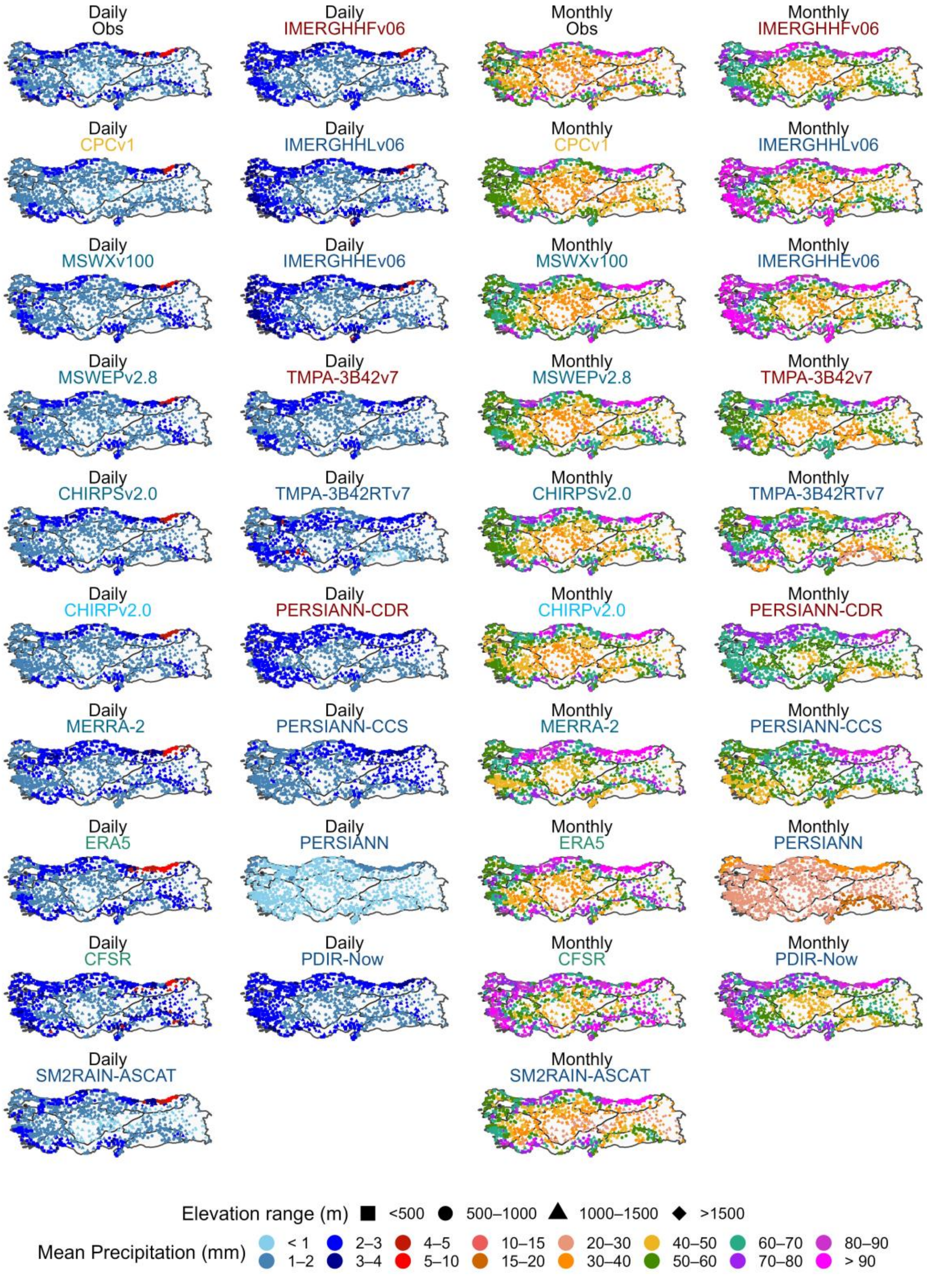

Figure 3.1. *Spatial distribution of mean daily and monthly precipitation at the station location obtained from observed (Obs) precipitation and 18 GPDs for the period of 2015–2019. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

However, ERA5 indicates significantly greater precipitation, in the eastern part of the Black Sea region, and CFSR seems to overestimate observed precipitation for most regions. In the same way, PERSIANN drastically underestimates mean daily and monthly precipitation for the entire area, where PERSIANN-CCS shows lower precipitation in the southwestern regions and greater precipitation in the eastern parts, and PDIR-Now provides a precipitation distribution that is comparable to PERSIANN-CDR. Both IMERGHHL(E)v06 exhibit greater precipitation in the western areas compared to their final product (IMERGHHFv06) and similar precipitation to that observed in the eastern parts. TMPA-3B42RTv7 shows less precipitation in the southeastern and northern parts of the country, while SM2RAIN-ASCAT shows mean daily and monthly distributions of precipitation closer to the observed than the other satellite-based precipitation datasets although it estimates higher precipitation in the southeastern part of the country.

Figures 3.2 and 3.3 show the frequency distribution of mean daily and monthly precipitation derived from observed and selected GPDs throughout the whole area for the entire period (2015–2019) and four seasons. Overall, observed mean precipitation frequently recurred between 0–3 mm for daily precipitation and 0–100 mm for monthly precipitation, while the observed mean precipitation frequency during the winter season reached up to 6 mm/day and 200 mm/month. According to the findings, the higher frequency of mean precipitation for observed and GPDs is around 2 mm/day and 60 mm/month. Among selected GPDs, PERSIANN shows the highest frequency within a narrow range of mean precipitation of 0–0.75 mm/day and 0–55 mm/month. Moreover, mean daily and monthly precipitation frequency during the spring season has a distribution close to the normal for many GPDs, and there is a positive skewed frequency distribution of mean daily and monthly precipitation for some GPDs which more intense for PERSIANN-CCS considering the summer and autumn seasons. The positive skewness of the frequency distribution can be attributed to the fact that high magnitude precipitation (extremes) is rarely occurred with lower frequency during the summer and autumn seasons, and during the winter and spring season the region often experiences precipitation with similar frequency and magnitude.

Figure 3.4 a and b summarize the mean daily and monthly precipitation at the regional scale, for the entire period and seasonal variability. According to observed mean daily and monthly precipitation, the region experiences 1.84 mm and 56.1 mm of mean daily and monthly precipitation for the entire period (2015–2019).

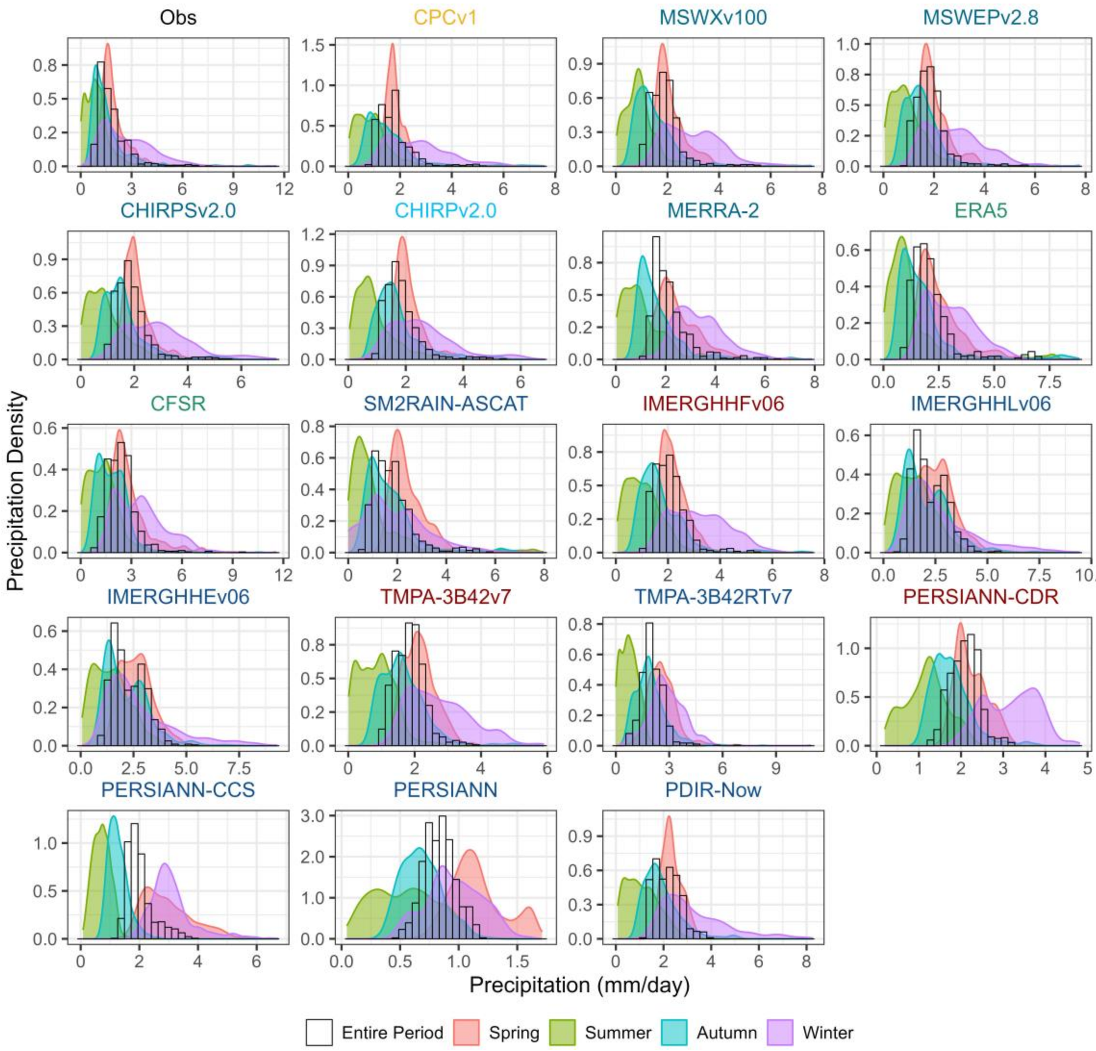

Figure 3.2. *Histogram density of mean daily precipitation from observed data and 18 GPDs for the entire period (2015–2019) and four seasons. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Furthermore, the region receives higher precipitation (2.73 mm/day and 82.14 mm/month) during the winter season, followed by the spring (1.96 mm/day and 60.22 mm/month) season, while the region experiences the least precipitation during the summer (1.04 mm/day and 31.9 mm/month) and autumn (1.65 mm/day and 50.16 mm/month) seasons. Overall, CPCv1 is able to present precipitation closer to the observed precipitation at the regional scale. Moreover, those multi-source GPDs, which combine information from three possible sources (S, R, and G), deliver accurate mean daily and monthly precipitation (except MERRA-2), respectively. In the same way, some GPDs, including PERSIANN-CCS, PERSIANN-CDR, PDIR-Now, and CFSR, highly overestimate mean precipitation.

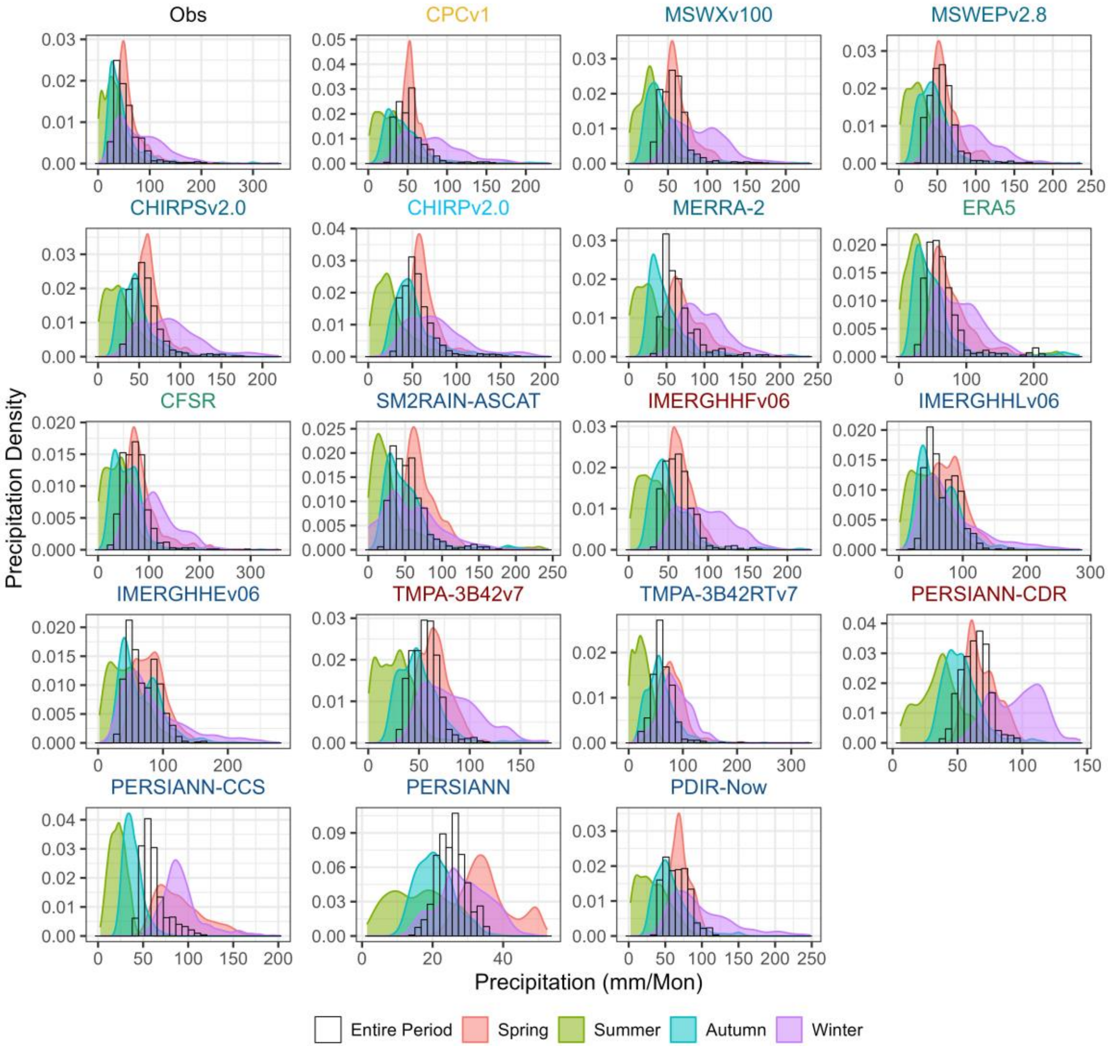

Figure 3.3. *Histogram density of mean monthly precipitation from observed data and 18 GPDs for the entire period (2015–2019) and four seasons. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Among satellite-based precipitation dataset, PERSIANN dataset significantly have the least mean precipitation amount, followed by SM2RAIN-ASCAT during the winter season. In the same way, Figure 3.4 c and d present mean daily and monthly precipitation over each year of the study. Generally, the region experiences greater precipitation during 2015 and 2019, but 2017 is known as the driest year due to it is receiving the least amount of precipitation. GPDs present distinct precipitation for different years. For instance, PDIR-Now shows close mean daily and monthly precipitation to observed precipitation from 2015 to 2017, while this dataset significantly overestimates observed precipitation from 2018 to 2019. On the other hand, some GPDs, including IMERGHHLv06 and IMERGHHEv06, consistently show higher mean precipitation than the observed over each year, and PERSIANN presents the lowest mean daily and monthly precipitation amount for the whole year, comparatively. Overall, CPCv1 and multi-source GPDs show daily and monthly precipitation close to the observed precipitation over each year which is followed by reanalysis dataset and satellite-based GPDs show distinct daily and monthly precipitation over selected five years.

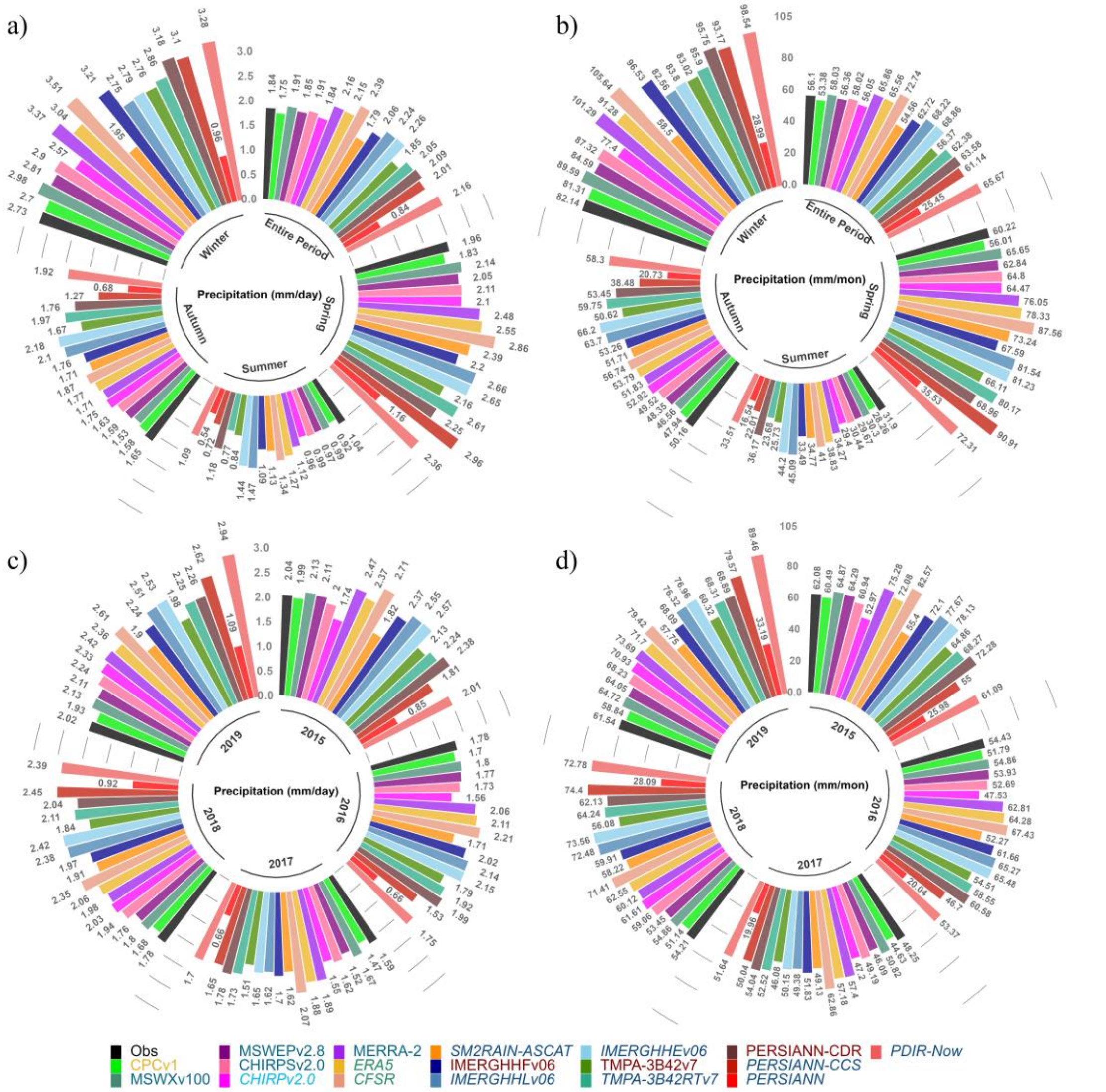

Figure 3.4. *a, b) Mean daily and monthly precipitation at the regional scale obtained from observed (Obs) precipitation and 18 GPDs for the period of 2015–2019 and four seasons. c, d) Mean daily and monthly precipitation at the regional scale over each year of the study obtained from observed (Obs) precipitation and 18 GPDs for the period of 2015–2019 and four seasons. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Figures 3.5 and 3.6 present the scatter plots of GPDs against observed precipitation for the entire period. In general, GPDs show concentrated point around the line with the unity slope for the larger time window, which means monthly precipitation from GPDs

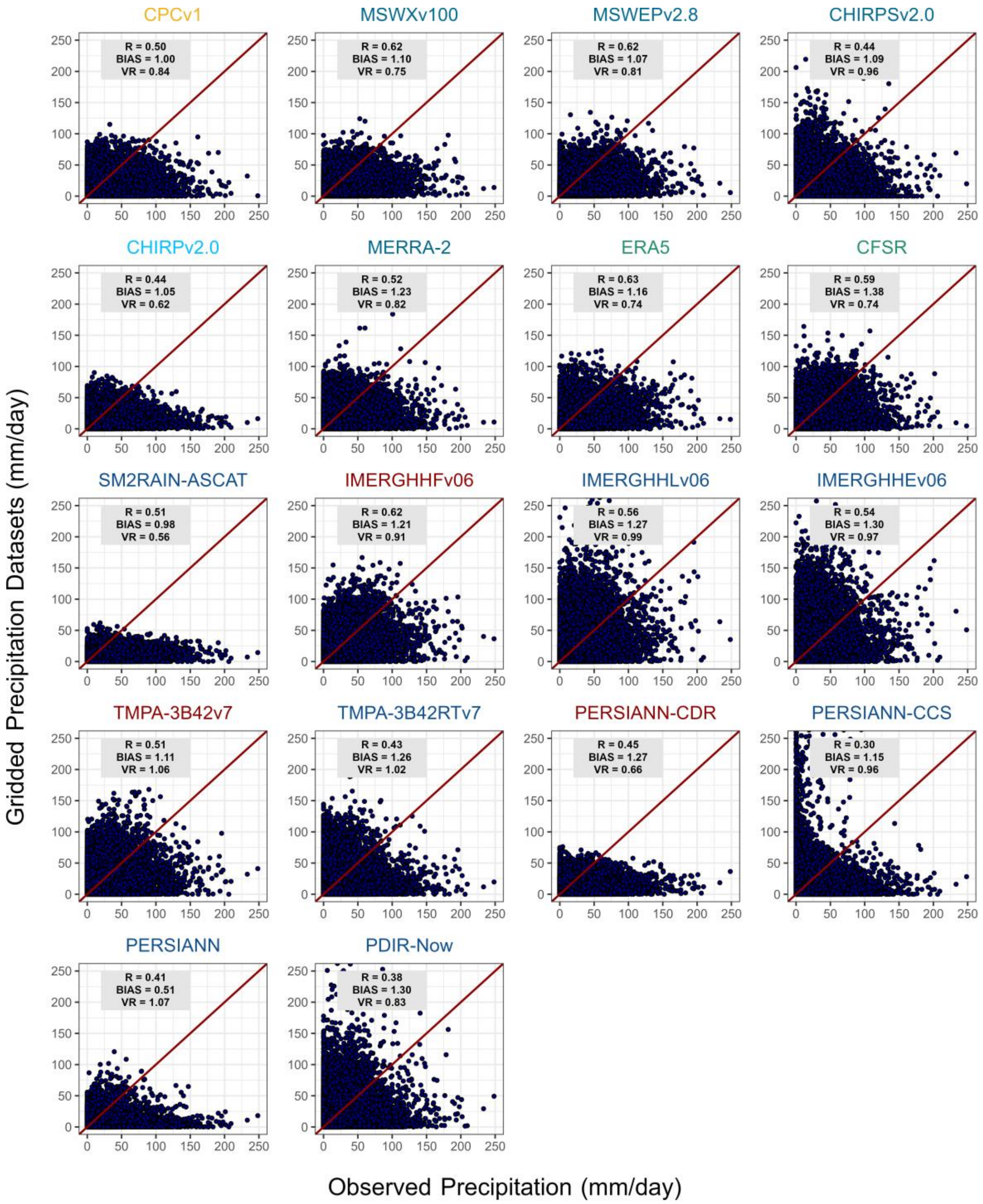

Figure 3.5. *Scatter plot of daily observed precipitation measurement versus 18 GPDs for the entire period. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

is closer to the observed precipitation than daily precipitation, which indicates a higher precipitation correlation between observed and GPDs. The observed daily precipitation events were mostly concentrated between 0–200 mm/day and 0–500 mm/month, where GPDs show different daily precipitation magnitudes.

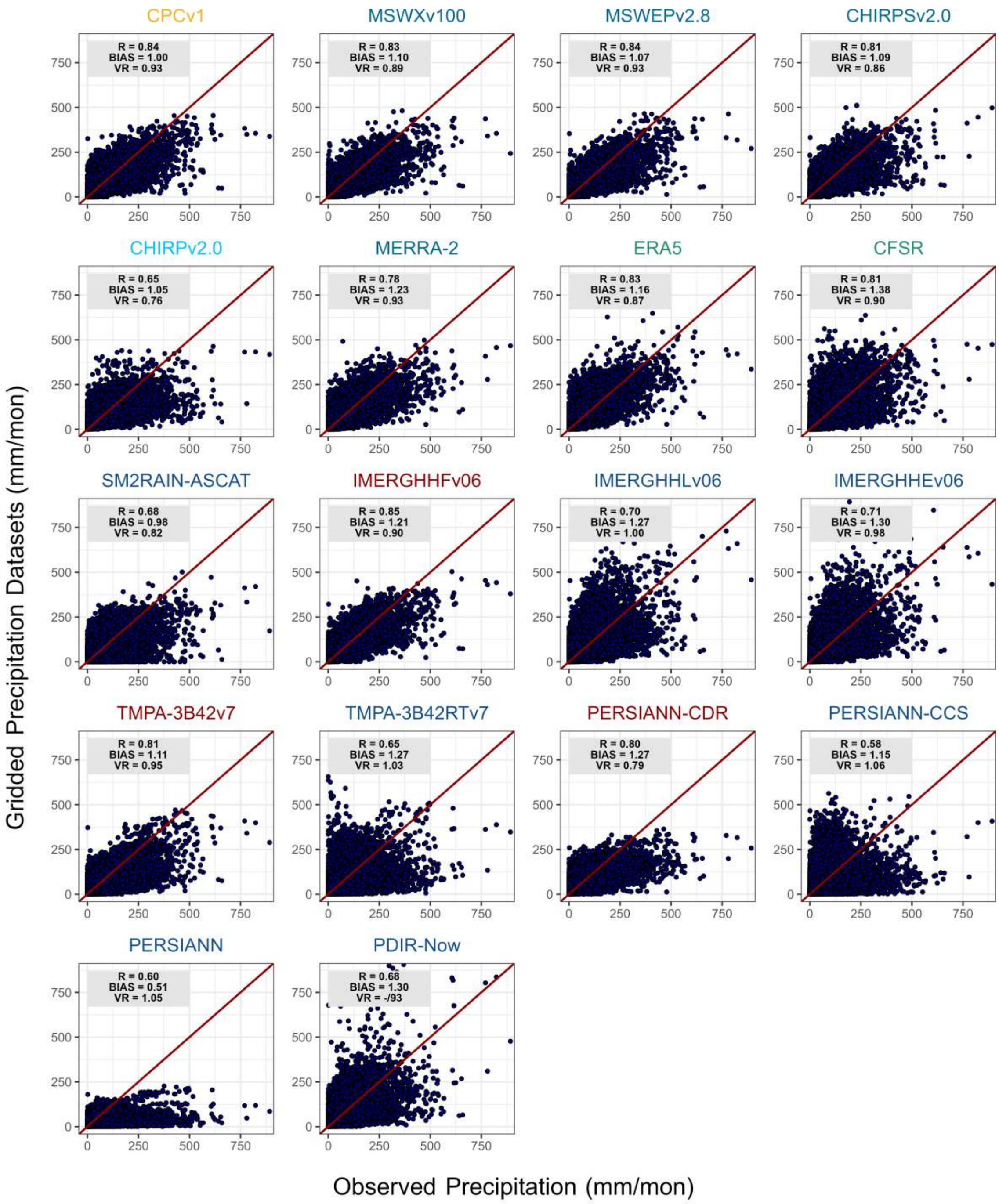

Figure 3.6. *Scatter plot of daily observed precipitation measurement versus 18 GPDs for the entire period. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

For instance, daily precipitation events recorded by CPCv1 are mostly confined between 0 and 100 mm/day, and the same condition is applicable for many GPDs, although it is more critical for SM2RAIN-ASCAT and PERSIANN-CDR.

Overall, ERA5 has the highest correlation (R = 0.63) with observed and PERSIANN-CCS demonstrates the lowest correlation (R = 0.30) with daily observed precipitation. Furthermore, the correlation between GPDs and observed precipitation is increased for monthly precipitation, where IMERGHHFv06 shows the highest correlation to the observed precipitation (R = 0.85), which is followed by CPCv1 and MSWEPv2.8 with a correlation of 0.84, and PERSIANN-CCS presents the lowest correlation (R = 0.58) to the observed for the monthly time step. It is important to keep in mind that the daily and monthly scatter plots only depict the distribution of precipitation events rather than their density. For example, PERSIANN-CDR and SM2RAIN-ASCAT show similar precipitation scatters. Where PERSIANN-CDR overestimates the ratio of bias (Bias = 1.27) and SM2RAIN-ASCAT underestimates the ratio of bias (Bias = 0.98), while for some GPDs, such as PERSIANN, the variability ratio is overestimated with an underestimation of the ratio of bias, and the opposite is also observable for some other GPDs (e.g., PERSIANN-CDR).

3.1.2. Performance of GPDs over the entire region

The reliability of GPDs at the station level is expressed in the form of KGE (Figure 3.7) and its Pearson correlation (Figure 3.8), the ratio of bias (Figure 3.9), and variability ratio (Figure 3.10) components for daily and monthly precipitation over the entire period (2015–2019). In general, GPDs perform better for the monthly time step than daily precipitation, which is distinguishable by a higher KGE and its correlation components. The GPDs demonstrate higher performance in the western parts of the country and lower performance in the eastern regions, which is typically described by complex topography and present weak performance over regions experiencing extremes such as the central and northeastern regions. From the results, CPCv1, which takes advantage of the spatial information of available gauges, demonstrates relatively higher performance for both daily and monthly precipitation, especially in the western regions, where this dataset is outperformed by some multi-source GPDs, such as MSWEPv2.8, IMERGHHFv06, and MSWXv100, while other gauge-corrected GPDs, such as MERRA-2 near the Black Sea region and CHIRPSv2.0 in the central areas and in the northern parts of the country, demonstrate weaker performance for daily time step but their performance increased for monthly time step. In the same way, TMPA-3B42v7 shows relatively higher performance in the western and southern regions than in the other parts, where PERSIANN-CDR displays poor performance for many regions considering daily precipitation, and this dataset carries out higher performance in the northern parts for the monthly time step, respectively.

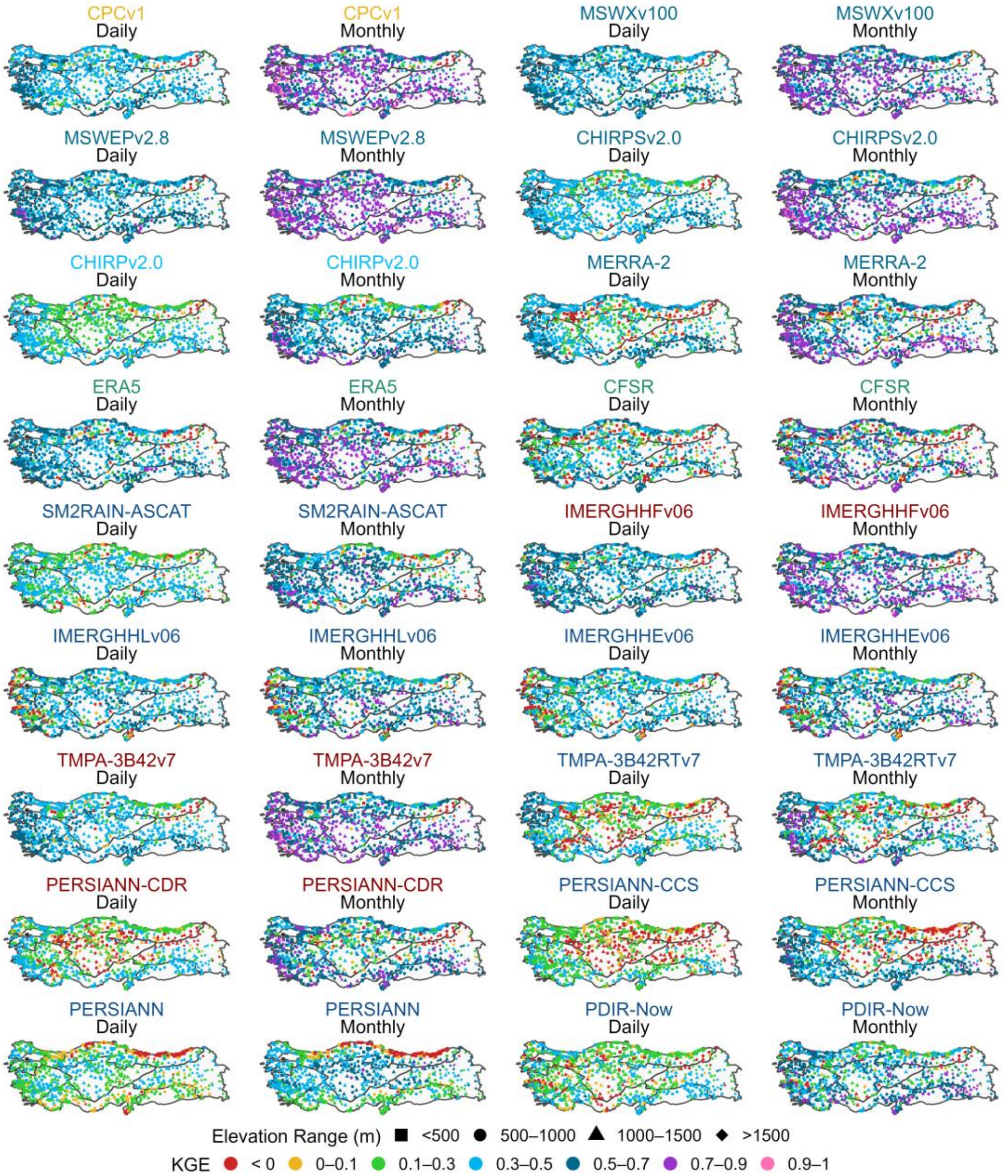

Figure 3.7. *GPDs reliability at the station level is expressed in the form of KGE for the daily and monthly precipitation considering data from 2015 to 2019. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Among gauge-uncorrected GPDs, some reanalysis datasets, such as ERA5 shows relatively higher performance compared to CFSR and satellite-based precipitation datasets and even performed higher than some gauge-corrected GPDs such as CHIRPSv2.0, MERRA-2, TMPA-3B42v7 especially during the daily time step, where CFSR, and IMERGHHE(L)v06 presents higher performance compared to TMPA-3B42RTv7, PERSIANN-CCS, PDIR-Now, and PERSIANN for the daily time step and SM2RAIN-ASCAT is the only satellite-based GPD that displays higher performance for the monthly time step, where PERSIANN presents the lowest performance for the daily time step, although PERSIANN depicts higher performance in the central regions compared to the other PERSIANN datasets (i.e., PERSIANN-CDR, PERSIANN-CCS, and PDIR-Now). On the other hand, CHIRPv2.0, which takes advantage of reanalysis and satellite information, presents relatively higher performance than satellite-based datasets and lower performance than reanalysis for both daily and monthly time steps, comparatively. Finally, some GPDs such as PERSIANN and PERSIANN-CCS still show lower performance over the northern parts of the Black Sea region and central regions, which is followed by TMPA-3B42RTv7 over the central and southeastern parts of the country considering the monthly time step.

Generally, GPDs have higher correlation (Figure 3.8) for the monthly time step than for the daily time step, which is more observable for GPDs whose temporal variations entirely (i.e., CPCv1) or to some degree depend on the ground data. For instance, CHIRPSv2.0, which has lower correlation for daily precipitation, shows significantly higher correlation to the observed precipitation for the monthly time step. However, the correlation increment is not only observed in the gauge-corrected GPDs but is also observable in gauge-uncorrected GPDs when the time step is shifted from daily to monthly. Among selected GPDs, ERA5 and IMERGHHFv06 deliver relatively higher correlation of precipitation to the observed, which is followed by some multi-source datasets such as MSWEPv2.8 and MSWXv100, where other gauge-corrected datasets demonstrate lower precipitation correlation in different regions of the country. Except for IMERGHHL(E)v06, the rest of the gauge-uncorrected GPDs demonstrate weak correlation to the observed precipitation at the station location, which is more observable for the daily time step.

It’s worth mentioning that the ratio of bias component of KGE (Figure 3.9) demonstrates the precipitation total volume from a certain GPD to the observed precipitation over a specific study time period, and it is more important that the ratio of bias does not depend on the selected time steps within that study time period, where each GPD shows the same bias over different time steps.

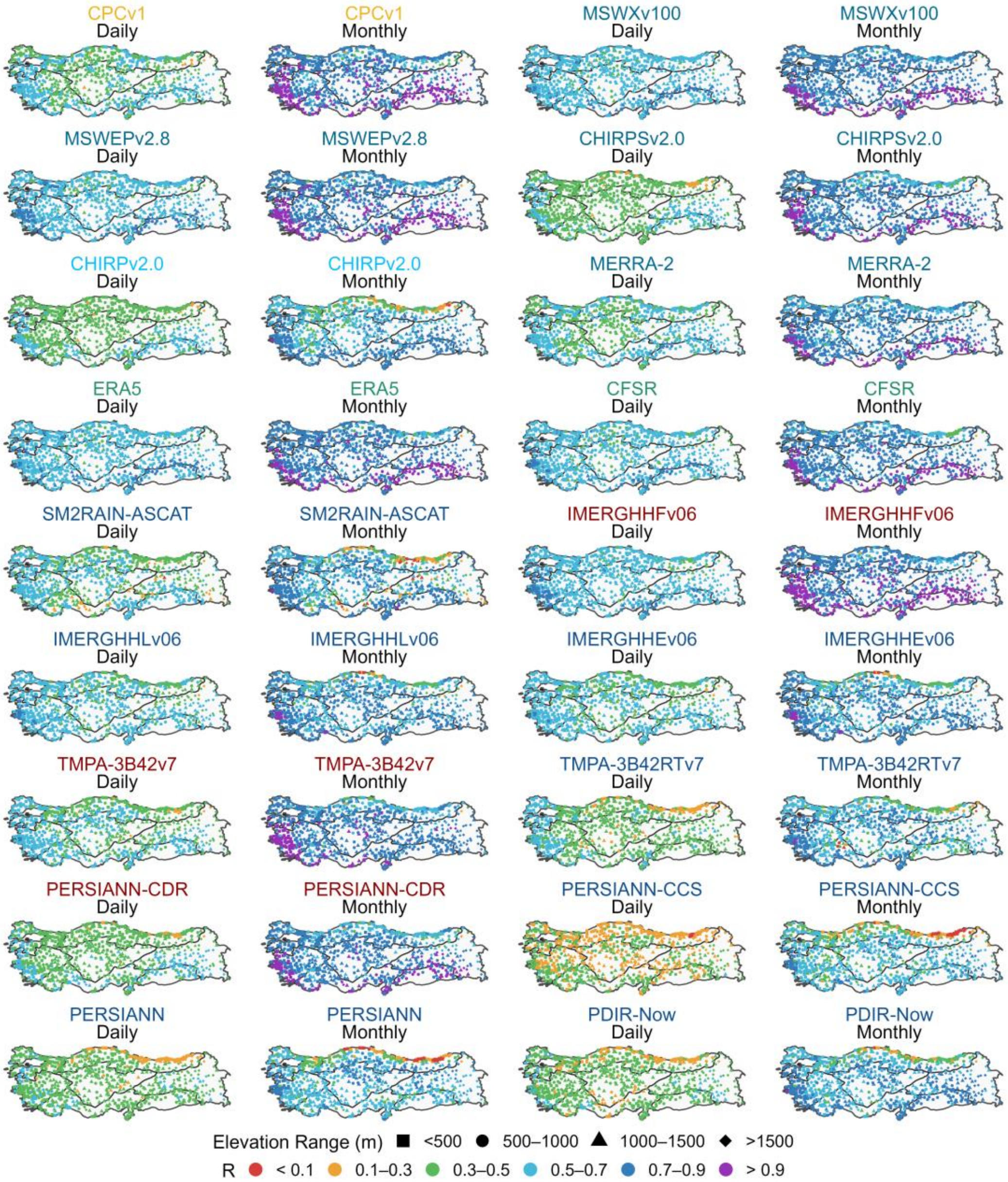

Figure 3.8. *GPDs reliability at the station level is expressed in the form of Pearson correlation coefficient (R) for the daily and monthly precipitation considering data from 2015 to 2019. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

In other words, the ratio of bias for a specific GPD is not changed as the time step shifts from the daily to the monthly. When compared to ground-based, gauge-corrected, and certain reanalysis GPDs, satellite-based GPDs have a ratio of bias variation at the station location that is more intense and distributed than the other three types. Overall, the ratio of bias of 0.8–1.2 is observable for many GPDs in the western parts of the country, whereas each GPD shows a different ratio of bias in the central regions and eastern mountainous parts.

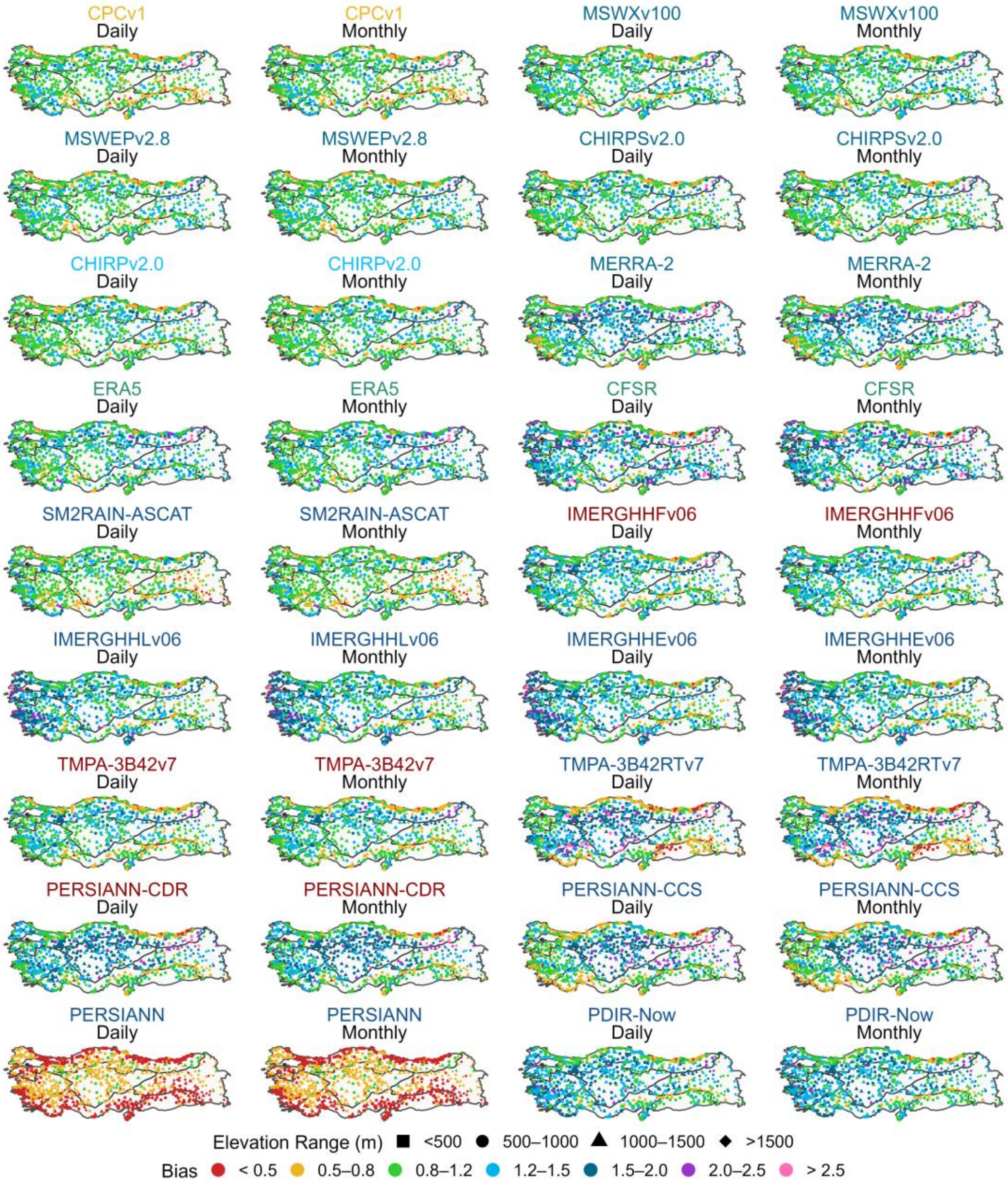

Figure 3.9. *GPDs reliability at the station level is expressed in the form of ratio of bias (Bias) for the daily and monthly precipitation considering data from 2015 to 2019. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Among GPDs, PERSIANN dataset shows much lower ratio of bias, which can be attributed to the fact that the dataset significantly underestimates precipitation compared to observed data, especially in the coastal areas and southeastern regions. In the same way, IMERGHHL(E)v06 shows distinct ratio of bias and significant overestimate of observed precipitation in the western part of the country compared to other GPDs, which is followed by CFSR dataset, respectively.

Unlike the ratio of bias component of KGE, the variability ratio (Figure 3.9), which shows the variability or dispersion of GPDs magnitudes relative to their mean compared to the observed precipitation variability relative to its mean, is a function of time. In other words, the variability ratio of GPDs changes when the time step shifts from daily to monthly. In general, most of the GPDs show lower dispersion of precipitation relative to their means compared to the observed variability for the daily time step. In other words, most of the multi-source datasets, ground-based dataset, and gauge-uncorrected datasets underestimate the variability ratio for the daily time step compared to the monthly time step, which is more observable for CPCv1, MSWEPv2.8, MSWXv100, MERRA-2, PERSIANN-CDR, CHIRPv2.0, ERA5, CFSR, SM2RAIN-ASCAT, and PDIR-Now, where SM2RAIN-ASCAT significantly underestimates the variability ratio in the coastal regions close to the Black Sea for the daily precipitation, and the rest of the selected GPDs overestimate the variability ratio. On the other hand, some GPDs, such as CHIRPSv2.0, PERSIANN, PERSIANN-CCS, IMERGHHL(E)v06, TMPA-3B42(RT)v7, highly overestimate the variability of precipitation compared to observed precipitation over different parts of the country. In the same way, GPDs, including PERSIANN-CCS and TMPA-3B42RTv7, still drastically overestimate the variability ratio in the northern parts, particularly in the northwest of the Black Sea region for the monthly time step.

Figures 3.11 and 3.12 present the box-and-whisker plots of KGE and its Pearson correlation (R), the ratio of bias (Bias), and variability ratio (VR) components for 18 GPDs at the regional scale for daily and monthly precipitation considering the entire period (2015–2019) and four seasons.

In general, the mean of median scores for KGE over all GPDs is estimated at 0.37 and 0.55 for daily and monthly precipitation for the entire period, while this amount decreases in the spring season to 0.34 and 0.45, and the lowest mean of median KGE comes with the summer season precipitation, which varies between 0.20 and 0.38 for daily and monthly precipitation. Similarly, the mean median of KGE for the autumn season is estimated at 0.38 and 0.49, and the winter season mean of median KGE varies between 0.32 and 0.46 for daily and monthly precipitation.

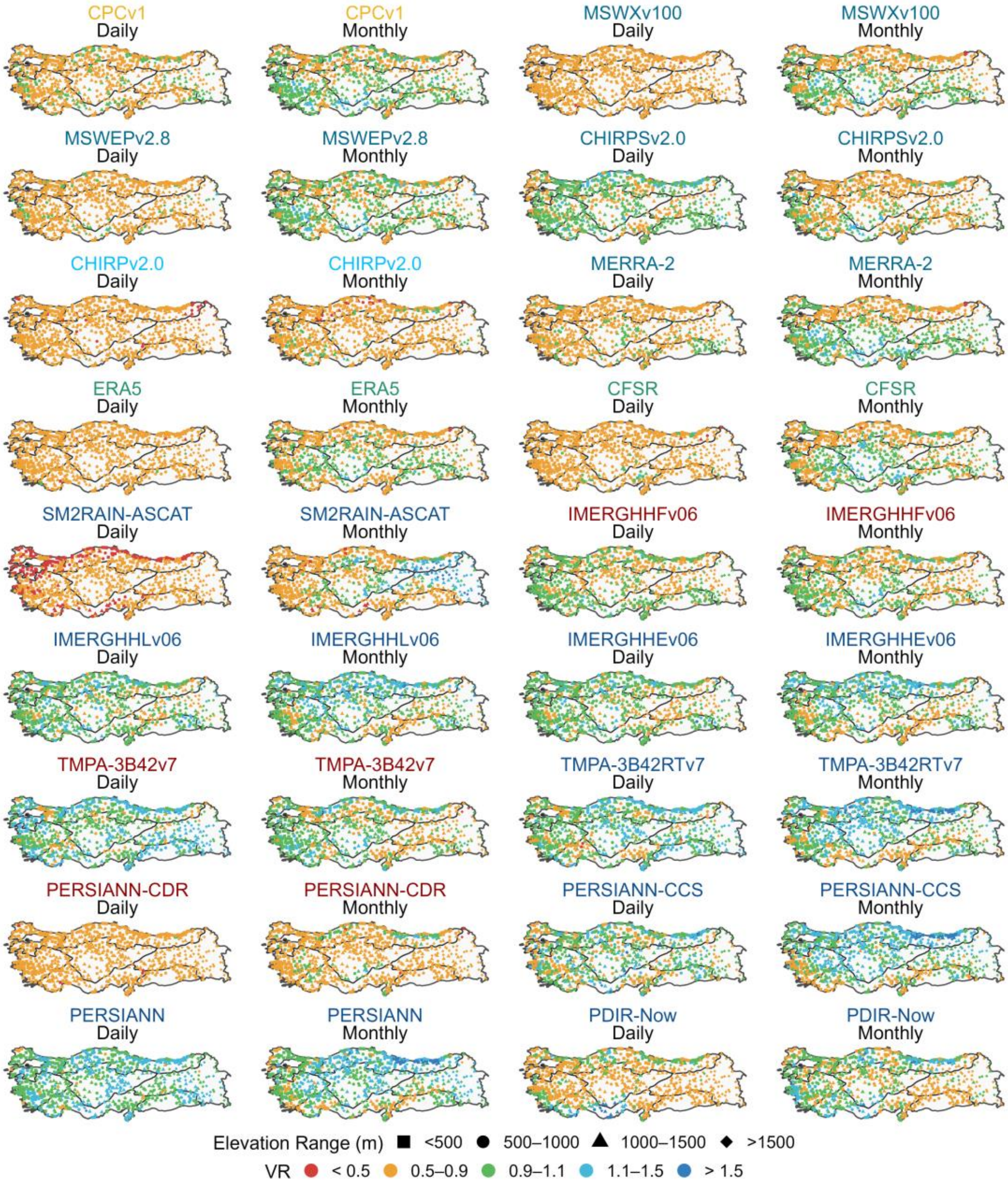

Figure 3.10. *GPDs reliability at the station level is expressed in the form of variability ratio (VR) for the daily and monthly precipitation considering data from 2015 to 2019. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

In the same way, the mean of the central tendency scores for Pearson correlation, the ratio of bias, and the variability ratio components of the KGE for daily and monthly precipitation are 0.50:0.75, 1.14:1.14, and 0.85:0.92 for the entire period; 0.50:0.71, 1.20:1.20, and 0.81:0.83 for the spring season; 0.40:0.73, 1.08:1.08, and 0.78:0.98 for the summer season; 0.56:0.74, 1.15:1.15, and 0.85:0.90 for the autumn season; and 0.51:0.73, 1.11: 1.11, and 0.89:0.91 for the winter season, respectively.

