



POTSDAM-INSTITUT FÜR
KLIMAFOLGENFORSCHUNG

Originally published as:

[Fernandez Palomino, C. A.](#), [Hattermann, F. F.](#), [Krysanova, V.](#), [Lobanova, A.](#), Vega-Jácome, F., Lavadao, W., Santini, W., Aybar, C., Bronstert, A. (2022): A Novel High-Resolution Gridded Precipitation Dataset for Peruvian and Ecuadorian Watersheds: Development and Hydrological Evaluation. - Journal of Hydrometeorology, 23, 3, 309-336.

DOI: <https://doi.org/10.1175/JHM-D-20-0285.1>

A Novel High-Resolution Gridded Precipitation Dataset for Peruvian and Ecuadorian Watersheds: Development and Hydrological Evaluation

CARLOS ANTONIO FERNANDEZ-PALOMINO,^{a,b} FRED F. HATTERMANN,^a VALENTINA KRYSANOVA,^a ANASTASIA LOBANOVA,^a FIORELLA VEGA-JÁCOME,^c WALDO LAVADO,^c WILLIAM SANTINI,^d CESAR AYBAR,^e AND AXEL BRONSTERT^b

^a *Research Department II—Climate Resilience, Potsdam Institute for Climate Impact Research, Potsdam, Germany*

^b *Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany*

^c *Hidrología—Estudios e Investigaciones Hidrológicas, Servicio Nacional de Meteorología e Hidrología del Perú, Lima, Peru*

^d *Institut de Recherche pour le Développement, Laboratoire GET (IRD, CNRS, UPS, CNES), Toulouse, France*

^e *Department of Geoinformatics—Z_GIS, University of Salzburg, Salzburg, Austria*

(Manuscript received 2 December 2020, in final form 20 November 2021)


ABSTRACT: A novel approach for estimating precipitation patterns is developed here and applied to generate a new hydrologically corrected daily precipitation dataset, called RAIN4PE (Rain for Peru and Ecuador), at 0.1° spatial resolution for the period 1981–2015 covering Peru and Ecuador. It is based on the application of 1) the random forest method to merge multi-source precipitation estimates (gauge, satellite, and reanalysis) with terrain elevation, and 2) observed and modeled streamflow data to first detect biases and second further adjust gridded precipitation by inversely applying the simulated results of the ecohydrological model SWAT (Soil and Water Assessment Tool). Hydrological results using RAIN4PE as input for the Peruvian and Ecuadorian catchments were compared against the ones when feeding other uncorrected (CHIRP and ERA5) and gauge-corrected (CHIRPS, MSWEP, and PISCO) precipitation datasets into the model. For that, SWAT was calibrated and validated at 72 river sections for each dataset using a range of performance metrics, including hydrograph goodness of fit and flow duration curve signatures. Results showed that gauge-corrected precipitation datasets outperformed uncorrected ones for streamflow simulation. However, CHIRPS, MSWEP, and PISCO showed limitations for streamflow simulation in several catchments draining into the Pacific Ocean and the Amazon River. RAIN4PE provided the best overall performance for streamflow simulation, including flow variability (low, high, and peak flows) and water budget closure. The overall good performance of RAIN4PE as input for hydrological modeling provides a valuable criterion of its applicability for robust countrywide hydrometeorological applications, including hydroclimatic extremes such as droughts and floods.


SIGNIFICANCE STATEMENT: We developed a novel precipitation dataset RAIN4PE for Peru and Ecuador by merging multisource precipitation data (satellite, reanalysis, and ground-based precipitation) with terrain elevation using the random forest method. Furthermore, RAIN4PE was hydrologically corrected using streamflow data in watersheds with precipitation underestimation through reverse hydrology. The results of a comprehensive hydrological evaluation showed that RAIN4PE outperformed state-of-the-art precipitation datasets such as CHIRP, ERA5, CHIRPS, MSWEP, and PISCO in terms of daily and monthly streamflow simulations, including extremely low and high flows in almost all Peruvian and Ecuadorian catchments. This underlines the suitability of RAIN4PE for hydrometeorological applications in this region. Furthermore, our approach for the generation of RAIN4PE can be used in other data-scarce regions.

KEYWORDS: Amazon region; Complex terrain; South America; Streamflow; Precipitation; Hydrology; Water budget/balance; Inverse methods; Mountain meteorology; Machine learning

1. Introduction

Precipitation is an essential component of the water cycle and reliable and accurate information about its spatiotemporal distribution is decisive for a multitude of scientific studies

 Denotes content that is immediately available upon publication as open access.

 Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-20-0285.s1>.

Corresponding author: Carlos Antonio Fernandez-Palomino, cafpxl@gmail.com

and operational applications. Rain gauge observations are the most used and—on a local scale—direct and accurate precipitation data sources. In addition, precipitation data can be derived from other sources such as rainfall radar stations, satellites, reanalysis products, or based on merging procedures (Sun et al. 2018). In many developing countries like Peru and Ecuador, rain gauges are unevenly and sparsely distributed (Scheel et al. 2011; Ochoa et al. 2014; Manz et al. 2016; Hunziker et al. 2017; Aybar et al. 2020). These features limit the precise estimation of spatial and temporal variability of precipitation using only gauge-based measurements in the tropical Andes.

Precipitation information derived from satellite data, e.g., CMORPH (Joyce et al. 2004), TMPA (Huffman et al. 2007), CHIRP (Funk et al. 2015a), and IMERG (Huffman et al.

DOI: 10.1175/JHM-D-20-0285.1

© 2022 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

2019), with a high spatiotemporal resolution, near-global coverage, and near-real-time availability have been produced in the last decades (see appendix A for abbreviations). These satellite-based precipitation products are promising alternative sources for regions with sparse observations. However, previous studies for the Andes domain (Scheel et al. 2011; Kneis et al. 2014; Mantas et al. 2014; Ochoa et al. 2014; Zulkafli et al. 2014; Satgé et al. 2016; Chavez and Takahashi 2017; Manz et al. 2017; Baez-Villanueva et al. 2018; Erazo et al. 2018) have reported that precipitation estimates from satellites can be erroneous or biased, and that ground-based data are often needed to reduce their bias. Furthermore, the current short length of satellite records in this region constitutes an important restriction for the use of most of these products for long-term applications.

Reanalysis precipitation data, such as CFSR (Saha et al. 2010), JRA-55 (Kobayashi et al. 2015), and MERRA (Reichle et al. 2017), rely on uncertain parameterizations, and their spatial resolution is too coarse to represent orographic precipitation (Beck et al. 2020b). Recently, the state-of-the-art climate reanalysis ERA5 (Hersbach et al. 2020) was released, which has been shown to outperform previous reanalyses for precipitation estimation (Beck et al. 2019a; Tall et al. 2019; Xu et al. 2019a; Fallah et al. 2020; Gleixner et al. 2020), and has shown acceptable performance for hydrological modeling over North America (Tarek et al. 2020), the Amazon River basin (Towner et al. 2019), and at the global scale (Alfieri et al. 2020).

In recent years, global merged precipitation products that incorporate satellite and reanalysis information with gauge-based datasets such as CHIRPS (Funk et al. 2015a) and MSWEP (Beck et al. 2017, 2019b) have been published and are available. Many studies worldwide have shown that these products have higher accuracy than precipitation estimates based on one source only (e.g., either satellite- or reanalysis-based precipitation products) and have significant potential for hydro-meteorological studies (Bai and Liu 2018; Beck et al. 2019a; Wu et al. 2019; Xu et al. 2019b). CHIRPS has been used successfully to understand the precipitation variability over the Andes (Segura et al. 2019) and Amazonia (Paccini et al. 2018; Espinoza et al. 2019; da Motta Paca et al. 2020). In South America, the accuracy of merged precipitation products has been tested only in a few studies using ground-based precipitation (Zambrano-Bigiarini et al. 2017; Baez-Villanueva et al. 2018; Satgé et al. 2019) and hydrological modeling (Wongchuig Correa et al. 2017; Satgé et al. 2019). CHIRPS and MSWEP showed good performance for streamflow simulation in the Amazon River basin (Wongchuig Correa et al. 2017) and in catchments draining into Titicaca Lake (Satgé et al. 2019, 2020). To the best of our knowledge, there are no case studies in the literature on hydrological evaluation of CHIRPS and MSWEP in Peruvian and Ecuadorian watersheds, which is addressed in this study.

At the regional scale, recently a high-resolution (0.1°) daily gridded precipitation dataset for Peru as part of PISCO datasets was developed (PISCO hereafter) (Aybar et al. 2020). PISCO is based on the merging of satellite estimates (CHIRP) and ground-based observations. It is used by SENAMHI for operational purposes in Peru for droughts and floods monitoring at the national scale, and was applied for

hydrological modeling of the Andean Vilcanota River catchment (Fernandez-Palomino et al. 2020), catchments draining into the Pacific Ocean (Asurza-Véliz and Lavado-Casimiro 2020), and Peruvian catchments (Llauca et al. 2021). As the method used to generate PISCO mainly corrects the biases of CHIRP using in situ precipitation data, the higher accuracy of precipitation estimates is constrained to gauged regions such as the Pacific coast and the eastern and western slopes of the Andes of Peru (Aybar et al. 2020; Llauca et al. 2021). Hence, the application of PISCO for Peruvian Amazon and trans-boundary river catchments is limited. This motivated us to generate a new rainfall dataset for hydro-meteorological applications at the national scale of Peru and Ecuador, exploiting the lessons learned from precipitation estimates derived not only from gauges and satellites but also from the state-of-the-art reanalysis ERA5. Indeed, ERA5 and CHIRP, which has long-term daily precipitation data available (1981–present) and hence appropriate for long-term hydrological applications, were used for the precipitation merging procedure in this study. Moreover, terrain elevation which was reported to be a key physical variable with a strong influence on precipitation patterns in mountainous regions (Chavez and Takahashi 2017; Bhuiyan et al. 2019; Beck et al. 2020b) was considered as an additional predictor variable.

Besides sparseness and uncertainty of rainfall observations in complex tropical mountain ranges, in some of those regions depositing fog and clouds may contribute significantly to precipitation, but cannot be recorded with conventional measurements. In páramos (grassland ecosystems extending from northern Peru to Venezuela and occurring between the tree line and glaciers) and tropical montane cloud forest (TMCF) such precipitation plays a key role in the water cycle as the cloud/fog interception by the páramos/forest constitutes an important water source to the system (Gomez-Peralta et al. 2008; Bruijnzeel et al. 2011; Clark et al. 2014; Cárdenas et al. 2017; Strauch et al. 2017). Modeled contributions of cloud water varying from less than 5% of total precipitation in wet areas to more than 75% in low-rainfall areas in TMCF were reported by Bruijnzeel et al. (2011). Fog water contribution of up to 30% of bulk precipitation (rainfall plus fog water) was estimated in tropical montane forests in the eastern Andes of Central Peru using fog gauges (Gomez-Peralta et al. 2008). Cloud water contribution of up to 15% of streamflow was reported for the montane Kosñyapata catchment in the eastern Peruvian Andes using an isotopic mixing model (Clark et al. 2014). Fog water contribution of up to 28% of the total precipitation to páramos in the Colombian Andes was measured using fog gauges (Cárdenas et al. 2017). To correct the underestimation of precipitation by gridded precipitation products, adjustment of precipitation data for regions covered by cloud forests has been proposed (Strauch et al. 2017) with reported increases of up to 50% of the precipitation values in the WFDEI dataset (Weedon et al. 2014) required to improve streamflow simulation in the tropical montane watersheds.

However, the cloud/fog water component is not represented in the aforementioned precipitation data sources. This lacuna, together with the dearth of precipitation gauges could explain some of the poor hydrologic model performances and

problems with water budget closure reported in previous studies in páramo and/or montane catchments draining into the Amazon River (Zulkafli et al. 2014; Zubieta et al. 2015; Manz et al. 2016; Zubieta et al. 2017; Strauch et al. 2017; Aybar et al. 2020). Thus, for reliable and accurate estimation of precipitation in regions such as the TMCF and páramos, it is important to consider the contribution of cloud/fog water to the terrestrial hydrological system.

Correcting potential errors in gridded precipitation datasets for these areas requires application of other types of observations and estimates. Corrected estimates of precipitation using satellite soil moisture products have been derived in recent years (Brocca et al. 2013; Román-Cascón et al. 2017; Brocca et al. 2019). However, the utility of these products could be limited due to their low accuracy in regions with dense forests (Brocca et al. 2020), such as TMCF and rain forest areas. Streamflow observations, which are spatially integrative and could be another source of data supplementing information from sparse rain gauges, offer an additional method to infer precipitation patterns and evaluate precipitation datasets (Le Moine et al. 2015; Henn et al. 2018). In this study, we applied regional streamflow observations inversely to infer or correct the precipitation input for the corresponding regional hydrological simulations. This approach has been termed “hydrology backwards” or “reverse hydrology” by Kirchner (2009) and has so far been applied in mountainous catchments like Rietholzbach in Switzerland (Teuling et al. 2010), Alzette in Luxembourg (Krier et al. 2012), Schlieffau and Krems in Austria (Herrnegger et al. 2015), and the Sierra Nevada mountain range of California (Henn et al. 2015, 2018). These studies used a simple lumped hydrological model to do reverse hydrology. In our case, we applied a process-based hydrological model to correct precipitation biases using streamflow data. We hypothesize that correction of precipitation using streamflow data can improve closing the observed water budget gap over complex tropical mountainous catchments such as páramo and montane watersheds.

This study is the first attempt to generate a precipitation dataset for Peru and Ecuador by merging different sources of precipitation and correcting precipitation estimates through reverse hydrology. Furthermore, we evaluate the applicability of the precipitation dataset generated in this study, uncorrected precipitation datasets used for merging procedure (CHIRP and ERA5), and current state-of-the-art local (PISCO) and global (CHIRPS and MSWEP) merged precipitation products for hydrological modeling of Peruvian and Ecuadorian river catchments. This will demonstrate the effectiveness of the new methods combined here, and will help illustrate the appropriateness of multiple precipitation datasets for the countrywide hydrometeorological applications both in Peru and Ecuador. The objectives of this study are 1) to generate a high-spatial-resolution and hydrologically adjusted precipitation dataset for Peru and Ecuador, and 2) to assess and compare the applicability of this precipitation data and the current state-of-the-art uncorrected and merged precipitation products for hydrological modeling.

2. Study area and data

a. Study area

The study area covers Peru and Ecuador with elevation ranging from 0 to 6518 m MSL (Fig. 1). The new precipitation dataset [Rain for Peru and Ecuador (RAIN4PE)] is generated for the terrestrial land surface between 19°S–2°N and 82°–67°W. The study area has complex hydroclimatic conditions related to its variable climate zones and the Andes Cordillera, which acts as a topographic barrier between the cold and dry eastern Pacific and the warm and moist Amazon region. The Andes divides the study area into three natural drainage basins (Fig. 1): (i) the Pacific basin (watersheds located on the western side of the Andes that convey water to the Pacific Ocean); (ii) Amazon basin (watersheds located on the eastern side of the Andes that drain to Amazon River); and (iii) Titicaca Lake basin (catchments draining into Titicaca Lake).

In the region, the great spatial variability of precipitation patterns is modulated by the interplay among large-scale (e.g., latitudinal migration of Atlantic intertropical convergence zone, South American monsoon systems, Hadley, Walker cell, marine currents, Bolivian high) and local circulation patterns (e.g., upslope and downslope moisture transport) and the complex Andean orography (Laraque et al. 2007; Tobar and Wyseure 2018; Segura et al. 2019; Espinoza et al. 2020). Furthermore, El Niño–Southern Oscillation (ENSO) is a major modulator of hydroclimatology at interannual time scales along the Andes (Poveda et al. 2020). The study area hosts a diversity of ecosystems such as deserts, punas (high mountain grasslands), páramos, glaciers, mountain forests, TMCFs, and rain forests. From these, páramo and TMCF (Fig. 1) are ecosystems where an important cloud/fog water input to the system was reported (Gomez-Peralta et al. 2008; Bruijnzeel et al. 2011; Clark et al. 2014; Cárdenas et al. 2017). This is an important precipitation source to consider in hydrological modeling of páramo and montane watersheds, as it was carried out herein.

b. Data

1) GROUND-BASED PRECIPITATION DATA

The precipitation data of a total of 804 precipitation gauges with record length greater than ten years for the 1981–2015 period were used for this study (Fig. 1), out of which 587 (217) gauges have daily (only monthly) precipitation data. The data were collected from different sources such as national hydrometeorological institutions and previous studies in the region. The data for Peru were obtained from the Peruvian ANA (Autoridad Nacional del Agua) and Aybar et al. (2020); for Ecuador from Morán-Tejeda et al. (2016), Tamayo (2017), Tobar and Wyseure (2018); for Brazil from Xavier et al. (2016, 2017); and for Colombia from IDEAM (Instituto de Hidrología, Meteorología y Estudios Ambientales). We used 587 (804) precipitation gauges with daily (monthly) data for the merging of precipitation datasets at the daily (monthly) time step.

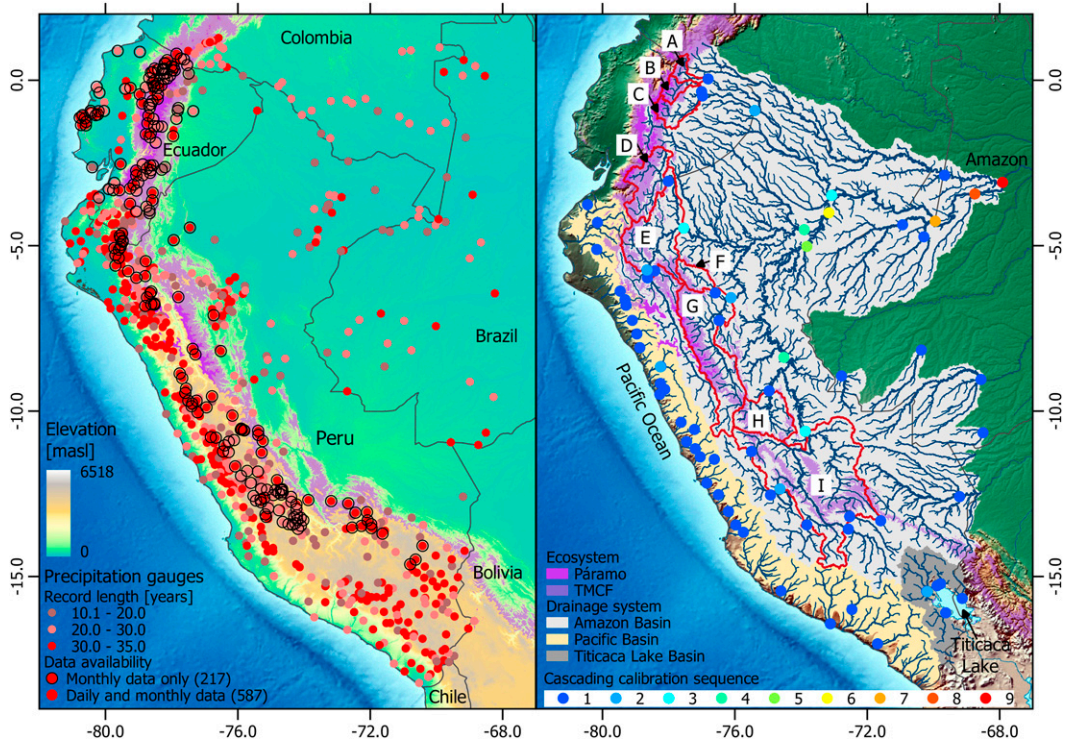


FIG. 1. (left) Study area and spatial distribution of precipitation gauges with record length greater than 10 years for the 1981–2015 period used for the merging procedure. (right) Drainage systems, river networks, and streamflow stations used for hydrological model calibration based on the cascading calibration approach. Red polygons show the gauged catchments with water budget imbalance where gridded precipitation datasets are corrected using streamflow data through reverse hydrology. Nueva Loja station gauges the catchment “A”, San Sebastian (B), Francisco De Orellana (C), Santiago (D), Borja (E), Shanao (F), Chazuta (G), Puerto Inca (H), and Lagarto (I). Boundaries of the páramo and tropical montane cloud forest (TMCF) ecosystems were obtained from Helmer et al. (2019).

