

Spatio-temporal Evaluation of Gridded Precipitation and Evapotranspiration Products Over Balochistan Province in Pakistan

Aman Jan (✉ aman.wrm@luawms.edu.pk)

Lasbela University of Agriculture Water and Marine Sciences <https://orcid.org/0000-0002-5181-4341>

Oscar Manuel Baez-Villanueva

Lars Ribbe

Research Article

Keywords: Balochistan, precipitation products, evaporation products, Budyko framework

Posted Date: January 5th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1158419/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Title

Spatio-temporal evaluation of gridded precipitation and evapotranspiration products over Balochistan province in Pakistan

Authors

1. Aman Jan (**Corresponding author**)
Department of Water Resources Management, Lasbela University of Agriculture, Water and Marine Sciences, Uthal, Pakistan.
ORCID identifier; 0000-0002-5181-4341
aman.wrm@luawms.edu.pk; aman.akram1991@gmail.com
2. Oscar M. Baez-Villanueva
Institute for Technology and Resources Management in the Tropics and Subtropics, Technische Hochschule Köln
Faculty of Spatial Planning, TU Dortmund University, Dortmund, Germany
ORCID identifier; 0000-0002-2262-1698
obaevzil@gmail.com
3. Lars Ribbe
Institute for Technology and Resources Management in the Tropics and Subtropics, TH Köln
ORCID identifier; 0000-0003-3052-1782
lars.ribbe@th-koeln.de

Abstract

This study evaluated the performance of four state-of-the-art precipitation (P ; CHIRPSv2, MSWEPv2.2, PERSIANN-CDR, and ERA5) and three state-of-the-art evaporation (E_a ; GLEAMv3.3a, SSEBopv4.0, and ERA5) products in the data-scarce province of Balochistan, Pakistan. The P products were evaluated through a point-to-pixel comparison at the monthly, annual, and seasonal temporal scales using data from 17 gauges. The modified Kling-Gupta efficiency (KGE') and the root mean square error (RMSE) were used as statistical indices to evaluate the performance of the P products. We used the Budyko framework to evaluate the E_a component due to unavailability of ground-based E_a data. For this purpose, we calculated the multi-annual mean E_a — defined as E_{a_WB} — for 12 defined catchments for a period of 16 years (2003–2018). Subsequently, the estimated E_{a_WB} was used as the reference to evaluate the performance of the E_a products. The results of the study revealed that among the P products generally MSWEPv2.2 closely followed by ERA5 presented the best spatio-temporal KGE' performance over most of the stations and across the province. CHIRPSv2 and PERSIANN-CDR relatively showed poor KGE' performance. Whereas, for the E_a component, ERA5 generally outperformed the other two products — i.e., GLEAMv3.3a and SSEBopv4.0 — in most of the evaluated catchments except for those located over the north-eastern region where GLEAMv3.3a performed the best. The results of this study can be beneficial to shift towards a proactive water management approach that can serve as a basis to improve the decision-making process in the region.

Keywords: Balochistan, precipitation products, evaporation products, Budyko framework

1 Introduction

Management of water resources is one of the key global challenges mankind has to face in the 21st century (Mehta, 2000; Sheffield et al., 2018). The occurrence of weather extremes — such as floods and droughts — are increasing and their impacts are inevitable (Yang et al., 2018). Therefore, water has been recognized as one of the key components of the Sustainable Development Goals (SDGs) of the Agenda 2030 (UN General Assembly, 2015). The SDG 6 aims to “ensure availability and sustainable management of water and sanitation for all” (UN General Assembly, 2015, p. 20). Therefore, ensuring water availability and managing water resources will help to deal with other SDGs, for example, poverty alleviation (SDG1), global food security (SDG2), and global health and well-being (SDG3; UNDP, 2004). Hence, it is crucial to characterize the spatio-temporal patterns of the key components of the water cycle (Trezza, 2002).

Precipitation (P) is the largest water input, while the actual evaporation (E_a) — including transpiration from vegetation (see Miralles et al., 2020) — is the major output of the water cycle (Jaber, 2017). An accurate representation of the spatio-temporal variability of these components remains a global challenge (Oliveira et al., 2014). Until a few decades ago, ground-based meteorological data have been the only source to provide relatively accurate measurements of key components of the water cycle such as P and E_a (Oliveira et al., 2014). Even though these measurements are relatively accurate, they are often sparsely distributed and are subject to uncertainties due to the lack of infrastructural and human resources (Voss et al., 2013; Sheffield et al., 2018). Fortunately, since the availability of remote sensing data and reanalysis datasets, it is now relatively easy to account for these climatological variables with low latency even in data-scarce regions (Gao et al., 2010). The use of these datasets is an emerging and viable mean to monitor the water cycle components (Tang et al., 2009).

Many satellite-based rainfall estimates and as well as reanalysis datasets have recently been used in many studies to estimate P (e.g., Zambrano-Bigiarini et al., 2017; Nogueira et al., 2018; Beck et al., 2019a; Torres-Batllo et al., 2020), to evaluate extreme events (e.g., Katsanos et al., 2016; Gupta et al., 2020), and to characterize droughts (e.g., Satgé et al., 2017; Wu et al., 2019; Zambrano-Bigiarini & Baez-Villaneuva, 2019). These products include Artificial Neural Networks- Climate Data Record (PERSIANN-CDR; Ashouri et al., 2015), Tropical Rainfall Measuring Mission (TRMM; Huffman et al., 2007), Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS; Funk et al., 2015), Multi-Source Weighted-Ensemble Precipitation (MSWEP; Beck et al., 2017a; Beck et al., 2019b), and ECMWF Reanalysis 5th Generation (ERA5; Hersbach et al., 2018) and so forth. The recent technological improvement of sensors and the availability of different methods to merge various data sources have resulted in continuous spatial and temporal improvement of these products (Baez-Villanueva et al., 2020).

Similarly, satellite-based E_a products have been developed since the availability of remotely sensed thermal infrared data (Brown & Rosenberg, 1973). Recently, many approaches have been developed to estimate E_a by remote sensing data (e.g., Jaber, 2017; Brenner et al., 2018; Sharma & Tare, 2018). Since E_a cannot be measured directly through remote sensing, its often indirectly estimated through diverse methodologies (McNamara et al., 2021) such as the Breathing Earth System Simulator (BESS; Ryu et al., 2011), the Surface Energy Balance (SEB; Jin et al., 2013), the Penman–Monteith equation (Mu et al., 2007; Adnan et al., 2009), and the Priestley-Taylor equation (Miralles et al., 2011), among others. The products developed based on such methods include the

Moderate Resolution Imaging Spectroradiometer's (MODIS) Global Evapotranspiration Project (MOD16; Mu et al., 2011), the Global Land Evaporation Amsterdam Model (GLEAM; Miralles et al., 2011), the operational Simplified Surface Energy Balance (SSEBop; Senay et al., 2013), and ERA5 among others.

However, P and E_a estimates are always subject to errors and uncertainties; and therefore, they must be evaluated before any hydrological application (Oliveira et al., 2014; Zambrano-Bigiarini et al., 2017). For example, Duan et al. (2012) evaluated two TRMM products (3B42v6 and 3B43v6) over the Caspian Sea region and found that both overestimated low P events and underestimated high P events. Similarly, Maggioni et al. (2016) evaluated the accuracy of different satellite P products including TRMM 3B42v7, the Global Satellite Mapping of Precipitation (GSMaP), the Climate Prediction Center Morphing method (CMORPH), and the PERSIANN-Cloud Classification System (PERSIANN-CCS). The results of this study showed that topography, seasonality, and climatological patterns influenced the P products' performance. Long et al. (2014) studied uncertainties in two remote-sensing based E_a products, including MOD16 and Advanced Very-High-Resolution Radiometer (AVHRR), in varying climate and land cover conditions in the south-central United States. They found that there was a trade-off between the spatial resolution of the E_a products and the level of uncertainty. That is to say, the higher was the resolution of the product the greater was the level of uncertainty; and lower was the resolution of the product the less was the level of uncertainty. In a recent study, Hugo et al. (2019) evaluated the performance of satellite-based E_a products, including GLEAM, SSEBop, Atmospheric Land Exchange Inverse (ALEXI), MOD16, Surface Energy Balance System (SEBS), and CMRSET, in the complex environment of the Amazon. They found that the uncertainty of the products was very low and sometimes negligible at the monthly temporal scale, but as they evaluated the products at higher temporal scales (e.g., more than one month), the uncertainty of the products increased

It's, therefore, important to study the uncertainties embedded in these products when used to account for P and E_a ; especially, over the arid and semi-arid regions due to the reduced performance of these products over arid and semi-arid areas (Emam et al., 2015). The arid region of Balochistan province in Pakistan is not an exception but is considered as one of the most water-scarce and drought-prone regions in the country (Ahmed et al., 2017). It has been observed that on average the province experiences four years of droughts out of ten years (Anjum et al., 2012) exerting severe economic impacts on the livelihood of the people who majorly depend on agro-pastoral activities (Ahmed et al., 2015b). For example, the prolonged drought of 1998–2002 resulted in a 60–80% reduction in crop yield and caused around two million livestock death in the province (Sarwar, 2008). In addition, the projections for the upcoming decades suggest an increase in drought frequency and a greater P variability due to climate change (Ahmed et al., 2005), which will place unprecedented pressure on the already scarce water resources of the area.

There have been many studies in Balochistan Region to understand the water cycle components and different climatic variables employing a variety of approaches. For example, Ahmed et al. (2015a) developed a model based on multilayer perceptron neural networks to downscale P in Balochistan Region using reanalysis datasets over the study area and calibrated and validated the data using monthly ground-based meteorological stations. They found a good correlation between observed and downscaled time series of P during both the calibration and the verification periods. However, the downscaled product underestimated the P variance during both the calibration and verification periods. In a recent study, Iqbal and Athar (2018) evaluated the performance of

TRMM's Multi-satellite Precipitation Analysis (TMPA) using ground-based P data over the entire country of Pakistan. They found that TMPA underestimates P in high elevations of the country, whereas the results were relatively satisfactory in medium elevated regions and lowlands. However, they do not discuss the performance of TMPA over specifically Balochistan province. Furthermore, in a recent comparative study, Ahmed et al. (2019) evaluated the performance of gauged-based gridded P products against ground-based meteorological station data in the region. They evaluated four products including the Global Precipitation Climatology Centre (GPCC), Center for Climatic Research-University of Delaware (UDel), Asian Precipitation Highly Resolved Observational Data Integration towards Evaluation (APHRODITE), and Climatic Research Unit (CRU). Their study concluded that the performance of P products varied in different climatic regions. However, the agreement between observed P and GPCC was much better in all climatic regions under all the statistical indices used in the study.

Howbeit, no study have evaluated the key meteorological components of the water balance (P and E_a) using the state-of-the-art satellite-based and reanalysis datasets over the Balochistan Region. Thus, the general objective of this study is to assess the ability of four P products and three E_a products of representing the spatio-temporal patterns of these key water balance components over the data-scarce Balochistan Region with a methodology that can be replicated in other data-scarce settings. The specific objectives of the study are given below.

1. To evaluate the performance of P and E_a products using ground-based measurements and to select the products that represent the best the spatio-temporal patterns of these variables.
2. To evaluate, whether the Budyko framework can be used to assess the performance of actual evaporation products in data-scarce regions.

Additionally, this work aims to provide the necessary information and knowledge about the water resources of the region and their distribution to relevant stakeholders and decision-makers. This knowledge can be particularly beneficial to shift towards a proactive water management approach that can serve as a basis to improve the decision-making process.

2 Study area

Balochistan is the largest province of Pakistan in terms of area (44% of the total land surface of the country) and is by far the smallest in terms of population (5% of the total population of the country; P&D- GoB, 2020). It covers a total land

Fig. 1. (a) Location of the study area. The figure also shows the location of the 14 selected rain gauge stations and the Digital Elevation Model (DEM) obtained from the US Geographical Survey's (USGS) Shuttle Radar Topography Mission (SRTM). (b) Location of the major catchments in the study area.

surface area of 347,190 km² with a population of 12.3 million inhabitants. The province lies on the south-eastern part of the Iranian plateau between the latitudes 25° and 32° N and between the longitudes 61° and 71° E (Fig. 1 [a]). It borders Iran to the west, Afghanistan to the north, the province of Khyber Pakhtunkhwa to the north-east, provinces of the Sindh and Punjab to the east, and the Arabian Sea to the south. The majority of the population in the province lives in the rural areas (85%) and depends on agriculture and livestock raising for their livelihood (Shafiq, 2008; Ashraf et al., 2014). These interdependent agro-pastoral activities provide around 67% of the employment to the total workforce in the region (Baloch & Tanik, 2008). Rangelands are the main source (90%) of the livestock feed requirements (Shafiq & Kakar, 2008).

The landscape of the province is composed of barren mountains, deserts, floodplains, and coastal plains (PDMA, 2020). As shown in Figure 1(a), the elevation ranges from 0 m.a.s.l. (in the southern coastal areas) to 3,558 m a.s.l. (in the northern highlands). There is a pronounced gradient of elevation which has a huge impact on the climate of the region (Ahmed et al., 2014). The northern highlands have very cold winters and warm summers while the plains in the east and lowlands in the south have very hot summers and mild winters (GoB, 2020).

The climate of the province is classified as arid-desert and arid-steppe based on the most updated Köppen-Geiger climate classification (Beck et al., 2018). The P is very low and unevenly distributed throughout the region (Adnan et al., 2009). The annual mean P ranges from 30 mm in the south-west desert to around 397 mm in the northeastern highlands (Ahmed et al., 2018). E_a losses are huge (e.g., in Quetta valley varies from 1 to 7.5 mm/day) across the province (Rasul & Mahmood, 1993). The only perennial surface water source comes from the Indus Basin Irrigation System (IBIS; 3.75 billion m³/yr.) that constitutes around 14% of total accessible water which irrigates around 1% of the total land surface area of the province (Aftab et al., 2018). The rest of the agriculture (around 6% of the total surface area of the province) is practiced at valley bottoms through floodwater harvesting, spate irrigation, dug wells, qanats, and the tubewells (Mustafa & Qazi, 2007).

Rainfall regimes are predominantly controlled by western depressions and the monsoon (Hussain & Lee, 2014). According to Ahmed et al. (2015a), winter P (December–March) is caused by the western disturbances that originate in the Mediterranean Sea and are responsible for around 58% of the total rainfall in the region. On the other hand, the summer monsoon winds, originated in the Bay of Bengal, are responsible for around 31% of the total rainfall that occurs from June to September. The south-eastern part of the province receives most of the P during the monsoon season, while the rest of the region receives most of the P during winter (Ahmed et al., 2018). As the monsoon winds advance from east to west, the moisture content in the air decreases and as a result, the P gradually reduces (Ahmed et al., 2019).

The major catchments in the study area are shown in figure 1(b). The eastern catchments (i.e., Zhob, Nari, Bolan, Mula, and Guj) flow to the Indus River basin in the east, and the southern catchments (i.e., Karradi, Porali, Hingol, Basol, Shadikor, and Kech) empty into the Arabian Sea in the south. Whereas the western and north-western Balochistan is entirely covered by the Kharan-Pashin-Lora (K.P.L.) closed basin which flows into numerous large (e.g., Hamun-e-Mashkel, Hamun-e-Lora, and Zangi Nawar lakes) and small desert lakes.

3 Materials and Methods

3.1 Data

3.1.1 Ground-based P data

Time series of monthly P data for 22 stations were obtained from the Pakistan Meteorological Department (PMD). The temporal coverage of the data ranges from 15 to more than 30 years. This study used 17 stations that had at least 18 years of record (2000–2017). The selection of the study period for the P analysis was selected because of

two main reasons: (i) to include the maximum number of stations in the study, and (ii) to include a period in which all the selected P products were available. Figure 1(a) shows the selected stations. Four stations (vis. Karachi, Larkana, Jacobabad, and D. I. Khan) out of the seventeen selected stations are located outside of the province of Balochistan (closer than 100 kilometers from the border of the province). These stations were included to have the maximum number of stations available for this study assuming that useful information can be obtained from stations lying in the proximity of the province.

