Naeem Saddique, Muhammad Muzammil, Istakhar Jahangir, Abid Sarwar, Ehtesham Ahmed, Rana Ammar Aslam, Christian Bernhofer

Hydrological evaluation of 14 satellite-based, gauge-based and reanalysis precipitation products in a data-scarce mountainous catchment

Erstveröffentlichung in / First published in:

Hydrological Sciences Journal. 2022, 67 (3), S. 436-450. Taylor & Francis. ISSN 2150-3435.

DOI: https://doi.org/10.1080/02626667.2021.2022152

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

Hydrological evaluation of 14 satellite-based, gauge-based and reanalysis precipitation products in a data-scarce mountainous catchment

Naeem Saddique, Muhammad Muzammil, Istakhar Jahangir, Abid Sarwar, Ehtesham Ahmed, Rana Ammar Aslam & Christian Bernhofer

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

Hydrological evaluation of 14 satellite-based, gauge-based and reanalysis precipitation products in a data-scarce mountainous catchment

Naeem Saddique[a,b], Muhammad Muzammil[b,c], Istakhar Jahangir[d], Abid Sarwar[b], Ehtesham Ahmed[e], Rana Ammar Aslam[f] and Christian Bernhofer[a]

[a]Institute of Hydrology and Meteorology, Technische Universität Dresden, Tharandt, Germany; [b]Department of Irrigation and Drainage, University of Agriculture, Faisalabad, Pakistan; [c]Institute for Landscape Ecology and Resources Management (ILR), Research Centre for Bio Systems, Land Use and Nutrition (IFZ), Justus Liebig University, Giessen, Germany; [d]Department of Environmental Remote Sensing and Geoinformatics, University of Trier, Trier, Germany; [e]Institute of Urban and Industrial Water Management, Technische Universität Dresden, Dresden, Germany; [f]Department of Structures and Environmental Engineering, University of Agriculture, Faisalabad, Pakistan

ABSTRACT

Availability of high-quality data is a major problem for climate and hydrological studies, especially in basins with complex topography where gauge network is typically limited and unevenly distributed. This study investigates the performance of 14 precipitation products – seven satellite-based (SPPs), two gauge-based (GPPs) and five reanalysis products (RPPs) – against ground observations (1998–2007) in the transboundary Jhelum River basin (33 397 km^2). Among the seven SPPs (bias corrected), five demonstrate a significantly high correlation coefficient (CC > 0.7) with observed rainfall. However, most of the products tend to underestimate the seasonal precipitation amount, particularly in winter and spring. Likewise, Asian Precipitation – Highly Resolved Observational Data Integration Towards Evaluation of water resources APHRODITE (GPPs) and Japanese 55-year Reanalysis JRA-55 (RPPs) are the best-performing products in daily streamflow predictions, with Nash-Sutcliffe efficiency values of 0.68 and 0.62, whilst MSWEP (Multi-Source Weighted-Ensemble Precipitation), AgMERRA (Climate Forcing Dataset for Agricultural Modeling) and CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) have also good potential in flow prediction. Generally, our results indicate that APHRODITE and JRA-55 could be used as alternative sources of precipitation data in the Himalayas region.

ARTICLE HISTORY
Received 11 February 2021
Accepted 18 November 2021

EDITOR
A. Fiori

ASSOCIATE EDITOR
S. M. Pingale

KEYWORDS
evaluation; satellite-based precipitation products; Jhelum River basin; APHRODITE; JRA-55; Himalayas region

1 Introduction

Precipitation is the vital component of the hydrological cycle and the key forcing element of hydrological models. Thus, accurate precipitation assessment is important for hydrological simulation and water resources management (Liu *et al.* 2017). Traditionally, in developing countries land-based stations are unevenly distributed, and the quality of historical precipitation records can be questionable. For climatological purposes, the World Meteorological Organization (WMO) provides recommendations for the density of raingauges: from one gauge per 900 km^2 in lowland areas to one every 250 km^2 in complex topography (WMO 2008). Using around one gauge per 15 000 km^2, the Pakistan Meteorological Department has recently published 50 years' worth of area-weighted precipitation on a monthly scale based on 56 gauges (Faisal and Gafar 2012). These factors affect the accuracy of hydrological model simulations and related research work (Zhu *et al.* 2016). Therefore, it is crucial to look at spatial and temporal variability based on observation gauge data. Last but not least, land-based data are often difficult to obtain due to data policies and boundary disputes in some countries.

To overcome the above shortcomings, researchers in recent years have developed many globally and regionally gridded precipitation datasets based on reanalysis and remote sensing data that are convenient for hydrological simulation in high-land and poorly gauged catchments. Additionally, these datasets are freely available for scientific research. These gridded datasets have been widely used, especially where high-quality in situ rainfall measurements are not available.

Some of the most extensively used satellite precipitation products (SPPs) are those from the Tropical Rainfall Measuring Mission (TRMM) (Huffman *et al.* 2007) and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network climate data record (PERSIANN-CDR) (Ashouri *et al.* 2015). Compared with the TRMM satellite data products, the integrated multisatellite retrievals for GPM (IMERG) (Global Precipitation Measurements (GPM) Integrated Multi-satellite Retrievals) satellite data has a relatively finer spatial resolution (0.1° × 0.1°). However, the precipitation radar and the channels of the passive microwave (PMW) imager have been upgraded and expanded, which has enhanced the detection capability of weak and solid precipitation. Further, this makes it possible to detect rainfall in both arid and cold regions (Hamada and Takayabu 2016). Additionally, Multi-Source Weighted-Ensemble Precipitation (MSWEP), the AgMERRA Climate Forcing Dataset for Agricultural Modeling and the Global Land Data Assimilation System (GLDAS) using satellite- and ground-based observations are available at high spatial and

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

temporal resolution (Beck *et al.* 2017b). Climate Hazards Group InfraRed Precipitation (CHIRP) and CHIRP with station version (CHIRPS) are new quasi-global precipitation products with a spatial resolution of 0.05° (Funk *et al.* 2015). Global reanalysis-based climate products such as Climate Forecast System Reanalysis (CFSR) and Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) are widely used in hydrology and climate change research. Another dataset (gauge-based) with higher spatial and temporal coverage is the Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) (Yatagai *et al.* 2012). However, RPPs and SPPs are generally very uncertain due to complex climate dynamics and local topography. The discrepancies in the SPPs commonly originate from the platform and sensor characteristics. For the RPPs, uncertainties and errors are mainly caused by the inadequacy of interpolation methods and the data assimilation system (Zhu *et al.* 2016). Therefore, it is essential to evaluate the suitability of these grid-based precipitation products before further utilization.

Recently, many studies have been conducted for the evaluation of grid-based rainfall products over the different regions of the world (Lauri *et al.* 2014, Wong *et al.* 2017, Gebrechorkos *et al.* 2018, Senent-Aparicio *et al.* 2018, Wu *et al.* 2018, Tang *et al.* 2019b, Zandler *et al.* 2019, Ahmed *et al.* 2020). Tan *et al.* (2015) evaluated five SPPs and one ground-based rainfall product and found APHRODITE and TRMM3B42V7 (Tropical Rainfall Measuring Mission, Version 7) performed well, while other products slightly overestimate (PERSIAN-CDR) and underestimate (CMORPH: CPC (Climate Prediction Center) MORPHing technique) the rainfall. Moreover, several studies have focused on the evaluation of grid-based rainfall datasets in the hydrological modelling framework, which will provide useful information for water resources management and hydrological studies (Bui *et al.* 2019, AL-Falahi *et al.* 2020, Dembélé *et al.* 2020, Talchabhadel *et al.* 2020). For instance, TRMM and GPCC (Global Precipitation Climatology Centre) datasets are suitable for simulating the streamflow in the Dong Nai River basin, Vietnam (Nhi *et al.* 2018), MSWEP shows good performance for hydrological simulation of the Upper Huaihe River basin, China, and APHRODITE rainfall data is a good option for simulating the discharge of Amu Darya River basin (Sidike *et al.* 2016). Lauri *et al.* (2014) analysed the feasibility of five grid-based products (APHRODITE, TRMM3B42 (V6, V7), NCEP (National Centers for Environmental Prediction)-CFSR and the ERA (ECMWF (European Centre for Medium-Range Weather Forecasts) Reanalysis) Interim datasets) over the Mekong River basin and concluded that APHRODITE and TRMM (V7) simulate the flow very similarly to observed data. Birylo *et al.* (2018) computed the different components of the water budget using GLDAS data over Poland. Tang *et al.* (2019a) assessed the uncertainties of four precipitation products (AgMERRA, PERSIAN-CDR, MSWEP and TMPA (near-real time Multi-satellite Precipitation Analysis)) for Soil and Water Assessment Tool (SWAT) modelling in Mekong River basin. Generally, the results of the above studies show that the performance of precipitation products varies with different topography and climate conditions.

A few studies were conducted in Himalayan basins for the evaluation of SPPs and GPPs. Li *et al.* (2018b) evaluated four precipitation products (India Meteorological Department (IMD), APHRODITE, ERA-Interim and Weather Research & Forecasting (WRF)) and recommended the APHRODITE dataset for hydrological studies in the western Himalayas. Khan *et al.* (2018) assessed the suitability of the TRMM-3B42-V7 rainfall product over Upper Indus basin (UIB) and found underestimation of rainfall at most stations across the UIB; still, they suggested that TRMM is the best option in data-scar ce regions. Faiz *et al.* (2020) validated the performance of five precipitation products in two cryosphere catchments of UIB. Their study showed that the CHIRPS and APHRODITE precipitation datasets perform well in simulating discharge, with NSE (Nash-Sutcliffe efficiency) values greater than 0.80 at both river basins.

Several other precipitation products such as MSWEP, GLDAS, AgMERRA, CFSR, Japanese 55-year Reanalysis (JRA-55) and Princeton Global Forcing (PGF) are available at fine resolution and have yet to be tested in this region. However, significant improvement in regional runoff modelling using the different precipitation datasets (SPPs, GPPs and RPPs) is still needed. These precipitation datasets have been generally evaluated in different parts of the world, but the region, elevation and local climate dynamics could influence their accuracy. So, a detailed analysis of these data products may improve runoff modelling in the areas for which there is no or scarce climatic data. Based on a literature review and to the best of the authors' knowledge, it was found that not a single study has been conducted to date that uses SPPs and GPPs for streamflow prediction in Jhelum River basin. Such a study is very important as 50% of the basin lies in a disputed area between Pakistan and India (Kashmir). This part of the basin has only a few weather stations, and data acquisition is difficult for Pakistan.

