Performance of Gridded Precipitation Products in the Black Sea Region for Hydrological Studies

SEAD AHMED SWALIH (seahhu@yahoo.com)
Istanbul Technical University https://orcid.org/0000-0003-3830-554X
Ercan Kahya
Istanbul Technical University

Research Article

Keywords: Gridded Precipitation, CFSR, ERA-Interim land, APHRODITE, MSWEP, SWAT, Black Sea

Posted Date: June 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-592842/v1

Version of Record: A version of this preprint was published at Theoretical and Applied Climatology on April 25th, 2022. See the published version at https://doi.org/10.1007/s00704-022-04054-z.

Performance of Gridded Precipitation Products in the Black Sea Region for hydrological studies

Sead Ahmed Swalih[1] and Ercan Kahya [2]

[1] Istanbul Technical University, Civil Engineering Department, Istanbul, Turkey, swalih@itu.edu.tr

[2] Istanbul Technical University, Civil Engineering Department, Istanbul, Turkey, kahyae@itu.edu.tr

(Corresponding author: *Sead Ahmed Swalih, swalih@itu.edu.tr, +905075457037*)

Abstract

Gridded precipitation products are becoming good alternative data sources for regions with limited weather gauging stations. In this study, four climate gridded precipitation products were utilized, namely: Climate Forecast System Reanalysis (CFSR), European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim/land, the Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE), and Multi-Source Weighted-Ensemble Precipitation (MSWEP). The key novelty of this study is to fill the gap in one of important areas of the transcontinental region of Eurasia, namely Rize province in the Black Sea region being selected as a study area since it has complex topography and climatology in addition to a limited number of gauging stations. A set of precipitation products were assessed for performance with the observed precipitation data before using a hydrological model (SWAT) to evaluate the basin response for the climate products. Three methods were considered in this study: (i) spatial comparison; (ii) hydrological, and (iii) statistical evaluation. Along with precipitation forcing, the SWAT model simulations were analysed in conjunction with streamflow observations. In an overall evaluation, the percentage bias of ERA-Interim/land, CFSR, APHRODITE, and MSWEP mean monthly precipitation is 19.9%, 33.4%, 41.4% and 85.0% respectively. For the flow simulations, the CFSR and MSWEP have resulted in exaggerated peak flows in the high flow season due to overestimated precipitation forcing (Nash Sutcliffe efficiency [NS] equal to 0.22 and -0.73, respectively). On the contrary, the APHRODITE underestimated the peak flows due to lower precipitation estimates (NS = 0.38). The ERA-Interim land showed good agreement with the observed flows (NS = 0.53). From these readings we stated that the ERA-Interim land exhibited improved performance with the observed precipitation whereas the CFSR showed the worst performance. The study suggests that gridded precipitation products

could supplement observed precipitation data for observational data scarcity in mountainous regions.

Keywords*: Gridded Precipitation, CFSR, ERA-Interim land, APHRODITE, MSWEP, SWAT, Black Sea*

Declarations

Funding – We would like to acknowledge TUBİTAK (Turkish Scientific and Technological Research Council) for providing the fund to accomplish this study. We are also indebted to the Turkish State Meteorological Service (MGM) and General Directorate of State Water Works (DSI) for the observed weather and flow data access. In addition, the license for ArcGIS software was provided by Istanbul Technical University (İTÜ).

Conflicts of interest/Competing interests – The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and material – Data used in this study could be provided upon request.

Code availability – Not applicable

Authors' contributions – Both authors have contributed in this manuscript preparation

Ethics approval – Not applicable

Consent to participate – Not applicable

Consent for publication – Not applicable

1. Introduction

A good quality of precipitation data is essential in all climatological and hydrological studies. Unless a sufficient number of gauging stations are available, it is difficult to assess the spatial and temporal variability of precipitation for a basin (Sharma et al., 2013). Insufficient density of observation stations is a major problem in many parts of the world. It is known that a poor

data coverage makes estimate of all hydrologic variables a difficult task. The World Meteorological Organization (WMO) recommends the distance between precipitation gauging stations at surface level not being more than 500km, and between stations measuring temperature, humidity, and wind at upper levels not to be greater than 1000km (WMO, 1956).

In recent years, sophisticated high-resolution satellite based reanalysis data and the merger of different precipitation products have been developed to help scientists and researchers in areas with sparse rain gauges. It is evident that such data could be an essential asset helping researchers who work with basins having a scarce coverage of precipitation stations (Bitew, 2011). It is important to emphasize that not all precipitation products are climate reanalysis data, but also merged observed and satellite precipitation data compiled from a number of meteorological organizations. Some of the well-known publicly available gridded precipitation data include: the National Centre for Environmental Prediction (NCEP) - Climate Forecast System Reanalysis (CFSR) (NCEP, 2015), the Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) (Yatagai, et al., 2012), Multi-Source Weighted-Ensemble Precipitation (MSWEP) (Beck et al., 2017), European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim land precipitation data (Balsamo et al., 2012; Balsamo et al., 2015), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Ashouri et al., 2014) and Integrated Multi-satellitE Retrievals for GPM (IMERG) (NASA, 2019); which are examples of precipitation data produced by the merger of different precipitation products. In this study, we used the first four precipitation products that all are known with state-of-the-art algorithms, as well as are freely available. In addition, these products have relatively high spatio-temporal resolutions for the study area at hand.

79 Reanalysis data and merger of different precipitation products have been increasingly used
80 recently by many researchers in hydro-climatological studies. For example, Renner et al.
81 (2009) presented the outcomes of flow prediction for the Rhine River in Europe through the
82 rainfall–runoff models with probabilistic weather forecasting. He reiterated the importance of
83 downscaling climate forecast products prior to enforcing rainfall-runoff models in flow
84 forecasting. On the other hand, Verkade et. al, (2013) advised to remove the bias in the forcing
85 ensemble weather products containing significant bias that cascades into the flow forecasts.
86 Richard (2008) compared two reanalysis data (ECMWF-ERA40 and NCEP-NDRa2)
87 globally, and detected significant disagreements between the two products in regions of the
88 world with higher topography. Feidas et al., 2009 and Feidas 2010 validated a number of
89 Tropical Rainfall Measuring Mission (TRMM) precipitation products for the Mediterranean
90 region. Inconsistency of the CFSR precipitation products was reported after a comparative
91 study of model and satellite-derived long-term precipitation data for Mediterranean basins
92 (Aznar, 2010). Moreover, Manzato (2015) compared ECMWF-ERA reanalysis data with 104
93 gauging stations data in Italy. He found out that ECMWF significantly underestimates
94 precipitation and the results of Bias and Root Mean Square Error (RMSE) were not acceptable.
95 Husain et al. (2017) evaluated two gridded precipitation datasets (TMPA and APHRODITE)
96 in the Himalaya mountainous basin. He found that the biggest error in the gridded data arose
97 mainly from elevation and TMPA dataset showed poor correlation with ground observations
98 especially for higher altitudes.
99 In the western hemisphere, Ward et al. (2011) analysed spatially averaged precipitation
100 products (ERA-interim land, NCEP-R1 hindcast, PERSIANN, and TRMM reanalysis
101 products) and compared with average precipitation observations in the Ecuadorian Andes and
102 in Chilean Patagonia. They concluded that the observations were consistent over the two study
103 regions which made extrapolation to other mountainous regions possible. In a number of

basins on the North-American continent, reanalysis weather data (CFSR and ERA-Interim land) were used to force hydrological models. They found that the reanalysis data cou successfully compensate the deficiency of surface observation records (Essou et al., 2016, Essou et al. 2017a, and Essou et al., 2017b). In the case of hydrological modelling for regions with high spatial variability of precipitation such as mountainous regions, these studies also pointed out that reanalysis data have good performance when the density of the observation station network is low.

Moving to the east side of the globe, Kamiguchi et al., (2010) constructed historical daily precipitation data with high-resolution (0.05° × 0.05°) over Japanese land area as part of the product of the APHRODITE project, which were used for validation of basin models. After examining long-term extreme precipitation trends in Japan, they found that a change in gauge density does not affect the trend of total precipitation, but does the trend of extreme values in Japan. In a study by Zhu et al. (2017), the performance of two reanalysis datasets (the twentieth century reanalysis - 20CR and ERA-Interim land) were evaluated in the context of reproducing the persistent weather extremes in China. They showed that the two datasets capture the intensity indices better than the frequency indices of the weather. They also found that the ERA Interim land reanalysis data is able to depict the relationship among persistent precipitation extremes and local persistent temperature extremes.

At last, Worqlul et al. (2014) used three weather products, namely TRMM product 3B42, Multi-Sensor Precipitation Estimate-Geostationary (MPEG) and CFSR, in the northern highland basin of Ethiopia to assess the capacity of these data in supplementing ground observational stations, and stated that both MPEG and CFSR worked well to obtain good precipitation estimates. Among other studies in Africa, Seyoum et al. (2013) studied flow forecasting in poorly gauged, flood-prone sub-basins of the Blue Nile to test the performance of four Quantitative Precipitation Forecasts (QPFs). He concluded that freely available

129 atmospheric forecasting products could provide additional information on precipitation and
130 peak flow events in areas where precipitation data are not available. In a mountainous basin
131 in Ethiopia, four climate products (the CPC MORPHing technique (CMORPH), TRMM,
132 TMPA- 3B42 and PERSIANN) were employed by Bitew et al. (2011) to assess their use in
133 hydrological modelling. Although a significant bias was observed in the reanalysis and merger
134 of different precipitation products, they explored that such data improved the calibration of
135 hydrological model after bias correction.
136 All the aforementioned studies used reanalysis and gridded satellite precipitation estimations
137 with surface precipitation observations. Knowledge on the validity of gridded precipitation
138 products in any part of the globe is of great scientific importance. Therefore, an investigation
139 is critical to fill the gap in some parts of Eurasia like the Eastern Black Sea region which is
140 known by its sparsely gauged highly elevated basins. Moreover, this region has been
141 frequently exposed to floods during the rainy season, so effective hydrological modelling and
142 assessments is crucial to mitigate flood damages. For this purpose, we selected Rize province
143 in the Eastern Black Sea region where only 5 stations with at least 20-year uninterrupted
144 observations are present. In the province, there is one gauging station corresponding to each
145 606-km² area, lower than the WMO's standard requirement (WMO, 1956). Therefore, it is
146 necessary to investigate the feasibility of gridded precipitation products data to supplement
147 the surface observations used in hydrological studies. For this, we selected commonly used
148 gridded precipitation products (CFSR, MSWEP, APRODITE and ERA Interim land). Hence,
149 the objective of this study is to examine the performance of these four gridded reanalyses and
150 merged precipitation products in estimating the spatio-temporal distribution of precipitation
151 in the Rize province of Black Sea region, Turkey. To achieve this goal, the products were
152 subjected to comparison with surface precipitation observations at each station, where daily
153 precipitation data were available. The performance of these products in estimating river flow

154 was tested using a calibrated physically based hydrological model, namely Soil and Water
155 Assessment Tool (SWAT). Annual and seasonal data performance of the climate reanalysis
156 products were then assessed before testing the quality of prediction with various statistical
157 analysis metrics. The study results are presented in three categories: spatial comparison,
158 hydrological evaluation, and statistical evaluation.

159 **2. Data**

160 **2.1. Study Area**

161 The Rize province is located in the north eastern Anatolian mainland of Turkey, so-called the
162 Eastern Black Sea region, with an area coverage of 3,920 km^2. An average annual precipitation
163 in the region is nearly 2,250 mm and total annual runoff is about 2,745 million m^3 (Sen and
164 Kahya 2017). The Black Sea borders Rize province from northern side. It is a mountainous
165 region with elevations reaching up to 3,000m. The borders of the study area are delineated in
166 Figure 1. In addition, the area modelled with SWAT (Ikizdere basin) with the subbasins and
167 flow gauging stations is depicted. The climate of Rize province is in class A (very wet)
168 category based on the Thornthwaite climate classification technique (Sensoy, 2008). In this
169 technique, climates are subdivided into 9 subcategories starting from A (very humid) to E
170 (dry) based on their Thornthwaite climate index values. At the same time, it is a part of the
171 homogeneous streamflow Region 7 (as a result of principle component analysis) of Kahya et
172 al. (2008a). In the same context, Rize province is covered by the homogeneous streamflow
173 Regions B, C or E (as a result of cluster analysis) defined by Kahya et al. (2008b) depending
174 on the month demostrating a distinctive pattern of hydrologic characteristics.
175
176 **Figure 1.**

2.2. Data Sources

The observational weather and streamflow data supplied by the Turkish Meteorological Directorate (MGM) comprises of daily total precipitation (mm) and long term average daily precipitation (mm) recorded at gauging stations in the Rize province (TÜMAS, 2015). In addition, monthly average precipitation and daily flow data in the basin were supplied by the Turkish General Directorate of State Hydraulic Works (Devlet Su Isleri –DSI) (DSI, 2015). Detailed information concerning the input data and the SWAT model setup for the study area is described in Swalih & Kahya (2021). All the input data of SWAT was first prepared prior to the model setup using ArcGIS 10.1, and later the SWAT model was set using its ArcGIS version (ArcSWAT). The model was run and calibrated using observed precipitation and flow data of the Ikizdere basin. The running period is considered between 1979 and 1996 since some of observed precipitation series only extended to 1996. Consequently, the data interval spanning from 1979 to 1990 (1990 to 1996) was used for calibration (validation) phase.