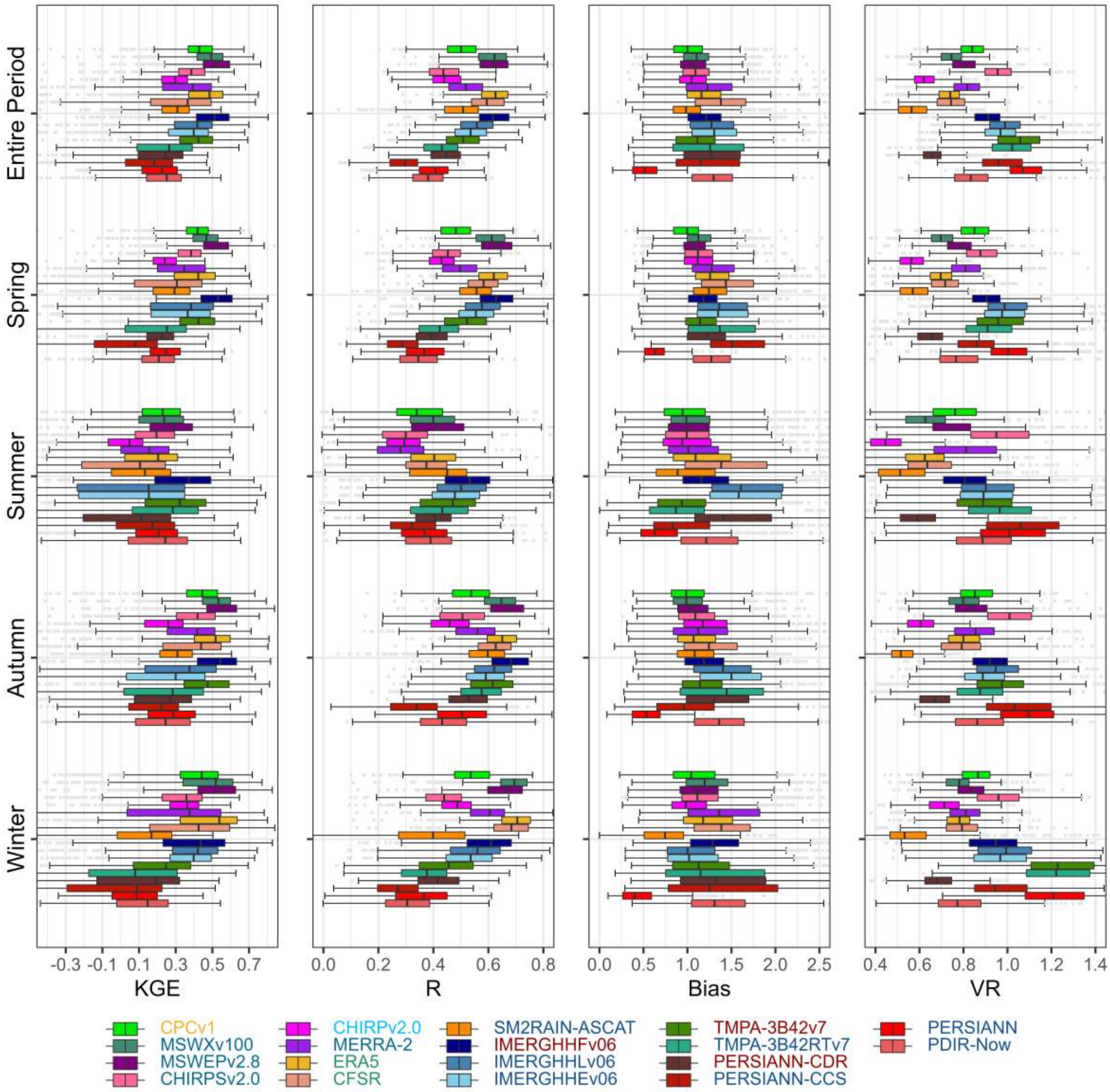

Figure 3.11. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the daily precipitation, considering data from 2015 to 2019 and four seasons. The boxes represent the 25th, 50th, and 75th percentile values, respectively. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

In addition, the findings show that there is higher agreement between observed precipitation and selected GPDs throughout the winter, spring, and autumn seasons than during the dry season (summer). Since the KGE combines the correlation, the ratio of bias, and the variability ratio, it is important to know which component has the most effect on the final KGE score. Looking at the daily and monthly performance of GPDs, most of the selected precipitation datasets performed fairly well in terms of correlation for daily precipitation compared to the monthly time step, which is understandable considering that long-term climatological precipitation statistics are easier to estimate than day-to-day precipitation dynamics.

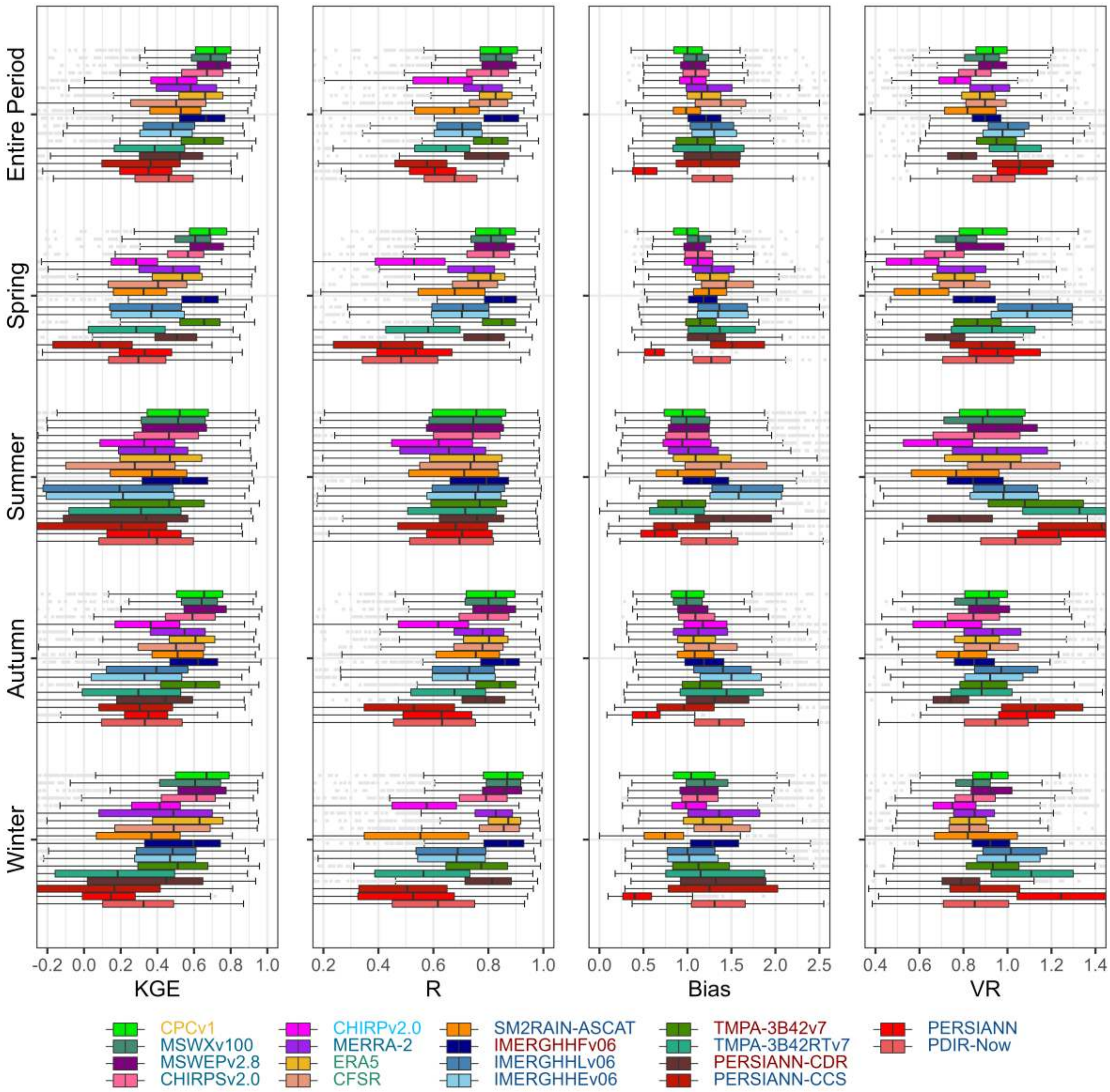

Figure 3.12. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the monthly precipitation, considering data from 2015 to 2019 and four seasons. The boxes represent the 25th, 50th, and 75th percentile values, respectively. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The correlation values have the most influence on the final KGE value due to the squaring of the three components in the KGE equation (see Equation 2.1). Hence, the

KGE performance ranking aligns well with the correlation performance ranking (Figures 3.11 and 3.12). These findings imply that the correlation is the most essential component value to enhance in order to improve the KGE value. This indicates that, for existing daily GPDs, changes to precipitation event timing (which dominates the correlation value) are more important and useful than improvements to precipitation total volume (which dominates the bias value) or the intensity distribution (which dominates the variability value).

Figure 3.13 depicts the performance of GPDs at the regional scale, expressed in the form of KGE and its individual components for both daily and monthly time steps over the entire period as well as seasonal variability. At the regional scale, GPDs show higher performance in terms of KGE for monthly time step than for daily time step, and it is because of the improvement in correlation, where the ratio of bias for both time steps remains unchanged, while underestimation of the variability ratio mostly occurs for the daily time step. Among GPDs, the IMERGHHL(E)v06 show higher overestimation of bias during the summer season while it is underestimated by PERSIANN and SM2RAIN-ASCAT during the winter season for both daily and monthly steps. Generally, underestimation of the variability ratio by many GPDs is more observable for the daily time step than for the monthly time step, and it is more intense during the spring season. In the same way, many GPDs (i.e., it is more visible for PERSIANN-CDR, SM2RAIN-ASCAT, and CHIRPv2.0 datasets) underestimate the variability of observed precipitation for both daily and monthly precipitation, with the exception of TMPA-3B42RTv7, PERSIANN-CCS, and PERSIANN for different time scales. Overall, GPDs show relatively poor performance during the summer season, with the exception of TMPA-3B42RTv7 and PDIR-Now, which display higher performance during this season when comparing their performance over other seasons for both time steps. From the results, MSWEPv2.8 demonstrated the highest performance, followed by IMERGHHFv6, MSWXv100, ERA5, and CPCv1 datasets. Overall, most of multi-source GPDs, which combine information from satellite, reanalysis, and ground data, perform substantially well, followed by reanalysis datasets, while satellite-based GPDs show poor performance. It’s worth mentioning that some exceptions exist; for instance, IMERGHHL(E)v06 demonstrates higher performance and outperforms even some gauge-corrected GPDs such as MERRA-2 and CHIRPSv2.0, and also shows higher performance than some reanalysis datasets such as CFSR dataset.

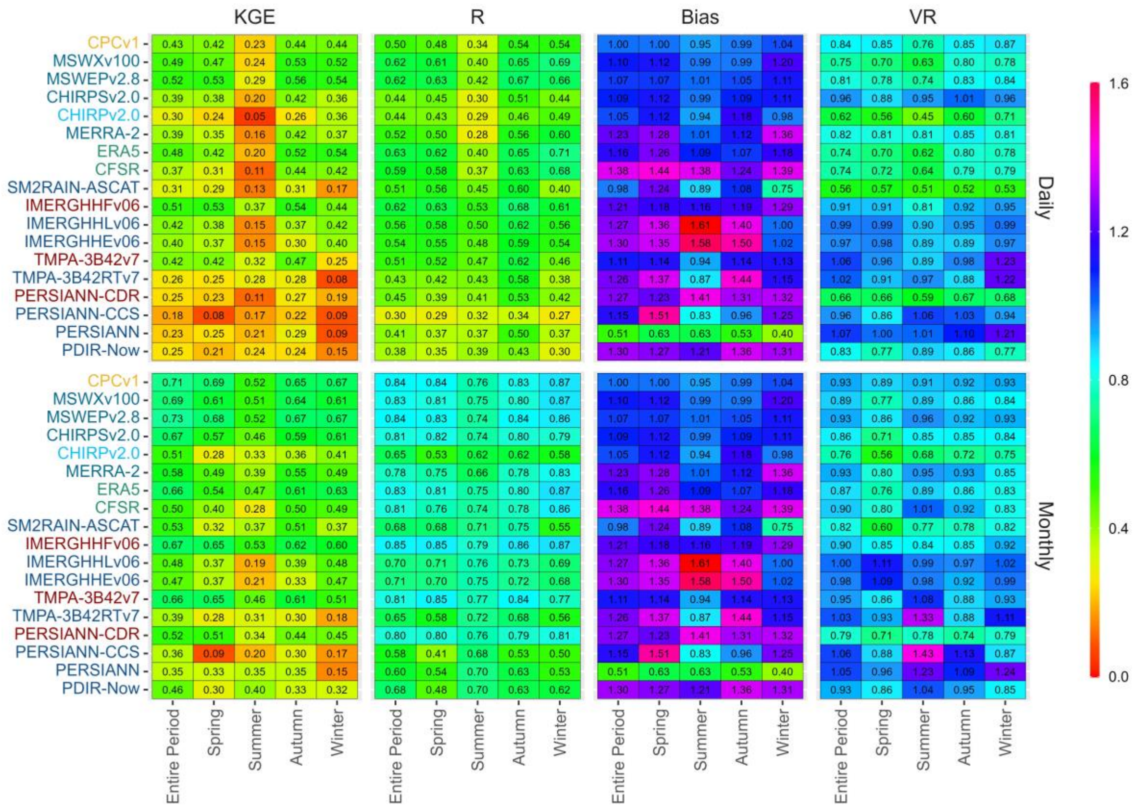

Figure 3.13. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the daily and monthly precipitation, considering data from 2015 to 2019 and four seasons. The heat table represent the 50th percentile values of figure 3.11 and 3.12, respectively. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

3.1.3. Performance of GPDs over different elevation ranges

Figures 3.14 and 3.15 depict the daily and monthly performance of GPDs at the regional scale over different elevation ranges for the entire period and four seasons.

In general, the combined effect of elevation and seasonal variability on the performance of GPDs is mostly distinguishable during the winter season, when most GPDs perform poorly as elevation increases, and their performance is worse during the summer season as elevation decreases. Moreover, GPDs show higher performance for the monthly time step than for the daily time step almost over all elevation ranges, and most of the GPDs demonstrate relatively higher performance for lower-elevated regions, where their performance decreases as the elevation increases over the entire period.

Overall, the highest performance is obtained by ERA5, followed by MSWEPv2.8 and MSWXv100 during the winter season, while some GPDs, such as CHIRPv2.0, PERSIANN, and PERSIANN-CDR during the summer season and PERSIANN during the winter season, present the lowest performance for the daily time step over regions with elevations less than 500 m.

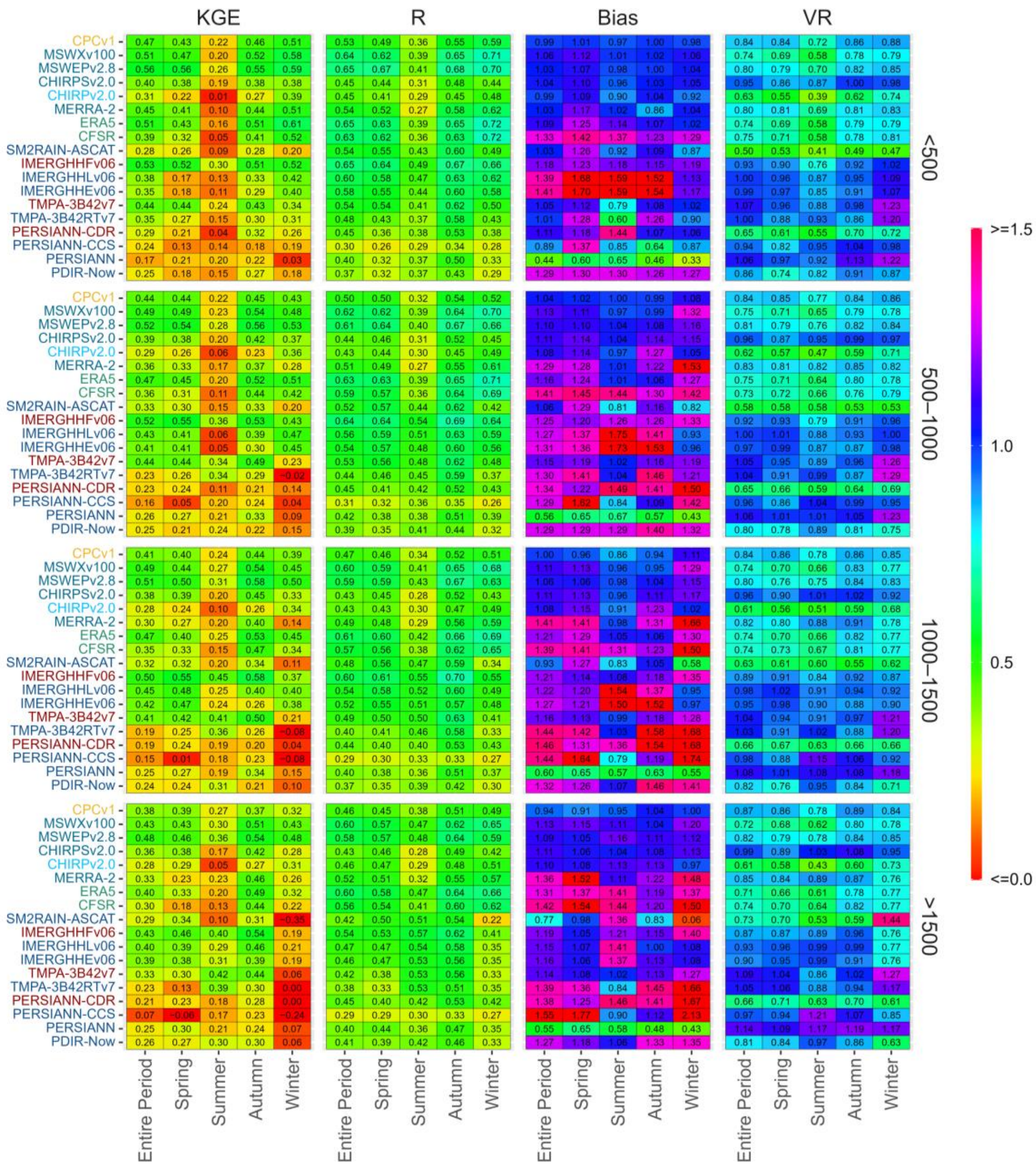

Figure 3.14. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the daily precipitation, considering data from 2015 to 2019 and four seasons. The heat table represent the median of KGE and its components over four distinct elevation ranges, respectively. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

For the monthly time step, CPCv1 and MSWEPv2.8 over the entire period and in the summer and spring seasons demonstrate the highest performance, while PERSIANN during the winter season, IMERGHHEv06 during the summer season, and PERSIANN-CCS during the spring season show the lowest performance among selected GPDs over areas with elevation less than 500 m. Considering the individual daily and monthly time steps, GPDs display distinct reliability over each elevation range, and even the performance of a certain GPD may vary over a distinct elevation range considering the entire period and different seasons. For instance, MSWEPv2.8 shows the highest performance over regions with elevation less than 500 m (i.e., this elevation mostly represents coastal areas) considering the entire period, and its performance decreases as the elevation increases. The dataset presents the same performance trend over distinct elevation ranges during the spring season and displays higher performance during the summer season over regions with elevation more than 1500 m, where its performance decreases as the elevation decrease. In the autumn season, it shows higher performance over regions located between 500 and 1500 m elevation, and finally, its performance during the winter season decreased as the elevation increased and is even outperformed by some reanalysis datasets such as ERA5 over areas having elevation less than 500 m. In the same way, MSWEPv2.8 performs differently over each elevation range, considering the entire period and each season (Figure 3.15) for the monthly time step, respectively. Overall, ground-based and multi-source GPDs perform well over selected four elevation ranges for both daily and monthly time steps. However, some reanalysis datasets such as ERA5 demonstrate higher performance than CPCv1 and many multi-source GPDs such as MSWXv100, MERRA-2, TMPA-3B42v7, CHIRPSv2.0, and PERSIANN-CDR over areas with elevation less than 500 m, especially for the daily time step over the entire period. In the same way, among gauge-uncorrected datasets, satellite-based GPDs show lower performance over different elevation ranges and seasons. While some exceptions exist, for example, IMERGHHL(E)v06 and CHIRPv2.0 demonstrate higher performance than other satellite-based datasets and even outperform some gauge-corrected GPDs such as PERSIANN-CDR over areas having elevation less than 500 m, where IMERGHHLv06 shows the fourth-best performance after MSWEPv2.8, MSWXv100, and IMERGHHF06 over regions having elevation more than 1500 m for the daily time step over the entire period where the dataset performance decreases for the monthly time step. From the satellite-based GPDs, PERSIANN datasets demonstrate lower performance, which is more intense for PERSIANN-CCS as the elevation increases for both daily and monthly time steps.

In the same way, GPDs' correlations are mostly depicted by their KGE values, and any increase or decrease of correlation over a distinct elevation range directly affects the

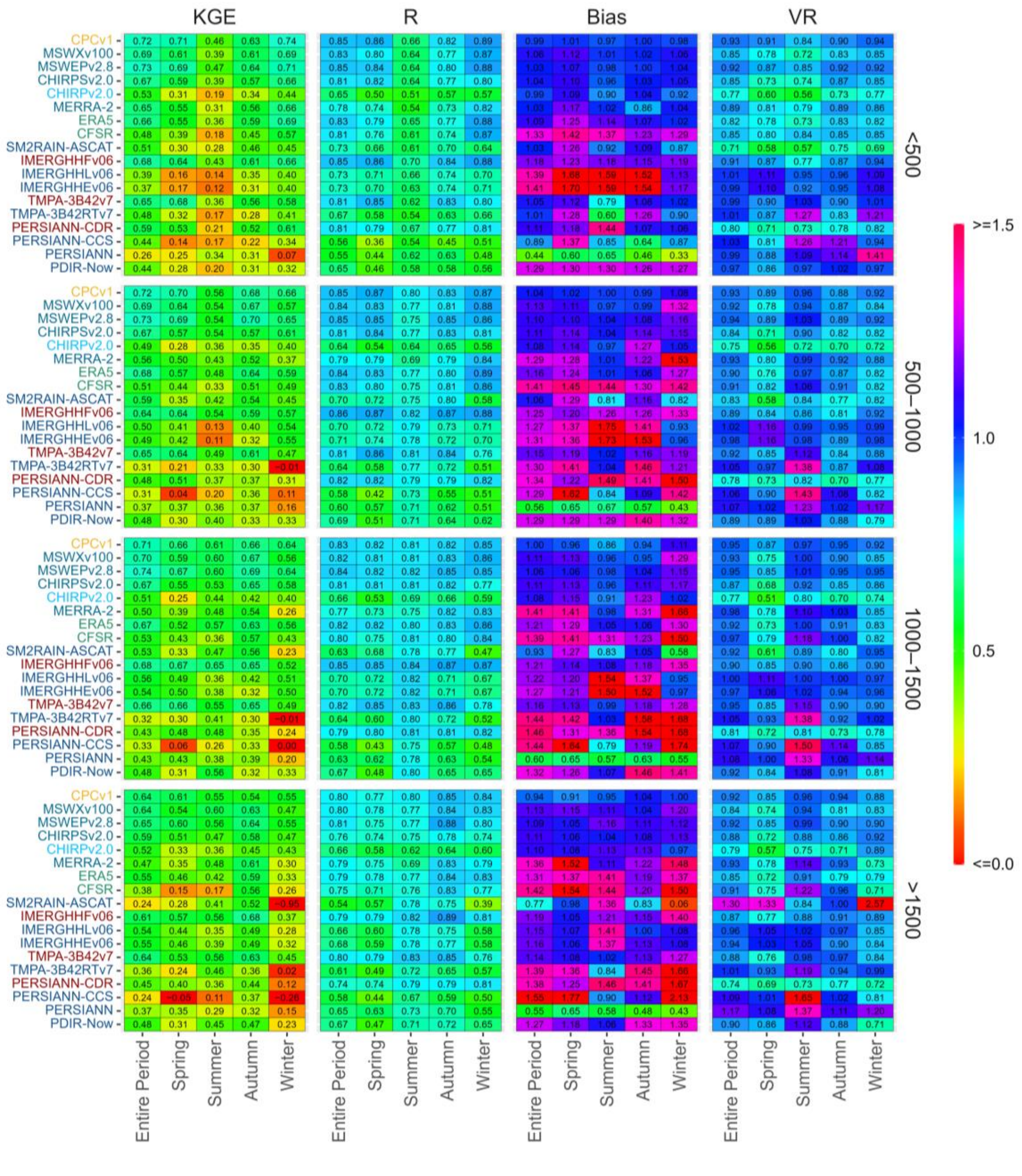

Figure 3.15. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the monthly precipitation, considering data from 2015 to 2019 and four seasons. The heat table represent the median of KGE and its components over four distinct elevation ranges, respectively. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

final performance, where the highest correlation is obtained by CFSR and ERA5 over regions having elevation less than 500 m and the lowest correlation to the observed precipitation is presented by SM2RAIN-ASCAT over regions with elevation more than

1500 m during the winter season as well as PERSIANN-CCS over areas with elevation less than 500 m during the spring season for the daily time step. For the monthly time step, CPCv1 demonstrates the highest correlation during the winter season, which is followed by IMERGHHFv06 over areas with elevation greater than 1500 m during the spring season, where ERA5 comes in third-best over areas with elevation 500–1000 m. Among selected GPDs, PERSIANN-CCS shows the lowest correlation during the spring season over areas having elevation less than 500 m, followed by SM2RAIN-ASCAT over areas with elevation more than 1500 m, respectively.

It is important to mention that the GPDs bias stays unchanged over different elevation ranges as the time step shifts from daily to monthly. Considering the combined effect of elevation changes and temporal variability (entire period and four seasons) at the daily and monthly time steps; for the entire period, among selected GPDs, PERSIANN datasets over different elevation ranges, and SM2RAIN-ASCAT and CPCv1 over regions with elevation greater than 1500 m as well as PERSIANN-CCS over areas with less than 500 m elevation highly underestimate bias, whereas PERSIANN-CCS over areas with elevation more than 1500 m and PERSIANN-CDR and TMPA-3B42RTv7 over areas with elevation 1000–1500 m drastically overestimate bias. Taking the effect of different seasons and elevation ranges into consideration, PERSIANN-CCS and IMERGHHL(E)v06 highly overestimate the ratio of bias during the winter and spring seasons over areas having elevation less than 1500 m, whereas SM2RAIN-ASCAT over areas having elevation more than 1500 m and PERSIANN over regions with elevation less than 500 m highly underestimate the ratio of bias during the winter season for both daily and monthly steps.

Finally, GPDs demonstrate distinct variability of precipitation compared to observed precipitation, where the underestimation of precipitation variability is more intense for the daily time step than for the monthly time step, and GPDs show highly different variability ratios for different seasons. Looking at the effect of elevation and temporal variability (entire period and four season) over GPDs variability for the daily time step, among datasets, CHIRPv2.0 and SM2RAIN-ASCAT highly underestimate the variability ratio over areas having elevation less than 500 m during the summer season, and SM2RAIN-ASCAT over areas having elevation more than 1500 m and TMPA-3B42RTv7 over areas with elevation 500–1000 m drastically overestimate the variability ratio during the winter season, respectively. For the monthly time step, still CHIRPv2.0

highly underestimates the variability ratio over areas with elevations between 500 and 1500 m, whereas SM2RAIN-ASCAT during the winter season and PERSIANN-CCS during the summer season over areas with elevations greater than 1500 m highly overestimate variability ratio compared to other selected GPDs, respectively.

3.1.4. Performance GPDs over different wetness classes

Figures 3.16 and 3.17 present the daily and monthly performance of GPDs over four distinct wetness classes for the entire period and four seasons.

Overall, GPDs show higher performance over moderately dry and moderately wet regions than over wet and dry areas, where the datasets demonstrate lower performance during the summer and winter seasons compared to other seasons over all wetness classes, and GPDs demonstrate weak performance during the winter season especially over dry and wet regions. From the results, among multi-source GPDs, IMERGHHFv06 delivers the highest performance over moderately wet regions, which is followed by MSEWPv2.8 and MSWXv100 for the entire period considering the daily time step. Among gauge-uncorrected GPDs, ERA5 and CFSR show higher performance over moderately dry and wet regions, and even ERA5 outperforms CPCv1 and some gauge-corrected GPDs such as TMPA-3B42v7, MERRA-2, PERSIANN-CDR, and CHIRPSv2.0 over different wetness classes for the entire period at the daily time step. Moreover, CHIRPv2.0 is also able to present higher performance than most satellite-based GPDs over different wetness classes, where IMERGHHL(E)v06 demonstrate relatively higher performance than the other satellite-based datasets and even outperformed CHIRPv2.0 and some gauge-corrected GPDs such as MERRA-2 over the dry regions and PERSIANN-CDR over all wetness regions for the entire period. SM2RAIN-ASCAT and TMPA-3B42RTv7 outperform PERSIANN, PERSIANN-CCS, and PDIR-Now, although PERSIANN over the dry regions outperforms TMPA-3B42RTv7, and PDIR-Now over the wet regions shows higher performance than SM2RAIN-ASCAT and TMPA-3B42RTv7 for the entire period. Consider the seasonal precipitation at the daily time step: among selected GPDs, ERA5 shows the highest performance over moderately wet regions, followed by MSWXv100, where some satellite-based datasets such as PERSIANN-CCS, TMPA-3B42RTv7, and even PERSIANN-CDR, which is a gauge-corrected dataset, show the lowest performance during the winter season for the daily time step. Considering the monthly time step for the entire period, MSWEPv2.8 presents the highest performance in terms of KGE and outperforms IMERGHHFv06 over moderately wet regions, where MSWXv100 comes in with the third-best performance among selected GPDs.

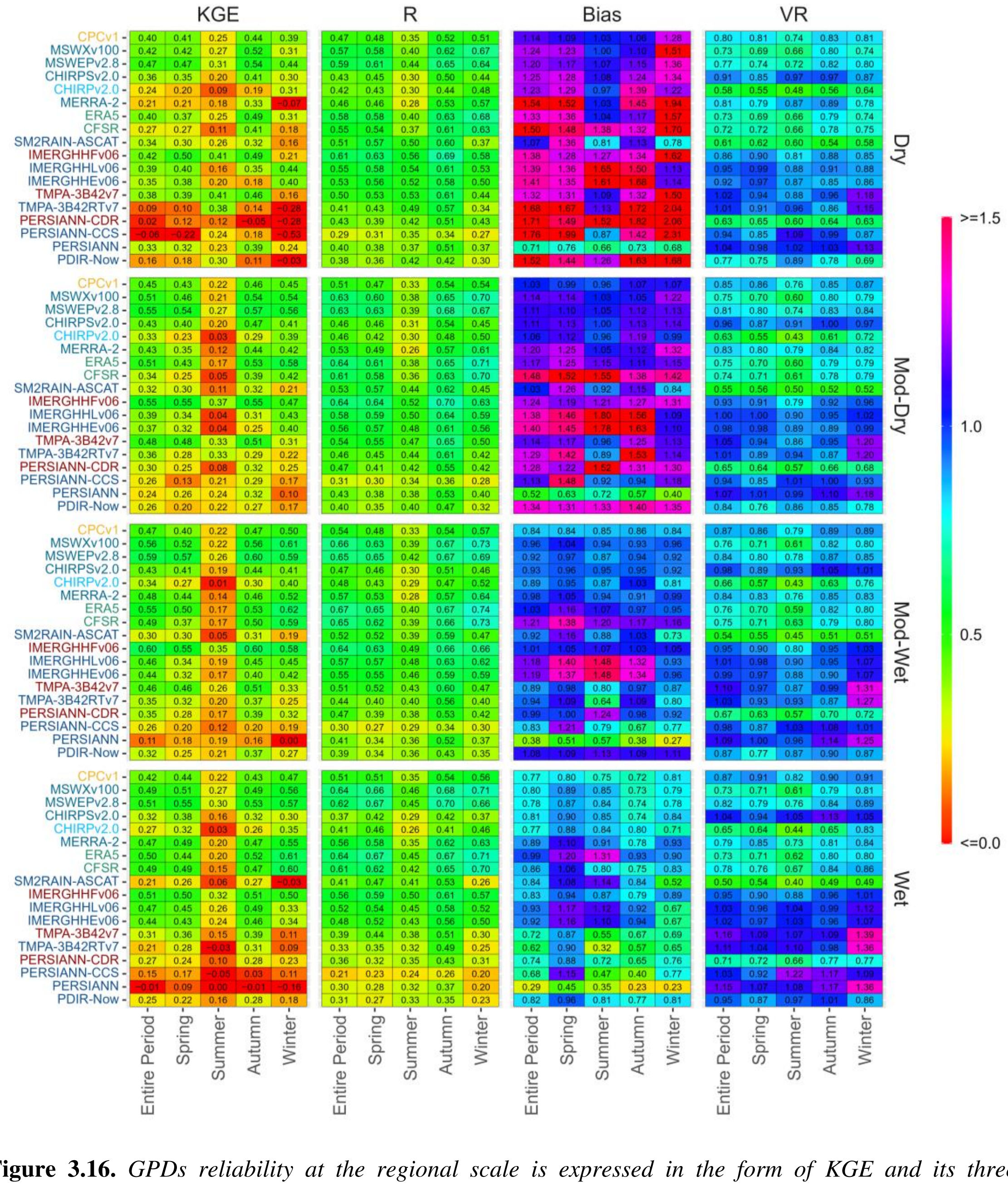

Figure 3.16. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the daily precipitation, considering data from 2015 to 2019 and four seasons. The heat table represent the median of KGE and its components over four distinct wetness classes, respectively. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

CPCv1 also demonstrated higher performance over moderately dry and dry regions and outperformed most of the reanalysis, gauge-corrected, and satellite-based datasets. On the other hand, satellite-based GPDs demonstrate lower performance over different wetness classes, and the lowest performance was obtained by PERSIANN over the wet regions and PERSIANN-CCS over the dry regions for the entire period at the daily time step. Considering the seasonal precipitation for the monthly time step, some multi-source GPDs such as MSWXv100 demonstrate the highest performance over moderately wet regions during the winter season, followed by MSWEPv2.8, ERA5, and IMERGHHFv06, where PERSIANN-CCS over dry regions and PERSIANN over wet regions during the winter season and PERSIANN-CCS and TMPA-3B42RTv7 over wet regions during the summer season demonstrate the lowest performance for the monthly time step.

As mentioned in previous statements, the correlation component highly affects the final KGE result, and any increase in the GPDs' correlation with observed precipitation results in higher final KGE performance. The datasets demonstrate a higher correlation for the monthly time step than the daily time step over different wetness classes. Considering the combined effect of wetness classes and seasonal variability for the daily time step, among selected GPDs, ERA5 shows the highest correlation over moderately wet regions during the winter season, followed by the CFSR and MSWXv100 datasets, while PERSIANN-CCS during the winter and summer seasons and PERSIANN and PDIR-Now during the winter season show the lowest correlation to the observed precipitation over wet regions. Considering the monthly time step, still ERA5 and CFSR show the highest correlation to observed precipitation during the winter season, followed by MSWEPv2.8 over moderately dry regions, where PERSIANN-CSS present the lowest monthly correlation to the observed precipitation compared to other GPDs over dry and wet regions during the winter and summer seasons.

Likewise, the ratio of bias components of KGE shows that this component is unchanged over all wetness classes as the time step shifts from daily to monthly, where most selected GPDs demonstrate overestimation of bias over the dry regions and underestimate bias over the wet regions, which makes sense, and GPDs show higher precipitation in regions that receive insufficient precipitation and lower precipitation amounts in regions that experience significant precipitation. Looking at the entire period for both daily and monthly time steps, among selected GPDs, some satellite-based datasets such as PERSIANN-CCS, TMPA-3B42RTv7, and PDIR-Now; a few gauge-corrected GPDs, including PERSIANN-CDR and MERRA-2; and some reanalysis datasets such as CFSR exhibit the highest overestimation of bias over the dry regions, where PERSIANN over all wetness classes, PERSIANN-CCS, TPMA-3B42(RT)v7, PERSIANN-CDR, CHIRP(S)v2.0, CPCv1, MSWEPv2.8, and MSWXv100 highly underestimate bias over wet regions.

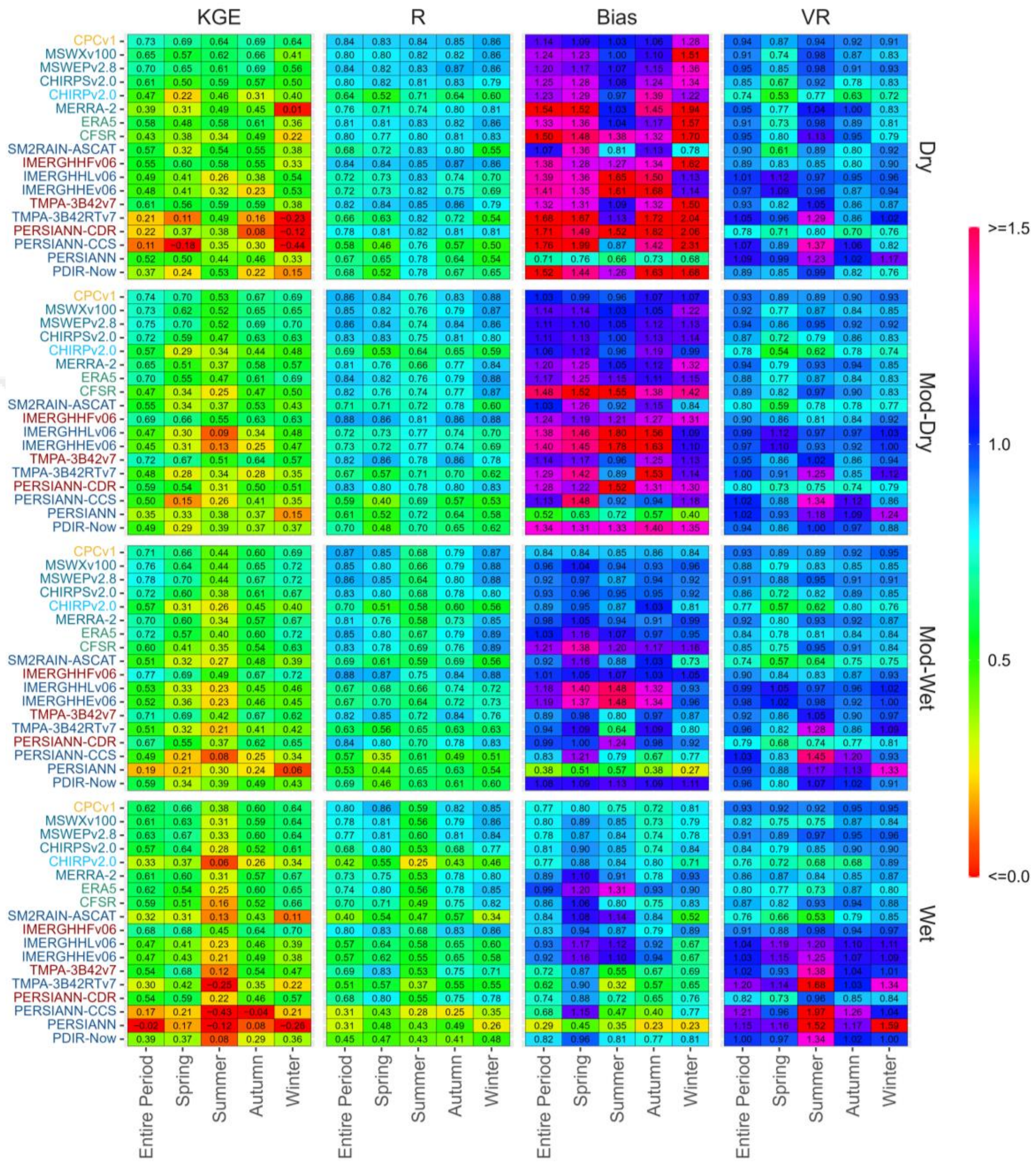

Figure 3.17. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the monthly precipitation, considering data from 2015 to 2019 and four seasons. The heat table represent the median of KGE and its components over four distinct wetness classes, respectively. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Considering the seasonal precipitation over different wetness classes, among selected GPDs, the PERSIANN dataset drastically underestimates the bias over wet and moderately wet regions during the winter and autumn seasons, whereas PERSIANN-CCS, PERSIANN-CDR, and TMPA-3B42RTv7 highly overestimate the bias over dry regions during the winter season, comparatively.

Finally, the variability of GPDs compared to the variability of observed precipitation varies over different wetness classes when considering different time steps and seasons. In general, most selected GPDs demonstrate a higher underestimation of variability ratio over different wetness classes for the daily time step compared to the monthly time step, and the variability ratio of GPDs is highly eye-catching during the spring, summer, and winter seasons for both time steps. Considering the effect of wetness classification and different time scales on GPD variability for the daily time step, among selected GPDs, TMPA-3B42(RT)v7 and PERSIANN highly overestimate the variability ratio over wet and moderately wet regions during the winter season, whereas PERSIANN-CCS shows a higher overestimation of the variability ratio over wet regions during the summer season. In the same way, SM2RAIN-ASCAT over wet and moderately wet regions during the summer season and over wet regions during the autumn season, and CHIRPv2.0 over all wetness classes during the summer season show highly underestimation of variability ratio. For the monthly time step, PERSIANN-CCS and TMPA-3B42 during the summer and PERSIANN during the winter season over wet regions highly overestimate the variability ratio, whereas, among GPDs, CHIRPv2.0 over dry regions during the spring season and SM2RAIN-ASCAT over wet regions during the summer season drastically underestimate the variability ratio, respectively.

3.1.5. Performance of GPDs over different climate zones

Figures 3.18 and 3.19 present reliability of GPDs at the regional scale over seven climate zones at daily and monthly time steps, expressed in the form of KGE and individual components for the entire period (2015–2019) and four seasons. In general, GPDs show distinct performances over different climate zones, considering different time steps and temporal scales (entire period and four seasons). Overall, most multi-source GPDs, ground-based datasets (CPCv1), and some reanalysis datasets depict higher performance compared to other datasets over different climate zones. In the same way, many GPDs show higher performance for the monthly time step than the daily time step, and the summer season is one of the most critical seasons for GPDs' performance, which is more intense over Southeastern Anatolia.

From the results, considering the performance of GPDs over different climate zones for the entire period at the daily time step, some multi-source GPDs, such as MSWEPv2.8, IMERGHHFv06, and MSWXv100 show the highest performance over Southeastern Anatolia, Aegean, Mediterranean, and Marmara regions, where some reanalysis datasets such ERA5 is able to show higher performance than most of the GPDs especially over Southeastern Anatolia where among GPDs, some satellite-based GPDs such as the PERSIANN datasets including its gauge-corrected version (PERSIANN-CDR), TMPA-3B42RTv7, and CHIRPv2.0 present lower performance over regions which experience extremes such Black Sea, Central Anatolia, and Eastern Anatolia. While considering the seasonal effect and climate condition at the same time, some GPDs, such as MSWEPv2.8 during the winter and autumn seasons and ERA5 during the winter season, show the highest performance over Marmara and Aegean regions, which is followed by IMERGHHFv06 over Southeastern Anatolia during the spring season, while some satellite-based GPDs, such as the PERSIANN datasets, IMERGHHL(E)v06, and CHIRPv2.0 over Southeastern Anatolia during the summer, TMPA-3B42RTv7 over Central Anatolia, and SM2RAIN-ASCAT over Eastern Anatolia during the winter, demonstrate the lowest performance compared to other GPDs. Considering the monthly time step for the entire period, some multi-source GPDs such as MERRA-2 and MSWEPv2.8 demonstrate the highest performance among selected GPDs over Southeastern Anatolia, while ERA5, which is one of the reanalysis datasets, comes in with the third-best performance over Southeastern Anatolia, which is followed by CPCv1 over Aegean regions. While looking at the lowest-performing GPDs at the monthly time step, among gauge-uncorrected GPDs, some satellite-datasets such as PERSIANN-CCS and PERSIANN show the weakest performance over the Black Sea region, and many satellite-based GPDs show lower performance over regions that experience extremes, such as the Black Sea, Central Anatolia, Southeastern Anatolia, and Aegean climate zones, respectively. In the same way, considering the combined effect of different seasons and climate zones on GPDs performance for the monthly time step, it is interesting that the highest performance and the lowest performance occur in the same climate zones at different seasons; for example, CPCv1 and TMPA-3B42v7 demonstrate the highest performance among selected GPDs over Southeastern Anatolia during the autumn season, whereas some gauge-corrected GPDs such as PERSIANN-CDR and satellite-based datasets, including PRESIANN-CCS and IMERGHHEv06, demonstrate the lowest performance over southeastern Anatolia during the summer season. In the same way, during the winter season, some GPDs such as SM2RAIN-ASCAT over Eastern Anatolia and MSWEPv2.8 over the Marmara region show the lowest and highest performance, respectively.

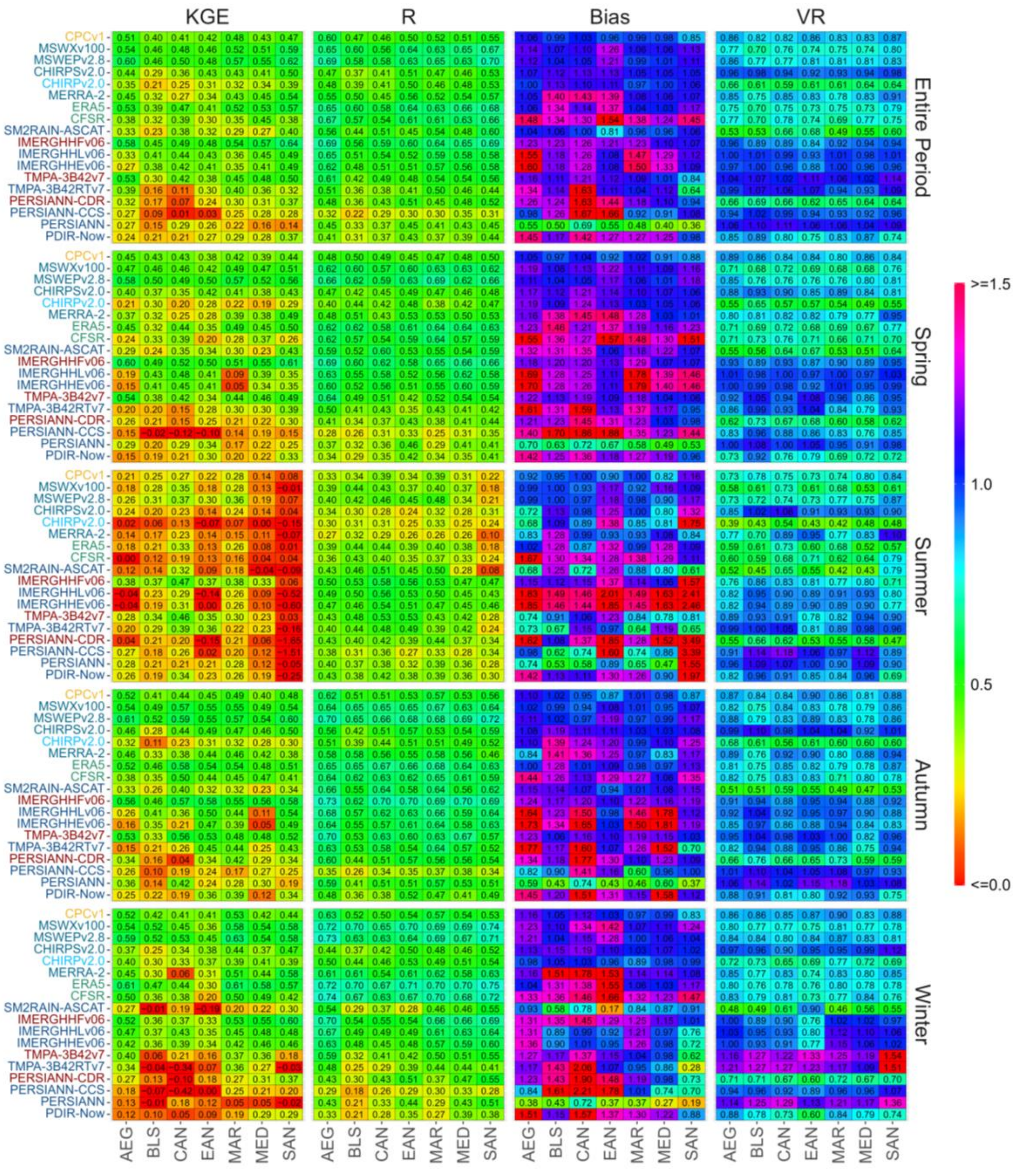

Figure 3.18. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the daily precipitation, considering data from 2015 to 2019 and four seasons. The heat table represent the median of KGE and its components over seven distinct climate zones, respectively. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

It's worth mentioning that looking at the correlation of selected GPDs to observed precipitation over different climate zones, and individual time steps, GPDs demonstrate higher correlation for the monthly time step than for the daily time step, and the summer and winter seasons are relatively critical for GPDs correlation to the observed, comparatively. Taking the combined effect of climate zones and different temporal scales for the daily time step, ERA5 shows the highest correlation over southeastern Anatolia, which is followed by CFSR over the Aegean region during the winter season, while MSWXv100 over Southeastern Anatolia and MSWEPv2.8 over the Aegean come in with the third and fourth best correlations during the winter season. Where the lowest correlation was obtained by SM2RAIN-ASCAT, MERRA-2, and MSWXv100, and ERA5 over Southeastern Anatolia during the summer season. For the monthly time step, IMERGHHFv06 and CPCv1 show the highest monthly correlation to the observed over Southeastern Anatolia and the Aegean regions for the winter season and entire period. Where TMPA-3B42v7 over the southeastern Anatolia region during the winter season and PERSIANN-CCS over the Aegean region during the spring season show the lowest correlation to observed precipitation compared to other GPDs.

It is also important to see the GPDs bias distribution compared to observed precipitation considering different climate zones and individual temporal scales such as the entire period and four seasons at the daily and monthly time steps, which is expressed in the form of the ratio of bias component of KGE. As it was mentioned earlier, the ratio of bias stays unchanged when the time step changes from daily to monthly, where there is some variation of bias considering the seasonal variability of precipitation and GPDs, which seems to show higher underestimations of bias during the rainy seasons such as the spring and winter and overestimations of bias for the summer and autumn seasons over distinct climate zones for both time steps. Considering the GPDs ratio of bias distribution over different climate zones for the entire period at both time steps, among selected GPDs, satellite-based GPDs either highly overestimate or drastically underestimate the ratio for bias. For example, PERSIANN-CCS shows the highest overestimation of bias over Central Anatolia and Eastern Anatolia, whereas TMPA-3B42RT demonstrates a relatively higher ratio of bias over Central Anatolia, which is followed by IMERGHHL(E)v06 over the Aegean region. However, some exceptions exist; for example, PERSIANN-CDR, which is known as one of the gauge-corrected GPDs and supposed to show lower bias, but the dataset highly overestimates bias over Central

Anatolia, which is more intense than the CFSR ratio of bias over Eastern Anatolia and Aegean regions or IMERGHHL over Marmara region in terms of magnitude.

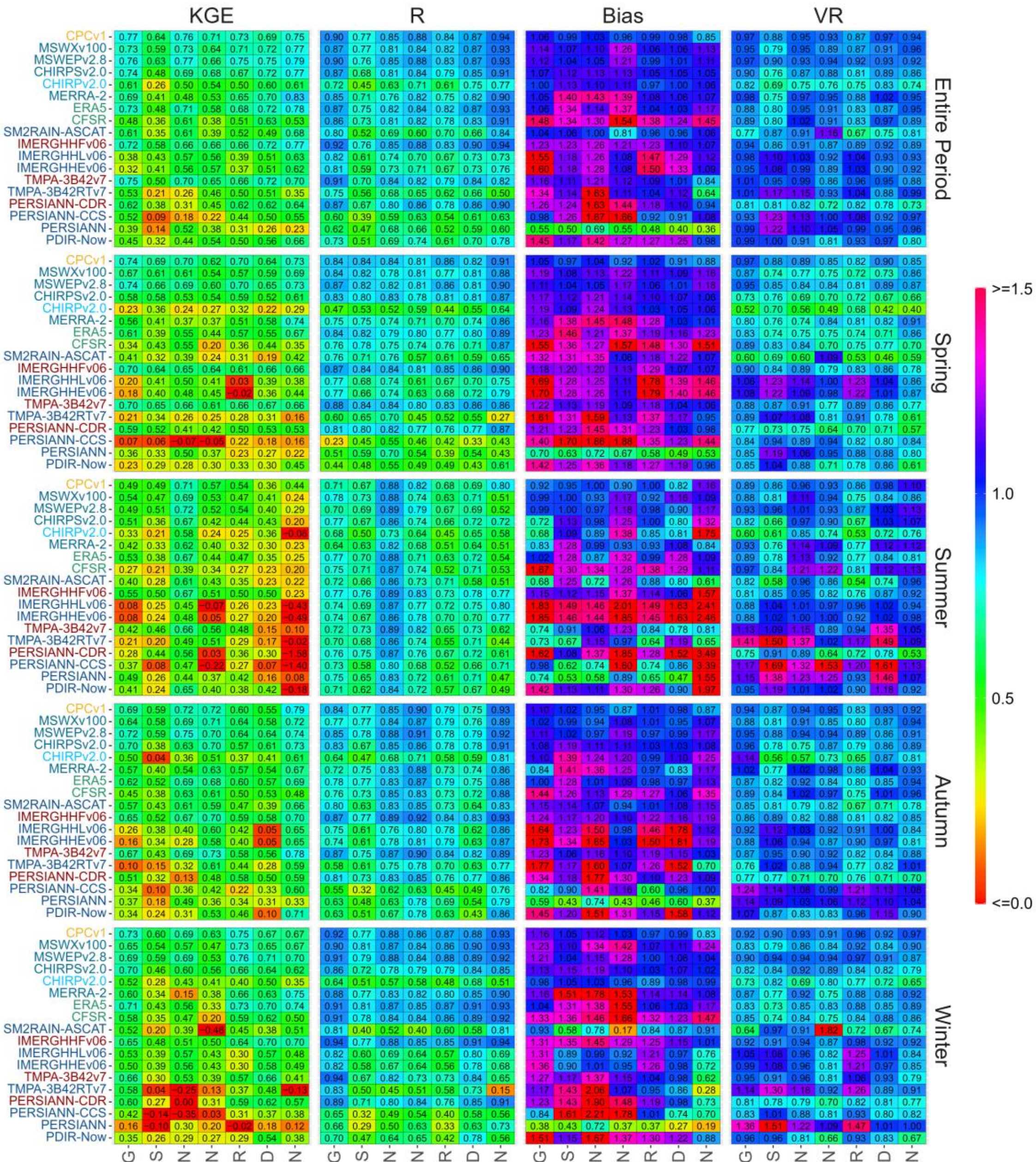

Figure 3.19. *GPDs reliability at the regional scale is expressed in the form of KGE and its three components (R, Bias, and VR) for the monthly precipitation, considering data from 2015 to 2019 and four seasons. The heat table represent the median of KGE and its components over seven distinct climate zones, respectively. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

On the other hand, among selected GPDs, it is again satellite-based datasets that highly underestimate the ratio of bias. For example, the PERSIANN dataset, except for the Mediterranean region, shows the highest underestimation of bias over most climate zones, followed by TMPA-3B42RTv7 over Southeastern Anatolia and SM2RAIN-ASCAT over Eastern Anatolia regions. Considering the effect of climate conditions and seasonal variability of precipitation for both time steps, among multi-source GPDs, PERSIANN-CDR highly overestimates bias over Southeastern Anatolia, followed by some satellite-based GPDs, such as PERSIANN-CCS, PDIR-Now, and IMERGHHL(E)v06, over the Southeastern Anatolia region during the summer season, while SM2RAIN-ASCAT over Eastern Anatolia during the winter season, PERSIANN datasets over almost all regions and all seasons highly underestimate the ratio of bias, whereas the reset of GPDs shows slightly overestimation or underestimating bias over different climate zones, respectively.