2) DISCHARGE DATA

Discharge data of 72 streamflow stations (Fig. 1) with record lengths ranging from one to 33 years for 1983–2015 were obtained from different sources, such as the Peruvian ANA and SENAMHI for catchments draining into the Pacific Ocean and located in the Andes. For the Amazon lowland, data were obtained from the Critical Zone Observatory HYBAM (Hydrogéochimie du Bassin Amazonien, www.sohybam.org). This hydrological network has been operated by an international team from IRD (Institut de Recherche pour le Développement; France), SENAMHI (Peru), INAMHI (Instituto Nacional de Meteorología e Hidrología; Ecuador), and the Brazilian ANA (Agência Nacional de Águas; Brazil) since 2003 (Armijos et al. 2013; Santini et al. 2019).

3) GRIDDED PRECIPITATION DATA

Table 1 presents the five precipitation datasets used in this study. We used the non-gauge-corrected datasets (CHIRP and ERA5) for the merging procedure to generate RAIN4PE dataset. The satellite-based CHIRP precipitation dataset (Funk et al. 2015a) is obtained by considering infrared-based precipitation estimates and corresponding monthly precipitation climatology

generated for Funk et al. (2015b). We selected CHIRP since it has high spatial resolution and long-term (from 1981 onward) daily precipitation data, which is appropriate for long-term hydrometeorological applications. ERA5 (Hersbach et al. 2020) is the latest climate reanalysis dataset produced by the European Centre for Medium Weather Forecasts (ECMWF). Compared with its predecessor ERA-Interim (Dee et al. 2011) that became operational in 2006, ERA5 is based on the ECMWF’s Integrated Forecasting System Cycle 41r2 which was operational in 2016. ERA5 thus benefits from a decade worth of numerical weather prediction developments in model physics, core dynamics, and data assimilation relative to ERA-Interim. Moreover, ERA5 has a much higher temporal and spatial resolution than previous global reanalyses. The hourly ERA5 precipitation data were downloaded and aggregated to obtain daily time step records matching the local gauge observations (from 0700 to 0700 local time).

To compare RAIN4PE against other gauge-corrected precipitation datasets besides the uncorrected ones (CHIRP and ERA5), we selected three merged products (CHIRPS, MSWEP, and PISCO) widely used in data evaluation and hydrometeorological applications in the region (Wongchuig Correa et al. 2017; Paccini et al. 2018; Bhuiyan et al. 2019;

TABLE 1. List of gridded precipitation datasets used in this study. In uncorrected datasets, their temporal dynamics depend entirely on satellite (S) or reanalysis (R) data, while in gauge-corrected datasets, their temporal dynamics depend at least partly on gauge (G) data. In the spatial coverage column, “Global” means fully global coverage including oceans, while “Land” indicates that the coverage is limited to the terrestrial land surface.

Dataset (version)	Data source(s)	Spatial resolution	Spatial coverage	Temporal resolution	Temporal coverage	Reference
Non-gauge-corrected datasets						
CHIRP (V2.0)	S	0.05°	Land, 50°N/S	Daily	1981–present	Funk et al. (2015a)
ERA5	R	0.25°	Global	Hourly	1950–present	Hersbach et al. (2020)
Gauge-corrected datasets						
CHIRPS (V2.0)	G, S	0.05°	Land, 50°N/S	Daily	1981–present	Funk et al. (2015a)
MSWEP (V2.2)	G, S, R	0.1°	Global	3-hourly	1979–present	Beck et al. (2017, 2019b)
PISCO (V2.1)	G, S	0.1°	Peru	Daily	1981–2016	Aybar et al. (2020)

Espinoza et al. 2019; Satgé et al. 2019, 2020; Asurza-Véliz and Lavado-Casimiro 2020; Baez-Villanueva et al. 2020; da Motta Paca et al. 2020; Fernandez-Palomino et al. 2020; Llauca et al. 2021). CHIRPS (Funk et al. 2015a) and PISCO (Aybar et al. 2020) are obtained by merging CHIRP and gauge estimates through deterministic and geostatistical interpolation methods. Finally, MSWEP is derived by optimally merging a range of gauge, satellite, and reanalysis precipitation estimates, where satellite and reanalysis datasets are merged using weights for each one based on the coefficient of determination between 3-day mean gauge- and grid-based precipitation time series (Beck et al. 2017, 2019b). The daily MSWEP precipitation data were provided for this study.

4) ADDITIONAL DATA

In addition to various precipitation products, Table 2 presents other datasets that were used for the hydrological modeling process. The surface elevation data were used both for the merging procedure and setting up the hydrological model.

3. Methods

The framework of this study involves three main steps (Fig. 2): (i) merging procedure through a machine learning technique at the daily and monthly scales; (ii) calibration of model parameters and hydrological adjustment through the reverse hydrology

TABLE 2. Data used for hydrological modeling.

Data type	Resolution	Description/source
Elevation	90 m	Surface elevation (m MSL) from Multi-Error-Removed Improved Terrain (MERIT; Yamazaki et al. 2017)
Land use	100 m	Land use classification representative for the year 2015 obtained from Copernicus Global Land Service (Buchhorn et al. 2019)
Soil	1000 m	Soil parameters for SWAT based on the Harmonized World Soil Database version 1.21 soil data (Abbaspour and Ashraf Vaghefi 2019)
Soil thickness	1000 m	Soil thickness data (Pelletier et al. 2016) were used to implement variable soil thicknesses at hydrological response units (HRUs)
Groundwater table depth	1000 m	Groundwater table depth data (Fan et al. 2013) were used to constrain soil thickness in shallow water tables across the rain forest region
Temperature	Daily/10 km (1981–2016)	Gridded temperature (maximum and minimum) dataset for Peru (Huerta et al. 2018) as provided by SENAMHI (ftp://publi_dgh2:123456@ftp.senamhi.gob.pe/)
Solar radiation	3-hourly/10 km (1983–2018)	Long-term monthly averages of solar radiation based on the global surface solar radiation data (Tang et al. 2019; Tang 2019) were used
Evapotranspiration	Daily/0.25° (1980–2020)	Evapotranspiration data from the Global Land Evaporation Amsterdam Model (GLEAM v3.5a; Miralles et al. 2011; Martens et al. 2017)
Evapotranspiration	8-day/1 km (2000–14)	Evapotranspiration data from the Moderate Resolution Imaging Spectroradiometer Global Evaporation (MOD16; Mu et al. 2011)

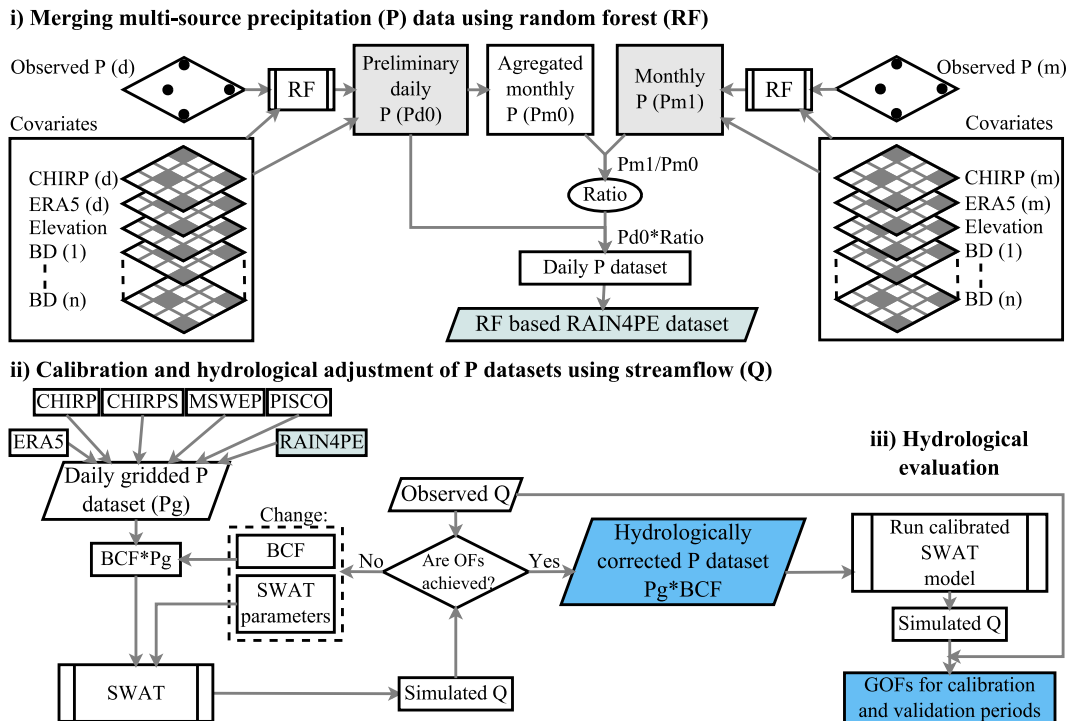


FIG. 2. Flowchart for (i) the generation of gridded precipitation dataset, (ii) hydrological model calibration and adjustment of precipitation datasets, and (iii) hydrological evaluation. Here d (m) indicates the daily (monthly) time step, $BD_{(1), \dots, (n)}$ are buffer distances (distance from any point to all precipitation gauges), BCF is the bias correction factor, OFs are the objective functions for hydrological model calibration, and GOFs are the goodness of fit measures. BCF is optimized only over catchments with water budget imbalance. Note that for hydrological evaluation (step iii), the model was rerun using the respective corrected precipitation data and optimum model parameters values with BCF set to 1.

concept; and (iii) evaluation of all precipitation products through hydrological modeling.

a. Merging procedure

In this section, the merging procedure to obtain RAIN4PE at 0.1° spatial resolution for the 1981–2015 period is described; see Fig. 2 for a scheme.

1) COVARIATES

For the merging procedure at the daily (monthly) scale, we used daily (monthly) precipitation estimates of CHIRP and reanalysis ERA5, surface elevation (Yamazaki et al. 2017), and buffer distances from observation points as covariates. The latter is to account for geographical proximity effects in the prediction process using the random forest (RF) method as suggested by Hengl et al. (2018). The elevation is taken into account because it is a key physical variable with a strong influence on precipitation patterns (Chavez and Takahashi 2017; Beck et al. 2020b). We selected these covariates: satellite precipitation, reanalysis precipitation, and elevation, all of them based on recent studies (Bhuiyan et al. 2019; Baez-Villanueva et al. 2020; Beck et al. 2020b; Hong et al. 2021). To match the 0.1° spatial resolution of the final precipitation product, the covariates with grid cell size $< 0.1^\circ$ ($> 0.1^\circ$) were regridded

to 0.1° spatial resolution applying the bilinear interpolation (nearest neighbor) method.

2) RANDOM FOREST MODELING TO COMBINE DIFFERENT DATA SOURCES

In this study, the RF method (Breiman 2001) was applied to produce a gridded precipitation dataset by merging multiple precipitation sources (gauge, satellite, and reanalysis). RF has been used and proved recently to have similar or superior performance in the interpolation of environmental variables such as precipitation, temperature, and evapotranspiration compared to traditional spatial interpolation techniques, e.g., regression kriging and inverse distance weighting (Hengl et al. 2018; da Silva Júnior et al. 2019; Sekulić et al. 2020). Last, RF-based methodologies (Bhuiyan et al. 2019; Baez-Villanueva et al. 2020) to merge precipitation products with ground-based measurements were developed and applied successfully in data-scarce and complex terrain regions such as the Peruvian and Colombian Andes (Bhuiyan et al. 2019) and Chilean territory (Baez-Villanueva et al. 2020).

RF is a multivariate and nonparametric machine learning algorithm, in which the prediction is generated as an ensemble estimate from a number of regression trees (Breiman 2001) as shown in Eq. (1):

$$\hat{f}(x) = \frac{1}{N} \sum_{b=1}^N T_b(x), \quad (1)$$

where $\hat{f}(x)$ is the final prediction, b is the individual bootstrap sample, N is the total number of trees, and T_b is the individual regression tree.

In RF, each tree is constructed from the random selection of covariates which ensures that trees are decorrelated with each other and a bootstrap sample of the observations (Breiman 2001). The unsampled data, called out-of-bag, can be used to test the prediction accuracy and the importance of input variables, and thus no extra independent validation dataset is needed (Breiman 2001).

We implemented RF using the R package randomForest (Liaw and Wiener 2002) and the following RF parameters: 1) the number of trees (set at 1000); 2) the number of predictor variables randomly selected at each node (set at one-third of the number of variables, default value); 3) the minimum number of observation in a tree's terminal node (set at 5, default value); and 4) the out-of-bag portion to test the accuracy of the predictions (set at one-third of the total number of observations). These parameter values were successfully used in other studies (Baez-Villanueva et al. 2020; Fox et al. 2020; Sekulić et al. 2020).

In the merging procedure (see Fig. 2), an RF model was trained using ground-based observations as the dependent variable and the selected covariates as predictor variables for each day and month in the 1981–2015 period. The trained RF models were then applied to covariates, yielding preliminary daily precipitation data (Pd0) and monthly precipitation data (Pm1). Finally, Pd0 was corrected to match Pm1. For that, the ratio of Pm1 over the monthly precipitation derived from Pd0 was computed on each grid cell for each month, and this ratio was then applied to multiply the Pd0 values on the grid for the month to generate the RF-based RAIN4PE dataset. This correction was because the interpolation of precipitation patterns at a monthly scale is more reliable and accurate than the daily interpolation (Aybar et al. 2020; He et al. 2020).

b. Hydrological modeling and adjustment of precipitation datasets

This section describes the approaches applied for hydrological model calibration and validation and hydrological adjustment of precipitation datasets using streamflow data through the reverse hydrology concept. The hydrological correction is applied only for nine catchments (Fig. 1) having a water budget imbalance due to underestimated streamflow in the simulations with uncorrected precipitation inputs, as reported in previous studies (Zulkafli et al. 2014; Zubieta et al. 2015, 2017; Strauch et al. 2017). We applied the reverse hydrology using the Soil and Water Assessment Tool (SWAT; Arnold et al. 1998) model in which both the bias correction factor (BCF) for precipitation fields and model parameters were calibrated jointly (Fig. 2).

1) SWAT MODEL

Hydrological simulations were performed with the SWAT 2012 model (Arnold et al. 1998), updated for improved representation of tropical vegetation dynamics (Alemayehu et al.

2017). SWAT is one of the most widely used ecohydrological models in the world (Gassman et al. 2014; Tan et al. 2020), and had been applied successfully already for ecohydrological modeling of an Andean basin of Peru (Fernandez-Palomino et al. 2020). SWAT is a process-oriented, semidistributed and time-continuous river basin model used to simulate hydrological processes as well as vegetation dynamics, nutrients, pesticides, and sediment loads within a basin (Arnold et al. 1998; Neitsch et al. 2011). SWAT divides a basin into subbasins, which are then further subdivided into hydrological response units (HRUs) representing unique combinations of land use, soil type, and slope classes (Neitsch et al. 2011). The water balance computation is performed at the HRU level considering four water storage types (snow, soil profile, and shallow and deep aquifers), as follows:

$$\Delta S = \sum_{i=1}^N (\text{BCF} \times P - \text{WYLD} - \text{ET} - \text{GWL}), \quad (2)$$

where ΔS is the change in water storage (mm); N is the time in days; and P , WYLD, ET, and GWL are the amount of precipitation (mm), water yield (mm), evapotranspiration (mm), and groundwater losses (mm), respectively. BCF introduced herein is the bias correction factor to infer the precipitation fields from observed streamflow data.

In SWAT, flow routing in river channels can be computed using the Muskingum or the variable storage method, considering the flow velocity to be the same across the channel and floodplain section (Neitsch et al. 2011). This approach has been shown to be inefficient for flow routing in Amazon rivers (Santini 2020), where flows are largely affected by floodplains that act as reservoirs, causing significant flood peak delay and attenuation (Paiva et al. 2011; Yamazaki et al. 2011; Santini et al. 2015; Santini 2020). To exclude this limitation, Santini (2020) has implemented a new flow routing method for SWAT to consider the river-floodplain dynamics, where the associated floodplain of each river reach was treated as a simple storage model, as in other hydrological models (Paiva et al. 2011; Yamazaki et al. 2011). This approach was used in our study.

2) SWAT MODEL SETUP, CALIBRATION, AND VALIDATION

The SWAT model was set up for Peruvian and Ecuadorian catchments (total of 1 638 793 km²) based on the input data listed in Table 2. The model includes 2675 subcatchments and 6843 HRUs. Channel cross-section parameters such as the bankfull width (B) and channel depth (CHD) were estimated using geomorphologic equations based on upstream drainage areas derived for Amazon rivers (Paiva et al. 2011). Floodplain width is estimated by multiplying the bankfull width by a factor (set at 5, default value). We assigned Manning's n values of 0.03 (0.10) for channels (floodplains). The modified Soil Conservation Service curve number, the Priestley–Taylor equation, and the variable storage methods were used to simulate surface runoff and infiltration, potential evapotranspiration, and river flow routing, respectively.

TABLE 3. Parameters and their ranges for model calibration for evapotranspiration (ET), streamflow (Q), and precipitation (P). In the “Change type” column, R (V) refers to a relative (absolute) change of parameter values during the calibration. Parameter set 1 was applied for Andean catchments draining into the Pacific Ocean and Titicaca Lake and for Andean catchments upstream the montane watersheds. Montane watersheds having a water budget imbalance were calibrated using parameter set 2. Catchments downstream the montane watersheds were calibrated using parameter set 3. Note that BCF is applied only for catchments with water budget closure problems to infer precipitation from streamflow data. See Neitsch et al. (2011) for detailed parameter definitions.