The predicted satellite precipitation estimate and observed gauged precipitation data have different spatial and temporal scales. The ground gauging network consists of 5 daily precipitation stations and 14 daily streamflow gauging stations, which are not uniformly distributed across Rize province (Figure 1). In our study, a size of 11, 16, 12 and 14 grid points were used for the CFSR, ERA-Interim land, APHRODITE and MSWEP reanalyses and merged satellite precipitation products, respectively. These grid points were given for our study area by default from the respective dataset product providers (grid size resolution). In order to give uniformity for our study, we chosen 0.5°× 0.5° horizontal resolution across the precipitation datasets since not all of these datasets have a higher plot resolution. Among the ECMWF products, we used the ERA Interim land and interchangeably termed as ECMWF throughout the manuscript. The grid meshes constructed for the gridded precipitation products

are presented on Figure 2, which illustrates the grid schemes of the four precipitation products over the study domain. The observed flow data at the outlet of Ikizdere basin was employed as check for the simulation accuracy using each product set. Thus, one observed and four reanalyses and merged satellite precipitation product sets were used in our analysis.

Figure 2.

The **Climate Forecast System Reanalysis (CFSR)** is a global, high-resolution, coupled atmosphere-ocean-land surface-sea ice system designed to give the best reanalysis estimate of these coupled domains over the 32-year period of record from January 1979 to March 2014. (Saha et.al., 2010). The CFSR dataset production was executed in such a way that each run of atmospheric and ocean analysis was repeated at every 6-hour (0000, 0600, 1200, and 1800 UTC) using a 9-hour guess forecast with 30-min coupling to the ocean. The land analysis using observed precipitation with the Noah land model was made only at 0000 UTC. Finally, a coupled 5-day forecast from every 0000 UTC was delivered with an identical horizontal resolution version of the atmosphere for check-up mechanism. The CFSR weather data was accessed from the official site of Global SWAT weather data with a special resolution of 0.5°X 0.5° and daily time step (https://globalweather.tamu.edu/).

The Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources project (APHRODITE) was an international cooperative project that generated a 57-year daily precipitation dataset for the Asian subcontinent by first collecting and analysing rain gauge observations from thousands of stations in Asia adding to the gauges reporting to the WMO Global Telecommunications System (Yatagai, et al., 2012). The gridded fields of the daily precipitation are defined by interpolating gauge observations obtained from meteorological stations. The main goal of this project was to produce a quantitative high-resolution product for the validation of high-resolution models and for

studies on hydrological applications. In the production of the APHRODITE dataset, gridded analysis of global daily and monthly precipitation have been interpolated on a 0.05-degree latitude–longitude grid by merging several kinds of information sources with different characteristics, including gauge observations and estimates inferred from a variety of satellite observations (Xie and Arkin, 1997; Yatagai et al., 2012). The APHRODITE used the WorldClim technique for its data generation to get data throughout the world. The steps used in producing the daily global climatology include: i) summing daily data as monthly values, ii) gathering the monthly data and averaging the value if the station has recorded data for more than 5 years, iii) preparing the climatology at 0.05°, iv) taking the ratio of station climatology (step 2) to the climatology at 0.05° (step 3) for each month, v) interpolating the ratio in step 4 interpolated using Sheremap technique (Willmott et al. 1985) at a resolution of 0.05°, vi) multiplying the interpolated values of step 5 with the world climatology (step 3), and vii) calculating the daily climatology by using the Fourier transform technique from the values obtained in step 6. The precipitation data was downloaded with a daily temporal resolution and 0.5°X 0.5° spatial resolution from the official data site of APHRODITE project (http://www.chikyu.ac.jp/precip/english/products.html).

The numerical weather prediction system at European Centre for Medium-Range Weather Forecasts (ECMWF) - ERA Interim/Land weather product is designed to use data assimilation systems to 'reanalyse' archived observations, creating global data sets describing the recent history of the atmosphere, land surface, and oceans. The quality of the product is controlled with ground and remote-sensing observations. The ERA-Interim/Land is a weather reanalysis product, which has a horizontal resolution of 80 km, and 3-hourly surface parameters. The ERA-Interim/Land is the result of a single 32-year period (Jan 1979 - Dec 2010) simulation with the latest ECMWF land surface model driven by meteorological forcing from the ERA-Interim atmospheric reanalysis and precipitation adjustments based on monthly

GPCP v2.1 (Global Precipitation Climatology Project). The monthly GPCP data set merges satellite and rain gauge data from a number of satellite sources including the global precipitation index. In addition, rain gauge data from the combination of the Global Historical Climate Network (GHCN), Climate Anomaly Monitoring System (CAMS) and Global Precipitation Climatology Centre (GPCC) data set (consisting of approximately 6700 quality controlled stations around the globe interpolated into monthly area averages) are used over the land. The technique used was stand-alone land simulation both for global and point scales given the complexity involved in the coupled land-atmosphere assimilation systems. It is stated to be one of the most accurate meteorological forcing to drive the land surface numerical schemes (Balsamo et al., 2012; Balsamo et al., 2015). The ERA Interim land precipitation data was downloaded with a daily temporal resolution, 0.7° X 0.7° spatial resolution from the official website (http://apps.ecmwf.int/datasets/data/interim-land/type=fc/).

At last, the **Multi-Source Weighted-Ensemble Precipitation (MSWEP)** is a new global precipitation dataset (1979–2016) with a high 3-hourly temporal and 0.1° spatial resolution. The dataset is unique so that it takes advantage of a wide range of data sources, including gauges, satellites, and atmospheric reanalysis models, to obtain the best possible precipitation estimates at global scale. The long-term mean of MSWEP was based on the recently released Climate Hazards Group Precipitation Climatology (CHPclim-version 1.0) dataset with 0.05 resolution (Beck et al., 2017). It is global precipitation climatology based on gauge observations and satellite data. The procedure used in the dataset preparation include the following steps: (i) calculating the long-term bias-corrected climatic mean; (ii) evaluating several gridded satellite and reanalysis precipitation datasets in terms of temporal variability in order to assess their potential inclusion in the MSWEP; and finally (iii) downscaling the long-term climatic mean temporally in a stepwise manner first to monthly, then daily, and finally to 3-hourly timescales using weighted averages of precipitation anomalies derived

from the gauge, satellite, and reanalysis datasets to yield the final MSWEP dataset. For accuracy, basin-ratio equations and streamflow values were used to correct the bias in the datasets. It has been reported that the MSWEP products outperformed other precipitation datasets, such as the CMORPH-CRT, PERSIANN-CCS, PERSIANN-CDR, and so on (Beck et al., 2017). The MSWEP precipitation dataset was downloaded with a daily temporal resolution, and 0.5°X 0.5° spatial resolution from the official website (http://www.gloh2o.org/).

The long term monthly average precipitation scatter plots for the four weather products versus the observed precipitation in the basin are depicted in Figure 3. The analysis was performed for the period 1979-1996. The precipitation regime of the area has been presented earlier in Figure 1. The coastal region typically has a high precipitation compared with the mainland (mountainous inland) region. This phenomenon has been previously documented by Sensoy (2008) and Sen (2013). The high peaks of mountains in the eastern Black Sea region are predominantly covered with snow during most of the time (especially in winter). Table 1 summarises the datasets used in the study.

Table 1.

Figure 3.

3. Methodology

The observational precipitation data was used to assess the performance of four state-of-the-art reanalyses and merged satellite based precipitation product. We used three step evaluations in this study: (i) spatial comparison, (ii) hydrological assessment, and (iii) statistical evaluation. The arithmetic weight of individual grid stations was used to calculate annual average precipitation values. Moreover, a calibrated and validated physically based

298 hydrological model (SWAT) based on the study area was employed to assess the quality of
299 these climate products.

300 In general, process-based models focus to formulate the entire physical process from
301 precipitation to flow in the hydrologic cycle by balancing the amount of water on daily,
302 monthly and seasonal time scales. The input data are usually temperature, humidity, soil
303 moisture, soil texture, precipitation, evapotranspiration, lateral flow and percolation rate. A
304 major drawback of these models is the requirement of large number of parameters. This
305 condition and other difficulties are limiting the use of the physical based models to a very
306 small number of river basins (Demirel et al. 2009). One of the most popular process-based
307 models is SWAT, which is used to examine the impacts of land use changes on the runoff and
308 groundwater, production of sediment and water quality; for example, flow in the tributaries or
309 agricultural issues. Readers are referred to Arnold (1996) and Arnold et al (2012) for further
310 information concerning the details of SWAT model. We set up the SWAT model for Ikizdere
311 basin with a total area of 731.4 km^2 (located in Rize province). We then run the model with
312 all the four precipitation data products in order to be compared with the observed river flow
313 records at the outlet of the basin. The general scheme used in this study is depicted in Figure
314 4 .

315 **Figure 4.**

316 **3.1. Spatial comparison**

317 We applied the Krigging geostatistical technique to interpolate point precipitation data over
318 the study domain. This technique is categorised under the interpolation methods consisting of
319 geostatistical methods, which are based on statistical models including autocorrelation; that is
320 to say, statistical relationships among measured points (Oliver, 1990). Kriging is usually
321 preferred due to the fact that it has not only the capability of prediction, but also can provide

some degree of accuracy of the predicted values unlike the deterministic interpolation techniques like the inverse distance weighting (IDW), which only estimate the unknown value based on the distance of neighboring points. Kriging calculation is based on the weighted sum of the data as follows:

$$P(Z_o) = \sum_{i=1}^{N} \mu_i P(Z_i) \quad (1)$$

where $P(Z_i)$: measured value at the *i*th point, μ_i: weight for the measured value at the *i*th point, $P(Z_o)$: value at the prediction location, *N*: number of measured values. The term weight (μ_i) depends not only on the distance of the prediction point to the measured points, but also on the overall arrangement of the measured points in the study area (Oliver, 1990). We used the ordinary Kriging method from the spatial interpolation methods (e.g., Kriging, IDW, and Spline) to create interpolated surfaces over the study area (available in ArcGIS interface). The statistical results presented in our work are based on the grid values not the Kriging interpolation surfaces.

3.2. Statistical Evaluation

The statistical estimates were calculated using the daily time series data for each grid station. In this section, we went through statistical evaluation of the gridded climate products and the streamflow simulation of SWAT (with gridded precipitation forcing). We prepared histogram plot of the daily time series, probabilistic distribution function (PDF) and cumulative distribution function (CDF). In order to quantify the performance simulated flow compared with the observational flow values, Moriasi et al. (2007) recommended to apply four quantitative statistics. They are namely: (i) Nash-Sutcliffe efficiency (NS), (ii) ratio of the root mean square error to the standard deviation of measured data (RSR), (iii) percent bias (PBIAS) and (iv) coefficient of determination (R^2). NS is used to assess the degree of fitness exhibited by the satellite data with that of the observational data, and implying the closeness of the value

of interest with the observation (Eq. 2). RSR is expressed to standardize the root mean square error (RMSE) statistic by the standard deviation of observations. RMSE is the difference between the distribution of the ground precipitation observations and that of satellite precipitation estimations. The lower the RMSE score, the closer the satellite precipitation estimations represents the observed ground precipitation observations. RSR is calculated as a ratio of the RMSE and standard deviation of measured data (Eq. 3). PBIAS measures the average tendency of the simulated data to be larger or smaller than their observed counterparts (Moriasi et al., 2007). The optimal value of PBIAS is zero (low values indicate accurate model simulation) (Eq. 4). Positive values indicate model underestimation and negative values overestimation bias. Initially, the bias for the inland area and coastal area was calculated separately to incorporate regional impact. Then the average of these values was taken as the overall dataset bias.

Our final quantitative statistic R^2 is used to evaluate the goodness of fit of the relation. R^2, imply the degree of linear association between the two variables, addresses the question on how well the gridded precipitation estimates correspond to the ground precipitation observations (Eq. 5).

In the decision stage, the indications of Moriasi et al. (2007) could be of practical criteria for evaluation. He stated that model simulations could be judged as satisfactory if NS > 0.50 and RSR < 0.70, and if PBIAS < 25% for flow analysis

$$NS = 1 - \left[\frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{n}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{mean}\right)^2}\right] \quad (2)$$

$$RSR = \frac{RMSE}{STDEV_{obs}} = \left[\frac{\sqrt{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{n}\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(Y_i^{obs} - Y^{mean}\right)^2}}\right] \quad (3)$$

$$PBIAS = \left[\frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{n}\right)*100}{\sum_{i=1}^{n}\left(Y_i^{obs}\right)}\right] \quad (4)$$

$$R^2 = \left[\frac{n\sum(G_i S_i) - (\sum G_i)(\sum S_i)}{\sqrt{\left(n(\sum G_i^2) - (\sum G_i)^2\right)\left(n(\sum S_i^2) - (\sum S_i)^2\right)}}\right]^2 \quad (5)$$

where: Y_i^{obs} is the i^{th} observation; Y_i^{n} is the i^{th} precipitation/ simulated flow value; Y_i^{mean} is the mean of observed data; STDEV: standard deviation of observed data; R^2: coefficient of determination; G_i is ground observation; S_i is precipitation estimates/simulated flow; and n is total number of data pairs or the total number of observations.