Considering the variability ratio of GPDs over different climate zones and different temporal scales at both time steps, the variability ratio is underestimated by many GPDs for the daily time step, especially during the spring and summer seasons, compared to the monthly time step. Looking at the combined effects of climate variability and different temporal scales (entire period and four seasons) for the daily time step, some gauge-uncorrected GPDs such as CHIRPv2.0 highly underestimate the variability of precipitation over different climate zones, which is intense over Aegean region during the summer seasons, followed by some other GPDs, such as PERSIANN-CDR, MSWXv100, SM2RAIN-ASCAT, and PERSIANN datasets, over different climate zones, especially during the spring and summer seasons, whereas some GPDs, such as TMPA-3B42(RT)v7and PERSIANN, show the highest overestimation of the variability ratio over Southeastern Anatolia during the season at the daily time step. For the monthly time step, CHIRPv2.0 still shows highly underestimated variability over different climate zones, which is intense over Southeastern Anatolian and Mediterranean regions during the spring season, followed by SM2RAIN-ASCAT and PERSIANN-CDR over different climate zones, especially during the spring and summer seasons, comparatively.

3.1.6. GPDs frequency and detectability analysis

3.1.6.1. Intensity-frequency analysis and detectability of GPDs over entire region

Figure 3.20 presents the results of PDFs of daily precipitation with different intensities at the regional scale for the entire region, obtained from observed data and 18 GPDs for the entire period and four seasons. Overall, observed precipitation and selected

GPDs show that the frequency of precipitation occurrence decreases as the intensity increases. From the results, it demonstrates that 80% of total observed precipitation events are within the 0–1 mm/day range for the entire period, where this amount decreases to 75% during the spring season and reaches its maximum of 89% during the summer season, while 82% of total observed precipitation is placed within the intensity threshold (0–1 mm/day) during the autumn season, and finally it reaches its lowest frequency (70%) during the winter season. Furthermore, observed precipitation shows a lower frequency (less than 15%) for light (1–5 mm/day) and moderate (5–20 mm/day) precipitation over the entire period and each season, while the lowest frequency of precipitation occurrences is obtained for heavy (20–40 mm/day) and extreme (more than 40 mm/day) precipitation events. As a result, GPDs demonstrate different frequencies of precipitation occurrences over distinct precipitation intensity thresholds. For example, PERSIANN displays higher frequency for precipitation less than 1 mm/day, which underestimates the frequency of observed precipitation for light, moderate, heavy, and extreme precipitation. On the other hand, PDIR-Now shows a lower frequency of precipitation within 0–1 mm/day and overestimates the frequency of observed precipitation for light precipitation. CHIRPv2.0 underestimates the observed precipitation frequency for intensities less than 1 mm/day and shows higher frequency for light precipitation, which demonstrates a close frequency for moderate precipitation, and again underestimates the observed frequency for heavy and extreme precipitation. PERSIANN-CDR overestimates heavy precipitation during the spring season, and SM2RAIN-ASCAT overestimates light precipitation during the summer season, while SM2RAIN-ASCAT also overestimates light precipitation for the rest of the time periods. Moreover, the frequency of observed precipitation for extreme precipitation is only overestimated by IMERGHHL(E)v06 for all temporal periods (except the summer season), where ERA5 shows higher frequency of extreme precipitation during the summer season. Hence, a high overestimate/underestimate of observed precipitation frequency by GPDs within a certain intensity threshold directly affects their detectability strength, especially if larger frequency bias occurs for precipitation of higher magnitude.

The Hanssen-Kuiper Skill (HKS) is utilized to assess the detection ability of GPDs for different precipitation events occurrence for the entire period and individual seasons at a daily time step (Figure 3.21).

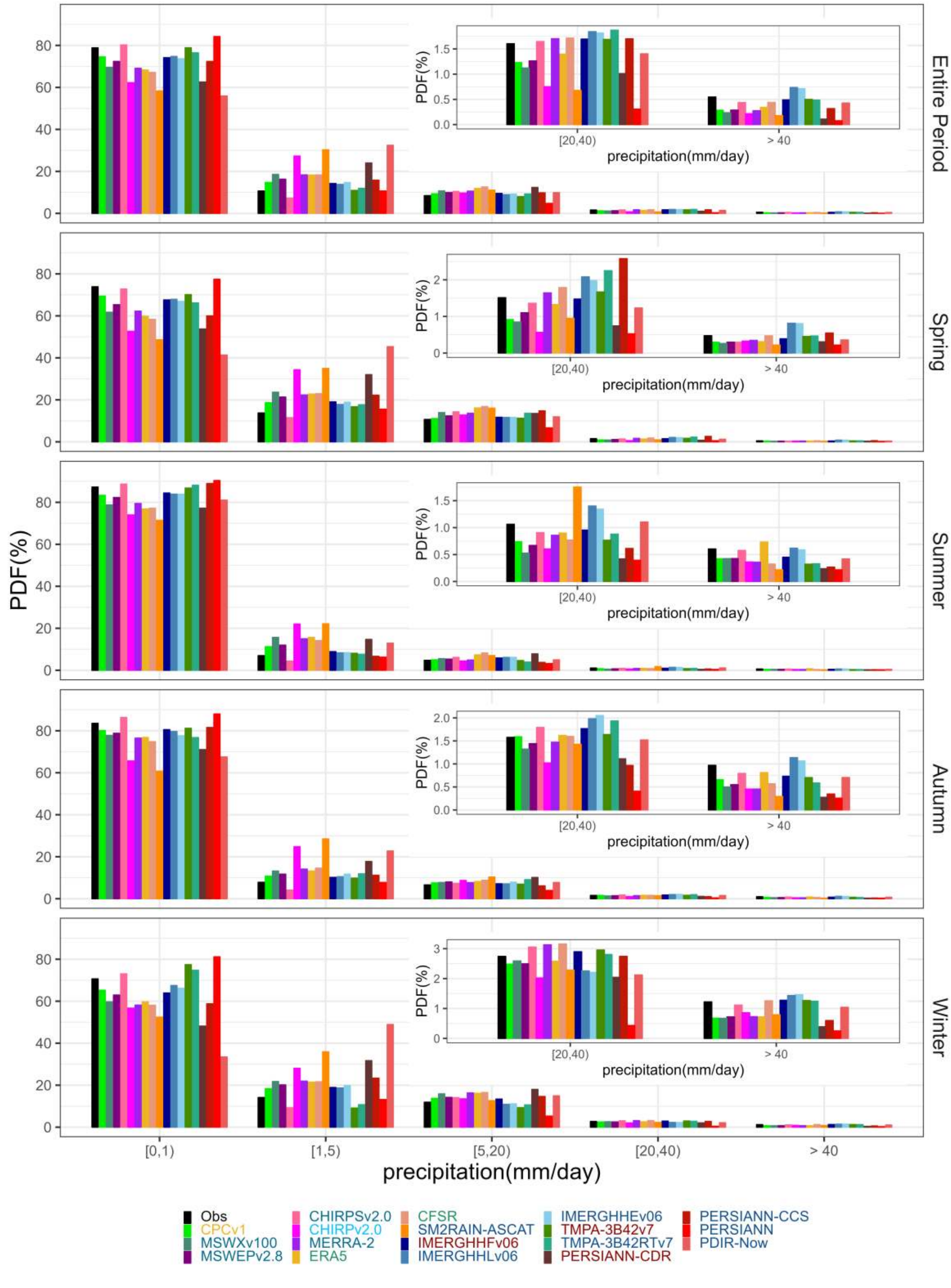

Figure 3.20. *Daily precipitation frequency for five different intensities from observed and 18 GPDs for the entire period and four seasons. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Overall, GPDs show higher detectability for catching precipitation events with different intensities during the autumn and present relatively lower detection ability in the summer seasons. Furthermore, the detection strength of GPDs decreases as the precipitation intensity increases, which is generally the case in literature. This is due to the classification of several intensity classes, which makes it difficult to differentiate between them rather than a straightforward rain/no rain scenario. This division gets even more problematic when the occurrence probability of a certain event is seen much more rarely compared to other classes (heavy and extreme events), as can be seen from Figure 3.20. As a result, detecting a rarely occurring very intense precipitation event has a weaker performance when assessed against a more frequently occurring light storm. Furthermore, many GPDs show comparatively higher detectability of moderate precipitation (5–20 mm/day) than light storms (1–5 mm/day). This can be attributed to the fact that the underestimation/overestimation of observed precipitation frequency over a frequently occurring light storm is more sensible than that frequency bias over rarely occurring classes (Figure 3.20). In other words, the same amount of frequency bias for a certain GPD has a higher effect on high-frequency events than on lower-frequency events. In general, considering the detection ability of GPDs for the entire period and four seasons, MSWEPv2.8 demonstrates the highest detectability of no-precipitation events and outperforms all GPDs during the autumn season and entire period, followed by ERA5 during the autumn and winter seasons, and MSWXv100 during the autumn season and entire periods, while CHIRPSv2.0 presents the lowest detectability of no-precipitation events among selected GPDs during the summer season, followed by PERSIANN, SM2RAIN-ASCAT, and TMPA-3B42(RT)v7 datasets during the winter season. In the same way, still MSWEPv2.8 presents the highest detectability of light precipitation among selected GPDs during the entire period, autumn, and spring seasons, followed by PDIR-Now during the summer season and MSWXv100 for the entire period, while CHIRPSv2.0, and TMPA-3B42(RT)v7 show the lowest detectability of light precipitation during the winter and autumn seasons. It is important to see that ERA5 shows the highest detection strength for moderate precipitation during the winter season, where MSWXv100 during the winter season and SM2RAIN-ASCAT during the autumn season come in with the second- and third-best detection abilities, respectively, while some GPDs, such as CHIRP2.0 and MERRA-2 in the summer and PERSIANN during the entire period, spring and winter seasons, show the lowest detectability of moderate precipitation compared to other GPDs. Furthermore, it looks like both reanalysis datasets show the highest ability of detecting heavy precipitation events during the winter season, followed by IMERGHHFv06 and TMPA-3B42(RT)v7 in the autumn and MSWXv100, MERRA-2, and MSWEPv2.8 during the winter season, while CHIRPv2.0, SM2RAIN-ASCAT, and MSWXv100 in the summer and SM2RAIN-ASCAT and PERSIANN in the winter demonstrate the lowest detectability of heavy precipitation events. Finally, it is the IMERGHHv06 final and near real-time datasets that show the highest detection ability of extreme precipitation for the entire period and different seasons, and CHIPRv2.0 and PERSIANN-CDR demonstrate the lowest detection abilities of extreme precipitation, respectively.

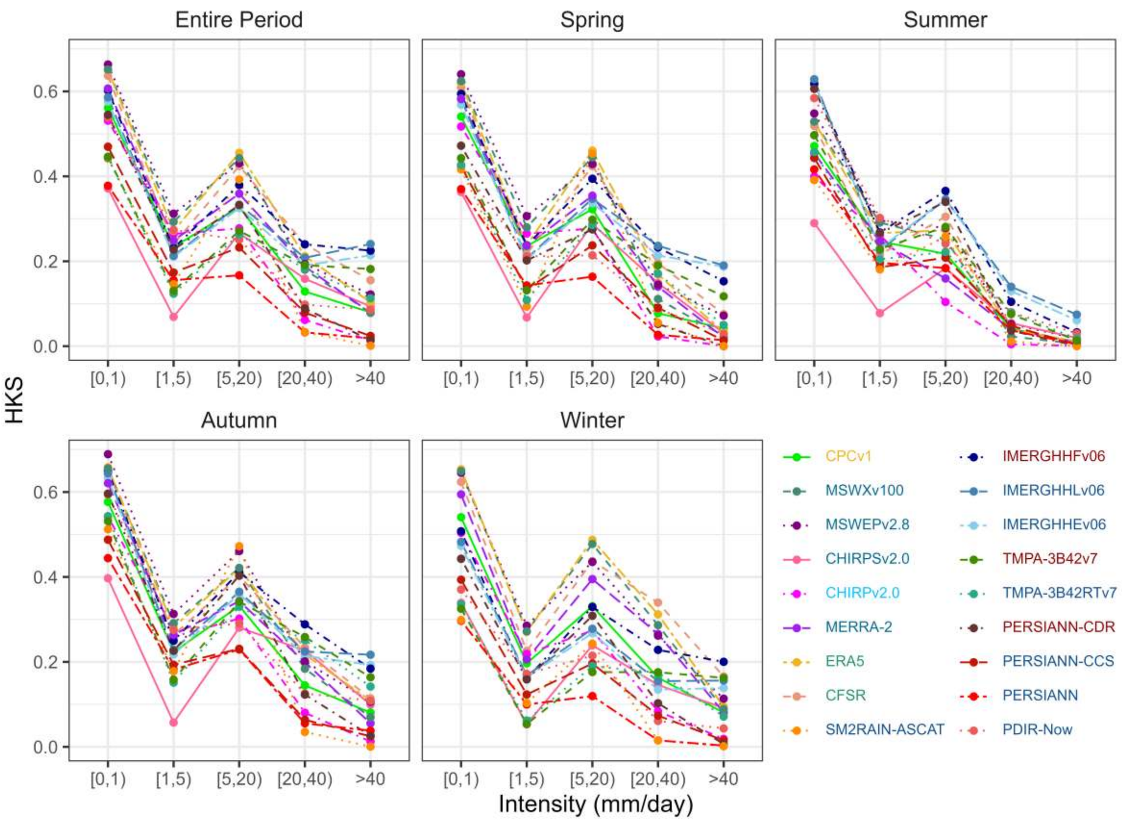

Figure 3.21. *Detectability of GPDs at the regional scale expressed in the form of Hanssen-Kuiper Skill score for the entire period and four seasons. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

3.1.6.2. Intensity-frequency analysis and detectability of GPDs over elevation ranges

Figure 3.22 shows the frequency of different precipitation events from observed precipitation and selected GPDs for four elevation ranges at the regional scale, considering the entire study period and seasonal precipitation. In general, observed precipitation and GPDs follow the same trend of PDFs as the entire region (section

3.1.6.1), where their frequency decreases as the intensity increases. Some GPDs, including CHIRPv2.0, SM2RAIN-ASCAT, PDR-Now, and PERSIANN-CDR, often underestimate the observed frequency for the intensity between 0 and 1 mm/day almost over the entire period and each season for all elevation ranges. Moreover, the overestimation of frequency by GPDs occurs for light precipitation (1–5 mm/day).

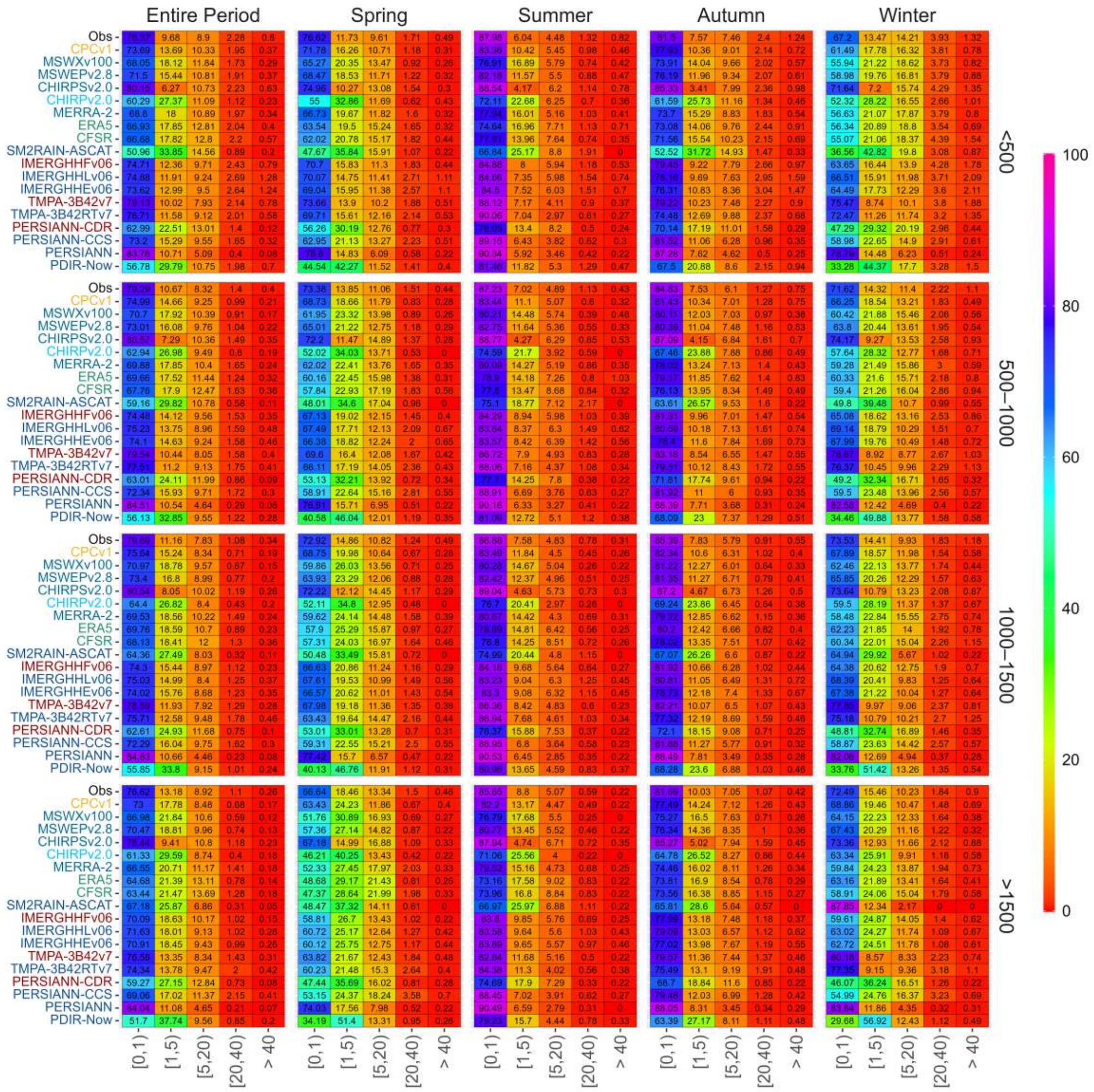

Figure 3.22. *Daily precipitation frequency for five different intensities from observed and 18 GPDs for the entire period and four seasons over four elevation ranges. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Considering the seasonal precipitation, there is a large difference between GPDs frequencies for light precipitation during the winter season. Among GPDs, PDIR-Now is highly affected by elevation and shows a higher frequency over highly elevated regions (areas with more than 1500 m), whereas SM2RAIN-ASCAT overestimates the observed frequency for light precipitation in low-elevation lands (with less than 500 m), particularly during the winter season.

Figure 3.23 presents the detection ability of GPDs over four elevation ranges, which is expressed in the form of Hanssen-Kuiper Skill (HKS) score for the entire study period and seasonal precipitation at the regional scale for different precipitation intensities. Overall, GPDs show slightly higher detectability strengths for low-elevated lands (with elevation of less than 500 m), and their detectability decreases as the elevation increases. Considering the combined effect of elevation change and temporal such as entire period and different season on GPDs detectability, MSWEPv2.8 demonstrate the highest detection ability of no-precipitation over areas having elevation less than 500 m and 500-100 m during the entire period, autumn and winter seasons, followed by MSWXv100, CFSR, and ERA5 over area with elevation less than 500 m during the autumn season, while some satellite-based GPDs such as SM2RAIN-ASCAT and PERSIANN during the winter season, and some multi-source GPDs, including TMPA-3B42v7 and CHIRPSv2.0 during the winter and summer season show the lowest detection ability of no-precipitation events over regions with elevation more 1500 m compared to other selected datasets. In the same way, PDIR-Now shows the highest skill of detecting light precipitation over areas with elevation more than 1500 m and 1000-1500 m during the summer season, while MSWXv100 during the summer season and MSWEPv2.8 during the autumn season come in with the second and third best detection abilities of light precipitation over areas with elevation more than 1500 m, while some GPDs, such as CHIRPSv2.0 over areas with elevation less than 500 m and SM2RAIN-ASCAT and TMPA-3B42(RT) over regions having elevation more than 1500 m, present the lowest detectability of light precipitation compared to other selected GPDs during the winter autumn seasons. Furthermore, it’s worth mentioning that some gauge-uncorrected GPDs such as SM2RAIN-ASCAT shows the highest detection ability of moderate precipitation over areas with elevation less than 500 m and 500–1000 m during the autumn season, followed by ERA5 and MSWXv100 over areas with elevation 500–1000 m and 1000–1500 m during the winter season and MSWEPv2.8 over areas having elevation 500–1000 m during the autumn season, while SM2RAIN-ASCAT over areas with elevation more than 1500 m during the winter season, CHIRPv2.0 over all elevation ranges during the summer season and PERSIANN dataset over different elevation range and seasons shows drastically lower detectability of moderate precipitation compared to other selected GPDs.

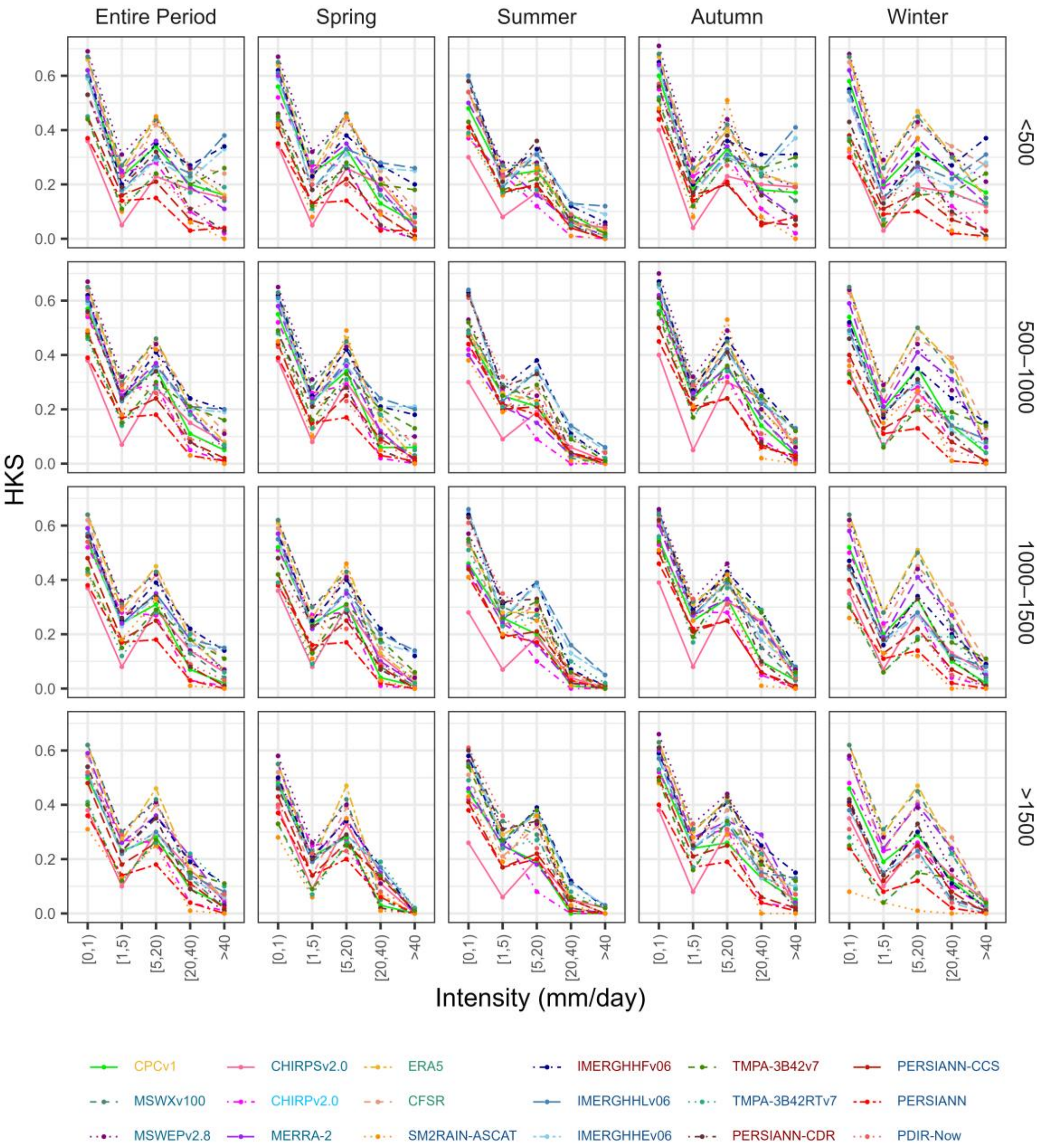

Figure 3.23. *Detectability of GPDs at the regional scale expressed in the form of Hanssen-Kuiper Skill score for the entire period and four seasons over four elevation classes. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The highest detection ability of heavy precipitation is obtained by CFSR over areas with elevations of 500–1000 m, followed by ERA5 and MSWXv100 during the winter season, where IMERGHHFv06 is also able to present a higher detection ability over areas with elevations less than 500 m during the autumn season, while CHIRPv2.0, SM2RAIN-ASCAT, and sometimes CPCv1 show the lowest heavy precipitation detection over different elevation ranges considering the entire period and four seasons. Finally, it is the IMERGHHv06 final and its near real-time datasets that mostly show the highest detectability of extremes precipitation, especially over regions with elevation less than 500 m for different temporal, which is followed by TMPA-3B42v7 dataset, while CFSR, CHIRP(S)v2.0, CPCv1, and ERA5 show the lowest detection ability of extremes precipitation compared to other GPDs over different elevation ranges and temporal, respectively.

3.1.6.3. Intensity-frequency analysis and detectability of GPDs over wetness classes

Figure 3.24 presents the results of PDFs of daily precipitation events over different intensities for four wetness classes at the regional scale, obtained from observed data and 18 GPDs for the entire period and four seasons. Overall, the frequency of observed and GPDs decreases as the intensity increases. Many GPDs and observed precipitation show a high frequency of no-rainy days (0–1 mm/day), which becomes more intense for the summer season, and some GPDs overestimate light precipitation frequency over all wetness classes, comparatively. Generally, most GPDs show a different frequency of no-rain and light precipitation over wet regions for the entire period and each season. Furthermore, some GPDs, including CHIRPv2.0, MERRA-2, ERA5, CFSR, SM2RAIN-ASCAT, PERSIANN-CDR, and PDIR-Now, show higher rainy days and significantly underestimate no-rainy days (0–1 mm/day), especially over moderately dry and moderately wet regions during the winter season. Hence, GPDs show relatively close frequency trend for the spring and winter seasons as well as summer and autumn seasons over all wetness classes. Considering the entire study period and four seasons, wet regions are identified as the most critical wetness class for GPD frequency analysis.

Figure 3.25 displays the detection ability of 18 GPDs for five distinct precipitation intensity thresholds at the regional scale over four wetness classes for the entire study period and four seasons. The detectability strength of many GPDs decreases as the intensity increases with a positive slope transition from light to moderate precipitation. In other words, GPDs are able to identify no-rainy days with higher detectability, and their detectability significantly decreases for light precipitation and increases for moderate precipitation, then extends with a negative or positive slope for heavy and extreme precipitation events over all wetness classes and time scales. However, some GPDs are grouped exceptionally, including PERSIANN, which shows a constant negative slope for precipitation detection from the no-rain threshold to extreme precipitation.

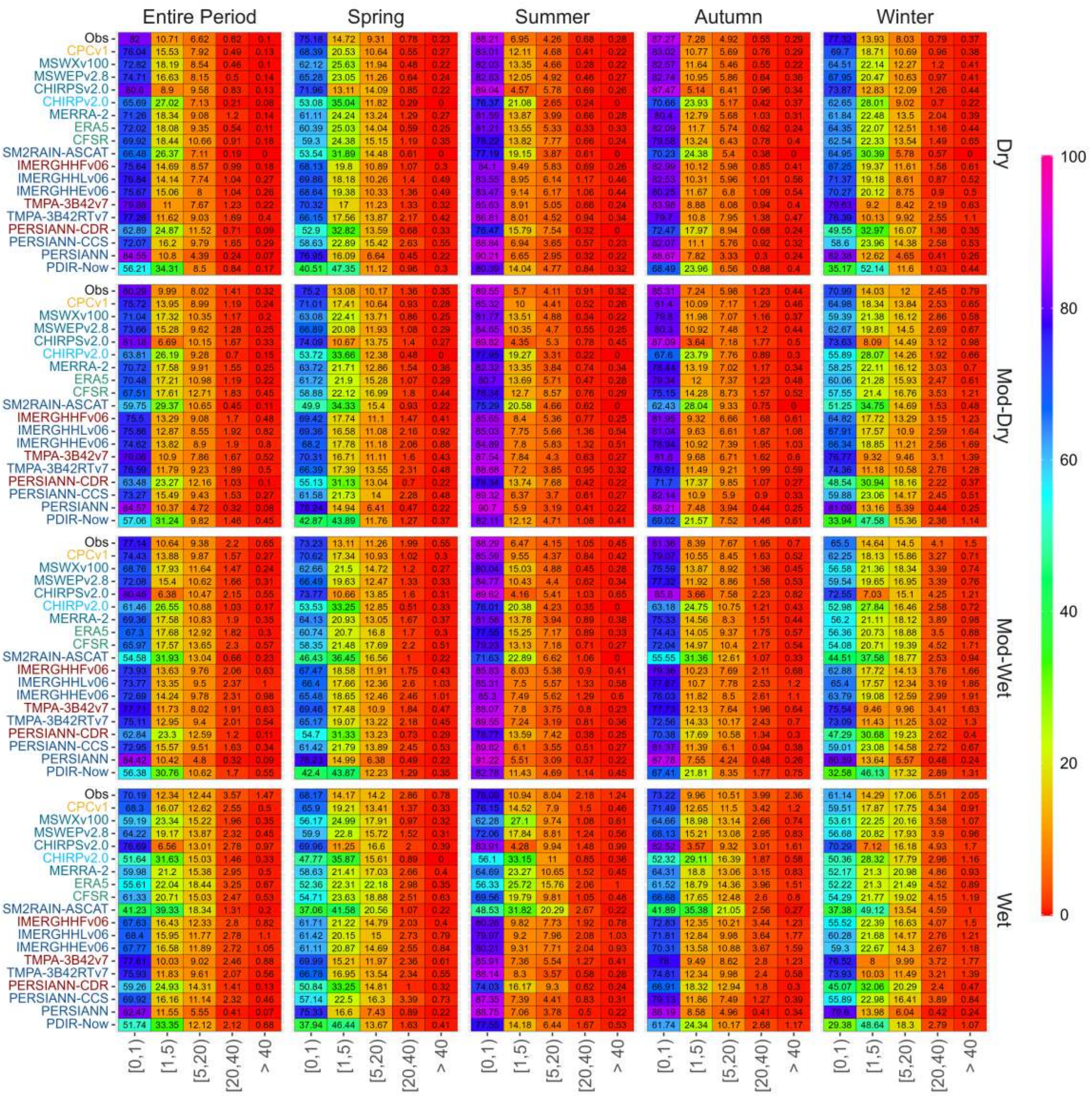

Figure 3.24. *Daily precipitation frequency for five different intensities from observed and 18 GPDs for the entire period and four seasons over four wetness classes. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

In general, GPDs show higher detectability during the autumn seasons, followed by the winter and spring seasons, and GPDs' detectability is close to each other and weaker during the summer season over all wetness classes. Considering the detection ability of GPDs over all four wetness classes under different temporal scales such as the entire period and four seasons, MSWEPv2.8 dataset shows the highest detection ability of no-

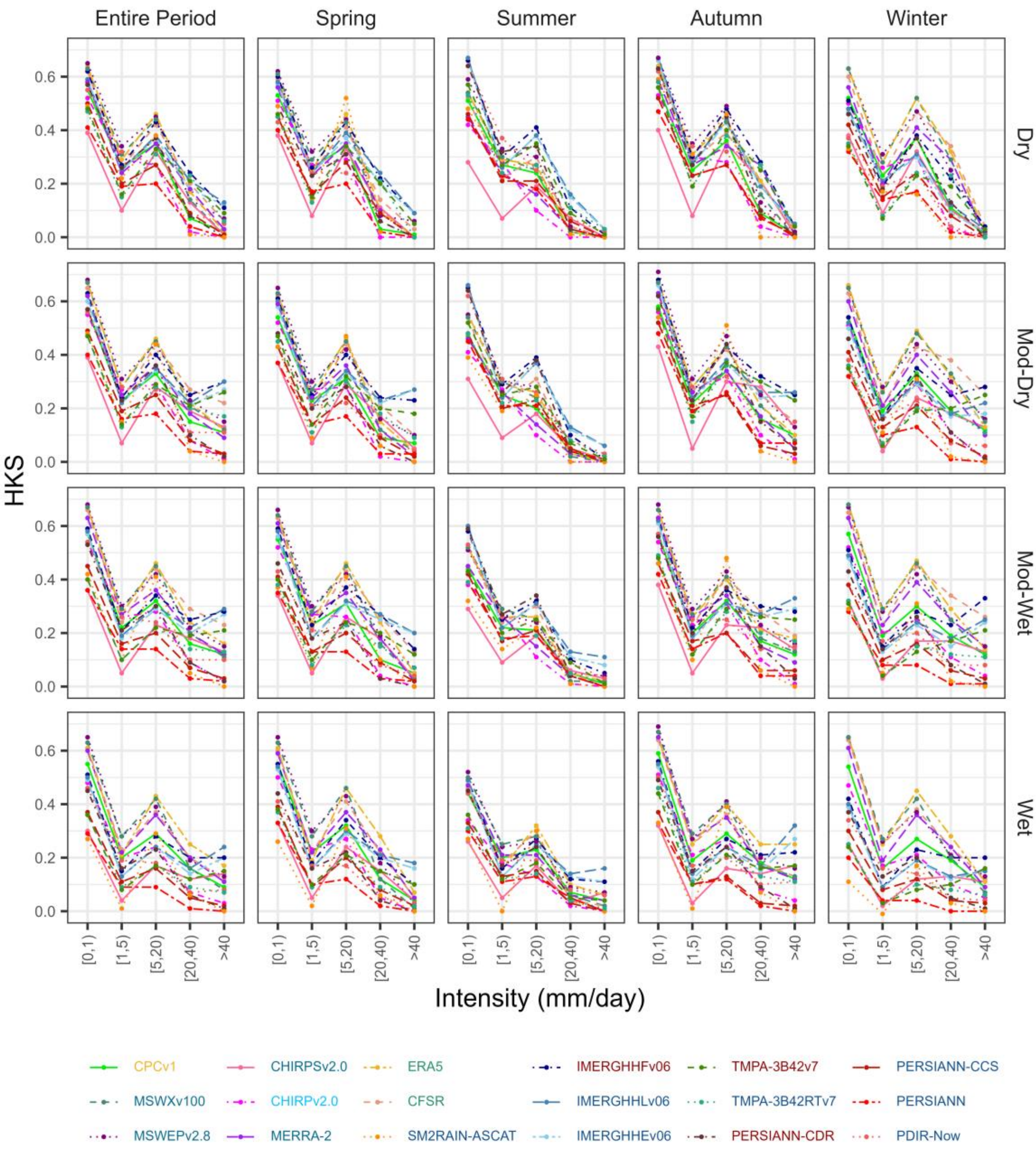

Figure 3.25. *Daily precipitation frequency for five different intensities from observed and 18 GPDs for the entire period and four seasons over four elevation ranges. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

precipitation event over moderately dry and wet region during the autumn season, where ERA5 by showing higher detection ability of no-precipitation over moderately dry and moderately wet region come in with the second-based detection ability during the autumn and winter seasons followed by IMERGHHFv06 over moderately dry regions during the autumn season, while some GPDs such as SM2RAIN-ASCAT, PERSIANN, TMPA-

3B42(RT)v7 and CHIRPSv2.0 show lowest detection ability of no-precipitation event over wet regions especially during the summer season. Some GPDs such as PDIR-Now and MSWEPv2.8 outperform other datasets and show the highest detection ability of light precipitation for the entire period and summer and autumn seasons, followed by PERSIANN-CDR during the summer season and ERA5 during the autumn season over dry regions, whereas some datasets such as SM2RAIN-ASCAT, CHIRPSv2.0, TMPA-3B42(RT)v7, and PERSIANN show lower detection abilities of light precipitation compared to other GPDs, especially over wet regions, for different seasons. In the same way, it is ERA5 and MSWXv100 during the winter season and MSWEPv2.8 and IMERGHHFv06 during the autumn season that show higher detection abilities of moderate precipitation over dry and some time over moderately dry regions, while PERSIANN, TMPA-3B42RTv7, CHIRPv2.0, and PERSIANN-CCS show lower detectability of moderate precipitation, especially over wet regions, considering different seasons. Furthermore, among selected GPDs, CFSR was able to show the highest detection ability of heavy precipitation during the winter season over different wetness classes except for the wet regions, which is followed by MSWXv100 over moderately dry and ERA5 over dry and moderately dry regions, where some GPDs, such as CHIRPv2.0 during the spring and summer over dry and moderately dry regions, PERSIANN over wet regions during the winter, and SM2RAIN-ASCAT over dry regions during the autumn season show lower detectability of heavy precipitation compared to other selected GPDs. Finally, the IMERGHHv06 datasets were able to demonstrate higher detection of extreme precipitation over different wetness classes and seasons, while CFSR and CHIRPSv2.0 show lower detectability of extreme precipitation, especially over dry regions for different seasons.

3.1.6.4. Intensity-frequency analysis and detectability of GPDs over climate zones

Figure 3.26 presents the frequency of observed and GPDs daily precipitation for five intensity thresholds at the regional scale over seven climate zones for the entire study period and four seasons. Many GPDs, including MSWXv100, CHIRPv2.0, MERRA-2, ERA5, CFSR, PERSIANN-CDR, and PDIR-Now, underestimate no-rainy days over the Black Sea region for all temporal scales. Moreover, GPDs and observed precipitation show the highest frequency of no-rain over Southeastern Anatolia during the summer season.

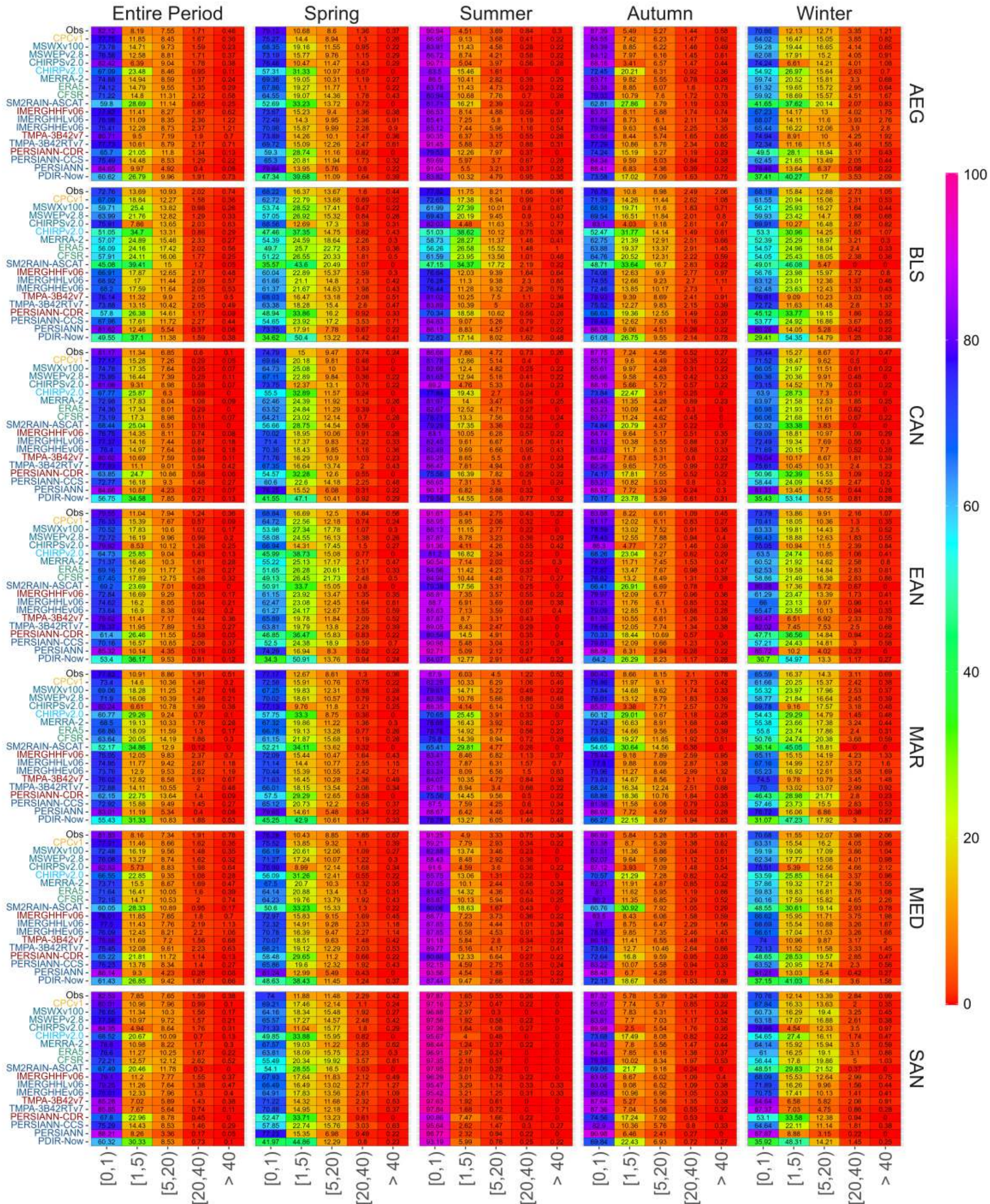

Figure 3.26. *Daily precipitation frequency for five different intensities from observed and 18 GPDs for the entire period and four seasons over seven climate zones. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The CHIRPSv2.0 dataset consistently underestimates the frequency of light observed precipitation almost across all climate zones and temporal, whereas reanalysis datasets such as ERA5 and CFSR overestimate the frequency of moderate observed precipitation over the Black Sea, and it becomes more intense over highly elevated lands such as Eastern Anatolia during the winter season, where many satellite-based GPDs underestimate moderated precipitation across all climate zones.

GPDs demonstrate different frequencies for the selected five precipitation intensity thresholds, and even a particular GPD may demonstrate a different frequency for a specific intensity when the temporal window is changed. For example, CHIRPv2.0 over Marmara underestimates the observed no-rainy days' frequency for the entire study period and spring season and overestimates for the rest of the time. The detection ability of each GPD over different climate zones at the regional scale is expressed in the form of Hanssen-Kuiper Skill (HKS) score for five daily precipitation intensity thresholds considering the entire study period and four seasons (Figure 3.27). Generally, GPDs follow the same trend of detectability as over wetness classes and elevation ranges, which means many of them show higher detectability on no-rain days, then the detectability decreases for light precipitation and extends to moderated precipitation with a positive slope, and finally their detectability decreases for heavy and extreme precipitation, where there is some negative and positive slope of detectability between heavy and extreme events due to the classification demands for precipitation intensities considering five different thresholds instead of a single rain/no-rain classification. However, some exceptions do exist; there is higher detection ability of extreme precipitation events than heavy for most of the GPDs over some climate zones, such as Aegean, Marmara, Mediterranean, and Southeastern Anatolia, over individual temporal scales, such as the entire period and four different seasons, where the detection ability of CHIRPSv2.0 dataset is mostly eye-catching over Southeastern Anatolia during the autumn season, where the dataset shows higher detection ability of no precipitation and the detectability decreases for the light precipitation, then it increases up to the heavy precipitation and again shows lower detection ability for the extreme precipitation. Considering the detection ability of GPDs over seven climate zones under different temporal constraints, such as the entire period and four seasons, some GPDs, such as CFSR, MSWEPv2.8, ERA5, and IMERGHHF(L)v06, show a higher detection ability of no-precipitation events over Southeastern Anatolia and Aegean, especially for the autumn season and entire period, whereas some datasets, including SM2RAIN-ASCAT, MERRA-2, PERSIANN, and CHIRPv2.0, show a lower detectability of no-precipitation events over regions experiencing extremes such as Southeastern Anatolia, Black Sea, and Central Anatolia during the summer and winter seasons. While PDIR-Now, PERSIANN-CDR, and MSWXv100 over Eastern and Southeastern Anatolia during the summer season and

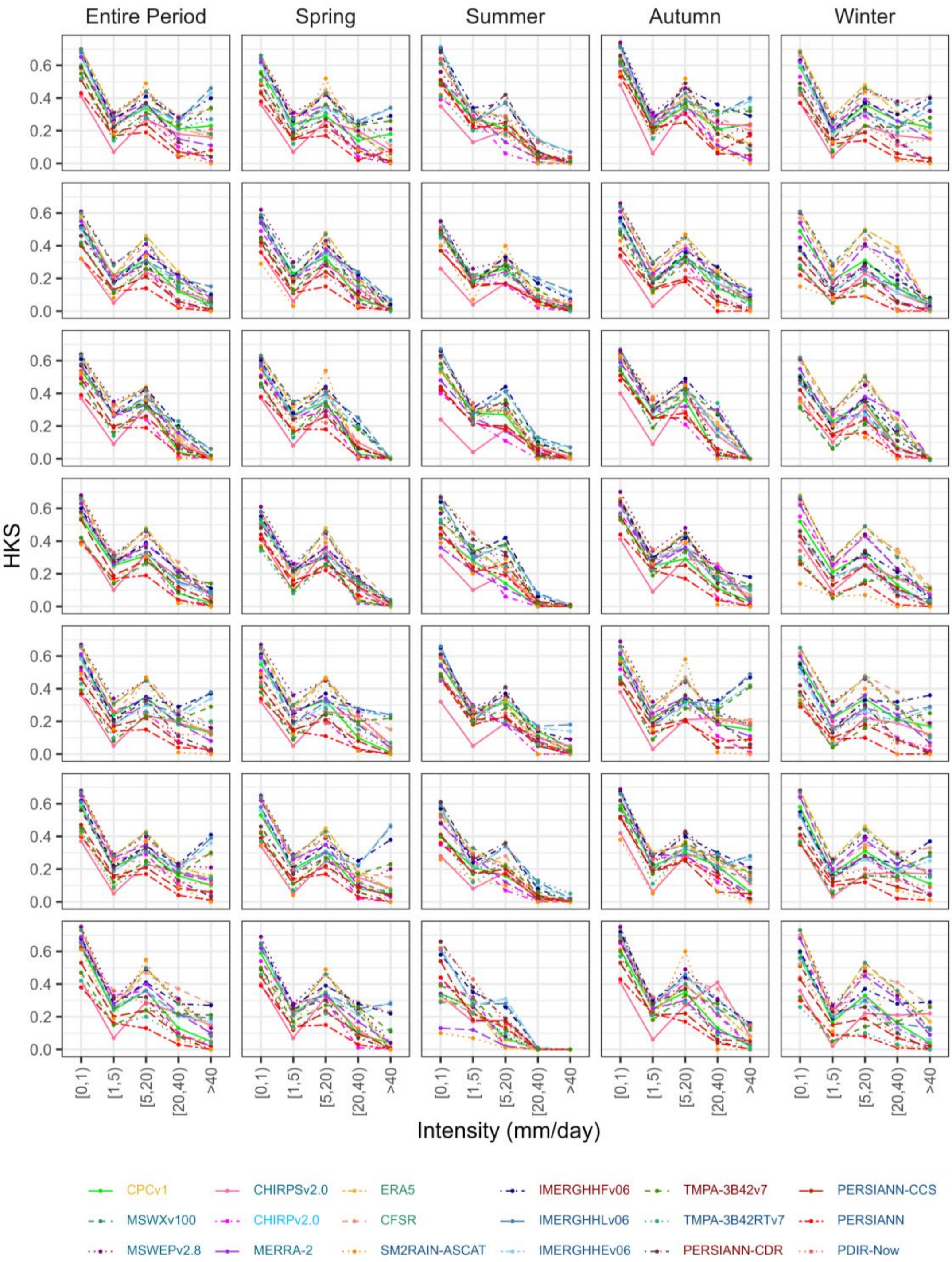

Figure 3.27. *Daily precipitation frequency for five different intensities from observed and 18 GPDs for the entire period and four seasons over seven climate zones. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

SM2RAIN-ASCAT and MSWEPv2.8 over Central Anatolia during the autumn season demonstrate the highest detectability of light precipitation compared to other selected GPDs, where some datasets, such as CHIRPSv2.0, show a lower detectability of light precipitation in most climate zones under different temporal conditions. It's worth mentioning that SM2RAIN-ASCAT is able to present relatively higher detection skills for moderate precipitation over Aegean, Marmara, Southeastern, and Central Anatolia regions in different seasons, while ERA5 and MSWXv100 are also able to show higher detectability of moderate precipitation over Southeastern Anatolia during the winter season. The same datasets (ERA5, MSWXv100, and SM2RAIN-ASCAT) also show the lowest detection ability of moderate precipitation, followed by CHIRPv2.0, MERRA-2, and CFSR datasets over Southeastern Anatolia during the summer season. For the heavy precipitation, it is interesting to see that CFSR and CHIRPSv2.0 show the highest detection ability over Southeastern Anatolia during the winter and autumn seasons and show the lowest detectability compared to other GPDs during the summer season. Finally, it is the IMERGHHv06 datasets that demonstrate the highest detection ability of extreme precipitation over Marmara, Mediterranean, and Aegean regions considering different seasons, which is followed by the TMPA-3B42(RT)v7 dataset, while TMPA-3B42RTv7 over Central Anatolia and CFSR and CHIRP(S)v2.0 over different climate zones mostly show the lowest detectability of extreme precipitation considering individual seasons.

3.1.7. Summarizing GPDs performance

Figures 3.28 and 3.29 display the top-six and worst-six GPDs over the entire region, seven climate zones, four wetness classes, and four elevation ranges for the entire study period and individual seasons, which are ranked based on their performance under the KGE performance indicator.

The GPDs' best and worst performance rankings differed at daily (Figure 3.28) and monthly (Figure 3.29) time steps. For the daily time step, many multi-source GPDs, such as MSWEP2.8, MSWXv100, IMERGHHFv06, and TMPA-3B42v7, show the best performance for most of the time over different regional classifications, followed by some reanalysis such as ERA5 and the ground-based dataset (CPCv1), where among gauge-uncorrected GPDs, satellite-based precipitation demonstrates the worst performance, respectively. However, there are some exceptions; for example, IMERGHHLv06, which is a satellite-based dataset, demonstrates higher performance than reanalysis and ground data over Black Sea region or outperforms some gauge-corrected GPDs such as TMPA-3B42v7 and MERRA-2 over Central Anatolia, where it is also one of the six best over regions with elevation greater than 1000 m considering the entire period.

Moreover, PDIR-Now has become one of the best GPDs over the Mediterranean region during the summer season, whereas ERA5 demonstrates the highest performance among the six best GPDs over many regional classifications during the winter season. However, PERSIANN datasets consistently ranked as the worst performing GPDs across multiple regional classification and temporal scales.