Parameter	Description (unit)	Calibrated output	Range	Change type	Set		
					1	2	3
SOL_AWC	Soil available water capacity (mm H ₂ O/mm soil)	ET	[−0.8, 0.8]	R	X		
GW_REVAP	Groundwater “revap” coefficient	ET	[0, 0.2]	V			X
SURLAG	Surface runoff delay coefficient	Q	[0.1, 2]	V	X	X	X
GW_DELAY	Groundwater delay time (days)	Q	[1, 100]	V	X	X	X
RCHRG_DP	Deep aquifer percolation fraction	Q	[0, 1]	V	X	X	X
GWQMN	Threshold for return flow from shallow aquifer (mm)	Q	[500, 1000]	V	X	X	X
ALPHA_BF	Baseflow recession constant	Q	[0.01, 1]	V	X	X	X
CH_K2	Hydraulic conductivity of main channel (mm h ^{−1})	Q	[0, 50]	V			X
CHD	Main channel depth (m)	Q	[−0.1, 0.5]	R			X
FP_W_F	Ratio of floodplain width over bankfull width	Q	[1, 5]	V			X
BCF	Bias correction factor	P, Q	[0, 1]	R		X	

The simulation period was from 1981 to 2015. The first two years were considered for the model spinup. For the model calibration, all flow data were used for stations with a record lower than 10 years, and for those with longer, two-third of the data were used. In the latter case, the remaining flow data were used for model validation (53 out of 72 streamflow stations). The model calibration for each precipitation product was performed applying the multisite cascading calibration approach (Xue et al. 2016) in nine sequences (Fig. 1), where the calibrated discharge from the upstream catchments was used as input for the downstream. The model parameters and BCFs for each (sub)catchment were calibrated using the respective set of parameters defined in Table 3 for Andean, montane, and lower Amazon catchments. Moreover, plant parameters were adopted from our previous study (Fernandez-Palomino et al. 2020).

The optimum values of model parameters and BCFs were obtained through multiobjective calibration. For that, the model was calibrated against observed discharge using the Nash–Sutcliffe efficiency log (INSE) and aggregated flow duration curve signature (FDC_{sign}) as objective functions (see Table 4). We selected INSE and FDC_{sign} since these have been shown sufficient to test the model for simulating all hydrograph aspects in the calibration (Fernandez-Palomino et al. 2020). Moreover, the application of FDC-based signatures provides more information about the hydrological behavior of the modeled basin (Yilmaz et al. 2008; Hrachowitz et al. 2014) and leads to better parameter identifiability, more accurate discharge simulation, and reduction of predictive uncertainty (Yilmaz et al. 2008; Pokhrel and Yilmaz 2012; Hrachowitz et al. 2014; Pfannerstill et al. 2014, 2017; Chilkoti et al. 2018; Fernandez-Palomino et al. 2020; Sahraei et al. 2020). Following Chilkoti et al. (2018) and Fernandez-Palomino et al. (2020), we estimated percent bias for four segments of the FDC [peak flow (0%–2%), high flow (2%–20%), midsegment (20%–70%), and low flow (70%–100%)], and then the absolute values of the bias percentages were

averaged to obtain the FDC_{sign} to take into consideration the hydrological signatures for model calibration. The respective FDC segmentation represents peak flow events occurring rarely, quick runoff (due to snowmelt and/or rainfall), the flashiness of a basin’s response, and the streamflow’s baseflow components. The Borg multiobjective evolutionary algorithm (Borg MOEA; Hadka and Reed 2013) was used to optimize the objective functions (maximization of INSE and minimization of FDC_{sign}) with 1000 iterations as maximum. The Borg MOEA parameterization was the same as in Fernandez-Palomino et al. (2020). The parameters for ungauged catchments (at HRU level) were obtained applying the spatial proximity approach (Guo et al. 2021) using the inverse distance weighting (Shepard 1968). For regionalization of parameters, donor catchments (gauged) within a radius of 150 km were used to avoid the influence of Amazonian catchments in the estimation of parameters for Andean basins draining into the Pacific Ocean and Titicaca Lake.

3) HYDROLOGICAL ADJUSTMENT OF PRECIPITATION DATASETS

The optimum BCFs obtained for each catchment with water budget imbalance (Fig. 1) in the calibration procedure were applied to the respective daily gridded precipitation data to obtain the hydrologically corrected daily precipitation dataset (Fig. 2). For that, a continuous BCF map at 0.1° spatial resolution was produced where grid cells within the respective catchment retained the respective BCF, and for cells on the boundary, the area-weighted BCFs were estimated. It is noteworthy that applying the resulting BCF map to gridded precipitation data can result in spatial discontinuities of precipitation patterns at the border of the catchments. To reduce such discontinuities, we further applied a 5 × 5 mean filter to the BCF map. Finally, the corrected precipitation data were used as input to SWAT to run the model with the

TABLE 4. Mathematical formulation of the goodness of fit metrics and hydrological signatures. Here, O and S are observed and simulated flow ($\text{m}^3 \text{s}^{-1}$), respectively; EP is exceedance probability; P , H , and L are the indices of the minimum flow of the peak flow, high flow, and low flow segments, respectively. In the optimization process for hydrological model calibration, INSE was maximized, whereas FDC_{sign} was minimized.

Criterion (reference)	Equation	Description
Discharge-related performance measures		
Nash–Sutcliffe efficiency (Nash and Sutcliffe 1970)	$\text{NSE} = 1 - \frac{\sum_{i=1}^n (S_i - O_i)^2}{\sum_{i=1}^n (O_i - O_a)^2}$	O_a is the average of the observed flow and n is the number of observations under evaluation
Nash–Sutcliffe efficiency log (Krause et al. 2005)	$\text{INSE} = 1 - \frac{\sum_{i=1}^n [\ln(S_i) - \ln(O_i)]^2}{\sum_{i=1}^n [\ln(O_i) - \ln(O_p)]^2}$	
Percent bias (Gupta et al. 1999)	$\text{PBIAS} = \frac{\sum_{i=1}^n (S_i - O_i)}{\sum_{i=1}^n O_i} \times 100$	
Kling–Gupta efficiency (Gupta et al. 2009; Kling et al. 2012)	$\text{KGE} = \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2}$	r is the Pearson product-moment correlation coefficient and beta (gamma) indicates the bias (relative dispersion) between observed and simulated flows
Signature measures based on flow duration curve (FDC)		
Percent bias in FDC peak segment volume (Yilmaz et al. 2008)	$S_{\text{peak}} = \frac{\sum_{p=1}^P (S_p - O_p) \times 100}{\sum_{p=1}^P O_p}$	$p = 1, 2, \dots, P$ are flow indices located within the FDC peak flow segment (EP lower than 2%)
Percent bias in FDC high segment volume (Yilmaz et al. 2008)	$S_{\text{high}} = \frac{\sum_{h=1}^H (S_h - O_h) \times 100}{\sum_{h=1}^H O_h}$	$h = 1, 2, \dots, H$ are flow indices located within the high flow segment (2%–20% flow EP)
Percent bias in FDC midsegment slope (Yilmaz et al. 2008; van Werkhoven et al. 2009)	$S_{\text{mid}} = \frac{[(S_{m1} - S_{m2}) - (O_{m1} - O_{m2})] \times 100}{(O_{m1} - O_{m2})}$	$m1$ and $m2$ are the lowest and highest flow EP, respectively, within the midsegment (20%–70%)
Percent bias in FDC low segment volume (Yilmaz et al. 2008)	$S_{\text{low}} = \frac{\sum_{l=1}^L (S_l - O_l) \times 100}{\sum_{l=1}^L O_l}$	$l = 1, 2, \dots, L$ are flow indices located within the low flow segment (70%–100% flow EP)
FDC signature (Chilkoti et al. 2018)	$\text{FDC}_{\text{sign}} = \frac{1}{4} (S_{\text{peak}} + S_{\text{high}} + S_{\text{mid}} + S_{\text{low}})$	FDC_{sign} is the aggregated FDC signature

respective optimum parameters for the simulation period to compute the model performance measures for the hydrological evaluation of precipitation datasets.

c. Evaluation methods

1) EVALUATION USING OUT-OF-BAG SAMPLE

The prediction accuracy of preliminary daily precipitation data (Pd0) and monthly precipitation data (Pm1) produced by the RF method (see Fig. 2) was assessed using the mean

absolute error (MAE) and determination coefficient (R^2) based on the out-of-bag sample.

2) HYDROLOGICAL EVALUATION

We evaluated the accuracy of precipitation estimates through hydrological modeling for the three drainage systems in the study area. It is an adequate approach evaluating gauge-corrected precipitation datasets since streamflow observations are independent from ground precipitation observations that are

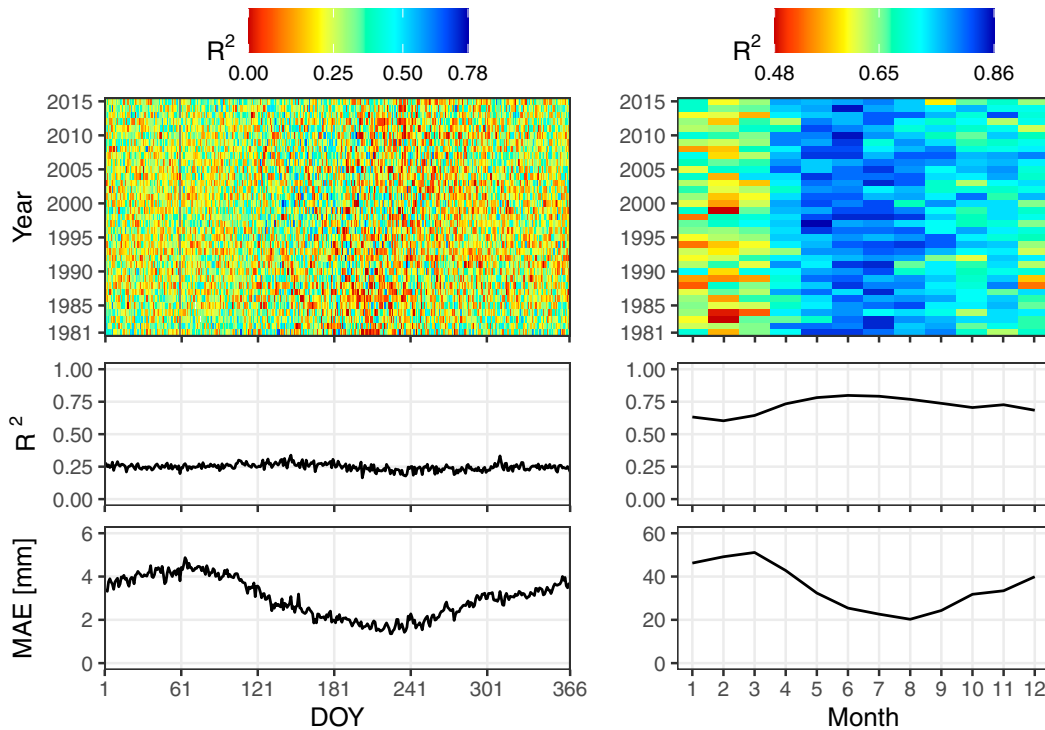


FIG. 3. Performance of the random forest algorithm for spatial interpolation of (left) daily and (right) monthly precipitations. Here, R^2 is the coefficient of determination, and MAE is the mean absolute error. The middle and bottom graphs show the performance measures averaged for each day or month in the 1981–2015 period.

used in these datasets (Beck et al. 2020a; Brocca et al. 2020; Satgé et al. 2020).

For hydrological evaluation, a multicriteria evaluation of SWAT-simulated streamflow using all precipitation products was carried out. For that, both hydrograph goodness of fit metrics and hydrological signatures (Table 4) were considered for both calibration and validation periods. The modified Kling–Gupta efficiency (KGE) and percent bias (PBIAS) were used for assessing model skills in representing general discharge dynamics and over or underestimation tendencies, respectively; INSE and percent bias in FDC low segment volume (S_{low}) for low flows; Nash–Sutcliffe efficiency (NSE) and percent bias in FDC high segment volume (S_{high}) for high flows; and percent bias in FDC peak segment volume (S_{peak}) for extremely peak flow conditions. This multicriteria evaluation aims to assess model skills representing all aspects of the observed FDC and hydrographs, which is important for assessing the reliability of precipitation products for hydrometeorological applications such as the analysis of water budget and hydroclimatic extremes (floods and droughts). The hydrological model performance was ranked based on the rating performance criteria of Moriasi et al. (2007). Thus, for simplicity, the absolute values of PBIAS, S_{low} , S_{high} , and $S_{peak} < 10$ were considered as very good, (10–15) good, (15–25) satisfactory, and (>25) unsatisfactory, and KGE, NSE, and INSE > 0.75 were considered very good, (0.65–0.75) good, (0.50–0.65) satisfactory, and (<0.50) unsatisfactory.

Furthermore, in this study, we analyzed the distribution of model parameters and compared the evapotranspiration (ET) simulated by SWAT with remotely sensed ET from Global Land Evaporation Amsterdam Model (GLEAM) and Moderate Resolution Imaging Spectroradiometer Global Evaporation (MOD16). The ET estimates from MOD16 and GLEAM are based on the Penman–Monteith and Priestly–Taylor equations, respectively. This comparison is to verify the plausibility of ET estimates which is one of the largest components of the water budget besides precipitation and difficult to estimate over complex terrain. Results of the analysis of parameter distribution and ET estimates are described in appendices B and C.

4. Results

a. Performance of the merging algorithm

The skill of the RF method for predicting daily and monthly precipitation patterns was evaluated using performance measures (R^2 and MAE) based on the out-of-bag sample. Figure 3 shows that based on the temporal distribution of R^2 , the RF performance does not have a seasonal pattern for the daily precipitation prediction, whereas it exhibits better performance in the period from April to December for monthly prediction. Furthermore, R^2 shows that prediction is better for the monthly (mean $R^2 = 0.72$) than the daily (mean

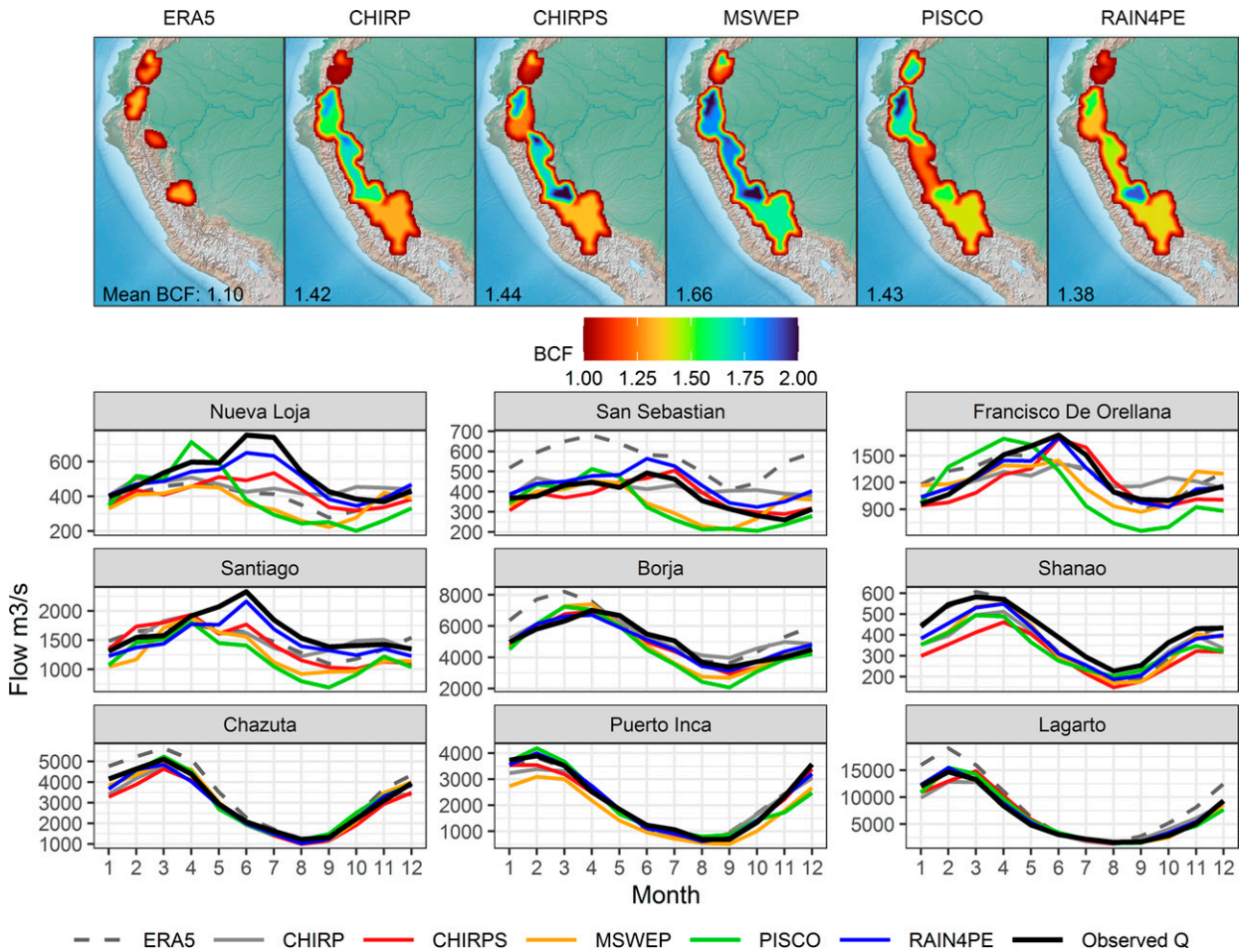


FIG. 4. (top) Bias correction factors (BCFs) for six precipitation datasets and (bottom) long-term mean seasonal streamflow (Q) dynamics in the period 1983–2015 after SWAT model calibration over nine catchments with underestimation of precipitation amounts in comparison with the observed mean seasonal discharge. The mean BCF was computed using the catchment areas as weights. Note that both observed and seasonal streamflow were computed only for the months with available streamflow data.

