Journal of Hydroinformatics

Journal of Hydroinformatics Vol 25 No 4, 1153 doi: 10.2166/hydro.2023.141

Statistical blending of global-gridded climatological products: an approach to inverse hydrological model

Rahimeh Mousavi [a], Mohsen Nasseri [b,*] and Saeed Abbasi [a]

[a] Department of Civil Engineering, Faculty of Engineering, University of Zanjan, Zanjan, Iran
[b] School of Civil Engineering, College of Engineering, University of Tehran, Tehran, Iran
*Corresponding author. E-mail: mnasseri@ut.ac.ir

RM, 0000-0002-4301-5400; MN, 0000-0002-7584-7631; SA, 0000-0003-1847-2105

ABSTRACT

The growing use of global-scale environmental products in hydro-climatic modeling has increased the variety of their applications and the complications of their uncertainties and evaluations. Researchers have recently turned to statistical blending of these products to achieve optimal modeling. The proposed statistical blending in this study includes five large-scale and satellite precipitation (CHIRPS, ERA5-Land of ECMWF, GPM (IMERG), TRMM, and Terra) and evapotranspiration (GLEAM, SSEBop, MODIS, Terra, and ERA) products committed in three modeling scenarios. The blending procedures are organized using a conceptual water balance model to achieve the best precipitation and evapotranspiration results for the conceptual production of streamflow using hydrological inverse modeling. Based on the results, the proposed blending procedures of precipitation and evapotranspiration improved the performance of the model using different statistical metrics. In addition, the results show the conformity of the pattern and behavior of the blended precipitation calculated using the moving least square method in the study area. This happened by changing the estimation based on in situ values, particularly in cold months considering the orographic/snow effects. The combining method provides a good fusion procedure to improve the realistic estimation of precipitation and evapotranspiration in ungagged watersheds as well.

Key words: evapotranspiration products, global-gridded products, inverse modelling, precipitation products, statistical blending, water balance modeling

HIGHLIGHTS

- Five precipitation and evapotranspiration products have been used to blend as new input sets.
- Statistical linear blending method has been applied to determine the model's inputs via the concept of inverse hydrology.
- 100 combinations of precipitation–evapotranspiration sets have been statistically combined to calibrate the models.
- The proposed method provides a good fusion method to improve the estimation of climatological signals.

1. INTRODUCTION

The calculation of water balance components and their quantification are two challenges in water resources management. Previously, the main sources for calculating and estimating the components were ground measurements and statistical estimates (the experimental formula). Recently, most researchers tend to use large-scale or remotely sensed datasets or both for this purpose, and different research projects have been conducted to investigate using such data in hydrological models (Moreira *et al.* 2019; Guo *et al.* 2022). Most of these research projects have focused on evaluating the accuracy of the products and comparing them with observed and *in situ* values (Shayeghi *et al.* 2020). In addition, a few studies have also tested the performance of these datasets as input or state variables in hydrological models (Duan *et al.* 2019).

These products are practical tools to estimate the components of hydrological models, especially in areas with gapped, missed, or scattered ground information (Huang *et al.* 2020). However, this information sometimes has considerable uncertainty and error because of the low resolution of spatiotemporal ground information, which can be due to the remotely sensed information and the algorithms used in their estimations (Nguyen *et al.* 2021). Therefore, extensive efforts have been made to solve this problem lately. Following this issue, one of the most useful actions is to blend different remote sensing or large-scale

products to achieve higher accuracy, lower error, and improved spatiotemporal estimates (Aires 2014; Munier *et al.* 2014; Schoups & Nasseri 2021). The studies conducted in the field of blending remotely sensed information are divided into three categories:

- Blending ground and satellite data to obtain improved, consolidated information.
- Comparing different statistical and probabilistic methods to blend the initial datasets.
- Evaluation of the blended information's effectiveness in the secondary environment, like hydrological models.

Ground data and one or more large-scale datasets have been merged in the first category. They are used to improve the reliability of information coverage using various statistical and probabilistic methods, such as Xie and Arkin's maximum likelihood estimation method, nonparametric kernel smoothing, the Random Forest-based MErging Procedure (RF-MEP), and the Bayesian approach (Huffman *et al.* 1997; Xie & Arkin 1997; Li & Shao 2010; Rozante *et al.* 2010; Xie & Xiong 2011; Manz *et al.* 2016; Yang *et al.* 2017; Kimani *et al.* 2018; Baez-Villanueva *et al.* 2020; Lu *et al.* 2021; Ma *et al.* 2021; Nguyen *et al.* 2021; Zandi *et al.* 2022). In the second category, several titles can be mentioned compared to the first one. For instance, simple averaging, statistical collocation, and the Bayesian approach are the most implemented methods for blending large-scale datasets without ground information (Jongjin *et al.* 2016; Baik *et al.* 2018; Yumnam *et al.* 2022).

The third category included studies in which the second model (the water balance model in the current research) was used to validate (or calibrate) the blended products. This type of blending approach is the focus of the current research. Sahoo *et al.* (2011) and Pan *et al.* (2012) used various satellites and observed precipitation and EvapoTranspiration (ET) products to blend multiple sources of water cycle variables, using large-scale water balance and the assimilation approach. Aires (2014) proposed four different statistical methods to blend different datasets of multiple water balance components. Schoups & Nasseri (2021) suggested a new probabilistic method to combine different climatological variables to close a large-scale water balance model.

The noble point of the current research is to blend various ET and precipitation products using inverse conceptual hydrological modeling. In the previous studies (Pan *et al.* 2012; Aires 2014; Schoups & Nasseri 2021), a large-scale water balance model (a macro-scale terrestrial water balance model) has been used, but in the current research, the terminology has been developed for the conceptual model in a mid-size watershed scale. The main idea of inverse hydrological modeling (or backward hydrology) has been introduced and developed by Vrugt *et al.* (2008), Kirchner (2009), and Herrnegger *et al.* (2015). It means going backward from streamflow as the output variable to define (or tune) precipitation, ET, or both as the input variables via statistical and probabilistic methods. In this research, different precipitation products (Climate Hazards Group Infrared Precipitation with Stations (CHIRPS), ERA5-Land of the European Center for Medium-Range Weather Forecasts (ECMWF) (ERA), Integrated Multi-Satellite Retrievals for GPM (IMERG), Tropical Rainfall Measuring Mission (TRMM), and Terra) and Actual EvapoTranspiration (AET) datasets (Global Land Evaporation Amsterdam Model (GLEAM), SSEBop, Moderate Resolution Imaging Spectroradiometer (MODIS), Terra, and ERA) are used to combine linearly through an inverse modeling approach. The following section presents and discusses the used information, the study area, and the modeling subject.

2. STUDY AREA AND DATA DESCRIPTION

2.1. Study area

The mountainous basin of Gheshlagh is one of the sub-basins of the border river watershed in the west of Iran and has an area of 1,062.12 km^2. The geographic coordinates of this basin are located between 46° 46′ and 47° 20′ east longitude and 35° 24′ and 35° 43′ north latitude (Figure 1). Due to its mountainous nature, the watershed is snow-covered during December to March. In addition, the case study has no alluvial aquifer (based on the official report of Iran Water Resources Management Company-IWRMC). The average annual rainfall of the basin is 454.4 mm, and the average annual temperature is 13.3 °C.

2.2. Ground data

Since the conceptual and monthly water balance models are used, it is necessary to aggregate the daily *in situ* information of precipitation, streamflow, temperature, and pan evaporation datasets collected by the IWRMC climatological and hydrometric stations. The quality of the datasets has been checked using run-test and double-mass curve methods. Then, climatological information (including precipitation, temperature, and pan evaporation datasets) was spatially estimated over the watershed with a 5 km resolution using the Moving Least Square (MLS) method on the computational grids

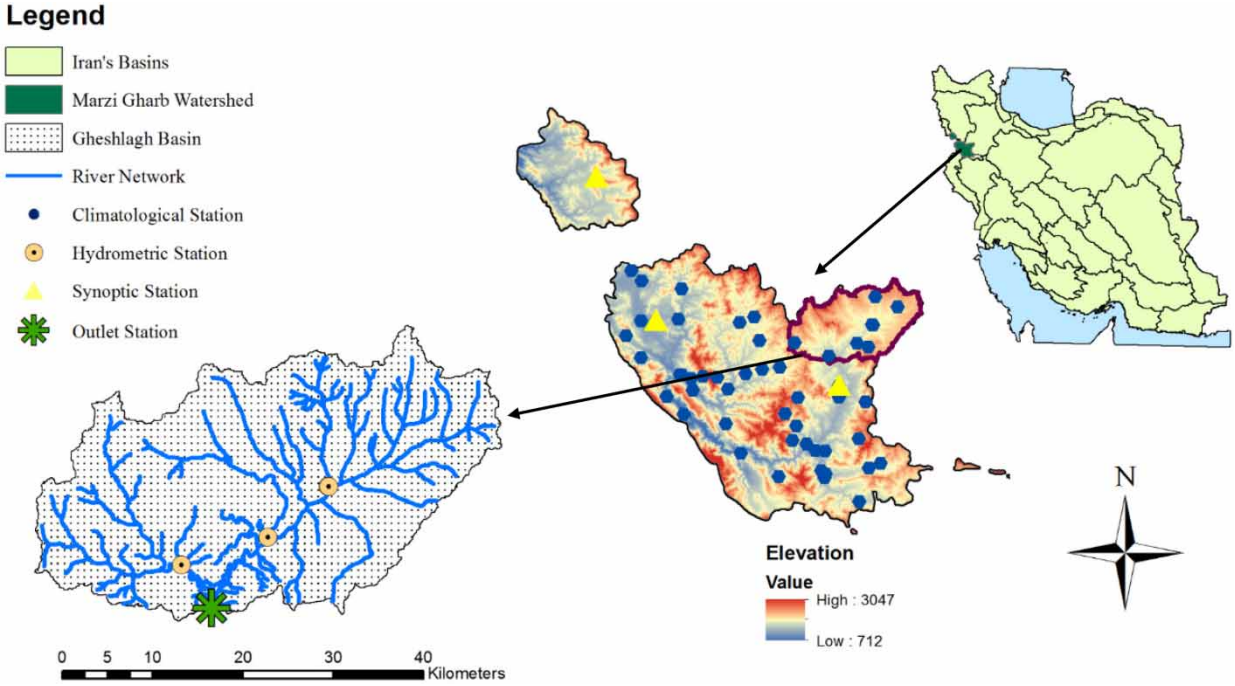

Figure 1 | Location of the Gheshlagh watershed and its river network, climatological, and hydrometric stations.