When the time step shifts from the daily to the monthly time step, still multi-source GPDs rank among the six best-performing datasets, where CHIRPSv2.0 and MERRA-2 also look among the best-ranked GPDs. On the other hand, CPCv1, which takes advantage of the spatial information of available ground data, is most of the time located

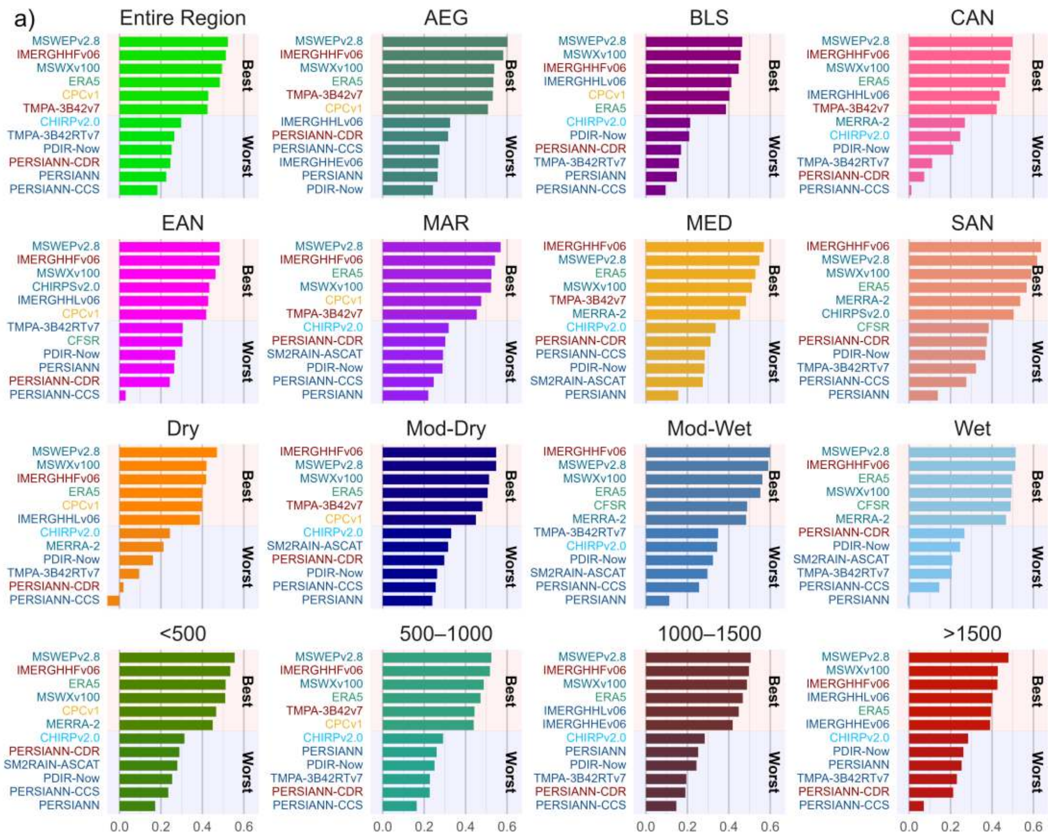

Figure 3.28. *GPDs best-six and worst-six ranking based on KGE for daily precipitation over the entire region, seven climate zones, four wetness classes, and four elevation ranges. a) entire study period, b) Spring season, c) Summer season, d) Autumn season, e) Winter season. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

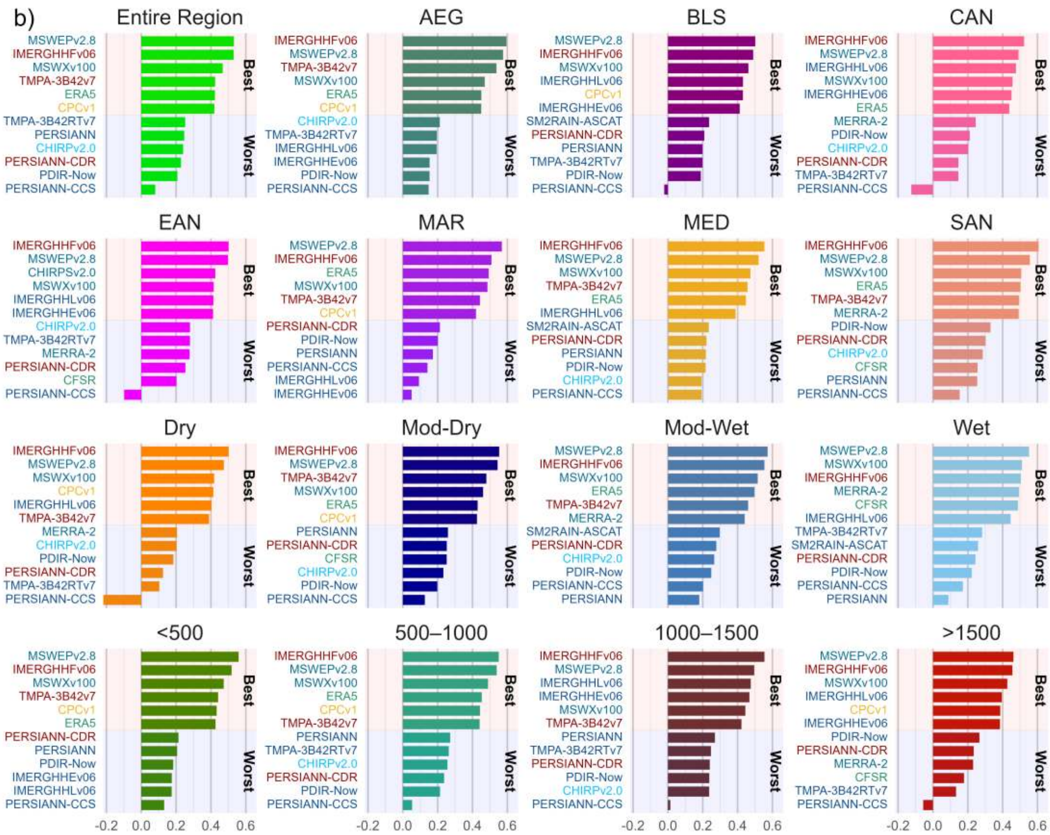

Figure 3.28. *Continued…*

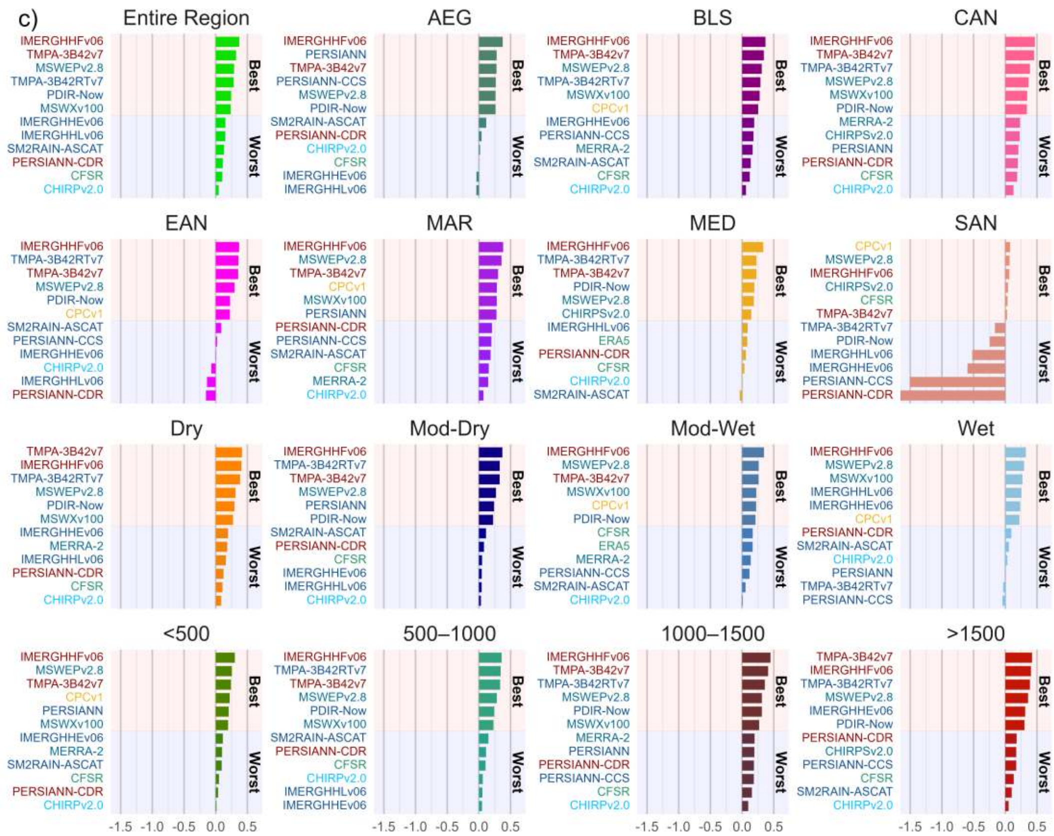

Figure 3.28. *Continued…*

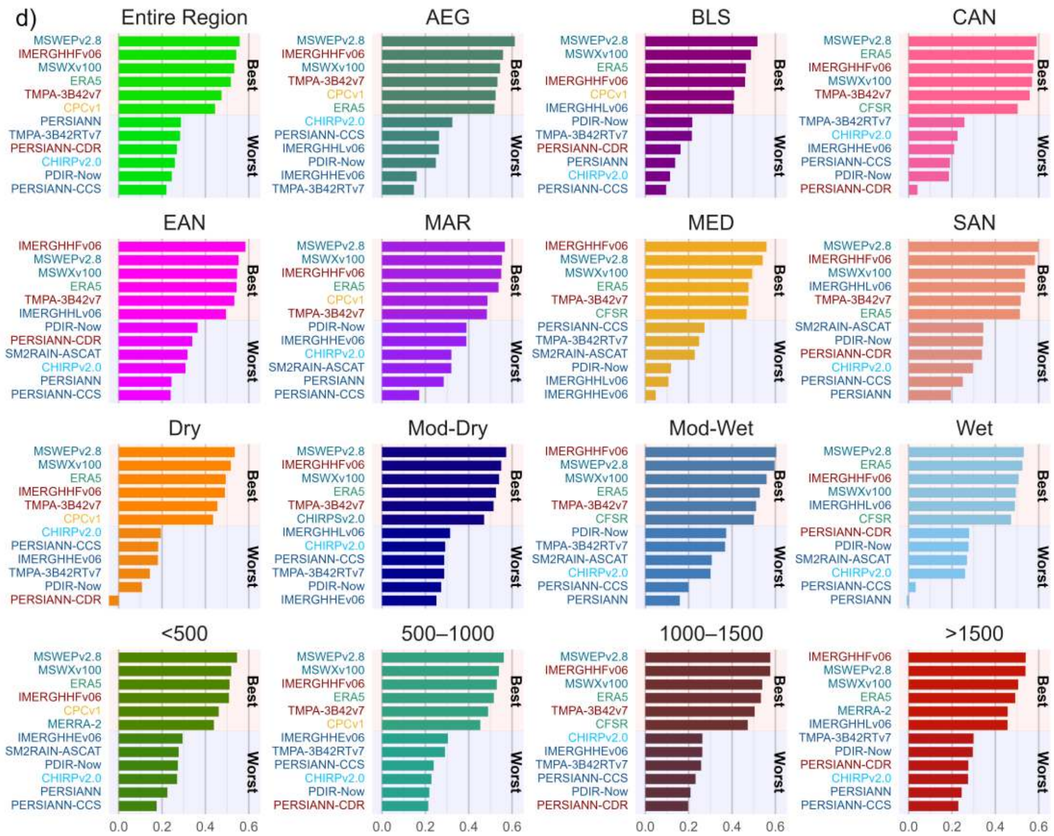

Figure 3.28. *Continued…*

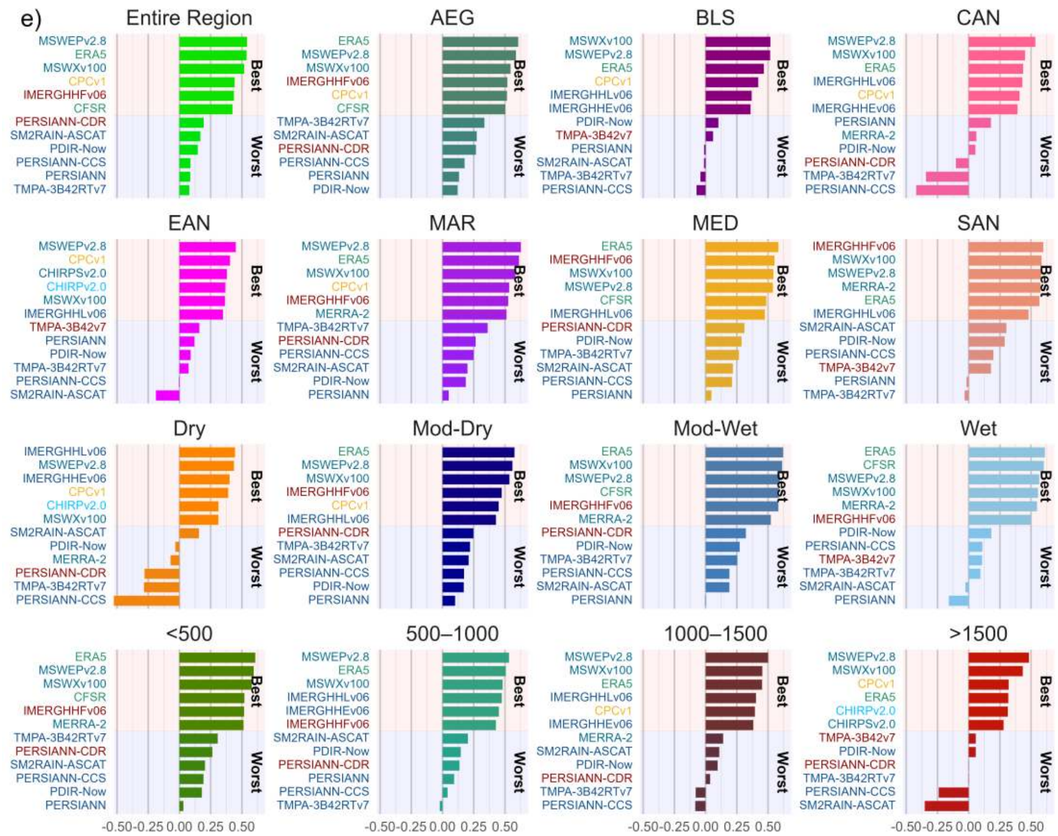

Figure 3.28. *Continued…*

at the top of the best-performing GPDs ranking over different regional classifications and temporal scales. Still, satellite-based GPDs, especially PERSIANN datasets, are mostly ranked as one of the worst-performing GPDs in different conditions for the monthly precipitation. However, some GPDs, such as CFSR, is ranked as some of the best performers and even outperform ERA5 and some multi-source GPDs over wet regions, where IMERGHHL(E)v06 is also ranked as one of the best performers over Central Anatolia during the winter season for the monthly time step.

Apart from GPDs performance over different region classifications under distinct conditions such as temporal window or considering seasonal variability, the findings also suggest that GPDs that employ ground data in their retrospective algorithms might not always be able to get realistic precipitation predictions. The source of information, rectification time scale, and available number of gauges used are crucial aspects that have

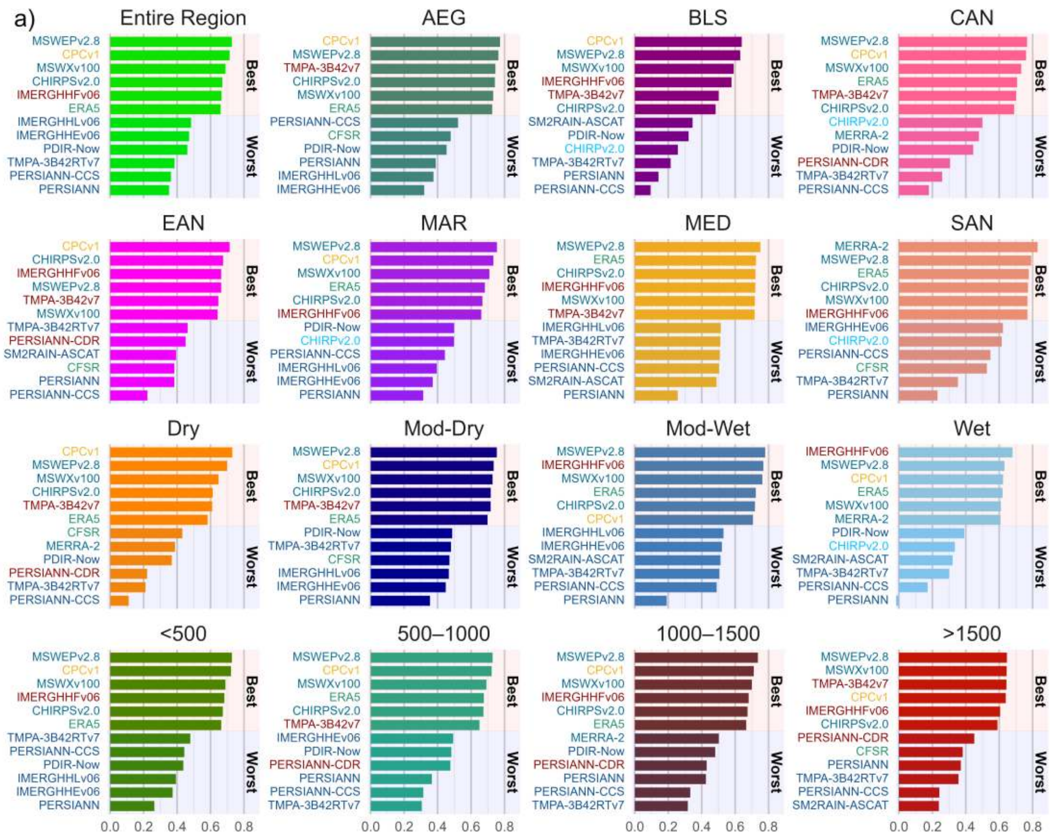

Figure 3.29. *GPDs best-six and worst-six ranking based on KGE for Monthly precipitation over the entire region, seven climate zones, four wetness classes, and four elevation ranges. a) entire study period, b) Spring season, c) Summer season, d) Autumn season, e) Winter season. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

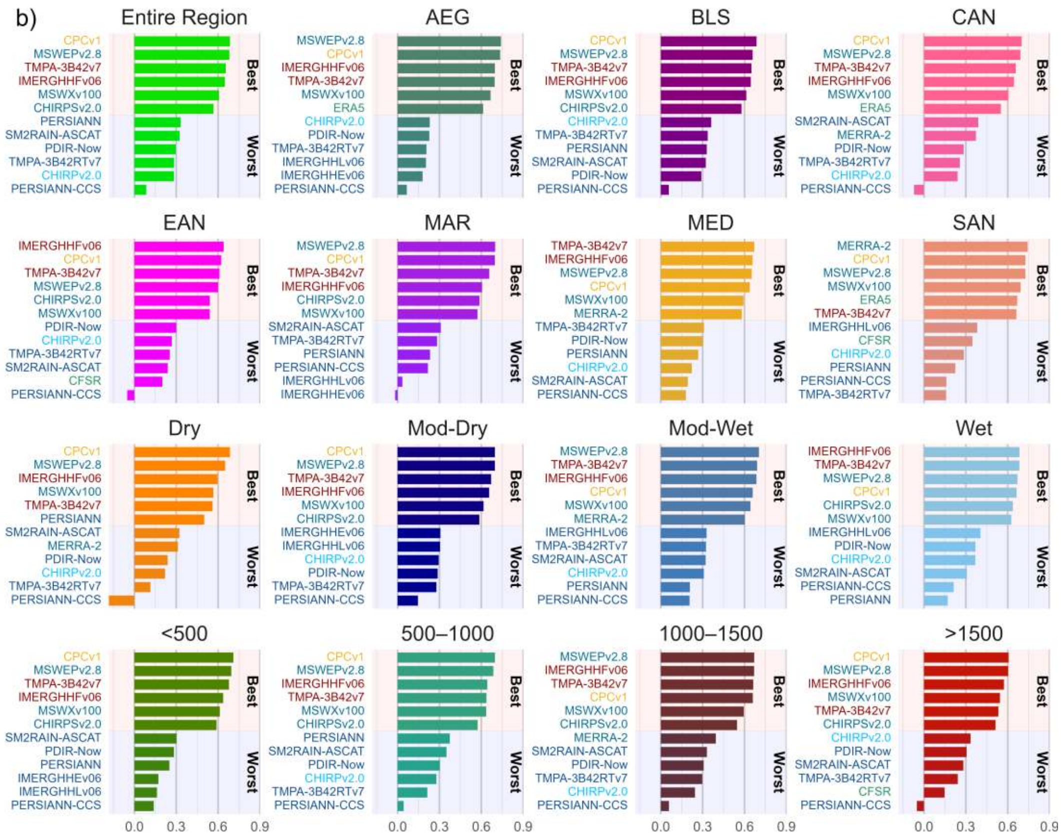

Figure 3.29. *Continued…*

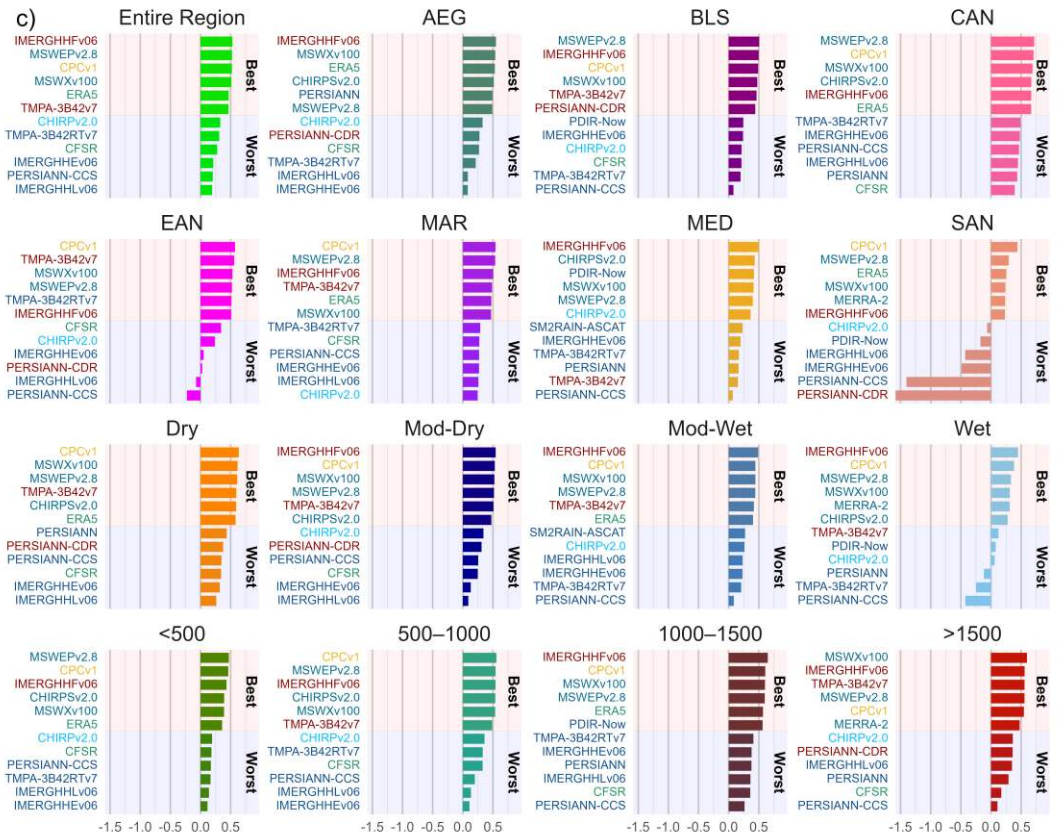

Figure 3.29. *Continued…*

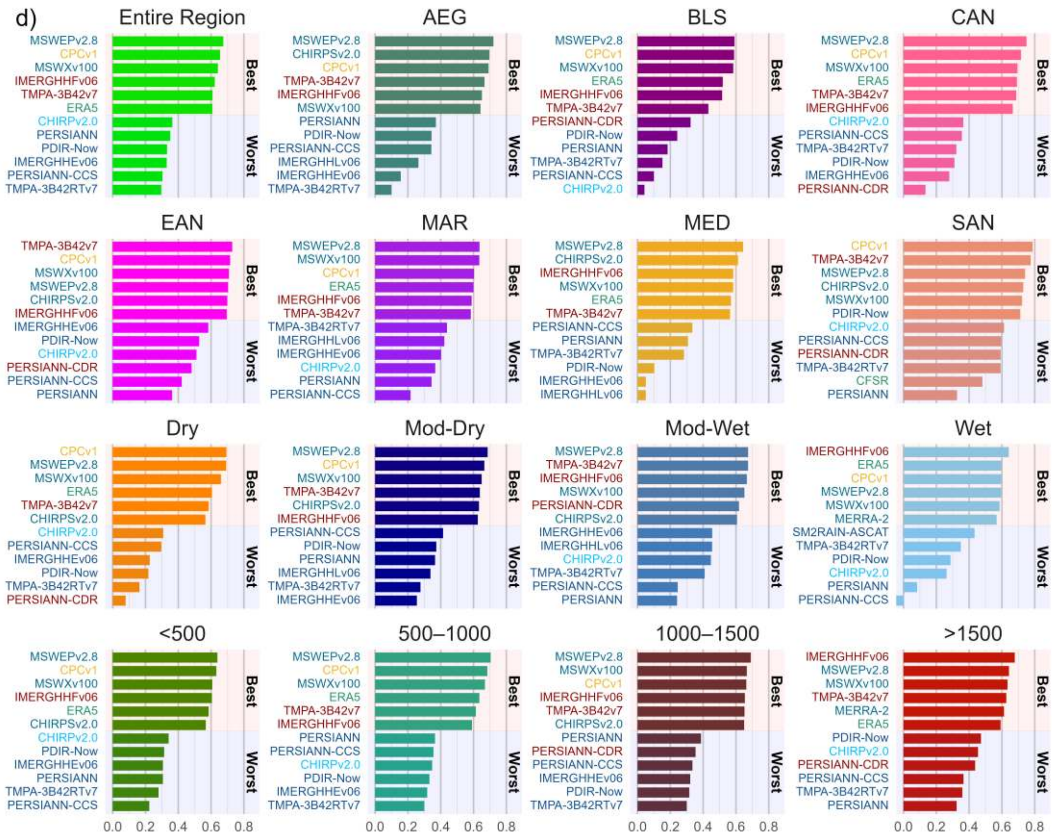

Figure 3.29. *Continued…*

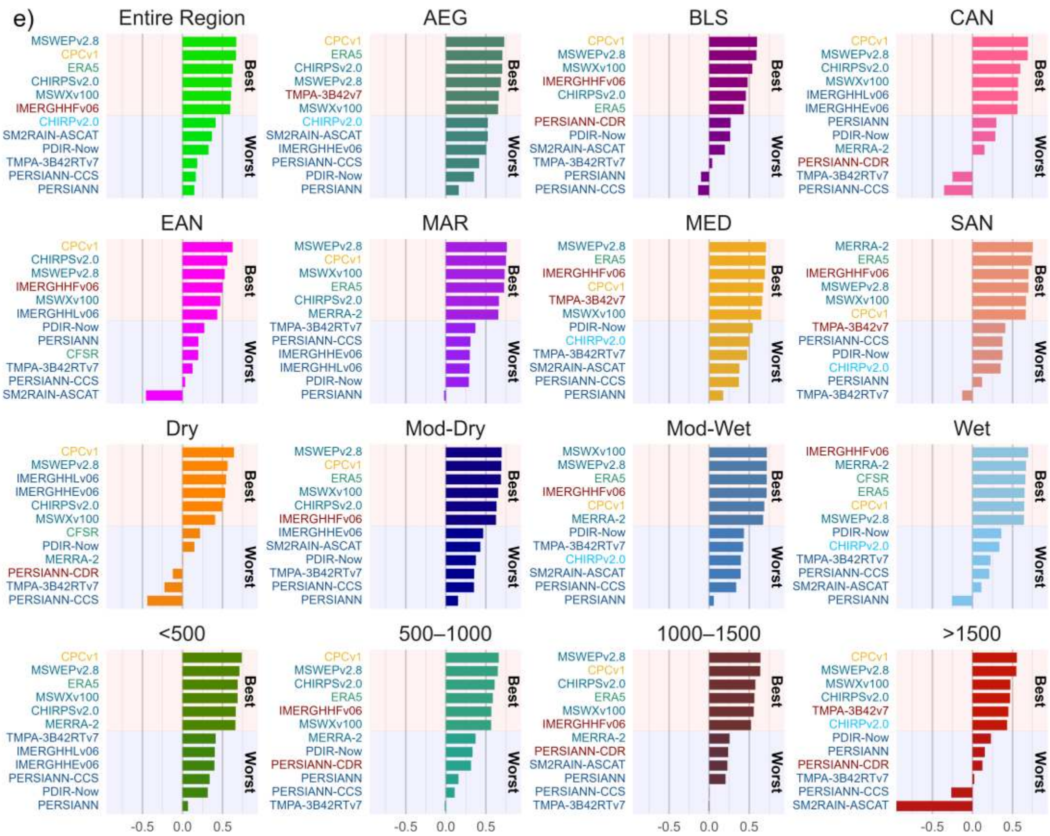

Figure 3.29. *Continued…*

a significant impact on the final product's accuracy.

In general, there are three types of gauge-based data that are used in gauge corrected GPDs: (1) point-based precipitation measurements based on rain gauge records; (2) GPDs estimates based on interpolation of rain gauge records; and (3) point-based or gridded gauge precipitation estimates that are combined with different satellite-based precipitation datasets. For instance, TMPA-3B42V7 (G. J. Huffman et al., 2007) utilizes the gridded form of monthly observed precipitation from GPCC with 1^{o} spatial resolution, while IMERGHHFv06 (G. J. Huffman et al., 2020) and PERSIANN-CDR (Ashouri et al., 2015) are adjusted with monthly data from GPCP, which is a merged product of ground observations and satellite-based precipitation estimates and has a 2.5^{o} spatial resolution. In the same way, CHIRPSv2.0 (Funk et al., 2015) uses the CHPclim (i.e., gridded precipitation estimates merged with satellite precipitation estimates) dataset (0.05^{o}) and gauge precipitation estimates from various public data networks, where MSWEPv2.8 (Beck, Wood, et al., 2019) and MSWXv100 (Beck et al., 2022) employ observed monthly precipitation estimates from GPCC and daily precipitation from various sources. Finally, MERRA-2 (Reichle, Liu, et al., 2017) utilizes gridded observed daily precipitation from CPCv1.

3.2. Hydrological evaluation of GPDs

In this context, selected GPDs are validated by utilizing them as a meteorological forcing in the TUW hydrologic model under different model calibration scenarios. In the direct comparison of GPDs with ground stations (section 3.1), it has been found that the performance of each GPD varies in spatial and temporal dimensions. In other words, it is usually recommended to investigate GPDs' performance against ground stations before employing them for streamflow simulation. As a result, stations within and near the basin are selected for direct evaluation of GPD performance and hydrological utility over that particular basin.

3.2.1. GPDs over the upper Euphrates (Karasu) basin

3.2.1.1. Direct comparison of GPDs

Figure 3.30 maps the mean daily precipitation and bias at the station level, obtained from ground observation and 18 GPDs for the entire period (2015–2019). According to the observed data, the mean daily precipitation increases from the north (1–1.5 mm) to the south (2.5–3 mm) of the basin. CPCv1, which is developed based on spatial information collected from ground stations, displays mean daily precipitation close to what is observed inside the basin and overestimates precipitation in the northeast and underestimates it in the southwest parts of the basin.

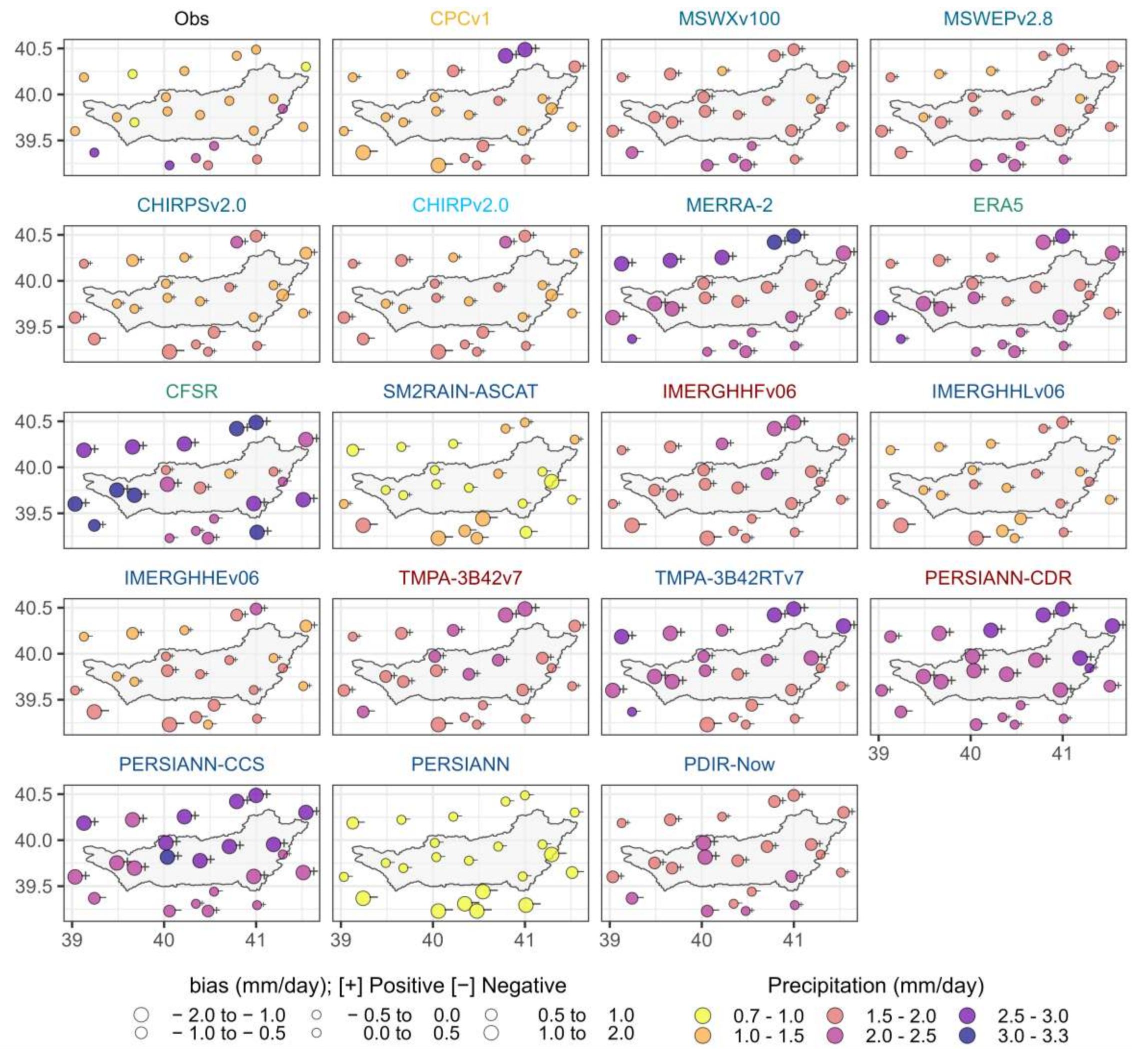

Figure 3.30. *Mean daily precipitation and bias at the station level over the Karasu river basin from gauge observations (Obs) and 18 GPDs for the period 2015–2019.The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Among gauge corrected GPDs, MSWXv100, MSWEPv2.8 and CHIRP(S)v2.0 deliver precipitation close to observed especially inside the basin where MERRA-2, IMERGHHFv06, TMPA-3B42v7 and PERSIANN-CDR overestimates mean daily precipitation with different bias. Both reanalysis GPDs (ERA5 and CFSR) show higher precipitation outside of the basin. Among satellite-based GPDs, PERSIANN and

SM2RAIN-ASCAT highly underestimate precipitation, where PERSIANN-CCS and TMPA-3B42RTv7 show higher precipitation in the region compared to observed precipitation, and IMERGHHE(L)v06 presents mean daily precipitation close to observed precipitation, followed by PDIR-Now.

Figure 3.31 displays the mean daily precipitation and bias for the selected GPDs at the regional scale, considering the entire period and four seasons (spring, summer, autumn, and winter). According to observed mean daily precipitation, the region receives more precipitation during the spring (2.24 mm) followed by winter (1.78 mm), while the summer (0.74 mm) and autumn (1.2 mm) seasons show less precipitation, with a 1.5 mm mean daily precipitation estimated over the Karasu basin for the entire period.

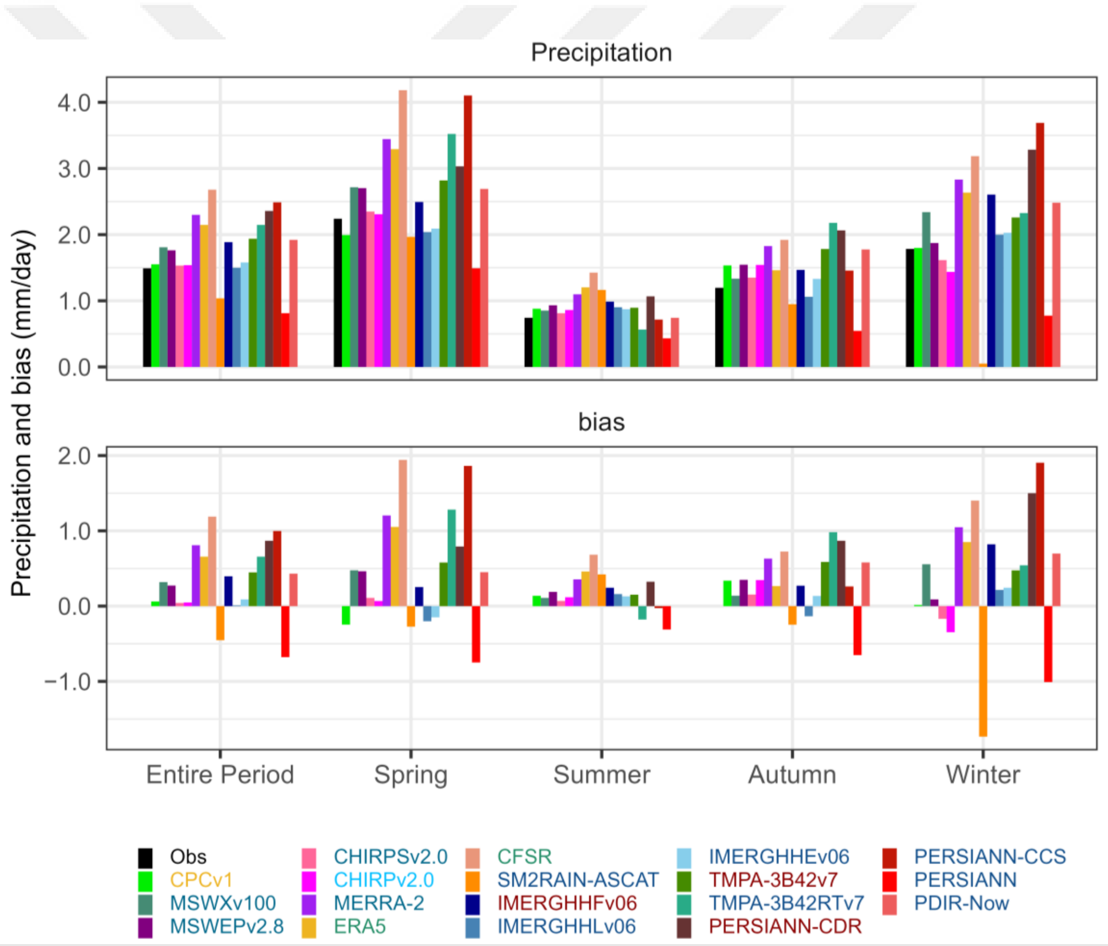

Figure 3.31. *Mean daily precipitation and estimated bias at the regional scale over Karasu basin for the entire period and four seasons. The color of the legend text presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

CPCv1 presents mean precipitation that is close to the observed for the entire period and all seasons where the estimated mean daily precipitation bias does not exceed ±0.3

mm. Moreover, both CHIRPv2.0 and CHIRPSv2.0 reproduce mean daily precipitation well and only underestimate precipitation during the winter season, where MSWXv100, MSWEPv2.8, and IMERGHHFv06 present mean daily precipitation close to the observed, and MSWEPv2.8 shows higher accuracy than MSWXv100 during the winter season. Many gauge-corrected GPDs, including MERRA-2, TMPA-3B42v07, and PERSIANN-CDR, present relatively higher precipitation for the entire study period and four seasons. Among reanalysis GPDs, ERA5 outperforms CFSR, but both datasets overestimate observed precipitation, which is more intense for CFSR. Among satellite-based GPDs, both IMERGHHEv06 and IMERGHHLv06 reproduce mean daily precipitation close to observed, while TMPA-3B42RTv7, PERSIANN-CCS, PDIR-Now overestimate, and PERSIANN and SM2RAIN-ASCAT significantly underestimate considering the entire period and four seasons. As a result, SM2RAIN-ASCAT, PERSIANN for the entire period; CPCv1, SM2RAIN-ASCAT, IMERGHHL(E)v06, and PERSIANN in the spring season; PERSIANN-CCS, PERSIANN, and TMPA-3B42RT in the summer season; PERSIANN, SM2RAIN-ASCAT, and IMERGHHLv06 in the autumn; and CHIRP(S)v2.0, SM2RAIN-ASCAT, and PERSIANN during the winter season underestimate observed mean daily precipitation at the regional scale over Karasu basin where the rest of GPDs shows higher precipitation than the observed.

The reliability of GPDs at the station location over Karasu basin is expressed in the form of KGE along with its correlation coefficient (R), the ratio of bias (Bias), and variability ratio (VR) (Figure 3.32). Overall, CPCv1 shows relatively higher performance compared to many GPDs and is the only GPD whose temporal variation is entirely dependent on ground observation. Among multi-source GPDs, MSWEPv2.8 has higher performance and even outperforms the CPCv1 dataset, followed by CHIRP(S)v2.0, MSWXv100, and IMERGHHFv06, while MERRA-2, TMPA-3B42RTv7, and PERSIANN-CDR perform relatively poorly inside and north of the basin and relatively better in the southern regions. In the same way, CHIRPv2.0 was able to present close performance to its gauge-corrected version (CHIRPSv2.0), while ERA5 and CFSR show higher performance in the southern part of the basin than in the inside and northern parts. Among satellite-based GPDs, IMERGHHL(E)v06 and SM2RAIN-ASCAT demonstrate higher performance, and even IMERGHHL(E)v06 outperforms its final version (IMERGHHFv06), where PERSIANN and PDIR-Now outperform TMPA-3B42RTv7 and PERSIANN-CCS and show higher performance than PERSIANN-CDR datasets. The correlation component also follows directly the final KGE most of the time, where CPCv1 and some multi-source and reanalysis datasets show higher correlation than the other GPDs at the station location, where the ratio of bias and variability ratio vary from one GPD to another and even a single dataset shows different values over individual stations.

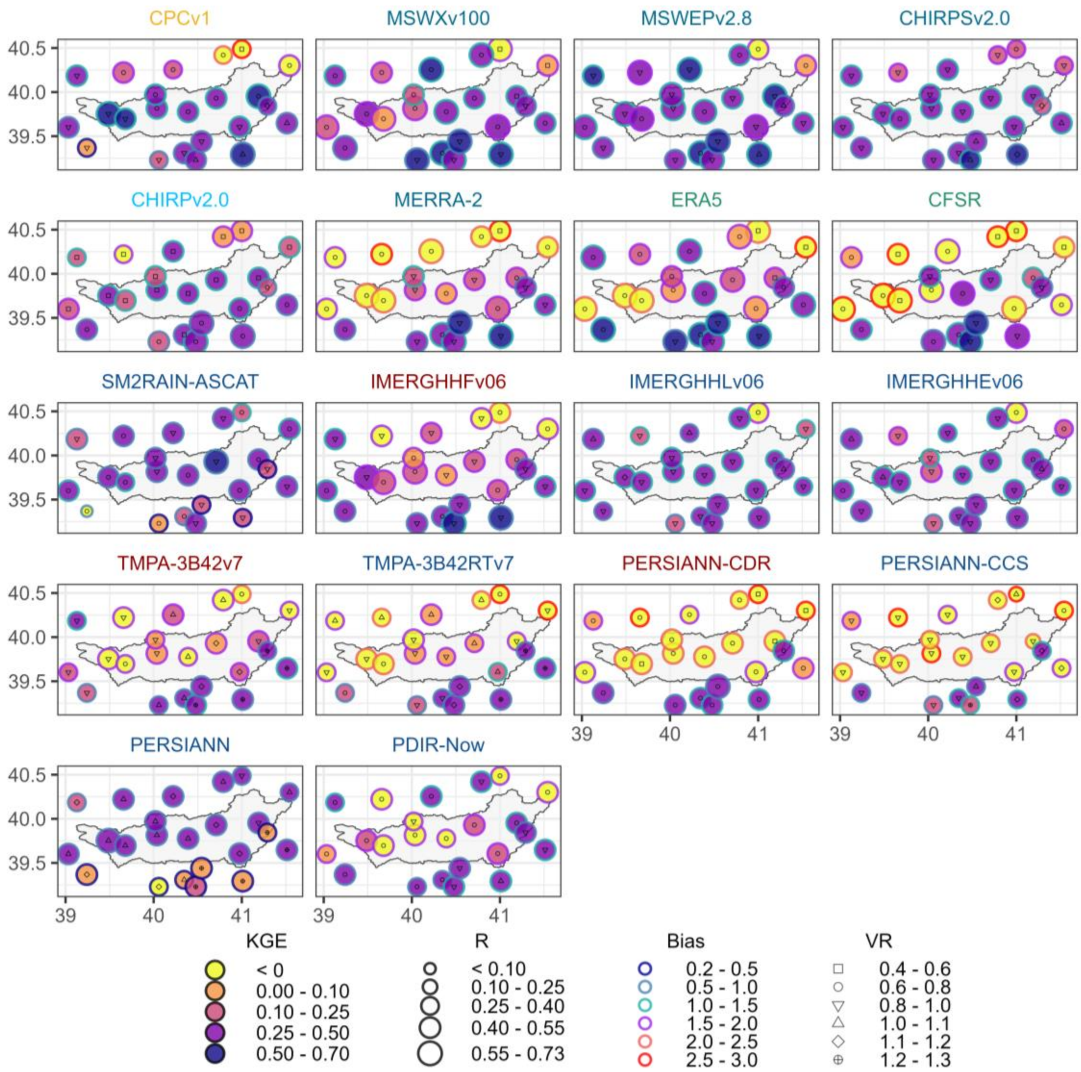

Figure 3.32. *Reliability of GPDs at the station level expressed in the form of KGE and its individual components at a daily time step over the Karasu river basin. The title color presents the source (s) of GPDs: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Figure 3.33 displays the performance of GPDs at the regional scale expressed in the form of the KGE and its three components over Karasu basin for the entire study period and four seasons. In general, GPDs show higher performance during the spring and autumn seasons than during the summer and winter seasons. As a result, MSWEPv2.8 outperforms other selected GPDs for the entire study period (median of KGE: 0.43) and in the spring (median of KGE: 0.47) season, which is followed by CPCv1 and other multi-source GPDs. Overall, some gauge-corrected GPDs such as MERRA-2, IMERGHHFv06, TMPA-3B42v7, and PERSIANN-CDR, reanalysis datasets, and many satellite-based GPDs show relatively lower performance, and the performance of many GPDs becomes worse during the winter season. However, some exceptions exist, for example, ERA5, which performs relatively lower but shows the second-best performance during the autumn season, or PERSIANN, PDIR-Now, and SM2RAIN-ASCAT, which show relatively higher performance than other gauge-uncorrected and some gauge-corrected GPDs for different temporal. In the same way, GPDs show higher correlation to the observed precipitation during the autumn season and relatively lower during the summer season, where datasets highly overestimate the ratio of bias during the winter season and GPDs show mostly underestimation of variability ratio at the regional scale.

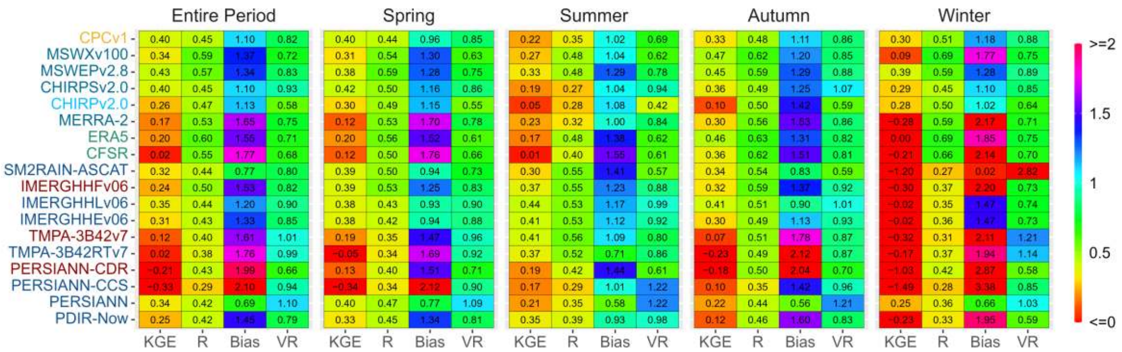

Figure 3.33. *Reliability of GPDs at the regional scale expressed in the form KGE and its individual components for daily and monthly precipitation over Karasu basin for the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Figure 3.34 summarizes the frequency of observed and selected GPDs for five different precipitation intensities over Karasu basin for the entire study period and four seasons. According to the observed data, around 77% of precipitation events occur in the range of 0–1 mm/day for the entire period, while this amount decreases for spring and winter precipitation and shows higher frequency during summer and autumn. Furthermore, as expected, the frequency of precipitation events decreases as the intensity of precipitation increases. GPDs display varied frequency of precipitation intensities, which are particularly obvious in the spring. Among GPDs, PERSIANN exhibits a higher precipitation frequency with intensities ranging from 0 to 1 mm/day compared to the observed and presents less frequency of light precipitation (1–5 mm/day) events during the entire period and all seasons, where the opposite is the general case for PDIR-Now. Many GPDs present higher precipitation frequencies of light precipitation during the spring season followed by the winter season, and overestimation of light frequency is more observable for CHIRPv2.0, SM2RAIN-ASCAT, and PDIR-Now during the summer and autumn seasons.

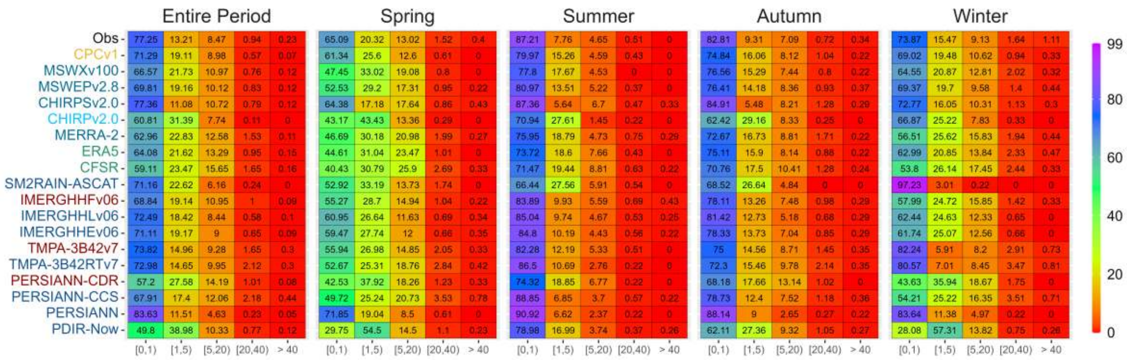

Figure 3.34. *Frequency of GPDs for different daily precipitation intensities over Karasu basin for the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The detectability of GPDs for five different precipitation intensities is evaluated by employing the Hanssen-Kuiper Skill (HKS) score over Karasu basin for the entire study period and for seasons (Figure 3.35). Generally, GPDs show higher detectability during autumn and lower detectability during summer. Additionally, the detection strength of GPDs decreases as the precipitation intensity increases, which is generally the case in literature. This can be attributed to the classification of several intensity classes, which makes it hard to differentiate between them instead of a simple rain/no rain scenario. This division gets even more problematic when the occurrence probability of a certain event is substantially lower in comparison to other classes (heavy and extremes intensities), as can be followed from Figure 3.34. Hence, detecting a rarely occurring very intense precipitation event has a weaker performance when assessed against a more frequently occurring light storm. From the results, among selected GPDs, MSWEPv2.0 shows the highest detectability for no-precipitation events during the autumn season, followed by ERA5 and MSWXv100 during the winter season. In the same way, PDIR-Now and MSWXv100 during the summer season and SM2RAIN-ASCAT and MSWEPv2.8 during the autumn season for the light precipitation, whereas MSWXv100, ERA5, and MERRA-

2 for the moderate precipitation, show the highest detection skill compared to other GPDs. Finally, some multi-source GPDs such as MERRA-2, IMERGHHFv06, PERSIANN-CDR, and CHIRPSv2.0 during the autumn season and CFSR and ERA5 during the winter season for the heavy precipitation, while CFSR during the autumn season and entire period and MSWEPv2.8 during the entire period and spring season for the extreme precipitation demonstrate the higher detection ability, comparatively.