The available precipitation products fall into three different types: SPPs, GPPs and RPPs. However, all of these datasets produce very uncertain rainfall estimations due to complex topography and local weather conditions. SPPs have an inherent advantage due to their higher spatial observations, but they also have certain limitations due to platforms and sensor characteristics. Reflectance from land surface, particularly snow and ice, can cause distinctive biases (Huffman *et al.* 2007). The GPPs are usually thought to be the most reliable, but users should be cautious when employing such data because of the inadequacy of interpolation methods and the unavoidable inferiority inherited from gauge measurements (Yang *et al.* 1998). RPPs are a combination of outputs from different observations and models. However, great caution must be exercised when using such data due to continuous changes in observing systems and model errors (Dee *et al.* 2011). Therefore, before such precipitation datasets can be used with confidence, it is important to evaluate their accuracy and error characteristics by comparing them with data from ground-based observations.

The main objective of this study is to evaluate the performance of seven SPPs, two GPPs and five RPPs in simulating streamflow with the SWAT hydrological model – using as a case study the Jhelum River basin (JRB; Mangla Dam Watershed),

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

which is one of the three major rivers of the Indus basin irrigation system (IBIS). The accuracy of these products was evaluated at daily, monthly and seasonal scales over the period 1998 to 2007 using the most widely applied statistical and graphical methods. Section 2 below provides a brief introduction to the study area together with a discussion of the observed raingauge and precipitation datasets. The methodology of the SWAT hydrological model is described in Section 3. In Section 4, evaluation results of precipitation products and streamflow simulation are provided, and conclusions are given in Section 5.

2 Study area and data

2.1 Description of study area

The JRB (Mangla Dam Watershed) is located between 33 and 35°N and between 73 and 75.62°E, and has a total drainage area of about 33 397 km^2. Figure 1 shows the location of the study area and the spatial distribution of grid points of precipitation products at 0.25° resolution within and around the study area. The watershed topography is mountainous with elevation varying from 232 m in lowland areas to 6287 m in highland areas. The catchment drains its whole flow into Mangla reservoir, which is the second largest reservoir of Pakistan. The water of this reservoir is mainly used for two purposes: to irrigate 6 Mha of agricultural land and to generate 1000 MW of electricity (which is 15% of the country's total electricity produced through hydropower plants). About 70% of rainfall in the basin occurred from March to August in the period 1961–2012 (Saddique *et al.* 2019a). Maximum and minimum rainfall occurs in July and November, respectively, whereas peak and low flows happen in June (more water due to snow melting) and January, respectively. The average annual flow at Azad Pattan is about 835 m^3/s. The climate of the basin is mild, with monthly average temperatures ranging from 4.9°C in January (coldest month) to 24.3°C in July (hottest month) (Saddique *et al.* 2020). The temperature of the basin decreases with increasing elevation (from south to north) but precipitation does not follow a specific trend in such a complex topography.

The average annual precipitation of each station is shown at its respective location in Fig. 1. The JRB is characterized by highly heterogeneous soil and land cover; the main types of soil include Gleyic Solonacks (49%), Calaric Phaoeozems (23%) and Mollic Planosols (21%) (Saddique *et al.* 2019b). The basin drainage area includes diverse land cover types: agriculture (30.81%), grass-sparse vegetation (36.79%), forest (27.88%), water (2.07%) and settlement (2.43%) (Saddique *et al.* 2019b).

2.2 Data description

2.2.1 Observed data

In situ daily precipitation and maximum and minimum temperature datasets for 16 stations were obtained from the Pakistan Meteorological Department (PMD), the Water and Power Development Authority (WAPDA) of Pakistan, and the

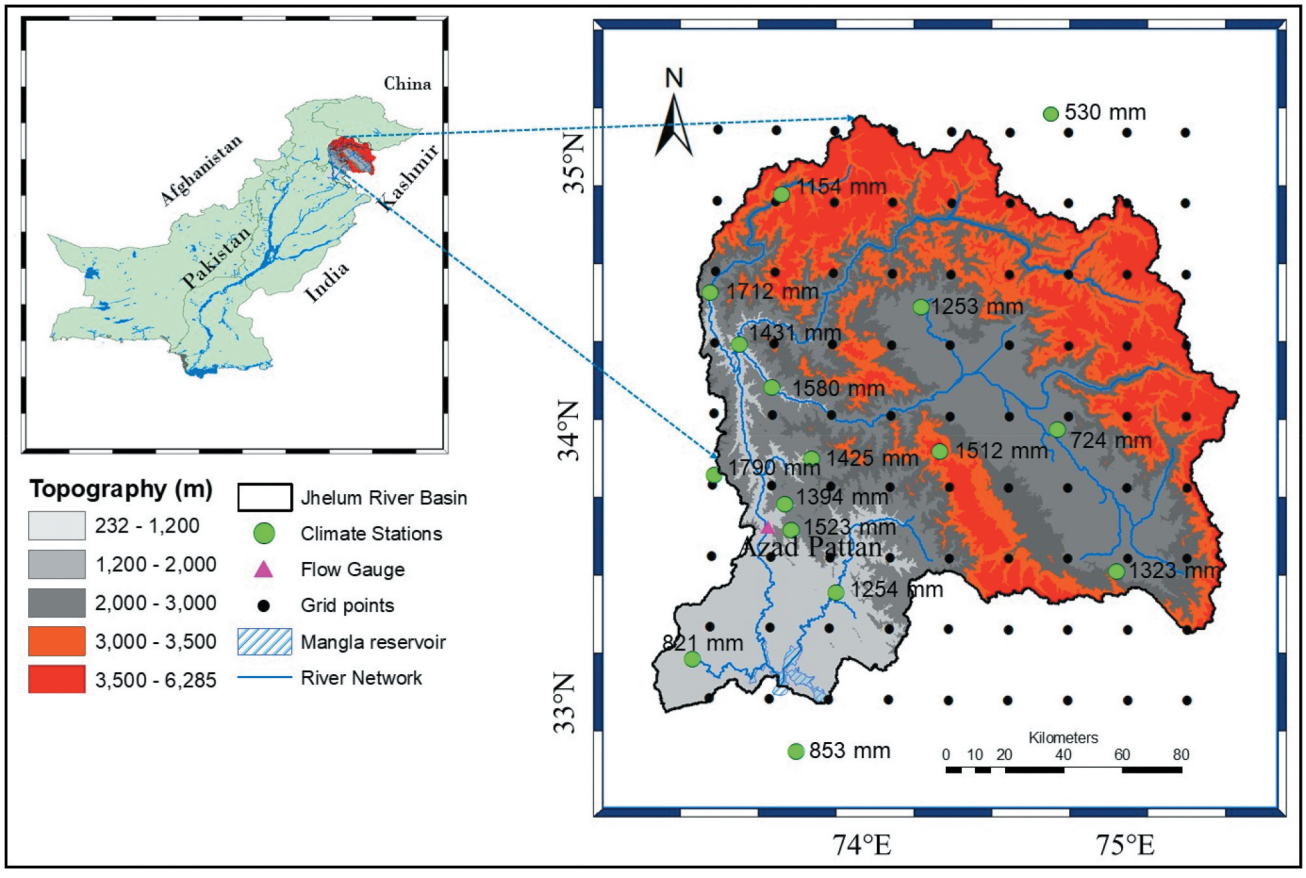

Figure 1. Location of Jhelum River basin, raingauges and grid points of precipitation products.

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

IMD. Relative humidity, wind speed and solar radiation data were simulated from the CFSR dataset available on the SWAT website. Daily flows for JRB at seven gauging stations were made available for the period 1995–2007, obtained from WAPDA.

2.2.2 Satellite-based precipitation products (SPPs)

CHIRPS is a quasi-global satellite dataset developed by the US Geological Survey (USGS) and the Climate Hazards Group for drought and environmental monitoring. First, infrared precipitation pentad (five-day) estimates are created from satellite data using information on cloud temperature and are calibrated with respect to the TMPA pentads. These pentads are divided by the long-term (1981–2013) normal values, and the fractions, multiplied against the corresponding Climate Hazards Precipitation, provide CHIRP estimates. Finally, pentad CHIRP values are redistributed to daily precipitation estimates based on daily National Oceanic and Atmospheric Administration – Climate Forecast System (NOAA-CFS) data (Zandler *et al.* 2019). CHIRPS was established by blending ground station data with rainfall derived from cold cloud duration (CCD) by the synergistic use of satellite infrared (IR). The rainfall product is available at a high spatial resolution (0.05°) from 1981 to the present at daily, pentad and monthly scale (Funk *et al.* 2015). In the current study, we have used both products (CHIRP and CHIRPS) of the Climate Hazard Group to evaluate their accuracy in the high-elevation basin.

AgMERRA high-resolution meteorological forcing datasets were designed to observe the impacts of climate variability and climate change on agriculture. AgMERRA daily temperature, precipitation and solar radiation were produced by combining MERRA daily resolution data with the 0.5° × 0.5° precipitation product of Climate prediction center unified (CPCU) from 1980 to 2005 and the real-time precipitation product of CPC from 2006 to 2010 with satellite observational datasets (Ruane *et al.* 2015). These datasets are available at a high spatiotemporal resolution (0.25° and daily), for the period 1980–2010.

PERSIAN-CDR daily rainfall data are available at 0.25° spatial resolution for the period 1983–present. Ashouri *et al.* (2015) estimated rainfall using the Gridsat-B1 IR data as input into an artificial neural network (ANN) algorithm and adjusted the high-resolution data using monthly data from the Global Precipitation Climatological Product Project (GPCP, 2.5°). Thereby, satellite estimates are rescaled to GPCP resolution and a correction factor is calculated using the ratio of the two products (Ashouri *et al.* 2015).

MSWEP version 2.1 is a three-hourly global precipitation dataset for the period 1979–2016 with a spatiotemporal resolution of 0.1°, specifically designed for hydrological modelling (Beck *et al.* 2019). This data includes a bias-corrected long-term climate mean obtained from the Climate Hazards Group Precipitation Climatology (CH-Pclim), streamflow data from the USGS Geospatial Attributes of Gages for Evaluating Streamflow (GAGES)-II, Precipitation-elevation Regressions on Independent Slopes Model (PRISM) climatic precipitation data, and data from the Global Runoff Data Centre (GRDC). MSWEP is based on merging the highest quality satellite, gauge and reanalysis data as a function of location and time. The temporal variability of the dataset was determined by the weighted average of global precipitation (P) anomalies of seven datasets such as NCEP-CFSR and JRA-55 and many gridded satellite-based precipitation products like PERSIANN, CMORPH, TMPA 3B42RT, SM2RAIN-ASCAT (Soil Moisture to Rain-Advanced SCATterometer), and GSMaPMVK (Global Satellite Mapping of Precipitation moving vector with Kalman filter).