$R^2 = 0.25$) precipitation. This result supports the correction of daily-predicted precipitation values to match the monthly predictions performed in our study as described in the methods. MAE is much lower in the period June–September for both daily and monthly precipitation prediction, indicating that precipitation is more easily predictable when most of the study area experiences lower precipitation during the dry season. It is important to mention that satellite precipitation (CHIRP) was often the most important covariate in the merging procedure both at daily and monthly scale, followed by reanalysis precipitation (ERA5) and terrain elevation, while buffer distances were negligible (Fig. S1 in the online supplemental material).

b. Hydrological correction of the gridded precipitation datasets

The spatial variation of the obtained bias correction factors (BCFs) for six precipitation datasets is shown in Fig. 4. This differs from the method of Strauch et al. (2017), who applied

a unique correction factor to WFDEI (Weedon et al. 2014) dataset for all montane regions. The lower values of BCFs for ERA5 are related to significant precipitation overestimation along the Andes by ERA5 (Figs. 5 and 6). The results for the other datasets (Fig. 4) show that higher BCFs were the result for MSWEP (mean BCF = 1.66) and lower for RAIN4PE (mean BCF = 1.38). For a BCF of 1.38, on average, 28% of total precipitation is the precipitation underpredicted in páramo and montane watersheds in the study area which falls in the range (0%–30%) of cloud/fog water contribution to total precipitation reported in previous studies of the region (Gomez-Peralta et al. 2008; Cárdenas et al. 2017). Figure 4 also shows that significant benefits of precipitation correction made for RAIN4PE are obvious in a good representation of streamflow seasonality for all nine catchments. The correction of CHIRPS also works relatively well in most of the catchments in terms of seasonal streamflow prediction, although it fails over the southern Ecuadorian Amazon (at Santiago station). The hydrological correction of the other datasets

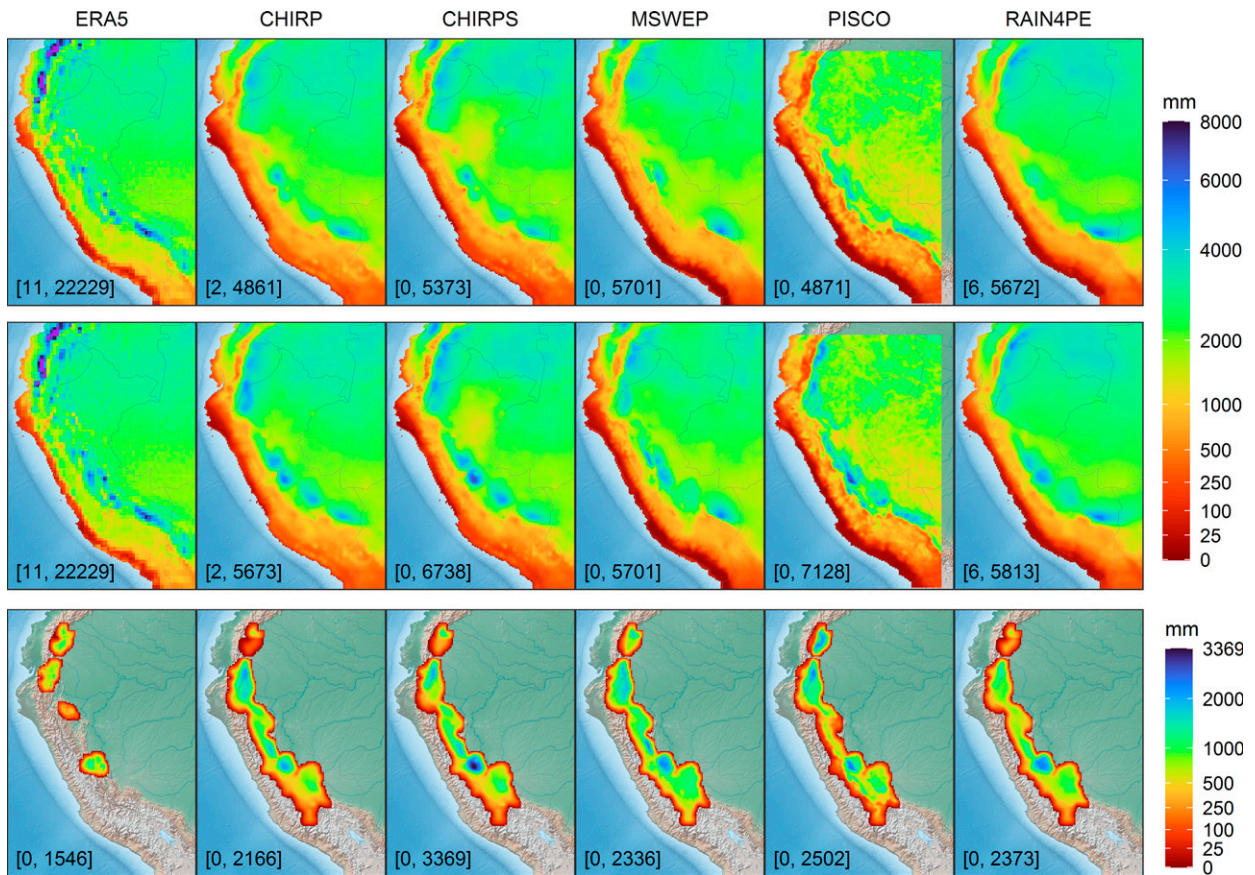


FIG. 5. The spatial patterns of average annual precipitation for the period 1981–2015 based on (top) raw and (middle) hydrologically adjusted precipitation data of ERA5, CHIRP, CHIRPS, MSWEP, PISCO, and RAIN4PE. (bottom) The underestimated precipitation fields for each precipitation dataset. The numbers in brackets represent the precipitation ranges. In the case of ERA5, precipitation values exceeding 8000 mm are in purple (distributed over the Ecuadorian Andes mainly).

(CHIRP, ERA5, MSWEP, and PISCO) performs well for southern catchments (from Borja to Lagarto station) but not in Ecuadorian catchments (from Nueva Loja to Santiago station) since the streamflow seasonality change is underestimated, indicating a serious drawback of these datasets.

c. Spatial patterns of precipitation

In general, the spatial variability of the long-term average annual precipitation (1981–2015) portrayed by all precipitation datasets looks quite similar (Fig. 5), although PISCO shows distinct precipitation patterns and magnitudes in the rain forest regions. Figure 5 also shows the spatial patterns of the estimated precipitation underestimates for each precipitation dataset. As can be seen, these patterns look quite similar over the Peruvian Amazon for five datasets (CHIRP, CHIRPS, MSWEP, PISCO, and RAIN4PE) but vary over the northern Amazon basin in Ecuador. The substantial precipitation underestimation (ranging from 0 to 3369 mm, Fig. 5) found here suggests that precipitation correction was necessary to achieve the closure of the water budget and appropriate hydrological modeling of the páramo and montane watersheds.

In addition, a comparison of the unadjusted precipitation data with gauge observation was done (Fig. 6) to assess precipitation datasets' reliability or critical shortcomings. It shows that ERA5 overestimates precipitation significantly over the Andes. CHIRP, CHIRPS, MSWEP, and PISCO (CHIRP and CHIRPS) underestimate (overestimate) precipitation over the northern (arid southern) Pacific coastal areas. Furthermore, ERA5, CHIRP, CHIRPS, MSWEP, and PISCO have inconsistent temporal distribution of precipitation over the northern Amazon, which is confirmed by low values of correlation and determination coefficients that result from comparing these products with gauge observations at a monthly scale (Fig. 6) and SWAT-simulated seasonal streamflow using these datasets (Fig. 4). Therefore, these datasets are less suitable for characterizing spatiotemporal variability of precipitation over the Ecuadorian Amazon than RAIN4PE. However, it should be kept in mind that the comparison measures in Fig. 6 could be biased toward datasets (CHIRPS, MSWEP, PISCO, and RAIN4PE) that used data from the assimilated precipitation gauges in their production (see Table 1).

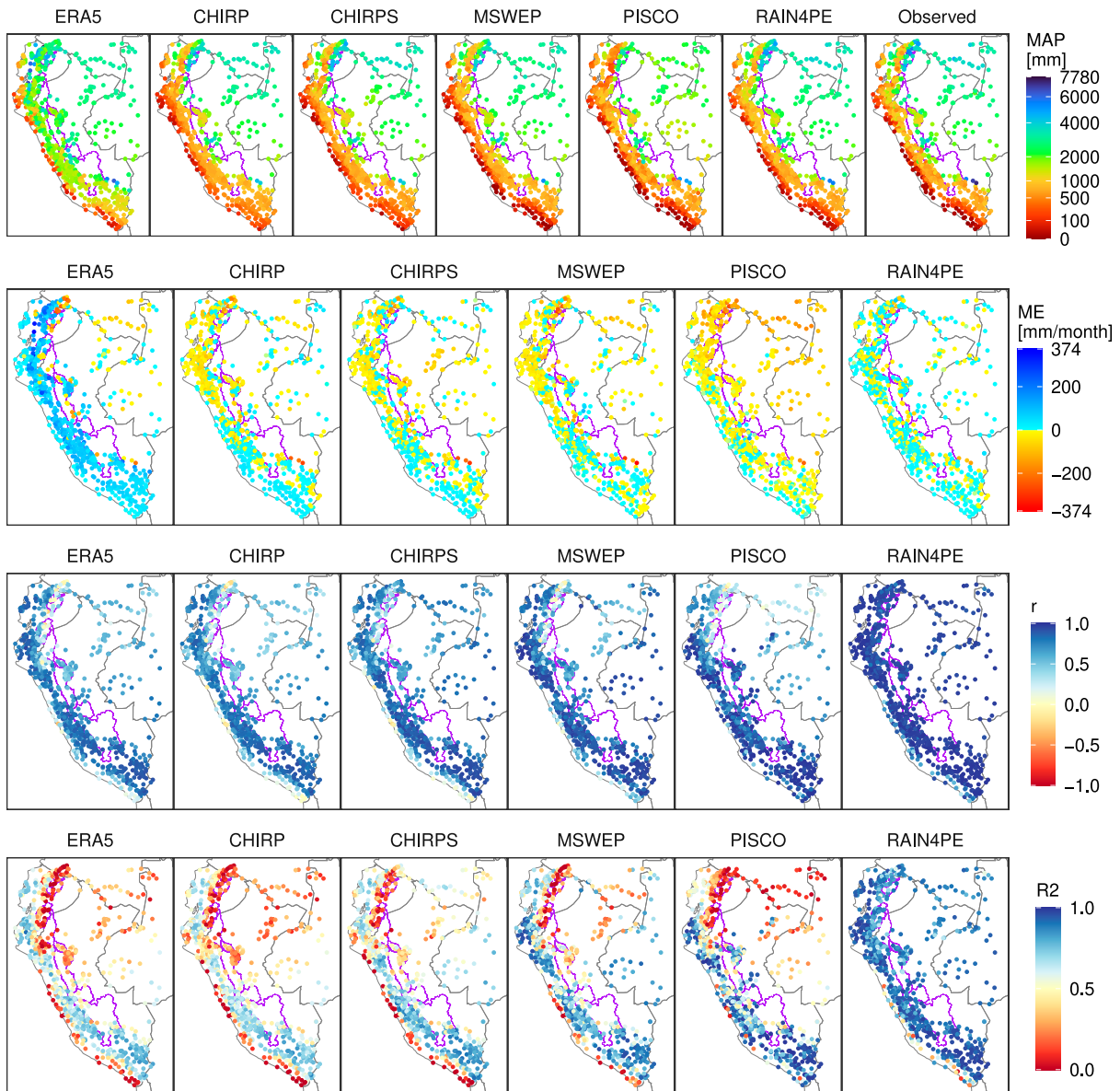


FIG. 6. Performance of the unadjusted precipitation datasets in comparison with gauge observations: MAP is the mean annual precipitation, ME is the mean error, r is the Pearson's correlation coefficient, and R^2 is the coefficient of determination. The comparison measures (ME, r , and R^2) were computed using monthly precipitation time series for 1981–2015.

d. Hydrological evaluation

In this section, we evaluate the performance of the SWAT model driven by the hydrologically adjusted CHIRP (CHIRP-SWAT), ERA5 (ERA5-SWAT), CHIRPS (CHIRPS-SWAT), MSWEP (MSWEP-SWAT), PISCO (PISCO-SWAT), and RAIN4PE (RAIN4PE-SWAT) for calibration and validation periods. For that, we used multiple performance measures to assess the model skills in representing discharge dynamics including all flow conditions (low, high, and peak flows). It is important to mention that temporal mismatches in the daily precipitation accumulation may influence the model performance at the daily

scale since CHIRP, CHIRPS, and MSWEP were delivered using different daily time window aggregation than the local one (from 0700 to 0700 local time). Furthermore, our analyses are based on the results of the only one hydrological model, SWAT, and the application of other hydrological models could be done in future to verify and refine the obtained results.

1) PERFORMANCE EVALUATION FOR DAILY STREAMFLOW AND EXTREMES

We investigated the spatial variability of hydrological model performance for streamflow simulation forced by six

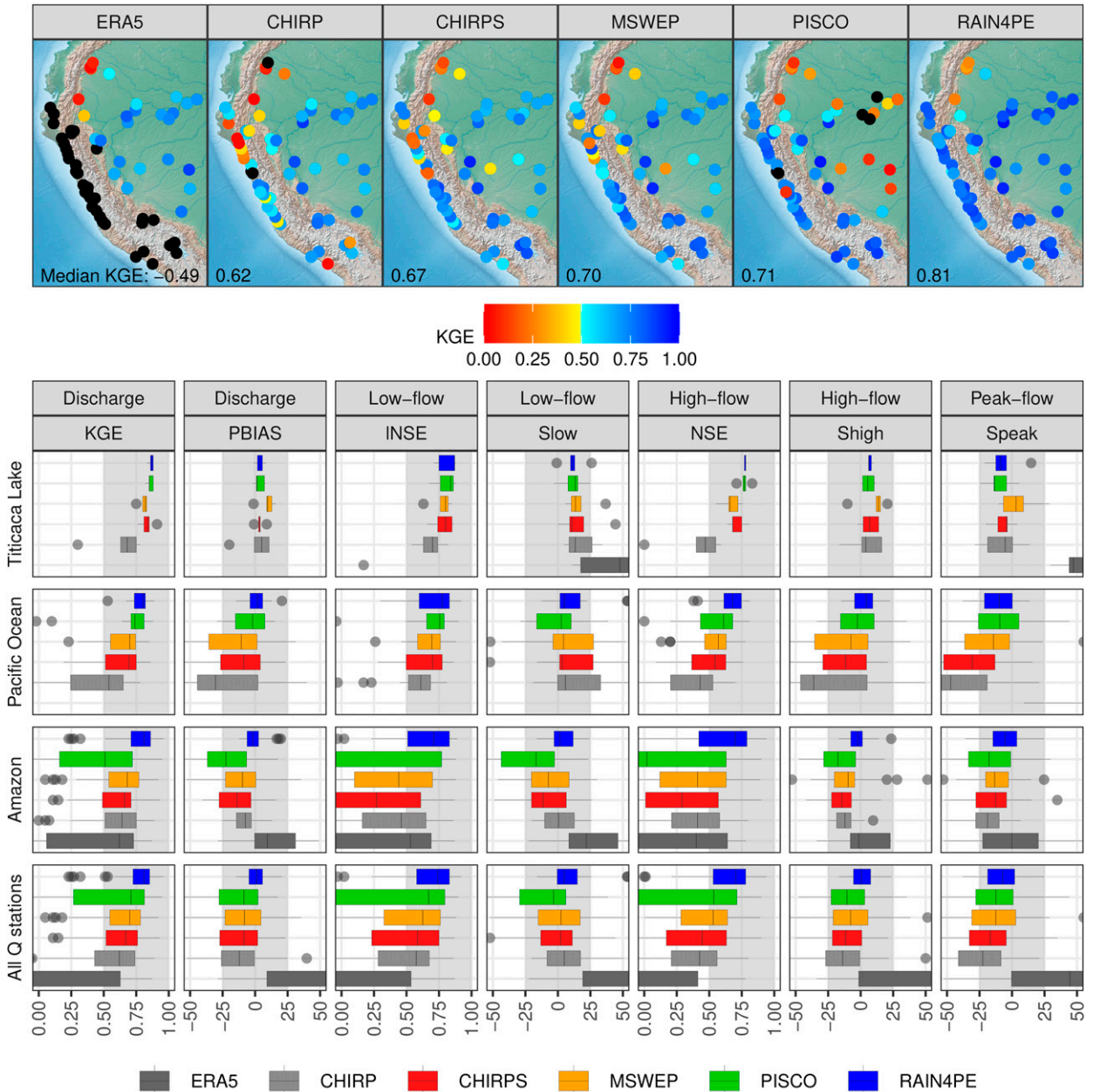


FIG. 7. Hydrological model performance metrics for daily streamflow simulations by SWAT driven by six precipitation datasets in the calibration period: (top) spatial patterns of KGE and (bottom) boxplots showing seven criteria for all streamflow (Q) stations and stations located in catchments draining into the Amazon River, Pacific Ocean, and Titicaca Lake. The datasets are sorted in ascending order of the median KGE for all Q stations. Values exceeding 0.5 (between $\pm 25\%$) for KGE, INSE, and NSE (PBIAS, S_{low} , S_{high} , and S_{peak}) are considered skillful (marked by light gray background in boxplots). Black points in the upper part represent negative values of KGE. Note that the x axis starts at 0 for KGE, NSE, and INSE to improve visualization, whereas PBIAS, S_{low} , S_{high} , and S_{peak} were constrained between $\pm 50\%$.

precipitation products in calibration (Fig. 7, Table S1) and validation (Fig. S2, Table S2) periods. These figures present the Kling–Gupta efficiency spatial distribution and show results in terms of seven criteria for all streamflow stations and catchments draining into the Titicaca Lake, the Pacific Ocean, and the Amazon River as boxplots. Table 5 shows each criterion’s median values for each drainage system and

precipitation product for the simulation period (1981–2015). The results described in this section are based on the outputs for calibration period (Fig. 7) but they are also valid for the validation period (Fig. S2), as results for both periods are similar.

Results for catchments draining into Titicaca Lake show that SWAT driven by gauge-corrected precipitation datasets

TABLE 5. Median values of each performance measure for daily streamflow simulation for the period 1983–2015 (without the spinup period). Values in bold denote the best performing product in each drainage system and the study area according to the specific score on the left.