In addition, we have used two statistical error metrics, which have been widely used by climatologists to evaluate precipitation dataset, namely probability of detection (POD) and false alarm rate (FAR) (Wilks, 1995). These metrics explain the ability of the gridded precipitation dataset to detect the occurrence of rain and no-rain events without considering the amount of rainfall in these precipitation events. If *a, b, c*, and *d* are defined by:

	Observed	
	Rain	No-rain
Gridded rain	***a***	***b***
Gridded no-rain	***c***	***d***

Then,

$$POD = \frac{a}{(a+c)} \quad (6)$$

$$FAR = \frac{a}{(a+b)} \quad (7)$$

The POD is the ratio of the gridded precipitation event occurred in which it was observed (Eq. 6). It is the likelihood that the event would be forecasted provided that it was observed. Again the best value of POD for a gridded precipitation is one and its worst value is zero. On the other hand, the FAR is the proportion of the gridded precipitation estimate which was not actually observed (Eq. 7). We prefer FAR to be zero for the best gridded dataset since higher value indicate high uncertainty of the gridded precipitation dataset.

3.3. Hydrological Evaluation

In this study, the semi-distributed hydrological model SWAT was used to model the Ikizdere basin of Rize province. The model was then used to assess impact of these various precipitation datasets on the hydrology of the basin. The weather data were used from available Turkish State Meteorological Service (MGM) stations in the SWAT model build up and calibration. ArcSWAT 2012 tool was used in this study to delineate the DEM, Soil & Lanuse maps (which were provided by the Istanbul Technical University, Civil Engineering Department) and setup the SWAT model for the basin. Once the hydrological response unit (HRU) and subbasin analysis is done, the ArcSWAT interface then calls the weather data of the basin to finalize the SWAT model setup. For model optimizations, the DSI observed flow data of Camlıkdere station of Ikizdere basin was implemented for our model calibration and validation (Swalih and Kahya, 2021).

The SWAT model setup for our study area was simulated for the period 1976-1996 (21 years) based on the available precipitation data in the basin. For the simulations and calibration phases, we employed observed precipitation data as it is only dominant climate factor, which has direct and significant impacts on the basin flow. In addition, we could not find alternative weather data with required level of quality in the study area. The mean monthly average values of temperature, wind speed, solar radiation and humidity were generated from the Climate Forecast System Reanalysis data (downloaded from the US National Centers for Environmental Prediction (NCEP - CFSR)). The simulations were run on a monthly time step and a warm-up period of 3 years was provided to give time for the model to adjust itself to the data and the basin. The output of the model in the warm-up period was not used for our analysis. After successfully setting up the SWAT model for the basin, it was calibrated and validated using historical flow data of the basin.

The Parasol routine of SWAT-CUP auto-calibration tool was used to calibrate the model (Abbaspour, 2015). The first three years (1976-1978) were used for model warm-up, next twelve years (1978-1990) for model calibration and six years (1991-1996) for model validation. The sensitivity analysis method implemented in SWAT is called the Latin Hypercube One-factor-At-a-Time (LH-OAT) design as proposed by Morris (1991). Details on the LH-OAT technique could be referred from Van Griensven et al. (2006). Based on recommendations from literature and result of sensitivity analysis, the top 26 sensitive model parameters were chosen for model calibration (Table 2). Each of these parameters represent a physical process affecting the flow (Neitsch et al., 2011). Since it is not possible to determine all these parameters, they must be adjusted by tuning their default values until the most appropriate model output (flow in this case) is achieved without violating the threshold boundary values. The most sensitive model parameters will be essential in model calibration since these parameters will affect our model the most.

Table 2.

Precipitation is the major component of the hydrological cycle. The streamflow and water yield of a basin are directly proportional to precipitation (rainfall and snowfall). In hydrology, the total amount of water flowing on the surface (streamflow) and subsurface (groundwater flow) is represented by the total water yield of a basin. It is calculated by subtracting evapotranspiration from precipitation. The daily time series was used to force the SWAT model to get the simulated flow from the basin on a monthly time scale. Therefore, the river flow analysis was carried out for monthly data for each gridded dataset. In order to avoid the confusion on the analysis period, we restricted our analysis period from 1979 to 1996 since some observational stations recording ceased starting from the year 1997.

The spatio-temporal performances of the reanalyses and merged satellite based precipitation product as compared to the observed precipitation data were evaluated using different graphs for visual assessment. The ArcSWAT interface of ArcGIS software has been implemented during the running phase of the model with various input data (Winchell et al., 2013). In the end, the performance of all the weather products were subjected to testing via the calibrated and validated SWAT model, which was parameterized according to our study basin using observed flow data at the basin outlet. Readers are referred to Neitsch et al. (2011) and Arnold et al. (2012) for further details on the hydrological model, theoretical and input-output documentation of the model.

4. Results and Discussion

4.1. Spatial comparison

As discussed in section 3.1, the spatial average precipitation of the four reanalyses and merged satellite based precipitation products were computed using the Kriging technique as depicted in Figure 5. The analysis was done using daily precipitation data from 1979 – 1996. The annual average precipitation estimation of the products was used in the analysis. Initial evaluation (yearly as well as seasonal) of the APHRODITE and ERA Interim/land precipitation were satisfactory compared with the other datasets. The MSWEP estimate seems good for the north-east region, however, it was poor for the south-west region. On the other hand, CFSR exhibits poor performance since the annual estimates are contradicting with the observed values. Similar conditions could be observed for the wet season (Sep. –Dec.) and the dry season (Mar.-May). In an overall evaluation, the percentage bias for the mean monthly precipitation for the datasets was calculated by using the observed precipitation as reference. Initially, the bias for the inland area and coastal area was calculated to incorporate the regional effect, then the average of these values was taken as the overall dataset bias. Hence, the mean

462 monthly precipitation bias for ERA-Interim/land, APHRODITE, CFSR and MSWEP was
463 found to be 8% and 23%, 42% and 22% respectively. It should be noted that the observations
464 are far from being perfect due to the low density of gauging stations in the province. It is
465 possible that the gridded precipitation products could give a better description of precipitation
466 pattern since they have improved grid density than the measuring gauges. Our analysis showed
467 that gridded precipitation products could serve as a good substitute to observation
468 precipitation data for sparsely gauged basins with insufficient weather data.

469 **Figure 5.**

470 **4.2. Statistical Evaluation**

471 The statistical evaluation of the gridded climate products and the river flow simulation of
472 SWAT (with climate products forcing) is discussed in this section. The daily time series
473 precipitation histogram, probabilistic distribution function (PDF) and cumulative distribution
474 function (CDF) plots for grid points are presented in Figure 6. The histogram plots of ERA
475 Interim and APHRODITE are comparable with the observed histogram except for the low
476 precipitation values. However, the statistical analysis for CFSR and MSWEP datasets
477 exhibited poor performance. The CDF plot of the observed precipitation is quite different
478 compared with all the gridded precipitation stations. All the measurement stations have a
479 distinct CDF for the observation whereas the CDFs for the gridded precipitation products have
480 little or no variability. MSWEP has captured the PDF of the observed precipitation much
481 better that the other datasets. The ERA Interim and APHRODITE exhibited poor performance
482 of PDF. Here we can observe that the performance of the various precipitation datasets
483 displayed different performance unlike what we observed in the spatial evaluation (section
484 4.1).

485 **Figure 6.**

Overall statistical performance of the SWAT model is presented in the sense of comparing the model simulation runs with the observed and reanalyses and merged satellite based precipitation products in the Table 3. The observed flow at the outlet of Ikizdere basin was used in our statistical analysis. In the table, MGM represents flow simulated by the SWAT using observed precipitation data. ERA Interim land precipitation data exhibited good performance (NS=0.53, PBIAS=19.9). These values are acceptable with the criteria set in the literature as discussed in Section 3.3. A simulation can be accepted as satisfactory if NS > 0.5 and PBIAS < 25% (Moriasi et al., 2007). On the other hand, the simulation of the model with the CFSR, APHRODITE, and MSWEP precipitation products have a very low performance (NS = 0.22, 0.38, and -0.73, respectively) and percentage bias (PBIAS = 33.40, 41.40 and 85.0, respectively), which are below the minimum criteria needed for a good simulation. Once again, the simulations with the CFSR, APHRODITE, and MSWEP precipitation data showed low performance.

Table 3.

The reliability of the gridded precipitation data sets was also checked by other metrics (POD and FAR). These metrics evaluate the underlying reason for differences in the observed and gridded precipitation data. We could see in table 3 that the POD values are far or less high. The preferred POD value is one for the perfect scenario. This indicates that the probability that an observed precipitation to be forecasted by the gridded precipitation data was high for all the datasets. However, the FAR values for all the gridded datasets were also higher, that is not desirable. We expect this value to be close to zero for a good dataset. A big FAR value means that the proportion of the gridded precipitation event which was not observed is higher for the datasets (i.e., all the datasets tend to give “false-alarm” when it comes to the precipitation estimation in our study area).

510 Overall performance of the various model simulations using the reanalyses and merged
511 satellite based precipitation product with that of the observed flow at the outlet of the study
512 basin is presented in Table 4. The notation MGM here represents the simulation run with
513 observed precipitation data. CFSR and MSWEP simulations depict overestimation while
514 APHRODITE depicts underestimation of the observed flow. Once again, ERA Interim land
515 has performed well compared with the other reanalysis and merged precipitation products and
516 its simulation flow statistics are comparable with the observed flows of the basin.

517 **Table 4.**

518 **4.3. Hydrological Evaluation**

519 When we come to the precipitation datasets that were selected for this study, first we decided
520 to analyse the annual and seasonal distribution on the whole study area. The yearly average
521 values of precipitation products were plotted in Figure 7 and 8. The weighted averages of the
522 grid values were used for both figures. There is a visible change in trend for the coastal and
523 mainland precipitation values. In general, the coastal areas have higher precipitation than that
524 of the mainland mountainous regions, which is in agreement with the previous studies
525 (Sensoy, 2008; Şen, 2013). The mountains are predominantly covered with snow in winter
526 retaining a significant amount of the precipitation. The CFSR results are the only exception
527 in which the estimates were inversely estimated. For our study area, the effect of elevation
528 difference on precipitation are noticeable in Figures 7 and 8, showing that precipitation is
529 significantly higher for the coastal region for the both wet season (September -November) and
530 dry season (March-May), opposite to our expectation. Our study period spans from 1979 to
531 1996 since the observation records were not available after 1996.t is well known that
532 precipitation increases with increase in altitude. Our results are in agreement with Sensoy
533 (2008) who reported mountain influences on the precipitation distribution of the country. Due

534 to the presence of Taurus mountains along the coastal areas, the rain clouds cannot penetrate
535 into the interior parts of Turkey, causing that the majority of precipitation falls along the
536 coastal regions. For this reason, the land of Rize province becomes the most humid coastal
537 regions all over the country.

538 Since the records of most gauging stations in the mainland region end in 1996; therefore, we
539 decided to present the analysis results only for the period 1979-1996 to avoid confusion. Our
540 findings (Fig. 7 and 8) once again confirm the indications of Şen (2013), who reported annual
541 precipitation higher than 1500 mm for the majority of Rize province. The CFSR reanalysis
542 exhibited low performance so that the precipitation is underestimated by 39% for the wet
543 season in the coastal region. For the mainland region, the wet season was underestimated by
544 49% and dry season was overestimated by over *200%*. On the other hand, the ERA
545 Interim/land underestimated the wet season by 27% for the coastal and overestimated by 18%
546 for the mainland region. But it gives good estimation of mean for the dry season (less than
547 20% underestimation). The APHRODITE underestimated precipitation in the mainland
548 region for the wet and dry season by 58% and 14 %, respectively. The estimates for the coastal
549 region exhibits small deviation from the observed climatology. The MSWEP overestimates
550 the precipitation in the coastal (mainland) region by 29% and 41% (64% and 42%) for the wet
551 and dry season, respectively. These results are in good agreement with Figure 3. The observed
552 precipitation value suddenly drops on the year 1988 for both regions due to drought conditions
553 prevalent in that particular year throughout Rize province. Another striking result is that a
554 reversed spatial distribution of annual precipitation exists compared to the observed data for
555 the CFSR dataset. It has been reported by the CFSR data provider that there have been
556 confirmed error reports of CFSR precipitation in some regions of the globe
557 (https://swat.tamu.edu/news/2015/cfsr-weather-data/).

Figure 7.

Figure 8.

The long-term monthly average precipitation plots for the four weather products versus the observed precipitation both for the coastal and mainland regions of the basin are depicted in Figure 9 (a). The overall comparison of all the datasets is plotted on figure 9 (b). The results herein are consistent with those in the previous analysis in section 4.1.

The ERA Interim land comparatively captures the observed precipitation pattern quite well for the coastal and mainland areas, except its underestimation of the coastal average precipitation. Likewise, the APHRODITE estimates well for the coastal area, but underestimates the mainland average precipitation by 49%. In contrast, the MSWEP demonstrates an overestimation for the average precipitation on the both coastal and mainland regions by 30% and 55%, respectively. The worst average precipitation estimates were made by the CFSR in this analysis as having an underestimation level up to 30% for the coastal area and overestimation level up to 80% for the mainland. As an overall evaluation, it can be said that the seasonal pattern of precipitation could be satisfactorily captured by the MSWEP, APHRODITE and, by ERA Interim land in our study area. Nevertheless, the CFSR products poorly resemble to the observed seasonal pattern.