TRMM data are based on microwave observations, precipitation radar and IR data. A number of precipitation datasets are produced by these sensors, such as 2A25 and 2B31 at level 2 and 3A25, 3B21, 3B42 and 3B43 at level 3. TRMM 3B42 data is available at a spatial resolution of 0.25° and temporal scales of three hours and daily. TRMM product uses multiple microwave and IR satellite precipitation estimates that are recalibrated with different GPCC datasets. Satellite data are adjusted using the large-scale means of the gauge analysis and combined by applying an inverse estimated-random-error variance weighting. In areas with high numbers of stations, the station values have a high effect on the resulting precipitation (Talchabhadel *et al.* 2020). However, in regions with poor gauge coverage (like our research area), the satellite input has much higher weight than the gauge adjustment (Huffman *et al.* 2007).

GLDAS is an extension of the North American Land Data Assimilation System (NALDAS) project (Rodell *et al.* 2004). The GLDAS fine-resolution precipitation product was generated by integrating satellite and ground station datasets with land surface models (LSMs). At present, GLDAS is generated by four different LSMs: Noah, Catchment, Community Land Model (CLM) and Variable Infiltration Capacity (VIC) model. In the present study, GLDAS V2.0/Catchment Land Surface Model daily 0.25° × 0.25° resolution was used to simulate the discharge.

2.2.3 Gauge-based precipitation products (GPPs)

APHRODITE is a state-of the-art high-resolution daily precipitation dataset. In the current study, we used the latest version for monsoon Asia at a spatial resolution of 0.25° available for the period 1951–2007. The precipitation data from a dense network of raingauges is first interpolated onto a grid of 0.05° using the modified version of the distance-weighting interpolation method, which considers sphericity and orography by the Spherema and the Mountain Mapper methods, respectively. This dataset is then regridded to 0.25° and 0.5° products using the area-weighted mean. The algorithm is improved in that the weighting function considers the local topography between the raingauge and interpolated point (Yatagai *et al.* 2012). APHRODITE uses the maximum number of observed gauges among interpolated available rainfall products, and this is therefore the most accurate dataset for Asia (Ménégoz *et al.* 2013).

CPCU is a global daily precipitation dataset from the NOAA CPC produced by using optimal interpolation (OI) on information from national and international agencies gauge reports that is collected from multiple sources, including the Global Telecommunications System and Cooperative Observer network (Chen *et al.* 2008). For data interpolation,

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

Table 1. Summary of satellite-based, gauge-based and reanalysis precipitation products. S: satellite, G: gauge, R: reanalysis.

	Datasets	Data source	Spatial resolution (°)	Temporal coverage	Temporal resolution	Reference
SPPs	CHIRPS	S, G, R	0.05	1981–present	Daily	Funk *et al.* (2015)
	AgMERRA	S, G	0.25	1980–2010	Daily	Ruane *et al.* (2015)
	CHIRP	S, R	0.05	1981–present	Daily	Funk *et al.* (2015)
	PERSIAN CDR	S, G	0.25	1983–present	Daily	Ashouri *et al.* (2015)
	MSWEP	S, G, R	0.25	1979–2015	Daily	Beck *et al.* (2017a)
	TRMM 3B42	S, G	0.25	1998–present	Daily	Huffman *et al.* (2007)
	GLDAS	S, G	0.25	1948–present	Daily	Rodell *et al.* (2004)
GPPs	APHRODITE	G	0.25	1951–2007	Daily	Yatagai *et al.* (2012)
	CPCU	G	0.5	1979–present	Daily	Chen *et al.* (2008)
RPPs	PGF	R	0.25	1948–2016	Daily	Sheffield *et al.* (2005)
	CFSR	R	0.313	1979–2017	Daily	Saha *et al.* (2014)
	MERRA-2	R	0.5/0.625	1980–2017	Daily	Gelaro *et al.* (2017)
	MERRA-1	R	0.5/0.66	1979–2010	Daily	Rienecker *et al.* (2011)
	JRA-55	R	0.5	1958–present	Daily	Ebita *et al.* (2011)

data from more than 30 000 gauges was used. CPCU version 1.0 was used in this study, and data was available at a spatial resolution of 0.5° over the period 1979–present.

2.2.4 Reanalysis precipitation products (RPPs)

PGF datasets are available at three different spatial (0.25°, 0.5° and 1°) and temporal (three-hourly, daily and monthly) resolutions from January 1948 to December 2016. The dataset was developed by the Terrestrial Hydrology Group at Princeton University by combining observation-based data with National Centers for Environmental Prediction – National Center for Atmospheric Research (NCEP-NCAR) reanalysis (Sheffield *et al.* 2005). We have used the latest version (V3) of PGF at 0.25° spatial resolution and at daily temporal resolution.

CFSR is a global coupled atmosphere–ocean–land surface–sea ice system. The CFSR includes (1) coupling of atmosphere and ocean during the generation of the six-hour guess field, (2) an interactive sea-ice model, and (3) assimilation of satellite radiances. CFSR uses the Noah land surface model, which is forced with the NOAA pentad CPC Merged Analysis of Precipitation and the CPCU daily gauge analysis instead of using the precipitation generated by the atmospheric model, which is considered too biased (Saha *et al.* 2010). The dataset is available at a daily time scale and a global atmosphere resolution of 0.31° (~38 km) from 1979 to the present. In this study, we have used CFSR version DS094.1.

MERRA-2 products are based on the assimilation of a vast number of in situ and remote sensing observations into an atmospheric general circulation model (AGCM) (Gelaro *et al.* 2017) and provide land and atmospheric conditions for the whole world. MERRA-2 is the most recent reanalysis dataset, and not many studies have been published analysing the improvement from its predecessor MERRA-1 (available from 1979 to 2010). MERRA-2 data are available at daily temporal and 0.5° spatial resolution over the period 1980–present.

The JRA-55 dataset was developed by the Japanese Meteorological Agency (JMA) to overcome the deficiencies in the previous generation, the Japanese 25-year Reanalysis (JRA-25). JRA-55 adopts a high-resolution model as well as an advanced data assimilation scheme (4D-Var) for the production of long-term atmospheric data (Ebita *et al.* 2011). In addition, JRA-55 uses several new observational datasets and greenhouse gases to improve the data quality. Table 1 provides a summary of the SPPs, GPPs and RPPs.

3 Methodology

3.1 Hydrological modelling

SWAT is a physically based semi-distributed hydrological model operating on sub-daily and daily time steps. It can be used to simulate the runoff, sediment and modelling of best management practices (BMPs) in meso- and macro-scale river basins (Arnold *et al.* 1998). It discretizes the watershed into sub-basins based on elevation and slope, which will be further divided into hydrological response units (HRUs) based on land use, topography and soil data. Soil Convention Service (SCS) runoff curve and Hargreaves methods were used to calculate the surface runoff process and evapotranspiration (ET), respectively. Table 2 provides a description of the datasets used in the SWAT model.

The SUFI-2 (Uncertainty in Sequential Uncertainty Fitting) algorithm in SWAT-
CUP (Calibration and Uncertainty Programs) was used for calibration and validation of model. Before calibration, global sensitivity analysis (GSA) was conducted in SWAT-CUP to analyse the sensitivity of 30 parameters selected from other studies. All information related to parameters (initial ranges and best-fit values) and calibration can be found in Saddique *et al.* (2019b). After the parameter sensitivity analysis, the model was run with the initial ranges of more sensitive parameters, and 3–4 iterations were used to calibrate the model (Saddique *et al.* 2019b). The lower and upper bounds of the

Table 2. Description of different datasets (DEM, land-use map, soil map) used in the SWAT model.

Data type	Data description	Resolution	Data source
Topography	Digital elevation map (DEM)	30 m	Shuttle Radar Topography Mission, NASA
Land-use types	Land-use/land-cover map	30 m	USGS Landsat 7
Soil types	Soil map	30 arc-seconds	Food and Agriculture Organization of the United Nations (FAO)

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

parameters are qualified as physically reasonable based on the SWAT model official documentation (Neitsch *et al.* 2005). One thousand simulations were run in each iteration; after each iteration, a new set of parameter ranges was given by SUFI-2 for the next simulation. The next iteration was performed based on the new parameter ranges. Detailed information about SUFI-2 and the protocol to calibrate the SWAT can be found in Abbaspour (2015).

3.2 Performance evaluation of precipitation datasets

Point-to-pixel and basin-averaged analysis was used to examine the raingauge observations and the corresponding precipitation products (Shayeghi *et al.* 2020). For the comprehensive evaluation, all the precipitation products are interpolated at a common grid resolution (0.25° × 0.25°) to overcome the spatial differences by using bilinear interpolation in climate data operators (CDOs). The most commonly used statistical indices, such as Pearson correlation coefficient (CC), percent bias (Pbias), root mean square error (RMSE) (Moazami *et al.* 2013), probability of detection (POD), false alarm ratio (FAR) and critical skill index (CSI) (Wilks 2011) are used for the performance evaluation of products (SPPs, GPPs and RPPs) with respect to station data. CC (Equation (1)) measures the strength of the relationship between products and observed values, and ranges between −1 and 1 (perfect fit). Pbias (Equation (2)) indicates the direction of the discrepancy bias and it varies between −∞ and ∞. Positive values of Pbias show overestimated precipitation product values, whereas negative values indicate underestimation. RMSE (Equation (3)) is a well-known goodness-of-fit indicator, which marks the differences between product outputs and observed values. POD (Equation (4)), FAR (Equation (5)), and CSI (Equation (6)) range from 0 to 1. A higher value represents a perfect fit for POD and CSI, whereas a lower value represents the ideal condition for FAR. Model simulation will be considered satisfactory if NSE values are greater than 0.5, PBIAS is within ± 25% (Moriasi *et al.* 2007) and R^2 is greater than 0.5 (Santhi *et al.* 2001).

$$CC = \frac{\sum_{i=1}^{n}(S_i - S).(O_i - O)}{\sqrt{\sum_{i=1}^{n}(S_i - S)^2}.\sqrt{\sum_{i=1}^{n}(O_i - O)^2}} \quad (1)$$

$$Pbias = \frac{\sum_{i=1}^{n}(S_i - O_i)}{\sum_{i=1}^{n} O_i} \times 100 \quad (2)$$

$$RMSE = \frac{\sqrt{\sum_{i=1}^{n}(O_i - S_i)^2}}{\sqrt{n}} \quad (3)$$

$$POD = \frac{H}{H + M} \quad (4)$$

$$FAR = \frac{F}{H + F} \quad (5)$$

$$CSI = \frac{H}{H + F + M} \quad (6)$$

Here S_i and O_i are the precipitation products and observed values, respectively; n is the number of paired values; S and $\bar{O}$ are the mean of product outputs and observed rainfall, respectively; H is the observed precipitation correctly detected; M is the observed precipitation not detected; and F is precipitation detected but not observed.