(Lancaster & Salkauskas 1981; Amini & Nasseri 2021). To do so, geographical characteristics of altitude, longitude, and latitude values of the stations (climatological and synoptic stations) are considered as input variables to calibrate the MLS model and the input variables have been estimated over the computational grids. In addition, due to the full coverage of information during the computational period (from 2003 to 2015), gap-filling methods are not applicable. The input variables of the selected monthly water balance model have been prepared by aggregating the spatially estimated precipitation, temperature, and pan evaporation datasets. It is worthwhile to mention that the watershed has been the subject of a few previous research projects by Sadeghi *et al.* (2013), Taheri *et al.* (2022), and Mousavi *et al.* (2022). The outlet of the selected basin is Gheshlagh Dam and the input streamflow to the reservoir is the target of the water balance modeling, as well.

2.3. Global-gridded climatological products

The specifications of the recommended large-scale and gridded climatological products of precipitation and AET used in the current research (based on Javadian *et al.* 2019; Moshir Panahi *et al.* 2021; Schoups & Nasseri 2021; Ghomlaghi *et al.* 2022; Nasseri *et al.* 2022) are briefly discussed as follows.

2.3.1. Global-gridded precipitation products

In the current research, five global-gridded and large-scale precipitation datasets (including CHIRPS, IMERG, ERA5-Land of the ECMWF, TRMM, and Terra Climate) with monthly temporal resolution have been used.

Table 1 shows their spatiotemporal specifications and descriptions. The CHIRPS, which combines satellite images and ground station data for generating rain time series to analyze precipitation trends and monitor the drought, was developed by the US Geological Survey (USGS) and the Climate Hazards Group at the University of California, Santa Barbara (UCSB) (Funk *et al.* 2014).

A combination of multiple satellites' microwave precipitation products, microwave-calibrated infrared (IR) satellite estimates, values of ground gauges, and other precipitation estimators on a suitable time and space scale forms the IMERG algorithm which is the GPM satellite's product for estimating surface precipitation (Huffman *et al.* 2019). The re-analyzed dataset named ERA5-Land (ERA), which combines historical observations of global estimates with meteorological and data-driven models, was developed by the ECMWF and has been available monthly since 1950 (Hersbach *et al.* 2019).

Table 1 | Descriptions of the used global-scale precipitation and ET products

Product		Description	Time coverage	Spatial resolution	Time resolution	Access link
Precipitation	CHIRPS		1981 to present	5 km	Monthly, 5-month, and daily	https://earlywarning.usgs.gov/fews/datadownloads/Global/CHIRPS%202.0
	IMERG		2000–2019	10 km	Monthly	https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGM_06/summary
	TRMM		2000–2016	25 km	Monthly	https://pmm.nasa.gov/dataaccess/downloads/trmm
	ERA-5	Reanalysis product	1950 to present	10 km	Monthly	https://www.ecmwf.int/
	Terra		1958–2021	4 km	Monthly	http://www.climatologylab.org/
Evapotranspiration	MOD16 ET	Penman–Monteith	2001 to present	500 m	8 day	https://search.earthdata.nasa.gov/search?q=C1000000524-LPDAAC_ECS
	GLEAM v3	Priestley–Taylor	2003–2015	25 km	Daily, monthly	https://www.gleam.eu/
	SSEBop	Penman–Monteith	2003 to present	1 km	10 day, monthly	https://earlywarning.usgs.gov/fews
	ERA5	Reanalysis product	1981 to present	9 km	Monthly	https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means?tab=overview
	Terra	Penman–Monteith	1981 to present	4 km	Monthly	http://thredds.northwestknowledge.net:8080/thredds/terraclimate_aggregated.html

The Terra product, available monthly since 1958, combines CRU Ts 4.0 and 55-year-old Japanese re-evaluated products (Abatzoglou *et al.* 2018). The TRMM integrates microwave and IR satellite information and is a joint space mission between NASA and the Japan Aerospace Exploration Agency (JAXA) (Huffman *et al.* 2010).

It should be noted that only the CHIRPS dataset has the same spatial resolution with the selected computational grid (5 km), and the other products have been resampled on the spatial grids.

2.3.2. Global-gridded ET products

Table 1 summarizes the specifications of the large-scale ET information used in this study. The product of MODIS (Moderate Resolution Imaging Spectroradiometer) with the name MOD16 has been produced on a global scale with a spatial accuracy of 500 m based on the Penman–Monteith equation logic since 2001 (Running *et al.* 2017). The GLEAM is a set of large-scale climatic results that separately estimate the various components of ET. The product is based on the Priestley–Taylor equation (Martens *et al.* 2017).

The Simplified Surface Energy Balance (SSEB) approach has been used to develop the SSEBop AET product since 2003. It is accomplished by combining the ET produced by MODIS 10-day remote sensing thermal images with the reference ET. The main structure used for estimating the ET product is based on the use of applied hot or cold pixels in SEBAL (Bastiaanssen *et al.* 1998) and METRIC (Allen *et al.* 2007) methods (Senay *et al.* 2020).

Since 1958, Terra Climate has been a collection of monthly climate datasets on global-scale terrestrial surfaces. With a monthly spatiotemporal resolution of 4 km, this dataset provides critical inputs for global-scale environmental and hydrological studies. Terra Climate also generates monthly surface water balance components using a water balance model (Abatzoglou *et al.* 2018).

ERA5-Land of ECMWF (ERA), the last used re-analyzed dataset, has been modified using observations from around the world into a complete and consistent global dataset via a physically based model since 1981. The spatial accuracy of the product is about 9 km (Hersbach *et al.* 2019).

3. METHODOLOGY

In this section, the proposed method of blending different climatological information using water balance models, the selected uncertainty assessment approach, and evaluation metrics have been introduced.

3.1. Proposed statistical blending method

3.1.1. Water balance model

The water balance model used in the current research is based on Wang's water balance model (Wang *et al.* 2011). The model focuses on monthly hydro-climatic structure and surface processes to simulate streamflow and its water balance components. The usual structure of Wang's water balance model was implemented in this research (in the Supplementary material, Figure S1). However, its snow hydrology because of snow-covered and mountainous nature of the area has been modified using the exponential degree-day relationship to separate rainfall and snow from the monthly precipitation values used by Guo *et al.* (2005). However, the four parameters in the original model and to better consideration of snow storage and initial storage values (soil and snow storages), the modified water balance model has seven (Mousavi *et al.* 2022; Taheri *et al.* 2022). A comprehensive simulation to detect the snow hydrology of the watershed with the same water balance model has been reported by Taheri *et al.* (2022). The final water balance model has a single-layer soil storage to play role in generating subsurface flow in the model, without considering groundwater and aquifer storage. Therefore, the final streamflow in the model is the sum of surface, subsurface, and snowmelt flows. The flow diagram of the used water balance model has been depicted in the supplementary materials (Supplementary material, Figure S1).

3.1.2. Statistical blending method

Blending the large-scale climate products provides more reliable and compatible inputs with the recorded streamflow of a river network (Aires 2014; Schoups & Nasseri 2021). For example, precipitation is calculated using recorded information from rain gauge stations, and its results have been used in lump, semi-distributed, and distributed hydrologic models. Because of the low density of stations or recorded values tied to different uncertainty sources (especially in high-altitude areas), these uncertain estimations cause inconsistency in the hydrological modeling response. Therefore, merging or combining raw precipitation datasets (without any correction with ground information) creates a chance to use the advantages of each precipitation product (even recorded values) in simulating the hydro-climatological behavior of the considered areas. On the other hand, the ET values are environmental variables that can be calculated or measured in different ways. Its appropriate determination is one of the most important challenges in hydrological and water resource modeling. Combining different ET products and determining their acceptable values are the goals of this research.

Against the usual hydrological modeling that goes from the main drivers to streamflow, the purpose of blending is fulfilled in the current study via inverse modeling, which means moving backward from observed streamflow to the main hydrologic drivers such as precipitation and ET. This hydrological modeling approach has formerly been used by other researchers, such as Herrnegger *et al.* (2015) and Vrugt *et al.* (2008).

To achieve the blending input variables for water balance modeling and their evaluations, three different scenarios have been introduced. The first scenario, known as the reference scenario, used the observed precipitation values and an experimental method to estimate ET and recorded streamflow at the basin outlet to calibrate the parameters of the selected water balance model. The calibrated parameters and AET time series have been used for the evaluation of the proposed scenarios in the next steps. It should be noted that the spatially averaged monthly precipitation in the basin of interest has been calculated using the MLS method.