The Expectation-maximization (EM) method (Dempster et al., 1977) was used to fill the monthly missing values — a total of four months — in Jacobabad, Jewani, Lasbela, and Quetta (S. Manda) stations (one month in each station). The EM method has been widely used in the literature to compute the missing P values (e.g., First et al., 2010; Ahmed et al., 2015a; Alamgir et al., 2015). A complete description of the EM algorithm can be found in Dempster et al. (1977).

3.1.2 P products

An increasing number of open-access P products are available for hydrological, ecological, and agricultural applications (e.g., Gehne et al., 2016; Zambrano-Bigiarini et al., 2017; Sun et al., 2018; Nicholson et al., 2019). This study evaluated three state-of-the-art satellite-based P products (i.e., CHIRPSv2, MSWEPv2.2, and PERSIANN-CDR) and one reanalysis dataset (i.e., ERA5) by comparing their estimates with ground-based measurements. The selection of these products was based on: (i) their long period of record, which is crucial in water balance applications, and (ii) their relatively high spatial resolution. Table 1 summarizes the main characteristics of the selected products.

CHIRPSv2

CHIRPSv2 is a satellite-based P product developed by the collaboration of the United States Geological Survey (USGS) and the University of California, Santa Barbara, and is available at https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_monthly/tifs/. This product has a quasi-global coverage (50°S–50°N), high spatial (0.05°) and temporal (daily) resolution, providing a long-term historical record (1981–present; Funk et al., 2015). This product was developed based on the integration of various other datasets including Climate Hazards group Precipitation climatology (CHPclim), TRMM 3B42, global Cold Cloud Duration (CCD) estimates, atmospheric model rainfall fields from the National Oceanic and Atmospheric Administration’s (NOAA) Climate Forecast System (CFS) and ground-based measurements from multiple sources (Tuo et al., 2016). This product has a latency of 2 days. Although the principal objective of developing this product was to monitor droughts, climate extremes, and environmental changes over land (Katsanos et al., 2016), it has been successfully used in a variety of hydrological applications (e.g., Katsanos et al., 2016; Tuo et al., 2016; Xian et al., 2019).

MSWEPv2.2

MSWEP version 2.2 is a global gridded P product spanning over more than 40 years (1979–present; Beck et al., 2019b) and is available at <http://www.gloh2o.org/>. MSWEPv2.2 features high spatial (0.1°) and temporal (3-hourly) resolutions (Beck et al., 2019b). This product was developed by combining a wide range of datasets such as rain gauges, satellite, and reanalysis P products aiming to provide an improved dataset worldwide (Beck et al.,

2019b). This product was specifically designed for hydrological applications (Beck et al., 2017a); taking into consideration the existing limitations of the P products to produce accurate estimates over mountainous, and snow-dominated regions around the globe (Zambrano-Bigiarini et al., 2017; Beck et al., 2020). It has been used in a range of hydrological studies such as hydrological modelling (Xiang et al., 2021), and P variability analysis (Chen & Dirmeyer, 2016; Zhang et al., 2017).

Table 1. Summary of P products used in this study.

Dataset	Spatial coverage	Spatial resolution	Temporal coverage	Temporal resolution	Reference
CHIRPSv2	50°S–50°N	0.05°	1981–present	daily	(Funk et al., 2015)
MSWEPv2.2	global	0.10°	1979–present	3-hourly	(Beck et al., 2019b)
PERSIANN-CDR	60°S–60°N	0.25°	1983–present	daily	(Ashouri et al., 2015)
ERA5	global	0.28°	1950–present	hourly	(Hersbach et al., 2018)

PERSIANN-CDR

PERSIANN-CDR is a quasi-global (60°S–60°N) satellite-based P product developed by the University of California, Irvine (UCI) to facilitate continuous monitoring of changing trends in daily P , particularly P extremes, caused by natural climate variability and climate change (Ashouri et al., 2015). PERSIANN-CDR is an open-source product available at <http://chrsdata.eng.uci.edu/>. The product features a relatively high spatial (0.25°) and temporal (daily) resolution covering more than 35 years (1983–present) of record. The dataset is generated by applying the PERSIANN algorithm on GridSat-B1 infrared data from NOAA’s National Centers for Environmental Prediction (NCEP; Ashouri *et al.*, 2015). The product is bias-corrected using monthly 2.5° data from the Global Precipitation Climatology Project (GPCP; Nguyen et al., 2019).

ERA5

ERA5 is an atmospheric reanalysis product produced by the European Centre for Medium-Range Weather Forecast (ECMWF) to replace ERA-Interim, which was released in 2006 (Hersbach et al., 2018). The product has global coverage with a spatial resolution of approximately 31km and is produced at 1-hourly temporal resolution with a coverage of more than 40 years (1979–present). This product is based on the Integrated Forecasting System (IFS) Cycle 41r2 and is available after five days (Hersbach et al., 2018). Furthermore, the dataset is based on an advanced numerical model and a vast variety of observation inputs and forcing datasets (Nogueira, 2020). Many recent studies have used ERA5 for many hydrological applications. These include modeling of global mean P (Hersbach et al., 2018), evaluating daily P products over the continental United States (Beck et al., 2019a), modelling of surface runoff (Matveeva & Sidorchuk, 2020), and catchment scale hydrological modelling (Tarek et al., 2020). The dataset can be obtained at <https://cds.climate.copernicus.eu/>.

3.1.3 *E* products

As ground-based E_a information is not available in the study area, we used the Budyko framework that uses potential evaporation (hereafter defined as E_p) information to calculate E_a which was then used as a reference to evaluate the E_a products (see section 2.3.2). Thus, this study used both E_a and E_p products. An overview of these products is given below. Table 2 summarizes the E products.

GLEAMv3.3

The GLEAMv3.3 is a satellite-based E and root-zone soil moisture dataset developed in 2011 (Miralles et al., 2011). This dataset features a relatively high spatial (0.25°) and temporal (daily) resolution with global coverage of 39 years (1980–2018). The Priestley and Taylor equation is used to calculate E_p based on surface radiation and near-surface air temperature (Miralles et al., 2011). The obtained E_p estimates are then converted to different components of terrestrial evaporation (i.e., interception, transpiration, bare soil evaporation, open water evaporation, and sublimation) separately considering four different land cover types (bare soil, low vegetation, tall vegetation, and open water; Martens et al., 2016). The land cover types are obtained from MODIS data (Martens et al., 2016). This study used two datasets from the GLEAMv3.3: i.e., E_p product and E_a product. The datasets are available in two latest versions i.e., GLEAMv3.3a and GLEAMv3.3b, which are available at www.gleam.eu. We used both E_p and E_a datasets from GLEAMv3.3a because this version uses a variety of observation inputs including gauged-based, satellite-based, and reanalysis datasets while the GLEAMv3.3b largely depends on satellite datasets alone (Martens et al., 2017).

SSEBopv4.0

The operational Simplified Surface Energy Balance (SSEBopv4.0; Senay et al., 2013), is one of the promising E_a products available in recent times (de Paula et al., 2019). The model is based on the Simplified Surface Energy Balance (SSEB) approach (Senay et al., 2011) with a unique simplified parameterization (Senay et al., 2013). The novelty of the SSEBopv4.0 parameterization is that for each pixel, the difference between the hot and the cold reference values are pre-defined and seasonally dynamic. It uses a thermal index approach to combine the local reference evaporation (E_o) with evaporation fractions obtained from remotely sensed thermal imagery from MODIS datasets (Senay et al., 2013). The pre-defined temperature difference is added to the cold reference value to obtain the hot reference value, whereas the cold reference value is obtained as a fraction of the air temperature (Senay, 2018). The SSEBopv4.0 product has been used in a variety of hydrological studies in data-scarce regions for famine early warning, drought monitoring, mapping E_a , and water accounting in different land-use types (Senay et al., 2011; Senay et al., 2016; Senay et al., 2013). The model is available at 10-day, monthly and annual temporal resolutions, and at a spatial resolution of 1 km. The SSEBopv4.0 covers more than 18 years of record (2003–present) and is available at <https://earlywarning.usgs.gov/fews/datadownloads>.

Table 2. Summary of the E products used in this study.

Dataset	Spatial coverage	Spatial resolution	Temporal coverage	Temporal resolution	Reference
GLEAMv3.3a (E_a and E_p)	global	0.25°	1980–2018	daily	(Miralles et al., 2011)
SSEBopv4.0 (E_a)	global	1 km	2003–present	10-day and monthly	(Senay et al., 2013)
ERA5 (E_a and E_p)	global	0.28°	1950–present	hourly and daily	(Hersbach et al., 2018)

ERA5

A general overview of ERA5 has already been discussed in section 2.2.2. However, besides the P product of the ERA5, this study also used E_a and E_p products from the same dataset. E_p parameter represents the extent to which the near-surface conditions are conducive to facilitate maximum possible E_a in a given time (ECMWF, 2020). This parameter is developed based on the surface energy balance calculations under certain underlying assumptions (ECMWF, 2020). One major assumption that may limit its performance in arid and semi-arid regions is that it assumes that the agricultural land use is well watered and that the atmospheric conditions are not affected by this artificial surface condition (ECMWF, 2020). E_a parameter represents the total evaporated water in a given time from the surface of the earth including the transpired water from the vegetation (ECMWF, 2020).

3.2 Methodology

3.2.1 Evaluation of P products

Four P products (CHIRPSv2, PERSIAN-CDR, MSWEPv2.2, and ERA5) were evaluated against 17 ground-based stations for 18 years (2000–2017) through a point-to-pixel approach (Duan et al., 2012; dos Reis et al., 2017; Paredes-Trejo et al., 2017; Zambrano-Bigiarini et al., 2017; Baez-Villanueva et al., 2018), which assumes that the ground-based measurements are representative of the P of the respective grid-cells of the products. The evaluation was performed on different temporal scales (i.e., monthly, seasonal, and annual) to analyse the performance of the P products for different hydrological applications (dos Reis et al., 2017). The P products were accumulated to monthly, seasonal, and annual totals for comparison. Following Zambrano-Bigiarini et al. (2017), all the products with a higher spatial resolution than 0.25° were unified to a grid cell of 0.25° by using the bilinear interpolation method to ensure a consistent evaluation. Bilinear interpolation technique averages the nearest four input cells to obtain the output resampled cell (Wu et al., 2008). According to Baez-Villanueva et al. (2018), resampling of the products using bilinear interpolation might impact the performance of the P products. However, this study assumes that the resampling of the products had no impact on the obtained results.

The ability of the selected products to represent the P distribution over the study area and at different temporal scales (monthly, seasonal and annual) was evaluated using five continuous statistical indices: the modified Kling-

Gupta efficiency (KGE' , Eq. 1; Gupta et al., 2009; Kling et al., 2012) along with its three individual components, i.e., the linear correlation (r ; Eq. 2), bias ratio (β ; Eq. 3) and variability ratio (γ ; Eq. 4), and the root mean square error (RMSE; Eq. 5). Following Zambrano-Bigiarini et al. (2017) and Muthoni et al. (2019), this study selected the KGE' because P products are generally expected to be able to reproduce temporal dynamics (measured by r) of P while also preserving the total volume (measured by β) and distribution (measured by γ ; Zambrano-Bigiarini et al., 2017). The KGE' and its components can be defined as follows:

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

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

$$\beta = \frac{\mu_s}{\mu_o} \quad (3)$$

$$\gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o}, \quad (4)$$

where n is the number of observations, O_i and S_i are the observed and estimated values at day i respectively; \bar{O} and \bar{S} are the arithmetic mean of the observed and estimated P , respectively; μ is the distribution mean; σ is the standard deviation and CV is the coefficient of variation.

The KGE' and its components have their optimum value at one (Kling et al., 2012). According to Baez-Villanueva et al. (2018), a value equal to 1 represents perfect linear correlation, a value of -1 represents a perfect negative linear correlation, while a 0 indicates the absence of linear correlation. β greater than 1 represents overestimation and β less than 1 represents underestimation. γ shows whether the dispersion of the estimated P values is higher or lower than the observed values at gauge stations.

Furthermore, the RMSE was applied to estimate the average error value between observed and estimated P (Prasetia et al., 2013). This index has widely been used in the evaluation of P products (e.g., dos Reis et al., 2017; Iqbal and Athar, 2018; Nicholson et al., 2019). RMSE value of 0 represents perfect agreement between observed and estimated datasets (Chen & Liu, 2012). The RMSE can be defined as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

where n is the number of observations; y is the observed value and \hat{y} is the corresponding P estimated value at day i .

All the aforementioned performance indices were computed in the R environment 3.3.1 (R Core Team, 2011) using the raster (Hijmans, 2020), hydroGOF (Zambrano-Bigiarini, 2020a), and hydroTSM (Zambrano-Bigiarini, 2020b) R packages. Furthermore, this study also used ArcMap 10.6.1. (ESRI, 2018) for processing the P products and creating the maps.

3.2.2 Evaluation of E_a products

Due to the unavailability of ground-based E_a data over the study area, this study firstly estimated E_a based on the simplified water balance approach (Cheng et al., 2011) and subsequently compared the results with the E_a products to evaluate their performance (for other examples of this approach the reader is referred to Cheng et al., 2011; Mianabadim et al., 2017). The simplified water balance approach is based on the water budget of a catchment and for any time can be expressed as:

$$E_a = P - R + \Delta S \quad 6$$

where E_a is actual evaporation — also referred as actual evapotranspiration (Miralles et al., 2020) — P is precipitation; R is surface runoff, interflow, and baseflow; and ΔS is the change in water storage. Commonly, over a long period of time, e.g., ≥ 5 -year (Tekleab et al., 2010), groundwater recharge and water accumulation in a natural catchment can be assumed to be negligible (Tekleab et al., 2010; Xiong & Guo, 2011; Du et al., 2016; Gunkel & Lange, 2017). As a result, a long-term steady-state water balance in a catchment can be expressed as:

$$P - R - E_a = 0 \quad 7$$

The long-term mean runoff R in a catchment can be calculated by the difference between the long-term mean P and the long-term mean E_a assuming that water accumulation at the catchment scale is negligible (Caracciolo et al., 2017), and can be expressed as:

$$R = P - E_a \quad 8$$

The steady-state water balance is controlled by water availability (P) and atmospheric water demand (E_p). The E_p in turn is controlled by available energy (Zhang et al., 2008). This relationship between water availability, atmospheric water demand, and energy availability is well established by the widely employed Budyko framework (Budyko, 1974). This framework assumes that over a large timescale under natural conditions water accumulation at the basin scale can be neglected and the long-term mean E_a is determined by long-term mean P and the atmospheric water demand (Tekleab et al., 2010). Under the given assumptions, the Budyko framework establishes a functional link (ψ) that relates the ratio between the long-term mean annual actual evaporation E_a and the long-term mean annual precipitation P with the aridity index; defined as the ratio between the long-term mean annual potential evaporation E_p and the long-term mean annual precipitation P (Caracciolo et al., 2017), which can be expressed as:

$$\frac{E_a}{P} = \psi \left(\frac{E_p}{P} \right) \quad (9)$$

The original Budyko framework has been widely used in different regions around the world to estimate the long-term water balance in different catchments (e.g., Porporato et al., 2004; Istanbuluoglu et al., 2012). However, it has always been observed that the approximation of the estimated mean E_a varies from catchment to catchment and in different regions due to several physical processes such as vegetation, soil characteristics, and seasonality in climate which are not taken into account in the original Budyko framework (Xiong & Guo, 2011; Xu et al., 2013). As a result, many researchers (e.g., Choudhury, 1999; Fu, 1981; Zhang et al., 2004) have proposed different analytical solutions to this limitation (e.g., ignoring the catchment characteristics) to the original Budyko framework. Fu's (1981) equation is one of the most widely used analytical solutions to the original Budyko framework in the literature and can be expressed as:

$$\frac{E_a}{P} = 1 + \frac{E_p}{P} - \left(1 + \left(\frac{E_p}{P} \right)^\omega \right)^{\frac{1}{\omega}} \quad (10)$$

where ω , known as Fu's parameter, has no analytical solution but it has been related to different measured data (e.g., vegetation, soil characteristics, seasonality in climate, and other catchment characteristics) or it can be adopted from other regional and global studies that have assessed ω (Caracciolo et al., 2017). Xu et al. (2013) developed a model to predict the spatial distribution of ω globally. They developed a neural network (NN) model using 224 small (100 – 10,000 km²) U.S. basins and 32 large (230,000 – 600,000 km²) global basins considering both local (i.e., slope and normalized difference vegetation index [NDVI]) and global (location) factors and mapped the ω globally. They also validated the predicted E_a over 36,600 large and small global basins with an independent satellite-based E_a product and found good agreement ($R^2=0.72$). Therefore, this study employed Fu's equation to estimate E_a and adopted ω values from Xu et al. (2013) for the selected catchments (see Figure 1(b)). Xu et al. (2013) suggest an average ω value of 1.8 for the catchments in the study area. The Fu's ω parameter, however, was used without any regional validation or calibration which may lead to certain uncertainties in estimating the E_a at catchment scale (Gunkel & Lange, 2017). Moreover, most of the catchments in the study area are not natural and experience at least some degree of anthropogenic influence (Baloch & Tanik, 2008) which may also lead to certain uncertainties.