Moreover, a linear scaling bias correction method was used to correct the biases present in selected precipitation datasets using the historical observations (1998–2007). Monthly correction factors (CF = observed/precipitation products) were calculated at each station with the respective precipitation product and then these factors were multiplied with raw daily precipitation values in the corresponding months. Only the results of bias-corrected precipitation are presented in the Results and discussion section.

4 Results and discussion

4.1 Comparing observed data with SPPs, GPPs and RPPs

The spatial distribution of mean annual precipitation of seven 0.25° SPPs, two GPPs and five RPPs during the period 1998–2007 are presented in Fig. 2. All the precipitation products showed low rainfall in the southern and north-eastern parts of the considered area except JRA-55 and MSWEP. In general, most of the products revealed high precipitation at elevation ranges from 600 to 2500 m a.s.l. Pang *et al.* (2014) and Dahri *et al.* (2016) also reported that the precipitation decreases significantly in Himalayas at elevations above 2400 m a.s.l. However, our results contrast with those of Immerzeel *et al.* (2012), who found a linearly increasing relationship between precipitation and altitude. Not surprisingly, two satellite (CHIRPS and CHIRP) and two reanalysis (MERRA-2 and MERRA-1) datasets show a similar spatial pattern with little difference in magnitude. Visually MSWEP, MERRA-2 and JRA-55 products showed high precipitation values in all grids compared to the other datasets.

Table 3 presents the performance evaluation of daily observed precipitation with 14 precipitation products (SPPs, GPPs and RPPs). The evaluation of each precipitation product showed a different degree of agreement with observed data. The CCs of the seven satellite datasets, two gauge datasets and five reanalysis datasets ranged from 0.68 to 0.77, 0.46 to 0.79 and 0.52 to 0.76, respectively, while RMSE values ranged from 1.55 mm to 2.65 mm for SPPs, 1.51 mm to 3.52 mm for GPPs and 1.61 mm to 3.65 mm for RPPs. Pbias values were low in magnitude for SPPs and high for RPPs for the period 1998–2007. The negative value of Pbias for MERRA-1 indicates that the data underestimated the observed daily precipitation (by −17.1%). In contrast, the positive value of Pbias for the JRA-55 product indicates it significantly overestimated the observed precipitation (by 11.4%). POD values ranged from 0.59 to 0.93 while FAR values ranged from 0.16 to 0.45. APHRODITE has the highest POD and the lowest FAR. The poor performance of CPCU

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

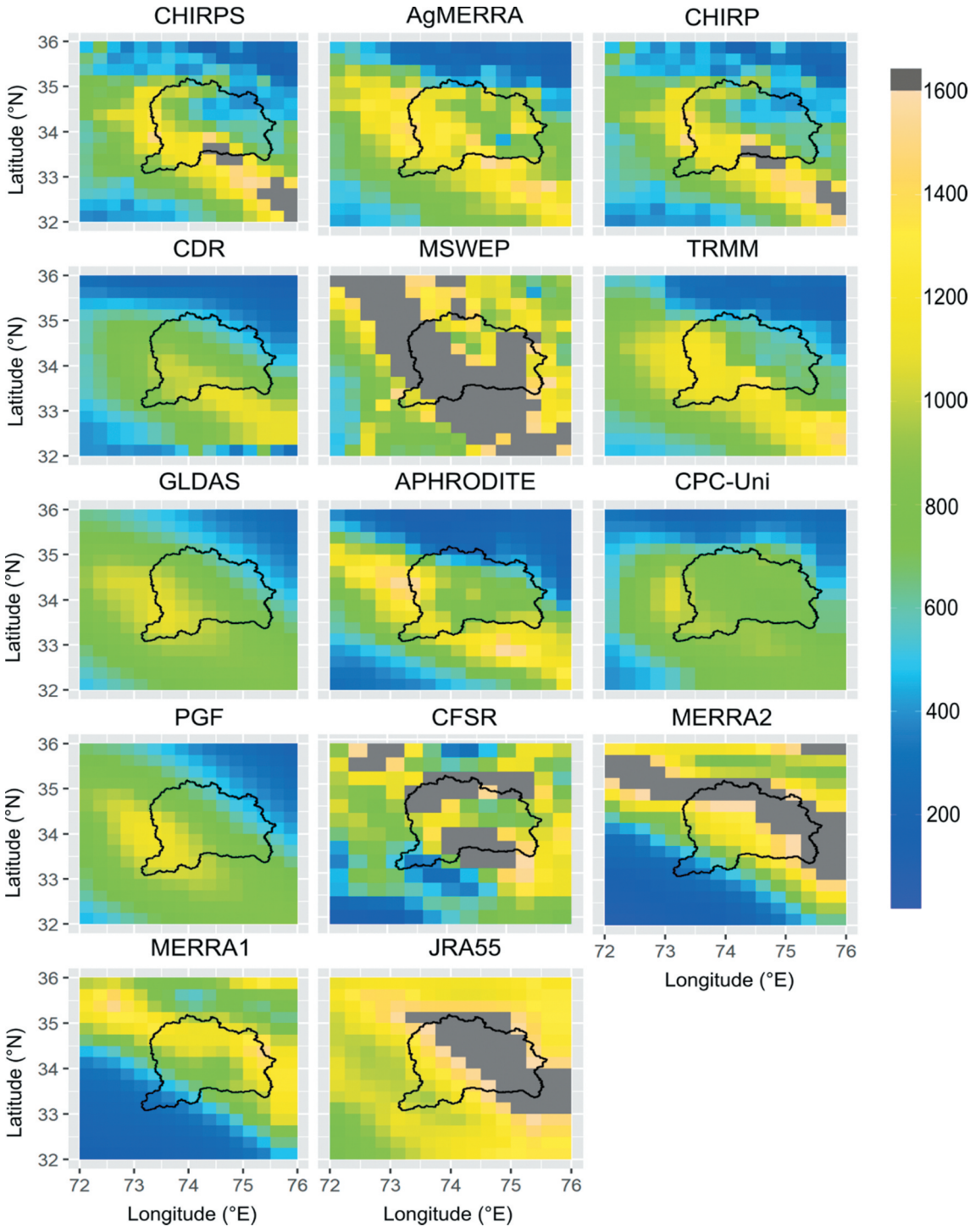

Figure 2. Spatial distribution of annual precipitation of SPP, GPP and RPP datasets. Values >1600 mm are assigned a grey colour to make the distribution more distinctive.

indicates that the number of raingauges used to construct this dataset was lower compared to APHRODITE (Rana *et al.* 2015). Among the seven SPP datasets, it is interesting to find that five products (CHIRPS, AgMERRA, CHIRP, MSWEP and CDR) have significantly higher CCs (>0.70) and lower RMSE and Pbias values. More interestingly, the performance of the CHIRPS dataset was found to be good compared to CHIRP, as ground station data (collected from different regional meteorological organizations) was blended during production of satellite output (Funk *et al.* 2015). In general, APHRODITE followed by AgMERRA, JRA-55, MSWEP and CHIRPS proved to be the most accurate precipitation datasets at a daily time scale. These findings are in agreement with those of similar studies (Gebrechorkos *et al.* 2018, Li *et al.* 2018, Iqbal *et al.* 2019, Tang *et al.* 2019).

The agreement of SPPs, GPPs and RPPs with observed data increases with decreasing temporal resolution, from daily to monthly and seasonal resolution. Appendix Table A1 provides

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

Table 3. Summary of statistical indices (CC, RMSE, Pbias) and categorical indices (POD, FAR, CSI) for the evaluation of SPPs, GPPs and RPPs on a daily time scale over the study area.

Precipitation products	Datasets	CC	RMSE (mm)	Pbias (%)	POD	FAR	CSI
SPPs	CHIRPS	0.73	1.63	−2.9	0.79	0.32	0.56
	AgMERRA	0.77	1.59	4.3	0.89	0.20	0.59
	CHIRP	0.70	1.76	−1.6	0.77	0.36	0.53
	CDR	0.71	1.55	−0.5	0.81	0.21	0.57
	MSWEP	0.75	2.65	1.2	0.72	0.26	0.61
	TRMM	0.68	1.84	5.3	0.70	0.35	0.53
	GLDAS	0.69	1.57	−0.7	0.76	0.25	0.56
GPPs	APHRODITE	0.79	1.51	−2.3	0.93	0.16	0.63
	CPCU	0.61	2.12	−20.1	0.65	0.24	0.47
RPPs	PGF	0.64	2.19	−3.3	0.75	0.23	0.49
	CFSR	0.59	2.49	1.1	0.75	0.29	0.44
	MERRA-2	0.52	2.86	8.1	0.59	0.45	0.39
	MERRA-1	0.62	3.15	−17.1	0.65	0.37	0.42
	JRA-55	0.76	1.61	11.4	0.90	0.18	0.61

the same statistical indices (CC, RMSE and Pbias) that are used in the evaluation of daily rainfall. Not surprisingly, all the indices improved at the monthly scale. Figure 3 presents a comparison of mean monthly-observed rainfall with the 14 precipitation products at the same spatial resolution of 0.25° over the JRB.

It can be seen that all of the datasets were able to capture the pattern of observed mean monthly precipitation and showed two distinct peaks of precipitation, one in March (produced due to western disturbance) and the other in July (produced due to summer monsoon) in the study area. In general, of the 14 precipitation datasets, 10 underestimate the peak in March (with different magnitudes). MSWEP and MERRA-2 exhibited overestimation in all months compared to observed data (about 80% in peak precipitation months). In addition, most of the SPPs overestimate the peak of precipitation in July (summer monsoon). Overall, SPPs and GPPs showed higher agreement with observations in producing monthly precipitation compared to RPPs. The agreement of all the rainfall products increases from daily to monthly resolution; this is consistent with other studies (Gebrechorkos *et al.* 2018, Shayeghi *et al.* 2020). As most of the SPPs were corrected by using the monthly correction factors from station data, the inclusion of monthly station data can be assumed to improve the performance of SPPs compared to other rainfall products (RPPs).