Basin	Product	KGE	NSE	INSE	PBIAS	S_{low}	S_{high}	S_{peak}
Titicaca Lake	ERA5	-1.1	-3.04	-0.07	205.9	47.8	197.5	44.5
Titicaca Lake	CHIRP	0.62	0.29	0.67	17.5	31.4	18.2	7.2
Titicaca Lake	CHIRPS	0.81	0.64	0.76	6.1	16	9	7.3
Titicaca Lake	MSWEP	0.79	0.67	0.8	14.6	20	15.2	8.8
Titicaca Lake	PISCO	0.84	0.74	0.82	7.2	18.5	7.6	13.9
Titicaca Lake	RAIN4PE	0.86	0.77	0.86	6.8	14.3	6.7	14.1
Pacific	ERA5	-2.29	-7.2	-1.47	325.6	238	293.9	117.9
Pacific	CHIRP	0.55	0.44	0.62	28.9	14.1	35.8	43.5
Pacific	CHIRPS	0.66	0.5	0.66	15.2	11.1	12.1	18.7
Pacific	MSWEP	0.68	0.55	0.66	17.4	14.8	21	27.5
Pacific	PISCO	0.74	0.57	0.73	10.2	11.4	10.7	17.3
Pacific	RAIN4PE	0.78	0.67	0.74	5.2	6	7.5	7.7
Amazon	ERA5	0.63	0.46	0.6	12.5	19.5	17.2	18.8
Amazon	CHIRP	0.63	0.4	0.45	7.7	12	9.6	12.6
Amazon	CHIRPS	0.67	0.31	0.38	13.7	15.3	14.5	14.6
Amazon	MSWEP	0.69	0.47	0.48	15.4	15.2	11.8	12.2
Amazon	PISCO	0.49	0	-0.11	21.9	23.9	18	18.9
Amazon	RAIN4PE	0.8	0.7	0.73	6.2	10.8	7.3	10.7
All watersheds	ERA5	-0.58	-2.26	-0.3	148.4	55	115.1	44.2
All watersheds	CHIRP	0.58	0.43	0.57	13.2	13.9	16.3	21.4
All watersheds	CHIRPS	0.67	0.41	0.58	13.6	15	13	15.8
All watersheds	MSWEP	0.7	0.53	0.63	15.3	16.5	15.2	15.1
All watersheds	PISCO	0.7	0.51	0.66	12.8	18.6	12.7	17.7
All watersheds	RAIN4PE	0.8	0.7	0.74	5.9	9.7	7.3	9.6

performs satisfactorily to very good for daily streamflow simulation (median KGE ≥ 0.79), including all flow conditions. The good performance of MSWEP and CHIRPS for hydrological modeling in the Titicaca Lake basin shown here coheres with the performance demonstrated in [Satgé et al. \(2019, 2020\)](#). However, RAIN4PE (median KGE = 0.86) was shown in our simulation to be the best choice for this drainage system. Regarding two non-gauge-corrected datasets, CHIRP-SWAT has unsatisfactory performances for high-flow dynamics, and ERA5-SWAT significantly overestimates streamflow (Fig. S4).

In the Pacific basin, CHIRPS-SWAT, MSWEP-SWAT, and even PISCO-SWAT have low KGE (≤ 0.5), high biases, and poor performance for high and peak flows for some stations. The outcome for MSWEP and PISCO aligns with the findings of previous studies ([Bhuiyan et al. 2019](#); [Derin et al. 2019](#); [Asurza-Véliz and Lavado-Casimiro 2020](#)). CHIRP-SWAT has more skill than ERA5-SWAT, which shows a significant overestimation of streamflow; however, they both are outperformed by the gauge-corrected precipitation datasets. The overall good performance of RAIN4PE-SWAT (median KGE = 0.78) allowed us to conclude that RAIN4PE is the most suitable precipitation product for daily streamflow simulation (including all flow conditions and water budget closure) in the catchments draining into the Pacific Ocean.

In the Amazon basin, among the six precipitation products driving SWAT, RAIN4PE (median KGE = 0.80) provided the best performance measures for daily streamflow simulation (including all flow conditions). PISCO (median KGE =

0.49) provided the worse measures, particularly over the lower Amazon catchments which is consistent with previous studies ([Aybar et al. 2020](#); [Llauca et al. 2021](#)). Despite the fact that median KGE (>0.5) is satisfactory for CHIRP, CHIRPS, ERA5, and MSWEP, the other measures such as the INSE and NSE show that they tend to perform unsatisfactorily for the simulation of low- and high-flow dynamics. However, KGE patterns (Fig. 7) show unsatisfactory scores over the Ecuadorian Amazon catchments, showing the limitations of all products (including RAIN4PE) in portraying the actual daily precipitation variability there.

In general, SWAT performance for all streamflow stations (Fig. 7 and Fig. S2 and Table 5, Tables S1 and S2) suggests that RAIN4PE (e.g., median KGE = 0.80) is the most appropriate product for daily streamflow simulation, including all flow conditions in the study area.

2) PERFORMANCE EVALUATION FOR MONTHLY STREAMFLOW

Fig. 8 and Fig. S3 display the spatial distribution of KGE, NSE, INSE, and PBIAS to assess the SWAT model skill for the monthly streamflow simulation in the calibration and validation periods. These figures show that results in both periods are quite similar, although the overall performance of PISCO-SWAT and MSWEP-SWAT is a bit lower in the validation period. Based on results of model performance in the validation period (Fig. 8), among the six precipitation products driving SWAT, overall RAIN4PE (median KGE = 0.86,

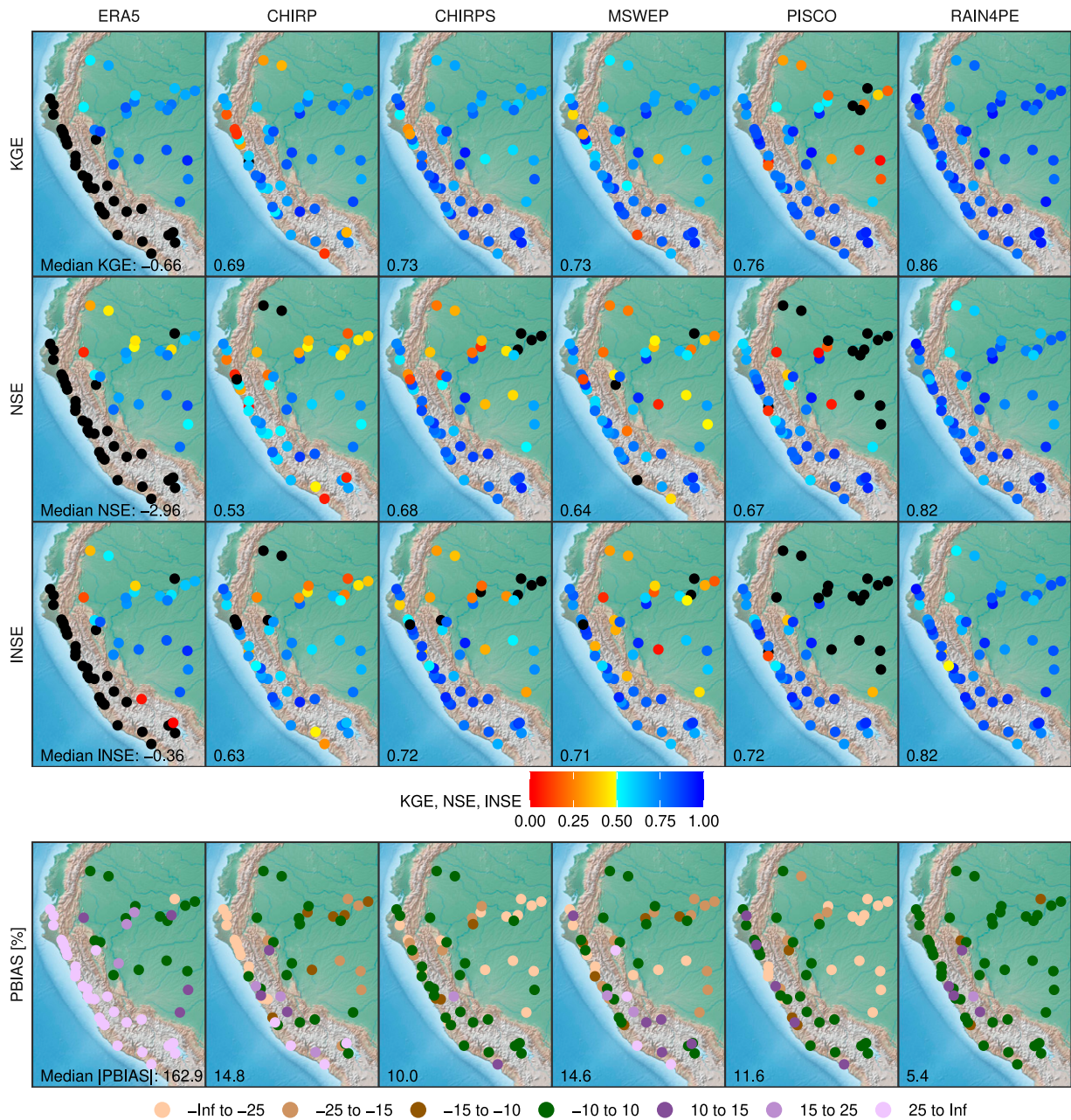


FIG. 8. Hydrological model performance metrics KGE, NSE, INSE, and PBIAS for monthly streamflow simulations by SWAT driven by six precipitation datasets in the validation period. Black points represent negative values of KGE, NSE, and INSE.

NSE = 0.82, INSE = 0.82, and $|PBIAS| = 5.4\%$) provided the best performance measures for monthly streamflow simulation in all evaluated catchments. Despite the median KGE, NSE, and INSE were satisfactory (>0.5 , Fig. 8) for CHIRP, CHIRPS, MSWEP, and PISCO, the spatial patterns of these measures show the limitation (e.g., NSE < 0.5) of these products for hydrological modeling over the Ecuadorian Amazon, lower Amazon, and some catchments draining into the Pacific

Ocean, which is in agreement with the results for the daily outputs. Otherwise, ERA5-SWAT was found to perform unsatisfactorily for Andean basins, although its performance improved for larger catchments in the Amazon basin. The overall very good performance in accordance with criteria by Moriasi et al. (2007) obtained by RAIN4PE-SWAT highlights the increased utility of RAIN4PE for countrywide hydrometeorological applications in Peru and Ecuador.

5. Discussion

a. Advantages of the merging methodology

This study demonstrates a successful method for merging multiple precipitation sources (based on gauge, satellite, and reanalysis data) with surface elevation using the RF method to generate a spatially gridded precipitation dataset RAIN4PE. This is supported by the significant improvement of RAIN4PE for hydrological simulations compared to the non-gauge-corrected datasets (CHIRP and ERA5) used for the merging procedure. Furthermore, the superiority of RAIN4PE regarding the gauge-corrected datasets (CHIRPS, MSWEP, and PISCO) for hydrological simulations suggests that the methodology applied herein to generate RAIN4PE is much more robust than that of the other merged precipitation products. This means that the RF method is more effective in merging multiple precipitation data sources than deterministic and geostatistical interpolation methods (Funk et al. 2015a; Aybar et al. 2020) and merging approaches that use weights for each source (Beck et al. 2017, 2019b). Compared to the aforementioned merging approaches, RF has the flexibility to include multiple precipitation sources and environmental variables (e.g., surface elevation) that explain precipitation patterns. Besides this advantage, RF can capture nonlinear dependencies and interactions of variables, such as the nonlinear interactions among the precipitation and terrain elevation due to the complex Andes morphology (Figs. 1 and 5; Chavez and Takahashi 2017), which could be challenging to model using geostatistical techniques. However, it is important to keep in mind the RF limitation for predicting value beyond the range in the training data (Hengl et al. 2018). Overall, the results of our study provide a reference for merging multi-source precipitation data and environmental variables using RF in complex data-scarce regions.

b. Hydrological correction of the gridded precipitation datasets

The high BCF values (Fig. 4) obtained to correct gridded precipitation biases make evident that most of the datasets evaluated (CHIRP, CHIRPS, MSWEP, PISCO, and RAIN4PE) often have precipitation underestimation over the páramo and montane watersheds in the Amazon (Fig. 5). This underestimation, especially by gauge-corrected datasets, could be caused by the low number of precipitation gauges available (Fig. 1), which is further amplified by the fact that the gauges do not account for the important cloud/fog water input into the system (Gomez-Peralta et al. 2008; Clark et al. 2014; Cárdenas et al. 2017).

A substantial precipitation underestimation over the páramo and montane watersheds is critical since it might even lead to physically unrealistic runoff ratios above 1 in water budget estimates as reported in previous studies (Zulkafli et al. 2014; Zubieta et al. 2015; Manz et al. 2016; Strauch et al. 2017; Builes-Jaramillo and Poveda 2018; Aybar et al. 2020). Furthermore, precipitation errors in the upstream catchments can negatively affect simulation results for downstream river catchments. For instance, the assignment of unrealistic model

parameter values to counterbalance precipitation uncertainty can lead to misrepresentation of the basinwide water budget (see more details in appendix B). To overcome these deficiencies, we used streamflow data to adjust precipitation biases.

Our results show that the hydrological correction of precipitation datasets was more efficient over the regions with the strongest rainfall seasonality such as the Peruvian catchments (Espinoza Villar et al. 2009; Segura et al. 2019). This suggests that actual spatiotemporal precipitation fields over these regions are well depicted by the assessed datasets, whereas the correction efficiency over the Ecuadorian Amazon catchments, which experience precipitation throughout the year with high spatial regime variability (Laraque et al. 2007; Tobar and Wyseure 2018), is more variable and depends more strongly on the precipitation product. For instance, the correction was not feasible for CHIRP, ERA5, MSWEP, PISCO, and even CHIRPS (at southern Ecuadorian Amazon) which led to the underrepresentation of the seasonal streamflow patterns and hence the true seasonal precipitation patterns as well (Figs. 4 and 6). This is a critical drawback of these products, and our findings here could be helpful for their revision and improvement. Even though the hydrological correction of CHIRPS resulted in the improvement of the seasonal streamflow simulations for the northern Ecuadorian Amazon, CHIRPS-SWAT still performed unsatisfactorily for the daily and monthly discharge dynamics (Figs. 7, 8 and Figs. S2 and S3), indicating that CHIRPS does not represent well the actual precipitation patterns over the Ecuadorian Amazon catchments. In these catchments, other datasets such as gauge-based ORE HYBAM (Guimberteau et al. 2012), gauge-corrected WFDEI, reanalysis ERA-Interim, and satellite-based PERSIANN (Hsu et al. 1997), TMPA, CMORPH, and IMERG have also been reported to perform unsatisfactorily for streamflow simulation (Zulkafli et al. 2014; Zubieta et al. 2015; Strauch et al. 2017; Zubieta et al. 2017; Towner et al. 2019). Overall, when comparing RAIN4PE to CHIRP, ERA5, MSWEP, PISCO, CHIRPS, and other datasets mentioned above, we can see that it shows satisfactory performance for monthly (Fig. 8) and seasonal (Fig. 4) streamflow simulations with SWAT over the Ecuadorian Amazon. However, its performance for daily simulation is still unsatisfactory (Fig. 7), which highlights that estimation of precipitation at a daily resolution over data-scarce regions such as the equatorial Amazon region is very challenging. The exposed shortcomings of precipitation datasets suggest the urgent implementation and densification of precipitation and cloud/fog gauge networks over the Ecuadorian Amazon and Peruvian montane watersheds. These could help to improve the depiction of rainfall amounts and their spatiotemporal distribution and hence could be useful for improving streamflow simulations. It is important to keep in mind that the correction of the proposed precipitation product through the reverse hydrology concept was performed using the SWAT hydrological model, and therefore the performance of the RAIN4PE dataset may change if another hydrological model is used. Though, as SWAT is a widely used comprehensively verified model, we expect only minor deviation.

c. Implications for hydrological modeling

The results of the hydrological evaluation clearly show the advantages and shortcomings of each evaluated precipitation dataset for streamflow simulation, including low, high, and peak flows. Moreover, we presented the comparison of SWAT-simulated seasonal streamflow using all evaluated datasets against observed seasonal streamflow for the three drainage systems (Titicaca Lake basin, Pacific basin, and Amazon basin) in Figs. S4–S6. These figures can assist practitioners in selecting the appropriate precipitation product for hydrological applications. In general, the hydrological evaluation highlighted RAIN4PE as the best precipitation dataset for hydrological modeling of the Peruvian and Ecuadorian watersheds. RAIN4PE is the only gridded precipitation product for Peru and Ecuador, which benefits from maximum available in situ observations, multiple precipitation sources, environmental variable (elevation data), and is supplemented by streamflow data to correct the precipitation underestimation over páramos and montane catchments. The exploitation of all these variables using state-of-the-practice methods to generate RAIN4PE proved that RAIN4PE-SWAT was capable of closing the (hitherto) observed water budget imbalance over Peruvian and Ecuadorian catchments which, eventually, makes the RAIN4PE a good candidate for hydrological applications in the region. Despite this, we consider that RAIN4PE is still subject to uncertainties, especially in regions where precipitation was inferred from the observed streamflow data. For these regions, precipitation estimates should be viewed with some care due to uncertainties in streamflow data, inferred evapotranspiration, gridded precipitation data, and hydrological model structure.

In this study, besides evaluating precipitation datasets for streamflow simulation, we show that uncertainties associated with precipitation estimates have implications in estimating hydrological model parameters (see appendix B) and water budget components (e.g., evapotranspiration, see appendix C). This is critical for the regionalization of parameters and reliable estimation of the water budget for water resources management. Furthermore, an aftermath verification of RAIN4PE-SWAT-simulated evapotranspiration with GLEAM and MOD16 estimates (appendix C) shows that GLEAM and MOD16 return higher estimated values of evapotranspiration which would not allow the water budget closure and bring inconsistencies in the temporal evapotranspiration distribution over northern Amazon in Ecuador. This suggests that evapotranspiration estimation is still a challenge for remotely sensed based evapotranspiration products in the region.

It is important to highlight that this study is the first applying SWAT updated for improved representation of tropical vegetation dynamics (Alemayehu et al. 2017) and river–floodplain dynamics. These improvements are crucial to model the hydrological processes of Andean and Amazonian river catchments appropriately. The benefits of appropriate representation of tropical vegetation dynamics were demonstrated in previous studies (Strauch and Volk 2013; Alemayehu et al. 2017; Fernandez-Palomino et al. 2020), while the benefit of flow water

routing that considers river–floodplain dynamics can be observed in the good representation of discharge dynamics of the Amazonian rivers in this study. For instance, in the Ucayali River (a tributary of the Amazon River), the significant observed flood peak delay (on a scale of months) from Lagarto to Requena station is well reproduced by SWAT (see Fig. S6), which is consistent with the findings of Santini (2020).

It is also important to highlight that this study is the first applying SWAT at the country-level of Peru and performing a multiobjective calibration and validation using hydrograph goodness of fit and FDC signatures for large-domain modeling (1.6 million km²) in a region with complex hydroclimatic conditions. Our results show the robustness of signature-based calibration guiding the model to reproduce not only one common objective function (e.g., high flows given by NSE) but all aspects of the hydrograph and FDC as supported by RAIN4PE-SWAT good performances reproducing all flow conditions. This is crucial for robust hydrometeorological applications including extremes such as droughts and floods as well as for the assessment of precipitation dataset reliability. Furthermore, our results reinforce previous study findings (Shafii and Tolson 2015; Chilkoti et al. 2018; Fernandez-Palomino et al. 2020), which proved the robustness of a signature-based calibration approach in the hydrological modeling of small watersheds. We consider that our approaches can be helpful for future studies related to precipitation estimates as well as to hydrological model calibration, evaluation, and application.

d. Future development and application

Based on the experiences we gained, our future investigations will focus on applying RAIN4PE-SWAT to analyze the water budget at the national scale of Peru, as well as climate change impacts on water resources using RAIN4PE as the basis for bias adjustment, and trends in frequency and intensity of meteorological and hydrological droughts. The current RAIN4PE data availability (1981–2015) is planned to be extended in the future. Moreover, the methodology presented in the paper will also be extended to the entire Amazon basin.