Figure 9.

4.3.1. Seasonal Performance

We adopted a hydrological model (SWAT) parameterized for the Ikizdere basin to make original assessments on the hydrological responses to precipitation inputs. The Ikizdere basin was categorized into 149 Hydrological Response Units (hru) and 22 subbasins based on the soil, Landuse and slope maps overlay. Then the observed precipitation and the other weather

581 data were implemented to simulate the model. The Muskingum routing method was used in
582 SWAT to route the river flow. The model calibration was done using the observed
583 precipitation data (MGM). As mentioned in section 3.3., the observed precipitation data was
584 divided into three periods: warm-up (1976-1978), calibration (1979-1990) and validation
585 (1991-1996). It is after completing the calibration and validation that we implemented the
586 model to assess the impact of the various precipitation datasets in our study area.

587 In this section, we evaluated the hydrological response of our study area to the precipitation
588 input using a calibrated SWAT model. Figure 10 illustrates the annual cycle of precipitation
589 (rainfall and snow fall) estimates of the gridded precipitation products compared with the that
590 of the observed precipitation. There are a number of important indications that can be readily
591 commented from Figure 10. All the cyclic behaviour of the seasonal APHRODITE product
592 staying below the measured annual cycle, implying underestimation of precipitation, while
593 the MSWEP output lies above the measured annual cycle indicating overestimation. A poor
594 performance is detected from the annual cycle of CFSR, fluctuating with excessive positive
595 deviations for the dry season (March-May) and preceding with negative deviations for the wet
596 season of year (Sep. – Nov). There is a dipole of product underperformance in the sense of
597 underestimation (overestimation) of precipitation with the APHRODITE (MSWEP) annual
598 cycles. It is worthwhile to note that the peaks of the annual precipitation cycle of CFSR and
599 APHRODITE products completely correspond to the peak timing of the observed rainfall. A
600 careful visual inspection reveals that, among others, ERA Interim land captures better the
601 annual cycle of the measured data with little deviations from the measured precipitation
602 values. These values are in agreement with the analysis in section 2.2 (figure 3). Considering
603 snowfall, both CFSR and MSWEP significantly overestimate in winter season while only the
604 MSWEP overestimate snowfall in fall. The APHRODITE slightly underestimate snowfall in
605 both winter and fall seasons. ERA Interim gave a better estimation of the snowfall compared

606 with the other datasets. In this particular study area, the snow cover is essential because
607 significant amount of precipitation will be stored in the mountainous parts of the mainland
608 region during the winter season.

609 **Figure 10.**

610 **Figure 11.**

611 The SWAT computes the various components of the hydrologic cycle in the basin including
612 water yield. The total amount of water flowing on the surface and subsurface is represented
613 by total water yield of the basin which is calculated by subtracting evapotranspiration from
614 precipitation. Figure 11 summarizes the output of seasonal water yield simulated by the
615 SWAT using the four gridded precipitation inputs for the Ikizdere Basin. It is clearly certain
616 that annual cyclic behavior of all data are quite consistent with each other, like having a peak
617 on May. From standpoint of peak water yield magnitude, simulation with the CFSR and
618 MSWEP data, the estimated flow in May is significantly greater than the observed water yield.
619 However, the water yield value was lower for the APHRODITE. Moreover, water yields were
620 highly overestimated during the period May-December by the MSWEP simulation, which is
621 due to overestimated precipitation by the MSWEP in the first four months (January-April)
622 when the majority of precipitation falls in the form of snow. The high snowmelt volume
623 simulated by SWAT is seemingly resulted from overestimated precipitation (snowfall)
624 estimation of CFSR and MSWEP (Figure 10). The water yield estimate of the ERA Interim
625 land is the most comparable with the observed value. In the case of surface streamflow
626 estimations, the both CFSR and MSWEP again overestimated flow out of the basin whereas
627 the APHRODITE tend to underestimate. As a result, most comparable flow with observed
628 value was that of the ERA Interim simulation. Our study gave similar result with what was
629 reported by Essou et al., (2016) (NS = 0.78) and Manzato et al., (2015) ($R \geq 0.82$) in such a

630 way that the ERA Interim land reanalysis data was found to successfully compensate the
631 deficiency of surface weather observation records. Other studies on CFSR has found that it
632 performs well in their study areas unlike what we have found for the Rize province (Essou et
633 al., 2016; Essou et al., 2017). These studies were done on similar areas to our study area with
634 complex topography and low density of weather gauging stations. In the study conducted by
635 Husain et al. (2017) for the evaluation of gridded precipitation in the Himalaya mountainous
636 basin, the largest error in the gridded data arose mainly from elevation, showing the impacts
637 of elevation on precipitation data measurement accuracy. Further study is needed to examine
638 the impacts of elevation difference on the precipitation distribution in the basin.

639 **4.3.2. Temporal Performance**

640 In this section, the performance of the reanalyses and merged satellite based precipitation
641 product at the temporal scale is analysed using the SWAT model simulations. The measured
642 monthly total flow at the outlet of Ikizdere basin is used to compare with the simulation of
643 gridded precipitation data (figure 12). The simulation run with the measured MGM
644 precipitation data is represented by 'Measured flow' in Figure 12. The simulations using the
645 MSWEP and CFSR precipitation resulted in overestimated flow, especially for the peak
646 seasons. The flow exaggeration is due to the overestimation of precipitation over the study
647 area by both these weather products. The simulation with the APHRODITE data
648 underestimates the flow of our study area. Precipitation underestimation of the APHRODITE
649 is once again the main cause of flow underestimation (Figure 10). The simulation with the
650 ERA Interim land reanalysis is the closest simulation flow with the observed flow at the outlet
651 of Ikizdere basin. The MSWEP and CFSR precipitation data generally overestimate the peak
652 flows; whereas, the ERA Interim land precipitation data can be observed to have satisfactory
653 simulation outcomes compared with the observed flow data. Our analysis results are in

654 agreement with Essou et al. (2016) and Zhu et al. (2017), who concluded that the ERA Interim
655 land reanalysis weather data could not only be successfully used in place of surface weather
656 observation records, but also improve hydrological modelling performance.

657 **Figure 12.**

658 In general, the MSWEP and CFSR results were not consistent with what have been reported
659 in other parts of the world and even in the Mediterranean basins. For example, the ERA
660 Interim dataset analysis in our study presented similar results with those of Manzato et al.,
661 (2015); Essou et al., (2016); and Essou et al., (2017) in such a way that the ERA Interim land
662 reanalysis data was found to successfully compensate the deficiency of surface weather
663 observation records. Some studies on CFSR (e.g., Worqlul et al., 2014; Essou et al., 2016;
664 Essou et al., 2017) reported good performance in their study areas unlike our case in the Rize
665 province. However, Worqlul et al. (2014) found that CFSR exhibited similar inconsistency in
666 regard to the bias over his study area (Lake Tana basin - Ethiopia). The possible reason could
667 be both studies were conducted on similar areas with complex topography and low density of
668 weather gauging stations.

669 **5. Conclusions**

670 Any area with limited weather gauging stations is difficult to study. Therefore alternative/
671 supplementary data source is essential to support the data scarcity. Reanalysis and gridded
672 precipitation products have been reported to be good alternative data source, which could
673 supplement precipitation data scarcity in climate studies. In this study, the four precipitation
674 products, namely (i) CFSR, (ii) APHRODITE, (iii) ERA Interim land, and (iv) MSWEP, were
675 subjected to performance tests in the Rize province of Eastern Black Sea Region, Turkey.

The precipitation products were compared using three analysis techniques: (1) spatial average graph, (2) statistical assessment, and (3) hydrological evaluation. To evaluate the performance of precipitation averaged over the study area (Rize province), spatial average of the precipitation datasets was plotted. The statistical evaluation of the gridded climate products and the streamflow simulation of SWAT (with the precipitation datasets as an input) was assessed in detail. The daily precipitation histogram, PDF, and CDF plots for each grid stations were plotted. Other statistical metrics were also used, including NSE, RSR, PBIAS, POD and FAR. The hydrological SWAT model was used to assess the basin's response for each precipitation product. The SWAT simulation flow outputs of the gridded precipitation products were assessed for seasonal and temporal performance in the study area.

In comparison with the observed precipitation data, the spatial annual average precipitation estimations of the APHRODITE and ERA Interim land were found to be satisfactory. The MSWEP estimations seem good for the Northern region. The CFSR estimates contradicted with the observed values exhibiting poor performance. When the wet (September-December) and dry (March-May) seasons were evaluated, again the APHRODITE and ERA-Interim/land overestimated the mean monthly precipitation by 8% and 23%, respectively. On the other hand, the CFSR and MSWEP overestimated the mean monthly precipitation by 42% and 22%, respectively. It should be noted that the observation data is far from being sufficient due to the low density of gauging stations in the province. It is possible that the gridded precipitation products could describe the precipitation pattern of a basin if they have improved grid density than the measuring gauges. This implies that gridded precipitation products could serve as a good supplement to observational data for sparsely gauged basins. However, the observed precipitation data is essential in the validation of the gridded precipitation data as well as the calibration of hydrological models for the basins where observed data is available.

The statistical evaluation of precipitation datasets and the river flow simulations (with input climate products) shows that the MSWEP captured the PDF and CDF of the observed precipitation much better than other datasets. The ERA Interim and APHRODITE exhibited poor performance. The histogram plots of ERA Interim and APHRODITE were found to be comparable with the observed precipitation whereas the CFSR and MSWEP precipitation were found to be less comparable. The CDF plot of the observed precipitation is quite different than those of all the gridded precipitation stations. Unlike what we observed in the spatial evaluation, the performance of the various gridded precipitation datasets using daily time series statistics was quite different. The ERA Interim and APHRODITE datasets did not exhibit good performance. The POD of all the datasets were found to be satisfactory. However, the FAR metric showed that all the precipitation datasets exhibit sustained "false-alarm". This shows that the performance of the gridded precipitation products could have different performance based on the resolution and type of data analysis technique used.

The SWAT model simulations with the CFSR, APHRODITE, and MSWEP precipitation datasets resulted in NS value of 0.2, 0.4, and -0.7, respectively; and percentage bias of 33.4%, 41.4% and 85.0%, respectively. The datasets of CFSR and MSWEP proved to overestimate the peak flows of the basin since the most dominant weather variable of SWAT model (precipitation) was overestimated. The APHRODITE simulation was found to underestimate streamflow (particularly low flows). These values are indicatives of the poor performance of these precipitation datasets for hydrological study of the area. Unlike the other simulations, the ERA Interim land SWAT simulation was in good agreement with the observed flow pattern of the basin. The ERA Interim exhibited a much better hydrological simulation performance with NS = 0.53 and PBIAS = 19.9, that is consistent with the suggested range by Moriasi et al. (2007) for a satisfactory model performance (section 3.3.).

In this study, we made a thorough analysis of four gridded precipitation products for Rize province of Eastern Black sea region. Our analysis results showed that gridded precipitation data could supplement weather observation records and they could improve hydrological modelling performance of the basin. The ERA Interim showed better performance than the datasets inputted to the hydrological model. However, the MSWEP and CFSR were proven to perform bad compared with the ground precipitation observations since they overestimated streamflow in the basin, in particular, during high flow seasons. The APHRODITE precipitation underestimated the streamflow of the study area. Our results were in agreement with those of Aznar (2010), Essou et al. (2016) and Husain et al. (2017) where ERA-ECMWF gridded precipitation product was proved to perform good in mountainous regions (section 1). Therefore, we could argue that gridded precipitation datasets tested to have good performance could be used to supplement surface precipitation observations in hydrological studies.

In a nutshell, it is plausible to encourage researchers to study other similar basins in Black Sea regions so as to put forward comprehensive assessments regarding the product performance in order to augment observed precipitation in the basins with low density gauges. It would be great to select a particular climate product and analyse the impact of the data on the spatial and temporal resolution. We also recommend to expand the climate dataset spectrum and incorporate more recently developed gridded precipitation data (for example, IMERG) and the elevation factor on the precipitation distribution in future studies. In addition, an in-depth further analysis of the precipitation datasets error arising from the data itself need to be investigated thoroughly.

6. Acknowledgements

We would like to acknowledge TUBİTAK (Turkish Scientific and Technological Research Council) for providing the fund to accomplish this study. We are also indebted to the Turkish

748 State Meteorological Service (MGM) and General Directorate of State Water Works (DSI)
749 for the observed weather and flow data access. In addition, the license for ArcGIS software
750 was provided by Istanbul Technical University (İTÜ).

751 **References**

752 Abbaspour K C (2015) SWAT-CUP: SWAT Calibration and Uncertainty Programs - A User
753 Manual. Eawag. Arabian Journal of Geosciences 11: 282

754 Arnold J, Williams J, Srinivasan R, King K, (1996) SWAT-soil and water assessment tool-
755 documentation and user's manual. USDA-ARS, Temple, Texas.