At the seasonal scale (Table 4), SPPs and GPPs showed low precipitation amounts except MSWEP, whereas RPPs (CFSR, MERRA-2 and JRA-55) produced high amounts. CPCU is characterized by the lowest estimates in all seasons, which led to an underestimation of annual precipitation. On the other hand, MERRA-2 produced the highest estimates, with a deviation of 138 mm and 388 mm in autumn and summer, respectively. More interestingly, SPPs (CHIRPS, CDR and AgMERRA) showed autumn precipitation amounts exactly the same as those observed. Our results are consistent with Hu *et al.* (2016), who reported significant overestimation of precipitation by reanalysis datasets (MERRA-2 and CFSR) and underestimation by the TRMM precipitation product in Central Asia at high elevations. MERRA-2 products give higher rainfall estimates than their counterparts by MERRA-1 in the JRB.

4.2 Calibration and validation of SWAT model

The SWAT model performance evaluation results are shown in Fig. 4. The model was calibrated and validated using the observed data of seven gauging stations in the JRB. The daily and monthly flow data for the period 1995–2007 (including a three-year warm-up period) were used during calibration (1998–2003) and validation (2004–2007). The results at Azad Pattan (where about 80% area of the basin contributes) revealed that the simulated flow well captured the observed flow pattern at daily and monthly scale. However, peak (monsoon) and very low (dry winter) flows are not well captured

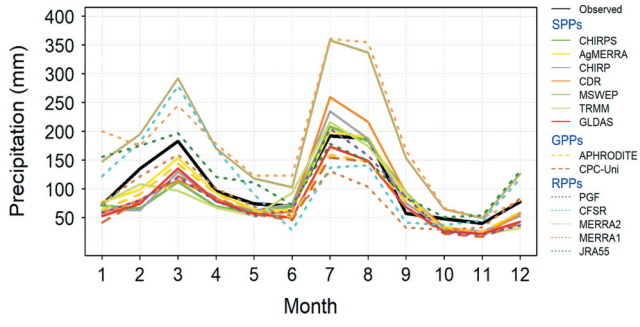

Figure 3. Comparison of observed mean monthly precipitation (basin-averaged) with 14 precipitation products (SPPs, GPPs, RPPs) over the Jhelum River basin for the period 1998–2007.

Table 4. Mean seasonal precipitation (in mm) estimated by three different products.

Precipitation products	Datasets	Winter	Spring	Summer	Autumn
	Observed	284	353	449	146
SPPs	CHIRPS	195	241	461	148
	AgMERRA	233	310	453	144
	CHIRP	179	270	496	133
	CDR	196	255	535	144
	MSWEP	468	581	797	268
	TRMM	214	218	483	142
	GLDAS	170	270	372	116
GPPs	APHRODITE	214	297	359	138
	CPCU	161	263	347	106
RPPs	PGF	176	251	426	131
	CFSR	421	539	305	123
	MERRA-2	455	546	837	284
	MERRA-1	280	327	288	95
	JRA-55	472	547	369	177

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

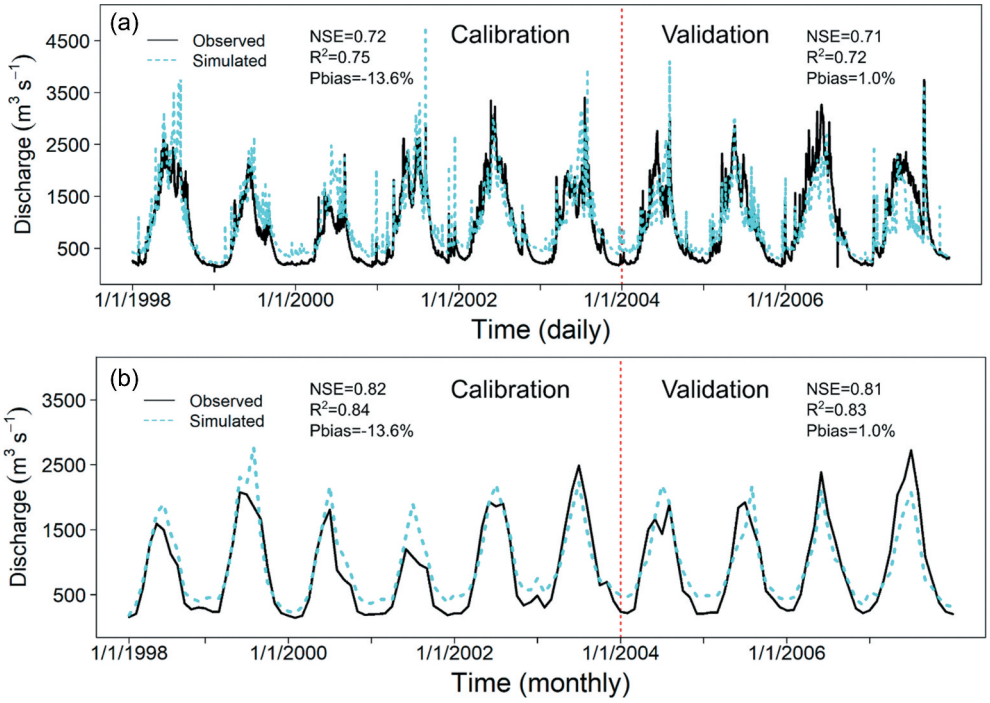

Figure 4. SWAT calibration and validation with observed meteorological data at (a) daily and (b) monthly temporal resolution at Azad Pattan (station) in the Jhelum River basin.

during the simulation period. Not surprisingly, this may have occurred due to the uneven distribution of meteorological stations across the basin, particularly at higher elevations. At the daily and monthly scale, simulated values show good agreement with observed streamflow, resulting in good NSE (0.72 and 0.82), Pbias (−13.6%) and R^2 (0.75 and 0.84), respectively, during calibration.

Similarly, NSE values of 0.71 and 0.81, Pbias of 1% and R^2 of 0.72 and 0.83 were obtained at daily and monthly temporal resolution, respectively, for the validation period. In addition, indices indicate that calibration results are slightly better than validation results. Overall, performance of the SWAT model falls within the "good category" according to Moriasi *et al.* (2007). Detailed information about calibration and validation at JRB can be found in Saddique *et al.* (2019b).

4.3 Hydrological evaluation of precipitation products

The discharge simulations from different precipitation products (SPPs, GPPs and RPPs) are calculated using a calibrated SWAT model over the study area. Figure 5 depicts the scatter plots of daily observed versus simulated discharge during 1998–2007. The statistical evaluation of SWAT driven by the results of 14 precipitation products is also presented in Fig. 5. It can be seen that APHRODITE (GPPs) displayed the best performance in simulating the flow, with higher NSE (0.68) and R^2 (0.71) and acceptable Pbias (−13%), followed by JRA-55 (NSE = 0.62, R^2 = 0.64, Pbias = −5%) and MSWEP (NSE = 0.59, R^2 = 0.61, Pbias = 6%). Other precipitation products such as CHIRPS, AgMERRA and CHIRP (SPPs) gave satisfactory performance as their NSE, Pbias and R^2 values lie within the range suggested by Moriasi *et al.* (2007). Among all the satellite datasets, TRMM's performance was poor with the lowest NSE (0.43). Generally, the results suggest that the performance of SPPs in simulating the daily streamflow was much better than that of GPPs and RPPs. Overall, out of 14 products, 11 showed underestimation in simulating daily flows as is clear from their negative Pbias values. Additionally, peak flows are overestimated by most of the precipitation product-driven simulations, as shown in Fig. 5.

To further validate the performance of the 14 precipitation products, flow duration curves at daily temporal resolution are presented in Fig. 6. The comparison shows that all the product-driven simulations overestimate the high (Q5) and low (Q95) flows while underestimating the median flows (Q50). Precipitation products are not able to capture the extreme events efficiently. These results are in line with a previous study by Try *et al.* (2020) who found that the SPPs, GPPs and RPPs had overestimated the high and low flows. Additionally, it can be seen that MERRA-2 overestimated all the flows compared to observations in the JRB. The performances of the precipitation products vary significantly over topographically complex regions and are complicated by significant elevation change, seasonality, and snow cover (Derin and Yilmaz 2014). Another study, by Stampoulis and Anagnostou (2012), found that over mountainous regions in Europe, SPP products significantly overestimate precipitation in the cold season because of snow/cold surface contamination.

In line with the comparison of daily simulations over the JRB, the evaluation was extended from daily to monthly temporal resolution using the same statistical indices as those defined by Moriasi *et al.* (2007). Figures 7 and 8 illustrate the monthly simulated results with observations. For this, daily streamflow simulations driven by the 14 precipitation products (SPPs, GPPs and RPPs) were aggregated to monthly