To estimate mean E_a from Eq. 10, mean annual E_p (2003–2018) were obtained from ERA5 and GLEAMv3.3a products for the major catchments in the study area. For each catchment two mean E_a were calculated; one from ERA5 and the other from GLEAMv3.3a, adopting a ω value of 1.8. The obtained mean E_a values were then compared against three E_a products, namely ERA5, GLEAMv3.3a, and SSEBopv4.0 to assess their performance based on the assumption that the less the difference between mean E_a obtained from Eq. 10 and mean E_a obtained from E_a products the better has the E_a product performed. The selection of the E_p products, i.e., ERA5 and GLEAMv3.3a, and E_a products, i.e., ERA5, GLEAMv3.3a, and SSEBopv4.0, was based on the maximum possible temporal coverage of the study period governed by P (ground-based and products) and E_a (products) data availability.

4 Results

4.1 Spatial performance of the P products

The spatial distribution was computed for the selected continuous indices (r , β , γ , KGE' and RMSE) at a monthly scale showing all the 17 stations. Figure 2(a) shows the KGE' values for the 4 selected P products compared against the ground observations at the monthly temporal scale. MSWEPv2.2 had the best KGE' performance followed by ERA5, CHIRPSv2, and PERSIANN-CDR, respectively. MSWEPv2.2 presented relatively good performance ($0.6 < \text{KGE}' \leq 0.8$) for most of the stations distributed within the study area except a few stations with relatively low ($0.4 < \text{KGE}' \leq 0.6$) KGE' values and one station (Lasbela) with very low KGE' value (0.02). ERA5 closely followed MSWEPv2.2 for most of the stations but showed a very low KGE' value (0.11) over the D. I. Khan station, which lay outside of the boundary of the study area. CHIRPSv2 showed better performance than PERSIANN-CDR. However, it presented KGE' values < 6 for all the stations located inside the boundary of the province except two stations located in the far northeastern part of the province. The performance of CHIRPSv2 was the lowest ($0.0 \leq \text{KGE}' < 0.4$) in the western part of the study area. PERSIANN-CDR showed the lowest KGE' performance overall. It had also lower KGE' values for the stations located in the southern part of the province as observed in the CHIRPSv2. In the case of PERSIANN-CDR, two stations (i.e., Nukkundi and Panjgur) had KGE' values in the negative range (-0.62 and -0.2 respectively).

The spatial distribution of r (related to the temporal dynamics of P) is shown in Figure 2(b) and revealed that MSWEPv2.2 had the best linear correlation followed by ERA5, PERSIANN-CDR, and CHIRPSv2, respectively. The values of r for MSWEPv2.2 varied from 0.63 to 0.95. Most of the stations had r values > 0.80 . ERA5 presented r values ranging from 0.68 to 0.88 and most of the stations lied in the range of 0.68 to 0.80. PERSIANN-CDR also depicted a good correlation ($0.66 \leq r \leq 0.92$) for all stations except Pasni station, which had a value of 0.59. CHIRPSv2 had the poorest performance in which r values were particularly low ($0.27 \leq r \leq 0.40$) in the south-western part of the province.

The spatial distribution of the β component (related to the total volume of P) of the KGE' at the monthly temporal scale is shown in Figure 3(a). CHIRPSv2 presented the best performance for the β component exhibiting values ranging from 0.87 to 1.35 for all the stations located within the study area and nearly a perfect agreement ($1 \leq \beta \leq 1.09$) between observed and estimated P (in terms of volume) for almost all the stations located in the southern part of the province. MSWEPv2.2 and ERA5 were the second and third best-performing products, respectively; with slightly different performances among each other. MSWEPv2.2 underestimated more stations than ERA5. PERSIANN-CDR was the product with the lowest performance and overestimated most of the stations particularly the ones located in the western part of the study area.

Fig. 2. (a) Modified Kling-Gupta Efficiency (KGE') and (b) linear correlation (r) evaluation between monthly P products and ground-based observations at the corresponding grid cells.

Fig. 3. (a) Bias ratio (β) and (b) variability ratio (γ) evaluation between monthly P products and ground-based observations at the corresponding grid cells of the P products.

Fig. 4. Root Mean Square Error (RMSE) evaluation between monthly P products and ground-based observations at the corresponding grid cell.

Figure 3(b) shows the spatial distribution of the γ component (related to the distribution of P) of the KGE' at the monthly temporal scale. As evident from this Figure, MSWEPv2.2 was the best performing product ($0.79 \leq \gamma \leq 1.13$) followed by ERA5, PERSIANN-CDR, and CHIRPSv2, respectively. ERA5 closely followed MSWEPv2.2 in most of the stations but relatively overestimated P over Panjgur station and underestimated P over Ormara station. PERSIANN-CDR and CHIRPSv2 underestimated the P values over the western part of the province. CHIRPSv2 also underestimated some stations over the eastern part of the study area.

Finally, the spatial distribution of the RMSE at the monthly scale is shown in Figure 4. Generally, RMSE performance for all the products was better in the drier stations compared to the wetter ones. As a result, the performance of the products was better in the stations located in the western part, which have less P , while the performance decreased towards the east where the stations receive relatively more P . However, MSWEPv2.2 followed by ERA5 had a relatively better performance for the stations located in the western part of the province compared to PERSIANN-CDR and CHIRPSv2, while PERSIANN-CDR followed by CHIRPSv2 had a relatively better performance for the stations located in the eastern part of the province compared to MSWEPv2.2 and ERA5.

4.2 Temporal performance of the P products

This study evaluated the performance of P products at the monthly, seasonal, and annual temporal scales. Following Iqbal and Athar (2018), and Sadiq (2013) the seasonal timescale was decomposed into 4 seasons, namely winter (December–March), pre-monsoon (April–June), Monsoon (July–September), and post-monsoon (October–November) based on the P occurrence and amount in the study area. The following paragraphs give a summary of the temporal performance of the P products measured by the selected continuous indices (KGE' r , β , γ , and RMSE) at corresponding grid cells for all the defined stations included in the study. The evaluation results are presented as boxplots in Figure 5.

Fig. 5. Modified Kling-Gupta Efficiency (KGE') along with its three individual components (r , β , and γ) and the Root Mean Square Error (RMSE) between satellite products and ground observations at the monthly, annual, and seasonal (winter, pre-monsoon, monsoon and post-monsoon) temporal scales. The horizontal green line indicates the optimum value of the performance indices.

The best r values of the evaluated P products were observed during the winter season ($0.71 \leq \text{median values} \leq 0.85$) followed by post-monsoon ($0.73 \leq \text{median value} \leq 0.8$). The products showed median values > 0.6 for all of the temporal scales except for the pre-monsoon in which PERSIANN-CDR exhibited a median value of 0.49. None of the P products performed the best in all of the temporal scales. However, MSWEPv2.2 was the best performing product at the monthly, monsoon, and post-monsoon temporal scales, while PERSIANN-CDR was the best performing product in winter and pre-monsoon seasons. ERA5 performed the best at the annual timescale and CHIRPSv2 was the least performing product in all the temporal scales except monsoon where it performed better than ERA5 and PERSIANN-CDR.

In the case of β , MSWEPv2.2 was almost unbiased (median value = 1) for all the temporal scales. CHIRPSv2 was also almost unbiased for monthly, annual, and monsoon temporal scales. ERA5 also showed its median value near

to 1 for the winter season. In general, all the products performed well over the winter season and worst over the post-monsoon. In overall temporal bias ratio performance, MSWEPv2.2 showed the best median performance followed by CHIRPSv2 and ERA5. PERSIANN-CDR was the worst-performing product and overestimated P at all temporal scales.

The evaluation of (γ) at different temporal scales revealed that the distribution of P at all the temporal scales was underestimated by all products except for MSWEPv2.2, which presented a median value close to 1 for the pre-monsoon season. However, the monthly timescale ($0.59 \leq \text{median} \leq 0.9$) followed by the annual timescale ($0.51 \leq \text{median} \leq 0.9$) was best represented by the P products compared to the rest of the temporal scales (i.e., winter, pre-monsoon, monsoon, and post-monsoon). In general, for all the temporal scales, MSWEPv2.2 was the best performing product followed by ERA5, PERSIANN-CDR, and CHIRPSv2.

The overall performance of the r , β , and γ is summarized in the KGE' as shown in Figure 5. All the P products performed the best during the monthly temporal scale except PERSIANN-CDR, which showed its best performance during winter. The monthly median KGE' values for MSWEPv2.2 and ERA5 varied from 0.68 to 0.62 respectively and for PERSIANN-CDR and CHIRPSv2 from 0.36 to 0.39, respectively. After the monthly timescale, the best performance of all P products was observed during the winter season followed by the annual temporal scale. All the products performed worst during post-monsoon followed by monsoon and pre-monsoon seasons. MSWEPv2.2 was the best performing product during monthly timescale and pre-monsoon, monsoon, and post-monsoon seasons; while ERA5 performed the best during the annual and winter temporal scales. In general, after MSWEPv2.2, ERA5 was the best performing product followed by CHIRPSv2 and PERSIANN-CDR. The performance of PERSIANN-CDR was particularly poor in pre-monsoon, monsoon, and post-monsoon where it showed negative KGE' values.

The best RMSE performance was observed in the post-monsoon ($6.5 \leq \text{median values} \leq 10.7$) followed by the monthly ($18 \leq \text{median values} \leq 20$), pre-monsoon ($21 \leq \text{median values} \leq 26.2$) winter ($31.6 \leq \text{median values} \leq 38.1$), and monsoon ($46.7 \leq \text{median values} \leq 64.3$) temporal scales. The annual temporal scale had the worst RMSE performance ($75.8 \leq \text{median values} \leq 89.2$) for all the P products. MSWEPv2.2 depicted the best median performance at monthly and post-monsoon temporal scales, CHIRPSv2 showed the best performance at annual, and monsoon temporal scales, and PERSIANN-CDR showed the best performance at winter and pre-monsoon temporal scales. Whereas ERA5 did not have the best median performance for any temporal scales. Generally, the values of MSWEPv2.2 and ERA5 had a larger spread to the median than the CHIRPSv2 and PERSIANN-CDR. Hence, it can be summarized that overall CHIRPSv2 followed by PERSIANN-CDR had the best RMSE performance.

4.3 Performance of the E_a products

Due to the unavailability of ground-based E_a information in the study area, this study first estimated mean E_a based on the simplified water balance approach, here defined as E_{a_WB} , (see section 2.3.2), and subsequently compared the results with the E_a products to evaluate their performance (as in Cheng et al., 2011; Mianabadi et al., 2017).

4.3.1 Estimation of E_{a_WB}

This study used Fu's (1981) equation (Eq. 10; see section 2.3.2) to estimate E_{a_WB} for the selected catchments. Fu's equation required P , E_p , and ω to calculate E_{a_WB} . P data were used from the best performing P product at the multi-annual temporal scale (i.e., ERA5), while E_p data were extracted from two different products, i.e., GLEAMv3.3a and ERA5; this study used two E_p products so that two aridity indices (AI) could be estimated and compared for each catchment, and employed a $\omega = 1.8$ adopted from Xu et al. (2013). The two estimated AIs were then compared to the climate classification of Köppen-Geiger (Beck et al, 2018) and the AI information available in the literature to select the E_p that made more sense in terms of climate and previous literature. Finally, the E_p product that was best representing the AI in the study area was used in the Fu's equation to calculate the mean annual E_{a_WB} for the selected catchments during 2003–2018. Figure 6 shows the evaporation index (EI; E_{a_WB}/P) versus aridity index (AI; E_p/P) and the Budyko curve calculated from both the GLEAMv3.3a and ERA5 E_p products. All the catchments fell within the water and energy limit boundaries (water and energy availability); and therefore, fulfill the Budyko assumptions. For the AI, ERA5 exhibited drier results ($4 < IA < 20$) compared to the GLEAMv3.3a ($1 < IA < 9$) for most of the catchments. However, the majority of the catchments lay in the range of $4 < AI \leq 6$ calculated from both the GLEAMv3.3a and ERA5.

The United Nations Environment Programme (1992) classifies climate zones into hyper-arid ($AI > 20$), arid ($5 < AI \leq 20$), semi-arid ($2 < AI \leq 5$), dry sub-humid ($1.5 < AI \leq 2$) and humid ($AI < 1.5$). Based on this classification, as shown in Figure 6, the catchments in the study area lay in the arid and semi-arid climate zones when AI was calculated using ERA5; while using GLEAMv3.3a the catchments lay in the arid, semi-arid, and dry sub-humid climate zones. More specifically, GLEAMv3.3a classified western catchments (Basol, Kech, Shadikor, and K.P.L. closed basin) as arid, south-eastern catchments (Guj, Hingol, Karadi, and Porali) as semi-arid and, central and north-eastern catchments (Bolan, Nari, and Zhob) as dry sub-humid (see Figure 1(b) for the location of the catchments). On the other hand, ERA5 classified all the catchments as arid except north-eastern catchments (Bolan, Nari, and Zhob) which were classified as semi-arid.

In order to decide which product, i.e., GLEAMv3.3a or ERA5, more accurately depicted the extent of the aridity in the study area, this study evaluated the climate classification results based on UNEP (1992) obtained from GLEAMv3.3a (arid, semi-arid, and dry sub-humid) and ERA5 (arid and semi-arid) in the study area in two ways; (i) the obtained climate results were compared to the Köppen-Geiger climate classification (Beck et al, 2018). According to this classification, most of the study area lies in the arid-desert (arid) climate zone except the north-eastern part which exhibits an arid-steppe (semi-arid) climate (Beck et al., 2018). (ii) the obtained climate results were compared to the results of the previous studies which have specifically analysed the aridity of the study area. These studies included Ahmed et al. (2015a), Ahmed et al. (2017), and Haider and Adnan (2014). In this sense, Haider and Adnan (2014) classified the Balochistan province as arid except the north-eastern part which was classified as semi-arid. Ahmed et al. (2017) and Ahmed et al. (2015a) had similar results as Haider and Adnan (2014) except with a difference that they classified extreme western Balochistan as hyper-arid.

The results obtained from ERA5 (i.e., arid and semi-arid) are in better agreement with the climate classification of the Köppen-Geiger (arid-desert and arid-steppe) and also to the results in the regional studies (arid and semi-arid) compared to the climate classification results obtained from the GLEAMv3.3a (i.e., arid, semi-arid and dry

sub-humid). Therefore, we selected ERA5 as the product to calculate the E_{a_WB} . These results suggest that ERA5 E_p estimates are of better quality to evaluate the performance of the E_a products, compared to the E_{a_WB} estimated with GLEAMv3.3a.

4.3.2 E_{a_WB} vs. E_a products

The estimated E_{a_WB} values from the Fu's equation were compared to three E_a products (i.e., GLEAMv3.3a, SSEBopv4.0, and ERA5) to evaluate their performance. The comparison was carried out at the multi-annual scale (2003–2018) and the long-term change in storage was neglected. As shown in Figure 7, none of the products reproduced the estimated E_{a_WB} satisfactorily. However, the performance of ERA5 was better in 6 catchments (i.e., Basol, Bolan, Hingol, Kech, Mula, and Porali) out of a total of 12 defined catchments; GLEAMv3.3a performed the best in 4 catchments (i.e., Guj, Karadi, Nari and Zhob) and SSEBopv4.0 in two (i.e., K.P.L. closed basin and Shadikor). Generally, GLEAMv3.3a performed the best over the north-eastern catchments which receive relatively high P ; ERA5 performed the best in the southern catchments which receive relatively low P compared to the north-eastern region; and SSEBopv4.0 performed the best in the north-western region which receive the lowest P in the study area.