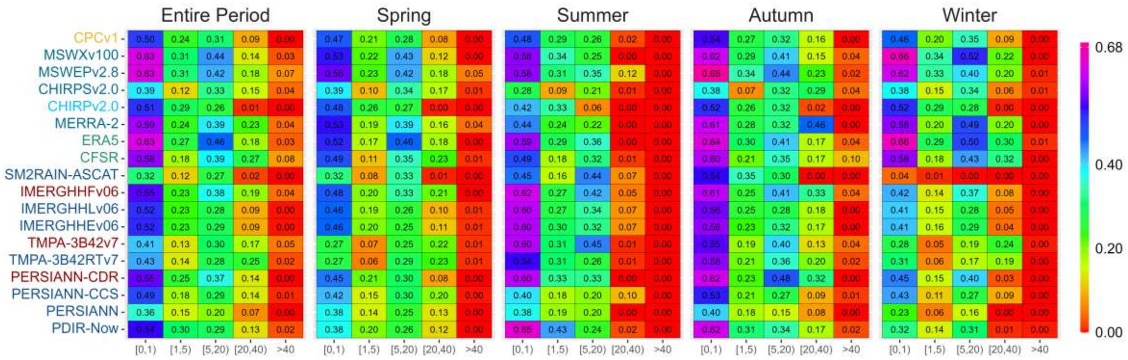

Figure 3.35. *Detection ability in reproducing daily precipitation intensities expressed in the form of the Hanssen–Kuiper Skill (HKS) score over Karasu basin considering the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

3.2.1.2. Hydrologic utility of GPDs over Karasu basin

As previously stated, the TUW hydrologic model is used for streamflow simulation at the outlet of Karasu basin (Kemah stream gauging station) utilizing observed precipitation and selected GPDs, which is done in two phases. In the beginning, the model parameters are primarily calibrated by observed precipitation, then the observed precipitation is replaced by each GPD (Scheme-1). Following that, the model parameters are calibrated for observed precipitation and each GPD independently (Scheme-2). Regardless of calibration scenarios, the model is calibrated over two hydrological years (October 2014–September 2016) and validated for three years (October 2016–September 2019).

Figure 3.36 depicts the observed streamflow hydrograph (gray shade area) and simulated streamflow hydrographs by observed precipitation and selected GPDs for two different calibration and validation scenarios (Scheme-1 and Scheme-2) at Kemah stream gauging station. From the results, observed precipitation draws a close hydrograph to observed streamflow time series for both calibration and validation periods. Moreover, GPDs show deviated hydrographs compared to observed streamflow when the model is run by an optimal parameter set, which is obtained by observed precipitation (Scheme-1) and presents better and closer hydrographs when the model simulates streamflow by optimal parameter sets of each GPD (Scheme-2).

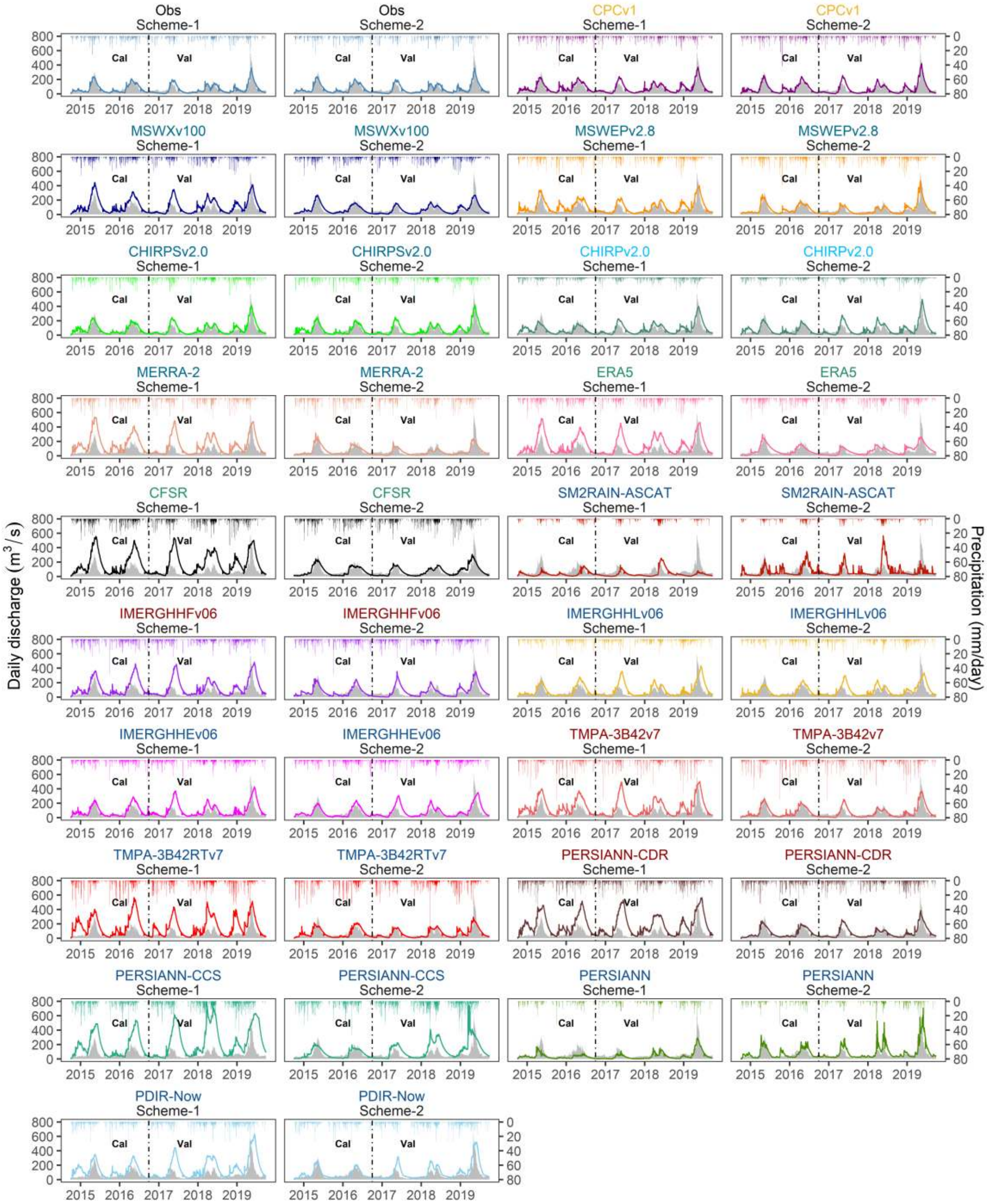

Figure 3.36. *Comparison of the observed and simulated daily hydrographs at the Karasu basin outlet based on the observed data and 18 GPD input for the calibration (October 2014 to September 2016) and validation period (October 2016 to September 2019). Title color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The difference between the two schemes for streamflow simulation can be easily differentiated by comparing the temporal dynamics of simulated streamflow against observed discharge in the form of a scatter plot (Figure 3.37).

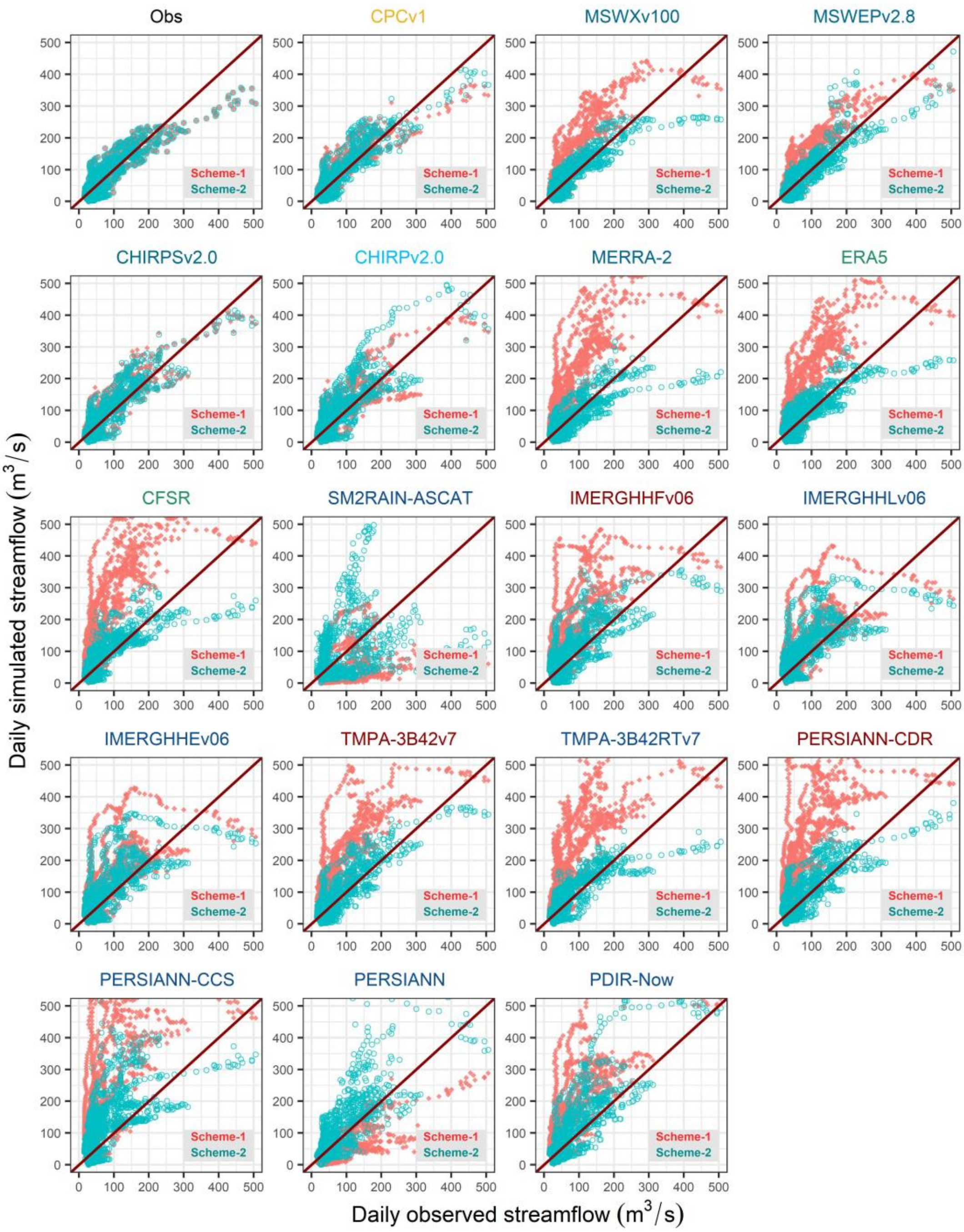

Figure 3.37. *Scatter plot of daily simulated streamflow against observed discharge for the observed precipitation and 13 GPDs at Karasu basin outlet. Title color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

In general, GPDs exhibit a higher amount of simulated streamflow in Scheme-1 when compared to Scheme-2. In other words, daily observed streamflow is overestimated by many GPDs when they are used as meteorological forcing along with the optimal model parameter set obtained by observed precipitation. Conversely, when the model parameters are calibrated either by observed precipitation (Scheme-1) or by the GPD itself (Scheme-2), certain GPDs, such as CHIRP(S)v2.0 and CPCv1, simulate streamflow that is relatively close to observed discharge. On the other hand, the SM2RAIN-ASCAT and PERSIANN datasets highly underestimate observed streamflow in Scheme-1 for high flows. Among gauge-corrected GPDs, MERRA-2, PERSIANN-CDR, TMPA-3B42v7, and IMERGHHFv06 reanalysis datasets (ERA5 and CFSR) highly overestimate observed streamflow in Scheme-1. Moreover, IMERGHHL(E)v06 shows a similar streamflow scatter for both schemes. The rest of the satellite-based GPDs show a higher ability for streamflow prediction in Scheme-2 than in Scheme-1. The scatter plot of some GPDs indicates the nonlinearity of simulated streamflow. This can be related to the existing high wet bias, which is magnified in the direct comparison of each GPD with observed precipitation (Figure 3.32 and Figure 3.33). Furthermore, the model parameters in Scheme-1 are calibrated based on observed precipitation, which is not an optimal parameter set for selected GPDs and can be the reason for the high overestimation of streamflow by some GPDs during the calibration and validation periods (especially for MERRA-2, ERA5, CFSR, IMERGHH datasets, TMPA datasets, PERSIANN-CDR, PDIR-Now, and PERSIANN-CCS). Lastly, it should be kept in mind that for Scheme-2 modeling, only precipitation values are taken from GPDs while PET and temperature values are still part of the observed meteorological forcing in the model during the calibration and validation.

Figure 3.38 shows the bias distribution of simulated streamflow compared to observed discharge in the form of a box-and-whiskers plot, obtained from observed precipitation and selected GPDs for the calibration (2015-2016) and validation (2017-2019) years. In general, GPDs show a higher wet bias for both calibration and validation years in Scheme-1 than in Scheme-2. However, looking at some GPDs' median bias, such as SM2RAIN-ASCAT and PERSIANN, highly underestimate observed streamflow for each year of calibration and validation periods, while CHIRPv2.0 only underestimates observed streamflow during 2016 and observed precipitation during 2019. Moreover, when streamflow is simulated by each GPD with its optimal parameter set (Scheme-2), the median of bias for simulated streamflow with different GPDs are close to zero in 2015, then the underestimation of observed streamflow is increased for the 2016 and 2017 years, and most GPDs overestimate observed streamflow for 2018 and 2019. Finally, the bias varies in a wider range for PERSIANN-CCS in all years in Scheme-1 and for validation years in Scheme-2, where SM2RAIN-ASCAT in the 2016 calibration year and IMERGHHFv06 in the 2017 validation year show wider bias variation, comparatively.

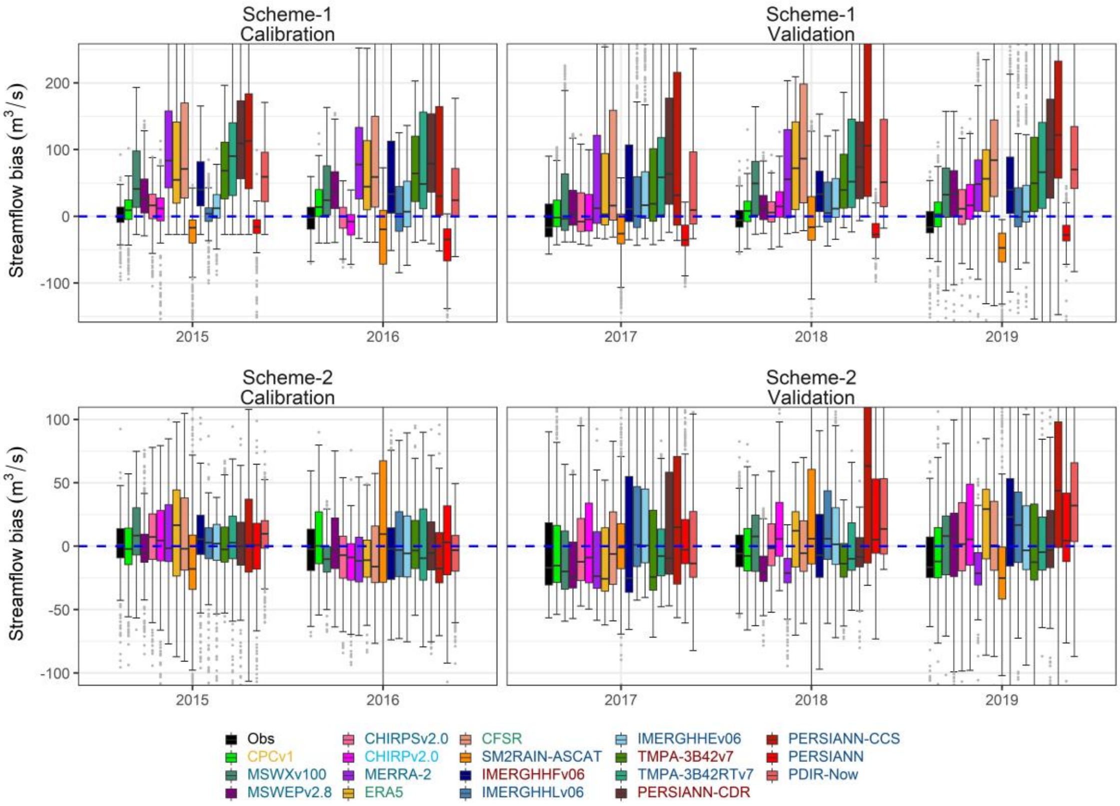

Figure 3.38. *Daily simulated streamflow bias distribution is expressed in the form of box and whiskers, considering the calibration and validation periods. The boxes represent the 25th, 50th, and 75th percentile values, respectively. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

According to the selected GPDs, the simulated streamflow time series bias is shown in Figure 3.39 in comparison to the observed discharge under two distinct model parameter calibration scenarios. In Scheme-1, many GPDs, including MSWXv100, MERRA-2, ERA5, CFSR, IMERGHHFv06, TMPA-3B42(R)v7, PERSIANN-CDR, and PERSIANN-CCS, exhibit higher streamflow bias than in Scheme-2 for both the calibration and validation periods, while the PERSIANN dataset underestimates observed streamflow in Scheme-1 and overestimates it in Scheme-2, and PDIR-Now displays higher simulated streamflow than the observed discharge for both schemes. Some GPDs, such as CHIRP(S)v2.0, SM2RAIN-ASCAT, and IMERGHHE(L)v06, exhibit streamflow biases that are comparable for both schemes. Furthermore, when the model is calibrated by each GPD separately, many GPDs underestimate observed streamflow during the 2019-year peak flow season, which is a significant problem in Scheme-2.

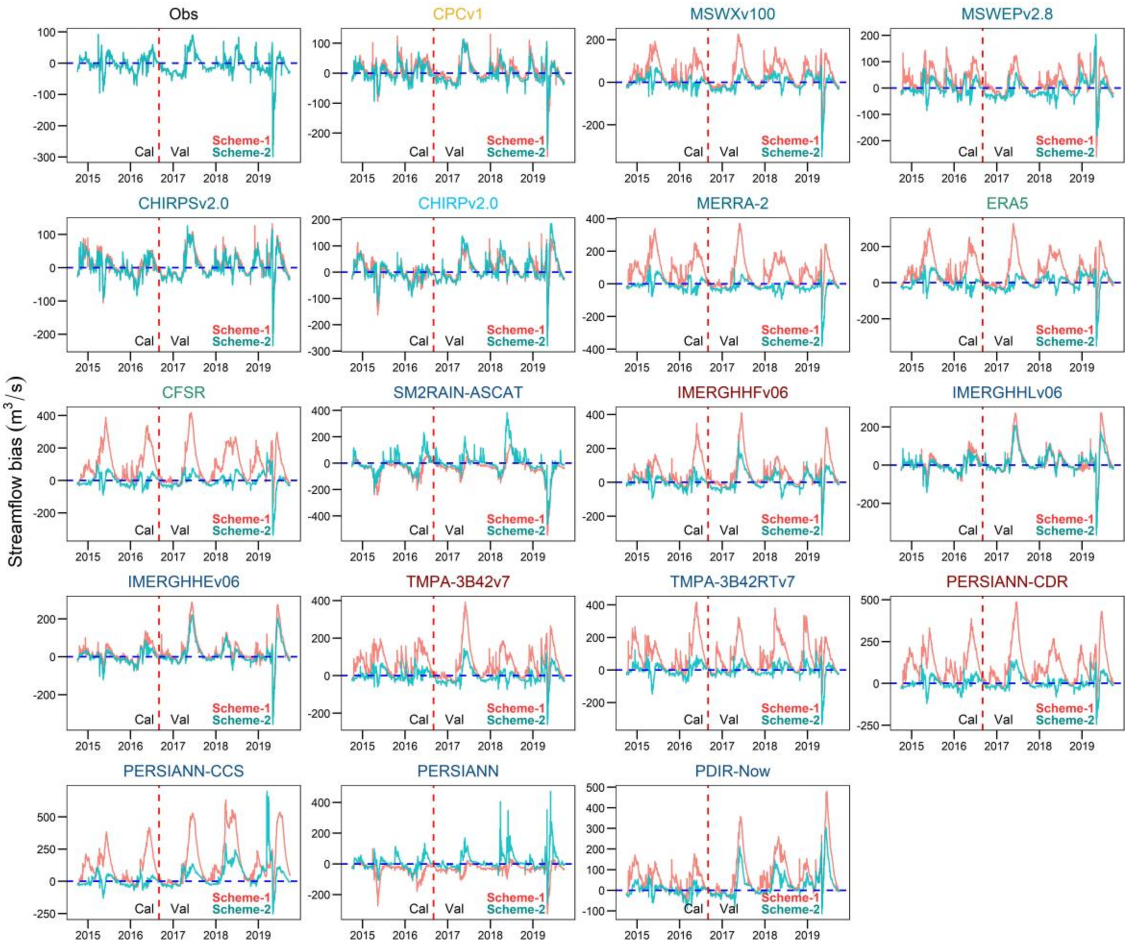

Figure 3.39. *Daily simulated streamflow timeseries bias for two schemes. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Figure 3.40 summarizes the performance of selected GPDs, including observed precipitation for streamflow simulation over the Karasu basin under NSE and KGE with its components (R, Bias, VR) for the calibration, validation, and entire periods considering two different model parameter calibration scenarios. Overall, GPDs show relatively higher correlation for both scenarios and the ratio of bias is highly overestimated by many GPDs and most of them underestimate variability ratio in Scheme-1 compared to Scheme-2. When the observed precipitation is utilized in the model to simulate streamflow, it reproduces relatively high KGE (0.92) and NSE (0.84) with almost no bias and variability ratio for the calibration period. In the same way, the model with observed precipitation is able to keep the higher performance for the validation period with KGE (0.83) and NSE (0.75) scores. In Scheme-1, CPCv1 performs close to the observed results for daily streamflow simulation with a slight overestimation in bias and underestimation in variability ratio for all phases. Among multi-source GPDs, CHIRP(S)v2.0 display a good performance for the calibration and validation periods as compared to MSWEPv2.8, MSWXv100, and MERRA-2, which, interestingly, presents better results for its validation stage than calibration.

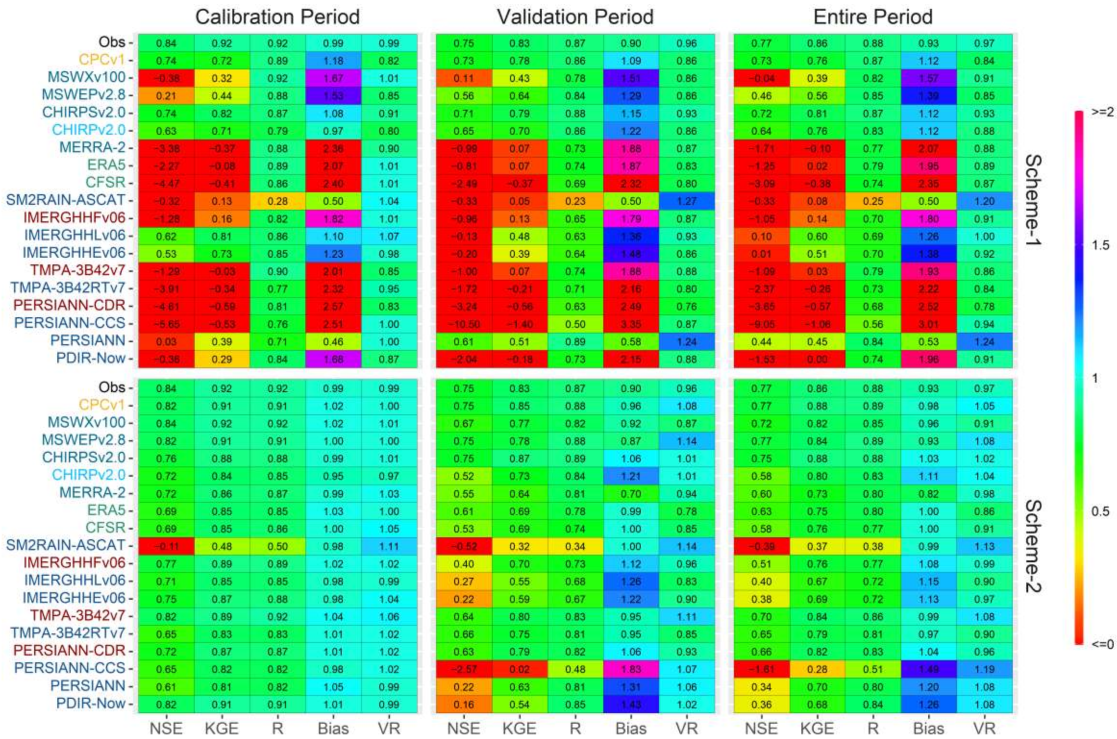

Figure 3.40. *Performance of daily streamflow simulations at the Karasu basin outlet using the gauge and 18 GPDs data considering calibration, validation, and entire period. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Moreover, IMERGHHFv06, TMPA-3B42v7, PERSIANN-CDR, and ERA5 perform poorly in simulating the streamflow for Scheme-1. Among satellite-based GPDs, both IMERGHHEv06 and IMERGHHLv06 show positive performance for the calibration period with their performance decreasing (with negative NSE) for validation and the entire period. TMPA-3B42RTv7 and PERSIANN-CCS perform poorly in simulating the streamflow for all periods and show negative KGE and NSE. The

PERSIANN dataset shows varying performance compared to other GPDs and underestimates bias and overestimates the variability ratio. When the model parameters are calibrated based on each GPD individually (Scheme-2), all GPDs simulate streamflow with high performance for the calibration period. MSWEPv2.8 shows close reproducibility of streamflow to CPCv1 and observed discharge for calibration, validation, and the entire period. Comparatively, PERSIANN-CCS shows high KGE (0.82) and NSE (0.65) during the calibration period but its reproducibility for streamflow generation is poor for validation and the entire period. Among all GPDs, CHIRPv2.0 and PERSIANN behave differently in Scheme-2. Both simulate streamflow well for the calibration phase while their performance decreases during validation and the entire period as compared to Scheme-1. PERSIANN-CCS significantly overestimates bias in both schemes. Overall, the TUW model shows good performance over snow dominant catchments as noted in the literature, and the relatively shorter time span for model calibration is one of the reasons for higher KGE. The optimum TUW model parameters for observed precipitation and selected GPDs are summarized in Figure 3.41.

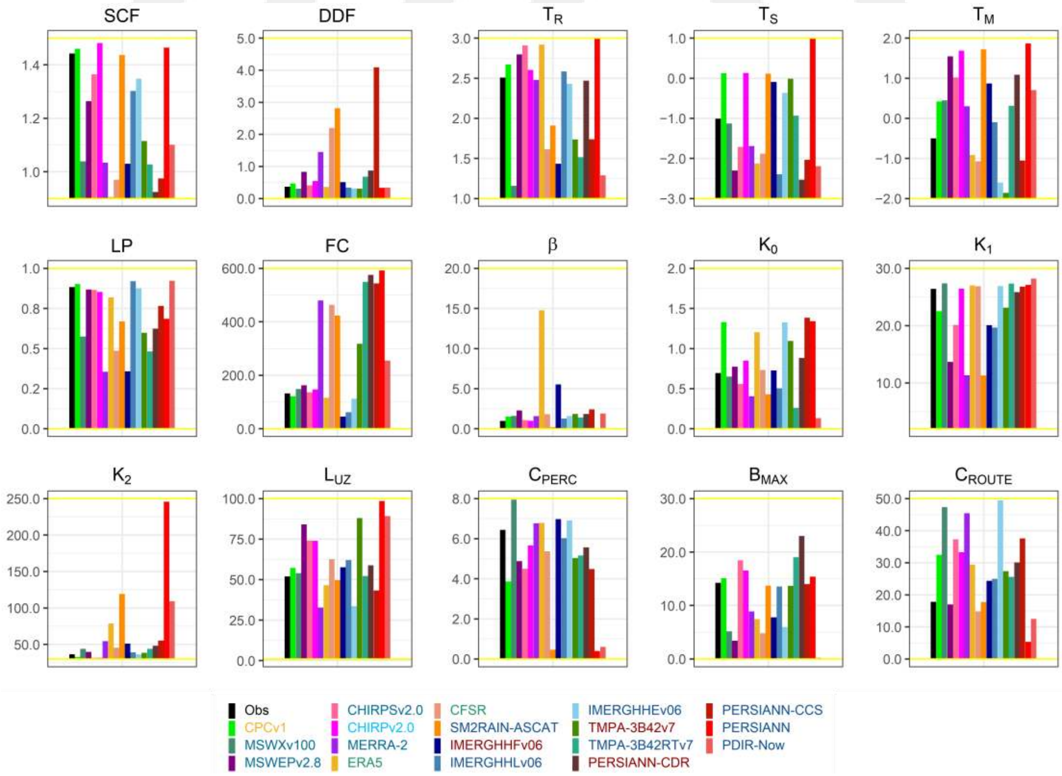

Figure 3.41. *Optimum TUW model parameters for the observed and 18 GPDs. Legend text color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Among GPDs, PERSIANN-CCS shows quite a different DDF compared to other GPDs. In the same way, ERA5 shows a high Beta (β), and PERSIANN displays a high K2 value. Additionally, TMPA and PERSIANN datasets show high FC compared to the rest of the GPDs. Further uncertainties arising from meteorological forcing and hydrological models may surely have an effect on streamflow simulations, but a detailed sensitivity/uncertainty analysis is not considered within the scope of this study.

3.2.2. GPDs over upper Aras (Kayabaşı) basin

The Upper Aras River basin (Kayabaşı) is the second basin which has been selected for GPDs hydrological utility and is located in the vicinity of Karasu basin in the eastern part of Turkey. Compared to Karasu basin, the Kayabaşı basin is relatively smaller in terms of drainage area (i.e., the catchment area of Kayabaşı is only 26.63% of Karasu basin drainage area). Hence, the spatial variation of precipitation in this basin is expected to be lower than that in the Karasu basin. However, ten meteorological stations that are located in or in the vicinity of the basin are selected for the direct comparison and hydrological utility of selected GPDs.

3.2.2.1. Direct comparison of GPDs

Figure 3.42 depicts the mean daily precipitation and precipitation bias at the grid-cell location for the observed precipitation and selected 18 GPDs over Kayabaşı basin for the entire period (2015–2019). According to the observed precipitation, the spatial variation of precipitation over Kayabaşı is less than in the Karasu basin. This may be due to the area difference between the two basins, and the mean daily precipitation of 1 to 1.5 mm is observable for most stations in the basin. In the same way, CPCv1 displays mean precipitation close to that observed inside the basin while showing higher precipitation in the southern part of the basin, which shows the opposite precipitation compared to the southern part (i.e., CPCv1 underestimates observed precipitation in the southwest parts of the Karasu basin). MSWXv100, MSWEPv2.8, CHIRP(S)v2.0, and MERRA-2 map precipitation that is close to the observed, especially inside the basin, where there are some discrepancies for precipitation outside the basin. Furthermore, IMERGHHFv06, TMPA-3B42v7, and PERSIANN-CDR overestimate mean daily precipitation with different biases over the Kayabaşı basin, which is also the case for these three GPD.

Among reanalysis GPDs, ERA5 shows precipitation close to observed where CFSR highly overestimates mean precipitation over the basin.

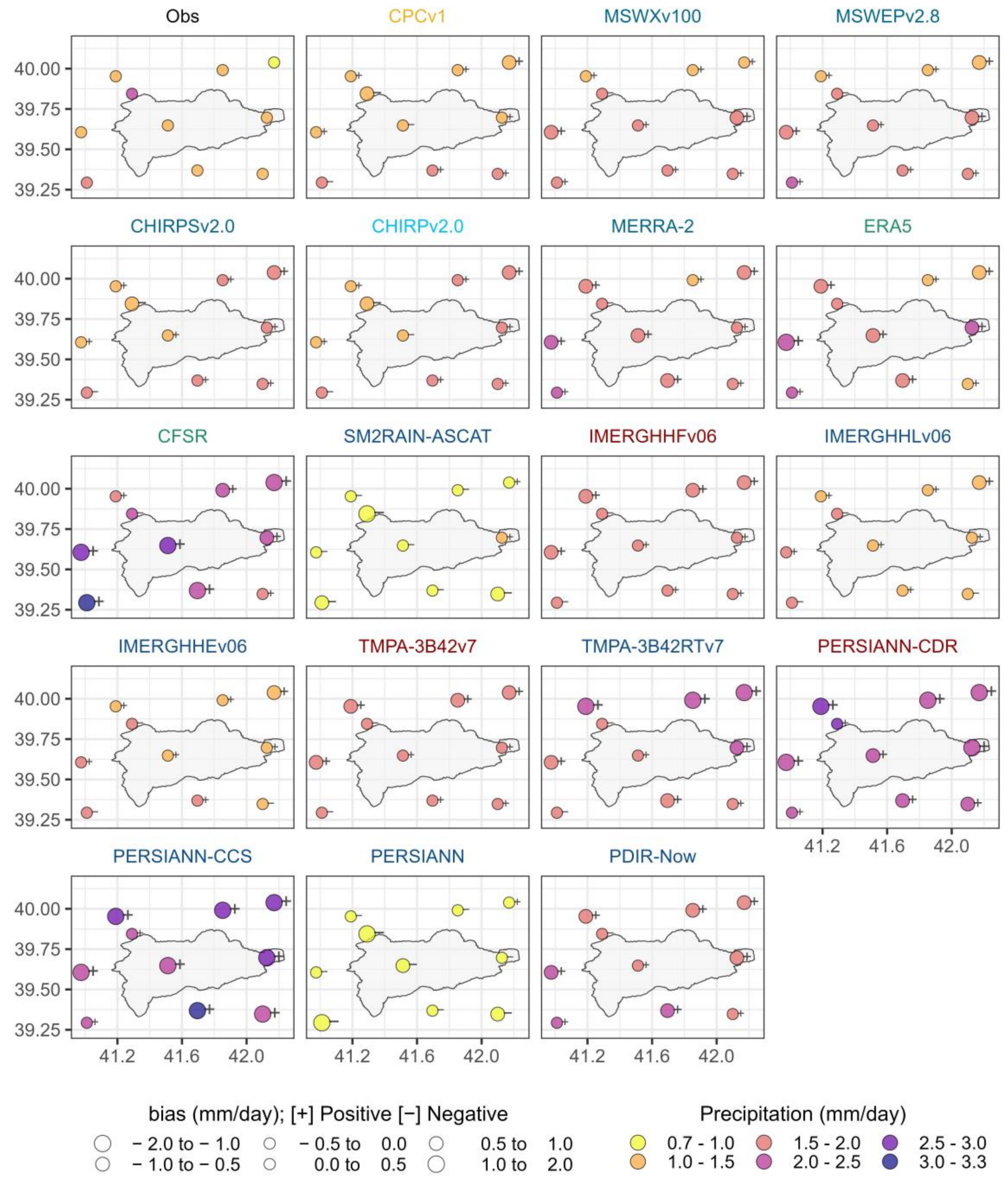

Figure 3.42. *Mean daily precipitation and bias at station level over the Kayabaşı river basin for the period of 2015–2019 from the gauge observation (Obs) and 18 GPDs. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Finally, satellite-based GPDs demonstrate a mean daily precipitation pattern that is similar to their performance over the Karasu basin. However, both PERSIANN and

SM2RAIN-ASCAT highly underestimate mean precipitation, where PERSIANN-CCS and TMPA-3B42RTv7 show higher precipitation in the region compared to observed precipitation, while IMERGHHE(L)v06 presents mean daily precipitation close to observed precipitation, followed by PDIR-Now.

Figure 3.43 presents the mean daily precipitation from observed and 18 GPDs and their bias at the regional scale for the entire study period and four seasons over Kayabaşı basin. From the observed precipitation, the region receives higher daily precipitation during the spring (2.03mm) season followed by the winter (1.34mm) season, while the region experiences a lower amount of daily precipitation during the summer (0.77 mm) season followed by the autumn (1.24 mm) season and the region receives 1.35 mm mean daily precipitation for the entire study period. Compared to Karasu basin, Kayabaşı receives lower precipitation during the entire study period and the spring and winter seasons, while this region receives slightly higher precipitation during the summer and autumn seasons.

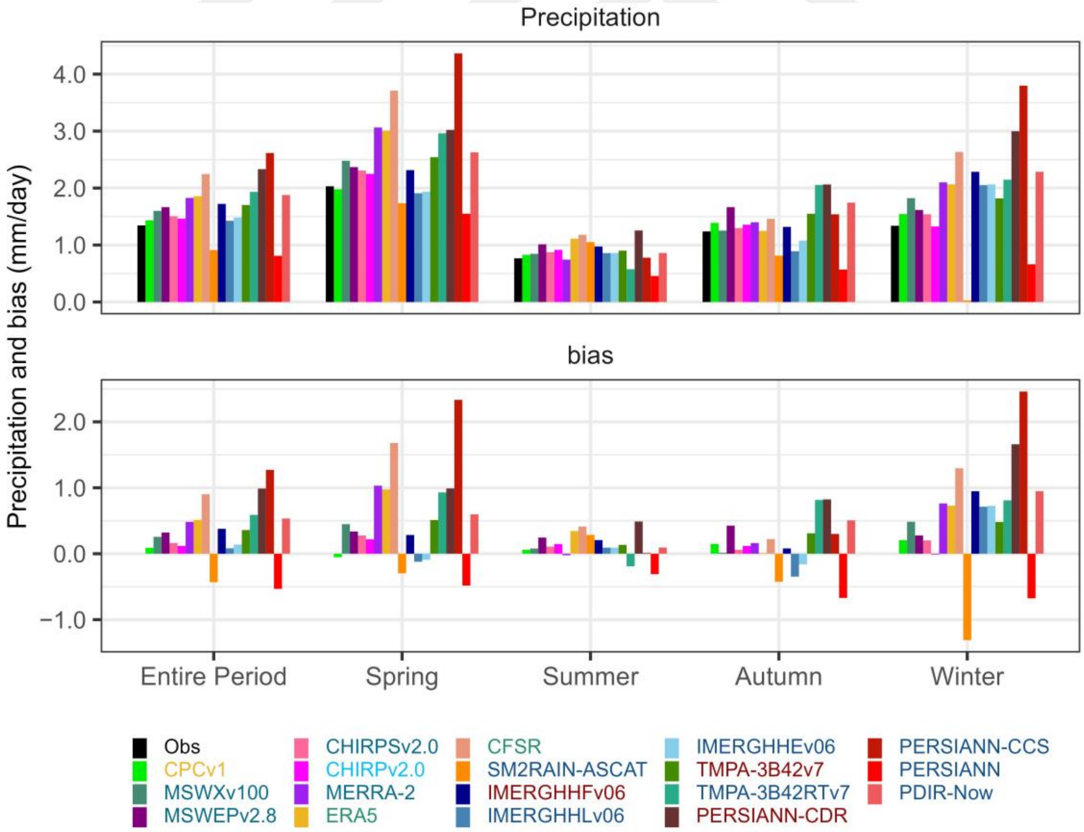

Figure 3.43 *Mean daily and monthly precipitation and estimated bias at the regional scale over Kayabaşı basin for the entire period and four seasons. The color of legend text present: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

CPCv1 displays mean daily precipitation close to the observed precipitation for the entire study period and four seasons where the highest mean daily precipitation bias is not exceeded by +0.2 mm, which is obtained for the autumn season. Among multi-source GPDs, CHIRP(S)v2.0 shows mean daily precipitation close to the observed with a slight overestimation for the entire study period and four seasons, followed by MSWXv100, MSWEPv2.8, IMERGHHFv06, and TMPA-3B42v7, while PERSIANN-CDR highly overestimates mean daily precipitation, followed by MERRA-2 dataset, comparatively. Both ERA5 and CFSR show relatively higher precipitation, especially for the spring and winter seasons. Among satellite-based GPDs, PERSIANN consistently shows lower mean daily precipitation and underestimates precipitation for the entire study period and four seasons, followed by SM2RAIN-ASCAT (except for the summer season), while IMERGHHL(E)v06 delivers mean daily precipitation close to the observed precipitation, whereas PDIR-Now consistently overestimates mean daily precipitation for the entire study period and four seasons.

Figure 3.44 maps the reliability of GPDs at the grid-cell level expressed in the form of Kling-Gupta Efficiency (KGE) and its Pearson correlation, the ratio of bias and variability ratio components for the entire study period over Kayabaşı basin. CPCv1 relatively shows higher performance over each station compared to other selected GPDs, followed by some gauge-corrected, such as MSWEPv2.8, MSWXv100, MERRA-2, CHIRP(S)v2.0, and IMERGHHFv06, while TMPA-3B42v7 and PERSIANN-CDR perform relatively poorly. Compared to Karasu basin, IMERGHHFv06 and MERRA-2 show higher performance over Kayabaşı basin. Among two reanalysis datasets, ERA5 outperforms CFSR and represents a higher KGE inside the basin than CFSR. Among satellite-based GPDs, IMERGHHL(E)v06 and PERSIANN outperforms PERSIANN-CCS and PDIR-Now, whereas TMPA-3B42RTv7 is outperformed by SM2RAIN-ASCAT, which was also the case for the Karasu basin. GPDs show relatively higher correlation at the station location and most of the datasets show overestimation of the ratio of bias and under estimation of variability ratio.

The reliability The reliability of selected GPDs at the regional scale over Kayabaşı basin is expressed in the form of the KGE and its three components for the entire study period and all four seasons (Figure 3.45). The findings over Karasu basin, where GPDs show higher performance during the spring and autumn seasons than during the summer

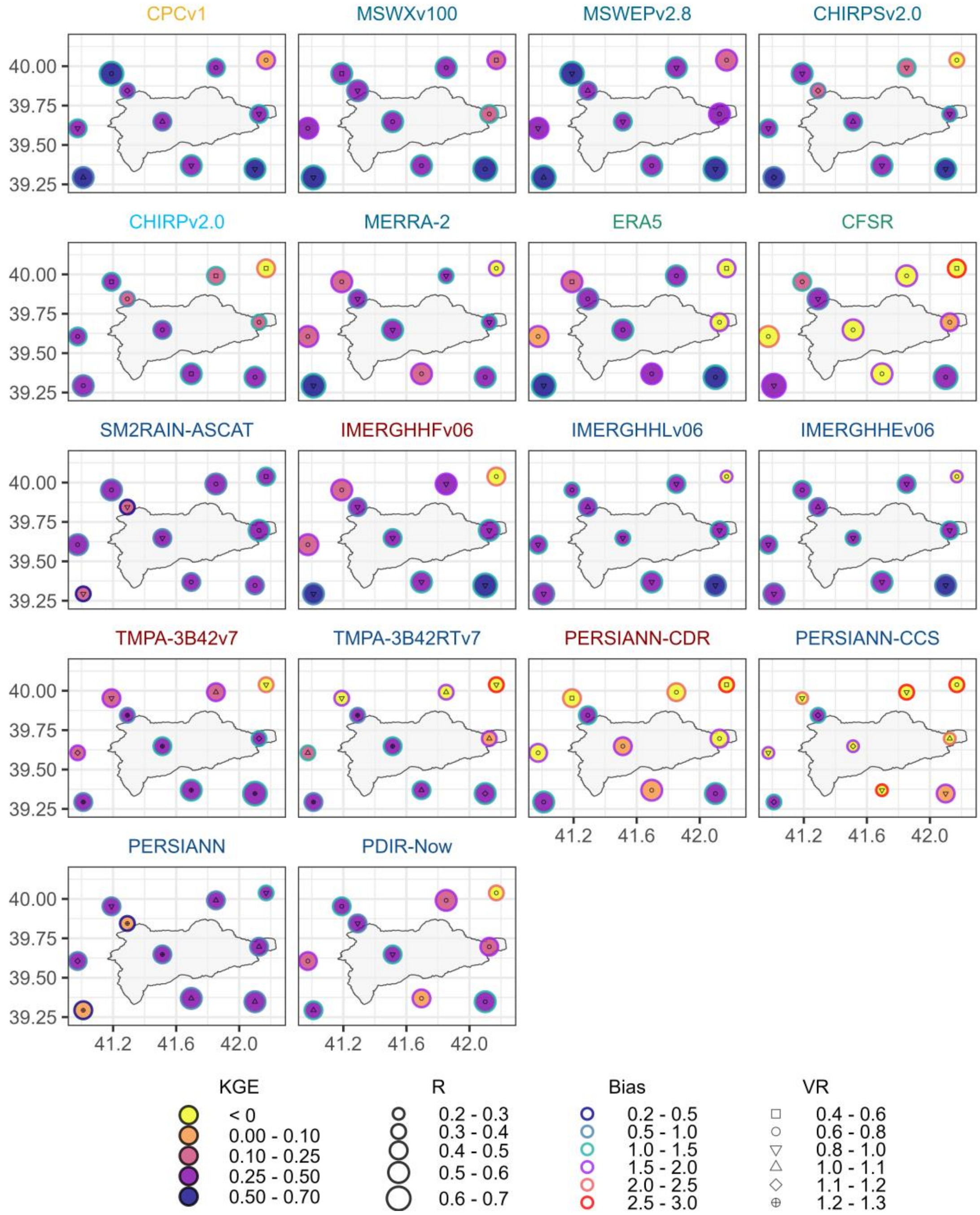

Figure 3.44 *Reliability of GPDs at the station level expressed in the form of KGE and its individual components at a daily time step over the Kayabaşı river basin. The title color presents the source (s) of GPDs: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

and winter seasons, can be applied to GPD performance over Kayabaşı basin. As a result, GPDs show slightly higher performance over Kayabaşı than the Karasu basin. Overall, IMERGHHFv06 outperforms other selected GPDs during the autumn season, followed by MERRA-2, IMERGHHEv06, and MSWXv100, where some GPDs such as PERSIANN-CCS, SM2RAIN-ASCAT, and PERSIANN-CDR show lower performance during the winter season, comparatively. On the other hand, GPDs show higher correlation during the autumn season, which is in line with final KGE values, while most datasets drastically overestimate the ratio of bias during the winter season, and the reliability of selected GPDs at the regional scale over Kayabaşı basin is expressed in the form of the KGE and its three components for the entire study period and all four seasons (Figure 3.45).

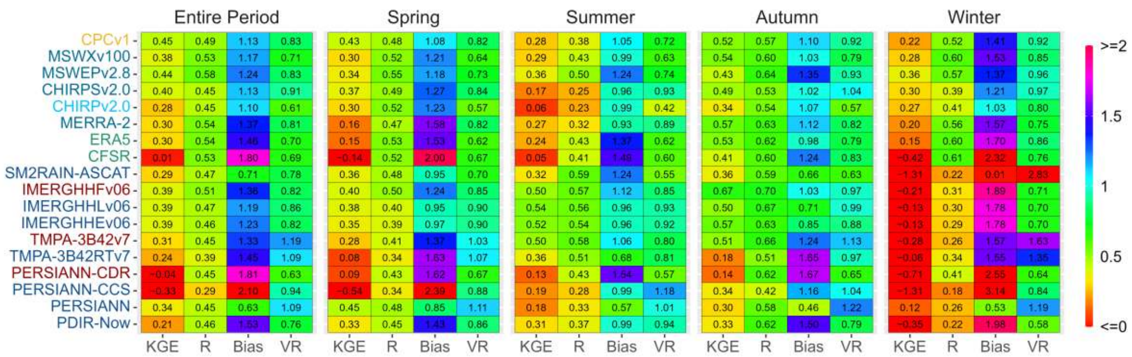

Figure 3.45. *Reliability of GPDs at the regional scale expressed in the form KGE and its individual components for daily precipitation over Kayabaşı basin for the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The findings over Karasu basin, where GPDs show higher performance during the spring and autumn seasons than during the summer and winter seasons, can be applied to GPD performance over Kayabaşı basin. As a result, GPDs show slightly higher performance over Kayabaşı than the Karasu basin. Overall, IMERGHHFv06 outperforms other selected GPDs during the autumn season, followed by MERRA-2, IMERGHHEv06, and MSWXv100, where some GPDs such as PERSIANN-CCS, SM2RAIN-ASCAT, and PERSIANN-CDR show lower performance during the winter season, comparatively. On the other hand, GPDs show higher correlation during the autumn season, which is in line with final KGE values, while most datasets drastically overestimate the ratio of bias during the winter season, and PERSIANN-CCS shows the highest overestimation of observed during the winter season. Finally, GPDs mostly underestimate the variability ratio for the entire period and different seasons, and only the

TMPA-3B42(RT)v7 and PERSIANN datasets show overestimation of the variability ratio over many temporal periods.

Figure 3.46 presents the frequency of observed precipitation and selected GPDs for five distinct daily precipitation intensities over Kayabaşı basin for the entire study period and four seasons. Considering observed daily precipitation frequencies, around 77% of events fall within the range of 0–1 mm for the entire study period, and this magnitude increases for the summer and autumn seasons, while the region experiences a high number of rainy days during the spring and winter seasons, which is also the general case for its adjacent basin (Karasu basin). Furthermore, for the entire period and each individual season, the probability of precipitation occurrence decreases as intensity increases. In general, GPDs exhibit distinct precipitation frequencies and, for some, it is hard to distinguish no-rainy days from light and moderate precipitation, especially during the rainy season such as the spring and winter seasons. For example, excluding a few GPDs, the rest of the datasets demonstrate higher light precipitation (1–5 mm/day) during the spring season, and PDIR-Now is one of the GPDs that shows higher light precipitation (55.91 mm/day) during the spring season. On the other hand, some GPDs, such as SM2RAIN-ASCAT, show lower light precipitation (2.06 mm/day) during the winter season, whereas PDIR-Now shows higher light precipitation during this season. Where the intensity increases, the frequency difference between GPDs and observed decreases.

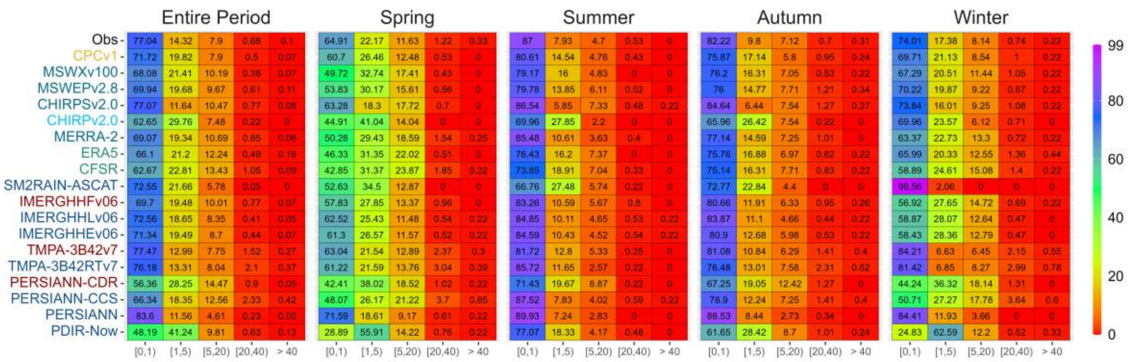

Figure 3.46. *Frequency of GPDs for different daily precipitation intensities over Kayabaşı basin for the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The Hanssen-Kuiper Skill (HKS) score is utilized to analyze the detectability strength of selected GPDs for five different precipitation intensities over Kayabaşı basin, considering the entire study period and four seasons (Figure 3.47). In the general case, which is spotted for GPDs' detectability strength in this study, it is also applicable over Kayabaşı, where many GPDs show higher detectability during the autumn season and lower in the summer season. Additionally, the detection strength of GPDs decreases as the precipitation intensity increases, which is generally the case in literature. However, some discrepancies exist in the detectability strength of GPDs for different precipitation events, considering light and moderate precipitation. From the results, MSWEPv2.8 and CFSR during the autumn season, PDIR-Now during the summer season, and ERA5 during the winter season show higher detection abilities of no precipitation events compared to other selected GPDs. While PDIR-Now and MSWXv100 in the summer season, and MSWEPv2.8 and SM2RAIN-ASCAT during the autumn season for the light precipitation, and SM2RAIN-ASCAT during the summer season, PERSIANN-CDR during the autumn season, CFSR in the winter season, and ERA5 during the spring season for moderate precipitation, show higher detectability skills compared to other GPDs. Finally, some GPDs, such as MERRA-2, TMPA-3B42RTv7, PERSIANN-CDR, PDIR-Now, and IMERGHHFv06 for heavy precipitation events and IMERGHHFv06 and TMPA-3B42v7 during the autumn season, show higher detection abilities and outperform other GPDs.