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

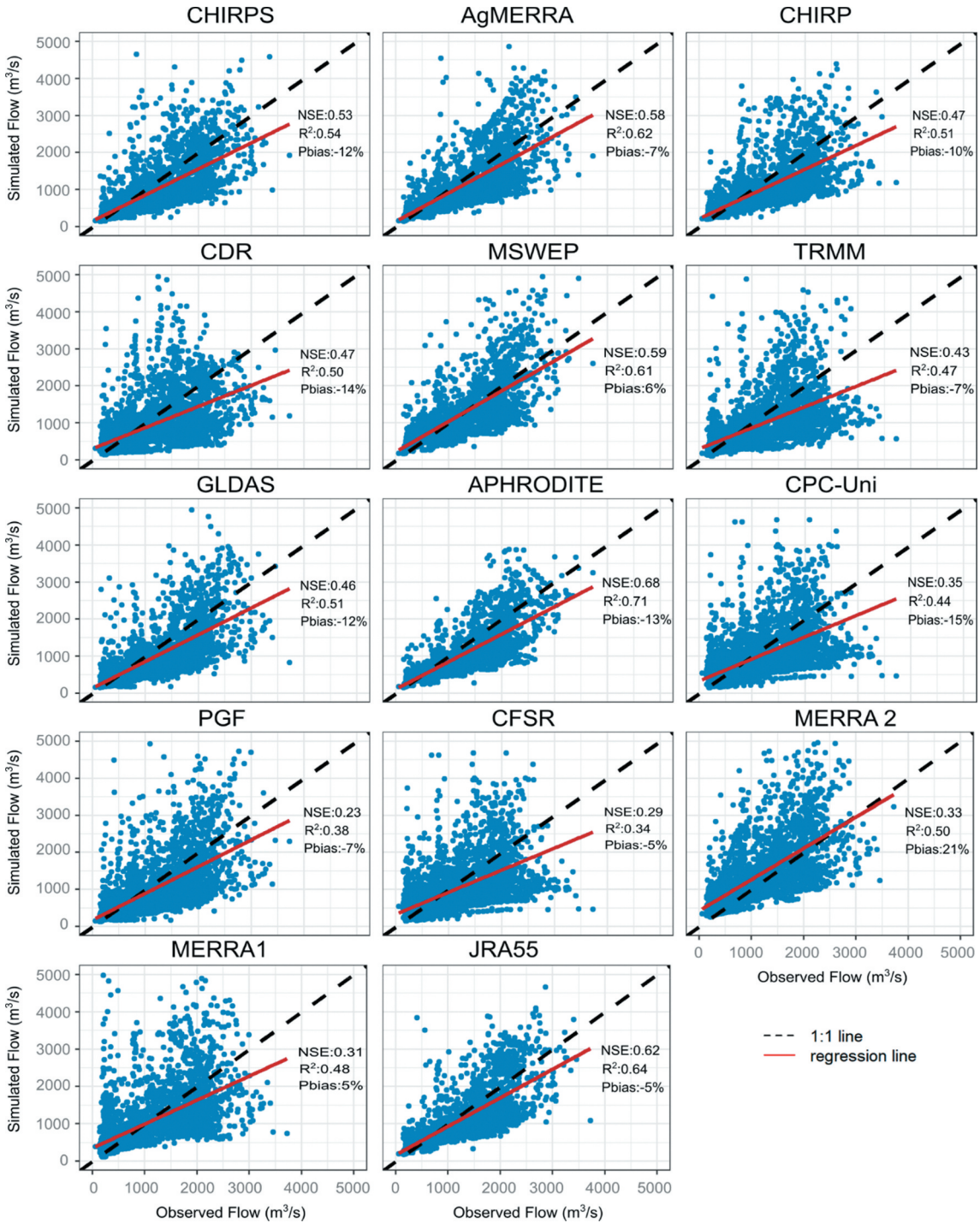

Figure 5. Scatter plot and evaluation indices (NSE, R^2, Pbias) of daily simulated flow driven by 14 precipitation products and observations.

resolution. At the monthly scale, the agreement of the product-derived simulations with the observed simulation remarkably improved during the simulation period.

In addition to increasing the CC, Pbias and RMSE values decreased at the monthly scale. Similar to its performance in the daily evaluation, APHRODITE displayed a remarkable performance at the monthly resolution, with NSE of 0.80, R^2 of 0.81 and Pbias of −13.8%, followed by JRA5, MSWEP, AgMERRA and CDR with similar NSE (0.77) and R^2 (0.88) but different Pbias values (−5.5%, 6%, −7.3% and −7.6%, respectively). Interestingly, the performance evaluation results of four SPPs (CHIRPS, CHIRP, TRMM, GLDAS), one GPPs (APHRODITE) and two RPPs (PGF, CFSR) showed agreement with Moriasi *et al.*'s (2007) definition of satisfactory results (NSE > 0.5, R^2 > 0.5 and Pbias < ±25%). In addition, the performance of MERRA-2 and MERRA-1 was found to be unsatisfactory in simulating the monthly discharge. Overall, all of the satellite product-driven simulations displayed good performance.

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

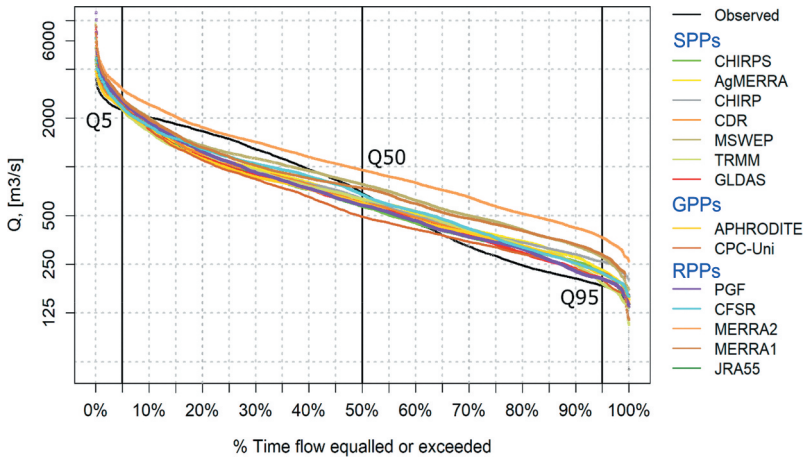

Figure 6. Flow duration curves (FDCs) of observed and simulated flows driven by 14 precipitation products (SPPs, GPPs, RPPs) on a daily time scale during the period 1998–2007. Q5: high flows, Q50: median flows, Q95: low flows.

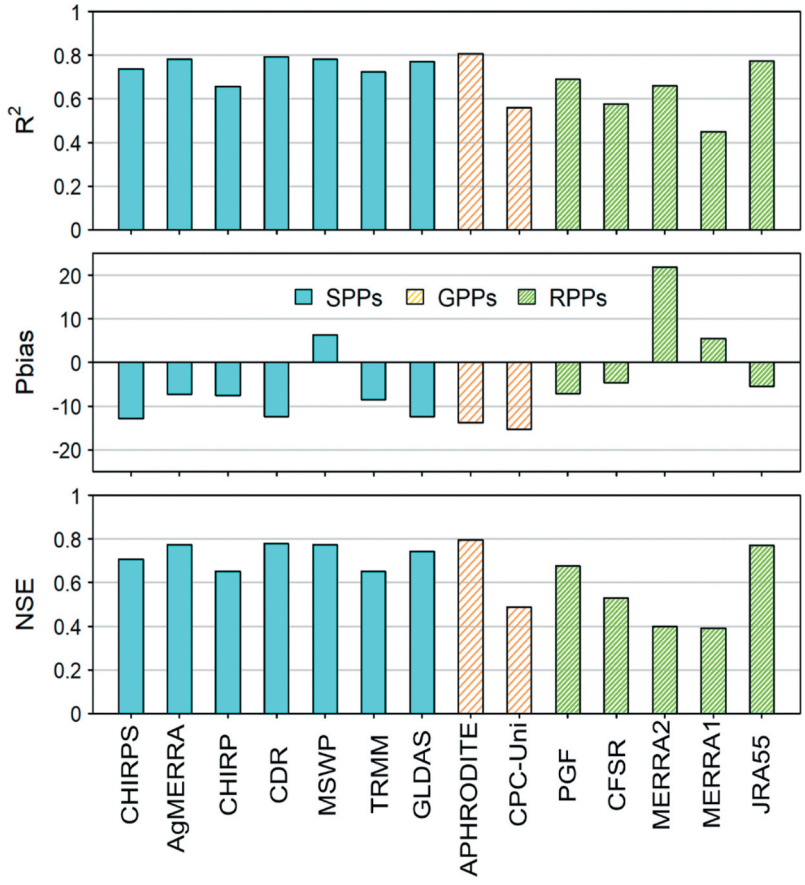

Figure 7. Monthly flow performance evaluation indices of satellite-based, gauge-based and reanalysis precipitation products over the Jhelum River basin.

4.4 Comparison with previous studies

Selection of the best SPPs or gridded precipitation products (GPPs and RPPs) is always challenging, particularly over regions with complex terrain. Therefore, the use of these products to simulate the discharge in catchments has always attracted scholarly interest, especially in areas with low density of raingauge stations (Tong *et al.* 2014, Nhi *et al.* 2018). We compared statistical and hydrological applications of SPPs, GPPs and RPPs with previous studies that applied these products over different regions of the world. Our results showed agreement with Thom *et al.* (2017), who used three satellite precipitation products (e.g. APHRODITE, PERSIAN-CDR, TRMM) in Vietnam and found that APHRODITE can be used as an alternative to observed rainfall for simulating daily flow. Faiz *et al.* (2020) used APHRODITE, CHIRPS, CPC, TRMM and PERSIAN-CDR in a hydrological model for a high-elevation basin and reported that the APHRODITE and CHIRPS displayed good performance in simulating the