Considering the performance of the products based on the area instead of the number of catchments, SSEBopv4.0 was the best performing product (125879.5 km² within the province as K.P.L. closed basin is a transboundary basin) followed by ERA5 (867250 km²) and GLEAMv3.3a (57474 km²). However, the performance of the SSEBopv4.0 was relatively low in the rest of the catchments where it was not the best performing product. On the other hand, ERA5 showed comparable results to SSEBopv4.0 in catchments where SSEBopv4.0 was the best performing product. For example, SSEBopv4.0 overestimated the Shadikor by 91.5 mm/yr, while ERA5 overestimated the catchment by 96.4 mm/yr. Similarly, SSEBopv4.0 underestimated the K.P.L. closed basin by 13.2 mm/yr, while ERA5 underestimated the catchment by 48 mm/yr. So, considering the comparable results of ERA5 to SSEBopv4.0 in Shadikor and K.P.L. closed basin, and ERA5's best performance in 6 catchments in the study area, it could be concluded that generally ERA5 showed the best performance in the province followed by SSEBopv4.0 and GLEAMv3.3a respectively.

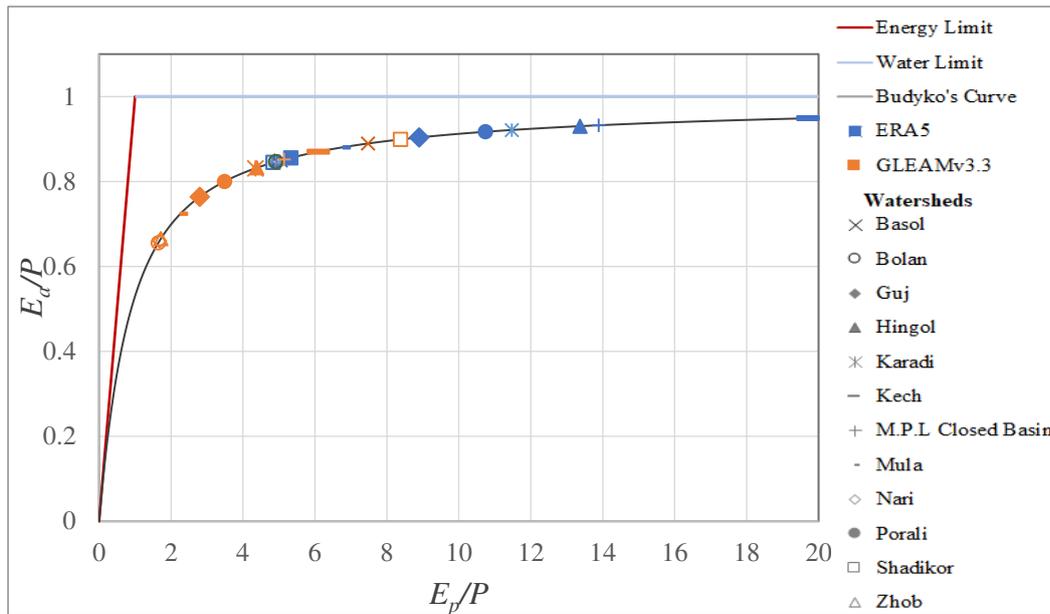


Fig. 6. The ratio of mean annual evaporation to mean annual precipitation (E_e/P) as a function of aridity index (E_p/P) for the selected catchments for 16 years (2003–2018). Scatter points represent the distribution of catchments over Budyko curve.

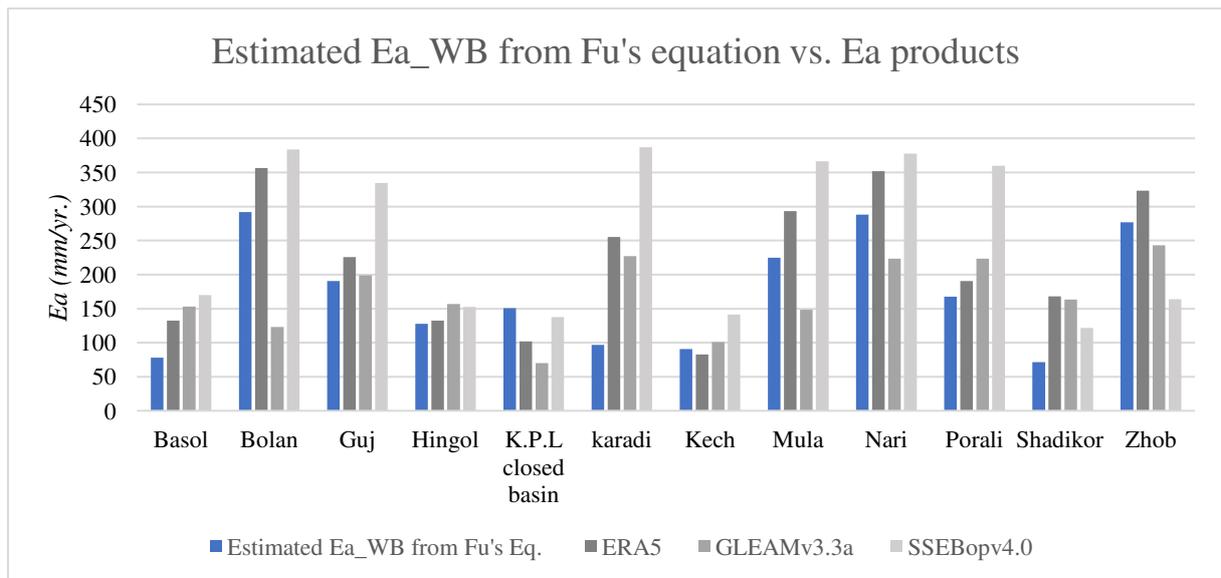


Fig. 7. Comparison of multi-annual means (2003–2018) of estimated E_aWB from Fu's equation to E_a products.

5 Discussion

This work aims to evaluate for the first time the key water balance components over the Balochistan province in Pakistan. The present study evaluated the spatio-temporal performance of 4 state-of-the-art P products (CHIRPSv2, MSWEPv2.2, PERSIANN-CDR, and ERA5) using 17 ground-based station data, and 3 state-of-the-art E_a products (GLEAMv3.3a, ERA5, and SSEBopv4.0) by applying the widely used Budyko framework over the data-scarce Balochistan province.

5.1 Performance of P products

The spatial distribution of the performance of the products evaluated with KGE' at the monthly timescale (Figure 2) revealed that MSWEPv2.2 was the best performing product followed by ERA5, CHIRPSv2, and PERSIANN-CDR, respectively. The performance of MSWEPv2.2 and ERA5 was relatively satisfactory ($0.6 < \text{KGE}' \leq 0.8$) for most of the stations and consistent across the province with a few exceptions. However, CHIRPSv2 and PERSIANN-CDR presented relatively poor performance (< 6). Their performance was particularly poor over the western part of the province, which receives most of the P during winter, compared to the stations located in the eastern part of the province that receives most of the P during the summer monsoon. Wu et al. (2019) obtained similar results for the performance of P products during winter and summer P over Yunnan province, China. This can be attributed to the fact that satellite P products generally tend to detect strong and convective P events (e.g., monsoon) more effectively than a drizzle or frozen P during the winter (Guo et al., 2015).

Similar results were obtained at annual and seasonal (winter, pre-monsoon, monsoon, and post-monsoon) temporal scales (Figure 5) for the KGE'. MSWEPv2.2 was the best performing product during the monthly, pre-monsoon, monsoon, and post-monsoon temporal scales, while ERA5 performed the best during annual and winter temporal scales. CHIRPSv2 also presented relatively good performance during monsoon and post-monsoon. However, PERSIANN-CDR presented poor performance, scoring negative KGE' values during pre-monsoon, monsoon, and post-monsoon. The poor performance of PERSIANN-CDR might be attributed to the fact that unlike the other satellite-based products (MSWEPv2.2 and CHIRPSv2) used in this study, that directly incorporate temporally finer gauge data for corrections, PERSIANN-CDR relies on the GPCP dataset to correct the monthly P bias (Beck et al., 2017b).

In general, the merged products (which combine gauges, reanalysis, and satellite-based data), that directly incorporate daily gauge data (e.g., MSWEPv2.2) performed the best, followed by ERA5, which is a reanalysis product. The ERA5 was then followed by CHIRPSv2, which also incorporates reanalysis data, and temporally coarser gauge data compared to MSWEPv2.2. Whereas Gauge-satellite-based product that indirectly incorporates gauge data (e.g., PERSIANN-CDR) through other multi-source datasets (e.g., PERSIANN-CDR is adjusted using gauge-based GPCP datasets) performed the worst. These findings are in agreement with Beck et al. (2017b), who studied the performance of 22 P products on a global scale and concluded that the products that directly incorporate daily gauge data perform better than the products that incorporate temporally coarser gauge data. In turn, these products outperform the products that indirectly incorporate gauge data.

Moreover, the products generally showed their best KGE' performance during the monthly timescale compared to the annual which is in agreement with Duan et al. (2016) and Zambrano-Bigiarini et al. (2017). The better performance of the products at monthly timescale is also due to the seasonality of the P at the monthly timescale compared to the annual which results in better r performance of the KGE'.

For the seasonal evaluation, the best KGE' performance was observed during the winter season followed by pre-monsoon and monsoon seasons. The poorest performance was observed during the post-monsoon season. Generally, the wetter the season was, the better the performance of the P products (e.g., best performance during winter). These findings are in agreement with the findings of Zambrano-Bigiarini et al. (2017), who concluded that the P products generally performed better in the wetter seasons.

RMSE results were different than those of the KGE' results which is in agreement with Baez-Villanueva et al. (2018). For example, contrary to the KGE' results where MSWEPv2.2 and ERA5 generally were the best performing products across the province and during all the temporal scales, in the case of the RMSE, PERSIANN-CDR followed by CHIRPSv2 had a relatively better performance. The RMSE performance was particularly better in the stations located in the eastern part of the province while MSWEPv2.2 followed by ERA5 showed a relatively better performance for the stations located over the western part of the study area.

According to Baez-Villanueva et al. (2018), the RMSE places more importance on mismatches of high P values (i.e., the rainy season) compared to the dry season. On the contrary, KGE' appears to be a better index to evaluate the performance of the P products due to two main reasons; (i) because its ability to reduce the excessive weight of the high P events; and (ii) because it decomposes the total performance of the P products into r (reproduces temporal dynamics), β (preserves the volume) and γ (preserves the distribution; Zambrano-Bigiarini et al., 2017).

The results of this study confirm that it is difficult to point out a single best-performing P product that performs the best across the province and during all temporal scales. However, MSWEPv2.2 performs the best, followed closely by ERA5. MSWEPv2.2 performed the best during the monthly, pre-monsoon, monsoon, and post-monsoon temporal scales, while ERA5 was the best performing product during the annual and winter temporal scales.

5.2 The capability of Budyko framework to estimate E_{a_WB}

As shown in figure 7, the E_{a_WB} were obtained using Fu's equation. Most of the E_{a_WB} results were not adequately comparable to those of the E_a products in most of the catchments. Although this study takes the E_{a_WB} as the reference E_a to evaluate the performance of the E_a products, the possibility that Fu's equation was not able to adequately estimate the E_{a_WB} in the defined catchments cannot be exempted. In other words, the poor agreement between E_{a_WB} and E_a products could be attributed either to the inherent uncertainties and poor performance of the Fu's equation or the poor performance of the E_a products. This could be investigated in future studies for the study area. However, the inadequate agreement between E_{a_WB} and E_a products is in agreement with the results obtained by Tekleab et al. (2010).

Furthermore, Tekleab et al. (2010) argue that poor performance of the Fu's equation could be attributed to many uncertainties including; ignoring the soil water storage in the catchments, catchment P and E_p , seasonality of the P , and non-stationary characteristics (e.g., NDVI, slope, land cover, etc.) of the catchments themselves measured by the Fu's ω parameter. Gunkel and Lange (2017) also found that regionally calibrated ω values lead to much better performance of the Fu's equation. Moreover, Cheng et al. (2011) noted that human interference to the catchment water cycle also leads to the poor performance of the Fu's equation. Our study notes that certain underlying assumptions in Budyko framework and Fu's equation may lead to inadequate performance of this framework which is in agreement with Tekleab et al. (2010), Cheng et al. (2011), and Gunkel and Lange (2017).

5.3 Possible biases in estimating the AI and the E_{a_WB}

Forcing datasets have a large influence over the accuracy of the E_a estimates (Mueller et al., 2013). Water input (measured by P) and energy availability (measured by E_p) are the two main variables affecting the E_a estimates (Zhang et al., 2016). In this study, annual P data from ERA5 (it performed the best at annual timescale) were

considered the best after point-to-pixel evaluation of the P products. However, For the E_p component, this study evaluated the performance of two E_p products including ERA5 and GLEAMv3.3a based on their ability to capture the aridity in the study area. The aridity results were compared to the climate classification of Köppen-Geiger (Beck et al., 2018) and also to the literature that has documented the aridity in the study area. The E_p product that best captured the aridity appeared to be ERA5; and therefore, this study used this product to calculate the E_a_WB for the major catchments within the study area. Finally, the calculated E_a_WB was used to evaluate the performance of the E_a products. As a result, once again ERA5 E_a product appeared to outperform the GLEAMv3.3a and the SSEBopv4.0 products. In other words, this study calculated aridity using E_p and P both from ERA5 and calculated E_a_WB . The estimated E_a_WB was then used to evaluate the performance of E_a products where once again ERA5 was the best performing E_a product. Thus, the evaluation of the E_a products could be biased towards ERA5 because both P and E_p were used to evaluate the E_a products. However, there are no ground-based measurements to evaluate these products more straightforwardly.

5.4 Overall performance of the E_a products

The results of the E_a products evaluation (Figure 7) revealed that generally, the products failed to reproduce the E_a_WB , as they tended to overestimate the E_a_WB in the majority of the catchments (except for GLEAM3.3a for a few catchments which were underestimated). Bai and Liu (2018), also reported an overestimation of the E_a products across the continental China. They attributed the relative failure of the E_a products to close the water balance to inherent uncertainties in the water balance approach itself, uncertainties in the algorithms of the E_a products, and the input variables used to develop the products.

However, in line with the results obtained by Mahto and Mishra (2019), the evaluation of the E_a products revealed that ERA5 outperformed the other two products (GLEAMv3.3a and SSEBopv4.0) in most of the catchments. Either it was the best performing product, or it showed comparable results in most of the catchments except for those located in the north-eastern part of the province where GLEAMv3.3a outperformed ERA5. The better performance of GLEAMv3.3a over these catchments could be attributed to the fact that these catchments receive relatively higher P compared to the rest of the catchments in the study area which was also observed by Bai and Liu (2018) who reported better performance of GLEAMv3.2 over the wetter catchments in China. In this way, SSEBopv4.0 also presented a relatively good performance in the north-western region. The improved performance of SSEBopv4.0 over this region could be attributed to its less complex topography as noted by Saxe et al. (2020) who found that SSEBopv4.0 performed better in regions with less pronounced topography across the continental United States. Therefore, our results suggest that it is difficult to point out a single best-performing product across the province; however, ERA5 performed the best, followed by SSEBopv4.0 and GLEAMv3.3a in most of the catchments in the study area.

6 Conclusions

P and E_a are the two major components of the water cycle. An accurate spatio-temporal evaluation of these components remains a global challenge. However, since the availability of satellite remote sensing data and the reanalysis datasets, it has been relatively easy to account for these climatological variables in near-real-time and make observations even in the most underdeveloped and difficult to access regions around the world. This study

evaluated the performance of 4 state-of-the-art P (CHIRPSv2, MSWEPv2.2, PERSIANN-CDR, and ERA5) and 3 state-of-the-art E_a (GLEAMv3.3a, ERA5, and SSEBopv4.0) products in the arid and semi-arid province of Balochistan, Pakistan. This study evaluated the performance of these P products employing a point-to-pixel approach at the monthly, annual, and seasonal (winter, pre-monsoon, monsoon, and post-monsoon) temporal scales for 2000–2017. The KGE' along with its three individual components, i.e., r (linear correlation), β , (bias ratio) and γ (variability ratio), and the RMSE, were used as statistical indices to measure the performance of the products. Due to the unavailability of ground-based E_a data, this study employed the widely used Budyko framework and calculated E_{a_WB} for 12 major catchments in the study area. Subsequently, the E_{a_WB} was used as the reference to evaluate the performance of the E_a products.