6. Summary and conclusions

We developed a new hydrologically adjusted daily precipitation dataset (1981–2015, 0.1° resolution) called RAIN4PE by merging three existing datasets for a domain covering Peru and Ecuador. This dataset takes advantages of ground-, satellite-, and reanalysis-based precipitation datasets, including CHIRP and ERA5, which are merged with terrain elevation using the random forest (RF) method to provide precipitation estimates. Furthermore, streamflow data was used to correct precipitation estimates over catchments with water budget closure problems (e.g., the páramo and montane watersheds) through the reverse hydrology methods, for which the SWAT model was applied for the first time herein. Moreover, a comprehensive hydrological evaluation of RAIN4PE, CHIRP, ERA5, and the existing state-of-the-art gauge-corrected precipitation datasets—CHIRPS, MSWEP, and PISCO—in the Peruvian and Ecuadorian

river catchments using a range of performance metrics was performed. For that, SWAT was calibrated and validated with each precipitation dataset in a number of catchments. We summarize our findings as follows.

- The good RAIN4PE-SWAT performance for streamflow simulation suggests the effectiveness of the RF method to merge multisource precipitation estimates with terrain elevation to develop a reliable spatially gridded precipitation dataset. As all datasets (CHIRP, ERA5, and terrain elevation) used to develop RAIN4PE are freely available, this approach can be used in other data-scarce regions.
- The utility of streamflow data to improve both precipitation and streamflow simulations over the páramo and montane watersheds with precipitation underestimation was demonstrated herein. This highlights that the reverse hydrology approach offers a new effective way of understanding the hydrological processes of the Andean–Amazon catchments, which have a key role in the hydrological variability of the entire Amazon basin.
- The hydrological evaluation results from uncorrected precipitation datasets forcing SWAT for streamflow simulation revealed that CHIRP outperformed ERA5, which significantly overestimate precipitation along the Andes. However, these products were outperformed by the gauge-based precipitation datasets.
- Among the gauge-corrected precipitation datasets forcing SWAT for streamflow simulation, all products performed well in the catchments draining into the Titicaca Lake. For catchments draining into the Pacific Ocean and Amazon River, CHIRPS, MSWEP, and PISCO performed unsatisfactorily in several catchments, indicating the limitations of these products for hydrological modeling over these drainage systems. In contrast, RAIN4PE was the only product that provided consistently good performance for the daily and monthly streamflow simulations, including all discharge conditions (low, high, and peak flows) and water budget closure in almost all Peruvian and Ecuadorian river catchments.
- We found that CHIRP, CHIRPS, ERA5, MSWEP, and PISCO cannot represent the seasonal distribution of precipitation and hence the seasonal streamflow over the Ecuadorian Amazon. This is a critical drawback that can have implications in hydrometeorological applications in the Amazon basin.
- We found that uncertainties in precipitation data in existing datasets affect the estimation of model parameters and water budget components, suggesting the importance of developing high-quality meteorological forcing datasets in mountainous regions. Our contribution is in line with this and marks progress in developing precipitation datasets in the region.

The overall good performance of the RAIN4PE highlights its utility as an important new gridded precipitation dataset, which opens new possibilities for numerous hydrometeorological applications throughout Peru and Ecuador. Examples are streamflow simulations, estimation of the water budget and its evolution, water resources management, understanding

spatiotemporal variations of droughts and floods, and exploring spatial variations and regimes of precipitation. We consider that RAIN4PE and our RAIN4PE-SWAT model can be adopted as a benchmark to evaluate precipitation datasets in Peru and Ecuador.

Acknowledgments. The authors thank the EPICC project that is part of the International Climate Initiative (IKI). The Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) supports this initiative on the basis of a decision adopted by the German Bundestag. We also thank SENAMHI, Peruvian ANA, observatory HYBAM, Enrique Morán-Tejeda, Guido G. Tamayo, Juan J. Nieto, Vladimiro Tobar, and Kevin J. Perez for providing the hydrometeorological dataset. The authors are thankful to CHIRP, CHIRPS, ERA5, MSWEP, and PISCO data generation teams for providing the precipitation data at free of cost. We are thankful to Dr. David Hadka and Dr. Patrick M. Reed for making their software “BORG: Many-Objective Evolutionary Computing Framework” available for this study. We are grateful to the Editor, Liz Stephens, Oscar M. Baez-Villanueva, and one anonymous reviewer for their constructive comments.

Data availability statement. The RAIN4PE data record is freely available at <https://doi.org/10.5880/pik.2020.010> (Fernandez-Palomino et al. 2021).

APPENDIX A

Glossary

CHIRP	Climate Hazards Group Infrared Precipitation
CHIRPS	CHIRP with Station data
CMORPH	Climate Prediction Center morphing technique
IMERG	Global Precipitation Measurement (GPM) Integrated Multisatellite Retrievals
MSWEP	Multi-Source Weighted-Ensemble Precipitation
PISCO	Peruvian Interpolated data of SENAMHI's Climatological and Hydrological Observations
SENAMHI	Servicio Nacional de Meteorología e Hidrología del Perú
TMPA	Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis
WFDEI	WATCH Forcing Data methodology applied to ERA-Interim data

APPENDIX B

Evaluating the Distribution of Model Parameters

In this section, we analyze the distribution of calibrated model parameters to see the regional parameter behavior and to elucidate potential input errors as they were identified to achieve the water budget closure using different precipitation datasets. Thus, unrealistic parameter values could be linked to input error. We advise readers to see Table 3 for the description of parameters and Neitsch et al. (2011)

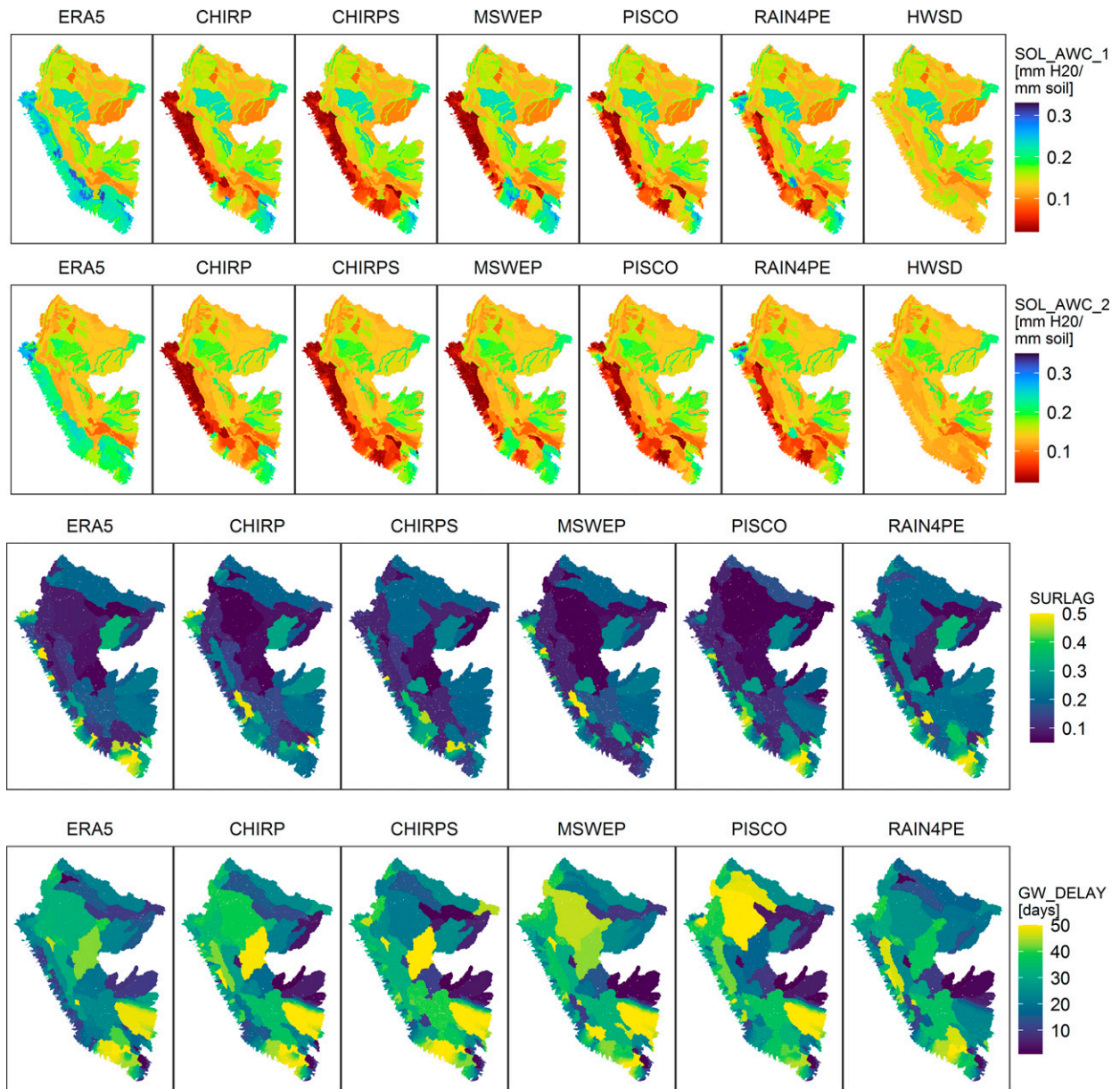


FIG. B1. Calibrated parameter values for the soil available water capacity (SOL_AWC) for topsoil (1) and subsoil (2), the surface runoff delay coefficient (SURLAG), and the groundwater delay time (GW_DELAY). The HWSD map shows SOL_AWC values derived from the Harmonized World Soil Database, which were used for setting up the SWAT model.

for detailed parameter definitions. Among the calibrated SWAT parameters, only two (SOL_AWC, GW_REVAP) can alter the water budget since they influence evapotranspiration and, subsequently, runoff estimation. The remaining parameters influence the surface runoff (SURLAG), groundwater (GW_DELAY, RCHRG_DP, GWQMN, ALPHA_BF), and flow routing (CH_K2, CHD, FP_W_F) not affecting water loss from the system. We illustrate in Figs. B1–B3 the spatial patterns of the calibrated parameters related to six precipitation datasets.

Figure B1 shows that the SOL_AWC, which constrains the maximum amount of plant available water a soil can provide and was derived from the Harmonized World Soil Database (HWSD; Abbaspour et al. 2019), was adjusted mainly for the Andean catchments. This is a critical parameter since higher values can lead to higher evapotranspiration and vice versa. The results show that high SOL_AWC values were identified for compensating the ERA5 precipitation overestimation (see positive errors in Fig. 6). However, despite this trade-off, discharge overestimation by ERA5-SWAT remains

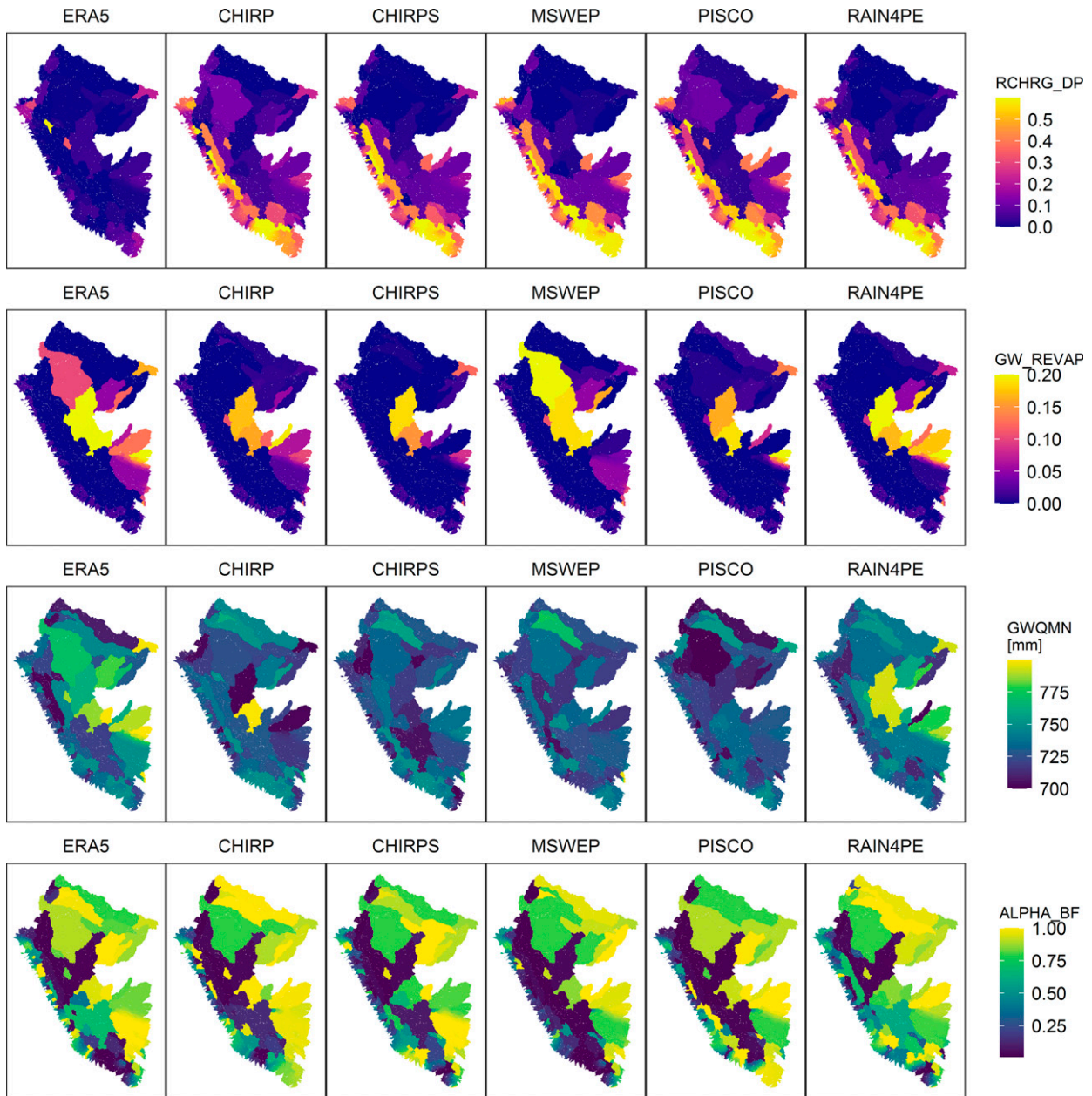


FIG. B2. Calibrated parameter values for the deep aquifer percolation fraction (RCHRG_DP), the threshold for return flow from the shallow aquifer (GWQMN), the groundwater “revap” coefficient (GW_REVAP), and the baseflow recession constant (ALPHA_BF).

(see PBIAS in Fig. 8), suggesting that ERA5 precipitation estimates must be bias-corrected for the Andean regions prior to hydrological applications. Otherwise, unrealistic low SOL_AWC values (≈ 0) and the prevalence of discharge underestimation (Fig. 8 and Fig. S5) over the northern pacific coastal catchments suggest that precipitation could be underestimated there, particularly by CHIRP, CHIRPS, MSWEP, and even PISCO (see negative errors in Fig. 6).

For the remaining parameters, we describe each one briefly based on the calibrated parameters for RAIN4PE-SWAT.

Figure B1 shows overall low values (ranging from 0.05 to 0.5) for SURLAG in the study area, which is important for smoothing the simulated hydrograph due to the delay in surface runoff released from the HRUs (Neitsch et al. 2011) to match the peaks in the observed hydrograph. The GW_DELAY values (ranging from 1 to 50 days) reflect the lag in time that water in soil profile needs to enter shallow aquifer; high (low) values are usual for most of the Andean (Amazonian) catchments.

Figure B2 shows spatial distribution of the calibrated groundwater-related parameters. The RCHRG_DP parameter

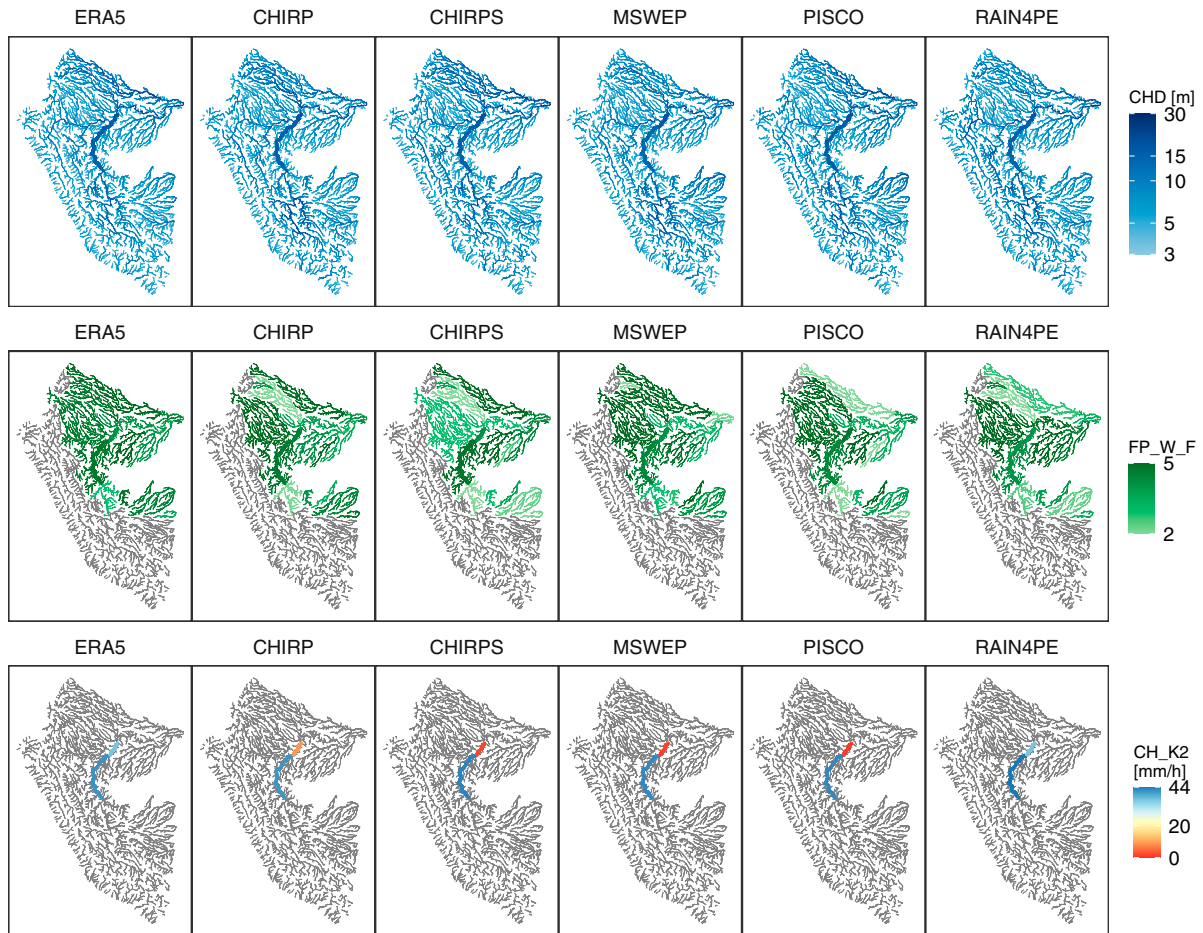


FIG. B3. Calibrated parameter values for the main channel depth (CHD), the ratio of floodplain width over bankfull width (FP_W_F), and the hydraulic conductivity of main channel (CH_K2). Reaches in gray indicate that the parameter was not important in these reaches.

reflects the water volume percolated into the deep aquifer relative to the total recharge entering aquifers (both shallow and deep). Therefore, the calibrated RCHRG_DP values provide an insight into the important recharge entering deep aquifers in Peruvian Andean catchments, which subsequently sustain the prolonged dry season flow in these catchments (Clark et al. 2014; Fernandez-Palomino et al. 2020). The GW_REVAP values greater than zero reflect the areas (lower Amazon) where water is re-evaporated from the shallow aquifer (water entering the soil for evaporation and transpiration). In these areas, deep-rooted evergreen forests can draw water from the shallow aquifer to meet their demands if available water in the soil profile is insufficient. All calibrated GWQMN values favor the return flow from aquifers and the re-evaporation from the shallow aquifer in areas (lower Amazon) where GWQMN values are greater than 750 mm (default water depth threshold in the shallow aquifer to allow re-evaporation). The high ALPHA_BF values (~1) show shallow aquifers quickly contributing return flow to streams

(e.g., lower Amazon catchments), whereas the low values (~0) show those with slow contributions (e.g., most of the Andean catchments draining into the Pacific Ocean).