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

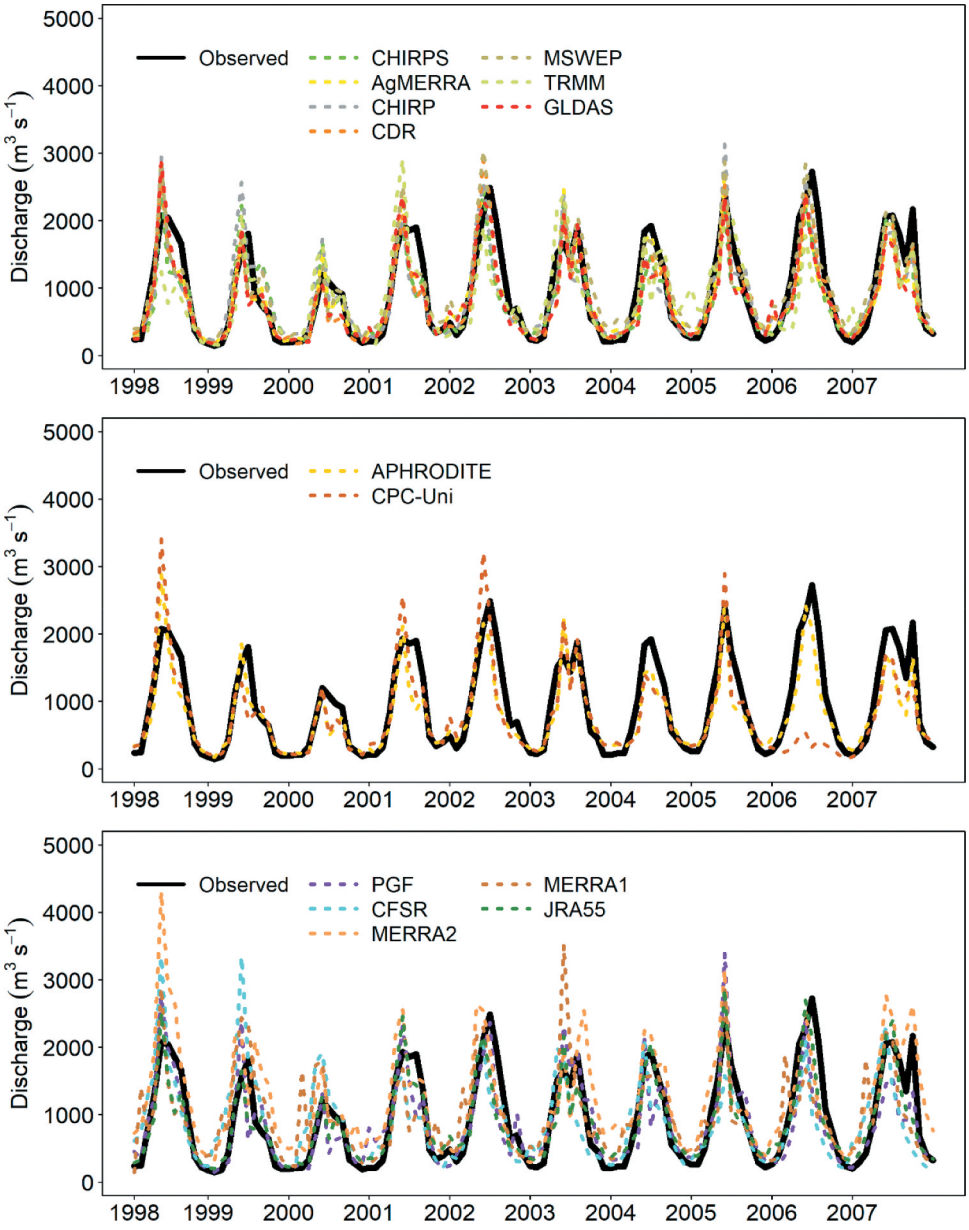

Figure 8. Comparison of simulated monthly flow driven by (a) SPPs, (b) GPPs and (c) RPPs with observed flow in the Jhelum River basin.

discharge. According to Wang *et al.* (2018) and Zhu *et al.* (2016), TRMM-driven simulations exhibited the lowest NSE and highest Pbias. Lakew *et al.* (2020) evaluated the hydrological performance of multiple precipitation products over the Nile basin, Ethiopia, and reported that MSWEP can be used alternatively to observations. Wang *et al.* (2020) analysed hydrological simulations in the Xihe River basin and found that satellite products performed better than reanalysis products.

5 Conclusion

Availability of high-quality data is a major problem for climate and hydrological studies, especially in a developing country like Pakistan where gauge network is typically limited and unevenly distributed. Recently, a number of freely available climate datasets have been developed for better estimation of precipitation and other atmospheric variables. This study investigates the applicability of 14 precipitation datasets with high spatiotemporal resolution over a transboundary high-altitude catchment (JRB) using both statistical and hydrological evaluations during the period 1998–2007.

At a daily time scale, APHRODITE (GPP), AgMERRA (SPP) and JRA-55 (RPP) showed remarkable results compared with observed rainfall data, with the highest CC and lowest RMSE and Pbias. Among the seven SPPs, five (CHIRPS, CHIRP, AgMERRA, CDR and MSWEP) have significantly high CC values (>0.7). As the temporal resolution becomes coarser (monthly), the agreement of precipitation products with observed data significantly increases but most of the products showed dry biases (underestimation) in the mean monthly rainfall amount.

We have found that observed rainfall well reproduced the observed discharge at both daily and monthly scales at Azad Pattan gauge (where about 80% area of the basin contributes) in

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

the JRB, indicating the SWAT hydrological model has the ability to capture hydrological processes and can be used with confidence for the evaluation of precipitation products. APHRODITE (GPP), JRA-55 (RPP) and MSWEP (SPP) demonstrated the best performance in simulating the daily flow, with NSE values of 0.68, 0.62 and 0.59, respectively, for the period 1998–2007. Interestingly, in simulating the monthly flow all the satellite-based products showed good performance, with NSE > 0.6. However, peak and low flow are overestimated by all of the product-driven simulations during the simulation period. Overall, the performance of SPPs was found to be good compared to the GPPs and RPPs. In conclusion, it can be said that APHRODITE, JRA-55, MSWEP, CHIRPS and AgMERRA have significant potential in hydrological studies where the observed raingauge network is limited and unevenly distributed.

Acknowledgements

The lead author thanks the Higher Education Commission of Pakistan (HEC) Pakistan and the German Academic Exchange Service (DAAD) for providing financial support for PhD studies. The authors are thankful to the Pakistan Meteorological Department (PMD), Water and Power Development Authority (WAPDA) and Indian Meteorological Department (IMD) for providing discharge and weather data used in this research.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abbaspour, K.C., 2015. *SWAT-CUP: SWAT Calibration and Uncertainty Programs- a user manual*. Duebendorf, Switzerland: Department of Systems Analysis, Intergrated Assessment and Modelling (SIAM), EAWAG. Swiss Federal Institute of Aqualtic Science and Technology, 100. doi:10.1007/s00402-009-1032-4

Ahmed, E., *et al.*, 2020. Hydrologic assessment of TRMM and GPM-based precipitation products in transboundary river catchment (Chenab River, Pakistan). *Water*, 12 (7), 1902. doi:10.3390/w12071902

AL-Falahi, A.H.A., *et al.*, 2020. Evaluation the performance of several gridded precipitation products over the highland region of Yemen for water resources management. *Remote Sensing*, 12 (18), 2984. doi:10.3390/rs12182984

Arnold, J.G., *et al.*, 1998. Large-area hydrologic modeling and assessment: part I. Model development. *Journal of the American Water Resources Association*, 34 (1), 73–89. doi:10.1111/j.1752-1688.1998.tb05961.x

Ashouri, H., *et al.*, 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. doi:10.1175/BAMS-D-13-00068.1

Beck, H.E., *et al.*, 2017a. 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. doi:10.5194/hess-21-589-2017

Beck, H.E., *et al.*, 2017b. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, 21 (1), 6201–6217. doi:10.5194/hess-21-6201-2017

Beck, H.E., *et al.*, 2019. MSWep v2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, 100 (3), 473–500. doi:10.1175/BAMS-D-17-0138.1

Birylo, M., Rzepecka, Z., and Nastula, J., 2018. Assessment of the water budget from GLDAS model. *In*: *2018 Baltic Geodetic Congress (BGC Geomatics)*, 86–90. doi:10.1109/BGC-Geomatics.2018.00022

Bui, H., Ishidaira, T., and Shaowei, H., 2019. Evaluation of the use of global satellite–gauge and satellite-only precipitation products in stream fow simulations. *Applied Water Science*, 9, 53. doi:10.1007/s13201-019-0931-y

Chen, M., *et al.*, 2008. Assessing objective techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research Atmospheres*, 113 (4), 1–13. doi:10.1029/2007JD009132

Dahri, Z.H., *et al.*, 2016. An appraisal of precipitation distribution in the high-altitude catchments of the Indus basin. *Science of the Total Environment*, 548–549, 289–306. doi:10.1016/j.scitotenv.2016.01.001

Dee, D.P., *et al.*, 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597. doi:10.1002/qj.828

Dembélé, M., *et al.*, 2020. Improving the predictive skill of a distributed hydrological model by calibration on spatial patterns with multiple satellite data sets. *Water Resources Research*, 56, e2019WR026085. doi:10.1029/2019WR026085

Derin, Y. and Yilmaz, K.Y., 2014. Evaluation of multiple satellite-based precipitation products over complex topography. *Journal of Hydrometeorology*, 1498-1516. doi:10.1175/JHM-D-13-0191.1

Ebita, A., *et al.*, 2011. The Japanese55-year reanalysis "JRA-55": an interim report. *SOLA*, 7, 149–152. doi:10.2151/sola.2011-038

Faisal, N. and Gafar, A., 2012. Development of Pakistan's new area weighted rainfall using Thiessen polygon method. *Pakistan Journal of Meteorology*, 9 (17), 107–116.

Faiz, M.A., *et al.*, 2020. Comprehensive evaluation of 0.25° precipitation datasets combined with MOD10A2 snow cover data in the ice-dominated river basins of Pakistan. *Atmospheric Research*, 231. March 2019. doi:10.1016/j.atmosres.2019.104653

Funk, C., *et al.*, 2015. The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. doi:10.1038/sdata.2015.66

Gebrechorkos, S.H., Hülsmann, S., and Bernhofer, C., 2018. Evaluation of multiple climate data sources for managing environmental resources in East Africa. *Hydrology and Earth System Sciences*, 22 (8), 4547–4564. doi:10.5194/hess-22-4547-2018

Gelaro, R., *et al.*, 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30 (14), 5419–5454. doi:10.1175/JCLI-D-16-0758.1

Hamada, A. and Takayabu, Y.N., 2016. Improvements in detection of light precipitation with the Global Precipitation Measurement Dual-Frequency Precipitation Radar (GPM DPR). *Journal of Atmospheric and Oceanic Technology*, 33, 653–667. doi:10.1175/JTECH-D-15-0097.1

Hu, Z., *et al.*, 2016. Evaluation of reanalysis, spatially interpolated and satellite remotely sensed precipitation data sets in Central Asia. *Journal of Geophysical Research: Atmospheres*, 121 (10), 5648–5663. doi:10.1002/2016JD024781

Huffman, G.J., *et al.*, 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8 (1), 38–55. doi:10.1175/JHM560.1

Immerzeel, W.W., Pellicciotti, F., and Shrestha, A.B., 2012. Glaciers as a proxy to quantify the spatial distribution of precipitation in the Hunza basin. *Mountain Research and Development*, 32 (1), 30–38. doi:10.1659/mrd-journal-d-11-00097.1

Iqbal, Z., *et al.*, 2019. Spatial distribution of the trends in precipitation and precipitation extremes in the sub-Himalayan region of Pakistan. *Theoretical and Applied Climatology*, 137 (3–4), 2755–2769. doi:10.1007/s00704-019-02773-4

Khan, A.J., Koch, M., and Chinchilla, K.M., 2018. Evaluation of gridded multi-satellite precipitation estimation (TRMM-3B42-V7) performance in the Upper Indus Basin (UIB). *Climate*, 6, 3. doi:10.3390/cli6030076

Lakew, H.B., Moges, S.A., and Asfaw D.H, 2020. Hydological performance evaluation of multiple satellite precipitation products in the upper Nile basin, Ethopia. *Journal of Hydrology: Regional Studies*, 27, 100664. doi:10.1016/j.ejrh.2020.100664