The second scenario is based on the use of linearly blended precipitation and ET products as the model's inputs (instead of using the recorded precipitation values and the usual experimental relationship of ET) to calibrate the observed streamflow values at the basin outlet. In this scenario to avoid overfitting in the internally calibration procedure and based on the literature (Aires 2014; Schoups & Nasseri 2021), the linear combination of precipitation and ET products is done by adding two new parameters to the calibration procedure of the reference scenario. The linear combination of the products is aimed as below,

$$v = \alpha \times v_i + (1 - \alpha) \times v_j \tag{1}$$

where v_i and v_j represent the same product values of different datasets (precipitation or ET), α represents the blending ratio with the purpose of combination. The blending ratio in the current modeling has a value range of [0, 1]. The third scenario is

similar to the second scenario. The only difference is that the best-blended products, considering the computed streamflow, are used in the third scenario for uncertainty assessment.

3.2. Uncertainty assessment of the water balance models

Most of the previous studies with the aim of modifying and combining large-scale remote sensing products are based on validation, analysis, comparison, and evaluation with *in situ* information. The effects of using these products (or their modified ones) on the uncertainty behavior of hydrological models and applications have not been studied. This type of evaluation is one of the other fundamental aspects of the current research. Generally, uncertainty originates from three sources, including model structure, datasets (including input and output), and model parameters (Chen *et al.* 2016). Various methods have been proposed in the literature to evaluate the uncertainty caused by the mentioned causes. The Generalized Likelihood Uncertainty Estimation (GLUE) method, proposed by Beven & Binley (1992), has been employed to assess the model parameters' uncertainty in this research. The basis of the method is the sequential sampling-simulation approach, considering their performance. GLUE has been widely used in evaluating the uncertainty of hydrological models individually or as a benchmark method (Ahmadi & Nasseri 2020). GLUE has no limit for the number of parameters and ultimately provides the overall uncertainty of the model structure and its parameters (Beven & Binley 1992). At this step, to evaluate the effectiveness of using the investigated scenarios in calibrating the balance model (the reference and modified models with the combination of remote sensing information), their parametric uncertainties were measured using the GLUE method. To maintain the effect of the generated random parameters, Nash-Sutcliffe Efficiency (NSE) has been considered to have an acceptance level threshold of 0.75. With this desirability limit, 10,000 acceptable random parameter sets (with an NSE of greater than 0.75) were included to form the statistical distribution of the parameters.

3.3. Modeling evaluation

In the following, different statistics are introduced to assess the behavior of the calibrated models and the effects of the proposed scenarios. In addition to the selected three metrics that have been used to evaluate the modeling uncertainty, two statistics have been used to measure the certain behavior of the models. NSE is the first examined criterion, indicating the consistency of the observational and computational data (Nash & Sutcliffe 1970). Equation (2) shows the mathematical formulation of NSE,

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} (\bar{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - y)^2} \tag{2}$$

where y_i, $\bar{y}_i$, and y represent the observed, computed, and observed mean values at the time i, respectively. NSE value varies between $[-\infty, 1]$, and when it equals 1, a complete similarity is observed. In addition, its values between [0, 1] are generally reported as acceptable performance levels. Its values less than zero indicate that the average observational values are better predictors of the simulated values, which means unacceptable performance (Moriasi *et al.* 2007). The Kling–Gupta Efficiency (KGE) metric is the second statistic to evaluate the deterministic behavior of the developed models. KGE, as the multifactorial similarity index, is based on correlation, bias changes, and bias mean, which eliminates NSE deficiencies and is used increasingly for model calibration and evaluation (Kouchi *et al.* 2017; Dembélé *et al.* 2020). Equation (3) shows the way that KGE is measured.

$$\mathrm{KGE} = 1 - \sqrt{(R(\bar{y}_i.y_i) - 1)^2 + \left(\frac{\mathrm{var}(\bar{y}_i)^{1/2}}{\mathrm{var}(y_i)^{1/2}} - 1\right)^2 + \left(\frac{\overline{y'}}{y} - 1\right)^2} \tag{3}$$

where $y_i, \bar{y}_i, y, y'$ represent the observed, computed, observed mean, and computed mean values of the streamflow at the time i, respectively. KGE values vary from negative i to the desired value of a variable. The values of these indicators regarding the model's performance have been categorized in articles by Moriasi *et al.* (2007) and Kouchi *et al.* (2017). The desirability ranges of the above evaluation indicators are presented in Table 2 based on experience in the literature.

Three statistical indicators were used to evaluate the uncertainty simulation. These metrics are Average Relative Interval Length (ARIL) (the uncertainty bandwidth of the model response to observed values), P_{level} (the percentage of observed values categorized by the uncertain bands), and Normalized Uncertainty Efficiency (NUE) (Jin *et al.* 2010; Li *et al.* 2010; Nasseri *et al.* 2013).

$$\text{ARIL} = \frac{1}{n}\sum \frac{\text{UPL}_t - LOL_t}{Q_{\text{obs},t}} \tag{4}$$

$$P_{\text{level}} = \frac{\text{NQ}_{\text{in}}}{n} \times 100 \tag{5}$$

$$\text{NUE} = \frac{P_{\text{level}}}{w \times \text{ARIL}} \tag{6}$$

which UPL and LOL are the upper and lower values of the simulated uncertainty. Q_{obs} is the observed streamflow, NQ_{in} is the number of observations placed within the uncertainty band, N is the number of the observations, and w is the effective weight between ARIL and P_{level} that is assumed to be 1 considering the same importance of the metrics in the current research.

4. RESULTS

4.1. Calibration of the water balance models

A decimal-based and single-objective genetic algorithm has been used to calibrate the water balance model. The computational period was from 2003 to 2015 due to the availability and coverage of ground and satellite data. This time range (149 months) was randomly divided into calibration (104 months) and validation (45 months) periods, which are considered the same throughout the evaluation of different scenarios of this model. The fitness function is set to maximize the NSE metric between the observed and calculated streamflow values of the outlet. The statistical metrics and parameters of the basic model and the second scenario models, which include 101 models in the study area, are presented in Supplementary material, Table S1. In addition, five statistics of the water balance models (Mean Squared Error (MSE), Normal Mean Squared Error (NMSE), Root Mean Square Error (RMSE), coefficient of determination (R^2), and Taylor skill's score (TS)) have been reported in Supplementary material, Table S1. The seven significant parameters of the model, the initial range, and its optimized value in the reference model, which is called the first scenario, are reported in Table 3. The initial limits of the parameter range are on loan from Wang *et al.* (2013).

According to Table 3, the acceptability limits of the models for simulating the streamflow surpass 0.5 with the values of the NSE and KGE indices in the monthly time step at the basin scale (Kouchi *et al.* 2017). In other words, models with more significant index values (more than 0.5) are acceptable. Therefore, these values were selected as adequate limits in the statistical evaluation to validate the models. Considering the combination of five precipitation and five ET products in the second scenario, 100 models were calibrated. According to the results achieved from the comparison of the statistical indicators (details in Supplementary material, Table S1), the first and second scenarios have almost the same level of performance. About 60% of the calibrated models have met the presented statistical requirements (see Table 2). The next two conditions have also been used to filter the accepted models.

The first condition for model selection is an NSE metric greater than the reference model during the calibration and validation periods (Figure 2(a)). Eighteen models had come across the condition among 100 calibrated models in which

Table 2 | Different classes of statistical performance (Kouchi *et al.* 2017)

Performance rating	NSE	KGE
Very good	$0.75 < \text{NSE} < 1$	$0.9 < \text{KGE} < 1$
Good	$0.65 < \text{NSE} < 0.75$	$0.75 < \text{KGE} < 0.9$
Satisfactory	$0.5 < \text{NSE} < 0.75$	$0.5 < \text{KGE} < 0.75$
Unsatisfactory	$\text{NSE} < 0.5$	$\text{KGE} < 0.5$

Table 3 | Parameters' descriptions, their available bounds, and optimum values

Parameter	Description	Maximum	Minimum	Optimum
K_s	Surface runoff coefficient	2	0	0.20
K_g	Subsurface runoff coefficient	2	0	0.03
K_{sn}	Snow melt coefficient	2	0	0.00
$S_{\max}$	Maximum soil moisture storage	300	200	269.87
T_{snow}	Snowfall threshold temperature	0	−12	−10.00
T_{rain}	Rainfall threshold temperature	12	0	7.40
SN_1	Initial snowpack	100	0	89.57

the value of the NSE index in both calibration and validation periods was higher than the achieved values in the reference model by values of 0.73 and 0.54, respectively.

The second condition is a KGE index greater than the reference model in the calibration and validation periods (Figure 2(b)). The KGE metrics for 35 models were greater than 0.75 and 0.77 for the calibration and validation periods (for the reference model), respectively.

In the next step, the models that met the aforementioned conditions were checked using other criteria. In this investigation, two parameters, alpha (α) and beta (β), which are respectively optimized as the blending coefficients for the precipitation and ET products of the models, should be calibrated in the range of (0, 1) (not exactly equal to zero or one). This means that the two products are combined, and neither is used alone in the model. By applying these two conditions to the previous models, 12 models were selected out of calibrated ones (Figure 2(c)).