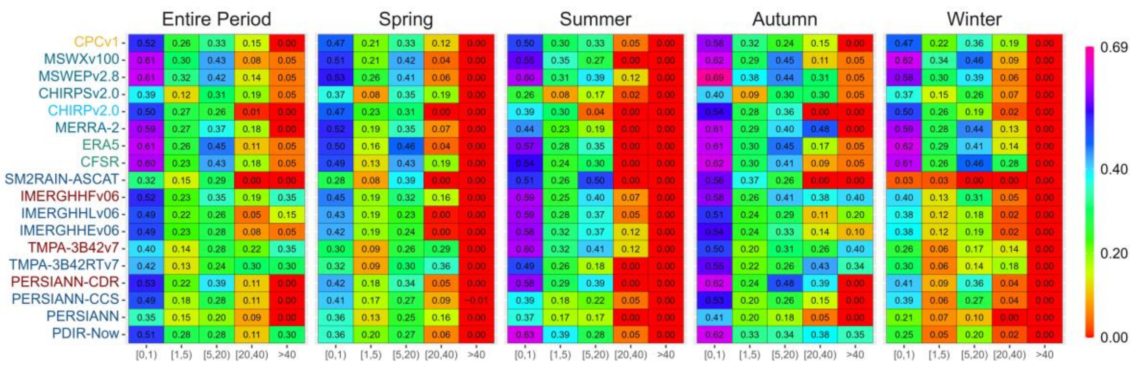

Figure 3.47. *Detection ability in reproducing daily precipitation intensities expressed in the form of the Hanssen–Kuiper Skill (HKS) score over Kayabaşı basin considering the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

3.2.2.2. Hydrologic utility of GPDs over Kayabaşı basin

The procedure for streamflow simulation over the Kayabaşı basin is the same as for the Karasu river basin; first observed precipitation over the basin is exploited for TUW model parameters calibration, and with the observed precipitation-based optimum model parameters, all GPDs are forced to simulate streamflow (Scheme-1). Then the model parameters are calibrated/validated and streamflow is simulated by each GPD individually (Scheme-2). Hence, in the second scenario, all GPDs have their optimum parameters set. Furthermore, the first two years (2015–2016) of the study period are selected for model calibration and the last three years (2016–2019) are utilized for model validation when they are forced by observed precipitation and each GPD separately.

Figure 3.48 maps the observed discharge hydrograph (gray shaded area) and simulated streamflow hydrograph at the outlet of Kayabaşı basin, obtained from observed precipitation and each GPD under two different model parameterization scenarios. The result indicates that, when the model is forced to simulate streamflow with observed precipitation-based optimum parameters (Scheme-1), observed precipitation demonstrates higher streamflow reproducibility than GPDs in terms of KGE and NSE, and GPDs are able to predict streamflow with higher accuracy with their own model parameters set (Scheme-2), especially for the calibration period.

Comparing the temporal dynamics of simulated streamflow against observed discharge by exploiting their scatter plots (Figure 3.49) for the entire study period provides detailed information on two model calibration scenarios and GPDs streamflow reproducibility. In general, GPDs show higher simulated streamflow for Scheme-1 than in Scheme-2, but the discrepancies between the two scenarios over Kayabaş basin are not as great as over the Karasu basin (Figure 3.37). However, some GPDs, including SM2RAIN-ASCAT and PERSIANN, underestimate observed streamflow in Scheme-1, which was also the general case for these datasets over Karasu basin. Finally, only PERSIANN-CDR and IMERGHHFv06 clearly overestimate observed streamflow in Scheme-1, where for the Karasu basin, in addition to these datasets, MERRA-2, ERA5, CFSR, IMERGHH datasets, TMPA datasets, PERSIANN-CDR, PDIR-Now, and PERSIANN-CCS also drastically overestimate the observed streamflow in Scheme-1. However, the stochastic nonlinearity between observed discharge and simulated streamflow, which was observable over Karasu basin, does not appear for the simulated streamflow in this basin. This may be due to the GPD bias difference in these two basins, climatological discrepancies such as the magnitude and form of precipitation, and the size of the basin.

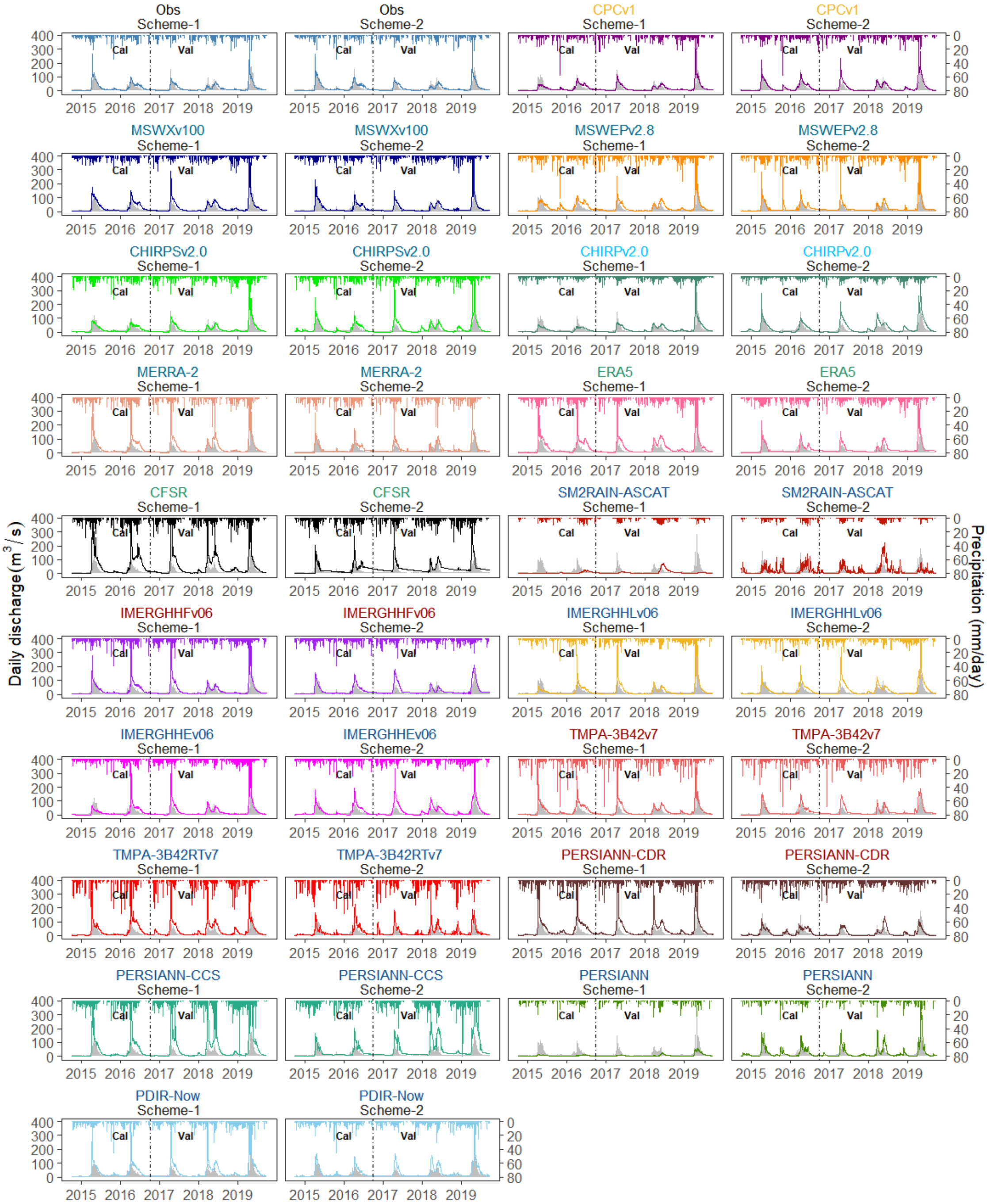

Figure 3.48. *Comparison of the observed and simulated daily hydrographs at the Kayabaşı basin outlet based on the observed precipitation and 18 GPD input for the calibration (October 2014 to September 2016) and validation period (October 2016 to September 2019). Title color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The simulated streamflow bias distribution based on observed precipitation and GPDs is expressed in the form of a box-and-whiskers plot for the two years of calibration and three years of validation periods over Kayabaşı basin. GPDs show a higher wet bias for Scheme-1 than in Scheme-2. Overall, GPDs show lower bias than over Karasu basin, which is definitely a function of observed discharge magnitude. As it stated in the direct comparison, the study area experienced higher precipitation in 2015 and 2019 than in 2017, which is known as the driest year in the selected study time window.

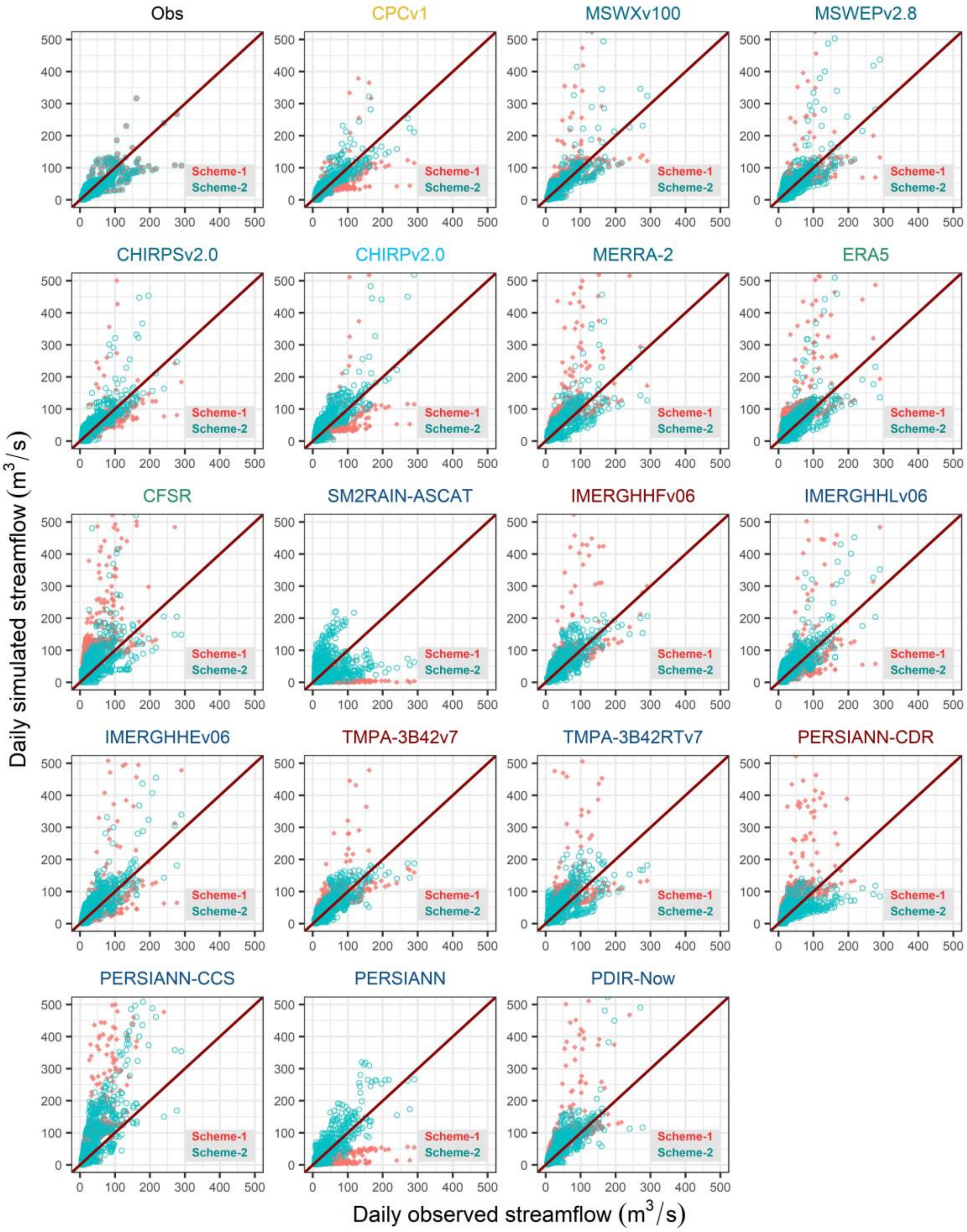

Figure 3.49. *Scatter plot of daily simulated streamflow against observed discharge for the observed precipitation and 13 GPDs at Kayabaşı basin outlet. Title color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Regardless of calibration and validation periods, many GPDs demonstrate lower streamflow bias during the wet years (2015 and 2019) and the dry (2017), where higher streamflow bias is observable for the rest of the study years. From the results, SM2RAIN-ASCAT constantly underestimates observed streamflow over each year for both scenarios (except in Scheme-2 during 2018), followed by the PERSIANN dataset. Both ERA5 and CFSR highly overestimate observed streamflow, especially in Shceme-1 where TMPA-3B42v7 shows relatively higher reproducibility than IMERGHHFv06 for both scenarios.

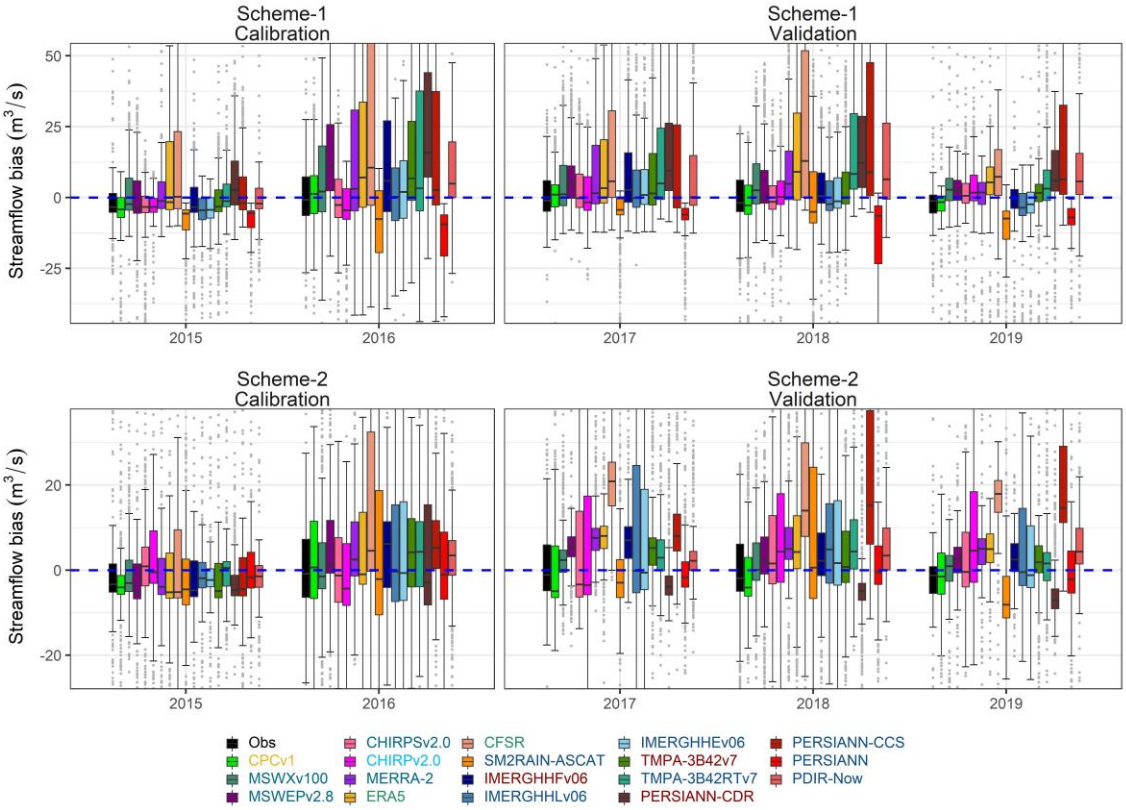

Figure 3.50. *Daily simulated streamflow bias distribution over Kayabaşı basin is expressed in the form of box and whiskers, considering the calibration and validation years. The boxes represent the 25th, 50th, and 75th percentile values, respectively. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Figure 3.51 displays the simulated streamflow timeseries bias of observed precipitation and GPDs for the entire study period under two different scenarios. As it was mentioned earlier (Figure 3.50 and 3.49), many GPDs, including MERRA-2, ERA5, CFSR, IMERGHHFv06, TMPA-3B42RTv7, PERSIANN-CDR, PERSIANN-CCS, and PDIR-Now highly overestimate observed streamflow, especially for peak discharge in each year, while SM2RAIN-ASCAT and PERSIANN exhibit lower simulated streamflow in Scheme-1. In general, CPCv1, MSWEPv2.8, MSWXv100, and CHIRP(S)v2.0 show lower streamflow bias for both scenarios. Furthermore, IMERGHHFv06 shows a higher streamflow bias compared to its predecessor (TMPA-3B42v7) over Kayabaşı basin.

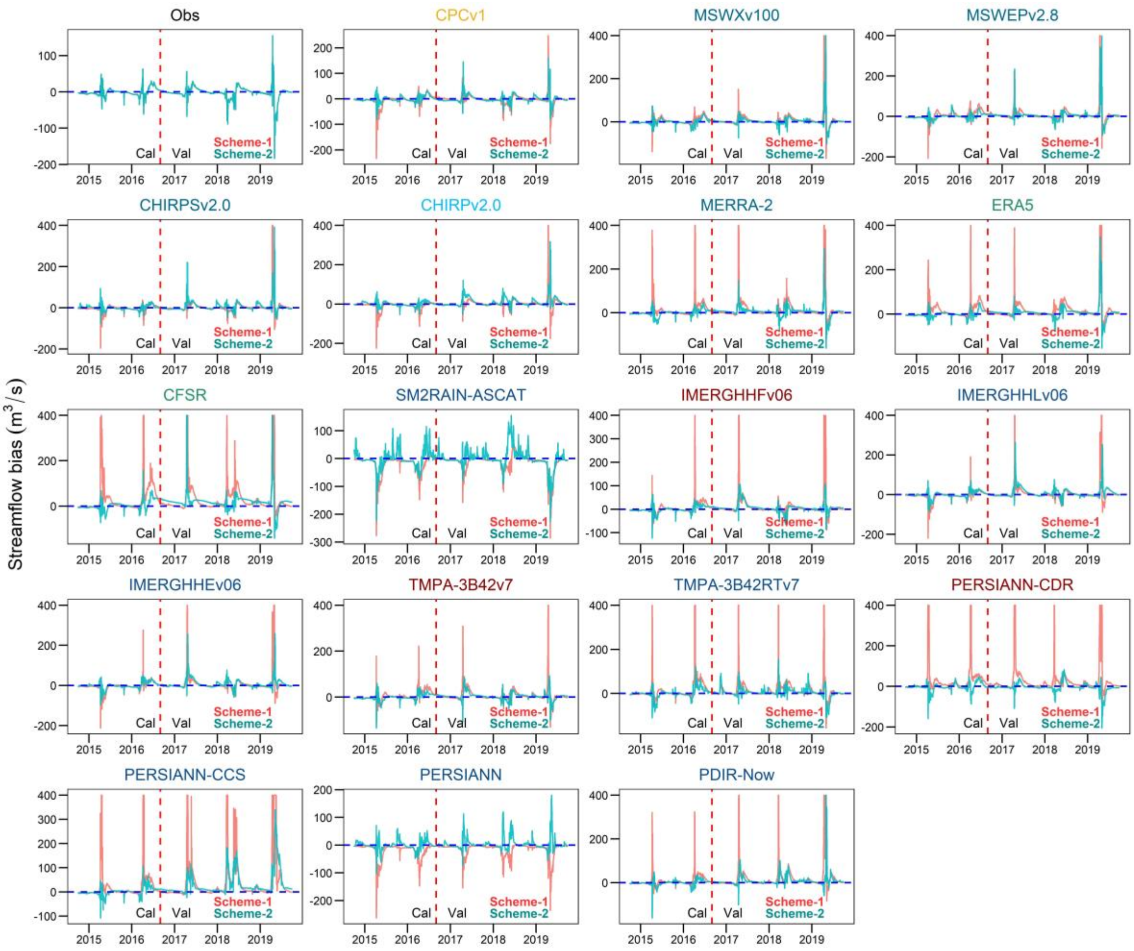

Figure 3.51. *Daily simulated streamflow timeseries bias for two schemes over Kayabaşı basin. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The reproducibility of streamflow by GPDs and observed precipitation are evaluated by NSE and KGE and its three components over Kayabaşı basin (Figure 3.52). Generally, GPDs show higher correlation for both Schemes, and overestimation of bias in Scheme-1 is common, although datasets overestimate variability ratio in Scheme-1 when compared with GPDs variability ratio over Karasu basin. Overall, observed precipitation and GPDs show streamflow reproducibility close to over Karasu basin, which is indicated by higher reproducibility in Scheme-2 than in Scheme-1 and better streamflow simulation during the calibration period for Scheme-2. However, all GPDs and observed precipitation show larger performance differences over the calibration and validation periods, and their performance sharply decreases for the validation period in both scenarios. Furthermore, TMPA-3B42v7 outperforms IMERGHHv06 datasets for both calibration and validation periods in Scheme-1 and for the validation period in Scheme-2. In the same way, all PERSIANN datasets are outperformed by PERSIANN (except PDIR-Now, which shows higher KGE for the calibration periods in both scenarios) in both scenarios. Finally, MERRA-2 shows higher performance over Kayabaşı basin in Scheme-1 compared to its performance over Karasu basin.

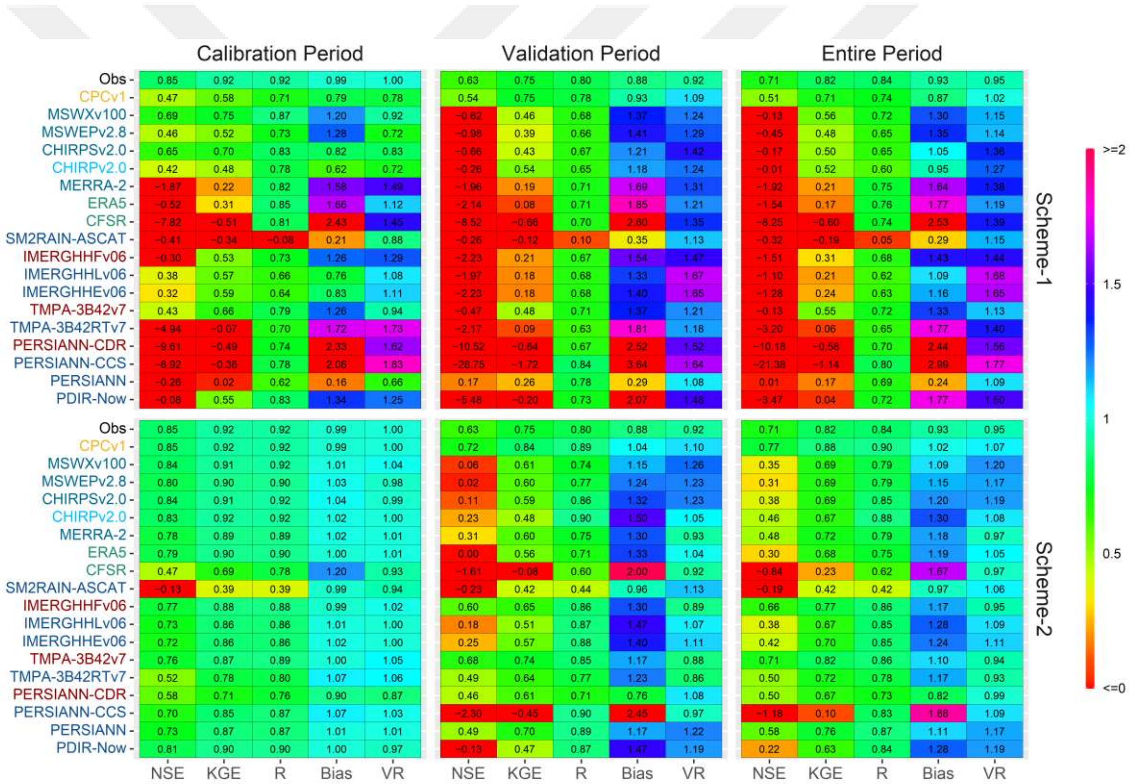

Figure 3.52. *Performance of daily streamflow simulations at the Kayabaşı basin outlet using the gauge and 18 GPDs data considering calibration, validation, and entire period. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Figure 3.52 presents the optimum model parameter set for the observed precipitation and selected GPDs for streamflow simulation over Kayabaşı basin. Compared to the optimal model parameter set of GPDs obtained for Karasu basin, some model parameters, including Snow Correction Factor (SCF), Degree Day Factor (DDF), Beta (β), and K2, show higher values for many GPDs over Kayabaşı basin, while some model parameters, such as melting temperature (TM), CPERC show lower values for some GPDs comparatively. Additional uncertainties coming from meteorological forcing and hydrological models will undoubtedly have an impact on streamflow simulations, but a comprehensive sensitivity/uncertainty analysis is not addressed within the scope of this study.

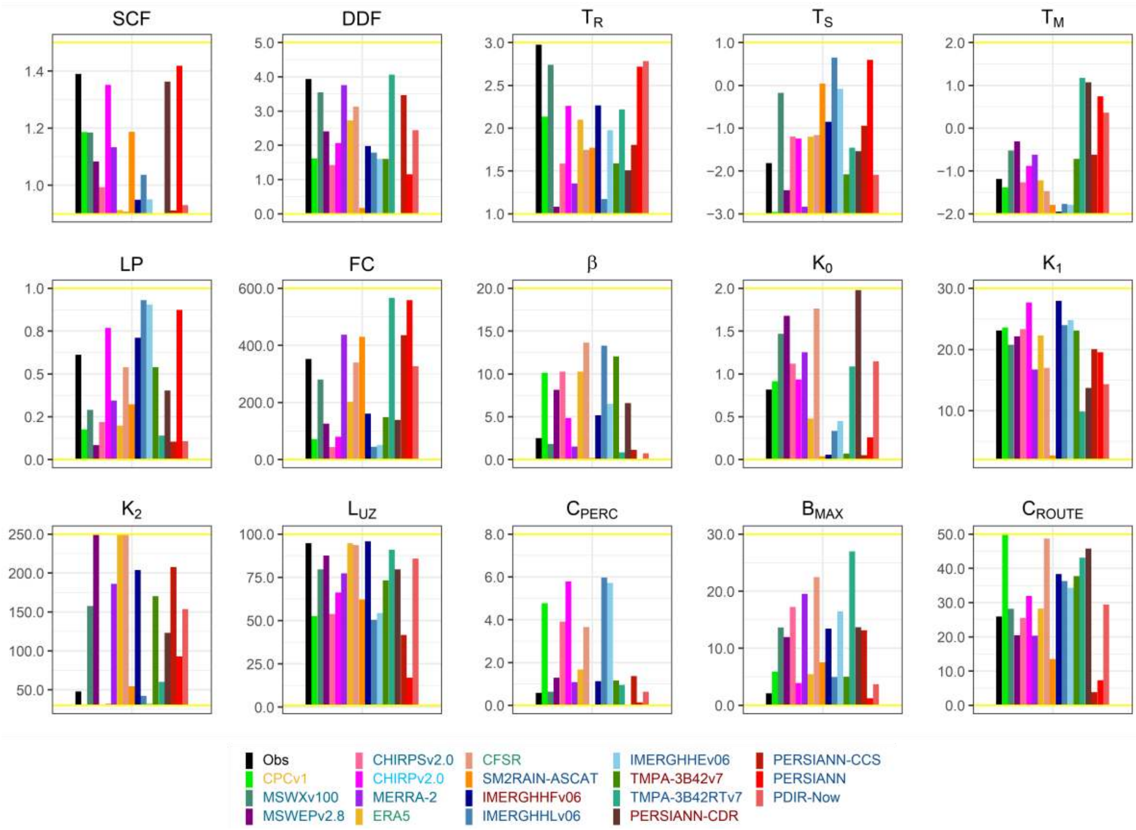

Figure 3.53. *Optimum TUW model parameters for the observed and 18 GPDs over Kayabaşı basin. Legend text color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

3.2.3. GPDs over upper Seyhan (Çukurkışla) basin

The Çukurkışla basin is the third and smallest basin in this study, which is selected for the hydrological utility of GPDs, situated in the southern part of the country with a drainage area of around 1510 km^2. The size of this basin is only about 14.73% of the size of Karasu basin and 55.31% of the size of Kayabaşı basin. Hence, it is expected that the spatial variation of precipitation over this basin is relatively lower than in the Karasu and Kayabaşı catchments. To assess the precipitation regime changes and hydrological utility of GPDs, 12 meteorological stations within and around the basin are utilized for both direct comparison and streamflow simulation.

3.2.3.1. Direct comparison of GPDs

Figure 3.54 displays the mean daily precipitation from observed and selected GPDs with their biases at the grid-cell level over Çukurkışla basin for the entire study period (2015–2019). Taking observed precipitation into consideration, the basin and surrounding area experience daily precipitation ranging from 1 to 1.5 mm, and this amount increases to 2.5-4.0 mm in the southern part of the basin.

The CPCv1, which uses ground spatial information, shows lower precipitation than observed and mostly underestimates precipitation with higher bias in the station levels at the southern part of the basin; the CPCv1 underestimation of mean daily precipitation rate is greater in this basin than in the Karasu and Kayabaşı basins. Among multi-source GPDs, MSWXv100, MSWEPv2.8, CHIRP(S)v2.0 deliver the close mean daily precipitation to observed precipitation followed by TMPA-3B42v7 (TMPA-3B42v7 shows better mean daily over this basin compared to Karasu and Kayabaşı basins) and IMERGHHFv06 where MERRA-2 and PERSIANN-CDR show lower mean daily precipitation in the southern parts and overestimates observed mean precipitation in the rest of region. Moreover, ERA5 displays mean daily precipitation better than CFSR, which is the general case for these GPDs in the other basins, where CFSR overestimates observed precipitation in the southeastern part of the basin. Among satellite-based GPDs, the PERSIANN dataset highly underestimates observed precipitation, followed by SM2RAIN-ASCAT dataset. Furthermore, compared to other satellite-based GPDs, PERSIANN-CCS and TMPA-3B42RTv7 show higher mean daily precipitation over the Çukurkışla basin, whereas IMERGHHL(E) v06 and PDIR-Now show precipitation close to the observed. Finally, almost all GPDs underestimate mean daily precipitation in the southeastern part of the basin.

Figure 3.55 presents the mean daily precipitation and bias at the regional scale, obtained from observed precipitation and selected 18 GPDs for the entire study period and four seasons. According to the observed precipitation, the region experiences higher daily precipitation (2.25 mm) during the winter season, followed by the spring season (2.05 mm) and the autumn season (1.1 mm), and receives lower mean daily precipitation amount during the summer season (0.6 mm).

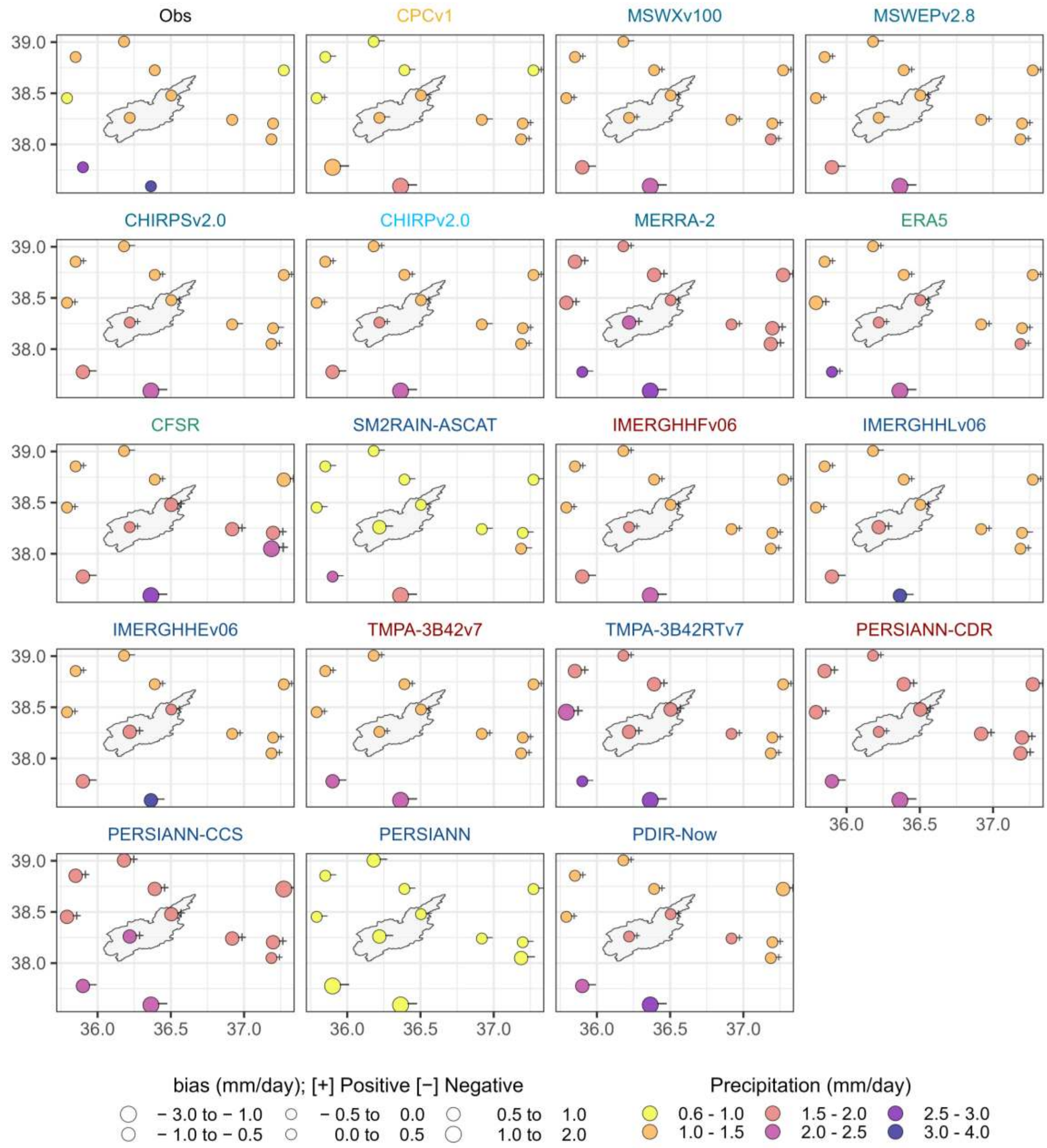

Figure 3.54. *Mean daily precipitation and bias at station level over the Çukurkışla basin for the period of 2015–2019 from the gauge observation (Obs) and 18 GPDs. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Among selected basins, Çukurkışla basin receives higher precipitation during the winter season, while Karasu and Kayabaşı catchments experience higher precipitation amounts in the spring season. From the results, many GPDs underestimated the mean daily precipitation of this basin compared to other selected basins. For instance, CPCv1 and PERSIANN, and MSWEPv2.8 datasets underestimate observed precipitation for the entire study period and four seasons, while SM2RAIN-ASCAT shows higher mean daily precipitation only during the spring season and highly underestimates winter precipitation.

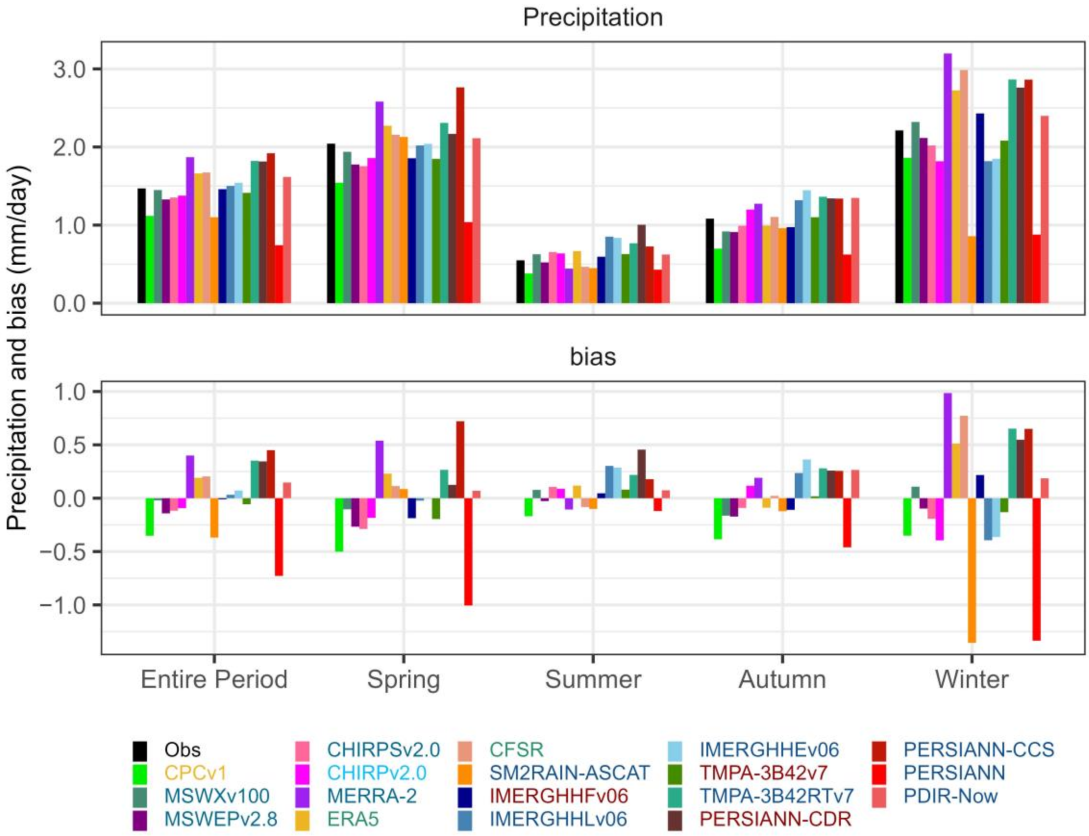

Figure 3.55. *Mean daily precipitation and estimated bias at the regional scale over Çukurkışla basin for the entire period and four seasons. The color of legend text present: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The reliability of GPDs at the grid-cell level over Çukurkışla basin is expressed in the form of KGE and its three components for the entire study period (Figure 3.56). CPCv1 shows relatively higher performance over different stations where some multi-source GPDs, including MSWEPv2.8 and MSWXv100, outperform the only ground-based dataset, and even IMERGHHFv06 and its near real-time version, IMERGHHL(E)v06, show higher performance than CPCv1, which is followed by CFSR.

In the same way, ERA5 was also able to present higher performance compared to some guag-corrected GPDs, such as MERRA-2, CHIPRSv2.0, TMPA-3B42v7, and PERSIANN-CDR, while SM2RAIN-ASCAT was able to show higher performance and outperform CHIRPv2.0, TMPA-3B42RTv7, PERSIANN, PDIR-Now, and PERSIANN-CCS, and even present higher performance compared to PERSIANN-CDR. In the same way, GPDs show a relatively higher correlation to the observed precipitation at the station locations, where most of the datasets overestimate bias and underestimate variability ratio at the grid-cell level.

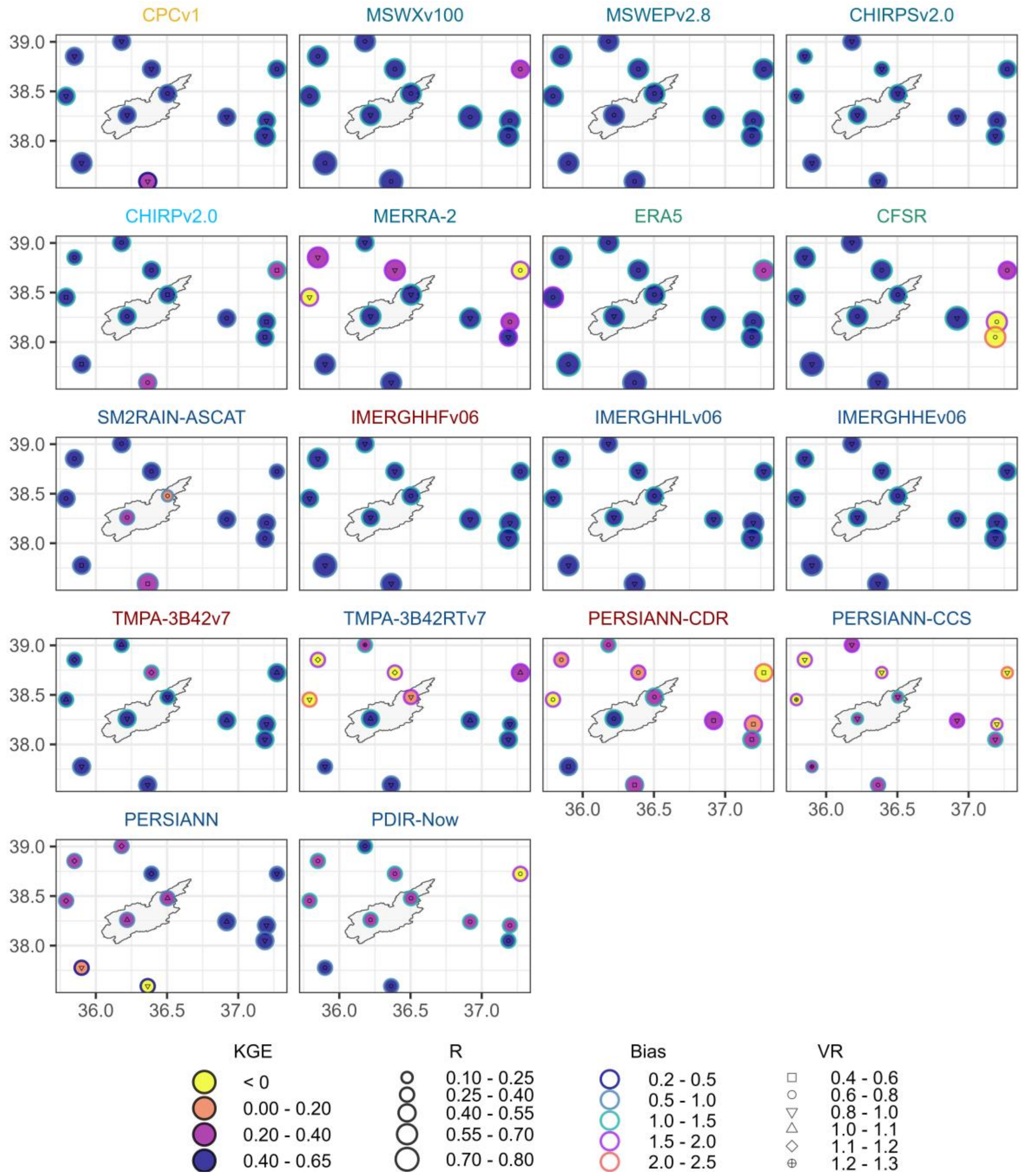

Figure 3.56. *Reliability of GPDs at the station level expressed in the form of KGE and its individual components at a daily time step over the Çukurkışla basin. The title color presents the source (s) of GPDs: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Figure 3.57 presents the reliability of selected GPDs at the regional scale over Çukurkışla basin for the entire study period and four seasons. When considering the seasonal performance, GPDs show higher performance during the winter and autumn seasons, followed by spring seasons, and exhibit lower performance during the summer. Compared to Karasu and Kayabaşı basins, GPDs relatively show higher performance in the winter season than in the spring season. According to the results, IMERGHHFv06, MSWEPv2.8, ERA5, and MSWXv100 during the autumn season outperformed other GPDs, while the IMERGHHL(E)v06 was also able to present higher performance than the other datasets during the spring season. Some GPDs, such as PERSIANN-CCS, TMPA-3B42RTv7, SM2RAIN-ASCAT, PERSIANN-CDR, and PDIR-Now, show relatively lower performance during the winter season. Furthermore, most GPDs show higher bias during the winter season, although the high bias was obtained by PERSIANN-CDR during the summer season. Finally, except for TMPA-3B42(RT)v7 and PERSIANN over some seasons, the rest of the GPDs underestimate the variability ratio at the regional scale over this basin.

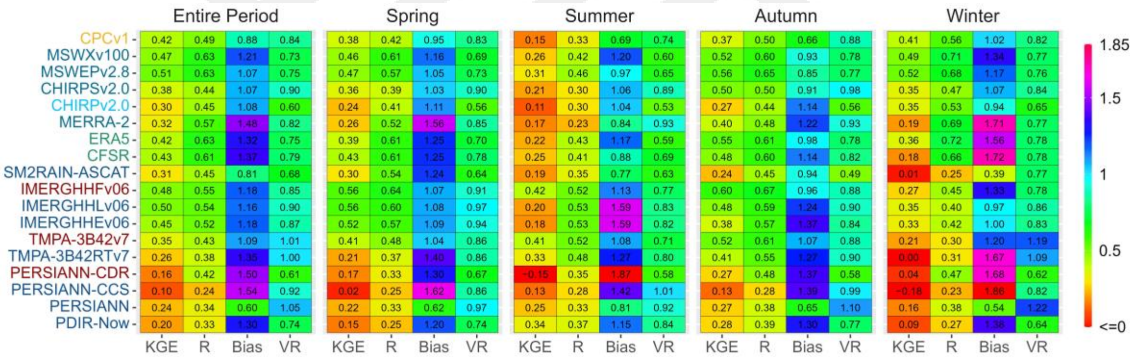

Figure 3.57. *Reliability of GPDs at the regional scale expressed in the form KGE and its individual components for daily precipitation over Çukurkışla basin for the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

The frequency of daily precipitation for five different precipitation intensities from observed precipitation and selected GPDs over Çukurkışla basin for the entire study period and four seasons is shown in Figure 3.58. The observed precipitation records exhibit higher no-rainy days (81.42%) compared to the frequency of no-rainy days over Karasu (77.25%) and Kayabaşı (77.04%) catchments for the entire study period, while the number of no-rainy days' increases during the summer season followed by the winter season and declines for the spring and winter seasons, which is the general case for the other selected basins. Furthermore, larger discrepancies between observed precipitation and selected GPD frequencies for five different precipitation intensities arise during the rainy seasons, and many GPDs are not able to differentiate between no-rain, light precipitation, and moderate precipitation, especially during the spring and winter seasons. In general, the precipitation frequency differences are decreasing as the intensity increases.

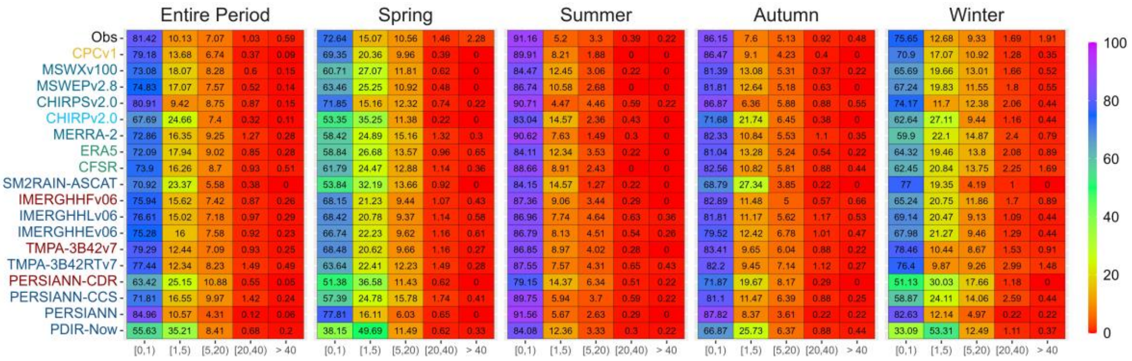

Figure 3.58. *Frequency of GPDs for different daily precipitation intensities over Çukurkışla basin for the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Figure 3.59 presents the detectability strength of GPDs expressed in the form of Hanssen-Kuiper Skill (HKS) score at the regional scale for the entire study period and four seasons. Overall, GPDs show higher detection ability for low-intensity precipitation and lower detectability for precipitation with higher intensities, and except for some GPDs, most datasets show higher detection ability for moderate precipitation than for light precipitation. From the results, PERISANN-CDR during the summer season and ERA5, MSWXv100, and MSWEPv2.8 during the winter season show higher detection skills for no precipitation events, while MSWEPv2.8, PDIR-Now, ERA5, and MSWXv100 during the summer season for light precipitation and ERA5, MSWXv100, CFSR, MERRA-2, and MSWEPv2.8 during the winter season for moderate precipitation demonstrate higher detectability compared to other GPDs. Finally, MERRA-2, TMPA-3B42v7, IMERGHHLv06, and its final version (IMERGHHFv06) during the autumn season for heavy precipitation, ERA5 and MERRA-2 during the winter seasons and entire period, and IMERGHHL(E)v06 during the autumn and entire period for extreme precipitation, demonstrate relatively higher detection ability compared to other datasets.

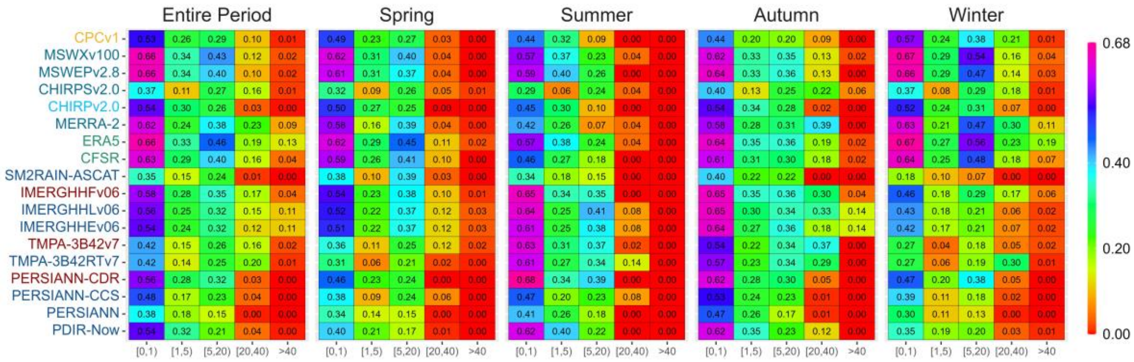

Figure 3.59. *Detection ability in reproducing daily precipitation intensities expressed in the form of the Hanssen–Kuiper Skill (HKS) score over Çukurkışla basin considering the entire period and four seasons. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

3.2.3.2. Hydrologic utility of GPDs over Çukurkışla basin

To assess the hydrological utility of GPDs over Çukurkışla basin, first the model is set up by observed precipitation and the observed precipitation-based model optimum parameter set is employed for GPDs to simulate streamflow (Scheme-1), where in the second scenario, the optimum model parameter set is obtained for each GPDs individually (Scheme-2). Hence, following the previous procedure, the entire study period is separated into two temporal windows for the model calibration (October 2014–September 2016) and validation (October 2016–September 2019), which indicates two years of calibration and three years of validation periods.

Figure 3.60 displays the observed discharge (gray shaded area) and simulate streamflow hydrographs from observed precipitation and selected GPDs for the calibration and validation periods. As it was mentioned earlier, Çukurkışla basin is the smallest basin in terms of drainage area compared to Karasu and Kayabaşı catchments which is appear from the observed hydrograph. From the results, the previous findings over other selected basins in this study are match over Çukurkışla basin where observed precipitation delivers the best reproducibility of streamflow simulate and all GPDs demonstrates higher streamflow reproducibility when the model is forced by each GPD' optimum parameter set (Scheme-2) especially during the calibration periods.

Figure 3.61 maps the temporal dynamics of simulated streamflow by observed precipitation and selected GPDs against observed discharge considering different model calibration scenarios for the entire study period. The results indicate that the simulated streamflow by observed precipitation shows higher linearity over Çukurkışla basin compared to Karasu and Kayabaşı catchments, where CPCv1 mostly underestimates daily observed discharge at different flow rates without considering the model calibration approaches.