The following paragraphs summarize the key findings of the study:

- MSWEPv2.2 closely followed by ERA5 presented the best spatio-temporal KGE' performance for most of the stations and across the province. CHIRPSv2 and PERSIANN-CDR relatively showed poor KGE' performance.
- ERA5 showed the best KGE' performance at the annual temporal scale and therefore was used as the precipitation product for calculating E_{a_WB} in the Fu's equation.
- RMSE results were different than those of the KGE'. Generally, CHIRPSv2 and PERSIANN-CDR outperformed MSWEPv2.2 and ERA5 at the temporal scales. Similar results were obtained for the spatial scale at the wetter stations that are located in eastern Balochistan. However, MSWEPv2.2 and ERA5 presented better spatial RMSE performance in the western drier stations.
- In general, the product that directly incorporate satellite, reanalysis, and daily gauge data (i.e., MSWEPv2.2) performed the best followed by the reanalysis product (i.e., ERA5), and the product that incorporate satellite, reanalysis, and monthly gauge data (i.e., CHIRPSv2). Whereas the satellite-based product that indirectly incorporate gauge data through other datasets (i.e., PERSIANN-CDR) performed the worst.
- Generally, all the P products performed the best at the monthly temporal scale compared to the annual.
- For the seasonal evaluation, the P products performed the best during winter followed by pre-monsoon, monsoon, and post-monsoon seasons.
- The results of this study confirm that it is difficult to point out a single best-performing P product that performs the best across regions with different climates and topographical gradients and at different temporal scales. However, MSWEPv2.2 performed the best followed closely by ERA5.
- The KGE' along with its three individual components (i.e., r , β , and γ) appears to be an improved performance index for evaluation of the P products compared to the RMSE due to its ability to decompose the total performance of the products into r (reproduces temporal dynamics), β (preserves the volume) and γ (preserves the distribution).

- The results for evaluation of the E_a products revealed that generally, the products failed to reproduce the E_a_{WB} of the selected catchments. However, ERA5 outperformed the other two products — i.e., GLEAMv3.3a and SSEBopv4.0 — in most of the catchments of the study area except for the north-eastern wetter catchments, where GLEAMv3.3a was the best performing product. SSEBopv4.0 presented the best performance over the north-western catchment of the K.P.L. closed basin.
- The E_a_{WB} values obtained using Fu's equation were not in good agreement with those of the E_a products in most of the catchments in the study area. This could be attributed to two main reasons: (i) the uncertainties from the Fu's equation, i.e., the soil water storage in the catchments is negligible, the catchments are natural, and the Fu's ω parameter represents the non-stationary characteristics (e.g., NDVI, slope, land cover, etc.) of the catchments, and (ii) the E_a products are not able to represent the E_a values over these data-scarce and arid catchments.

Finally, this study concludes that although there are certain underlying limitations in the evaluation results in the present study including the various assumptions discussed in the methodology section — sparsely distributed ground P measurements and unavailability of ground E_a measurements in the study area — are the major obstacles to study the water resources in the study area (Ahmed et al., 2019) — the following concrete findings of this study can be withdrawn and considered as important to the concerned decision-makers or might contribute in management of the water resources in the study area:

- MSWEPv2.2 and ERA5 presented a relatively good performance ($0.6 < KGE' \leq 0.8$) across the province of Balochistan and during most of the timescales with a few exceptions. However, a site-specific evaluation of these products is still recommended for any hydrological applications.
- For the E_a products, ERA5 generally outperformed the other two products — i.e., GLEAMv3.3a and SSEBopv4.0 — in most of the catchments except the north-eastern wetter catchments, where GLEAMv3.3a performed the best, and the north-western catchment of the K.P.L. closed basin, where SSEBopv4.0 presented the best performance. However, we recommend an evaluation of these products for specific applications in catchments within the study area.

Moreover, this study recommends that future studies might calibrate the P , E_a , and E_p products with local observations to achieve improved performance of these products (Baez-Villanueva et al., 2020; Dembélé et al., 2020). Furthermore, Fu's ω parameter could also be estimated and calibrated locally for better performance of Budyko framework as observed by Gunkel and Lange (2017) who reported improved estimation of renewable water resources in Jordan River Basin after estimating and calibrating the ω values locally. Similarly, future studies could also take into account the change in water storage (ΔS at catchment scale (relying on remotely-sensed data) as done by Bastiaanssen et al. (2014) to estimate improved multi-annual water balance at the catchment scale.

Statements & Declarations

Funding

The authors declare that no funding, grants or any other financial support was received for conducting this study.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Authors' contributions

All authors contributed to the study's conception and design. Data preparation, data analysis, and writing the original draft was done by AJ. OMBV guided the methodology, helped in writing and reviewing the manuscript, data preparation and analysis. LR was the supervisor of the study.

Availability of data and material

The dataset generated during and/or analyzed during this study can be categorized into two sections

1. Gauge data

The gauge data are available from Pakistan Meteorological Department, Meteorological Complex University Road, Karachi-75270, Pakistan but restrictions apply to the availability of these data. For the current study, the data were requested from the aforementioned department and are not publicly available.

2. Open source remotely sensed and reanalysis datasets

The following repositories were used to download the precipitation and evaporation data

CHIRPSv2: https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_monthly/tifs/

MSWEPv2.2: <http://www.gloh2o.org/mswep/>

CHRS data repository: <http://chrsdata.eng.uci.edu/>

ERA5: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form>

GLEAM: www.gleam.eu

USGS data repository: <https://earlywarning.usgs.gov/fews/datadownloads>; <https://earthexplorer.usgs.gov/>

Code availability

Not applicable

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent for publication

All Authors agree to publish this

7 References

- Adnan, S., Mahmood, R., & Khan, A. H. (2009). Water balance conditions in rainfed areas of Potohar and Balochistan plateau during 1931-08. *World Applied Sciences Journal*, 7(2), 162–169.
- Aftab, S. M., Siddiqui, R. H., & Farooqui, M. A. (2018). Strategies to manage aquifer recharge in Balochistan, Pakistan: An overview. In *IOP Conference Series: Materials Science and Engineering* (Vol. 414). <https://doi.org/10.1088/1757-899X/414/1/012023>
- Ahmed, K., Shahid, S., Ali, R. O., Harun, S. Bin, & Wang, X. J. (2017). Evaluation of the performance of gridded precipitation products over Balochistan province, Pakistan. *Desalination and Water Treatment*, 79(January), 73–86. <https://doi.org/10.5004/dwt.2017.20859>
- Ahmed, K., Shahid, S., & Haroon, S. Bin. (2014). Spatial interpolation of climatic variables in a predominantly arid region with complex topography. *Environment Systems and Decisions*, 34, 555–563. <https://doi.org/10.1007/s10669-014-9519-0>
- Ahmed, K., Shahid, S., Haroon, S. Bin, & Wang, X. (2015a). Multilayer perceptron neural network for downscaling rainfall in arid region: A case study of Baluchistan, Pakistan. *Journal of Earth System Science*, 124(6), 1325–1341. <https://doi.org/10.1007/s12040-015-0602-9>
- Ahmed, K., Shahid, S., Harun, S. Bin, & Wang, X. (2015b). Characterization of seasonal droughts in Balochistan Province, Pakistan. *Stochastic Environmental Research and Risk Assessment*, 30(2), 747–762. <https://doi.org/10.1007/s00477-015-1117-2>
- Ahmed, K., Shahid, S., Ismail, T., Nawaz, N., & Wang, X. J. (2018). Absolute homogeneity assessment of precipitation time series in an arid region of Pakistan. *Atmosfera*, 31(3), 301–316. <https://doi.org/10.20937/ATM.2018.31.03.06>
- Ahmed, K., Shahid, S., Wang, X., & Nawaz, N. (2019). Evaluation of gridded precipitation datasets over arid regions of Pakistan. *Water*, 11, 210. <https://doi.org/10.3390/w11020210>
- Alamgir, M., Shahid, S., Hazarika, M. K., Nashrullah, S., Harun, S. Bin, & Shamsudin, S. (2015). Analysis of meteorological drought pattern during different climatic and cropping seasons in Bangladesh. *Journal of the American Water Resources Association*, 51(3), 794–806. <https://doi.org/10.1111/jawr.12276>
- Anjum, S. A., Saleem, M. F., Chema, M. A., Bilal, M. F., & Kahliq, T. (2012). An assessment to vulnerability, extent, characteristics and severity of drought hazard in Pakistan. *Pakistan Journal of Science*, 64(2), 138–143.
- Ashouri, H., Hsu, K.-L., Sorooshian, S., & Braithwaite, D. K. (2015). PERSIANN-CDR: Daily precipitation climate data record from multi-satellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, 96, 6983. <https://doi.org/http://dx.doi.org/10.1175/BAMS-D-13-00068.1>

- Ashraf, M., Routray, J. K., & Saeed, M. (2014). Determinants of farmers' choice of coping and adaptation measures to the drought hazard in northwest Balochistan, Pakistan. *Natural Hazards*, 73(3), 1451–1473. <https://doi.org/10.1007/s11069-014-1149-9>
- 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>
- Baez-Villanueva, O. M., Zambrano-Bigiarini, M., Beck, H. E., McNamara, I., Ribbe, L., Nauditt, A., Birkel, C., Verbist, K., Giraldo-Osorio, J. D., & Xuan Thinh, N. (2020). RF-MEP: A novel Random Forest method for merging gridded precipitation products and ground-based measurements. *Remote Sensing of Environment*, 239, 111606. <https://doi.org/10.1016/j.rse.2019.111606>
- Bai, P., & Liu, X. (2018). Intercomparison and evaluation of three global high-resolution evapotranspiration products across China. *Journal of Hydrology*, 566(October), 743–755. <https://doi.org/10.1016/j.jhydrol.2018.09.065>
- Baloch, M. A., & Tanik, A. (2008). Development of an integrated watershed management strategy for resource conservation in Balochistan province of Pakistan. *Desalination*, 226(September 2006), 38–46. <https://doi.org/10.1016/j.desal.2007.02.098>
- Balochistan Provincial Disaster Management Authority. (2020). Balochistan Profile. Retrieved February 6, 2020, from <https://www.pdma.gob.pk/balochistan-profile/>
- Bastiaanssen, W. G. M., Karimi, P., Rebelo, L. M., Duan, Z., Senay, G., Muthuwatte, L., & Smakhtin, V. (2014). Earth observation based assessment of the water production and water consumption of Nile basin agro-ecosystems. *Remote Sensing*, 6(11), 10306–10334. <https://doi.org/10.3390/rs61110306>
- Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., ... Wood, E. F. (2019a). 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://doi.org/10.5194/hess-23-207-2019>
- Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & De Roo, A. (2017a). MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21, 589–615. <https://doi.org/10.5194/hess-21-589-2017>
- Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., ... Wood, E. F. (2017b). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, 21(12), 6201–6217. https://doi.org/10.1007/978-3-030-35798-6_9
- Beck, H. E., Wood, E. F., McVicar, T. R., Zambrano-Bigiarini, M., Alvarez-Garreton, C., Baez-Villanueva, O. M., ... Karger, D. N. (2020). Bias correction of global high-resolution precipitation climatologies using streamflow observations from 9372 catchments. *Journal of Climate*, 33(4), 1299–1315. <https://doi.org/10.1175/JCLI-D-19-0332.1>

- Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., ... Adler, R. F. (2019b). 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>
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*, 1–12. <https://doi.org/10.1038/sdata.2018.214>
- Brenner, C., Zeeman, M., Bernhardt, M., & Schulz, K. (2018). Estimation of evapotranspiration of temperate grassland based on high-resolution thermal and visible range imagery from unmanned aerial systems. *International Journal of Remote Sensing*, *39*(15–16), 5141–5174. <https://doi.org/10.1080/01431161.2018.1471550>
- Brown, K. W., & Rosenberg, N. J. (1973). A resistance model to predict evapotranspiration and its application to a sugar beet field. *Agronomy Journal*, *65*(May), 341–347. <https://doi.org/10.2134/agronj1973.00021962006500030001x>
- Budyko, M. I. (1974). *Climate and life*. New York: Academic Press.
- Caracciolo, D., Deidda, R., & Viola, F. (2017). Analytical estimation of annual runoff distribution in ungauged seasonally dry basins based on a first order Taylor expansion of the Fu's equation. *Advances in Water Resources*, *109*, 320–332. <https://doi.org/10.1016/j.advwatres.2017.09.019>
- Chen, F. W., & Liu, C. W. (2012). Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. *Paddy and Water Environment*, *10*(3), 209–222. <https://doi.org/10.1007/s10333-012-0319-1>
- Chen, L., & Dirmeyer, P. A. (2016). Impacts of land-use/land-cover change on afternoon precipitation over North America. *Journal of Climate*, *30*, 2121–2140. <https://doi.org/10.1175/JCLI-D-16-0589.1>
- Cheng, L., Xu, Z., Wang, D., & Cai, X. (2011). Assessing interannual variability of evapotranspiration at the catchment scale using satellite-based evapotranspiration data sets. *Water Resources Research*, *47*(9), 1–11. <https://doi.org/10.1029/2011WR010636>
- Choudhury, B. J. (1999). Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model. *Journal of Hydrology*, *216*(1–2), 99–110. [https://doi.org/10.1016/S0022-1694\(98\)00293-5](https://doi.org/10.1016/S0022-1694(98)00293-5)
- Copernicus Climate Change Service (C3S). (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Retrieved June 15, 2020, from <https://cds.climate.copernicus.eu/cdsapp#!/home>
- de Paula, A. C. P., da Silva, C. L., Rodrigues, L. N., & Scherer-Warren, M. (2019). Performance of the SSEBop model in the estimation of the actual evapotranspiration of soybean and bean crops. *Pesquisa Agropecuaria Brasileira*, *54*. <https://doi.org/10.1590/S1678-3921.PAB2019.V54.00739>

- Dembélé, M., Ceperley, N., Zwart, S. J., Salvadore, E., Mariethoz, G., & Schaepli, B. (2020). Potential of satellite and reanalysis evaporation datasets for hydrological modelling under various model calibration strategies. *Advances in Water Resources*, 143(March). <https://doi.org/10.1016/j.advwatres.2020.103667>
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). *Maximum Likelihood from Incomplete Data Via the EM Algorithm*. *Journal of the Royal Statistical Society: Series B (Methodological)* (Vol. 39). <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>
- dos Reis, J. B. C., Rennó, C. D., & Lopes, E. S. S. (2017). Validation of satellite rainfall products over a mountainous watershed in a humid subtropical climate region of Brazil. *Remote Sensing*, 9(12), 1240. <https://doi.org/10.3390/rs9121240>
- Du, C., Sun, F., Yu, J., & Liu, X. (2016). New interpretation of the role of water balance in an extended Budyko hypothesis in arid regions. *Hydrology and Earth System Sciences*, 20(1), 393–409. <https://doi.org/10.5194/hess-20-393-2016>
- Duan, Z., Bastiaanssen, W. G. M., & Liu, J. (2012). Monthly and annual validation of TRMM multisatellite precipitation analysis (TMPA) products in the Caspian Sea region for the period of 1999–2003. In *2012 IEEE International Geoscience and Remote Sensing Symposium* (pp. 3696–3699). Munich: IEEE. <https://doi.org/https://doi.org/10.1109/IGARSS.2012.6350613>
- Duan, Z., Liu, J., Tuo, Y., Chiogna, G., & Disse, M. (2016). Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. *Science of the Total Environment*, 573, 1536–1553. <https://doi.org/10.1016/j.scitotenv.2016.08.213>
- ECMWF. (2020). ERA5 hourly data on single levels from 1979 to present. Retrieved May 20, 2020, from <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview>
- Emam, A. R., Kappas, M., & Abbaspour, K. C. (2015). Simulation of water balance components in a watershed located in terrestrial drainage basin of Iran. In V. Lakshmi, D. Alsdorf, M. Anderson, S. Biancamaria, M. Cosh, J. Entin, ... C. Rüdiger (Eds.), *Remote Sensing of the Terrestrial Water Cycle* (pp. 463–478). Washington DC and New Jersey: American Geophysical Union and Wiley & Sons, Inc. Retrieved from [https://books.google.de/books?hl=en&lr=&id=B7wkBQAAQBAJ&oi=fnd&pg=PA463&dq=simulation+of+water+balance+components+in+a+watershed+located+in+central+drainage+basin+of+iran&ots=WzR9NXARMT&sig=Vc9ozSAQAM9fQtdqdxWI7YBNjkk&redir_esc=y#v=onepage&q=simulation of](https://books.google.de/books?hl=en&lr=&id=B7wkBQAAQBAJ&oi=fnd&pg=PA463&dq=simulation+of+water+balance+components+in+a+watershed+located+in+central+drainage+basin+of+iran&ots=WzR9NXARMT&sig=Vc9ozSAQAM9fQtdqdxWI7YBNjkk&redir_esc=y#v=onepage&q=simulation%20of)
- ESRI. (2018). ArcMap 10.6.1. Redlands: Environmental Systems Research Institute. Retrieved from <https://desktop.arcgis.com/en/arcmap/>
- Firat, M., Dikbas, F., Koç, A. C., & Gungor, M. (2010). Missing data analysis and homogeneity test for Turkish precipitation series. *Sadhana*, 35(6), 707–720. <https://doi.org/10.1007/s12046-010-0051-8>
- Fu, B. (1981). On the calculation of evaporation from land surface in mountainous areas (in Chinese). *Scientia Atmospherica Sinica*, 5(1), 23–31.
- Funk, C. C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., ... Verdin, A. P.