The diameter of the drawn circles in this figure is proportional to the multiplication root value of two parameters, alpha and beta. The smaller the diameter of these circles (circles with a darker color), the smaller the optimized parameter. The details of the indicators and optimized parameters of the 12 models are mentioned in the supplementary materials (Supplementary material, Table S2). The important point in these selected models is that the ERA product has participated in all 12 models, which shows the product's greater contributions in this way of modeling than the other products. In choosing the models, superiority over the reference model has been considered. Therefore, it was concluded that the performance of the model will be in a good range (considering Table 2) if it has these three conditions simultaneously,

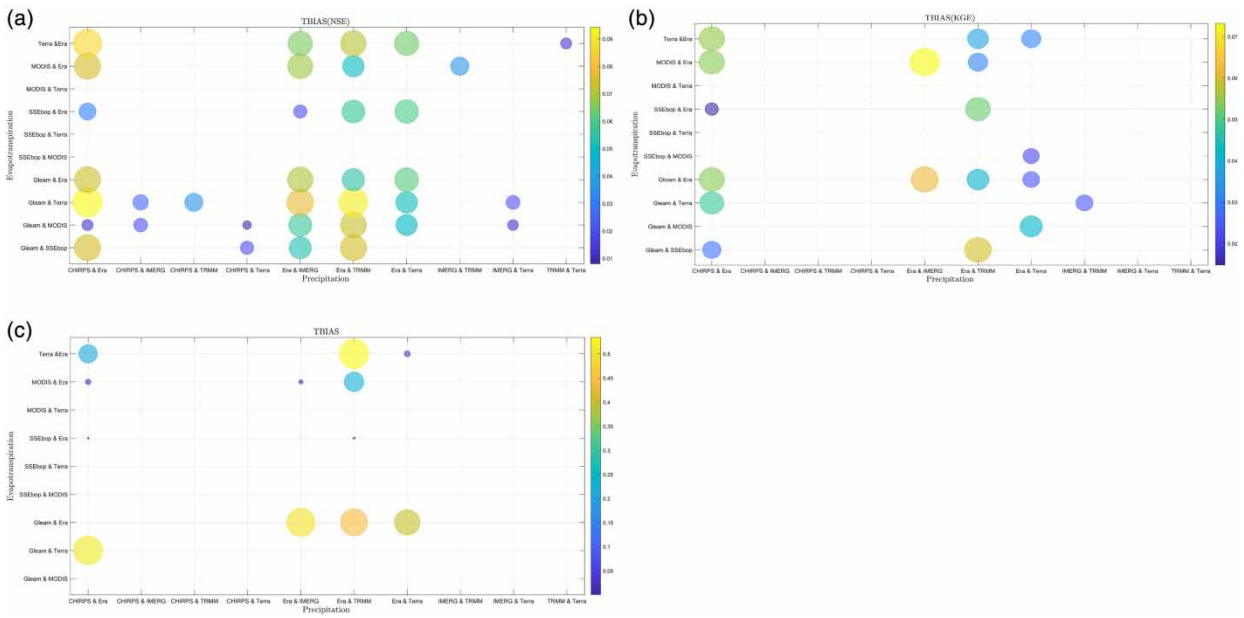

Figure 2 | The selected models based on (a) NSE, (b) KGE, and (c) alpha–beta values.

- An NSE value over 0.73 and 0.54 in the calibration and validation periods.
- A KGE value over 0.75 and 0.77 in the calibration and validation periods.
- The statistical superiority over the reference model.

Figure 3 shows the combined monthly precipitation and ET products in the selected moreover, reference models. The highest amount of *in situ* precipitation was 67.3 mm in April. In the same month, the lowest amount of precipitation (60.79 mm) belongs to the P ERA & TRMM-E SSEBop & ERA (blended precipitation of ERA and TRMM, and ET of SSEBop and ERA) model. The lowest amount of precipitation in the P ERA & Terra-E Terra & ERA model is 0.99 mm in August. The highest rainfall difference percentage is 263% in July and the smallest difference is 6% in January. In general, the combined precipitation amount agrees well with the lumped precipitation values based on *in situ* data. November and

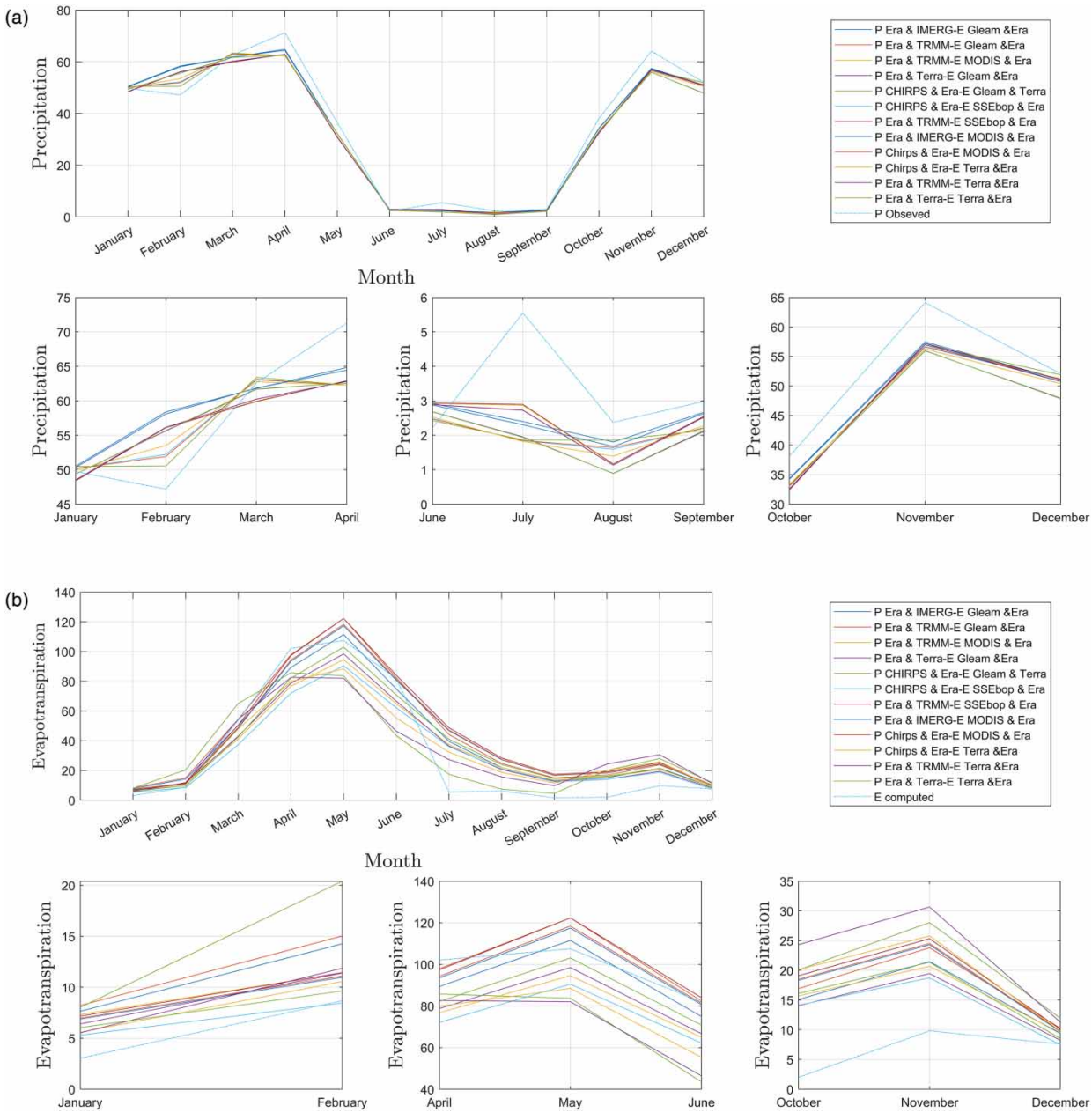

Figure 3 | Average monthly precipitation and ET values of the selected and reference model: (a) precipitation and (b) ET.

April have the largest differences between the combined observed and computed values. According to the figure, the months of February and July have the highest and lowest similarity between the amounts of products and observations, with a value of 19 mm difference related to IMERG precipitation and 0.01 mm difference related to TRMM precipitation.

Although the figures have a similar trend in the case of combined and calculated ET values in the reference model, different optimal combinations in different models have greater change ranges compared to the precipitation graphs. The highest amount of ET in May relates to the P ERA & TRMM-E Gleam & ERA model with a value of 121.84 mm, and the lowest value of 81.99 mm in this month relates to the P ERA & TRMM-E Terra & ERA model. Although the same procedure is followed, the ET peaked in the combination results this month. Notably, the depicted ET is achieved after applying the ET parameter's coefficient in the model.

It appears that based on the blending coefficients and their combinations, blended AET results have more diversity in their responses than blended precipitation results. This convergence in blended precipitation and variation in merged ET values represent the behavioral similarity and dissimilarity of precipitation and ET products. The values of the calibration and uncertainty evaluation metrics of the selected models (12 models) and the reference model are presented in Table 4.

Their KGE and NSE values are presented for two calibration and validation periods of the proposed scenarios (blended precipitation and ET products). Based on the table, the values of both metrics in the calibration and validation periods are well adjusted, and there is no significant difference between the models' performances in the simulation periods. The combinations of precipitation and ET values in all developed models have led to improved modeling results. Based on the results in Table 4, the P ERA and TRMM-E Gleam and ERA models have acceptable values in almost all metrics.