756 Arnold J G, Moriasi D N, Gassman P W, Abbaspour K C et al (2012) Soil & Water
757 Assessment Tool - Input/Output Documentation - version 2012. Texas Water Institute, Texas,
758 USA.

759 Ashouri H, Kuo-Lin H, Sorooshian S et al (2014) PERSIANN-CDR_Daily Precipitation
760 Climate Data Record from Multi-Satellite Observations for Hydrological and Climate Studies.
761 American Meteorological Society 69-84. DOI:10.1175/BAMS-D-13-00068.1

762 Aznar R, Sotillo M G, Martín M L, Somot S, Valero F (2010) Comparison of model and
763 satellite-derived long-term precipitation databases over the Mediterranean basin: A general
764 overview. Journal of Atmospheric Research, 170–184.

765 Balsamo G, Albergel C, Beljaars A et al (2012) ERA-Interim/Land: a global land surface
766 reanalysis based on ERA-Interim meteorological forcing - ERA Report Series. European
767 Centre for Medium Range Weather Forecasts. Shinfield Park, Reading, Berkshire RG2 9AX,
768 England.

769 Balsamo G, Albergel1 C, Beljaars A et al (2015) ERA-Interim/Land: a global land surface
770 reanalysis dataset. Hydrology and Earth System Sciences 19: 389–407

771 Beck H E, Albert I, van Dijk M et al (2017) MSWEP: 3-hourly 0.25_ global gridded
772 precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. Journal of
773 Hydrology and Earth System Sciences 21:589–615.

774 Bitew M M, G/Michael M, G/Michael L T, Bayissa Y A (2011) Evaluation of High Resolution
775 Satellite Rainfall Products through Streamflow Simulation in a Hydrological Modelling of a
776 Small Mountainous Watershed in Ethiopia. Journal of Hydrometeorology 13: 338-350.

777 Demirel M C, Venancio A and Kahya A (2009) Flow Forecast by SWAT Model and ANN in
778 Pracana Basin, Portugal. Advances in Engineering Software 40(7): 467-473. DOI:
779 10.1016/j.advengsoft.2008.08.002

780 DSI weather dataset, 2015. State Water Works General Directorate. http://www.dsi.gov.tr/.
781 Accessed 10/05/2017

782 Essou G R C, and Sabarly F (2016) Can Precipitation and Temperature from Meteorological
783 Reanalyses Be Used for Hydrological Modelling? Journal of Hydrometeorology 17: 1929-
784 1950.

785 Essou G R C, Brissette F, Lucas-Picher P (2017a) Impacts of combining reanalyses and
786 weather station data on the accuracy of discharge modelling. Journal of Hydrology 120–131.

787 Essou G R C, Brissette, F, Lucas-Picher P (2017b) The use of reanalyses and gridded
788 observations as weather input data for a hydrological model: Comparison of performances of
789 simulated river flows based on the density of weather stations. Journal of Hydrometeorology
790 497-513

791 Feidas H, Kokolatos G, Negri A, Manyin M, Chrysoulakis N, Kamarianakis Y (2009)
792 Validation of an infrared-based satellite algorithm to estimate accumulated rainfall over the
793 Mediterranean basin. Journal of Theoretical Applied Climatology 95: 91-109.

794 Feidas H (2010) Validation of satellite rainfall products over Greece. Journal of Theoretical
795 Applied Climatology 99: 193-2016.

796 Hussain S, Songa X, Renb G, Hussain I, Hanaand D, Zaman H H (2017) Evaluation of gridded
797 precipitation data in the Hindu Kush–Karakoram–Himalaya mountainous area. Hydrological
798 Sciences Journal–Journal Des Sciences Hydrologiques 62: 14 2393–2405.

799 Kahya E, Kalaycı S, and Piechota T C (2008a) Streamflow Regionalization: Case Study of
800 Turkey. Journal of Hydrologic Engineering 13(4): 205-214.

Kahya E, Demirel M C and Bég A O (2008b) Hydrologic Homogeneous Regions Using Monthly Streamflow in Turkey. Earth Sciences Research Journal 12(2): 181-193.

Kamiguchi K, Arakawa1 O, Kitoh1 A, Yatagai A, Hamada A and Yasutomi A (2010) Development of APHRO_JP, the first Japanese high-resolution daily precipitation product for more than 100 years. Journal of Hydrological Research Letters 4: 60–64.

Manzato A, Cicogna A, Pucillo A (2015) 6-hour maximum rain in Friuli Venezia Giulia: Climatology and ECMWF-based forecasts. Journal of Atmospheric Research 465–484.

Moriasi D N, Arnold J G, Van Liew M G et al (2007) Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. American Society of Agricultural and Biological Engineers, 50(3): 885−900.

Morris M D (1991) Factorial sampling plans for preliminary computational experiments. Technometrics 33 (2): 161–174.

NASA (2015) Two Decades of Rain, Snowfall from NASA's Precipitation Missions. NASA's Goddard Space Flight Center, Greenbelt, Md., https://www.nasa.gov/feature/goddard/2019/precipitation-missions-release-two-decades-of-rain-snow-data/. Accessed 03 January 2019

NCEP (2015) Global Weather Data for SWAT. https://globalweather.tamu.edu/. Accessed on 05-2017

Neitsch S L, Arnold J G, Kiniry J R, Williams J R (2011) Soil and Water Assessment Tool Theoretical Documentation Version 2009. Technical Report No. 406, Texas Water Resources Institute. College Station, Texas 77843-2118.

Oliver, M. A., 1990: Kriging: A Method of Interpolation for Geographical Information Systems. International Journal of Geographic Information Systems 4: 313–332.

Renner, M., Werner, M. G. F., Rademacher, S., Sprokkeree, E., 2009: Verification of ensemble flow forecasts for the River Rhine. Journal of Hydrology 463–475.

Richard G (2008) A comparison of selected fields in NCEP/DOE AMIP-II and ECMWF ERA-40 reanalyses. Journal of Dynamics of Atmospheres and Oceans 108–142.

828 Saha et al (2010) The NCEP Climate Forecast System Reanalysis. Bulletin of the American
829 Meteorological Society, 91: 1015–1057.

830 Sen O and Kahya E (2017) Determination of flood risk: A case study in the rainiest city of
831 Turkey, Environmental Modelling & Software 93: 296-309. DOI:
832 10.1016/j.envsoft.2017.03.030.

833 Sen,Ö L (2013) A Holistic View of Climate Change and its Impacts in Turkey, Istanbul Policy
834 Center (IPC), Sabancı University, Istanbul.

835 Sensoy et al (2008) Climate of Turkey, Turkish State Meteorological Service(MGM), Ankara,
836 Turkey.

837 Seyoum M, van Andel S J, Xuan Y, Amare K (2013) Precipitation forecasts for rainfall runoff
838 predictions. A case study in poorly gauged Ribb and Gumara catchments, upper Blue Nile,
839 Ethiopia. Physics and Chemistry of the Earth 43–51.

840 Sharma S, Isik S, Srivastava P & Kalin L (2013) Deriving Spatially Distributed Precipitation
841 Data Using the Artificial Neural Network and Multilinear Regression Models. Journal of
842 Hydrologic Engineering 18(2): 194-205

843 Swalih S A and Kahya E (2021): Hydrological model optimization using multi-gauge
844 calibration (MGC) in a mountainous region. Journal of Hydroinformatics. 23 (2): 340–351.
845 https://doi.org/10.2166/hydro.2020.034

846 TÜMAS (2015) Turkish Meteorological Data Management and Archiving System - TÜMAS.
847 http://www.mgm.gov.tr/site/urunler.aspx?u=tum/. Accessed 20 June 2015

848 Van Griensven A, Meixner T, Grunwald S, Bishop T, Diluzio M, Srinivasan R (2006) A
849 global sensitivity analysis tool for the parameters of multi-variable catchment models. Journal
850 of Hydrology. 324: 10–23.

851 Verkade V S, Brown J D, Reggiani P, Weerts A H (2013) Post-processing ECMWF
852 precipitation and temperature ensemble reforecasts for operational hydrologic forecasting at
853 various spatial scales. Journal of Hydrology 73–91.

Ward E, Buytaert W, Peaver W, Wheater H (2011) Evaluation of precipitation products over complex mountainous terrain: A water resources perspective. Advances in Water Resources 1222–1231.

Winchell M, Srinivasan R, Di Luzio M and Arnold J G (2013) Arcswat Interface for SWAT2012: User's Guide. Blackland Research Center, Texas AgriLife Research, College Station 1-464.

WMO (1956) Second Congress of the World Meteorological Organization – Final Report: Volume II, Technical Regulations, Secretariat of the World Meteorological Organization, Geneva – Switzerland.

Wilks D S (1995) Statistical methods in the atmospheric sciences. San Diego, Academic Press

Worqlul A W, Maathuis W B, Adem A A, Demissie S S, Langan S, and Steenhuis T S (2014) Comparison of TRMM, MPEG and CFSR rainfall estimation with the ground observed data for the Lake Tana Basin. Hydrology and Earth System Sciences 11: 8013–8038.

Yatagai A, Yatagai A, Kamiguchi K (2012) Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. American Meteorological Society 1401-1415.

Xie P and Arkin P A (1997) Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs. National Centers for Environmental Prediction, National Oceanic and Atmospheric Administration, Washington, D.C., Bulletin of the American Meteorological Society 78: 2541-2558.

Zhu J, Huang J Q, Yan P W, Huang Y, Kuang X Y (2017) Can reanalysis datasets describe the persistent temperature and precipitation extremes over China? Journal of Theoretical and Applied Climatology 130: 655-671.

880 # TABLES

881 ## Table 1. The Gridded Precipitation datasets used in the study

882

Acronym	Resolution	Source	Reference
CFSR	0.5°X 0.5°	https://globalweather.tamu.edu/	Saha et.al., 2010
ERA Interim	0.7° X 0.7°	http://apps.ecmwf.int/datasets/data/interim-land/type=fc/	Balsamo et al., 2012
APHRODITE	0.5°X 0.5°	http://www.chikyu.ac.jp/precip/english/products.html	Yatagai, et al., 2012
MSWEP	0.5°X 0.5°	http://www.gloh2o.org/	Beck et al., 2017

883

884 ## Table 2. Model calibration parameters used

No	SWAT Parameter	Description	Min.	Max.	Calibrated Value
1	CN2	SCS runoff curve number	-5	5	4.7
2	ALPHA_BF	Baseflow a factor for recession constant	-0.2	0.2	0.0
3	GW_DELAY	Groundwater delay	0	1	0.1
4	GWQMN	Minimum threshold depth of water in the shallow aquifer for return flow to occur	30	450	28.7
5	GW_REVAP	Groundwater ‘‘revap’’ coefficient	0	2	0.6
6	ESCO	Soil evaporation compensation factor	0	0.2	0.1
7	CH_N2	Manning’s ‘‘n’’ value for the main channel	0.8	1	0.9
8	CH_K2	Effective hydraulic conductivity in the main channel	0	0.3	0.2

9	OV_N	Manning's ''n'' value for overland flow	5	130	44.0
10	SLSUBBSN	Average slope length	0.01	30	17.5
11	HRU_SLP	Average slope steepness	10	150	72.8
12	ALPHA_BNK	Bank flow recession constant /const. of proportionality	0	0.6	0.3
13	SOL_AWC	Available soil water capacity	0	1	0.3
14	SOL_K(1)	Saturated Hydraulic Conductivity	-0.2	0.4	-0.1
15	SOL_BD(1)	Bulk density of the layer	-0.8	0.8	0.6
16	SFTMP	Snowfall temperature	-0.5	0.6	-0.4
17	SMFMX	Melt factor on June 21	0	10	2.3
18	SMFMN	Melt factor on December 21	0	10	0.5
19	TIMP	Snow Temprature Lag Factor	0.01	1	0.5
20	RCHRG_DP	Aquifer percolation coeficient	0	1	0.1
21	PLAPS	Precipitation laps rate	0	1000	496.3
22	TLAPS	Temprature laps rate	-10	-0.1	-4.3
23	SLSOIL	Hillslope length	0	150	95.9
24	LAT_TTIME	Lateralflow travel time	0	180	32.4
25	SNO50COV	Fraction of SNOCOVMX that provides 50% cover	0.01	0.9	0.4
26	SNOCOVMX	Threshold depth of snow above which there is 100% cover	0	500	197.0

885

886

887

888

889 **Table 3. Hydrological performance evaluation for the SWAT simulations using the**

890 **gridded precipitation products.**

891

Metrix	CFSR	ERA Interim	APHRODITE	MSWEP
NS	0.22	0.53	0.38	-0.73
RSR	0.42	0.64	0.63	0.51
PBIAS	33.40	19.90	41.40	85.00
R^2	0.77	0.69	0.75	0.61
POD	0.91	0.79	0.71	0.93
FAR	0.62	0.62	0.61	0.61

892

893

894

895

896

897

898

899 **Table 4. Statistics of comparison amongst the various simulations flows with the**

900 **gridded precipitations used as climate forcing.**

901

	OBSERVED (Flow)	MGM (simulation)	CFSR (simulation)	ERA (simulation)	APHRODITE (simulation)	MSWEP (simulation)
Mean	**13.32**	11.65	17.76	**14.90**	7.28	22.99
St. Dev.	**12.09**	9.34	18.71	**11.03**	6.39	17.74
Min	3.27	1.09	1.05	1.30	0.35	1.68
Max	56.92	45.23	90.14	53.74	25.10	78.88