Figure B3 shows the calibrated reach and floodplain parameters (CHD, FP_W_F, and CH_K2). Among these parameters, the FP_W_F values can reflect the occurrence of flow over floodplains during the high discharge season in the lower Amazon rivers. The CH_K2 values greater than zero show reaches where water is infiltrated at the floodplain surface from floodplain flow or ponded water during overbank flood events. Then water stored at floodplain alluvium flows back to the channel when flood wave has passed and water levels in the channel have dropped, and the hydraulic gradient is reversed. This interaction between floodplains and reaches can explain the significant observed flood peak delay (on a scale of months) from Lagarto to Requena station (see Fig. S6), which is consistent with the findings of Santini (2020).

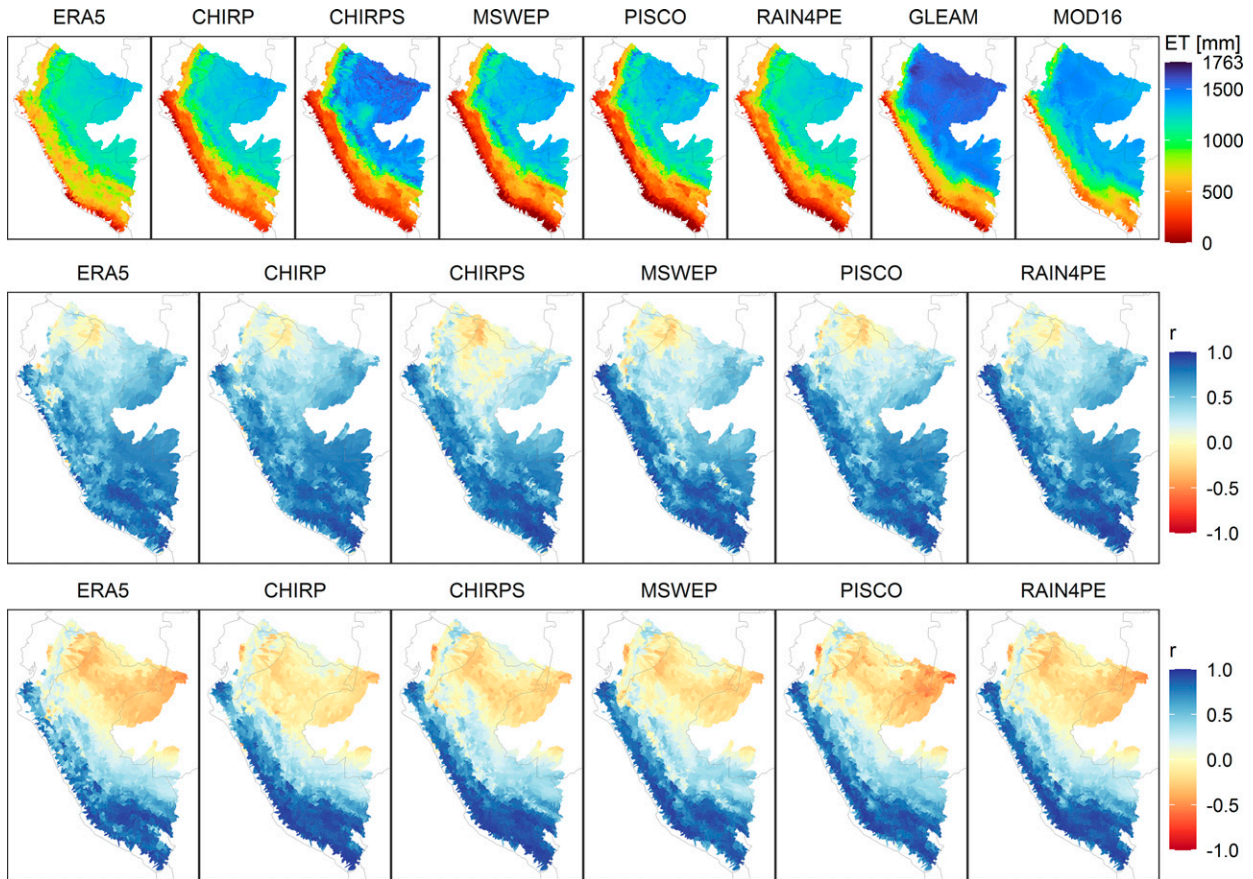


FIG. C1. Comparison of evapotranspiration (ET) estimates from the calibrated SWAT model using different precipitation datasets as input with the remotely sensed based ET estimates from GLEAM and MOD16: (top) average annual ET for the period 2000–14 and Pearson's correlation coefficient (r) between SWAT-simulated ET and ET estimates from (middle) GLEAM and (bottom) MOD16. The comparison measure (r) was computed using monthly ET time series at the subcatchment scale for 2000–14.

APPENDIX C

Comparison of SWAT and Remotely Sensed Based Evapotranspiration

Figure C1 compares evapotranspiration (ET) estimates from the calibrated SWAT model driven by different precipitation datasets with the GLEAM and MOD16 estimates. All ET estimates show similar spatial patterns with increasing ET gradients from west to east. The differences in SWAT-simulated ET volumes can be attributed to inappropriate parameter estimation due to precipitation biases and uncertainties. The higher ET values for CHIRPS-SWAT are likely due to the prevalence of dry conditions in CHIRPS which affect the estimation of daily relative humidity and subsequently potential evapotranspiration, vapor stress on plant growth, and ET.

We compared GLEAM and MOD16 against the simulated ET by RAIN4PE-SWAT since it represents the water budget well (see PBIAS in Fig. 8). Figure C1 shows a general tendency for GLEAM and MOD16 to overestimate ET in the study area, and even their estimates are greater than precipitation along the Andes (see Fig. C1 and Fig. 6), which would not allow the water budget closure in the

Andean catchments. The correlation coefficient (Fig. C1, middle panel) shows better agreement between GLEAM and SWAT-simulated ET, which are based on the same equation (Priestley–Taylor) for potential evapotranspiration estimation. Spatially, both GLEAM and MOD16 agree well with the SWAT-simulated ET in areas with strong seasonal precipitation variability, as the Peruvian Andes and southern region of the Peruvian Amazon. However, negative correlation values over the northern Amazon basin areas with a bimodal rainfall regime (Laraque et al. 2007) can indicate inconsistency in the temporal distribution in GLEAM and MOD16 ET estimates there. This is in line with the findings of Dile et al. (2020), who reported that remotely sensed based ET did not respond well to the rainfall in areas with a bimodal rainfall pattern in Ethiopia. Furthermore, the ET estimates could be affected by the inherent uncertainties in methods and input data. Our results demonstrate that the ET estimation by the remotely sensed ET products is still a challenge in the region, and the ground-based measurements are required for better understanding the ET spatio-temporal patterns and for a more reliable evaluation of the ET estimates.

REFERENCES

- Abbaspour, K. C., and S. A. Ashraf Vaghefi, 2019: Harmonized World Soil Database in SWAT Format. PANGAEA, accessed 30 March 2020, <https://doi.org/10.1594/PANGAEA.901309>.
- , S. A. Vaghefi, H. Yang, and R. Srinivasan, 2019: Global soil, landuse, evapotranspiration, historical and future weather databases for SWAT Applications. *Sci. Data*, **6**, 263, <https://doi.org/10.1038/s41597-019-0282-4>.
- Alemayehu, T., A. van Griensven, B. T. Woldegiorgis, and W. Bauwens, 2017: An improved SWAT vegetation growth module and its evaluation for four tropical ecosystems. *Hydrol. Earth Syst. Sci.*, **21**, 4449–4467, <https://doi.org/10.5194/hess-21-4449-2017>.
- Alfieri, L., V. Lorini, F. A. Hirpa, S. Harrigan, E. Zsoter, C. Prudhomme, and P. Salamon, 2020: A global streamflow reanalysis for 1980–2018. *J. Hydrol. X*, **6**, 100049, <https://doi.org/10.1016/j.hydroa.2019.100049>.
- Armijos, E., and Coauthors, 2013: Suspended sediment dynamics in the Amazon River of Peru. *J. S. Amer. Earth Sci.*, **44**, 75–84, <https://doi.org/10.1016/j.jsames.2012.09.002>.
- Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams, 1998: Large area hydrologic modeling and assessment part I: Model development. *J. Amer. Water Resour. Assoc.*, **34**, 73–89, <https://doi.org/10.1111/j.1752-1688.1998.tb05961.x>.
- Asurza-Véliz, F. A., and W. S. Lavado-Casimiro, 2020: Regional parameter estimation of the SWAT model: Methodology and application to river basins in the Peruvian Pacific drainage. *Water*, **12**, 3198, <https://doi.org/10.3390/w12113198>.
- Aybar, C., C. Fernández, A. Huerta, W. Lavado, F. Vega, and O. Felipe-Obando, 2020: Construction of a high-resolution gridded rainfall dataset for Peru from 1981 to the present day. *Hydrol. Sci. J.*, **65**, 770–785, <https://doi.org/10.1080/02626667.2019.1649411>.
- Baez-Villanueva, O. M., M. Zambrano-Bigiarini, L. Ribbe, A. Nauditt, J. D. Giraldo-Osorio, and N. X. Thinh, 2018: Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. *Atmos. Res.*, **213**, 34–50, <https://doi.org/10.1016/j.atmosres.2018.05.011>.
- , and Coauthors, 2020: RF-MEP: A novel random Forest method for merging gridded precipitation products and ground-based measurements. *Remote Sens. Environ.*, **239**, 111606, <https://doi.org/10.1016/j.rse.2019.111606>.
- Bai, P., and X. Liu, 2018: Evaluation of five satellite-based precipitation products in two gauge-scarce basins on the Tibetan Plateau. *Remote Sens.*, **10**, 1316, <https://doi.org/10.3390/rs10081316>.
- Beck, H. E., A. I. J. M. van Dijk, V. Levizzani, J. Schellekens, D. G. Miralles, B. Martens, and A. de Roo, 2017: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.*, **21**, 589–615, <https://doi.org/10.5194/hess-21-589-2017>.
- , and Coauthors, 2019a: Daily evaluation of 26 precipitation datasets using stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci.*, **23**, 207–224, <https://doi.org/10.5194/hess-23-207-2019>.
- , E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. van Dijk, T. R. McVicar, and R. F. Adler, 2019b: MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bull. Amer. Meteor. Soc.*, **100**, 473–500, <https://doi.org/10.1175/BAMS-D-17-0138.1>.
- , and Coauthors, 2020a: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Satellite Precipitation Measurement*, Vol. 2, V. Levizzani et al., Eds., Advances in Global Change Research, Vol. 69, Springer, 625–653, https://doi.org/10.1007/978-3-030-35798-6_9.
- , E. F. Wood, T. R. McVicar, M. Zambrano-Bigiarini, C. Alvarez-Garreton, O. M. Baez-Villanueva, J. Sheffield, and D. N. Karger, 2020b: Bias correction of global high-resolution precipitation climatologies using streamflow observations from 9372 catchments. *J. Climate*, **33**, 1299–1315, <https://doi.org/10.1175/JCLI-D-19-0332.1>.
- Bhuiyan, M. A. E., E. I. Nikolopoulos, and E. N. Anagnostou, 2019: Machine learning–based blending of satellite and reanalysis precipitation datasets: A multiregional tropical complex terrain evaluation. *J. Hydrometeorol.*, **20**, 2147–2161, <https://doi.org/10.1175/JHM-D-19-0073.1>.
- Breiman, L., 2001: Random forests. *Mach. Learn.*, **45**, 5–32, <https://doi.org/10.1023/A:1010933404324>.
- Brocca, L., T. Moramarco, F. Melone, and W. Wagner, 2013: A new method for rainfall estimation through soil moisture observations. *Geophys. Res. Lett.*, **40**, 853–858, <https://doi.org/10.1002/grl.50173>.
- , and Coauthors, 2019: SM2RAIN–ASCAT (2007–2018): Global daily satellite rainfall data from ASCAT soil moisture observations. *Earth Syst. Sci. Data*, **11**, 1583–1601, <https://doi.org/10.5194/essd-11-1583-2019>.
- , C. Massari, T. Pellarin, P. Filippucci, L. Ciabatta, S. Camici, Y. H. Kerr, and D. Fernández-Prieto, 2020: River flow prediction in data scarce regions: Soil moisture integrated satellite rainfall products outperform rain gauge observations in West Africa. *Sci. Rep.*, **10**, 12517, <https://doi.org/10.1038/s41598-020-69343-x>.
- Bruijnzeel, L. A., M. Mulligan, and F. N. Scatena, 2011: Hydrometeorology of tropical montane cloud forests: Emerging patterns. *Hydrol. Processes*, **25**, 465–498, <https://doi.org/10.1002/hyp.7974>.
- Buchhorn, M., B. Smets, L. Bertels, M. Lesiv, N.-E. Tsendbazar, M. Herold, and S. Fritz, 2019: Copernicus global land service: Land cover 100 m: Epoch 2015: Globe (V2.0.2). Zenodo, accessed 30 March 2020, <https://doi.org/10.5281/zenodo.3243509>.
- Builes-Jaramillo, A., and G. Poveda, 2018: Conjoint analysis of surface and atmospheric water balances in the Andes-Amazon system. *Water Resour. Res.*, **54**, 3472–3489, <https://doi.org/10.1029/2017WR021338>.
- Cárdenas, M. F., C. Tobón, and W. Buytaert, 2017: Contribution of occult precipitation to the water balance of páramo ecosystems in the Colombian Andes. *Hydrol. Processes*, **31**, 4440–4449, <https://doi.org/10.1002/hyp.11374>.
- Chavez, S. P., and K. Takahashi, 2017: Orographic rainfall hot spots in the Andes-Amazon transition according to the TRMM precipitation radar and in situ data. *J. Geophys. Res. Atmos.*, **122**, 5870–5882, <https://doi.org/10.1002/2016JD026282>.
- Chilkoti, V., T. Bolisetti, and R. Balachandar, 2018: Multi-objective autocalibration of SWAT model for improved low flow performance for a small snowed catchment. *Hydrol. Sci. J.*, **63**, 1482–1501, <https://doi.org/10.1080/02626667.2018.1505047>.
- Clark, K. E., and Coauthors, 2014: The hydrological regime of a forested tropical Andean catchment. *Hydrol. Earth Syst. Sci.*, **18**, 5377–5397, <https://doi.org/10.5194/hess-18-5377-2014>.
- da Motta Paca, V. H., G. E. Espinoza-Dávalos, D. M. Moreira, and G. Comair, 2020: Variability of trends in precipitation across the Amazon River Basin determined from the CHIRPS precipitation product and from station records. *Water*, **12**, 1244, <https://doi.org/10.3390/w12051244>.