4.2. Results of the inverse models

The uncertainties of the model parameters and simulated streamflow are among the other aspects that can be considered for the evaluation of hydrological models using remotely sensed and large-scale datasets. Determining the aforementioned uncertainties is the aim of this part of the paper, which is the second phase of comprehensive evaluation and behavior assessment. In the first step, the reference model based on the information from the calibration period has been employed to calibrate its parameters considering the observed streamflow at the basin outlet. The reference model has also been used as a benchmark scenario to compare the performance of other models from the perspective of evaluating the uncertainty of the calibrated model's parameters.

As mentioned before, parametric uncertainty has been evaluated using the GLUE method. Table 5 presents the values of the uncertainty evaluation metrics of the reference models. Interpreting the indicators' mean values, a lower ARIL and a higher P_{level} result in a lower effect of uncertainty sources on the model's behavior. Therefore, a more stable model causes this fact. The NUE metric provides a general representation of uncertain simulation performance using a combination of P_{level} and ARIL metrics. The higher value of this metric causes more efficiency and reliability in uncertainty simulation.

As explained, the new precipitation and ET datasets will be generated based on a linear combination of remotely sensed products. Figure 4 shows the histograms of the blending ratios (alpha and beta) in the selected models to evaluate the uncertainty of the models' parameters. As shown in Figure 4(a), the combined weight of CHIRPS and ERA is almost the same regardless of the type of ET product. Intuitively, the median of the statistical distributions of the blending coefficients is around 0.4 (to 0.5). In the combination of ERA and TRMM precipitation products (in all investigated cases with ET products), a much larger share is specific to the ERA product and TRMM is combined with a smaller coefficient. This is in contrast to the nearly uniform distributions of CHIPS & ERA with SSEBop & ERA and ERA & TERRA ET with TERRA & ERA ET.

According to the figure, it is necessary to explain that the ERA precipitation product contributed to the combined precipitation with a more significant coefficient. Moreover, the behavior of CHIRPS and ERA in the first and second categories is similar. The second parameter used to combine the ET products is the beta parameter (Figure 4(b)). Three different patterns can be inferred from the various combinations of the products. In the first one, that is the combination of MODIS & ERA and SSEBop & ERA products (regardless of the used precipitation products), MODIS and SSEBop datasets had less contributions than ERA product which played a vital role in the dual combination of the products. The second pattern is related to the category of blended information that GLEAM contributed to it. In these combinations, despite the wholly defeated combination with more participation of ERA, the GLEAM product has equal or almost equal participation. In the third pattern, which is the combination of TERRA & ERA, the contribution of ERA is more than that of TERRA, but its amount is less than in the first pattern.

Table 4 | Statistical metrics to evaluate streamflow of the blending scenarios

Datasets	Metrics	Base model	P CHIRPS & ERA E MODIS & ERA	P CHIRPS & ERA-E Gleam & Terra	P CHIRPS & ERA-E SSEBop & ERA	P CHIRPS & ERA-E Terra & ERA	P ERA & IMERG E MODIS & ERA	P ERA & IMERG-E Gleam & ERA	P ERA & Terra-E Gleam & ERA	P ERA & Terra-E Terra & ERA	P ERA & TRMM E MODIS & ERA	P ERA & TRMM-E Gleam & ERA	P ERA & TRMM-E SSEBop & ERA	P ERA & TRMM-E Terra & ERA
Calibration	KGE	0.746	0.847	0.862	0.807	0.860	0.846	0.832	0.861	0.857	0.867	0.860	0.850	0.863
	NSE	0.730	0.796	0.800	0.745	0.810	0.773	0.776	0.803	0.803	0.824	0.794	0.788	0.796
Validation	KGE	0.765	0.794	0.785	0.769	0.792	0.819	0.810	0.770	0.774	0.774	0.780	0.792	0.777
	NSE	0.537	0.623	0.668	0.604	0.633	0.647	0.648	0.591	0.595	0.566	0.586	0.595	0.616

Table 5 | Metrics of uncertainty assessment using GLUE

Metrics	Base model	P CHIRPS & ERA E MODIS & ERA	P CHIRPS & ERA-E Gleam & Terra	P CHIRPS & ERA-E SSEBop & ERA	P CHIRPS & ERA-E Terra & ERA	P ERA & IMERG E MODIS & ERA	P ERA & IMERG-E Gleam & ERA	P ERA & Terra-E Gleam & ERA	P ERA & Terra-E Terra & ERA	P ERA & TRMM E MODIS & ERA	P ERA & TRMM-E Gleam & ERA	P ERA & TRMM-E SSEBop & ERA	P ERA & TRMM-E Terra & ERA
ARIL	2.61	2.12	2.73	3.46	2.86	3.56	3.82	2.36	2.77	2.70	2.76	2.59	3.04
P_{level}	27.62	25.71	30.48	39.05	32.38	38.10	40.00	29.52	32.38	31.43	32.38	31.43	32.38
NUE	10.57	12.12	11.17	11.30	11.32	10.71	10.48	12.51	11.67	11.66	11.75	12.12	10.66

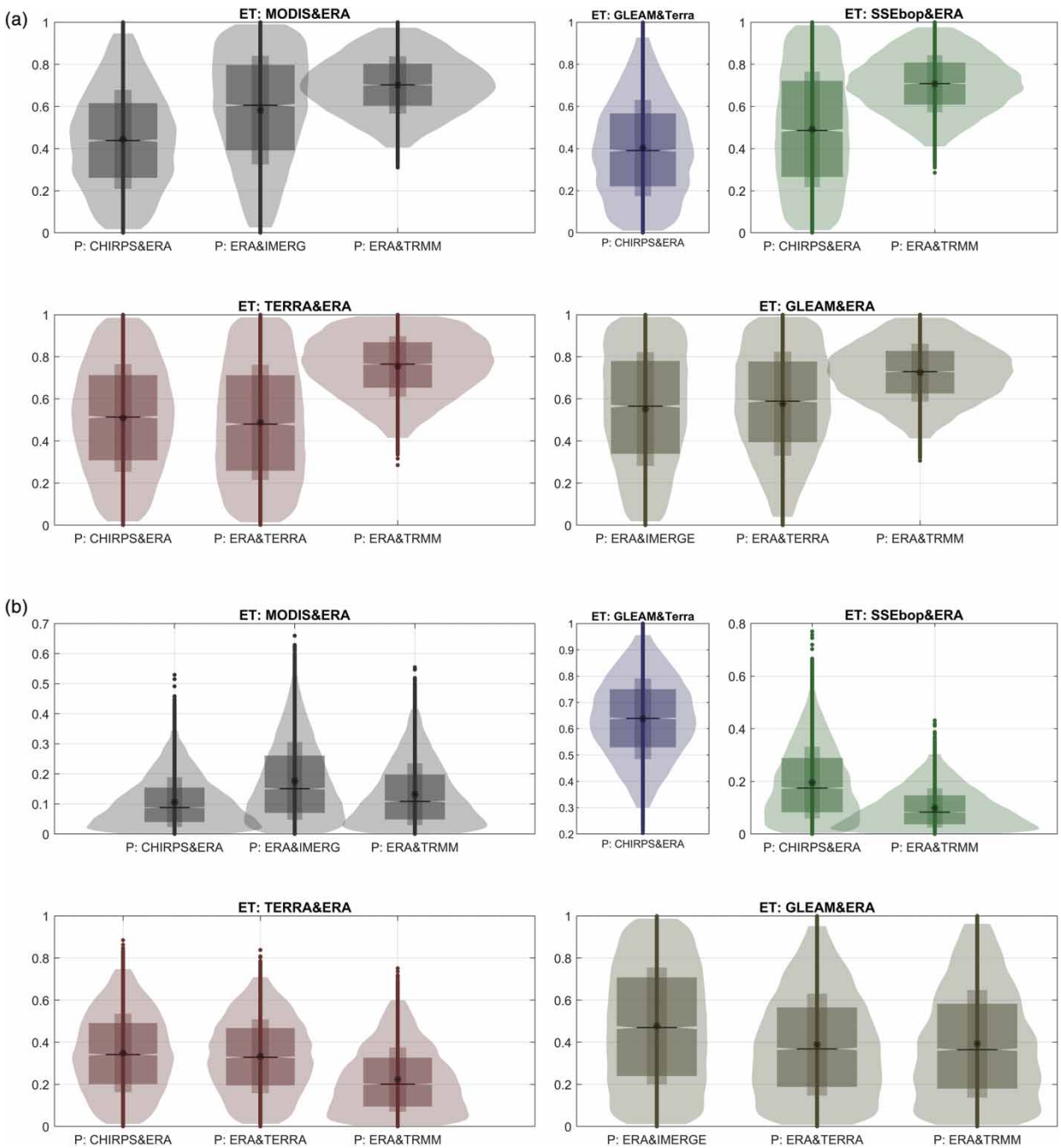

Figure 4 | Uncertainty for index (a) alpha and (b) beta in the selected models.

In Table 6, the uncertainty metrics resulting from the developed scenarios related to blended precipitation, ET, and streamflow have been reported. According to the table, the ET uncertainty metrics have the highest NUE and lowest ARIL values when compared to the P CHIRPS and ERA MODIS and ERA models, which are 3.88 and 9.01, respectively.