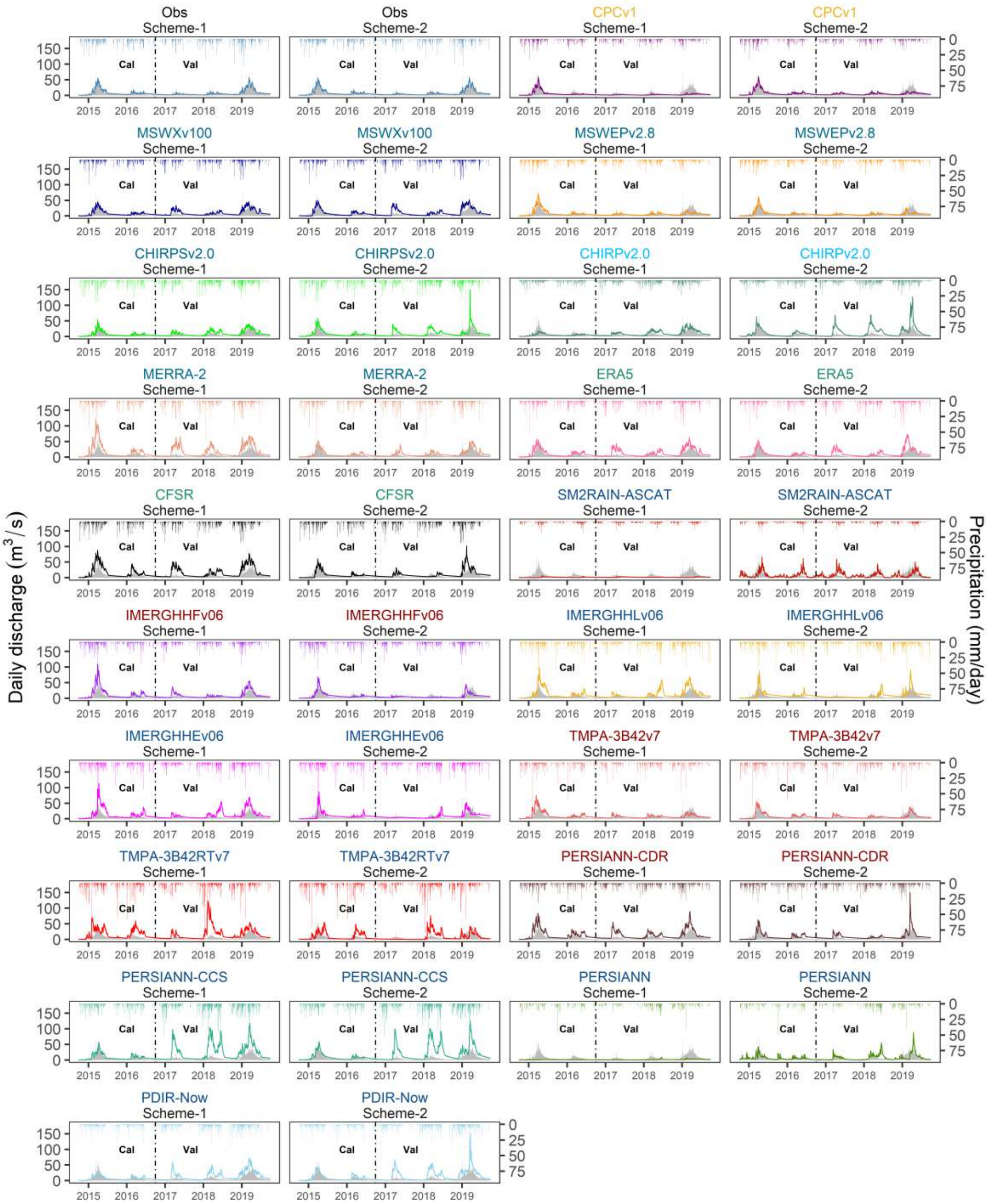

Figure 3.60. *Comparison of the observed and simulated daily hydrographs at the Çukurkışla basin outlet based on the observed data and 18 GPD input for the calibration (October 2014 to September 2016) and validation period (October 2016 to September 2019). Title color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Among multi-source GPDs, MSWXv100 seems to produce streamflow close to observed and outperforms MSWEPv2.8 (i.e., MSWEPv2.8 outperforms MSWXv100 over Karasu and Kayabaşı catchments) over this basin, where MSWEPv2.8 displays a scatter plot similar to CPCv1.

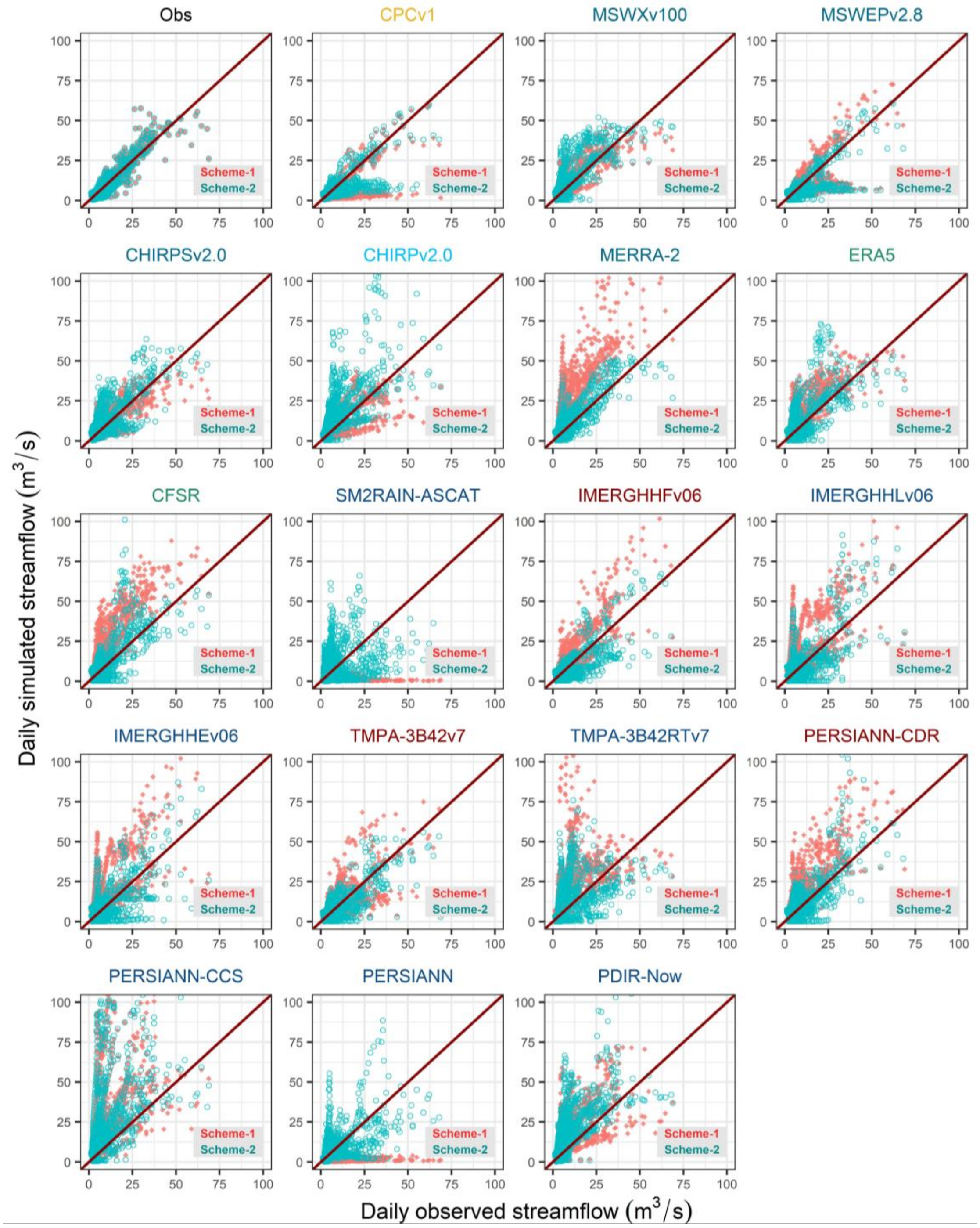

Figure 3.61. *Scatter plot of daily simulated streamflow against observed discharge for the observed precipitation and 13 GPDs at Çukurkışla basin outlet. Title color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Furthermore, CHIRPSv2.0 shows higher streamflow reproducibility compared to CHIRPv2.0, while MERRA-2 clearly draws different paths in both schemes and displays higher simulated streamflow compared to observed discharge, where the overestimation is more intense for Scheme-1. TMPA-3B42v7 outperforms IMERGHHFv06 in Scheme-1 and PERSIANN-CDR shows higher ability of streamflow simulation by its optimum parameter set compared to Scheme-1. In the same way, among reanalysis GPDs, ERA5 shows higher streamflow reproducibility compared to CFSR, while SM2RAIN-ASCAT and PERSIANN highly underestimate observed streamflow in Scheme-1 compared to other satellite-based GPDs.

Figure 3.62 presents the simulated streamflow bias distribution based on observed precipitation and GPDs, which is expressed in the form of a box-and-whiskers plot for the two years of calibration and three years of validation periods over Çukurkışla basin.

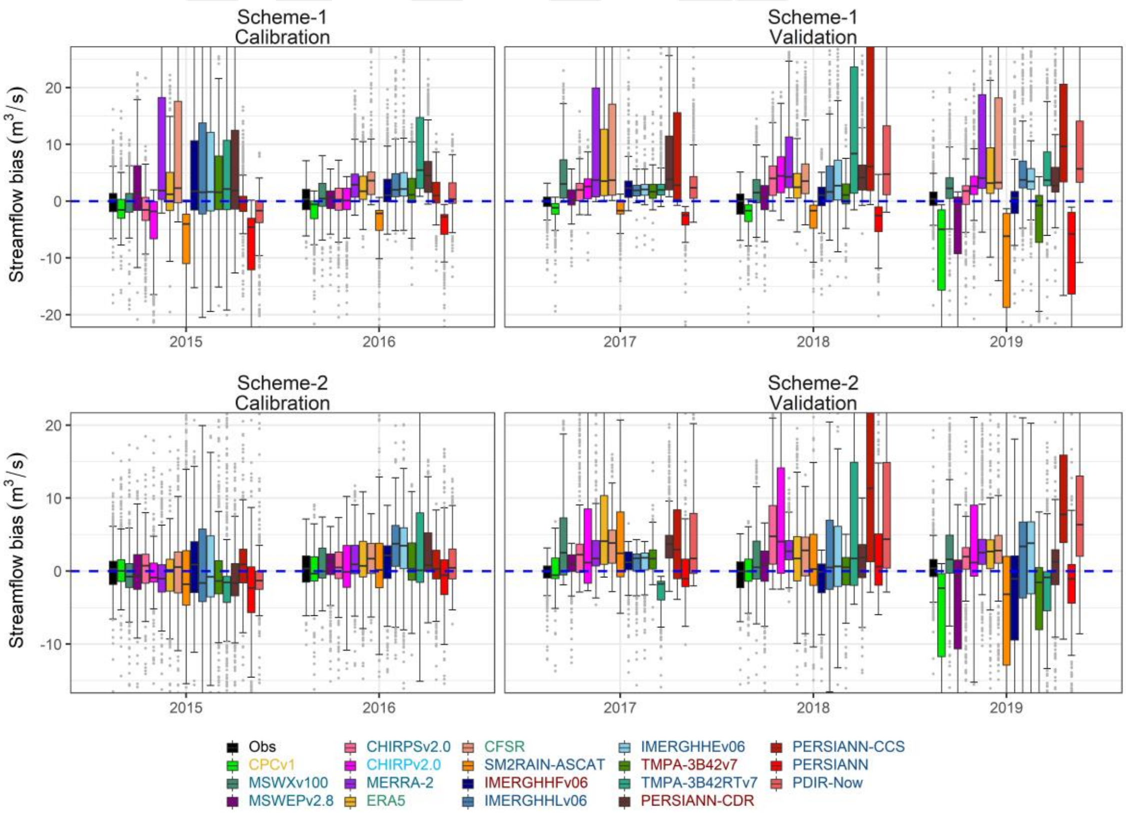

Figure 3.62. *Daily simulated streamflow bias distribution is expressed in the form of box and whiskers, considering the calibration and validation periods over Çukurkışla basin. The boxes represent the 25th, 50th, and 75th percentile values, respectively. The legend color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

In general, the range of streamflow bias changes a narrow range for both schemes in this basin compared to other selected basins, which indicates that the GPDs show lower bias for lower observed discharge. From the results, GPDs show low bias variation during the 2016 years and their bias distributions are quite dispersed for the rest of the study period in Scheme-1. Furthermore, GPDs show a higher wet and dry bias in 2019, where SM2RAIN-ASCAT and PERSIANN, like the previous, greatly underestimate streamflow for the entire period in Scheme-1, followed by CPCv1.

The simulated streamflow timeseries bias of observed precipitation and GPDs for the entire study period under two different scenarios is shown in Figure 3.63. The results indicate that CPCv1 highly underestimates observed discharge for the 2019 flow peak for both schemes.

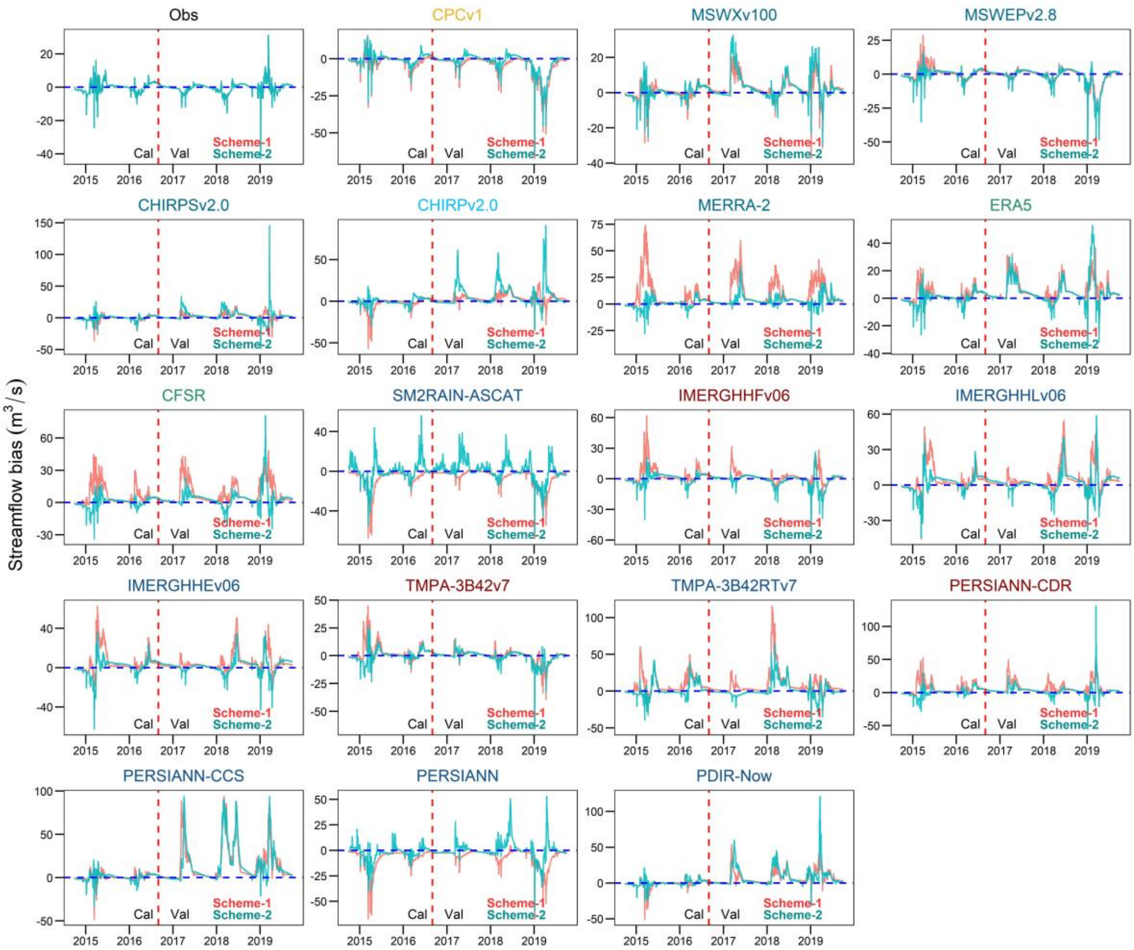

Figure 3.63. *Daily simulated streamflow timeseries bias for two schemes over Çukurkışla basin. The title color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Among gauge-corrected GPDs, MERRA-2 highly overestimates streamflow, and CHIRPSv2.0 shows unusual simulated streamflow for the last year (2019) of observation in Scheme-1. Furthermore, ERA5 shows lower bias variation over the selected study period compared to CFSR. Among satellite-based GPDs, PERSIANN and SM2RAIN-ASCAT show larger dry bias, comparatively.

Figure 3.64 displays the streamflow reproducibility of observed precipitation and selected GPDs, which is evaluated by NSE and KGE with its three components over Çukurkışla basin. Overall, GPDs shows relatively higher correlation for both schemes, whereas the most of the datasets overestimate the ratio of bias with different magnitudes in both schemes, while some GPDs underestimate variability ratio during the validation and entire period in Scheme-1. In general, observed precipitation is able to simulate streamflow closer to observed discharge compared to selected GPDs for both validation and calibration periods. CPCv1 shows lower performance over Çukurkışla basin compared to its performance over other selected catchments in this study. In the same way, MSWXv100 outperforms MSWEPv2.8 considering both schemes.

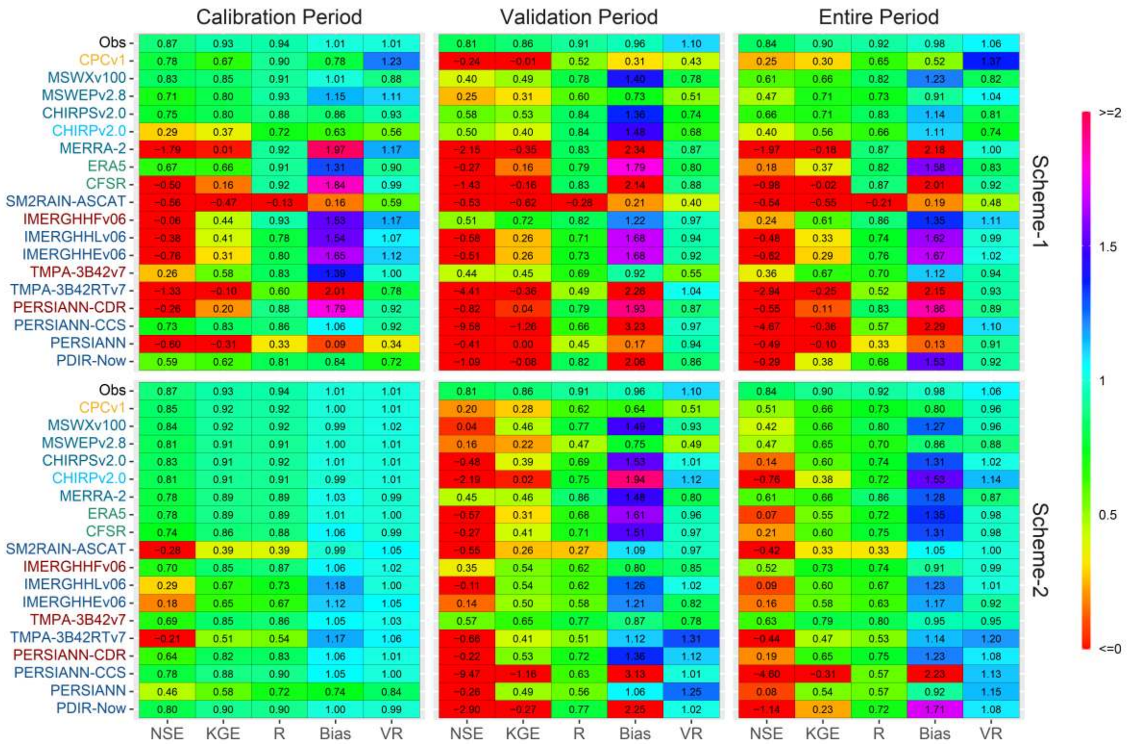

Figure 3.64. *Performance of daily streamflow simulations at the Çukurkışla basin outlet using the gauge and 18 GPDs data considering calibration, validation, and entire period. Y-axis color presents: satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

Both CHIRPSv2.0 and its gauge uncorrected version (CHIRPv2.0) show higher streamflow reproducibility for the validation period in Scheme-1 than in Scheme-2, where MERRA-2 is not able to simulate streamflow close to observed discharge in Scheme-1, while TMPA-3B42v7 shows higher accuracy than IMERGHHFv06 during the calibration but in the validation period in Scheme-1, both perform close to each other for the calibration period, and TMPA-3B42v7 outperforms IMERGHHFv06 during the validation period in Scheme-2. PERSIANN-CDR shows lower performance compared to other gauge-corrected GPDs. Among reanalysis GPDs, ERA5 outperforms CFSR while only PERSIANN-CCS and PDIR-Now show slightly higher streamflow reproducibility compared to other satellite-based GPDs in the calibration period in Scheme-1. Finally, many GPDs show higher streamflow reproducibility when they have their own optimum model parameter set.

Figure 3.65 displays the optimum TUW model parameter set of the precipitation and selected GPDs for streamflow simulation over Çukurkışla basin.

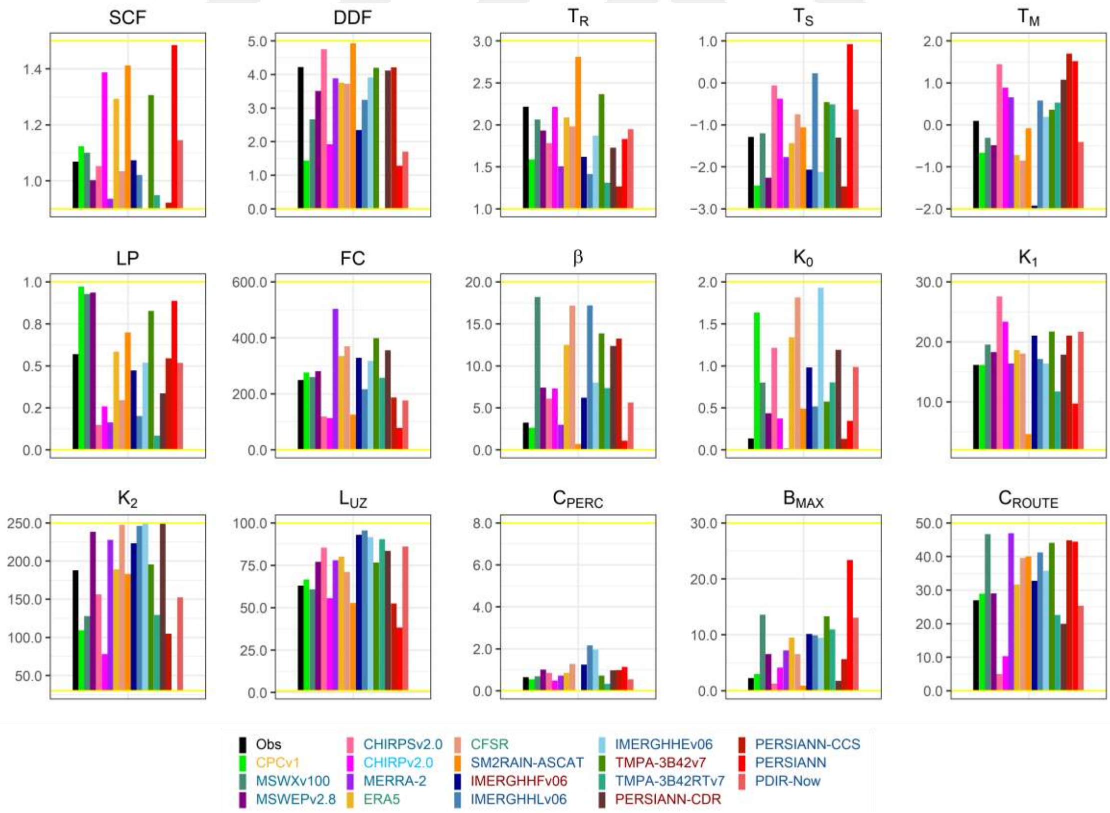

Figure 3.65. Optimum TUW model parameters for the observed and 18 GPDs *over Çukurkışla basin*. Legend text color presents: observed (black), satellite-based (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).

Overall, GPDs show higher Degree Day Factor (DDF) over Çukurkışla basin when compared to GPDs DFF over Karasu and Kayabaşı catchments. In the same way many GPDs have lower CPERC values in this basin. Finally, some multisource GPDs show higher LP compared to reanalysis and satellite-based GPDs.

3.2.4. Overall ability of GPDs for streamflow prediction

Figure 3.66 ranks six of the best and six of the poorest performing GPDs for streamflow reproducibility over three selected basins for the calibration (2015–2016), validation (2017–2019), and entire study periods (2015–2019). Overall, there is no such GPD which outperforms observed precipitation in streamflow reproducibility when the observed precipitation-based optimum model parameter set is taken into account (Scheme-1) for all selected three catchments. Furthermore, when each GPD has its own optimum model parameter set (Scheme-2), some GPDs show higher streamflow reproducibility than observed precipitation, which is more observable over the Karasu and Kayabaşı catchments, while observed precipitation still shows the highest streamflow reproducibility over Çukurkışla basin in Scheme-2. Among GPDs, CHIRP(S)v2.0 mostly ranked as the sixth best performing GPD for streamflow simulation over selected basins in Scheme-1, followed by MSWEPv2.8 and MSWXv100 while MERRA-2 showed

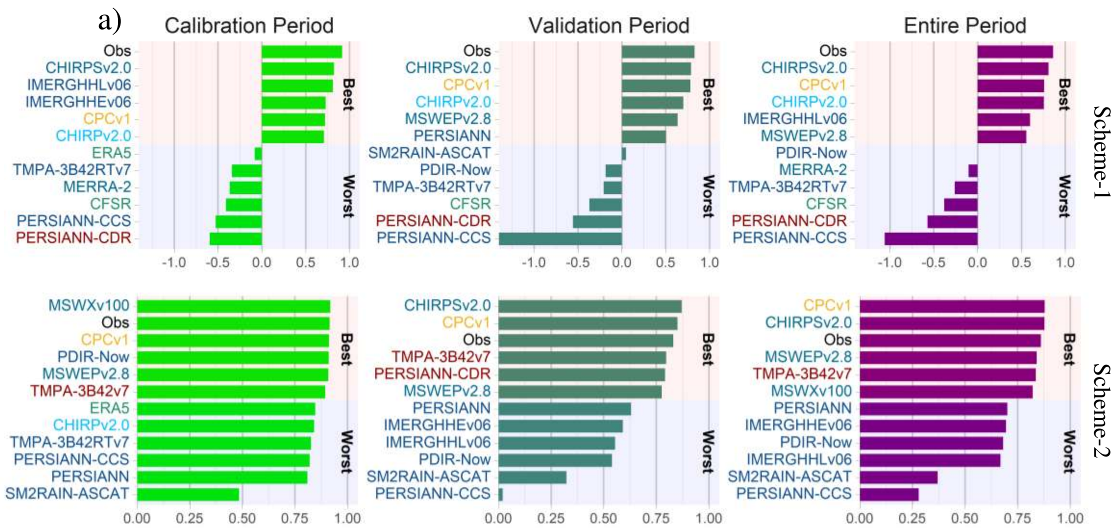

Figure 3.66. *GPDs best-six and worst-six ranking based on KGE for daily streamflow simulation considering two schemes over calibration, validation and entire periods. a) Karasu basin, b) Kayabaşı basin, and c) Çukurkışla basin. The y-axis color presents the source (s) of GPDs: satellite (blue), gauge and satellite (red), reanalysis and satellite (sky blue), reanalysis (green), reanalysis, ground, and satellite (steel blue), and ground (yellow).*

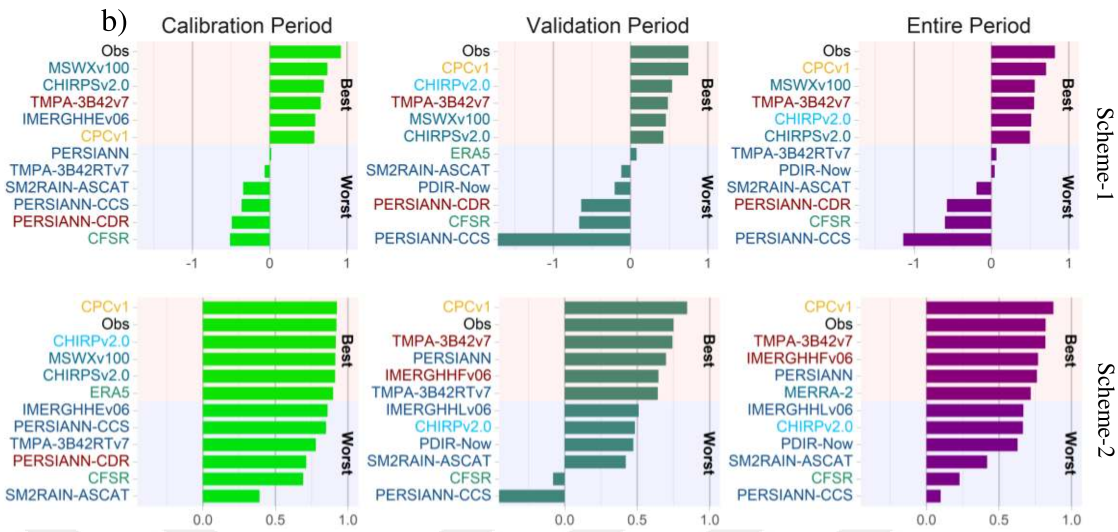

Figure 3.66. *Continued…*

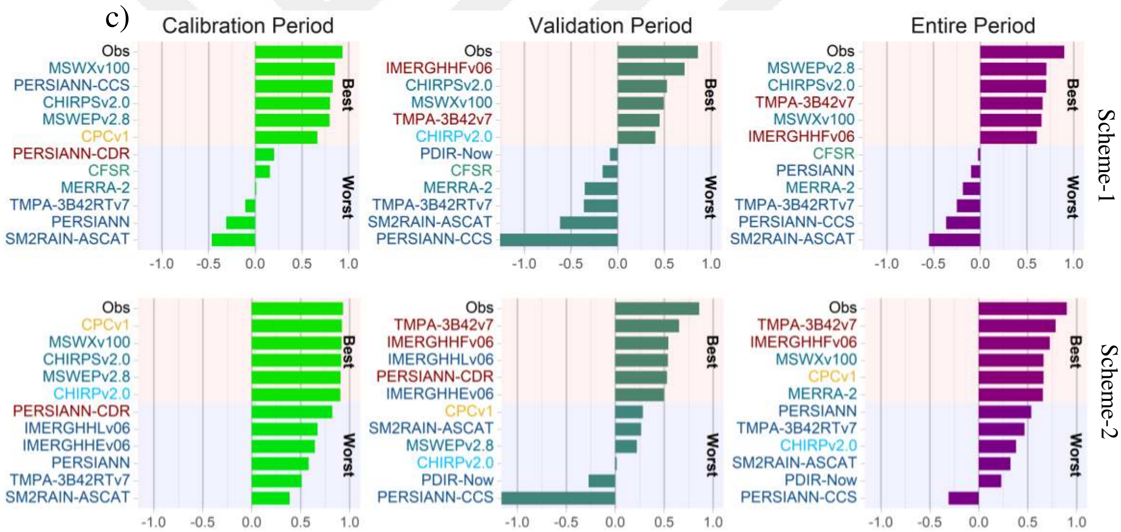

Figure 3.66. *Continued…*

weaker performance compared to GPDs which take advantage of satellite, reanalysis, and ground information. In the same way, among satellite-based GPDs, IMERGHHL(E)v06 shows the best reproducibility over the Karasu basin, even better than reanalysis and some gauge corrected GPDs. Many satellite-based GPDs are ranked as the sixth worst performed GPDs for streamflow reproducibility over all selected basins for both scenarios. Among reanalysis GPDs, CFSR appears more than ERA5 in the sixth poor performed GPDs. Hence, it is tricky to specify a certain GPD performance rank where GPD performances are changeable over each basin and selected temporal (calibration, validation, and entire periods).

4. DISCUSSIONS

In this study, the spatio-temporal consistency of 18 GPDs (i.e., one ground-based, seven satellite-based, two reanalysis, one reanalysis and satellite, and seven multi-source) is evaluated by direct comparison with ground data (Stations No. 757), and then the hydrological utility of selected GPDs is evaluated over three different snow-dominated basins (i.e., Karasu, Çukurkışla, and Kayabaşı basins). For the direct comparison, we imposed several regional classifications, such as the elevation, wetness, and climate zones, to consider the topography and climate variability effects on the performance of precipitation datasets at both daily and monthly time steps within seasonal variability for the period (2015–2019), where for hydrological utility of GPDs, two different scenarios are employed: calibrating the model parameters by observed precipitation and then replacing the observed precipitation with each GPD (Scheme-1); and calibrating the model parameters by each GPD individually and then comparing the results of both scenarios with observed streamflow data.

Based on daily and monthly precipitation patterns obtained from observed precipitation and selected GPDs over the entire region for the entire period (see Figure 3.1), all GPDs capture the typical precipitation gradient from the coast to the center of the country, with a shoreline-shaped precipitation gradient in the north extending from east to west near the Black Sea and the two precipitation hotspots located in the south-western regions near the Aegean Sea and in the protuberance part of the southern regions. These findings support the results of Darand and Khandu (2020), who compared 23 GPDs with rain gauge observations over Iran. They indicated that many GPDs are able to draw precipitation patterns, with high values in the west, close to the Persian Gulf, in the north, tangent to the Caspian Sea, and low values over the central parts of the country. Several other studies (Aksu & Akgül, 2020; Guo et al., 2016; Navidi Nassaj et al., 2022; Satgé et al., 2020) found similar results. Our results also demonstrate that many GPDs highly underestimate mean daily and monthly precipitation in the coastal regions or areas having less than 500 m elevation, which typically present coastal areas such as Black Sea region, and the results support previous GPDs studies (Amjad et al., 2020; Derin & Yilmaz, 2014). Furthermore, our results (see Figures 3.2–3.4) indicate that most GPDs demonstrate higher precipitation for the entire study period, and many GPDs highly overestimate mean precipitation during the spring and winter seasons, while GPDs show both underestimation and overestimation of precipitation during the autumn and summer seasons, which is highly in line with previous studies regarding GPD validations (K. Ahmed et al., 2019; Ayoub et al., 2020; Liu, Shangguan, et al., 2019). However, among GPDs, the PERSIANN dataset drastically underestimates precipitation at the regional scale under different temporal constraints, whereas SM2RAIN-ASCAT underestimates mean precipitation during the cold period, which is also confirmed in previous studies regarding GPDs validation (Liu, Xia, et al., 2019; Rahman et al., 2019; Saemian et al., 2021; Uysal, 2022). In the same way, considering the precipitation dynamics at both daily and monthly time steps (see Figures 3.5 and 3.6), precipitation events from GPDs show higher linearity with observed precipitation for the monthly time step compared to the daily time step, which confirms there is a higher correlation between observed and GPDs precipitation events for the greater time steps such as monthly, which is also confirmed in previous studies related to precipitation datasets evaluations (Navidi Nassaj et al., 2022; W. Sun et al., 2022).

Considering the performance of GPDs over different time steps (see section 3.1.2), all datasets performed relatively lower in terms of correlation for the daily time step, and it is rational that estimating precipitation dynamics over a longer time period (monthly) is much easier than over a smaller time window (daily). Since the correlation component (see equation 2.1) has a significant influence on the final KGE scores, the performance ranking of GPDs in terms of KGE and its correlation component (R) is highly synchronized. These findings are in line with the results presented by Beck et al. (2019a), which revealed that the low performance of 26 GPDs across the Continental United States (CONUS) was highly dependent on correlation rather than the ratio of bias and variability ratio, where our finding also explained it clearly when considering the daily time step for the entire period (see Figure 3.11); the mean of GPDs median correlation is 50% less than its optimum value, while the mean of median bias is 14%, and the variability ratio is 15% away from its optimum values. As a result, it is expected that the correlation would increase among other components of KGE, which is also confirmed by some previous studies (Centella-Artola et al., 2020; Dahri et al., 2021; Xu et al., 2022). Also, we demonstrated that when driving monthly accumulated precipitation from daily observations, the bias ratio component of KGE will stay the same for both daily and monthly time steps. These findings also support the results presented by Yu et al. (2021), who evaluated the performance of IMERGHHFv06 and IEMRGHHE(L)v06 over China. They showed that the final run of IMERGHHv06 performed better than the early runs for both daily and monthly time steps, while the bias ratio stayed the same with no variation for both runs at different time steps. This study also confirmed our findings that the IMERGHHFv06 shows higher performance than the near real-time versions, including IMERGHHEv06 and IMERGHHLv06, under different conditions, such as diverse regional classifications and time scales. Our findings also infer that the result could be different if only considering the monthly time step, since the correlation has relatively higher values and the ratio of bias would be the major source of error in the performance of GPDs for the monthly time step. The results also support the findings by Saemian et al. (2021), who evaluated the spatio-temporal consistency of 34 GPDs at the daily time step and 44 GPDs at the monthly time step over Iran, where similar findings were also reported in other studies (Ahmadebrahimpour et al., 2019; Nwachukwu et al., 2020; Satgé et al., 2020). However, they demonstrated that the variability ratio is the third-important component of KGE that we also spotted in this study too. Finally, the variability ratio highly varies from one GPD to the other, especially during the winter season for the daily time step, and it is more observable during the summer season for monthly time step, which is confirmed by previous findings (Beck, Pan, et al., 2019). In addition, our findings also show that there is a lower agreement between observed precipitation and selected GPDs throughout the dry season (i.e., summer season) compared to other seasons, which is also confirmed by Darand and Khandu (2020).

The results show that many multi-source GPDs and reanalysis datasets demonstrated higher performance than the satellite-based datasets in all areas of assessment (see sections 3.1.2–3.1.5), which explains the necessity of utilizing gauge information for the calibration and correction of various datasets and is in line with earlier studies (Hisam et al., 2023; Moazami et al., 2013). Moreover, our findings demonstrate that, regardless of considering regional classifications and looking at the seasonal variability of precipitation, the performance of all GPDs is greatly improved when moving from a smaller time step (daily) to a larger time step (monthly). When looking at selected datasets from their prospective sources, the performance of ground-based and multi-source GPDs, including MSWEPv2.8, MSWXv100, CHIRPSv2.0, IMERGHHFv06, TMPA-3B42v7, MERRA-2 and PERSIANN-CDR, demonstrate relatively higher changes from the daily to monthly time step than those of satellite-based (i.e., IMERGHHE/Lv06, SM2RAIN-ASCAT, TMPA-3B42RTv7, PERSIANN, PDIR-Now, and PERSIANN-CCS), reanalysis (i.e., ERA5, CFSR), and satellite-reanalysis

(CHIRPv2.0) GPDs which is in line with previous studies (Dembélé & Zwart, 2016; Gan et al., 2021; Hussain et al., 2018), while our results suggest that GPDs that use ground data as one of their input sources probably are not always able to give accurate estimates of precipitation. The number of gauges or the spatial resolution of ground-based GPD, source of information, and the correction timescale used are crucial aspects that significantly impact the accuracy of the final product. For example, among gauge-corrected GPDs, PERSIANN-CDR that is adjusted with monthly data from GPCP, which is a merged product of ground observations and satellite-based precipitation estimates and has a 2.5^{o} spatial resolution, shows relatively poor performance. Our findings are consistent with previous studies (Beck, Pan, et al., 2019; Liu, Shangguan, et al., 2019; Xiang et al., 2021).

Our results reveal that MSWEPv2.8 demonstrates the highest performance among other selected GPDs for the daily time step considering the entire period (see Figures 3.11–3.19), followed by some other multi-source datasets, reanalysis, and ground-based data, where many satellite-based and only satellite-reanalysis datasets show relatively lower performance while PERSIANN-CCS show the worst performance. For the monthly time step considering the entire period, MSWEPv2.8 still delivers the best performance, and CPCv1 comes in with the second-best performance, followed again by some gauge-corrected GPDs and reanalysis where satellite-based datasets show relatively lower performance. Our results support previous findings by Bai et al. (2020), who evaluated 12 GPDs over China and found that MSWEPv1 outperformed other GPDs and PERSIANN-CCS demonstrated poor performance, and some other studies also reported similar results over different regions of the world (S. hu Jiang et al., 2023; Lu et al., 2021; Nashwan et al., 2020; Sahlu et al., 2017; Satgé et al., 2020; Xu et al., 2022; Yang et al., 2016; Zeng et al., 2018). However, our findings demonstrate some distorted results. For example, some satellite-based GPDs, such as the IMERGHHEv06 and IMERGHHLv06, demonstrated higher performance than some gauge-corrected GPDs, including MERRA-2, CHIRPSv2.0, and PERSIANN-CDR, and also outperformed CFSR, which is one of the reanalysis dataset, for the daily time step. Also, the performance ranking may change if considering the seasonal variability, where CPCv1 outperforms all GPDs during the spring season for the monthly time step. In the same way, the results show that both IMERGHHFv06 and its near real-time versions IMERGHHL(E)v06, show relatively higher performance compared to their predecessors, such as TMPA-3B42(RT)v7 which is also confirmed in previous studies (Amjad et al., 2020; Ma et al., 2021; Morsy et al., 2021; Z. Shen et al., 2022), while ERA5 shows relatively higher performance than CFSR, which is in line with some other studies (Dubey et al., 2021; Ougahi & Mahmood, 2022). Among satellite-based datasets, PERSIANN datasets demonstrate lower performance for both daily and monthly time steps (Anjum et al., 2022; Hosseini-Moghari et al., 2018; Ray et al., 2022). Moreover, the performance of GPDs over different elevation ranges (see Figures 3.14–3.15) demonstrate that almost all datasets demonstrated higher performance over areas having elevation less than 500 m and show relatively lower performance over highly elevated regions, which are typically described by complex topography. These findings support the results obtained by Zambrano-Bigiarini et al. (2017), who evaluated the spatio-temporal of seven GPDs over different elevation ranges considering daily, monthly, annual, and seasonal temporal. In the same way, many GPDs show relatively lower performance over regions experiencing extremes (see Figures 3.16–3.17) such as dry and wet regions, which is in line with previous studies (Amjad et al., 2020), where the same lower performance of GPDs was obtained over the Black Sea, Central Anatolia and Southeastern Anatolia (see Figures 3.18–3.19), which is confirmed by Aksu and Akgül (2020).

The results indicate that the general trend of precipitation detectability for GPDs is higher for low precipitation intensities, and the detectability is decreased for highly intense precipitation (see section 3.1.6). This is because precipitation is being divided into many thresholds instead of being classified as simple rain and no rain events, which is consistent with previous studies (Amjad et al., 2020; Oreggioni Weiberlen & Báez Benítez, 2018; Satgé et al., 2020). Regardless of any regional classification, there is a bit of an exception in our results, in which most of the time GPDs demonstrate higher detectability for moderate precipitation than for light precipitation, although light precipitation exhibits a higher frequency of precipitation events than the moderate threshold. This can be illustrated by the fact that, usually, detecting an event with a low probability of occurrence is much harder than identifying an event with a high probability of occurrence. On the other hand, using five different intensity thresholds instead of a binary precipitation or no precipitation event can be considered a more demanding classification scheme. This, along with larger frequency differences between light and moderate intensities, could be attributed as the main reasons why most GPDs indicate a higher detection ability for moderate precipitation compared to light precipitation, and this division gets even more problematic when the occurrence probability of a certain event is seen much more rarely compared to other classes (heavy and extreme intensities). Such results are also observed in other GPD validations (Baez-Villanueva et al., 2018; Zambrano-Bigiarini et al., 2017).

From the hydrological utility of GPDs (see section 3.2), with the calibrated model parameters set by observed precipitation (Scheme-1), the behavior of streamflow simulation from many GPDs is worst and produces deteriorated biases, lower KGE, and lower NSE for both the calibration and validation periods, which also effect the entire period. However, with the model parameters recalibrated by each GPD individually (Scheme-2), the streamflow reproducibility of GPDs significantly increased. These results highly support the finding of Su et al. (2021), who evaluated the hydrological utility of three GPDs in China over Mishui basin and stated that the datasets demonstrated higher ability of streamflow simulation by their optimal model parameter sets, and only TMPA-3B42v6 performed fairly well for streamflow simulation with observed precipitation calibrated model parameters. Our results also demonstrate that not only TMPA-3B42v7 but also other GPDs, including CPCv1, CHIRP(S)v2.0, MSWEPv2.8, MSWXv100, ERA5, and IMERGHHFv07, even some satellite-based GPDs, such as IMERGHHEv06 and IMERGHHLv06, show relatively good streamflow reproducibility compared to other GPDs in Scheme-1 over three different basins. Similar results can also be found in other studies (Xue et al., 2013; Y. Zhang et al., 2022). Considering the calibration and validation periods for both scenarios, GPDs demonstrate low streamflow simulation ability for the validation periods compared to their calibration periods. This can be explained by the fact that the short calibration period may not be representative of long-term basin conditions, or the lower performance of GPDs during the validation period may be due to the difference between hydrological conditions for the calibration and validation periods. These results support the earlier finding by Jiang and Bauer-Gottwein (2019), who evaluated the hydrological utility of IMERGHHF(E)v06 and TMPA-3B42v7 over 300 catchments in China, and demonstrate that both IMERGHHv06 products show relatively higher performance than TMPA-3B42v7, where our results also confirmed the outperformance of TMPA-3B42(RT)v7 by IMERGHHv06 datasets. It is important to note that the calibrated model parameter sets for each GPD have a tendency to account for other aspects, such as model structure, forcing data, etc., and as a result, they may vary throughout different calibration periods, particularly for short periods of time, which is also confirmed by Merz et al. (2011). Moreover, the hydrological utility of GPDs shows that there is a higher agreement between observed streamflow and GPDs simulated streamflow in terms of correlation than to their direct comparison at the daily time step. This can be attributed to the fact that precipitation is the most uncertain part of the hydrologic cycle compared to streamflow, while its distribution varies over time and space faster than the variation of streamflow measured at a single point (the outlet). Moreover, for streamflow simulations, precipitation is probably the most critical meteorological forcing for hydrological models. However, other model forcing, such as temperature and potential evapotranspiration, may also have an influence on precipitation-runoff modeling, considering a significant snowmelt process in a mountainous basin. In the same way, our results show that the bias propagated in the GPDs is the main source of error that highly affects the final streamflow reproducibility skill of datasets, which is expressed in the form of KGE. For example, PDIR-Now demonstrates higher streamflow reproducibility during the calibration period for both schemes over Çukurkışla basin, whereas for the validation period performs poorly. This may be due to the change in bias in the dataset over the years, where in the direct comparison we spotted that the PDIR-Now bias increased for those last two years of the study period. These results support previous findings by Uysal and Şorman (2021), who evaluated PERSIANN and PERSIANN-CDR datasets by direct comparison with observed precipitation and utilized them for streamflow simulation over Karasu basin and demonstrated that applying a post- and pre-bias corrections to the selected GPDs had a highly significant effect on their performance for streamflow simulation.

The results from hydrological utility of GPDs show that the ground-based and multi-source GPDs show a higher skill of streamflow reproducibility over three snow-dominated basins, which is followed by reanalysis, and that, except for IMERGHHL(E)v06, other satellite-based datasets show a relatively lower skill of streamflow reproducibility. However, it is important to keep in mind that there are some assumptions and limitations. These findings only generalized the performance of GPDs for streamflow simulation over snow-dominated basins employing a conceptual model (TUW). Hence, the findings would be changed if a different rainfall-runoff model was utilized or the hydrological assessment of GPDs was done over basins with different climate conditions. The short period is another limitation of this study. However, we believe a timely assessment is crucial for further GPD improvement and future research.

5. SUMMARY AND CONCLUSIONS

In the study, the spatio-temporal consistency of 18 recent GPDs over variable topography and climate conditions in Turkey was analyzed, as well as their hydrological utility for streamflow simulation. The performance of GPDs over variable topography was obtained by discretizing the region into four elevation ranges where the effect of climate was considered by grouping the study area into seven climate zones and differentiating selected stations into four wetness classes. Furthermore, GPDs were validated with 757 meteorological stations, which were available for the five hydrologic years (2015–2019). The evaluations are done by considering individual metrics, including Kling-Gupta Efficiency (KGE) with its three components, Hanssen-Kuiper Skill (HKS) score, and Nash-Sutcliffe Efficiency (NSE). Finally, the TUW model is employed for streamflow simulation based on observed precipitation and selected GPDs under different model parametrization scenarios over three selected basins (Karasu, Kayabaşı, and Çukurkışla). The major conclusions are summarized as follows:

5.1. Conclusions from GPDs directed comparison

- CPCv1, whose temporal variation is entirely dependent on the spatial information of gauge observations, demonstrates close mean daily and monthly precipitation to the observed data, followed by some multi-source GPDs, including MSWXv100, MSWEPv2.8, and CHIRPSv2.0.
- Among gauge uncorrected GPDs, CHIRPv2.0 shows the same amount of mean daily and monthly precipitation as its corrected version (CHIRPSv2.0) and is close to observed precipitation, while ERA5 comes in with the second-best spatial distribution of mean precipitation, but ERA5 indicates significantly greater precipitation.
- Some GPDs, such as PERSIANN-CCS, PERSIANN-CDR, PDIR-Now, and CFSR, highly overestimate mean precipitation, whereas PERSIANN datasets significantly have the least mean precipitation amount, followed by SM2RAIN-ASCAT during the winter season.
- GPDs show higher performance for the monthly time step than daily precipitation, which is distinguished by a higher KGE and its correlation components.

- For the daily precipitation, it is MSWEPv2.8 that demonstrates the highest performance at the regional scale, followed by some other multi-source GPDs, whereas for the monthly time step, MSWEPv2.8 still outperforms other GPDs, followed by CPCv1 and other multi-source GPDs, while reanalysis datasets outperform satellite-based, satellite-reanalysis, and even some other datasets such as PERSIANN-CDR, CHIRPSv2.0, and MERRA-2, especially for the daily time step.
- GPDs relatively show higher performance for the autumn season and lower performance during the summer season, while their performance quite diversifies for the spring and winter seasons over different regions.
- In general, GPDs show weak performance as the elevation and amount of precipitation are increased or decreased, such as in the Black Sea region, Central, Southeastern, and Eastern Anatolia. and show lower performance over regions that experience extremes, such as wet and dry regions.
- GPDs show higher detectability for lower daily precipitation intensities, but many GPDs perform poorly at differentiating between light and moderate, and heavy and extreme precipitation events, where GPDs show distinct detectability of different precipitation events under various regional classifications and seasonal variability of precipitation.

5.2. Conclusions from GPDs hydrological utility

- GPDs show higher streamflow reproducibility when the model is run with their own optimal parameter sets. In other words, GPDs show higher performance in Scheme-2 than in Scheme-1..
- Still, observed precipitation-based streamflow simulations are highly similar to the observed discharge in Scheme-1, whereas in Scheme-2, some GPDs, including CPCv1, MSWXv100, and CHIRPSv2.0, outperform the observed precipitation.
- The streamflow reproducibility of GPDs vary in each basin. Both IMERGHHFv06 and TMPA-3B42v7 show lower streamflow reproducibility in the eastern part of Turkey (Karasu basin), whereas these datasets exhibit higher streamflow reproducibility in the southern basin (Çukurkışla basin), even outperforms CPCv1 and other GPDs.

- Some satellite-based GPDs are becoming available with high spatial resolution and short time lag (latency), which is very important for real-time operation, but the existing bias limits their reliability for hydro-meteorological studies. As an example, PERSIANN-CCS is available after a one-hour time lag with 0.04° spatial resolution, although its performance is not very high. On the other hand, IMERGHHLv06 presents precipitation after 14 hours with a coarser spatial resolution (0.1°) compared to PERSIANN-CCS and is more reliable among selected satellite-based GPDs.
- Finally, when satellite-based GPDs are merged with other sources, such as reanalysis and/or ground observation data, they become more accurate. For example, MSWEPv2.8 and CHIRPSv2.0 are the most reliable GPDs over the Karasu basin, but they have longer time lags varying from one month to a few months. It can be concluded that there are GPDs available for a near-real-time study, and as product merging from different sources is implemented, increasing latency, the reliability of the new product seems to increase.

5.3. Recommendations

- The gauge uncorrected GPDs can be merged either by available ground data or other gauge corrected GPDs to reduce the existence bias in these GPDs.
- The spatial merging of GPDs either with each other or with the possible dense gauge network may highly increase their performances over time and space.
- The mandatory regional classification in terms of elevation and wetness with different thresholds may provide different result for GPDs performance.
- Scheme-2 of hydrological modeling is recommended for ungauged basins where the calibration procedure uses GPD, ground temperature, and PET values. Since all model forcing will have their influence on optimal model parameter set, it would be more reliable if model calibration may be implemented from the same climate data.

5.4. Future Studies

The future studies will focus on:

- Spatial merging and bias correction of selected GPDs using machine learning.
- Performing GPDs evaluation with a distributed observed precipitation (grid-based) instead of point to grid method.
- Using available climate data for streamflow prediction instead of utilizing only precipitation as a meteorological forcing.
- Some GPDs, such MSWXv100 provide forecast data which will be used for flood forecasting, drought monitoring and early warning systems.