- (2014). A Quasi-Global Precipitation Time Series for Drought Monitoring. <https://doi.org/http://dx.doi.org/10.3133/ds832>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... Michaelsen, J. (2015). The climate hazards infrared precipitation with stations — a new environmental record for monitoring extremes. *Scientific Data*, 2(150066). <https://doi.org/10.1038/sdata.2015.66>
- Gao, H., Tang, Q., Ferguson, C. R., Wood, E. F., & Lettenmaier, D. P. (2010). Estimating the water budget of major US river basins via remote sensing. *International Journal of Remote Sensing*, 31(14), 3955–3978. <https://doi.org/10.1080/01431161.2010.483488>
- Gehne, M., Hamill, T. M., Kiladis, G. N., & Trenberth, K. E. (2016). Comparison of global precipitation estimates across a range of temporal and spatial scales. *Journal of Climate*, 29(21), 7773–7795. <https://doi.org/10.1175/JCLI-D-15-0618.1>
- Global Administrative Areas. (2012). GADM database of Global Administrative Areas, version 2.0. Retrieved June 6, 2020, from www.gadm.org.
- Gunkel, A., & Lange, J. (2017). Water scarcity, data scarcity and the Budyko curve—An application in the Lower Jordan River Basin. *Journal of Hydrology: Regional Studies*, 12(August), 136–149. <https://doi.org/10.1016/j.ejrh.2017.04.004>
- Guo, H., Chen, S., Bao, A., Hu, J., Gebregiorgis, A. S., Xue, X., & Zhang, X. (2015). Inter-comparison of high-resolution satellite precipitation products over Central Asia. *Remote Sensing*, 7(6), 7181–7211. <https://doi.org/10.3390/rs70607181>
- Gupta, V., Jain, M. K., Singh, P. K., & Singh, V. (2020). An assessment of global satellite-based precipitation datasets in capturing precipitation extremes: A comparison with observed precipitation dataset in India. *International Journal of Climatology*, 40(8), 3667–3688. <https://doi.org/10.1002/joc.6419>
- 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 modeling. *Journal of Hydrology*, 377, 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Haider, S., & Adnan, S. (2014). Classification and assessment of aridity over Pakistan provinces (1969-2009). *International Journal of Environment*, 3(4), 17–35. <https://doi.org/https://doi.org/10.3126/ije.v3i4.11728>
- Hersbach, H., De Rosnay, P., Bell, B., Schepers, D., Simmons, A., Soci, C., ... Zuo, H. (2018). *Operational global reanalysis: progress, future directions and synergies with NWP including updates on the ERA5 production status. ERA Report Series (Vol. 27)*. <https://doi.org/10.21957/tkic6g3wm>
- Hijmans, R. J. (2020). raster: Geographic Data Analysis and Modeling. Retrieved from <https://cran.r-project.org/web/packages/raster/index.html>
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., ... 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>

- Hugo, V., Espinoza-dávalos, G. E., Hessels, T. M., Moreira, D. M., Comair, G. F., & Bastiaanssen, W. G. M. (2019). The spatial variability of actual evapotranspiration across the Amazon River Basin based on remote sensing products validated with flux towers. *Ecological Processes*, 8(6). <https://doi.org/https://doi.org/10.1186/s13717-019-0158-8>
- Hussain, M. S., & Lee, S. (2014). Long-term variability and changes of the precipitation regime in Pakistan. *Asia-Pacific Journal of Atmospheric Sciences*, 50(271–282). <https://doi.org/10.1007/s13143-014-0015-8>
- Iqbal, M. F., & Athar, H. (2018). Validation of satellite-based precipitation over diverse topography of Pakistan. *Atmospheric Research*, 201, 247–260. <https://doi.org/10.1016/j.atmosres.2017.10.026>
- Istanbulluoglu, E., Wang, T., Wright, O. M., & Lenters, J. D. (2012). Interpretation of hydrologic trends from a water balance perspective: The role of groundwater storage in the Budyko hypothesis. *Water Resources Research*, 48, 1–22. <https://doi.org/10.1029/2010WR010100>
- Jaber, H. S. (2017). *Development of GIS-based SEBAL model for estimation of evapotranspiration in Iraq*. University Putra Malaysia. Retrieved from [http://psasir.upm.edu.my/id/eprint/71182/1/FK 2017 54 - IR.pdf](http://psasir.upm.edu.my/id/eprint/71182/1/FK%202017%2054-IR.pdf)
- Jin, X., Guo, R., & Xia, W. (2013). Distribution of actual evapotranspiration over Qaidam Basin, an arid area in China. *Remote Sensing*, 5, 6976–6996. <https://doi.org/10.3390/rs5126976>
- Katsanos, D., Retalis, A., Tymvios, F., & Michaelides, S. (2016). Analysis of precipitation extremes based on satellite (CHIRPS) and in situ dataset over Cyprus. *Natural Hazards*, 83, 53–63. <https://doi.org/10.1007/s11069-016-2335-8>
- 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>
- Long, D., Longuevergne, L., & Scanlon, B. R. (2014). Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. *Water Resources Research*, 50, 1131–1151. <https://doi.org/10.1002/2013WR014581>.Received
- Maggioni, V., Meyers, P. C., & Robinson, M. D. (2016). A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM). *Journal of Hydrometeorology*, 17, 1101–1117. <https://doi.org/10.1175/JHM-D-15-0190.1>
- Mahto, S. S., & Mishra, V. (2019). Does ERA-5 outperform other reanalysis products for hydrologic applications in India? *Journal of Geophysical Research: Atmospheres*, 124(16), 9423–9441. <https://doi.org/10.1029/2019JD031155>
- Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., ... Verhoest, N. E. C. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5), 1903–1925. <https://doi.org/10.5194/gmd-10-1903-2017>

- Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., ... Verhoest, N. E. C. (2016). GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development Discussions*, (August), 1–36. <https://doi.org/10.5194/gmd-2016-162>
- Matveeva, T., & Sidorchuk, A. (2020). Modelling of surface runoff on the Yamal Peninsula, Russia, using ERA5 reanalysis. *Water*, 12(8), 2099. <https://doi.org/10.3390/W12082099>
- McNamara, I., Baez-Villanueva, O. M., Zomorodian, A., Ayyad, S., Zambrano-Bigiarini, M., Zaroug, M., Mersha, A., Nauditt, A., Mbuliro, M., Wamala, S., & Ribbe, L. (2021). How well do gridded precipitation and actual evapotranspiration products represent the key water balance components in the Nile Basin? *Journal of Hydrology: Regional Studies*, 37(August), 100884. <https://doi.org/10.1016/j.ejrh.2021.100884>
- Mehta, L. (2000). *Water for the twenty-first century: Challenges and misconceptions* (ISBN. 1 85864 302 3 No. 111). Retrieved from https://entwicklungspolitik.uni-hohenheim.de/uploads/media/Water_for_the_Twenty-First_Century-ids-workingpaper_111_04.pdf
- Mianabadi, A., Alizadeh, A., Sanainejad, S. H., Ahahreman, B., & K, D. (2017). Assessment of the SEBAL algorithm to estimate actual evapotranspiration in Neishaboor-Rokh watershed using SWAT model. *Iranian Journal of Water Research in Agriculture*, 30(4), 525–541. Retrieved from <https://www.sid.ir/en/Journal/ViewPaper.aspx?ID=532185>
- Miralles, D. G., Brutsaert, W., Dolman, A. J., & Gash, J. H. (2020). On the use of the term “Evapotranspiration.” *Water Resources Research*, 56(11). <https://doi.org/10.1029/2020WR028055>
- Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., & Dolman, A. J. (2011). Magnitude and variability of land evaporation and its components at the global scale. *Hydrology and Earth System Sciences*, 15(3), 967–981. <https://doi.org/10.5194/hess-15-967-2011>
- Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, 15(2), 453–469. <https://doi.org/10.5194/hess-15-453-2011>
- Mu, Q., Heinsch, F. A., Zhao, M., & Running, S. W. (2007). Regional evaporation estimates from flux tower and MODIS satellite data. *Remote Sensing of Environment*, 111, 519–536. <https://doi.org/https://doi.org/10.1016/j.rse.2007.04.015>
- Mu, Q., Zhao, M., & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, 115(8), 1781–1800. <https://doi.org/10.1016/j.rse.2011.02.019>
- Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., ... Seneviratne, S. I. (2013). Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis. *Hydrology and Earth System Sciences*, 17(10), 3707–3720. <https://doi.org/10.5194/hess-17-3707-2013>
- Mustafa, D., & Qazi, M. U. (2007). Transition from karez to tubewell irrigation: development, modernization, and social capital in Balochistan, Pakistan. *World Development*, 35(10), 1796–1813. <https://doi.org/10.1016/j.worlddev.2007.06.002>
- Muthoni, F. K., Odongo, V. O., Ochieng, J., Mugalavai, E. M., Mourice, S. K., Hoesche-Zeledon, I., ... Bekunda, M. (2019). Long-term spatial-temporal trends and variability of rainfall over Eastern and

- Southern Africa. *Theoretical and Applied Climatology*, 137, 1869–1882. <https://doi.org/10.1007/s00704-018-2712-1>
- Nguyen, P., Shearer, E. J., Tran, H., Ombadi, M., Hayatbini, N., Palacios, T., ... Sorooshian, S. (2019). The CHRS Data Portal , an easily accessible public repository for PERSIANN global satellite precipitation data. *Scientific Data*, 6, 180296. <https://doi.org/10.1038/sdata.2018.296>
- Nicholson, S. E., Kotter, D., Zhou, L., & Hua, W. (2019). Validation of satellite precipitation estimates over the Congo Basin. *Journal of Hydrometeorology*, 20, 631–656. <https://doi.org/10.1175/JHM-D-18-0118.1>
- Nogueira, M. (2020). Inter-comparison of ERA-5, ERA-interim and GPCP rainfall over the last 40 years: Process-based analysis of systematic and random differences. *Journal of Hydrology*, 583, 124632. <https://doi.org/10.1016/j.jhydrol.2020.124632>
- Nogueira, S. M. C., Moreira, M. A., & Volpato, M. M. L. (2018). Evaluating precipitation estimates from Eta, TRMM and CHRIPS data in the South-Southeast region of Minas Gerais State-Brazil. *Remote Sensing*, 10(3)(313). <https://doi.org/10.3390/rs10020313>
- Oliveira, P. T. S., Nearing, M. A., Moran, M. S., Goodrich, D. C., Wendland, E., & Gupta, H. V. (2014). Trends in water balance components across the Brazilian Cerrado. *Water Resources Research*, 50, 7100–7114. <https://doi.org/10.1002/2013WR015202>
- Paredes-Trejo, F. J., Barbosa, H. A., & Lakshmi Kumar, T. V. (2017). Validating CHIRPS-based satellite precipitation estimates in Northeast Brazil. *Journal of Arid Environments*, 139, 26–40. <https://doi.org/10.1016/j.jaridenv.2016.12.009>
- Planning and Development Department - Government of Balochistan. (2020). *Balochistan Comprehensive Development Strategy 2013-2020*. Quetta. Retrieved from [https://www.undp.org/content/dam/pakistan/docs/PROCUREMENT/2018/July 2018/2\) BCDS Revised Final Draft August 27 2013.pdf](https://www.undp.org/content/dam/pakistan/docs/PROCUREMENT/2018/July%202018/2%20BCDS%20Revised%20Final%20Draft%20August%2027%202013.pdf)
- Porporato, A., Daly, E., & Todriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. *Journal of Chemical Education*, 164(5), 625–632. https://doi.org/10.1007/978-3-642-61320-3_6
- Potter, N. J., & Zhang, L. (2009). Interannual variability of catchment water balance in Australia. *Journal of Hydrology*, 369(1–2), 120–129. <https://doi.org/10.1016/j.jhydrol.2009.02.005>
- Prasetia, R., As-syakur, A. R., & Osawa, T. (2013). Validation of TRMM Precipitation Radar satellite data over Indonesian region. *Theoretical and Applied Climatology*, 112(3–4), 575–587. <https://doi.org/10.1007/s00704-012-0756-1>
- Qureshi, S. A. (2005). Climate change and water resources management in Pakistan. In M. Q. Mirza & Q. K. Ahmed (Eds.), *Climate change and water resources in South Asia* (1st edition, pp. 197–230). Taylor & Francis. <https://doi.org/10.1201/9780203020777.ch8>
- R Core Team. (2011). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>

- Rasul, G., & Mahmood, A. (1993). Performance evaluation of different methods for estimation of evapotranspiration in Pakistan's climate. *Pakistan Journal of Meteorology*, 05(10), 1–12. Retrieved from https://climateinfo.pk/frontend/web/attachments/data-type/Vol5_Issue10.pdf#page=27
- Ryu, Y., Baldocchi, D. D., Kobayashi, H., Van Ingen, C., Li, J., Black, T. A., Beringer, J., Van Gorsel, E., Knohl, A., Law, B. E., & Rouspard, O. (2011). Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales. *Global Biogeochemical Cycles*, 25(4), GB4017. <https://doi.org/10.1029/2011GB004053>
- Sadiq, N. (2013). Thunderstorm and rainfall frequencies over Sindh, Pakistan for the period 1961-2010. *The Nucleus*, 50(2), 125–133.
- Sarwar, A. (2008). Droughts in Pakistan—a socio-political perspective. In J. Jairath & V. Ballabh (Eds.), *Droughts and integrated water resource management in South Asia: issues, alternatives and futures* (pp. 200–230). New Dehli: SAGE Publications. Retrieved from https://books.google.de/books?hl=en&lr=&id=YbmGAwAAQBAJ&oi=fnd&pg=PA200&dq=Droughts+in+Pakistan—a+socio-political+perspective&ots=5qWGWxxK6g&sig=TpxNlwp_St7kxowb7vt4RrIJ7Z0&redir_esc=y#v=onepage&q=Droughts+in+Pakistan—a+socio-political+perspective&f=false
- Satgé, F., Espinoza, R., Zolá, R. P., Roig, H., Timouk, F., Molina, J., Garnier, J., Calmant, S., Seyler, F., & Bonnet, M. P. (2017). Role of climate variability and human activity on Poopó Lake droughts between 1990 and 2015 assessed using remote sensing data. *Remote Sensing*, 9, 218. <https://doi.org/10.3390/rs9030218>
- Saxe, S., Farmer, W., Driscoll, J., & Hogue, T. (2020). Implications of model selection: A comparison of publicly available, CONUS-extent hydrologic component estimates. *Hydrology and Earth System Sciences Discussions*, (June), 1–70. <https://doi.org/10.5194/hess-2020-216>
- Senay, G B. (2018). Satellite psychrometric formulation of the Operational Simplified Surface Energy Balance (SSEBop) model for quantifying and mapping evapotranspiration. *Applied Engineering in Agriculture*, 34(3), 555–566. <https://doi.org/10.13031/aea.12614>
- Senay, G B, Budde, M. E., & Verdin, J. P. (2011). Enhancing the Simplified Surface Energy Balance (SSEB) approach for estimating landscape ET: Validation with the METRIC model. *Agricultural Water Management*, 98(4), 606–618. <https://doi.org/10.1016/j.agwat.2010.10.014>
- Senay, Gabriel B., Friedrichs, M., Singh, R. K., & Velpuri, N. M. (2016). Evaluating Landsat 8 evapotranspiration for water use mapping in the Colorado River Basin. *Remote Sensing of Environment*, 185, 171–185. <https://doi.org/10.1016/j.rse.2015.12.043>
- Senay, Gabriel B, Bohms, S., Singh, R. K., Gowda, P. H., Velpuri, N. M., Alemu, H., & Verdin, J. P. (2013). Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach. *Journal of the American Water Resources Association*, 49(3), 577–591. <https://doi.org/10.1111/jawr.12057>