The highest P_{level} value of 45.83 was related to the P ERA & IMERG model as well. Regarding the precipitation variable, the maximum value of NUE and minimum value of ARIL related to P ERA & TRMM E MODIS & ERA and P ERA & TRMM E SSEBop & ERA models have been obtained at 28.85 and 0.76, respectively. Since, there is no reference model for these two variables; these metrics have only been compared with each other. Generally, ERA & TRMM and ERA & Terra precipitation models had less uncertainty.

Based on the statistics reported in Table 6, the ARIL values of streamflow uncertainties vary in the range of [2.12, 3.82] percent compared to the base model. P ERA & IMERG – E Gleam & ERA and P CHIRPS & ERA – E MODIS & ERA have the highest and lowest metric values, respectively. Furthermore, regarding the percentage of the observed values grouped with the upper and lower bounds of the simulated uncertainty (P_{level}), only the P CHIRPS & ERA-E MODIS & ERA combination is less than the reference model, and the rest models have greater values of the metric compared to the reference model. In the rest of the models, their NUE values are higher than the metrics of the reference model, except for one case of P ERA & IMERG E Gleam & ERA, which indicates the better performance of the models compared to the reference model. Table 6 shows the uncertainty of the simulated streamflow using blended P ERA & Terra – E GLEAM & ERA which is the most efficient model considering its NUE values (equal to 12.51).

On the other hand, the lowest efficiency (with an NUE value of 10.48) was related to the calibrated model using blended P ERA & IMERG E Gleam & ERA datasets. According to the contribution of ET products in these two models, the results show that the type of precipitation product combined with ERA has influenced the efficiency of the modeling uncertainty.

4.3. Blended precipitation and ET

Based on the current research goal (combination of ET and precipitation products using the inverse water balance model), five-blended precipitation, and ET datasets are depicted in Figures 5 and 6 to be compared with the same variables of the reference model. These combinations are P ERA & TRMM-E Terra & ERA, P ERA & Terra-E Terra & ERA, P ERA & IMERG E MODIS & ERA, P CHIRPS & ERA-E SSEBop & ERA, and P CHIRPS & ERA-E Gleam & Terra, respectively.

According to Table 6 and Figure 5, the developed blended precipitation datasets have significantly less uncertainty than the blended ET values. Comparing the results, the most important differences between the new precipitation datasets and ground information are their monthly peak values (in the months of February, March, and April). These scenarios are in good agreement in the middle and low monthly precipitation. In addition, the calculated uncertain peaks are randomly lower and higher than the precipitation amount of the reference model. Considering the values of the NUE reported in Table 6, uncertainty of the blended precipitation for the models P ERA & TRMM-E GLEAM & ERA and P ERA & TRMM-E SSEBop & ERA show the best performance and are innately the best inverse and uncertain simulated precipitation datasets. Based on the NUE metric, the PCHIRPS and ERA-E combination models produced the most disproportionate blended certain and uncertain precipitation dataset.

However, in the ET model, the simulated uncertain blended values are greater than the ET values of the reference model except in the model calibrated using P CHIRPS & ERA-E GLEAM & Terra input variables. Not only is the uncertain range of ET significant in the months with the most extensive range of changes (April, May, and June), but also in cold months, when ET has lower values, it is even remarkable in an uncertain way. Likewise, according to Figure 6, the combined uncertain ET in dry months shows a higher value than the reference model. Based on the NUE values presented in Table 6, the best uncertain blended dataset compared to the calculated ET values of the model would be the P CHIRPS & ERA – E MODIS & ERA, which is a scenario of product combinations.

The comparison of Table 6 and Figures 5 and 6 shows that due to the objective nature of precipitation and the conceptual nature of ET (without observed values), they have less uncertainty, and the correction resulting from their combination only includes cold months with high local amounts of precipitation. However, due to the non-objective and conceptual nature of ET, it has an estimated uncertain range in the combined model even in cold months, and naturally, this range will increase significantly in hot months. The addition of ET in the blended scenarios compared to the values of the reference model indicates the tendency of the model to remove moisture in cold months to simulate the streamflow behavior in these months with low values in contrast to the dry and wet months.

Table 6 | Statistical metrics to evaluate uncertainty responses of ET, uncertainty, and streamflow

Hydro-climatological variables	Statistical Metrics	Base model	P CHIRPS & ERA E MODIS & ERA	P CHIRPS & ERA-E Gleam & Terra	P CHIRPS & ERA-E SSEBop & ERA	P CHIRPS & ERA-E Terra & ERA	P ERA & IMERG E MODIS & ERA	P ERA & IMERG-E Gleam & ERA	P ERA & Terra-E Gleam & ERA	P ERA & Terra-E Terra & ERA	P ERA & TRMM E MODIS & ERA	P ERA & TRMM-E Gleam & ERA	P ERA & TRMM-E SSEBop & ERA	P ERA & TRMM-E Terra & ERA
E	ARIL		9.010	12.134	12.779	16.689	13.961	13.521	10.551	15.180	9.747	11.074	7.975	14.436
	P_{level}		35.000	20.833	40.833	33.333	45.833	40.000	39.167	33.333	35.833	40.833	30.833	34.167
	NUE		3.884	1.717	3.195	1.997	3.283	2.958	3.712	2.196	3.676	3.687	3.866	2.367
P	ARIL		1.604	1.551	1.735	1.696	1.545	1.510	0.911	0.956	0.764	0.775	0.764	0.786
	P_{level}		23.622	22.835	24.409	23.622	27.559	26.772	22.835	24.409	22.047	22.047	22.047	21.260
	NUE		14.728	14.723	14.067	13.932	17.842	17.734	25.065	25.530	28.846	28.437	28.846	27.055
Q	ARIL	2.61	2.12	2.73	3.46	2.86	3.56	3.82	2.36	2.78	2.70	2.76	2.59	3.04
	P_{level}	27.62	25.71	30.48	39.05	32.38	38.10	40.00	29.52	32.38	31.43	32.38	31.43	32.38
	NUE	10.57	12.12	11.17	11.30	11.33	10.71	10.48	12.51	11.67	11.66	11.75	12.12	10.66

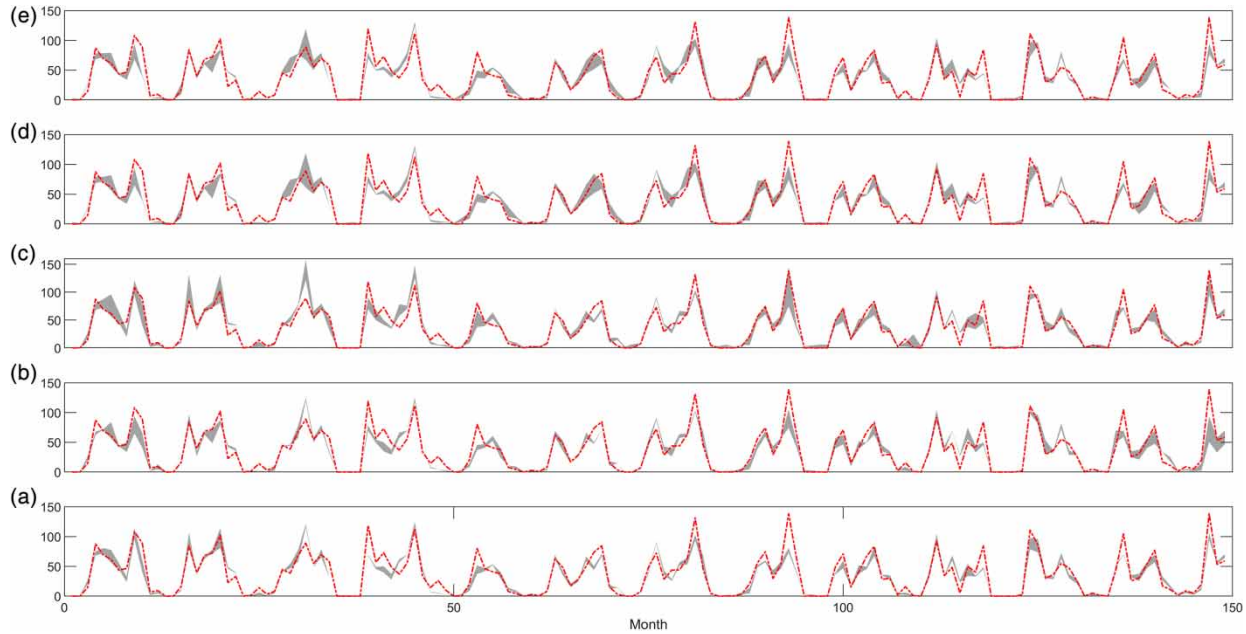

Figure 5 | Monthly observed precipitation of the benchmark model (– – –) versus blended scenarios: (a) P ERA & TRMM-E Terra & ERA, (b) P ERA & Terra-E Terra & ERA, (c) P ERA & IMERG E MODIS & ERA, (d) P CHIRPS & ERA-E SSEBop & ERA, and (e) P CHIRPS & ERA-E Gleam & Terra.