902

903

904

905

906

907

908

909

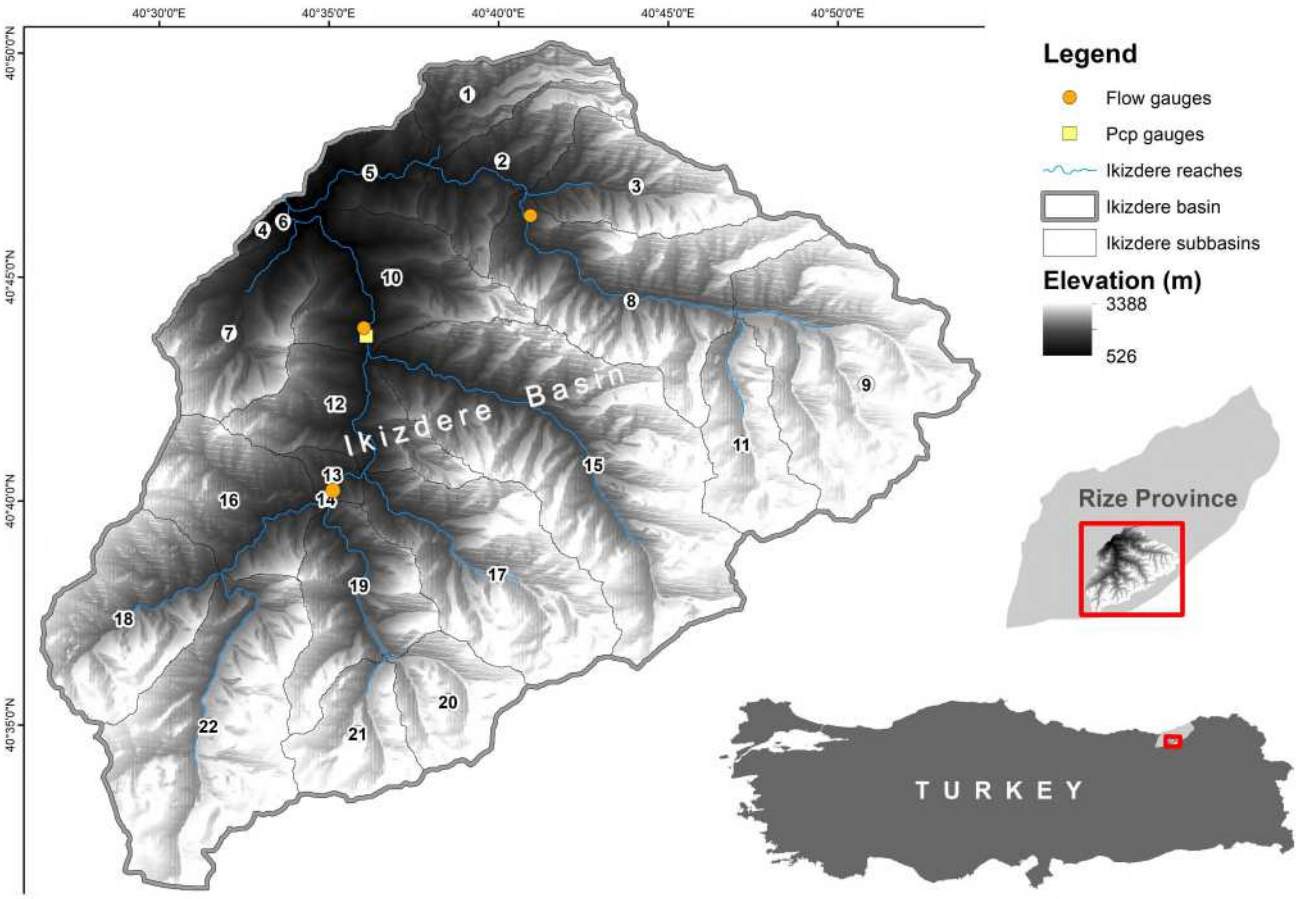
910

911

912

913

914 **FIGURES**



915

916 **Figure 1. The study area and flow observation stations in Rize province.**

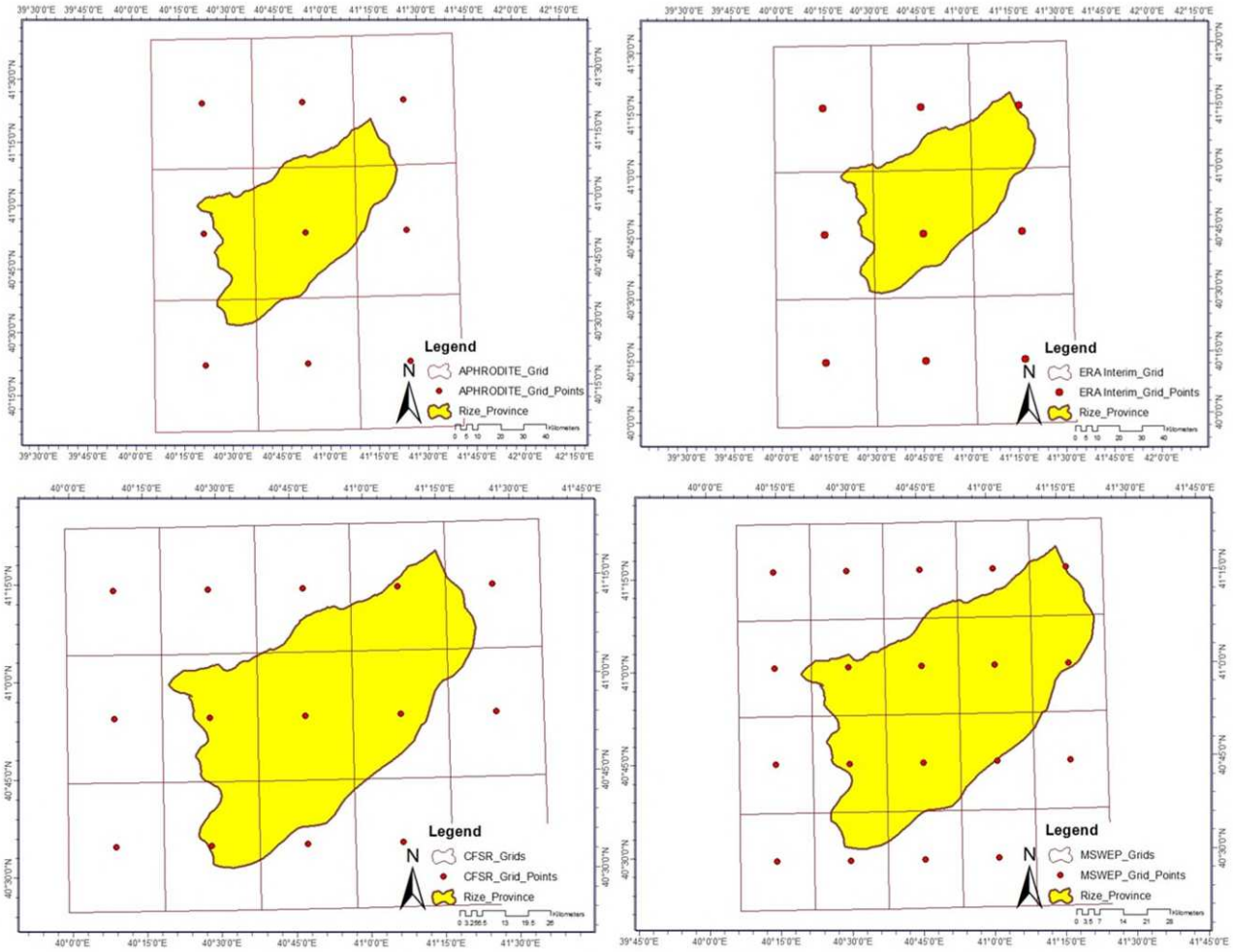


Figure 2. The plot of precipitation datasets grid mesh.

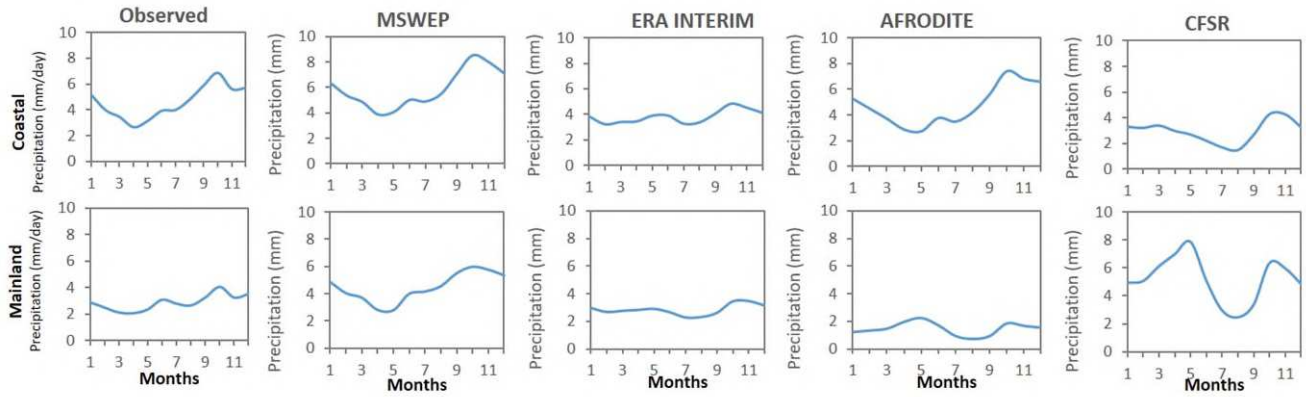


Figure 3. Monthly long term average graph for precipitation datasets.

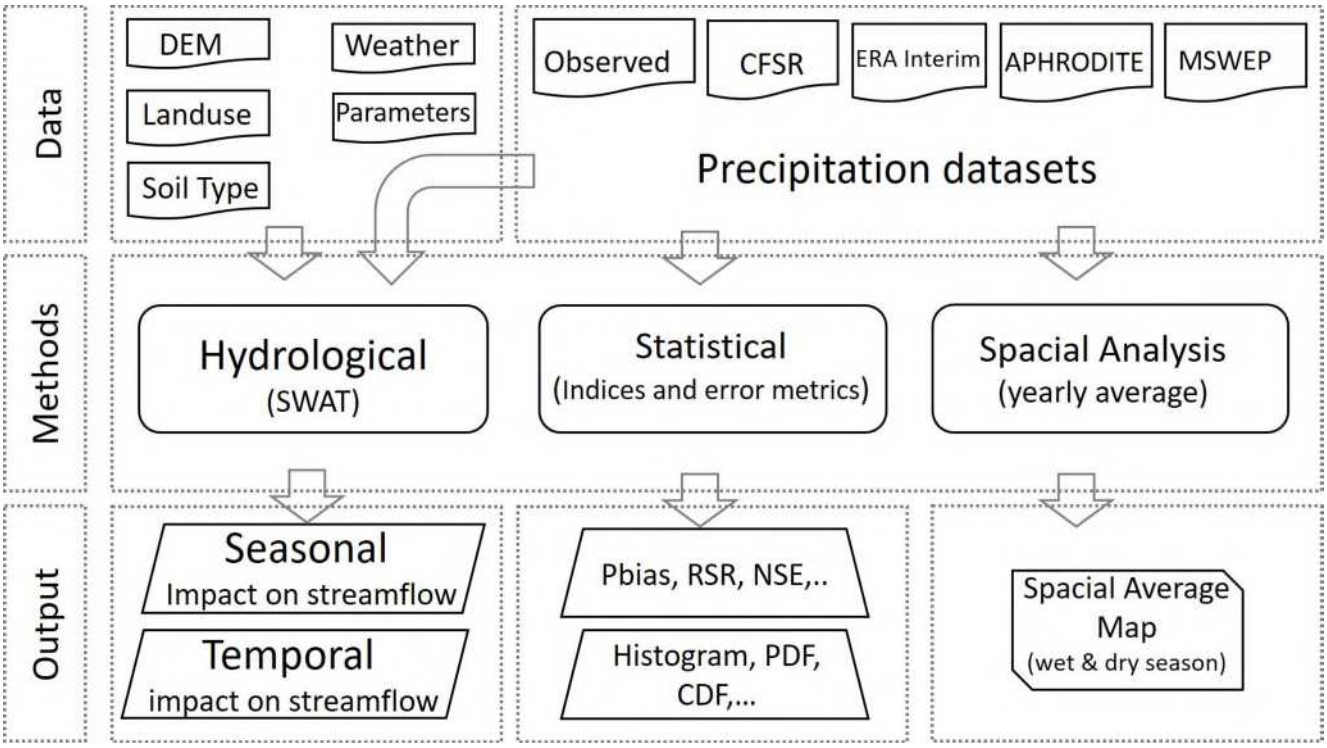


Figure 4. Flow chart of the input–output setup adopted in the study.