- Shafiq, M., & Kakar, M. A. (2008). Effects of drought on livestock sector in Balochistan province of Pakistan. *International Journal of Agriculture and Biology (Pakistan)*, 9(4), 657–665. Retrieved from <http://agris.fao.org/agris-search/search.do?recordID=PK2007001176>
- Shafiq, Muhammad. (2008). Analysis of the role of women in livestock production in Balochistan , Pakistan. *Journal of Agriculture & Social Sciences*, 4, 18–22. Retrieved from https://www.researchgate.net/profile/Muhammad_Shafiq58/publication/228907326_Analysis_of_the_role_of_women_in_livestock_production_in_Balochistan_Pakistan/links/58891006aca272f628d25902/Analysis-of-the-role-of-women-in-livestock-production-in-Balochistan-
- Sharma, D. N., & Tare, V. (2018). Evapotranspiration estimation using SSEBop method with SENTINEL -2 and LANDSAT-8 data set. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42(5), 1–5. <https://doi.org/10.5194/isprs-archives-XLII-5-563-2018>
- Shen, Y., Xiong, A., Hong, Y., Yu, J., Pan, Y., Chen, Z., & Saharia, M. (2014). Uncertainty analysis of five satellite-based precipitation products and evaluation of three optimally merged multi-algorithm products over the Tibetan Plateau. *International Journal of Remote Sensing*, 35(19), 6843–6858. <https://doi.org/10.1080/01431161.2014.960612>
- 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, 79–107. <https://doi.org/10.1002/2017RG000574>
- Tang, Q., Gao, H., Lu, H., & Lettenmaier, D. P. (2009). Remote sensing: hydrology. *Progress in Physical Geography*, 33(4), 490–509. <https://doi.org/10.1177/0309133309346650>
- Tarek, M., Brissette, F. P., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences*, 24(5), 2527–2544. <https://doi.org/10.5194/hess-24-2527-2020>
- Tekleab, S., Uhlenbrook, S., Mohamed, Y., Savenije, H. H. G., Ayalew, S., Temesgen, M., & Wenninger, J. (2010). Water balance modeling of Upper Blue Nile catchments using a top-down approach. *Hydrology and Earth System Sciences Discussions*, 7(5), 6851–6886. <https://doi.org/10.5194/hessd-7-6851-2010>
- Torres-Batllo, J., Marti-Cardona, B., & Pillco-zola, R. (2020). Mapping evapotranspiration, vegetation and precipitation trends in the catchment of the shrinking Lake Poopó. *Remote Sensing*, 12(73). <https://doi.org/10.3390/rs12010073>
- Trezza, R. (2002). *Evapotranspiration using a satellite-based surface energy balance with standardized ground control*. Utah State University Logan, Utah. Retrieved from <https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1711&context=etd>
- Tuo, Y., Duan, Z., Disse, M., & Chiogna, G. (2016). Evaluation of precipitation input for SWAT modeling in Alpine catchment : A case study in the Adige river basin (Italy). *Science of the Total Environment*, 573, 66–82. <https://doi.org/10.1016/j.scitotenv.2016.08.034>
- UN General Assembly. (2015). *Transforming our world: the 2030 agenda for sustainable development*.

Retrieved from [https://sustainabledevelopment.un.org/content/documents/21252030 Agenda for Sustainable Development web.pdf](https://sustainabledevelopment.un.org/content/documents/21252030_Agenda_for_Sustainable_Development_web.pdf)

- UNDP. (2004). *Water Governance for Poverty Reduction: Key Issues and the UNDP Response to Millennium Development Goals*. New York. Retrieved from [https://www.undp.org/content/dam/aplaws/publication/en/publications/environment-energy/www-ee-library/water-governance/water-governance-for-poverty-reduction/UNDP_Water Governance for Poverty Reduction.pdf](https://www.undp.org/content/dam/aplaws/publication/en/publications/environment-energy/www-ee-library/water-governance/water-governance-for-poverty-reduction/UNDP_Water_Governance_for_Poverty_Reduction.pdf)
- United Nations Environment Programme. (1992). *World atlas of desertification*. (E. Arnold, Ed.). London ; Baltimore: United Nations Environment Programme. <https://doi.org/https://nla.gov.au/nla.cat-vn624121>
- U.S. Geological Survey. (n.d.). FEWS Data Downloads. Retrieved June 15, 2020, from <https://earlywarning.usgs.gov/fews/datadownloads>
- U.S. Geological Survey. (2002). Shuttle Radar Topography Mission (SRTM) Elevation Dataset. Retrieved June 7, 2020, from <https://earthexplorer.usgs.gov/>
- Voss, K. A., Famiglietti, J. S., Lo, M., De Linage, C., Rodell, M., & Swenson, S. C. (2013). Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. *Water Resources Research*, 49, 904–914. <https://doi.org/10.1002/wrcr.20078>
- Wu, S., Li, J., & Huang, G. H. (2008). A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution. *Applied Geography*, 28(3), 210–223. <https://doi.org/10.1016/j.apgeog.2008.02.006>
- Wu, W., Li, Y., Luo, X., Zhang, Y., Ji, X., & Li, X. (2019). Performance evaluation of the CHIRPS precipitation dataset and its utility in drought monitoring over Yunnan Province, China. *Geomatics, Natural Hazards and Risk*, 10(1), 2145–2162. <https://doi.org/10.1080/19475705.2019.1683082>
- Xian, L. U. O., Wenqi, W. U., Daming, H. E., Yungang, L. I., & Xuan, J. I. (2019). Hydrological simulation using TRMM and CHIRPS precipitation estimates in the Lower Lancang-Mekong river basin. *Chinese Geographical Science*, 29(1), 13–25. <https://doi.org/https://doi.org/10.1007/s11769-019-1014-6>
- 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>
- Xiong, L., & Guo, S. (2011). Appraisal of Budyko formula in calculating long-term water balance in humid watersheds of southern China. *Hydrological Processes*, 26(9), 1370–1378. <https://doi.org/10.1002/hyp.8273>
- Xu, W., Zou, Y., Zhang, G., & Linderman, M. (2014). A comparison among spatial interpolation techniques for daily rainfall data in Sichuan Province, China. *International Journal of Climatology*, 35(10), 2898–2907. <https://doi.org/10.1002/joc.4180>
- Xu, X., Liu, W., Scanlon, B. R., Zhang, L., & Pan, M. (2013). Local and global factors controlling water-energy

- balances within the Budyko framework. *Geophysical Research Letters*, 40(23), 6123–6129.
<https://doi.org/10.1002/2013GL058324>
- Yang, P., Xia, J., Zhang, Y., Zhan, C., & Qiao, Y. (2018). Comprehensive assessment of drought risk in the arid region of Northwest China based on the global palmer drought severity index gridded data. *Science of the Total Environment*, 627, 951–962. <https://doi.org/10.1016/j.scitotenv.2018.01.234>
- Zambrano-Bigiarini, M. (2020a). hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. Retrieved from <https://cran.r-project.org/web/packages/hydroGOF/index.html>
- Zambrano-Bigiarini, M. (2020b). hydroTSM: Time series management, analysis and interpolation for hydrological modeling. Retrieved from <https://cran.r-project.org/web/packages/hydroTSM/index.html>
- Zambrano-Bigiarini, M., & Baez-Villaneuva, O. M. (2019). Characterizing meteorological droughts in data scarce regions using remote sensing estimates of precipitation. In V. Maggioni & C. Massari (Eds.), *Extreme Hydroclimatic Events and Multivariate Hazards in a Changing Environment* (pp. 221–246). Elsevier Inc. <https://doi.org/10.1016/b978-0-12-814899-0.00009-2>
- 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, 1295–1320. <https://doi.org/10.5194/hess-21-1295-2017>
- Zhang, D., Liu, X., Zhang, Q., Liang, K., & Liu, C. (2016). Investigation of factors affecting intra-annual variability of evapotranspiration and streamflow under different climate conditions. *Journal of Hydrology*, 543, 759–769. <https://doi.org/10.1016/j.jhydrol.2016.10.047>
- Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., & Briggs, P. R. (2004). A rational function approach for estimating mean annual evapotranspiration. *Water Resources Research*, 40(2), 1–14. <https://doi.org/10.1029/2003WR002710>
- Zhang, Lu, Potter, N., Hickel, K., Zhang, Y., & Shao, Q. (2008). Water balance modeling over variable time scales based on the Budyko framework - Model development and testing. *Journal of Hydrology*, 360, 117–131. <https://doi.org/10.1016/j.jhydrol.2008.07.021>
- Zhang, W., Brandt, M., Guichard, F., Tian, Q., & Fensholt, R. (2017). Using long-term daily satellite based rainfall data (1983–2015) to analyze spatio-temporal changes in the Sahelian rainfall regime. *Journal of Hydrology*, 550, 427–440. <https://doi.org/10.1016/j.jhydrol.2017.05.033>

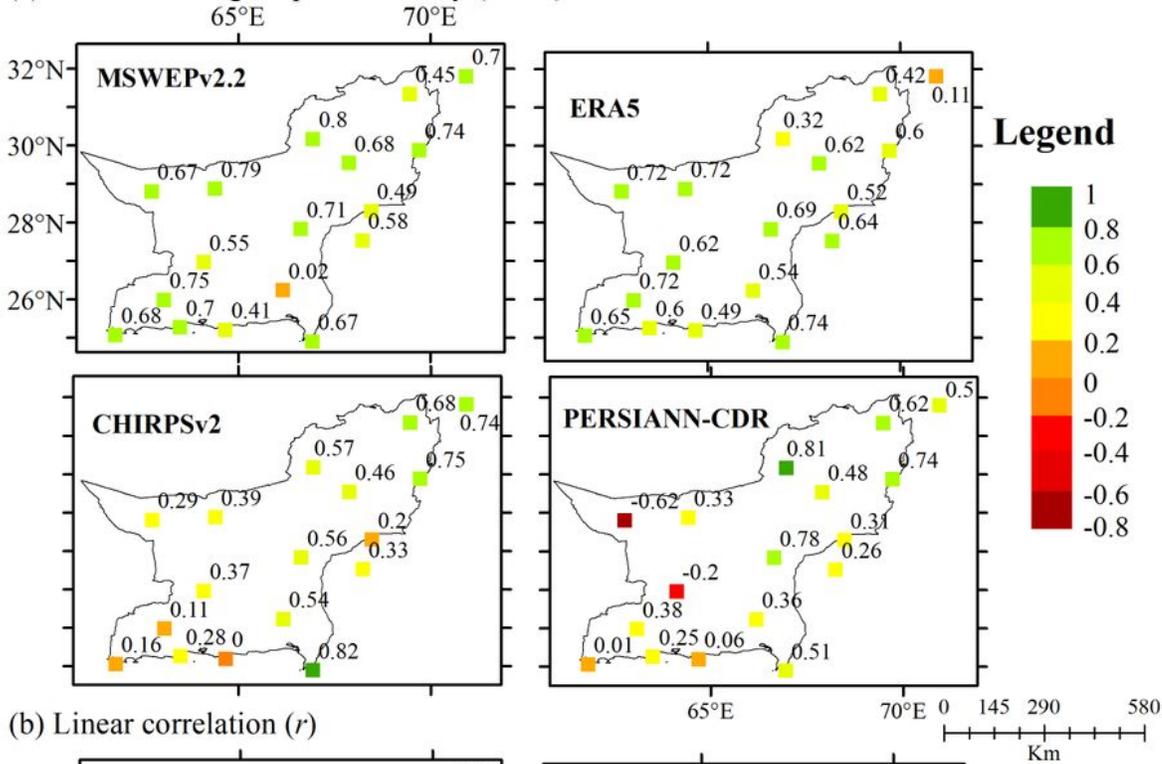
Figures

Figure 1

(a) Location of the study area. The figure also shows the location of the 14 selected rain gauge stations and the Digital Elevation Model (DEM) obtained from the US Geographical Survey's (USGS) Shuttle Radar Topography Mission (SRTM). (b) Location of the major catchments in the study area.

Modified Kling-Gupta Efficiency (KGE') and linear correlation (r) evaluation at monthly scale

(a) Modified Kling-Gupta Efficiency (KGE')



(b) Linear correlation (r)

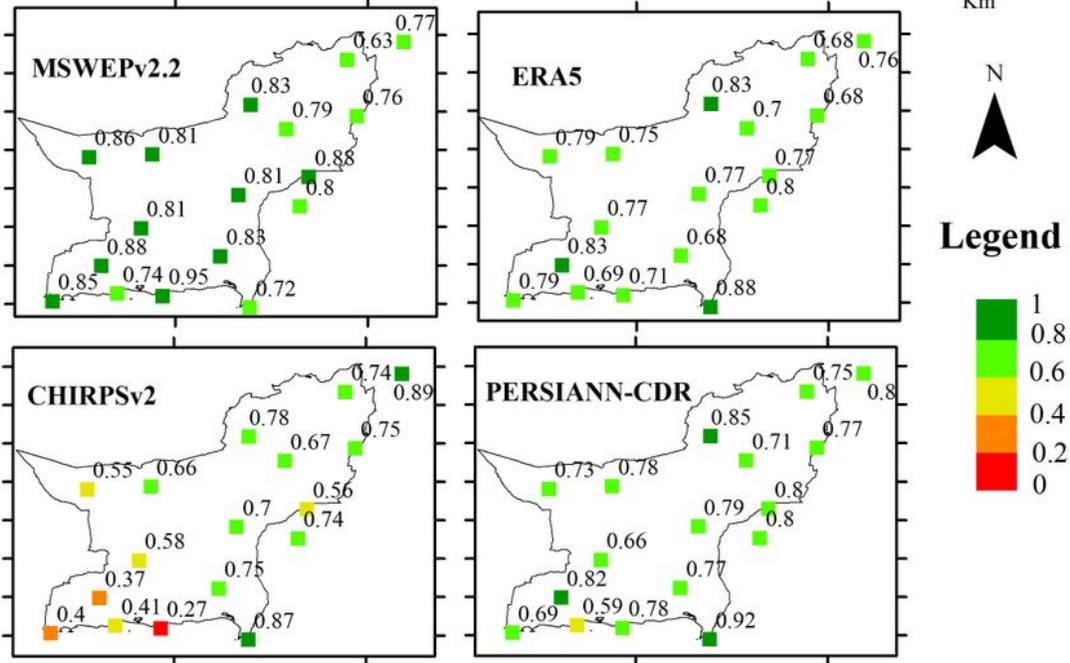
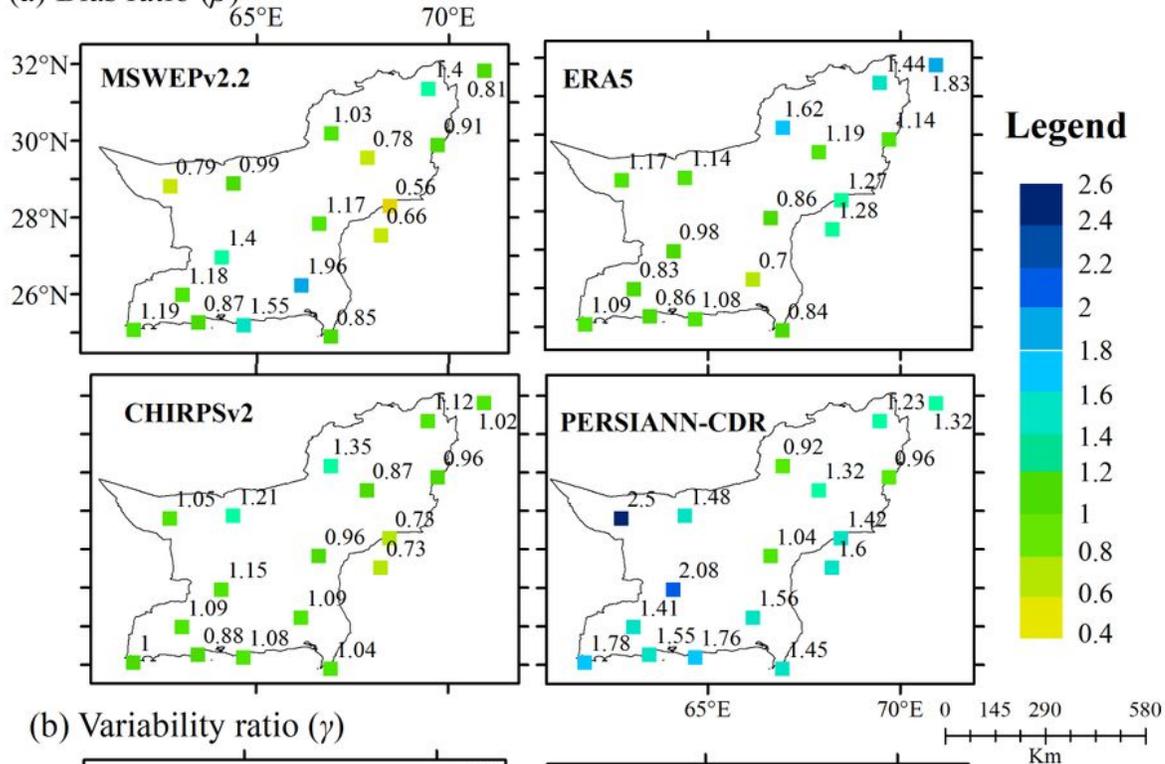


Figure 2

(a) Modified Kling-Gupta Efficiency (KGE') and (b) linear correlation (r) evaluation between monthly P products and ground-based observations at the corresponding grid cells.