Lauri, H., Räsänen, T.A., and Kummu, M., 2014. Using reanalysis and remotely sensed temperature and precipitation data for hydrological modeling in monsoon climate: Mekong River case study. *Journal of Hydrometeorology*, 15 (4), 1532–1545. doi:10.1175/jhm-d-13-084.1

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

Li, D., *et al*. 2018a. Adequacy of TRMM satellite rainfall data in driving the SWAT modeling of Tiaoxi catchment (Taihu Lake basin, China). *Journal of Hydrology*, 556, 1139–1152. January 2017. doi:10.1016/j.jhydrol.2017.01.006

Li, H., Haugen, J.E., and Xu, C.Y., 2018b. Precipitation pattern in the Western Himalayas revealed by four datasets. *Hydrology and Earth System Sciences*, 22 (10), 5097–5110. doi:10.5194/hess-22-5097-2018

Liu, X., *et al*., 2017. Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau. *Hydrology and Earth System Sciences*, 21 (1), 169–181. doi:10.5194/hess-21-169-2017

Ménégoz, M., Gallée, H., and Jacobi, H.W., 2013. Precipitation and snow cover in the Himalaya: from reanalysis to regional climate simulations. *Hydrology and Earth System Sciences*, 17 (10), 3921–3936. doi:10.5194/hess-17-3921-2013

Moazami, S., *et al*., 2013. Comparison of PERSIANN and V7 TRMM Multisatellite Precipitation Analysis (TMPA) products with rain gauge data over Iran. *International Journal of Remote Sensing*, 34, 8156–8171. doi:10.1080/01431161.2013.833360

Moriasi, D.N., *et al*., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50, 885–900. doi:10.13031/2013.23153

Neitsch, S.L., *et al*., 2005. *Soil and water assessment tool – theoretical documentation, version 2005*. Temple, TX.

Nhi, P.T.T., Khoi, D.N., and Hoan, N.X., 2018. Evaluation of five gridded rainfall datasets in simulating streamflow in the upper Dong Nai River basin, Vietnam. *International Journal of Digital Earth*, 12 (3), 311–327. doi:10.1080/17538947.2018.1426647

Pang, H., *et al*., 2014. Influence of regional precipitation patterns on stable isotopes in ice cores from the central Himalayas. *Cryosphere*, 8 (1), 289–301. doi:10.5194/tc-8-289-2014

Rana, S., Mcgregor, J., and Renwick, J., 2015. Precipitation seasonality over the Indian subcontinent: an evaluation of gauge, reanalyses, and satellite retrievals. *Journal of Hydrometeorology*, 16 (2), 631–651. doi:10.1175/JHM-D-14-0106.1

Rienecker, M.M., *et al.*, 2011. MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, 24 (2011), 3624–3648. doi:10.1175/JCLI-D-11-00015.1

Rodell, M., *et al*., 2004. The global land data assimilation system. *Bulletin of the American Meteorological Society*. doi:10.1175/BAMS-85-3-381

Ruane, A.C., Goldberg, R., and Chryssanthacopoulos, J., 2015. Climate forcing datasets for agricultural modeling: merged products for gap-filling and historical climate series estimation. *Agricultural and Forest Meteorology*, 200, 233–248. doi:10.1016/j.agrformet.2014.09.016

Saddique, N., *et al*., 2019a. Downscaling of CMIP5 models output by using statistical models in a data scarce mountain environment (Mangla Dam Watershed), Northern Pakistan. *Asia-Pacific Journal of Atmospheric Sciences*, 55, 719–735. doi:10.1007/s13143-019-00111-2

Saddique, N., Mahmood, T., and Bernhofer, C., 2020. Quantifying the impacts of land use/land cover change on the water balance in the afforested river basin, Pakistan. *Environmental Earth Sciences*, 79 (19), 448. doi:10.1007/s12665-020-09206-w

Saddique, N., Usman, M., and Bernhofer, C., 2019b. Simulating the impact of climate change on the hydrological regimes of a sparsely gauged mountainous basin, Northern Pakistan. *Water (Switzerland)*, 11, 10. doi:10.3390/w11102141

Saha, S., *et al*., 2010. The NCEP climate forecast system reanalysis. *American Meteorological Society*, 91 (8). doi:10.1175/2010BAMS3001.1

Saha, S., *et al*., 2014. The NCEP climate forecast system version 2. *Journal of Climate*, 27 (6), 2185–2208. doi:10.1175/JCLI-D-12-00823.134

Santhi, C., *et al*., 2001. Validation of the SWAT model on a large river basin with point and nonpoint sources. *Journal of the American Water Resources Association*, 37 (5), 1169–1188. doi:10.1111/j.1752-1688.2001.tb03630

Senent-Aparicio, J., *et al*., 2018. Using multiple monthly water balance models to evaluate gridded precipitation products over Peninsular Spain. *Remote Sensing*, 10 (6), 922. doi:10.3390/rs10060922

Shayeghi, A., Azizian, A., and Brocca, L., 2020. Reliability of reanalysis and remotely sensed precipitation products for hydrological simulation over the Sefidrood River basin, Iran. *Hydrological Sciences Journal*, 65 (2), 296–310. doi:10.1080/02626667.2019.1691217

Sheffield, J., Goteti, G., and Wood, E.F., 2005. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, 19 (13), 3088–3111. doi:10.1175/JCLI3790.1

Sidike, A., *et al*., 2016. Investigating alternative climate data sources for hydrological simulations in the upstream of the Amu Darya River. *Water (Switzerland)*, 8, 10. doi:10.3390/w8100441

Stampoulis, D. and Anagnostou, E.N., 2012. Evaluation of global satellite rainfall products over continental Europe. *Journal of Hydrometeorology*, 13 (2), 588–603. doi:10.1175/JHM-D-11-086.1

Talchabhadel, R., *et al*., 2020. 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. doi:10.1016/j.ejrh.2020.100768

Tan, M.L., *et al*., 2015. Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia. *Remote Sensing*, 7 (2), 1504–1528. doi:10.3390/rs70201504

Tang, X., *et al*., 2019a. Assessing the uncertainties of four precipitation products for SWAT modeling in Mekong River basin. *Remote Sensing*, 11 (3), 1–24. doi:10.3390/rs11030304

Tang, X., *et al*., 2019b. Evaluating suitability of multiple precipitation products for the Lancang River basin. *Chinese Geographical Science*, 29 (1), 37–57. doi:10.1007/s11769-019-1015-5

Thom, V.T., Khoi, D.N., and Linh, D.Q., 2017. Using gridded rainfall products in simulating streamflow in a tropical catchment - a case study of the Srepok River catchment, Vietnam. *Journal of Hydrology and Hydromechanics*, 65 (1), 18–25. doi:10.1515/johh-2016-0047

Tong, K., *et al*., 2014. Evaluation of satellite precipitation retrievals and their potential utilities in hydrologic modeling over the Tibetan Plateau. *Journal of Hydrology*, 519, 423–437. doi:10.1016/j.jhydrol.2014.07.044

Try, S., *et al*., 2020. Comparison of gridded precipitation datasets for rainfall-runoff and inundation modeling in the Mekong River basin. *PLoS ONE*, 15 (1), e0226814. doi:10.1371/journal.pone.0226814

Wang, Z., *et al*., 2018. Hydrologic assessment of the TMPA 3B42-V7 product in a typical alpine and gorge region: the Lancang River basin, China. *Hydrology Research*, 49 (6), 2002–2015. doi:10.2166/nh.2018.024

Wang, N., *et al*., 2020. Evaluating satellite-based and reanalysis precipitation datasets with gauge-observed data and hydrological modeling in the Xihe River basin, China. *Atmospheric Research*, 231, 2020. April. doi:10.1016/j.atmosres.2019.104746

Wilks, D.S., 2011. *Statistical methods in the atmospheric sciences*. 3rd ed., International Geophysics Series. Amsterdam, The Netherlands: Elsevier.

WMO, 2008. *Guide to hydrological practices, volume I, hydrology – from measurement to hydrological information* Geneva, Switzerland.

Wong, J.S., *et al*., 2017. Inter-comparison of daily precipitation products for large-scale hydro-climatic applications over Canada. *Hydrology and Earth System Sciences*, 21 (4), 2163–2185. doi:10.5194/hess-21-2163-2017

Wu, Z., *et al*., 2018. Hydrologic evaluation of multi-source satellite precipitation products for the Upper Huaihe River basin, China. *Remote Sensing*, 10 (6), 840. doi:10.3390/rs10060840

Yang, D., *et al*., 1998. Adjustment of daily precipitation data at 10 climate stations in Alaska: application of World Meteorological Organization intercomparison results. *Water Resources Research*, 34, 241–256. doi:10.1029/97WR02681

Yatagai, A., *et al*., 2012. Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, 93 (9), 1401–1415. doi:10.1175/BAMS-D-11-00122.1

Zandler, H., Haag, I., and Samimi, C., 2019. Evaluation needs and temporal performance diferences of gridded precipitation products in peripheral mountain regions. *Scientific Reports*, 9, 15118. doi:10.1038/s41598-019-5166

Zhu, Q., *et al*., 2016. Evaluation and hydrological application of precipitation estimates derived from PERSIANN-CDR, TRMM 3B42V7, and NCEP-CFSR over humid regions in China. *Hydrological Processes*, 30 (17), 3061–3083. doi:10.1002/hyp.10846

Final edited form was published in "Hydrological Sciences Journal" 67 (3), S. 436-450. ISSN: 2150-3435
https://doi.org/10.1080/02626667.2021.2022152

Appendix

Table A1. Summary of statistical indices (CC, RMSE, Pbias) for the evaluation of SPPs, GPPs and RPPs on a monthly scale over the study area.

Precipitation products	Datasets	CC	RMSE (mm)	Pbias (%)
SPPs	CHIRPS	0.87	37.4	−2.9
	AgMERRA	0.91	29.03	4.3
	CHIRP	0.88	31.45	−1.6
	CDR	0.89	30.67	−0.5
	MSWEP	0.90	30.23	1.2
	TRMM	0.82	44.0	5.3
	GLDAS	0.85	40.0	−0.7
GPPs	APHRODITE	0.94	27.51	−2.3
	CPCU	0.75	53.76	−20.1
RPPs	PGF	0.82	54.33	−3.3
	CFSR	0.78	28.43	1.1
	MERRA-2	0.72	48.25	8.1
	MERRA-1	0.54	74.23	−17.1
	JRA-55	0.89	36.57	11.4