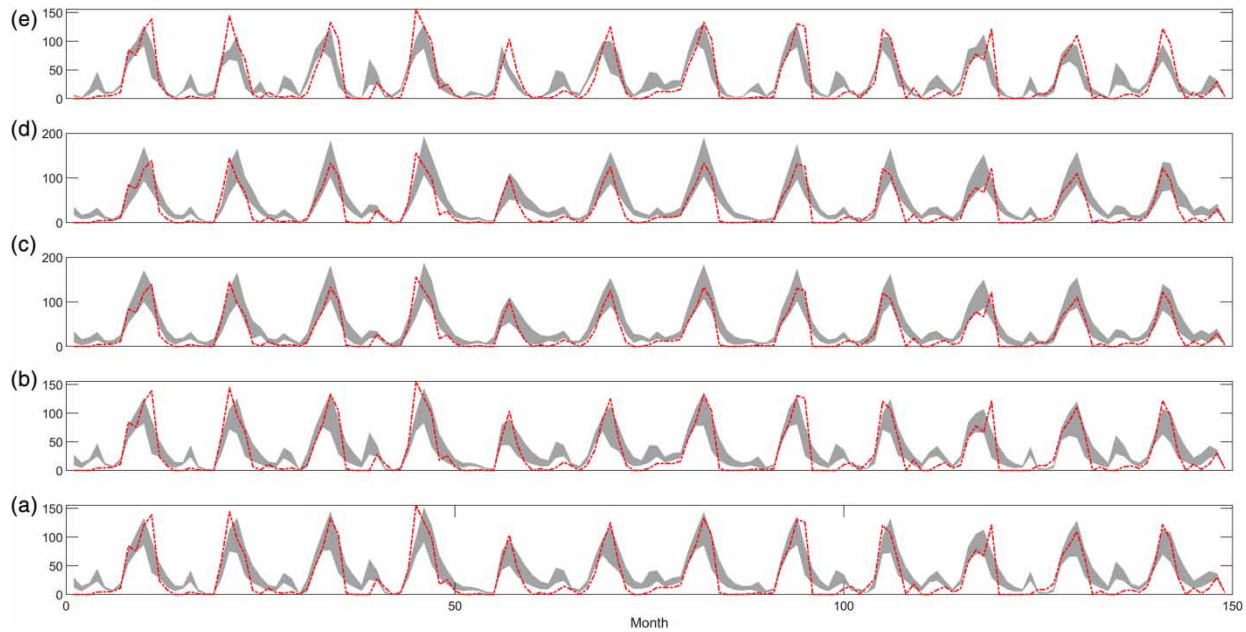

Figure 6 | Computed ET of the benchmark model (– – –) versus blended scenarios: (a) P ERA & TRMM-E Terra & ERA, (b) P ERA & Terra-E Terra & ERA, (c) P ERA & IMERG E MODIS & ERA, (d) P CHIRPS & ERA-E SSEBop & ERA, and (e) P CHIRPS & ERA-E Gleam & Terra.

5. CONCLUDING REMARKS

The current study's innovation is the creation of new precipitation and ET datasets based on linear combinations to be available in two certain and uncertain types for many large-scale products at the basin level.

The proposed blending method has been used to evaluate the conceptual structure of the water balance model and the inverse modeling procedure. In addition to evaluating the effectiveness of hydrologic modeling using the blended values

of large-scale precipitation (such as CHIRPS, ERA, TRMM, Terra, and IMERG) and ET products (such as GLEAM, MODIS, SSEBop, Terra, and ERA), a new criterion is presented in this article.

This criterion was introduced to measure the suitability of this large-scale information not with ground information but in a complex process such as the hydrological model. Although some products have been combined and modified in some research projects in the form of Terrestrial Water Storage (TWS) macro-water balance modeling (Schoups & Nasseri 2021), doing this process in the form of a model with many details and on a small spatial scale is the novel focus point of the present research.

In this procedure, the large-scale information is dually and linearly (uncertainly and certainly) blended to reach the best precipitation and ET signals to generate streamflow with an inverse approach. The results show the conformity of the pattern and behavior of the blended precipitation obtained using the MLS method in the study area. This type of blending precipitation led to slight differences in the cold months of February, March, April, and (roughly) May. Due to the mountainous nature of the study area, these differences can be caused by the insufficient expansion of the precipitation-monitoring network, especially in the highlands. This proposed method provides a clear suggestion for investigating and analyzing precipitation patterns in areas without rainfall recording stations or with variable topography, where the uncertainty of the rainfall amount on the ground is notable.

This combination process was carried out for the large-scale ET product data (in the form of two combination scenarios with precipitation). To evaluate the changes caused by this combination, the results of the corresponding scenarios were compared with the amount of ET calculated in the calibrated reference model, which was created using recorded environmental information (precipitation, temperature, pan evaporation, and streamflow). In addition to following the general procedure of ET in the base scenario, the combined values have mainly high uncertainty, and compared to the used reference balance model, they bring more ET in cold months. Using this method for measuring the optimal pattern of large-scale precipitation and ET products' combination results is a new improvement in the production of calibrated information of hydrological models at the basin scale. It also provides a controlled approximation in areas without input information or an information gap. This method can use large-scale snow-water equivalent products in its evaluation format, and by entering the temperature combination, snow-water equivalent supplies the main components of hydrological modeling at the watershed scale.

Examining these components as a leading research step plays a role in establishing a process model like a water balance model in areas without stations or calibrating large-scale information from a process path. The use of data mining methods to gradually verify the ratio of information combinations, especially in online hydrology modeling, can be considered the next step in modifying and calibrating large-scale information with a process purpose.

DATA AVAILABILITY STATEMENT

All relevant data are available from an online repository or repositories http://dx.doi.org/10.5281/zenodo.8062625.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. & Hegewisch, K. C. 2018 Terraclimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958 to 2015. *Scientific Data* **5**, 1–12.

Ahmadi, A. & Nasseri, M. 2020 Do direct and inverse uncertainty assessment methods present the same results? *Journal of Hydroinformatics* **22**, 842–855.

Aires, F. 2014 Combining datasets of satellite-retrieved products. Part I: methodology and water budget closure. *Journal of Hydrometeorology* **15**, 1677–1691.

Allen, R. G., Tasumi, M., Morse, A., Trezza, R., Wright, J. L., Bastiaanssen, W., Kramber, W., Lorite, I. & Robison, C. W. 2007 Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)–Applications. *Journal of Irrigation and Drainage Engineering* **133**, 395–406.

Amini, Y. & Nasseri, M. 2021 Improving spatial estimation of hydrologic attributes via optimized moving search strategies. *Arabian Journal of Geosciences* **14**, 723. doi: 10.1007/s12517-021-06961-3.

Baez-Villanueva, O. M., Zambrano-Bigiarini, M., Beck, H. E., Mcnamara, I., Ribbe, L., Nauditt, A., Birkel, C., Verbist, K., Giraldo-Osorio, J. D. & Thinh, N. X. 2020 RF-MEP: a novel random forest method for merging gridded precipitation products and ground-based measurements. *Remote Sensing of Environment* **239**, 111606.

Baik, J., Liaqat, U. W. & Choi, M. 2018 Assessment of satellite-and reanalysis-based evapotranspiration products with two blending approaches over the complex landscapes and climates of Australia. *Agricultural and Forest Meteorology* **263**, 388–398.

Bastiaanssen, W. G., Menenti, M., Feddes, R. & Holtslag, A. 1998 A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *Journal of Hydrology* **212**, 198–212.

Beven, K. & Binley, A. 1992 The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes* **6**, 279–298.

Chen, M., Senay, G. B., Singh, R. K. & Verdin, J. P. 2016 Uncertainty analysis of the operational simplified surface energy balance (SSEBop) model at multiple flux tower sites. *Journal of Hydrology* **536**, 384–399.

Dembele, M., Ceperley, N., Zwart, S. J., Salvadore, E., Mariethoz, G. & Schaefli, B. 2020 Potential of satellite and reanalysis evaporation datasets for hydrological modelling under various model calibration strategies. *Advances in Water Resources* **143**, 103667.

Funk, C. C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., Romero, B. E., Husak, G. J., Michaelsen, J. C. & Verdin, A. P. 2014 A quasi-global precipitation time series for drought monitoring. *US Geological Survey Data Series* **832**, 1–12.

Ghomlaghi, A., Nasseri, M. & Bayat, B. 2022 Comparing and contrasting the performance of high-resolution precipitation products via error decomposition and triple collocation: an application to different climate classes of the central Iran. *Journal of Hydrology* **612**. doi:10.1016/j.jhydrol.2022.128298.

Guo, S., Chen, H., Zhang, H., Xiong, L., Liu, P., Pang, B., Wang, G. & Wang, Y. 2005 A semi-distributed monthly water balance model and its application in a climate change impact study in the middle and lower Yellow River basin. *Water International* **30**, 250–260.

Guo, X., Wu, Z., He, H. & Xu, Z. 2022 Evaluating the potential of different evapotranspiration datasets for distributed hydrological model calibration. *Remote Sensing* **14**, 629.

Herrnegger, M., Nachtnebel, H. P. & Schulz, K. 2015 From runoff to rainfall: inverse rainfall–runoff modelling in a high temporal resolution. *Hydrology and Earth System Sciences* **19**, 4619–4639. https://doi.org/10.5194/hess-19-4619-2015.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D. & Thépaut, J.-N. 2019 ERA5 monthly averaged data on pressure levels from 1959 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. doi:10.24381/cds.6860a573.

Huang, Q., Qin, G., Zhang, Y., Tang, Q., Liu, C., Xia, J., Chiew, F. H. & Post, D. 2020 Using remote sensing data-based hydrological model calibrations for predicting runoff in ungauged or poorly gauged catchments. *Water Resources Research* **56** (8), e2020WR028205.

Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., Mcnab, A., Rudolf, B. & Schneider, U. 1997 The global precipitation climatology project (GPCP) combined precipitation dataset. *Bulletin of the American Meteorological Society* **78**, 5–20.