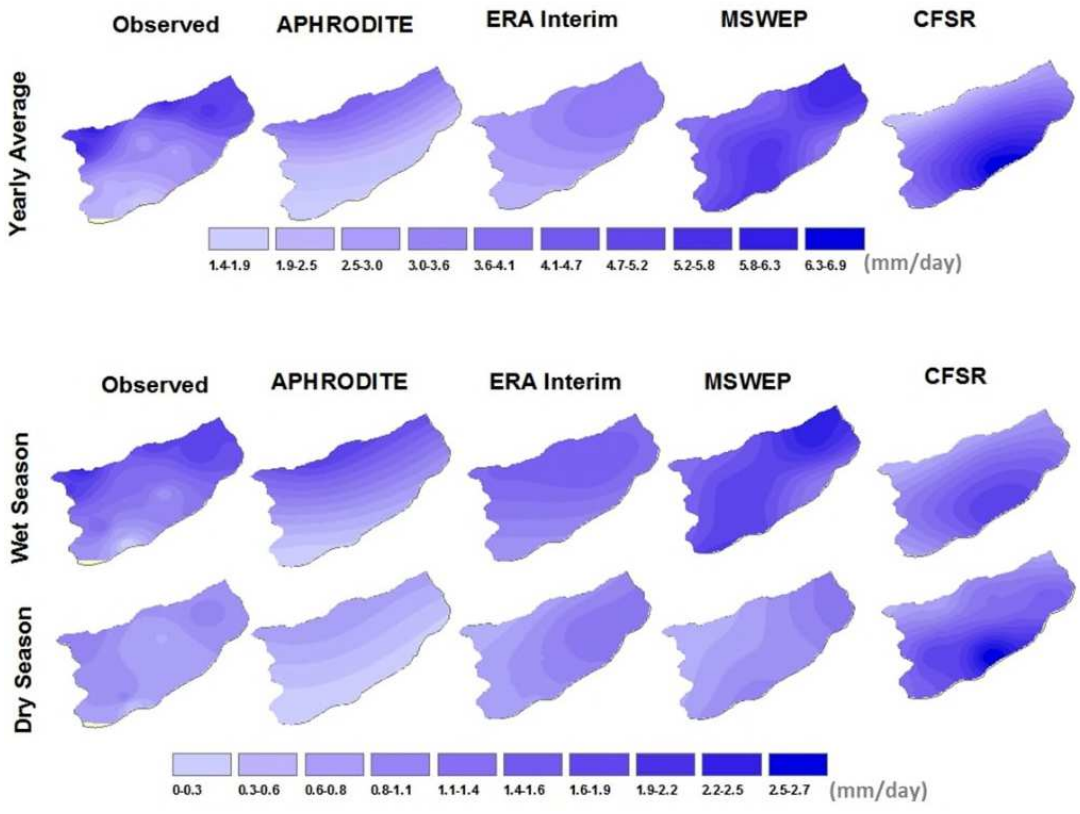
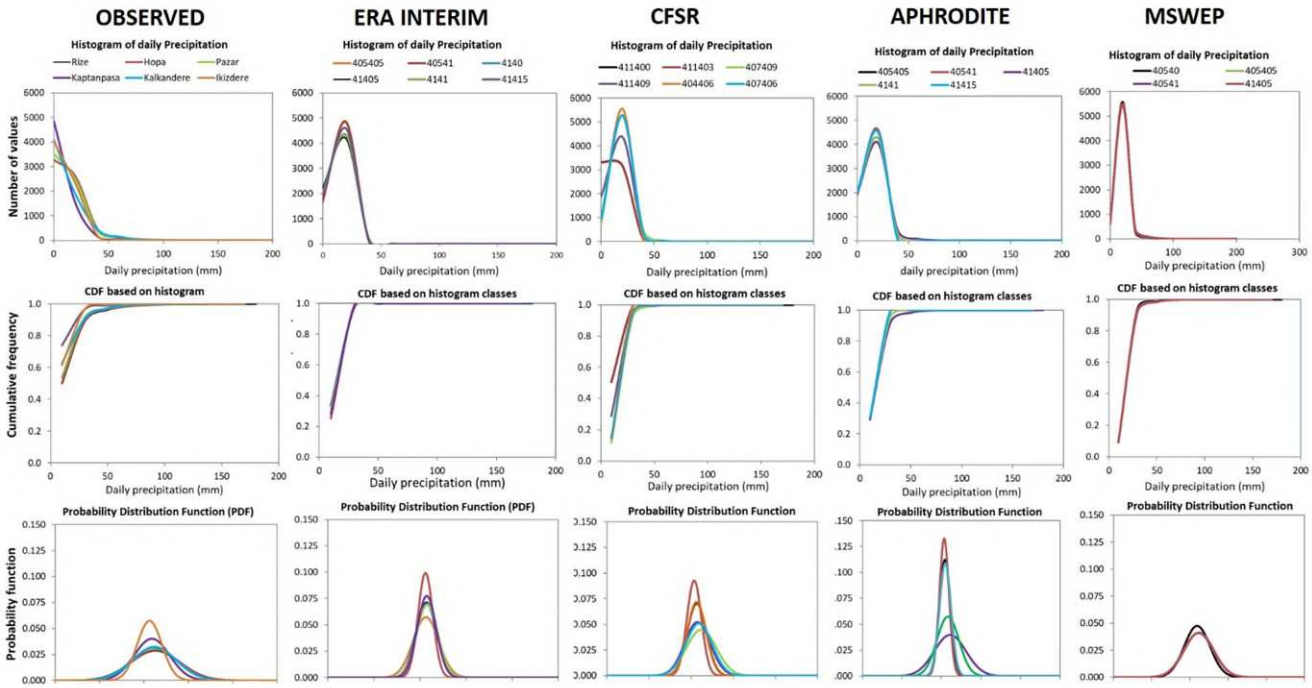


Figure 5. The long term average precipitation (mm/day) over the study area.



927

928 **Figure 6. The histogram (top), cumulative distribution function (middle), probabilistic**

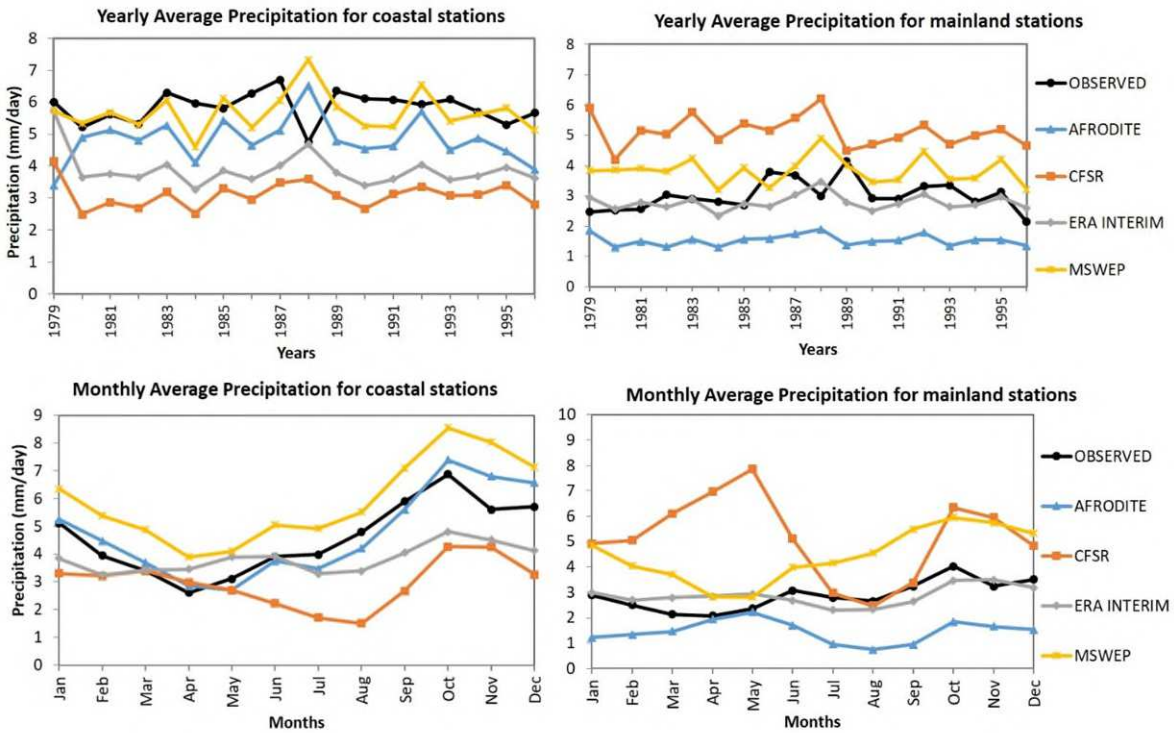
929 **distribution function (bottom) plots of the observed and the gridded precipitation**

930 **products.**

931

932

933



934

935 **Figure 7. Yearly average (top) and Monthly average (bottom) precipitation for the**

936 **coastal and mainland regions.**

937

938

939

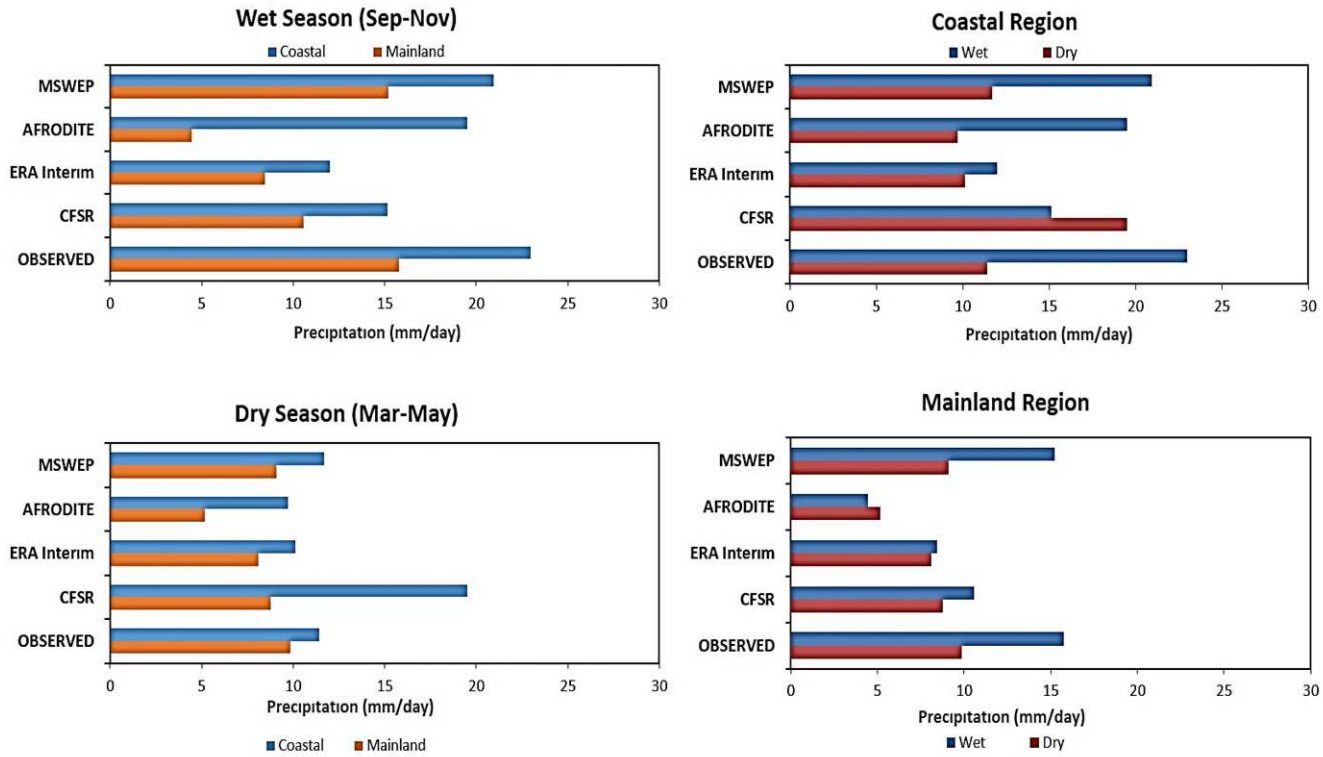
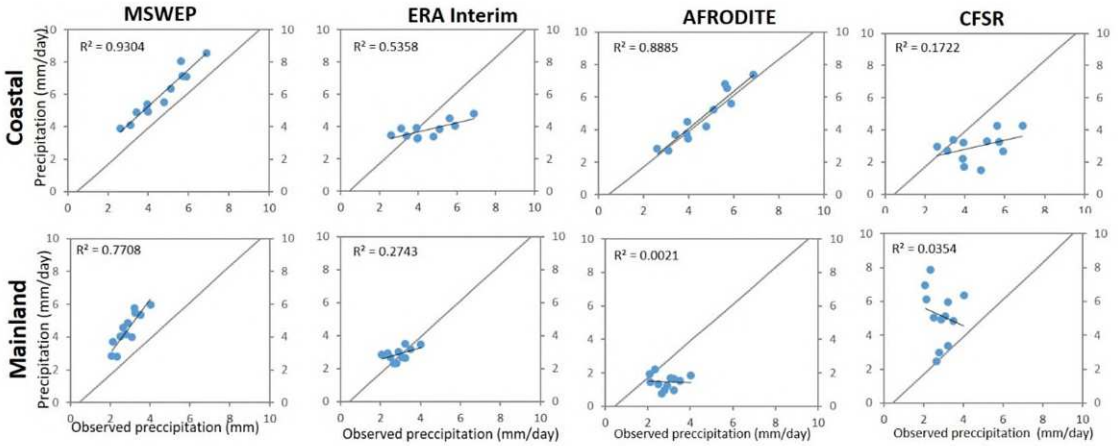
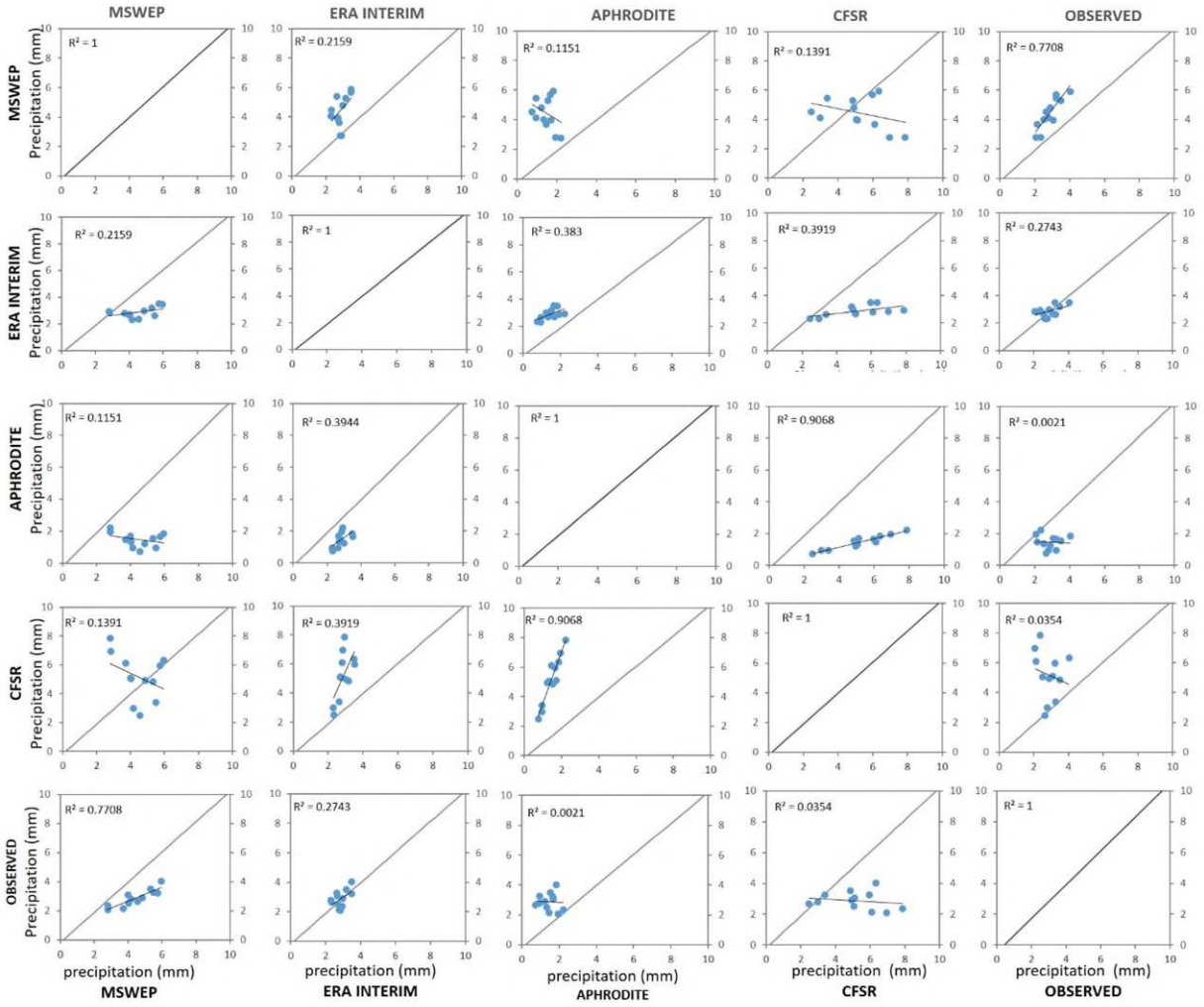