REFERENCES

Ahmadebrahimpour, E., Aminnejad, B., & Khalili, K. (2019). Assessment of the reliability of three gauged-based global gridded precipitation datasets for drought monitoring. *International Journal of Global Warming*, *18*(2), 103–119. https://doi.org/10.1504/IJGW.2019.100312

Ahmed, E., Al Janabi, F., Zhang, J., Yang, W., Saddique, N., & Krebs, P. (2020). Hydrologic Assessment of TRMM and GPM-Based Precipitation Products in Transboundary River Catchment (Chenab River, Pakistan). Water 2020, Vol. 12, Page 1902, 12(7), 1902. https://doi.org/10.3390/W12071902

Ahmed, K., Shahid, S., Wang, X., Nawaz, N., & Najeebullah, K. (2019). Evaluation of gridded precipitation datasets over arid regions of Pakistan. Water (Switzerland), 11(2), 210. https://doi.org/10.3390/w11020210

Akbulut, N. (Emir), Bayarı, S., Akbulut, A., Özyurt, N. N., & Sahin, Y. (2022). Rivers of Turkey. Rivers of Europe, 851–880. https://doi.org/10.1016/B978-0-08-102612-0.00017-1

Aksu, H., & Akgül, M. A. (2020). Performance evaluation of CHIRPS satellite precipitation estimates over Turkey. *Theoretical and Applied Climatology*, *142*(1–2), 71–84. https://doi.org/10.1007/S00704-020-03301-5/TABLES/4

Akyurek, Z., Hall, D. K., Riggs, G. A., & Sensoy, A. (2010). Evaluating the utility of the ANSA blended snow cover product in the mountains of eastern Turkey. *International Journal of Remote Sensing*, *31*(14), 3727–3744. https://doi.org/10.1080/01431161.2010.483484

Amjad, M., Yilmaz, M. T., Yucel, I., & Yilmaz, K. K. (2020). Performance evaluation of satellite- and model-based precipitation products over varying climate and complex topography. *Journal of Hydrology*, *584*, 124707. https://doi.org/10.1016/J.JHYDROL.2020.124707

Andermann, C., Bonnet, S., & Gloaguen, R. (2011). Evaluation of precipitation data sets along the Himalayan front. *Geochemistry, Geophysics, Geosystems*, *12*(7), 1631. https://doi.org/10.1029/2011GC003513

Anjum, M. N., Irfan, M., Waseem, M., Leta, M. K., Niazi, U. M., Rahman, S. U., Ghanim, A., Mukhtar, M. A., & Nadeem, M. U. (2022). Assessment of PERSIANN-CCS, PERSIANN-CDR, SM2RAIN-ASCAT, and CHIRPS-2.0 Rainfall Products over a Semi-Arid Subtropical Climatic Region. *Water (Switzerland)*, *14*(2), 147. https://doi.org/10.3390/w14020147

Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., & Prat, O. P. (2015). PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bulletin of the American Meteorological Society*, *96*(1), 69–83. https://doi.org/10.1175/BAMS-D-13-00068.1

Awange, J. L., Ferreira, V. G., Forootan, E., Khandu, Andam-Akorful, S. A., Agutu, N. O., & He, X. F. (2016). Uncertainties in remotely sensed precipitation data over Africa. *International Journal of Climatology*, *36*(1), 303–323. https://doi.org/10.1002/JOC.4346

Ayoub, A. B., Tangang, F., Juneng, L., Tan, M. L., & Chung, J. X. (2020). Evaluation of Gridded Precipitation Datasets in Malaysia. *Remote Sensing 2020, Vol. 12, Page 613*, *12*(4), 613. https://doi.org/10.3390/RS12040613

Baez-Villanueva, O. M., Zambrano-Bigiarini, M., Ribbe, L., Nauditt, A., Giraldo-Osorio, J. D., & Thinh, N. X. (2018). Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. *Atmospheric Research*, *213*, 34–50. https://doi.org/10.1016/J.ATMOSRES.2018.05.011

Bai, L., Wen, Y., Shi, C., Yang, Y., Zhang, F., Wu, J., Gu, J., Pan, Y., Sun, S., & Meng, J. (2020). Which precipitation product works best in the qinghai-tibet plateau, multi-source blended data, global/regional reanalysis data, or satellite retrieved precipitation data? *Remote Sensing*, *12*(4), 683. https://doi.org/10.3390/rs12040683

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., Van Dijk, A. I. J. M. J. M., Huffman, G. J., Adler, R. F., & 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. https://hess.copernicus.org/articles/23/207/2019/

Beck, H. E., Van Dijk, A. I. J. M., Larraondo, P. R., McVicar, T. R., Pan, M., Dutra, E., & Miralles, D. G. (2022). Global 3-Hourly 0.1 Bias-Corrected Meteorological Data Including Near-Real-Time Updates and Forecast Ensembles. *Bulletin of the American Meteorological Society*, *103*(3), E710–E732. https://doi.org/10.1175/BAMS-D-21-0145.1

Beck, H. E., Van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & 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. https://doi.org/10.5194/HESS-21-589-2017

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., McVicar, T. R., & Adler, R. F. (2019). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

Behrangi, A., Khakbaz, B., Jaw, T. C., AghaKouchak, A., Hsu, K., & Sorooshian, S. (2011). Hydrologic evaluation of satellite precipitation products over a mid-size basin. *Journal of Hydrology*, *397*(3–4), 225–237. https://doi.org/10.1016/J.JHYDROL.2010.11.043

Bergstrom, S. (1995). The HBV model. *Computer Models of Watershed*, 443–476. https://ci.nii.ac.jp/naid/10015209308/

Bergström, S., & Lindström, G. (2015). Interpretation of runoff processes in hydrological modelling—experience from the HBV approach. *Hydrological Processes*, *29*(16), 3535–3545. https://doi.org/10.1002/HYP.10510

Bitew, M. M., & Gebremichael, M. (2011). Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model. *Water Resources Research*, *47*(6), 6526. https://doi.org/10.1029/2010WR009917

Bıyık, G., Unal, Y., & Onol, B. (2009). *Assessment of Precipitation Forecast Accuracy over Eastern Black Sea Region using WRF-ARW. 11*, 178.

Blacutt, L. A., Herdies, D. L., de Gonçalves, L. G. G., Vila, D. A., & Andrade, M. (2015). Precipitation comparison for the CFSR, MERRA, TRMM3B42 and Combined Scheme datasets in Bolivia. *Atmospheric Research*, *163*, 117–131. https://doi.org/10.1016/J.ATMOSRES.2015.02.002

Bosilovich, M. G., J. Chen, F. R. Robertson, & R. F. Adler. (2015). MERRA-2: Initial evaluation of the climate. *National Aeronautics and Space Administration, Goddard Space Flight Center.*

Boushaki, F. I., Hsu, K. L., Sorooshian, S., Park, G. H., Mahani, S., & Shi, W. (2009). Bias Adjustment of Satellite Precipitation Estimation Using Ground-Based Measurement: A Case Study Evaluation over the Southwestern United States. *Journal of Hydrometeorology*, *10*(5), 1231–1242. https://doi.org/10.1175/2009JHM1099.1

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., & Wagner, W. (2019). SM2RAIN-ASCAT (2007-2018): Global daily satellite rainfall data from ASCAT soil moisture observations. *Earth System Science Data*, *11*(4), 1583–1601. https://doi.org/10.5194/ESSD-11-1583-2019

Centella-Artola, A., Bezanilla-Morlot, A., Taylor, M. A., Herrera, D. A., Martinez-Castro, D., Gouirand, I., Sierra-Lorenzo, M., Vichot-Llano, A., Stephenson, T., Fonseca, C., Campbell, J., & Alpizar, M. (2020). Evaluation of Sixteen Gridded Precipitation Datasets over the Caribbean Region Using Gauge Observations. *Atmosphere 2020, Vol. 11, Page 1334*, *11*(12), 1334. https://doi.org/10.3390/ATMOS11121334

Ceola, S., Arheimer, B., Baratti, E., Blöschl, G., Capell, R., Castellarin, A., Freer, J., Han, D., Hrachowitz, M., Hundecha, Y., Hutton, C., Lindström, G., Montanari, A., Nijzink, R., Parajka, J., Toth, E., Viglione, A., & Wagener, T. (2015). Virtual laboratories: New opportunities for collaborative water science. *Hydrology and Earth System Sciences*, *19*(4), 2101–2117. https://doi.org/10.5194/HESS-19-2101-2015

Chiaravalloti, F., Brocca, L., Procopio, A., Massari, C., & Gabriele, S. (2018). Assessment of GPM and SM2RAIN-ASCAT rainfall products over complex terrain in southern Italy. *Atmospheric Research*, *206*, 64–74. https://doi.org/10.1016/J.ATMOSRES.2018.02.019

Chua, Z. W., Kuleshov, Y., & Watkins, A. (2020). Evaluation of Satellite Precipitation Estimates over Australia. *Remote Sensing 2020, Vol. 12, Page 678*, *12*(4), 678.

https://doi.org/10.3390/RS12040678

Dahri, Z. H., Ludwig, F., Moors, E., Ahmad, S., Ahmad, B., Shoaib, M., Ali, I., Iqbal, M. S., Pomee, M. S., Mangrio, A. G., Ahmad, M. M., & Kabat, P. (2021). Spatio-temporal evaluation of gridded precipitation products for the high-altitude Indus basin. *International Journal of Climatology*, *41*(8), 4283–4306. https://doi.org/10.1002/joc.7073

Darand, M., & Khandu, K. (2020). Statistical evaluation of gridded precipitation datasets using rain gauge observations over Iran. *Journal of Arid Environments*, *178*, 104172. https://doi.org/10.1016/J.JARIDENV.2020.104172

de Leeuw, J., Methven, J., & Blackburn, M. (2015). Evaluation of ERA-Interim reanalysis precipitation products using England and Wales observations. *Quarterly Journal of the Royal Meteorological Society*, *141*(688), 798–806. https://doi.org/10.1002/QJ.2395

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., … Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/QJ.828

Dembélé, M., & Zwart, S. J. (2016). Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. *International Journal of Remote Sensing*, *37*(17), 3995–4014. https://doi.org/10.1080/01431161.2016.1207258

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C. H., Chen, H., Delrieu, G., Hsu, Y. C., Lavado-Casimiro, W., Manz, B., Moges, S., Nikolopoulos, E. I., Sahlu, D., Salerno, F., Rodríguez-Sánchez, J. P., Vergara, H. J., & Yilmaz, K. K. (2019). Evaluation of GPM-era Global Satellite Precipitation Products over Multiple Complex Terrain Regions. *Remote Sensing 2019, Vol. 11, Page 2936*, *11*(24), 2936. https://doi.org/10.3390/RS11242936

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C. H., Delrieu, G., Hong, Y., Hsu, Y. C., Lavado-Casimiro, W., Manz, B., Moges, S., Nikolopoulos, E. I., Sahlu, D., Salerno, F., Rodríguez-Sánchez, J. P., Vergara, H. J., & Yilmaz, K. K. (2016). Multiregional Satellite Precipitation Products Evaluation over Complex Terrain. *Journal of Hydrometeorology*, *17*(6), 1817–1836. https://doi.org/10.1175/JHM-D-15-0197.1

Derin, Y., & Yilmaz, K. K. (2014). Evaluation of Multiple Satellite-Based Precipitation Products over Complex Topography. *Journal of Hydrometeorology*, *15*(4), 1498–1516. https://doi.org/10.1175/JHM-D-13-0191.1

Dinku, T., Ruiz, F., Connor, S. J., & Ceccato, P. (2010). Validation and Intercomparison of Satellite Rainfall Estimates over Colombia. *Journal of Applied Meteorology and Climatology*, *49*(5), 1004–1014. https://doi.org/10.1175/2009JAMC2260.1

Dubey, S., Gupta, H., Goyal, M. K., & Joshi, N. (2021). Evaluation of precipitation datasets available on Google earth engine over India. *International Journal of Climatology*, *41*(10), 4844–4863. https://doi.org/10.1002/joc.7102

Ebert, E. E., Janowiak, J. E., & Kidd, C. (2007). Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bulletin of the American Meteorological Society*, *88*(1), 47–64. https://doi.org/10.1175/BAMS-88-1-47

El Kenawy, A. M., Lopez-Moreno, J. I., McCabe, M. F., & Vicente-Serrano, S. M. (2015). Evaluation of the TMPA-3B42 precipitation product using a high-density rain gauge network over complex terrain in northeastern Iberia. *Global and Planetary Change*, *133*, 188–200. https://doi.org/10.1016/J.GLOPLACHA.2015.08.013

Frenken, K. (2009). Irrigation in the Middle East region in figures AQUASTAT Survey-2008. *Water Reports*, *34*. https://www.cabdirect.org/cabdirect/abstract/20093247971

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data 2015 2:1*, *2*(1), 1–21. https://doi.org/10.1038/sdata.2015.66

Gan, F., Gao, Y., & Xiao, L. (2021). Comprehensive validation of the latest IMERG V06 precipitation estimates over a basin coupled with coastal locations, tropical climate and hill-karst combined landform. *Atmospheric Research*, *249*, 105293. https://doi.org/10.1016/j.atmosres.2020.105293

Gebere, S. B., Alamirew, T., Merkel, B. J., & Melesse, A. M. (2015). Performance of High Resolution Satellite Rainfall Products over Data Scarce Parts of Eastern Ethiopia. *Remote Sensing 2015, Vol. 7, Pages 11639-11663*, *7*(9), 11639–11663. https://doi.org/10.3390/RS70911639

Gebregiorgis, A. S., Tian, Y., Peters-Lidard, C. D., & Hossain, F. (2012). Tracing hydrologic model simulation error as a function of satellite rainfall estimation bias components and land use and land cover conditions. *Water Resources Research*, *48*(11), 11509. https://doi.org/10.1029/2011WR011643

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., … Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Ghaedamini, H. A., Morid, S., Nazemosadat, M. J., Shamsoddini, A., & Moghadam, H. S. (2021). Validation of the CHIRPS and CPC-Unified products for estimating extreme daily precipitation over southwestern Iran. *Theoretical and Applied Climatology*, *146*(3–4), 1207–1225. https://doi.org/10.1007/S00704-021-03790-

Y/TABLES/6

Ghulami, M., Babel, M. S., & Shrestha, M. S. (2017). Evaluation of gridded precipitation datasets for the kabul basin, afghanistan. *International Journal of Remote Sensing*, *38*(11), 3317–3332. https://doi.org/10.1080/01431161.2017.1294775

Gunathilake, M. B., Amaratunga, Y. V., Perera, A., Karunanayake, C., Gunathilake, A. S., & Rathnayake, U. (2020). Statistical evaluation and hydrologic simulation capacity of different satellite-based precipitation products (SbPPs) in the Upper Nan River Basin, Northern Thailand. *Journal of Hydrology: Regional Studies*, *32*, 100743. https://doi.org/10.1016/J.EJRH.2020.100743

Guo, H., Chen, S., Bao, A., Hu, J., Yang, B., & Stepanian, P. M. (2016). Comprehensive evaluation of high-resolution satellite-based precipitation products over China. *Atmosphere*, *7*(1), 6. https://doi.org/10.3390/atmos7010006

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/J.JHYDROL.2009.08.003

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/QJ.3803

Hisam, E., Danandeh Mehr, A., Alganci, U., & Zafer Seker, D. (2023). Comprehensive evaluation of Satellite-Based and reanalysis precipitation products over the Mediterranean region in Turkey. *Advances in Space Research*, *71*(7), 3005–3021. https://doi.org/10.1016/j.asr.2022.11.007

Hong, Y., Hsu, K. L., Sorooshian, S., & Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology*, *43*(12), 1834–1852. https://doi.org/10.1175/jam2173.1

Hosseini-Moghari, S.-M., Araghinejad, S., Ebrahimi, K., & Kanae, S. (2018). Spatio-temporal evaluation of global gridded precipitation datasets across Iran. *Hydrological Sciences Journal*, *63*(11), 1669–1688. https://doi.org/10.1080/02626667.2018.1524986

Hsu, J., Huang, W. R., & Liu, P. Y. (2021). Performance assessment of GPM-based near-real-time satellite products in depicting diurnal precipitation variation over Taiwan. *Journal of Hydrology: Regional Studies*, *38*, 100957. https://doi.org/10.1016/J.EJRH.2021.100957

Hsu, K. L., Gao, X., Sorooshian, S., & Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, *36*(9), 1176–1190. https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J., & Xie, P. (2015). NASA global precipitation measurement (GPM) Integrated multi-satellitE Retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD) version 06, Technical Report. *Technical Report*, *b*, 26.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., & Wolff, D. B. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, *8*(1), 38–55. https://doi.org/10.1175/JHM560.1

Huffman, G. J., Adler, R. F., Bolvin, D. T., & Nelkin, E. J. (2010). The TRMM Multi-satellite Precipitation Analysis (TMPA). In *Satellite Rainfall Applications for Surface Hydrology* (pp. 3–22). Springer, Dordrecht. https://doi.org/10.1007/978-90-481-2915-7_1

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., Wolff, D. B., & Xie, P. (2020). Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). *Advances in Global Change Research*, *67*, 343–353. https://doi.org/10.1007/978-3-030-24568-9_19

Hussain, Y., Satgé, F., Hussain, M. B., Martinez-Carvajal, H., Bonnet, M. P., Cárdenas-Soto, M., Roig, H. L., & Akhter, G. (2018). Performance of CMORPH, TMPA, and PERSIANN rainfall datasets over plain, mountainous, and glacial regions of Pakistan. *Theoretical and Applied Climatology*, *131*(3–4), 1119–1132. https://doi.org/10.1007/s00704-016-2027-z

Iizumi, T., Takikawa, H., Hirabayashi, Y., Hanasaki, N., & Nishimori, M. (2017). Contributions of different bias-correction methods and reference meteorological forcing data sets to uncertainty in projected temperature and precipitation extremes. *Journal of Geophysical Research: Atmospheres*, *122*(15), 7800–7819. https://doi.org/10.1002/2017JD026613

Irvem, A., & Ozbuldu, M. (2019). Evaluation of Satellite and Reanalysis Precipitation Products Using GIS for All Basins in Turkey. *Advances in Meteorology*, *2019*. https://doi.org/10.1155/2019/4820136

Jaber, A. (2020). *Evaluating SM2RAIN and WRF reanalysis precipitation datasets over Turkey and hydrological model application*. https://gcris.eskisehir.edu.tr/handle/20.500.13087/1015

Jiang, L., & 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. https://doi.org/10.1016/J.JHYDROL.2019.03.042

Jiang, S. hu, Wei, L. yong, Ren, L. liang, Zhang, L. qi, Wang, M. hao, & Cui, H. (2023). Evaluation of IMERG, TMPA, ERA5, and CPC precipitation products over mainland China: Spatiotemporal patterns and extremes. *Water Science and Engineering*, *16*(1), 45–56. https://doi.org/10.1016/j.wse.2022.05.001

Jiang, S., Liu, S., Ren, L., Yong, B., Zhang, L., Wang, M., Lu, Y., & He, Y. (2017). Hydrologic Evaluation of Six High Resolution Satellite Precipitation Products in Capturing Extreme Precipitation and Streamflow over a Medium-Sized Basin in China. *Water 2018, Vol. 10, Page 25*, *10*(1), 25. https://doi.org/10.3390/W10010025

Jiang, S., Ren, L., Hong, Y., Yong, B., Yang, X., Yuan, F., & Ma, M. (2012). Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. *Journal of Hydrology*, *452–453*, 213–225. https://doi.org/10.1016/J.JHYDROL.2012.05.055

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & 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. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Karakoc, U., & Patil, S. (2016). Comparison of TRMM satellite and ground-based precipitation data for predicting streamflow in Kucuk Menderes river basin, Turkey. *Geophysical Research Abstracts*, *18*, 2016–1467.

Karmouda, N., Kacimi, I., Elkharrim, M., Brirhet, · Hassan, & Hamidi, M. (2022). Geo-statistical and hydrological assessment of three satellite precipitation products over Ouergha basin (Northern Morocco). *Arabian Journal of Geosciences 2022 15:3*, *15*(3), 1–23. https://doi.org/10.1007/S12517-021-09124-6

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *424–425*, 264–277. https://doi.org/10.1016/J.JHYDROL.2012.01.011

Lei, H., Li, H., Zhao, H., Ao, T., & Li, X. (2021). Comprehensive evaluation of satellite and reanalysis precipitation products over the eastern Tibetan plateau characterized by a high diversity of topographies. *Atmospheric Research*, *259*, 105661. https://doi.org/10.1016/J.ATMOSRES.2021.105661

Lei, H., Zhao, H., & Ao, T. (2022). Ground validation and error decomposition for six state-of-the-art satellite precipitation products over mainland China. *Atmospheric Research*, *269*, 106017. https://doi.org/10.1016/J.ATMOSRES.2022.106017

Li, G., Yu, Z., Wang, W., Ju, Q., & Chen, X. (2021). Analysis of the spatial Distribution of precipitation and topography with GPM data in the Tibetan Plateau. *Atmospheric Research*, *247*, 105259. https://doi.org/10.1016/J.ATMOSRES.2020.105259

Li, N., Tang, G., Zhao, P., Hong, Y., Gou, Y., & 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. https://doi.org/10.1016/J.ATMOSRES.2016.07.020

Li, X. H., Zhang, Q., & Xu, C. Y. (2012). Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin. *Journal of Hydrology*, *426–427*, 28–38. https://doi.org/10.1016/J.JHYDROL.2012.01.013

Li, Z., Yang, D., & Hong, Y. (2013). Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. *Journal of Hydrology*, *500*, 157–169. https://doi.org/10.1016/J.JHYDROL.2013.07.023

Liu, J., Shangguan, D., Liu, S., Ding, Y., Wang, S., & Wang, X. (2019). Evaluation and comparison of CHIRPS and MSWEP daily-precipitation products in the Qinghai-Tibet Plateau during the period of 1981–2015. *Atmospheric Research*, *230*, 104634. https://doi.org/10.1016/J.ATMOSRES.2019.104634

Liu, J., Xia, J., She, D., Li, L., Wang, Q., & Zou, L. (2019). Evaluation of six satellite-based precipitation products and their ability for capturing characteristics of extreme precipitation events over a climate transition area in China. *Remote Sensing*, *11*(12), 1477. https://doi.org/10.3390/rs11121477

Lu, C., Ye, J., Fang, G., Huang, X., & Yan, M. (2021). Assessment of GPM IMERG Satellite Precipitation Estimation under Complex Climatic and Topographic Conditions. *Atmosphere 2021, Vol. 12, Page 780*, *12*(6), 780. https://doi.org/10.3390/ATMOS12060780

Ly, S., Charles, C., & Degré, A. (2013). Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review. *Biotechnology, Agronomy and Society and Environment*, *17*(2), 392–406.

Ma, Q., Li, Y., Feng, H., Yu, Q., Zou, Y., Liu, F., & Pulatov, B. (2021). Performance evaluation and correction of precipitation data using the 20-year IMERG and TMPA precipitation products in diverse subregions of China. *Atmospheric Research*, *249*, 105304. https://doi.org/10.1016/j.atmosres.2020.105304

Mengistu, D., Bewket, W., Dosio, A., & Panitz, H. J. (2021). Climate change impacts on water resources in the Upper Blue Nile (Abay) River Basin, Ethiopia. *Journal of Hydrology*, *592*, 125614. https://doi.org/10.1016/J.JHYDROL.2020.125614

Merz, R., Parajka, J., & Blöschl, G. (2011). Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resources Research*, *47*(2), 2531. https://doi.org/10.1029/2010WR009505

Milewski, A., Elkadiri, R., & Durham, M. (2015). Assessment and Comparison of TMPA Satellite Precipitation Products in Varying Climatic and Topographic Regimes in Morocco. *Remote Sensing 2015, Vol. 7, Pages 5697-5717*, *7*(5), 5697–5717. https://doi.org/10.3390/RS70505697

Moazami, S., Golian, S., Hong, Y., Sheng, C., & Kavianpour, M. R. (2016). Comprehensive evaluation of four high-resolution satellite precipitation products under diverse climate conditions in Iran. *Hydrological Sciences Journal*, *61*(2), 420–440. https://doi.org/10.1080/02626667.2014.987675

Moazami, S., Golian, S., Kavianpour, M. R., & Hong, Y. (2013). Comparison of PERSIANN and V7 TRMM multi-satellite precipitation analysis (TMPA) products with rain gauge data over Iran. *International Journal of Remote Sensing*, *34*(22), 8156–8171. https://doi.org/10.1080/01431161.2013.833360

Montero, R. A., Schwanenberg, D., Krahe, P., Lisniak, D., Sensoy, A., Sorman, A. A., & Akkol, B. (2016). Moving horizon estimation for assimilating H-SAF remote sensing data into the HBV hydrological model. *Advances in Water Resources*, *92*, 248–257. https://doi.org/10.1016/J.ADVWATRES.2016.04.011

Morsy, M., Scholten, T., Michaelides, S., Borg, E., Sherief, Y., & Dietrich, P. (2021). Comparative analysis of TMPA and IMERG precipitation datasets in the arid environment of El-Qaa plain, Sinai. *Remote Sensing*, *13*(4), 1–19. https://doi.org/10.3390/rs13040588

Nashwan, M. S., Shahid, S., Dewan, A., Ismail, T., & Alias, N. (2020). Performance of five high resolution satellite-based precipitation products in arid region of Egypt: An evaluation. *Atmospheric Research*, *236*, 104809. https://doi.org/10.1016/j.atmosres.2019.104809

Navidi Nassaj, B., Zohrabi, N., Nikbakht Shahbazi, A., & Fathian, H. (2022). Evaluating the performance of eight global gridded precipitation datasets across Iran. *Dynamics of Atmospheres and Oceans*, *98*, 101297. https://doi.org/10.1016/J.DYNATMOCE.2022.101297

Neri, M., Parajka, J., & Toth, E. (2020). Importance of the informative content in the study area when regionalising rainfall-runoff model parameters: The role of nested catchments and gauging station density. *Hydrology and Earth System Sciences*, *24*(11), 5149–5171. https://doi.org/10.5194/HESS-24-5149-2020

Nguyen, H. H., Recknagel, F., Meyer, W., Frizenschaf, J., Ying, H., & Gibbs, M. S. (2019). Comparison of the alternative models SOURCE and SWAT for predicting catchment streamflow, sediment and nutrient loads under the effect of land use changes. *Science of The Total Environment*, *662*, 254–265. https://doi.org/10.1016/J.SCITOTENV.2019.01.286

Nguyen, P., Ombadi, M., Gorooh, V. A., Shearer, E. J., Sadeghi, M., Sorooshian, S., Hsu, K., Bolvin, D., & Ralph, M. F. (2020). PERSIANN Dynamic Infrared–Rain Rate (PDIR-Now): A Near-Real-Time, Quasi-Global Satellite Precipitation Dataset. *Journal of Hydrometeorology*, *21*(12), 2893–2906. https://doi.org/10.1175/JHM-D-20-0177.1

Nikahd, A., Hashim, M., & Nazemosadat, M. J. (2016). A Review of Uncertainty Sources on Weather Ground-Based Radar for Rainfall Estimation. *Applied Mechanics and Materials*, *818*, 254–271. https://doi.org/10.4028/WWW.SCIENTIFIC.NET/AMM.818.254

Nwachukwu, P. N., Satge, F., Yacoubi, S. El, Pinel, S., & Bonnet, M. P. (2020). From trmm to GPM: How reliable are satellite-based precipitation data across Nigeria? *Remote Sensing*, *12*(23), 1–22. https://doi.org/10.3390/rs12233964

Omonge, P., Feigl, M., Olang, L., Schulz, K., & Herrnegger, M. (2022). Evaluation of satellite precipitation products for water allocation studies in the Sio-Malaba-Malakisi river basin of East Africa. *Journal of Hydrology: Regional Studies*, *39*, 100983. https://doi.org/10.1016/J.EJRH.2021.100983

Oreggioni Weiberlen, F., & Báez Benítez, J. (2018). Assessment of satellite-based precipitation estimates over Paraguay. *Acta Geophysica*, *66*(3), 369–379. https://doi.org/10.1007/s11600-018-0146-x

Ougahi, J. H., & Mahmood, S. A. (2022). Evaluation of satellite-based and reanalysis precipitation datasets by hydrologic simulation in the Chenab river basin. *Journal of Water and Climate Change*, *13*(3), 1563–1582. https://doi.org/10.2166/wcc.2022.410

Parajka, J., Bezak, N., Burkhart, J., Hauksson, B., Holko, L., Hundecha, Y., Jenicek, M., Krajčí, P., Mangini, W., Molnar, P., Riboust, P., Rizzi, J., Sensoy, A., Thirel, G., & Viglione, A. (2019). Modis snowline elevation changes during snowmelt runoff events in Europe. *Journal of Hydrology and Hydromechanics*, *67*(1), 101–109. https://doi.org/10.2478/JOHH-2018-0011

Parajka, J., Merz, R., & Blöschl, G. (2007). Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments. *Hydrological Processes*, *21*(4), 435–446. https://doi.org/10.1002/HYP.6253

Prakash, S. (2019). Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. *Journal of Hydrology*, *571*, 50–59. https://doi.org/10.1016/J.JHYDROL.2019.01.036

Rahman, K. U., Shang, S., Shahid, M., & Wen, Y. (2019). Performance assessment of SM2RAIN-CCI and SM2RAIN-ASCAT precipitation products over Pakistan. *Remote Sensing*, *11*(17), 2040. https://doi.org/10.3390/rs11172040

Ray, R. L., Sishodia, R. P., & Tefera, G. W. (2022). Evaluation of Gridded Precipitation Data for Hydrologic Modeling in North-Central Texas. *Remote Sensing*, *14*(16), 3860. https://doi.org/10.3390/rs14163860

Reichle, R. H., Draper, C. S., Liu, Q., Girotto, M., Mahanama, S. P. P., Koster, R. D., & De Lannoy, G. J. M. (2017). Assessment of MERRA-2 land surface hydrology estimates. *Journal of Climate*, *30*(8), 2937–2960. https://doi.org/10.1175/JCLI-D-16-0720.1

Reichle, R. H., Liu, Q., Koster, R. D., Draper, C. S., Mahanama, S. P. P., & Partyka, G. S. (2017). Land Surface Precipitation in MERRA-2. *Journal of Climate*, *30*(5), 1643–1664. https://doi.org/10.1175/JCLI-D-16-0570.1

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., … Woollen, J. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, *24*(14), 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Saber, M., & Yilmaz, K. K. (2018). Evaluation and Bias Correction of Satellite-Based Rainfall Estimates for Modelling Flash Floods over the Mediterranean region: Application to Karpuz River Basin, Turkey. *Water 2018, Vol. 10, Page 657*, *10*(5), 657. https://doi.org/10.3390/W10050657

Saddique, N., Muzammil, M., Jahangir, I., Sarwar, A., Ahmed, E., Aslam, R. A., & Bernhofer, C. (2022). Hydrological evaluation of 14 satellite-based, gauge-based and reanalysis precipitation products in a data-scarce mountainous catchment. *Hydrological Sciences Journal*, 1–15. https://doi.org/10.1080/02626667.2021.2022152

Saeed, F., Bethke, I., Fischer, E., Legutke, S., Shiogama, H., Stone, D. A., & Schleussner, C. F. (2018). Robust changes in tropical rainy season length at 1.5 °C and 2 °C. *Environmental Research Letters*, *13*(6), 064024. https://doi.org/10.1088/1748-9326/AAB797

Saemian, P., Hosseini-Moghari, S. M., Fatehi, I., Shoarinezhad, V., Modiri, E., Tourian, M. J., Tang, Q., Nowak, W., Bárdossy, A., & Sneeuw, N. (2021). Comprehensive evaluation of precipitation datasets over Iran. *Journal of Hydrology*, *603*, 127054. https://doi.org/10.1016/j.jhydrol.2021.127054

Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y. T., Chuang, H. Y., Juang, H. M. H., Sela, J., … Goldberg, M. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1057. https://doi.org/10.1175/2010BAMS3001.1

Sahlu, D., Moges, S. A., Nikolopoulos, E. I., Anagnostou, E. N., & Hailu, D. (2017). Evaluation of High-Resolution Multisatellite and Reanalysis Rainfall Products over East Africa. *Advances in Meteorology*, *2017*. https://doi.org/10.1155/2017/4957960

Salih, W., Chehbouni, A., & Epule, T. E. (2022). Evaluation of the Performance of Multi-Source Satellite Products in Simulating Observed Precipitation over the Tensift Basin in Morocco. *Remote Sensing 2022, Vol. 14, Page 1171*, *14*(5), 1171. https://doi.org/10.3390/RS14051171

Satgé, F., Defrance, D., Sultan, B., Bonnet, M. P., Seyler, F., Rouché, N., Pierron, F., & Paturel, J. E. (2020). Evaluation of 23 gridded precipitation datasets across West Africa. *Journal of Hydrology*, *581*, 124412. https://doi.org/10.1016/J.JHYDROL.2019.124412

Satgé, F., Pillot, B., Roig, H., & Bonnet, M. P. (2021). Are gridded precipitation datasets a good option for streamflow simulation across the Juruá river basin, Amazon? *Journal of Hydrology*, *602*, 126773. https://doi.org/10.1016/J.JHYDROL.2021.126773

Selek, B., & Aksu, H. (2020). Water Resources Potential of Turkey. In *Water Resources of Turkey* (pp. 241–256). Springer, Cham. https://doi.org/10.1007/978-3-030-11729-0_8

Şensoy, A., Uysal, G., & Şorman, A. A. (2023). Assessment of H SAF satellite snow products in hydrological applications over the Upper Euphrates Basin. *Theoretical and Applied Climatology*, *151*(1–2), 535–551. https://doi.org/10.1007/s00704-022-04292-1

Sensoy, S. (2004). The mountains influence on Turkey climate. *In Balwois Conference on Water Observation and Information System for Decision Support.*, 25–29.

Sevruk, B. (1987). Point precipitation measurements: why are they not corrected. *Water for the Future: Hydrology in Perspective*, 447–457.

Sevruk, B., Ondrás, M., & Chvíla, B. (2009). The WMO precipitation measurement intercomparisons. *Atmospheric Research*, *92*(3), 376–380. https://doi.org/10.1016/J.ATMOSRES.2009.01.016

Sharannya, T. M., Al-Ansari, N., Barma, S. D., & Mahesha, A. (2020). Evaluation of Satellite Precipitation Products in Simulating Streamflow in a Humid Tropical Catchment of India Using a Semi-Distributed Hydrological Model. *Water 2020, Vol. 12, Page 2400*, *12*(9), 2400. https://doi.org/10.3390/W12092400

Sharifi, E., Eitzinger, J., & Dorigo, W. (2019). Performance of the State-Of-The-Art Gridded Precipitation Products over Mountainous Terrain: A Regional Study over Austria. *Remote Sensing 2019, Vol. 11, Page 2018*, *11*(17), 2018. https://doi.org/10.3390/RS11172018

Shen, Y., Xiong, A., Wang, Y., & Xie, P. (2010). Performance of high-resolution satellite precipitation products over China. *Journal of Geophysical Research: Atmospheres*, *115*(D2), 2114. https://doi.org/10.1029/2009JD012097

Shen, Z., Yong, B., Yi, L., Wu, H., & Xu, H. (2022). From TRMM to GPM, how do improvements of post/near-real-time satellite precipitation estimates manifest? *Atmospheric Research*, *268*, 106029. https://doi.org/10.1016/j.atmosres.2022.106029

Singh, V. P. (2018). Hydrologic modeling: progress and future directions. *Geoscience Letters 2018 5:1*, *5*(1), 1–18. https://doi.org/10.1186/S40562-018-0113-Z

Sohn, B. J., Ryu, G. H., Song, H. J., & Ou, M. L. (2013). Characteristic features of warm-type rain producing heavy rainfall over the korean peninsula inferred from TRMM measurements. *Monthly Weather Review*, *141*(11), 3873–3888. https://doi.org/10.1175/MWR-D-13-00075.1

Sorman, A. A., Tas, E., & Dogan, Y. O. (2020). Comparison of hydrological models in upper Aras Basin Yukarı Aras Havzasında hidrolojik modellerin karşılaştırması. *Pamukkale University Journal of Engineering Sciences*, *26*(6), 1015–1022. https://doi.org/10.5505/pajes.2019.98852

Şorman, A. A., Uysal, G., & Şensoy, A. (2019). Probabilistic snow cover and ensemble streamflow estimations in the upper euphrates basin. *Journal of Hydrology and Hydromechanics*, *67*(1), 82–92. https://doi.org/10.2478/JOHH-2018-0025

Şorman, A. Ü., Akyürek, Z., Şensoy, A., Şorman, A. A., & Tekeli, A. E. (2007).

Commentary on comparison of MODIS snow cover and albedo products with ground observations over the mountainous terrain of Turkey. *Hydrology and Earth System Sciences*, *11*(4), 1353–1360. https://doi.org/10.5194/HESS-11-1353-2007

Sorman, A. U., & Beser, O. (2013). Determination of snow water equivalent over the eastern part of Turkey using passive microwave data. *Hydrological Processes*, *27*(14), 1945–1958. https://doi.org/10.1002/HYP.9267

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, *81*(9), 2035–2046. https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2

Stisen, S., & Sandholt, I. (2010). Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling. *Hydrological Processes*, *24*(7), 879–891. https://doi.org/10.1002/HYP.7529

Su, F., Hong, Y., & Lettenmaier, D. P. (2008). Evaluation of TRMM multisatellite precipitation analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *Journal of Hydrometeorology*, *9*(4), 622–640. https://doi.org/10.1175/2007JHM944.1

Su, J., Li, X., Ren, W., Lü, H., & Zheng, D. (2021). How reliable are the satellite-based precipitation estimations in guiding hydrological modelling in South China? *Journal of Hydrology*, *602*, 126705. https://doi.org/10.1016/J.JHYDROL.2021.126705

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. L. (2018). A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Reviews of Geophysics*, *56*(1), 79–107. https://doi.org/10.1002/2017RG000574

Sun, R., Yuan, H., Liu, X., & Jiang, X. (2016). Evaluation of the latest satellite–gauge precipitation products and their hydrologic applications over the Huaihe River basin. *Journal of Hydrology*, *536*, 302–319. https://doi.org/10.1016/J.JHYDROL.2016.02.054

Sun, W., Chen, R., Wang, L., Wang, Y., Han, C., & Huai, B. (2022). How do GPM and TRMM precipitation products perform in alpine regions?: A case study in northwestern China's Qilian Mountains. *Journal of Geographical Sciences*, *32*(5), 913–931. https://doi.org/10.1007/s11442-022-1978-5

Sunilkumar, K., Yatagai, A., & 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. https://doi.org/10.1029/2018EA000503

Talchabhadel, R., Aryal, A., Kawaike, K., Yamanoi, K., & Nakagawa, H. (2021). A comprehensive analysis of projected changes of extreme precipitation indices in West Rapti River basin, Nepal under changing climate. *International Journal of Climatology*, *41*(S1), E2581–E2599. https://doi.org/10.1002/JOC.6866

Talchabhadel, R., Aryal, A., Kawaike, K., Yamanoi, K., Nakagawa, H., Bhatta, B., Karki, S., & Thapa, B. R. (2021). Evaluation of precipitation elasticity using precipitation data from ground and satellite-based estimates and watershed modeling in Western Nepal. *Journal of Hydrology: Regional Studies*, *33*, 100768. https://doi.org/10.1016/J.EJRH.2020.100768

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A. T., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., & de Castro, M. (2012). Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research*, *104–105*, 70–97. https://doi.org/10.1016/J.ATMOSRES.2011.10.021

Tekeli, A. E., Akyürek, Z., Şorman, A. A., Şensoy, A., & Şorman, A. Ü. (2005). Using MODIS snow cover maps in modeling snowmelt runoff process in the eastern part of Turkey. *Remote Sensing of Environment*, *97*(2), 216–230. https://doi.org/10.1016/J.RSE.2005.03.013

Temimi, M., Zhan, X., Wen, J., Liu, R., Ramadhan, R., Yusnaini, H., Marzuki, M., Muharsyah, R., Suryanto, W., Sholihun, S., Vonnisa, M., Harmadi, H., Ningsih, A. P., Battaglia, A., Hashiguchi, H., & Tokay, A. (2022). Evaluation of GPM IMERG Performance Using Gauge Data over Indonesian Maritime Continent at Different Time Scales. *Remote Sensing 2022, Vol. 14, Page 1172*, *14*(5), 1172. https://doi.org/10.3390/RS14051172

Tian, Y., Peters-Lidard, C. D., Choudhury, B. J., & Garcia, M. (2007). Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. *Journal of Hydrometeorology*, *8*(6), 1165–1183. https://doi.org/10.1175/2007JHM859.1

Turkes, M. (1996). SPATIAL AND TEMPORAL ANALYSIS OF ANNUAL RAINFALL VARIATIONS IN TURKEY. *INTERNATIONAL JOURNAL OF CLIMATOLOGY*, *16*, 1057–1076. https://doi.org/10.1002/(SICI)1097-0088(199609)16:9

Uysal, G. (2022). Product- and Hydro-Validation of Satellite-Based Precipitation Data Sets for a Poorly Gauged Snow-Fed Basin in Turkey. *Water (Switzerland)*, *14*(17), 2758. https://doi.org/10.3390/W14172758/S1

Uysal, G., Alvarado-Montero, R., Şensoy, A., & Şorman, A. A. (2023). Comparison of Deterministic and Probabilistic Variational Data Assimilation Methods Using Snow and Streamflow Data Coupled in HBV Model for Upper Euphrates Basin. *Geosciences (Switzerland)*, *13*(3), 89. https://doi.org/10.3390/geosciences13030089

Uysal, G., & Şorman, A. Ü. (2021). Evaluation of PERSIANN family remote sensing precipitation products for snowmelt runoff estimation in a mountainous basin. *Hydrological Sciences Journal*, *66*(12), 1790–1807. https://doi.org/10.1080/02626667.2021.1954651

Wang, X., Ding, Y., Zhao, C., & Wang, J. (2019). 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. https://doi.org/10.1016/J.ATMOSRES.2018.12.011

Wang, Z., Zhong, R., Lai, C., & Chen, J. (2017). Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmospheric Research*, *196*, 151–163. https://doi.org/10.1016/J.ATMOSRES.2017.06.020

WMO. (2008). Volume I: Hydrology-From Measurement to Hydrological Information. In *Guide to Hydrological Practices* (p. 296). https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=WMO%2C+Guide+to+Hydrological+Practices.+Volume+I.+Hydrology–From+Measurement+to+Hydrological+Information.+World+Meteorological+Organization+%28WMO%29+Geneva+%282008%29.&btnG=

World Water Assessment Programme (United Nations). (2018). *The United nations world water development report 2018 (United nations educational, scientific and cultural organization, New York, United States)*. https://www.unwater.org/publications/world-water-development-report-2018/

Xiang, Y., Chen, J., Li, L., Peng, T., & Yin, Z. (2021). Evaluation of eight global precipitation datasets in hydrological modeling. *Remote Sensing*, *13*(14), 2831. https://doi.org/10.3390/rs13142831

Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C., & Yang, S. (2007). A Gauge-Based Analysis of Daily Precipitation over East Asia. *Journal of Hydrometeorology*, *8*(3), 607–626. https://doi.org/10.1175/JHM583.1

Xu, J., Ma, Z., Yan, S., & Peng, J. (2022). Do ERA5 and ERA5-land precipitation estimates outperform satellite-based precipitation products? A comprehensive comparison between state-of-the-art model-based and satellite-based precipitation products over mainland China. *Journal of Hydrology*, *605*, 127353. https://doi.org/10.1016/J.JHYDROL.2021.127353

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

Yan, J., & Bárdossy, A. (2019). Short time precipitation estimation using weather radar and surface observations: With rainfall displacement information integrated in a stochastic manner. *Journal of Hydrology*, *574*, 672–682. https://doi.org/10.1016/J.JHYDROL.2019.04.061

Yang, X., Yong, B., Hong, Y., Chen, S., & Zhang, X. (2016). Error analysis of multi-satellite precipitation estimates with an independent raingauge observation network over a medium-sized humid basin. *Hydrological Sciences Journal*, *61*(10), 1813–1830. https://doi.org/10.1080/02626667.2015.1040020

Yong, B., Ren, L. L., Hong, Y., Wang, J. H., Gourley, J. J., Jiang, S. H., Chen, X., & Wang, W. (2010). Hydrologic evaluation of Multisatellite Precipitation Analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China. *Water Resources Research*, *46*(7), 7542. https://doi.org/10.1029/2009WR008965

Yu, L., Leng, G., Python, A., & Peng, J. (2021). A comprehensive evaluation of latest GPM IMERG V06 early, late and final precipitation products across China. *Remote Sensing*, *13*(6), 1208. https://doi.org/10.3390/rs13061208

Yu, L., Zhang, Y., & Yang, Y. (2020). Using High-Density Rain Gauges to Validate the Accuracy of Satellite Precipitation Products over Complex Terrains. *Atmosphere 2020, Vol. 11, Page 633*, *11*(6), 633. https://doi.org/10.3390/ATMOS11060633

Yu, Z., Wu, J., & Chen, X. (2019). An approach to revising the climate forecast system reanalysis rainfall data in a sparsely-gauged mountain basin. *Atmospheric Research*, *220*, 194–205. https://doi.org/10.1016/J.ATMOSRES.2019.01.014

Yuan, F., Wang, B., Shi, C., Cui, W., Zhao, C., Liu, Y., Ren, L., Zhang, L., Zhu, Y., Chen, T., Jiang, S., & Yang, X. (2018). Evaluation of hydrological utility of IMERG Final run V05 and TMPA 3B42V7 satellite precipitation products in the Yellow River source region, China. *Journal of Hydrology*, *567*, 696–711. https://doi.org/10.1016/J.JHYDROL.2018.06.045

Yuan, X., Yang, K., Lu, H., He, J., Sun, J., & Wang, Y. (2021). Characterizing the features of precipitation for the Tibetan Plateau among four gridded datasets: Detection accuracy and spatio-temporal variabilities. *Atmospheric Research*, *264*, 105875. https://doi.org/10.1016/J.ATMOSRES.2021.105875

Yucel, I., Onen, A., Yilmaz, K. K., & Gochis, D. J. (2015). Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall. *Journal of Hydrology*, *523*, 49–66. https://doi.org/10.1016/J.JHYDROL.2015.01.042

Zambrano-Bigiarini, M., & Baez-Villanueva, O. (2020). *Tutorial for using hydroPSO to calibrate TUWmodel*. 1–60. https://doi.org/10.5281/ZENODO.3772176

Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., & Ribbe, L. (2017). Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrology and Earth System Sciences*, *21*(2), 1295–1320. https://doi.org/10.5194/HESS-21-1295-2017

Zambrano-Bigiarini, M., & Rojas, R. (2013). A model-independent Particle Swarm Optimisation software for model calibration. *Environmental Modelling & Software*, *43*, 5–25. https://doi.org/10.1016/J.ENVSOFT.2013.01.004

Zeng, Q., Wang, Y., Chen, L., Wang, Z., Zhu, H., & Li, B. (2018). Inter-comparison and evaluation of remote sensing precipitation products over China from 2005 to 2013. *Remote Sensing*, *10*(2), 168. https://doi.org/10.3390/rs10020168

Zhang, L., Chen, X., Lai, R., & Zhu, Z. (2022). Performance of satellite-based and reanalysis precipitation products under multi-temporal scales and extreme weather in mainland China. *Journal of Hydrology*, *605*, 127389. https://doi.org/10.1016/J.JHYDROL.2021.127389

Zhang, L., Lan, P., Qin, G., Mello, C. R., Boyer, E. W., Luo, P., & Guo, L. (2021). Evaluation of Three Gridded Precipitation Products to Quantify Water Inputs over

Complex Mountainous Terrain of Western China. *Remote Sensing 2021, Vol. 13, Page 3795*, *13*(19), 3795. https://doi.org/10.3390/RS13193795

Zhang, Y., Sun, A., Sun, H., Gui, D., Xue, J., Liao, W., Yan, D., Zhao, N., & Zeng, X. (2019). Error adjustment of TMPA satellite precipitation estimates and assessment of their hydrological utility in the middle and upper Yangtze River Basin, China. *Atmospheric Research*, *216*, 52–64. https://doi.org/10.1016/J.ATMOSRES.2018.09.021

Zhang, Y., Wu, C., Yeh, P. J. F., Li, J., Hu, B. X., Feng, P., & Jun, C. (2022). Evaluation and comparison of precipitation estimates and hydrologic utility of CHIRPS, TRMM 3B42 V7 and PERSIANN-CDR products in various climate regimes. *Atmospheric Research*, *265*, 105881. https://doi.org/10.1016/j.atmosres.2021.105881

Zhao, C., Ren, L., Yuan, F., Zhang, L., Jiang, S., Shi, J., Chen, T., Liu, S., Yang, X., Liu, Y., & Fernandez-Rodriguez, E. (2020). Statistical and Hydrological Evaluations of Multiple Satellite Precipitation Products in the Yellow River Source Region of China. *Water 2020, Vol. 12, Page 3082*, *12*(11), 3082. https://doi.org/10.3390/W12113082

Zhao, T., & Yatagai, A. (2014). Evaluation of TRMM 3B42 product using a new gauge-based analysis of daily precipitation over China. *International Journal of Climatology*, *34*(8), 2749–2762. https://doi.org/10.1002/JOC.3872

Zhu, H., Li, Y., Huang, Y., Li, Y., Hou, C., & Shi, X. (2018). Evaluation and hydrological application of satellite-based precipitation datasets in driving hydrological models over the Huifa river basin in Northeast China. *Atmospheric Research*, *207*, 28–41. https://doi.org/10.1016/J.ATMOSRES.2018.02.022

http-1: ftp://ftp.cpc.ncep.noaa.gov/precip/CPCUNI_PRCP (accessed on 15 March 2020).

http-2: http://www.gloh2o.org/mswx (accessed on 10 July 2022).

http-3: http://www.gloh2o.org/mswep (accessed on Januray2021).

http-4: ftp://ftp.chg.ucsb.edu/pub/org/chg/products/CHIRP (accessed on 10 April 2020).

http-5: ftp://ftp.chg.ucsb.edu/pub/org/chg/products/CHIRPS-2.0 (accessed on 10 April 2020).

http-6: https://daac.gsfc.nasa.gov (accessed on 20 June 2021).

http-7: https://cds.climate.copernicus.eu/ (accessed on 25 May 2020).

http-8: https://rda.ucar.edu/datasets/ds094.1/ (accessed on 18 May 2021).

http-9: https://doi.org/10.5281/zenodo.3635932 (accessed on 10 June 2020).

http-10: https://pmm.nasa.gov/TRMM (accessed on 10 May 2020).

http-11: https://pmm.nasa.gov/GPM (accessed on 10 May 2020).

http-12: https://chrsdata.eng.uci.edu/ (accessed on 10 May 2020).

APPENDIX

CURRICULUM VITAE

ORCID ID: 0000-0002-2744-9516

Name, Surname : Hamed HAFIZI

Foreign Languages : English, Turkish, and Urdu

EDUCATION

- 2023, Eskisehir Technical University, Institute of Graduate Programs, Department of Civil Engineering, Programme in Hydraulics (Ph.D.)
- 2016, Kabul Polytechnic University, Faculty of Water Resources and Environmental Engineering, Department of Hydraulics and Hydraulic Structures (M.Sc.)
- 2013, Kabul Polytechnic University, Faculty of Water Resources and Environmental Engineering, Department of Hydraulics and Hydraulic Structures (B.Sc.)

PUBLICATIONS

- Hafizi, H., & Sorman, A. A. (2023). Performance assessment of multi-source, satellite-based and reanalysis precipitation products over variable climate of Turkey. *Theoretical and Applied Climatology*, 1-14. https://doi.org/10.1007/s00704-023-04538-6
- Hafizi, H., & Sorman, A. A. (2022). Integrating meteorological forcing from ground observations and MSWX dataset for streamflow prediction under multiple parameterization scenarios. *Water*, 14(17), 2721. https://doi.org/10.3390/w14172721
- Hafizi, H., & Sorman, A. A. (2022). Assessment of 13 gridded precipitation datasets for hydrological modeling in a mountainous basin. *Atmosphere*, 13(1), 143. https://doi.org/10.3390/atmos13010143
- Hamed Hafizi, & Ali Arda Sorman. (2022) Assessing the potential of four near-real-time satellite-based precipitation products for streamflow simulation in a mountainous catchment. *International Mountain Conference (IMC2022)*, Innsbruk, Austria. https://doi.org/10.5281/zenodo.7495683

- Hafizi, H., & Sorman, A. A. (2022). Performance assessment of CHIRPSv2. 0 and MERRA-2 gridded precipitation datasets over complex topography of Turkey. *Environmental Sciences Proceedings*, 19(1), 21. https://doi.org/10.3390/ecas2022-12815
- Hamed H, & Sorman A. A. (2021). Statistical Assessment and Hydrological Utility of Multiple Gridded Precipitation Datasets in Upper Aras Basin. *International Symposium on Remote Sensing in Meteorology (METEO IRS)*, Istanbul, Turkey. https://doi.org/10.5281/zenodo.7422092
- Hamed H, & Sorman A. A. (2021). Evaluating the Potential of Gridded Precipitation Datasets for Reproducing Streamflow in a Mountainous Basin. *14th INTERNATIONAL CONGRESS ON ADVANCES IN CIVIL ENGINEERING, Yıldız Technical University*, Istanbul, Turkey. https://doi.org/10.5281/zenodo.5812766
- Hafizi, H., & Sorman, A. A. (2021). Assessment of satellite and reanalysis precipitation products for rainfall–runoff modelling in a mountainous basin. *Environmental Sciences Proceedings*, 8(1), 25. https://doi.org/10.3390/ecas2021-10345
- Uysal, G., Hafizi, H., & Sorman, A. A. (2021, April). Spatial and temporal evaluation of multiple gridded precipitation datasets over complex topography and variable climate of Turkey. In *EGU General Assembly Conference Abstracts* (pp. EGU21-14239). https://doi.org/10.5194/egusphere-egu21-14239, 2021
- Hafizi, H., & Kalkan, K. (2020). Evaluation of object-based water body extraction approaches using Landsat-8 imagery. *Journal of Aeronautics and Space Technologies*, 13(1), 81-89.

GRANTS AND AWARDS:

- 2021, Best Paper Award in the 4th International Electronic Conference on Atmospheric Sciences (ECAS), Manuscript ID: sciforum-0469944
- 2020, Eskişehir Technical University Scientific Research Project (Project No: 20DRP214), Eskişehir, Turkey
- 2018, Turkish Language Course (C2 level + Academic language) Anadolu University, Eskişehir, Turkey
- 2017, Ph.D. fellowship award, Presidency for Turks Abroad and Related Communities, Turkey