Bias ratio (β) and variability ratio (γ) evaluation at monthly scale

(a) Bias ratio (β)



(b) Variability ratio (γ)

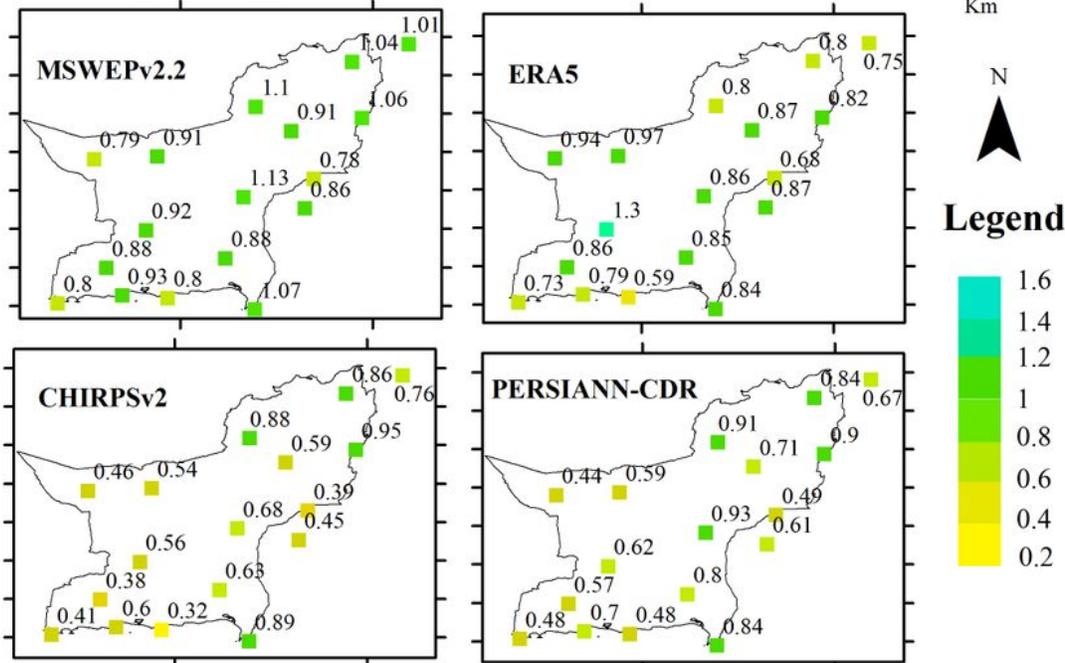


Figure 3

(a) Bias ratio (β) and (b) variability ratio (γ) evaluation between monthly P products and ground-based observations at the corresponding grid cells of the P products.

RMSE evaluation for the precipitation estimates at monthly scale

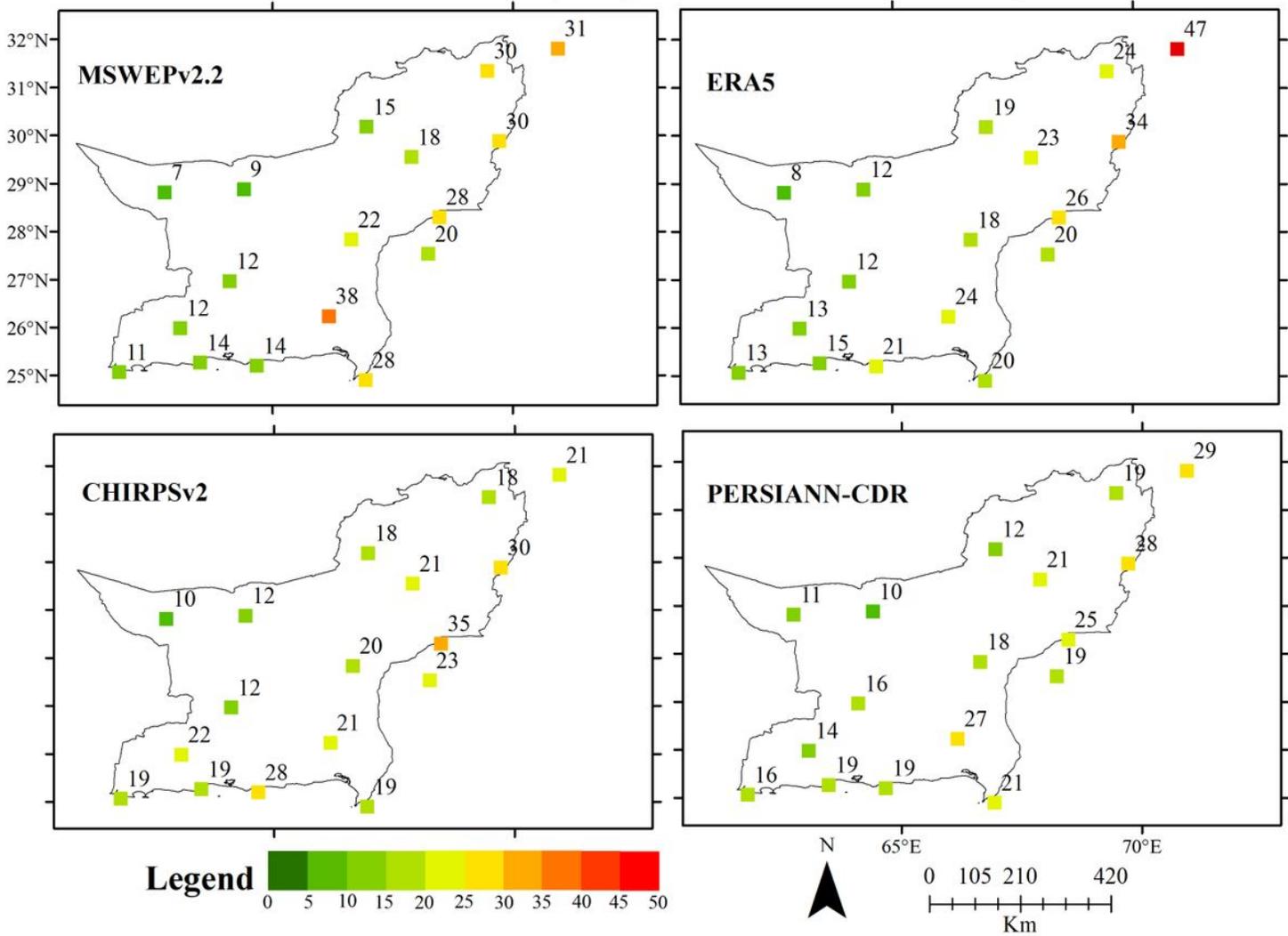


Figure 4

Root Mean Square Error (RMSE) evaluation between monthly P products and ground-based observations at the corresponding grid cell.

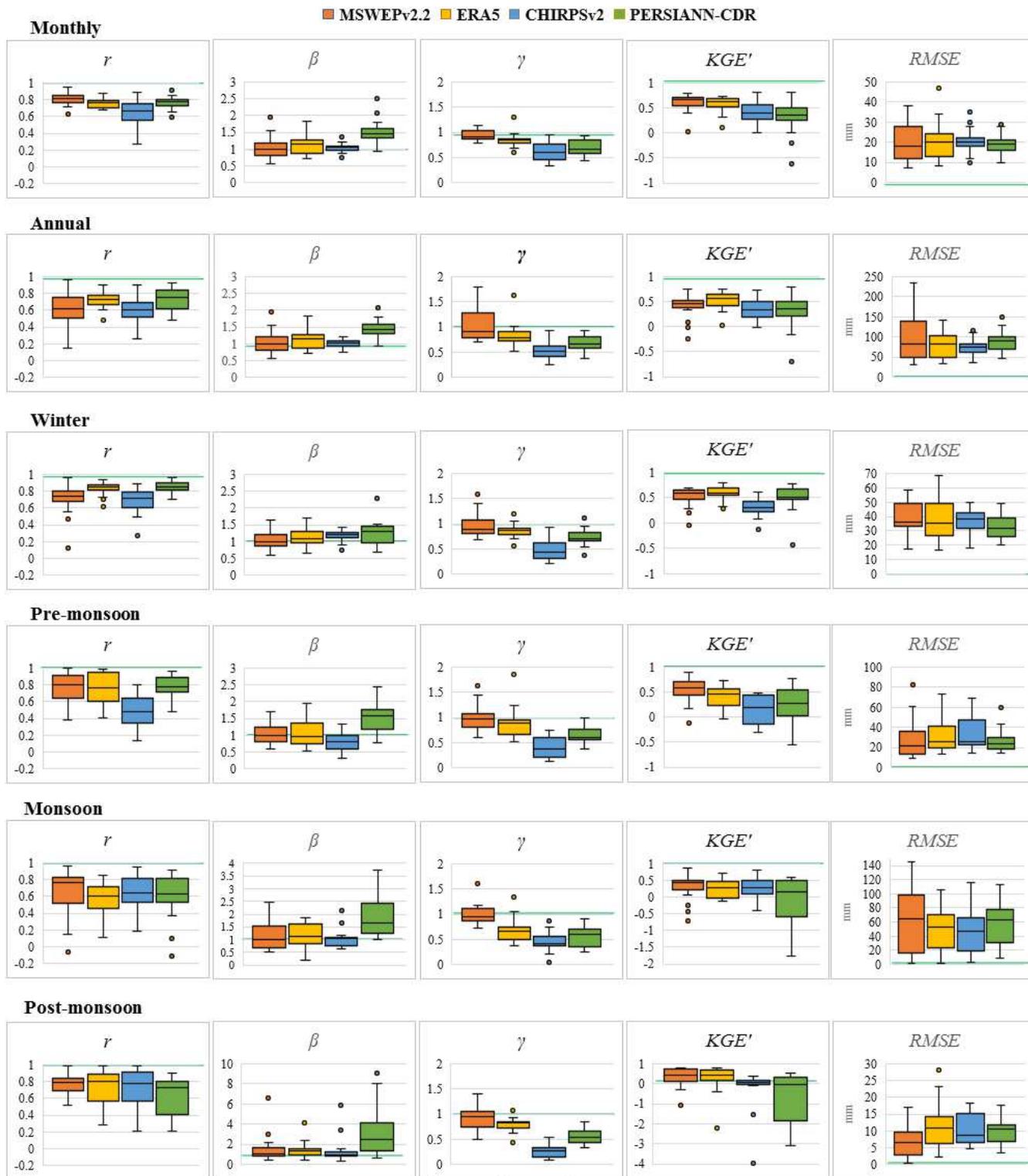


Figure 5

Modified Kling-Gupta Efficiency (KGE') along with its three individual components (r , β , and γ) and the Root Mean Square Error (RMSE) between satellite products and ground observations at the monthly, annual, and seasonal (winter, pre-monsoon, monsoon and post-monsoon) temporal scales. The horizontal green line indicates the optimum value of the performance indices.

Figure 6

The ratio of mean annual evaporation to mean annual precipitation (Ea/P) as a function of aridity index (Ep/P) for the selected catchments for 16 years (2003–2018). Scatter points represent the distribution of catchments over Budyko curve.

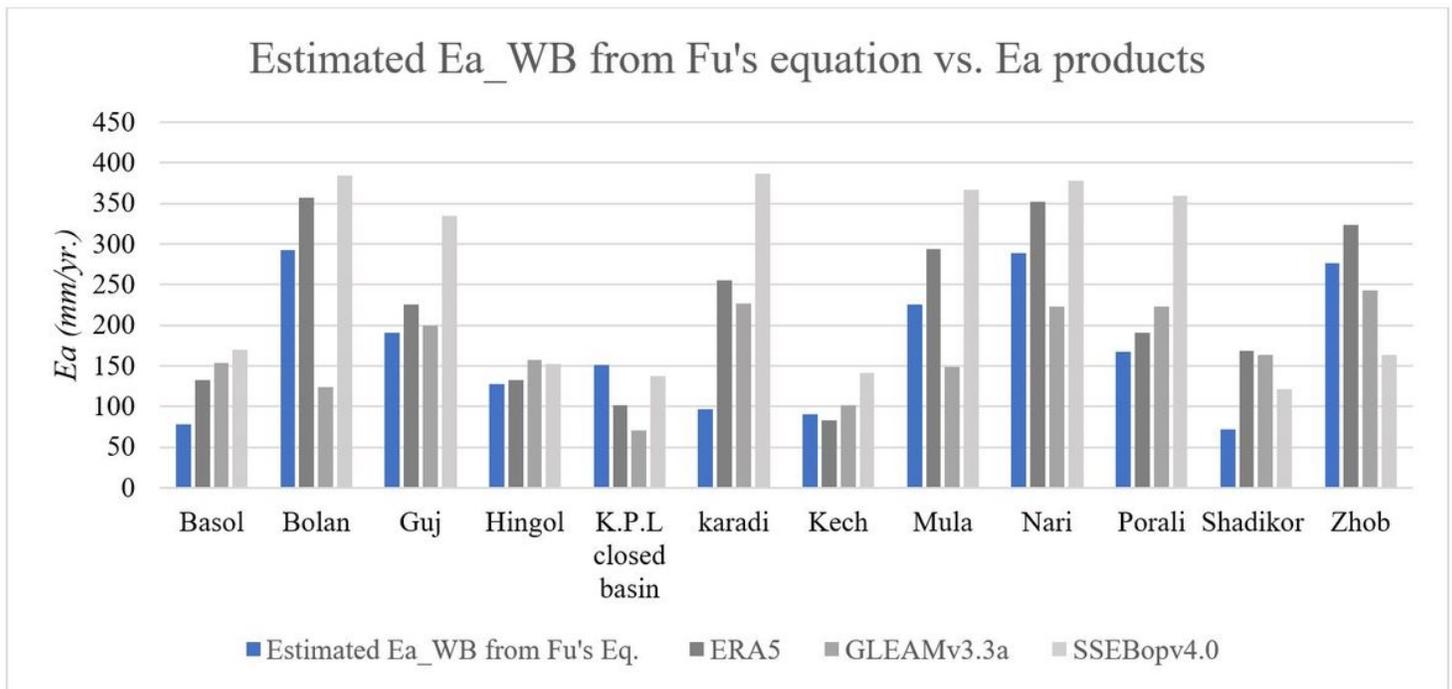


Figure 7

Comparison of multi-annual means (2003–2018) of estimated EaWB from Fu's equation to Ea products.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Suplimentarymaterial.docx](#)