Figure 8. The long year average precipitation of the reanalysis/merged products for the coastal/mainland regions (top), and wet/dry seasons (bottom).



950

951 **(a)**



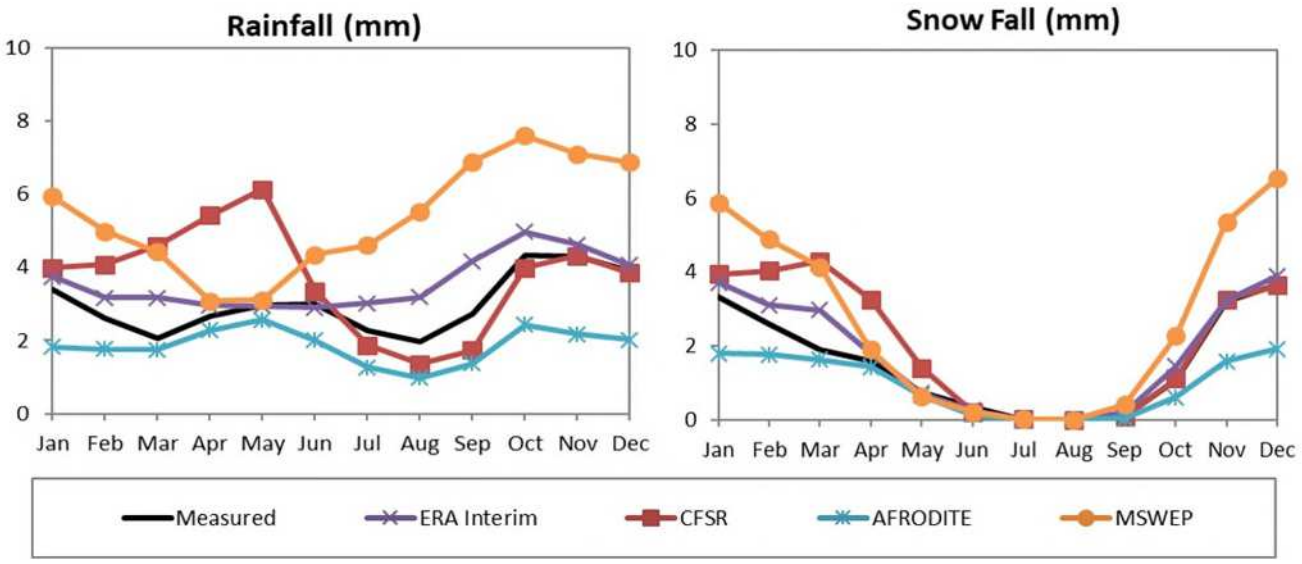
952

953 **(b)**

954 **Figure 9. Comparison plots for the monthly long term averages of precipitation for**

955 **mainland and coastal regions (a), Inter-comparison of all datasets (b).**

956



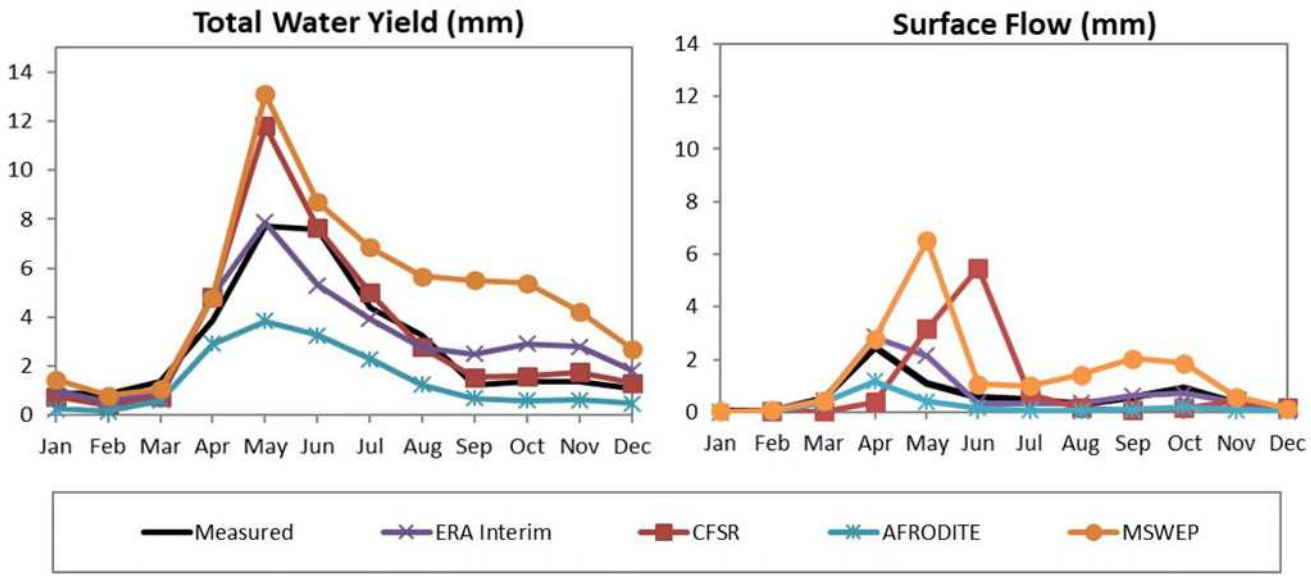
957

958 **Figure 10. Comparison of the total monthly rainfall and snowfall of the**

959 **reanalysis/merged products with the observations for the study area.**

960

961



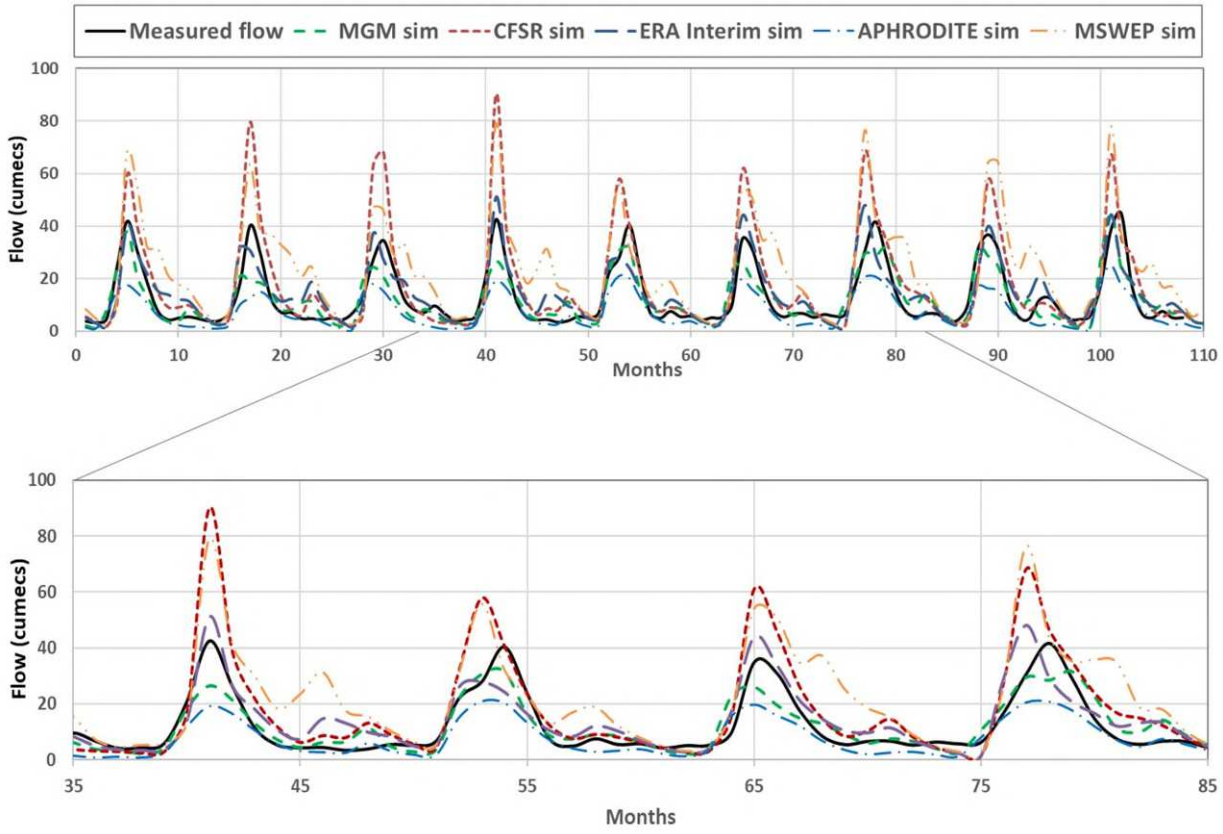
962

963 **Figure 11. The measured versus simulated monthly total water yield and streamflow**

964 **comparison (mm/day) computed using the various weather inputs in the SWAT**

965 **simulations.**

966



967

968 **Figure 12. SWAT simulations of the various precipitation products compared with**

969 **observed flow at Ikizdere basin outlet.**

970

971