Huffman, G. J., Adler, R. F., Bolvin, D. T. & Nelkin, E. J. 2010 The TRMM Multi-Satellite Precipitation Analysis (TMPA) (Chapter 1). In: *Satellite Rainfall Applications for Surface Hydrology* (F. Hossain & M. Gebremichael, eds.), Springer Verlag, Berlin, 3–22. https://doi.org/10.1007/978-90-481-2915-7_1.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Stocker, E. F. & Tan, J. 2019 *V06 IMERG Release Notes*. NASA/GSFC, Greenbelt, MD, USA.

Javadian, M., Kordi, F. & Tajrishy, M. 2019 Evaluation and comparison of estimation methods for actual evapotranspiration in the Urmia Lake Basin. *Iranian Journal of Ecohydrology* **6** (1), 125–136.

Jin, X., Xu, C.-Y., Zhang, Q. & Singh, V. 2010 Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *Journal of Hydrology* **383**, 147–155.

Jongjin, B., Jongmin, P., Dongryeol, R. & Minha, C. 2016 Geospatial blending to improve spatial mapping of precipitation with high spatial resolution by merging satellite-based and ground-based data. *Hydrological Processes* **30**, 2789–2803.

Kimani, M. W., Hoedjes, J. C. & Su, Z. 2018 Bayesian bias correction of satellite rainfall estimates for climate studies. *Remote Sensing* **10**, 1074.

Kirchner, J. W. 2009 Catchments as simple dynamical systems: catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research* **45**, W02429. doi:10.1029/2008WR006912.

Kouchi, D. H., Esmaili, K., Faridhosseini, A., Sanaeinejad, S. H., Khalili, D. & Abbaspour, K. C. 2017 Sensitivity of calibrated parameters and water resource estimates on different objective functions and optimization algorithms. *Water* **9**, 384.

Lancaster, P. & Salkauskas, K. 1981 Surfaces generated by moving least squares methods. *Mathematics of Computation* **37**, 141–158.

Li, M. & Shao, Q. 2010 An improved statistical approach to merge satellite rainfall estimates and rain gauge data. *Journal of Hydrology* **385**, 51–64.

Li, L., Xia, J., Xu, C.-Y. & Singh, V. 2010 Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models. *Journal of Hydrology* **390**, 210–221.

Lu, J., Wang, G., Chen, T., Li, S., Hagan, D. F. T., Kattel, G., Peng, J., Jiang, T. & Su, B. 2021 A harmonized global land evaporation dataset from model-based products covering 1980–2017. *Earth System Science Data* **13**, 5879–5898.

Ma, Y., Sun, X., Chen, H., Hong, Y. & Zhang, Y. 2021 A two-stage blending approach for merging multiple satellite precipitation estimates and rain gauge observations: an experiment in the northeastern Tibetan plateau. *Hydrology and Earth System Sciences* **25**, 359–374.

Manz, B., Buytaert, W., Zulkafli, Z., Lavado, W., Willems, B., Robles, L. A. & Rodríguez-Sánchez, J. P. 2016 High-resolution satellite-gauge merged precipitation climatologies of the Tropical Andes. *Journal of Geophysical Research: Atmospheres* **121**, 1190–1207.

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A. & Verhoest, N. E. 2017 GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development* **10**, 1903–1925.

Moreira, A. A., Ruhoff, A. L., Roberti, D. R., De Arruda Souza, V., Da Rocha, H. R. & De Paiva, R. C. D. 2019 Assessment of terrestrial water balance using remote sensing data in South America. *Journal of Hydrology* **575**, 131–147.

Moshir Panahi, D., Sadeghi Tabas, S., Kalantari, Z., Ferreira, C. S. S. & Zahabiyoun, B. 2021 Spatio-temporal assessment of global gridded evapotranspiration datasets across Iran. *Remote Sensing* **13** (9), 1816.

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

Mousavi, R., Nasseri, M., Abbasi, S., Taheri, M. & Shamsi Anboohi, M. 2022 Global gridded products efficiency in closing water balance models: various modeling scenarios for behavioral assessments. *Acta-Geophisica*. doi:10.1007/s11600-022-01004-1.

Munier, S., Aires, F., Schlaffer, S., Prigent, C., Papa, F., Maisongrande, P. & Pan, M. 2014 Combining data sets of satellite-retrieved products for basin-scale water balance study: 2. Evaluation on the Mississippi basin and closure correction model. *Journal of Geophysical Research: Atmospheres* **119**, 12,100–12,116.

Nasseri, M., Schoups, G. & Taheri, M. 2022 A spatiotemporal framework to calibrate high-resolution global monthly precipitation products: an application to the Urmia lake watershed in Iran. *International Journal of Climatology* **42** (4), 2169–2194.

Nasseri, M., Zahraie, B., Ansari, A. & Solomatine, D. P. 2013 Uncertainty assessment of monthly water balance models based on Incremental Modified Fuzzy Extension Principle method. *Journal of Hydroinformatics* **15** (4), 1340–1360. doi: https://doi.org/10.2166/hydro.2013.159.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I–A discussion of principles. *Journal of hydrology* **10**, 282–290.

Nguyen, G. V., Le, X.-H., Van, L. N., Jung, S., Yeon, M. & Lee, G. 2021 Application of random forest algorithm for merging multiple satellite precipitation products across South Korea. *Remote Sensing* **13**, 4033.

Pan, M., Sahoo, A. K., Troy, T. J., Vinukollu, R. K., Sheffield, J. & Wood, E. F. 2012 Multisource estimation of long-term terrestrial water budget for major global river basins. *Journal of Climate* **25**, 3191–3206.

Rozante, J. R., Moreira, D. S., De Goncalves, L. G. G. & Vila, D. A. 2010 Combining TRMM and surface observations of precipitation: technique and validation over South America. *Weather and Forecasting* **25**, 885–894.

Running, S. W., Mu, Q., Zhao, M. & Moreno, A. 2017 *MODIS Global Terrestrial Evapotranspiration (ET) Product (NASA MOD16A2/A3) NASA Earth Observing System MODIS Land Algorithm*. NASA, Washington, DC, USA.

Sadeghi, S. H. R., Gholami, L. & Khaledi Darvishan, A. V. 2013 Suitability of MUSLT for storm sediment yield prediction in Chehelgazi watershed, Iran. *Hydrological Sciences Journal* **58** (4), 892–897. doi: 10.1080/02626667.2013.782406.

Sahoo, A. K., Pan, M., Troy, T. J., Vinukollu, R. K., Sheffield, J. & Wood, E. F. 2011 Reconciling the global terrestrial water budget using satellite remote sensing. *Remote Sensing of Environment* **115**, 1850–1865.

Schoups, G. & Nasseri, M. 2021 GRACEfully closing the water balance: A data-driven probabilistic approach applied to river basins in Iran. *Water Resources Research*, **57**, e2020WR029071. https://doi.org/10.1029/2020WR029071

Senay, G. B., Kagone, S. & Velpuri, N. M. 2020 Operational global actual evapotranspiration: development, evaluation, and dissemination. *Sensors* **20**, 1915.

Shayeghi, A., Azizian, A. & Brocca, L. 2020 Reliability of reanalysis and remotely sensed precipitation products for hydrological simulation over the Sefidrood River Basin, Iran. *Hydrological Sciences Journal* **65**, 296–310.

Taheri, M., Shamsi Anboohi, M., Mousavi, R. & Nasseri, M. 2022 Hybrid global gridded snow products and conceptual simulations of distributed snow budget: evaluation of different scenarios in a mountainous watershed. *Frontiers of Earth Science*. doi: 10.1007/s11707-022-1005-2.

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M. & Robinson, B. A. 2008 Treatment of input uncertainty in hydrologic modeling: doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research* **44**, W00B09. doi:10.1029/2007WR006720.

Wang, Q. J., Pagano, T. C., Zhou, S. L., Hapuarachchi, H. A. P., Zhang, L. & Robertson, D. E. 2011 Monthly versus daily water balance models in simulating monthly runoff. *Journal of Hydrology* **404** (3–4), 166–175.

Wang, G., Zhang, J., Xuan, Y., Liu, J., Jin, J., Bao, Z., He, R., Liu, C., Liu, Y. & Yan, X. 2013 Simulating the impact of climate change on runoff in a typical river catchment of the Loess Plateau, China. *Journal of Hydrometeorology* **14**, 1553–1561.

Xie, P. & Arkin, P. A. 1997 Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society* **78**, 2539–2558.

Xie, P. & Xiong, A. Y. 2011 A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. *J. Geophys. Res.* **116**, D21106, doi:10.1029/2011JD016118.

Yang, Z., Hsu, K., Sorooshian, S., Xu, X., Braithwaite, D., Zhang, Y. & Verbist, K. M. 2017 Merging high-resolution satellite-based precipitation fields and point-scale rain gauge measurements – a case study in Chile. *Journal of Geophysical Research: Atmospheres* **122**, 5267–5284.

Yumnam, K., Guntu, R. K., Rathinasamy, M. & Agarwal, A. 2022 Quantile-based Bayesian model averaging approach towards merging of precipitation products. *Journal of Hydrology* **604**, 127206.

Zandi, O., Zahraie, B., Nasseri, M. & Behrangi, A. 2022 Stacking machine learning models versus a locally weighted linear model to generate high-resolution monthly precipitation over a topographically complex area. *Atmospheric Research* **272**, 623–641. doi:10.1016/j.atmosres.2022.106159.

First received 6 September 2022; accepted in revised form 31 May 2023. Available online 15 June 2023