- da Silva Júnior, J. C., V. Medeiros, C. Garrozi, A. Montenegro, and G. E. Gonçalves, 2019: Random forest techniques for spatial interpolation of evapotranspiration data from Brazilian's Northeast. *Comput. Electron. Agric.*, **166**, 105017, <https://doi.org/10.1016/j.compag.2019.105017>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Derin, Y., and Coauthors, 2019: Evaluation of GPM-era global satellite precipitation products over multiple complex terrain regions. *Remote Sens.*, **11**, 2936, <https://doi.org/10.3390/rs11242936>.
- Dile, Y. T., E. K. Ayana, A. W. Worqlul, H. Xie, R. Srinivasan, N. Lefore, L. You, and N. Clarke, 2020: Evaluating satellite-based evapotranspiration estimates for hydrological applications in data-scarce regions: A case in Ethiopia. *Sci. Total Environ.*, **743**, 140702, <https://doi.org/10.1016/j.scitotenv.2020.140702>.
- Erazo, B., L. Bourrel, F. Frappart, O. Chimborazo, D. Labat, L. Dominguez-Granda, D. Matamoros, and R. Mejia, 2018: Validation of satellite estimates (Tropical Rainfall Measuring Mission, TRMM) for rainfall variability over the Pacific slope and coast of Ecuador. *Water*, **10**, 213, <https://doi.org/10.3390/w10020213>.
- Espinoza, J. C., J. Ronchail, J. A. Marengo, and H. Segura, 2019: Contrasting North–South changes in Amazon wet-day and dry-day frequency and related atmospheric features (1981–2017). *Climate Dyn.*, **52**, 5413–5430, <https://doi.org/10.1007/s00382-018-4462-2>.
- , R. Garreaud, G. Poveda, P. A. Arias, J. Molina-Carpio, M. Masiokas, M. Viale, and L. Scaff, 2020: Hydroclimate of the Andes. Part I: Main climatic features. *Front. Earth Sci.*, **8**, 64, <https://doi.org/10.3389/feart.2020.00064>.
- Espinoza Villar, J. C., and Coauthors, 2009: Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador). *Int. J. Climatol.*, **29**, 1574–1594, <https://doi.org/10.1002/joc.1791>.
- Fallah, A., G. R. Rakhshandehroo, P. Berg, O. Sungmin, and R. Orth, 2020: Evaluation of precipitation datasets against local observations in southwestern Iran. *Int. J. Climatol.*, **40**, 4102–4116, <https://doi.org/10.1002/joc.6445>.
- Fan, Y., H. Li, and G. Miguez-Macho, 2013: Global patterns of groundwater table depth. *Science*, **339**, 940–943, <https://doi.org/10.1126/science.1229881>.
- Fernandez-Palomino, C. A., F. F. Hattermann, V. Krysanova, F. Vega-Jácome, and A. Bronstert, 2020: Towards a more consistent eco-hydrological modelling through multi-objective calibration: a case study in the Andean Vilcanota River basin, Peru. *Hydrol. Sci. J.*, **66**, 59–74, <https://doi.org/10.1080/02626667.2020.1846740>.
- , and Coauthors, 2021: Rain for Peru and Ecuador (RAIN4PE). V. 1.0. GFZ Data Services, accessed 14 January 2022, <https://doi.org/10.5880/pik.2020.010>.
- Fox, E. W., J. M. Ver Hoef, and A. R. Olsen, 2020: Comparing spatial regression to random forests for large environmental data sets. *PLOS ONE*, **15**, e0229509, <https://doi.org/10.1371/journal.pone.0229509>.
- Funk, C., and Coauthors, 2015a: The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Sci. Data*, **2**, 150066, <https://doi.org/10.1038/sdata.2015.66>.
- , A. Verdin, J. Michaelsen, P. Peterson, D. Pedreros, and G. Husak, 2015b: A global satellite-assisted precipitation climatology. *Earth Syst. Sci. Data*, **7**, 275–287, <https://doi.org/10.5194/essd-7-275-2015>.
- Gassman, P. W., A. M. Sadeghi, and R. Srinivasan, 2014: Applications of the SWAT model special section: Overview and insights. *J. Environ. Qual.*, **43**, 1–8, <https://doi.org/10.2134/jeq2013.11.0466>.
- Gleixner, S., T. Demissie, and G. T. Diro, 2020: Did ERA5 improve temperature and precipitation reanalysis over East Africa? *Atmosphere*, **11**, 996, <https://doi.org/10.3390/atmos11090996>.
- Gomez-Peralta, D., S. F. Oberbauer, M. E. McClain, and T. E. Philippi, 2008: Rainfall and cloud-water interception in tropical montane forests in the eastern Andes of Central Peru. *For. Ecol. Manage.*, **255**, 1315–1325, <https://doi.org/10.1016/j.foreco.2007.10.058>.
- Guimberteau, M., and Coauthors, 2012: Discharge simulation in the sub-basins of the Amazon using ORCHIDEE forced by new datasets. *Hydrol. Earth Syst. Sci.*, **16**, 911–935, <https://doi.org/10.5194/hess-16-911-2012>.
- Guo, Y., Y. Zhang, L. Zhang, and Z. Wang, 2021: Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. *Wiley Interdiscip. Rev. Water*, **8**, <https://doi.org/10.1002/wat2.1487>.
- Gupta, H. V., S. Sorooshian, and P. O. Yapo, 1999: Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *J. Hydrol. Eng.*, **4**, 135–143, [https://doi.org/10.1061/\(ASCE\)1084-0699\(1999\)4:2\(135\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135)).
- , H. Kling, K. K. Yilmaz, and G. F. Martinez, 2009: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.*, **377**, 80–91, <https://doi.org/10.1016/j.jhydrol.2009.08.003>.
- Hadka, D., and P. Reed, 2013: Borg: An auto-adaptive many-objective evolutionary computing framework. *Evol. Comput.*, **21**, 231–259, https://doi.org/10.1162/EVCO_a.00075.
- He, J., K. Yang, W. Tang, H. Lu, J. Qin, Y. Chen, and X. Li, 2020: The first high-resolution meteorological forcing dataset for land process studies over China. *Sci. Data*, **7**, 25, <https://doi.org/10.1038/s41597-020-0369-y>.
- Helmer, E. H., E. A. Gerson, L. S. Baggett, B. J. Bird, T. S. Ruzicky, and S. M. Voggeser, 2019: Neotropical cloud forests and páramo to contract and dry from declines in cloud immersion and frost. *PLOS ONE*, **14**, e0213155, <https://doi.org/10.1371/journal.pone.0213155>.
- Hengl, T., M. Nussbaum, M. N. Wright, G. B. M. Heuvelink, and B. Gräler, 2018: Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, **6**, e5518, <https://doi.org/10.7717/peerj.5518>.
- Henn, B., M. P. Clark, D. Kavetski, and J. D. Lundquist, 2015: Estimating mountain basin-mean precipitation from streamflow using Bayesian inference. *Water Resour. Res.*, **51**, 8012–8033, <https://doi.org/10.1002/2014WR016736>.
- , —, —, A. J. Newman, M. Hughes, B. McGurk, and J. D. Lundquist, 2018: Spatiotemporal patterns of precipitation inferred from streamflow observations across the Sierra Nevada mountain range. *J. Hydrol.*, **556**, 993–1012, <https://doi.org/10.1016/j.jhydrol.2016.08.009>.
- Herrnegger, M., H. P. Nachtnebel, and K. Schulz, 2015: From runoff to rainfall: Inverse rainfall-runoff modelling in a high temporal resolution. *Hydrol. Earth Syst. Sci.*, **19**, 4619–4639, <https://doi.org/10.5194/hess-19-4619-2015>.

- Hersbach, H., and coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Hong, Z., Z. Han, X. Li, D. Long, G. Tang, and J. Wang, 2021: Generation of an improved precipitation data set from multi-source information over the Tibetan Plateau. *J. Hydrometeorol.*, **22**, 1275–1295, <https://doi.org/10.1175/JHM-D-20-0252.1>.
- Hrachowitz, M., and Coauthors, 2014: Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water Resour. Res.*, **50**, 7445–7469, <https://doi.org/10.1002/2014WR015484>.
- Hsu, K., X. Gao, S. Sorooshian, H. V. Gupta, K. Hsu, X. Gao, S. Sorooshian, and H. V. Gupta, 1997: Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteor.*, **36**, 1176–1190, [https://doi.org/10.1175/1520-0450\(1997\)036<1176:PEFRSI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2).
- Huerta, A., C. Aybar, and W. Lavado-Casimiro, 2018: PISCO temperatura v.1.1. SENAMHI -DHI-2018, 16 pp.
- Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- , and Coauthors, 2019: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 06, 32 pp., https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf.
- Hunziker, S., and Coauthors, 2017: Identifying, attributing, and overcoming common data quality issues of manned station observations. *Int. J. Climatol.*, **37**, 4131–4145, <https://doi.org/10.1002/joc.5037>.
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, P. Xie, R. J. Joyce, J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, **5**, 487–503, [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2).
- Kirchner, J. W., 2009: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resour. Res.*, **45**, W02429, <https://doi.org/10.1029/2008WR006912>.
- Kling, H., M. Fuchs, and M. Paulin, 2012: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.*, **424–425**, 264–277, <https://doi.org/10.1016/j.jhydrol.2012.01.011>.
- Kneis, D., C. Chatterjee, and R. Singh, 2014: Evaluation of TRMM rainfall estimates over a large Indian River Basin (Mahanadi). *Hydrol. Earth Syst. Sci.*, **18**, 2493–2502, <https://doi.org/10.5194/hess-18-2493-2014>.
- Kobayashi, S., and Coauthors, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *Kisho Shushi Dai2shu*, **93**, 5–48, <https://doi.org/10.2151/jmsj.2015.001>.
- Krause, P., D. P. Boyle, and F. Bäse, 2005: Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.*, **5**, 89–97, <https://doi.org/10.5194/adgeo-5-89-2005>.
- Krier, R., P. Matgen, K. Goergen, L. Pfister, L. Hoffmann, J. W. Kirchner, S. Uhlenbrook, and H. H. G. Savenije, 2012: Inferring catchment precipitation by doing hydrology backward: A test in 24 small and mesoscale catchments in Luxembourg. *Water Resour. Res.*, **48**, W10525, <https://doi.org/10.1029/2011WR010657>.
- Laraque, A., J. Ronchail, G. Cochonneau, R. Pombosa, and J. L. Guyot, 2007: Heterogeneous distribution of rainfall and discharge regimes in the Ecuadorian Amazon basin. *J. Hydrometeorol.*, **8**, 1364–1381, <https://doi.org/10.1175/2007JHM784.1>.
- Le Moine, N., F. Hendrickx, J. Gailhard, R. Garçon, and F. Gottardi, 2015: Hydrologically aided interpolation of daily precipitation and temperature fields in a mesoscale Alpine catchment. *J. Hydrometeorol.*, **16**, 2595–2618, <https://doi.org/10.1175/JHM-D-14-0162.1>.
- Liaw, A., and M. Wiener, 2002: Classification and regression by randomForest. *R News*, **2**, 18–22.
- Llauca, H., W. Lavado-Casimiro, C. Montesinos, W. Santini, and P. Rau, 2021: PISCO_HyM_GR2M: A model of monthly water balance in Peru (1981–2020). *Water*, **13**, 1048, <https://doi.org/10.3390/w13081048>.
- Mantas, V. M., Z. Liu, C. Caro, and A. J. S. C. Pereira, 2014: Validation of TRMM Multi-satellite Precipitation Analysis (TMPA) products in the Peruvian Andes. *Atmos. Res.*, **163**, 132–145, <https://doi.org/10.1016/j.atmosres.2014.11.012>.
- Manz, B., W. Buytaert, Z. Zulkafli, W. Lavado, B. Willems, L. A. Robles, and J.-P. Rodríguez-Sánchez, 2016: High-resolution satellite-gauge merged precipitation climatologies of the tropical Andes. *J. Geophys. Res. Atmos.*, **121**, 1190–1207, <https://doi.org/10.1002/2015JD023788>.
- , and Coauthors, 2017: Comparative ground validation of IMERG and TMPA at variable spatiotemporal scales in the tropical Andes. *J. Hydrometeorol.*, **18**, 2469–2489, <https://doi.org/10.1175/JHM-D-16-0277.1>.
- Martens, B., and Coauthors, 2017: GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geosci. Model Dev.*, **10**, 1903–1925, <https://doi.org/10.5194/gmd-10-1903-2017>.
- Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman, 2011: Global land-surface evaporation estimated from satellite-based observations. *Hydrol. Earth Syst. Sci.*, **15**, 453–469, <https://doi.org/10.5194/hess-15-453-2011>.
- Morán-Tejeda, E., and Coauthors, 2016: Climate trends and variability in Ecuador (1966–2011). *Int. J. Climatol.*, **36**, 3839–3855, <https://doi.org/10.1002/joc.4597>.
- Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Binger, R. D. Harmel, and T. L. Veith, 2007: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE*, **50**, 885–900, <https://doi.org/10.13031/2013.23153>.
- Mu, Q., M. Zhao, and S. W. Running, 2011: Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.*, **115**, 1781–1800, <https://doi.org/10.1016/j.rse.2011.02.019>.
- Nash, J. E., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models part I —A discussion of principles. *J. Hydrol.*, **10**, 282–290, [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams, 2011: Soil & Water Assessment Tool Theoretical Documentation Version 2009. Texas Water Resource Institute, 647 pp., <https://swat.tamu.edu/media/99192/swat2009-theory.pdf>.
- Ochoa, A., L. Pineda, P. Crespo, and P. Willems, 2014: Evaluation of TRMM 3B42 precipitation estimates and WRF retrospective precipitation simulation over the Pacific–Andean region of Ecuador and Peru. *Hydrol. Earth Syst. Sci.*, **18**, 3179–3193, <https://doi.org/10.5194/hess-18-3179-2014>.
- Paccini, L., J. C. Espinoza, J. Ronchail, and H. Segura, 2018: Intra-seasonal rainfall variability in the Amazon basin related

- to large-scale circulation patterns: A focus on western Amazon-Andes transition region. *Int. J. Climatol.*, **38**, 2386–2399, <https://doi.org/10.1002/joc.5341>.
- Paiva, R. C. D., W. Collischonn, and C. E. M. Tucci, 2011: Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach. *J. Hydrol.*, **406**, 170–181, <https://doi.org/10.1016/j.jhydrol.2011.06.007>.
- Pelletier, J. D., and Coauthors, 2016: A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. *J. Adv. Model. Earth Syst.*, **8**, 41–65, <https://doi.org/10.1002/2015MS000526>.
- Pfannerstill, M., B. Guse, and N. Fohrer, 2014: Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. *J. Hydrol.*, **510**, 447–458, <https://doi.org/10.1016/j.jhydrol.2013.12.044>.
- , K. Bieger, B. Guse, D. D. Bosch, N. Fohrer, and J. G. Arnold, 2017: How to constrain multi-objective calibrations of the SWAT model using water balance components. *J. Amer. Water Resour. Assoc.*, **53**, 532–546, <https://doi.org/10.1111/1752-1688.12524>.
- Pokhrel, P., and K. K. Yilmaz, 2012: Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. *J. Hydrol.*, **418–419**, 49–60, <https://doi.org/10.1016/j.jhydrol.2008.12.004>.
- Poveda, G., J. C. Espinoza, M. D. Zuluaga, S. A. Solman, R. Garreaud, and P. J. van Oevelen, 2020: High impact weather events in the Andes. *Front. Earth Sci.*, **8**, 162, <https://doi.org/10.3389/feart.2020.00162>.
- Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. P. Mahanama, and G. S. Partyka, 2017: Land surface precipitation in MERRA-2. *J. Climate*, **30**, 1643–1664, <https://doi.org/10.1175/JCLI-D-16-0570.1>.
- Román-Cascón, C., and Coauthors, 2017: Correcting satellite-based precipitation products through SMOS soil moisture data assimilation in two land-surface models of different complexity: API and SURFEX. *Remote Sens. Environ.*, **200**, 295–310, <https://doi.org/10.1016/j.rse.2017.08.022>.
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, <https://doi.org/10.1175/2010BAMS3001.1>.
- Sahraei, S., M. Asadzadeh, and F. Unduche, 2020: Signature-based multi-modelling and multi-objective calibration of hydrologic models: Application in flood forecasting for Canadian Prairies. *J. Hydrol.*, **588**, 125095, <https://doi.org/10.1016/j.jhydrol.2020.125095>.
- Santini, W., 2020: Caractérisation de la dynamique hydro-sédimentaire du bassin de l'Ucayali (Pérou), par une approche intégrant réseau de mesures, télédétection et modélisation hydrologique. Université Toulouse III - Paul Sabatier, 479 pp., <https://hal.archives-ouvertes.fr/tel-03276962/>.
- , and Coauthors, 2015: Sediment budget in the Ucayali River basin, an Andean tributary of the Amazon River. *Proc. Int. Assoc. Hydrol. Sci.*, **367**, 320–325, <https://doi.org/10.5194/piahs-367-320-2015>.
- , and Coauthors, 2019: An index concentration method for suspended load monitoring in large rivers of the Amazonian foreland. *Earth Surf. Dyn.*, **7**, 515–536, <https://doi.org/10.5194/esurf-7-515-2019>.
- Satgé, F., M.-P. Bonnet, M. Gosset, J. Molina, W. Hernan Yuque Lima, R. Pillco Zolá, F. Timouk, and J. Garnier, 2016: Assessment of satellite rainfall products over the Andean plateau. *Atmos. Res.*, **167**, 1–14, <https://doi.org/10.1016/j.atmosres.2015.07.012>.
- , D. Ruelland, M.-P. Bonnet, J. Molina, and R. Pillco, 2019: Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region. *Hydrol. Earth Syst. Sci.*, **23**, 595–619, <https://doi.org/10.5194/hess-23-595-2019>.
- , Y. Hussain, J. Molina-Carpio, R. Pillco, C. Laugner, G. Akhter, and M. Bonnet, 2020: Reliability of SM2RAIN precipitation datasets in comparison to gauge observations and hydrological modelling over arid regions. *Int. J. Climatol.*, **41**, E517–E536, <https://doi.org/10.1002/joc.6704>.
- Scheel, M. L. M., M. Rohrer, C. Huggel, D. Santos Villar, E. Silvestre, and G. J. Huffman, 2011: Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrol. Earth Syst. Sci.*, **15**, 2649–2663, <https://doi.org/10.5194/hess-15-2649-2011>.
- Segura, H., C. Junquas, J. C. Espinoza, M. Vuille, Y. R. Jauregui, A. Rabatel, T. Condom, and T. Lebel, 2019: New insights into the rainfall variability in the tropical Andes on seasonal and interannual time scales. *Climate Dyn.*, **53**, 405–426, <https://doi.org/10.1007/s00382-018-4590-8>.
- Sekulić, A., M. Kilibarda, G. B. M. Heuvelink, M. Nikolić, and B. Bajat, 2020: Random Forest spatial interpolation. *Remote Sens.*, **12**, 1687, <https://doi.org/10.3390/rs12101687>.
- Shafii, M., and B. A. Tolson, 2015: Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives. *Water Resour. Res.*, **51**, 3796–3814, <https://doi.org/10.1002/2014WR016520>.
- Shepard, D., 1968: A two-dimensional interpolation function for irregularly-spaced data. *ACM '68: Proceedings of the 1968 23rd ACM National Conference*, Association for Computing Machinery, 517–524, <https://doi.org/10.1145/800186.810616>.
- Strauch, M., and M. Volk, 2013: SWAT plant growth modification for improved modeling of perennial vegetation in the tropics. *Ecol. Modell.*, **269**, 98–112, <https://doi.org/10.1016/j.ecolmodel.2013.08.013>.
- , R. Kumar, S. Eisner, M. Mulligan, J. Reinhardt, W. Santini, T. Vetter, and J. Friesen, 2017: Adjustment of global precipitation data for enhanced hydrologic modeling of tropical Andean watersheds. *Climatic Change*, **141**, 547–560, <https://doi.org/10.1007/s10584-016-1706-1>.
- Sun, Q., C. Miao, Q. Duan, H. Ashouri, S. Sorooshian, and K. Hsu, 2018: A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.*, **56**, 79–107, <https://doi.org/10.1002/2017RG000574>.
- Tall, M., and Coauthors, 2019: Towards a long-term reanalysis of land surface variables over western Africa: LDAS-Monde applied over Burkina Faso from 2001 to 2018. *Remote Sens.*, **11**, 735, <https://doi.org/10.3390/rs11060735>.
- Tamayo, G. G., 2017: Evaluación de los caudales líquidos y de producción de sedimentos estimados con el modelo Soil Water Assessment Tool (SWAT) y su relación con los caudales líquidos y sólidos observados. Escuela Politécnica Nacional Quito, accessed 22 October 2020, <http://bibdigital.epn.edu.ec/handle/15000/17067>.
- Tan, M. L., P. W. Gassman, X. Yang, and J. Haywood, 2020: A review of SWAT applications, performance and future needs for simulation of hydro-climatic extremes. *Adv. Water Resour.*, **143**, 103662, <https://doi.org/10.1016/j.advwatres.2020.103662>.
- Tang, W., 2019: Dataset of high-resolution (3 hour, 10 km) global surface solar radiation (1983–2018). National Tibetan Plateau

- Data Center, accessed 17 January 2021, <https://doi.org/10.11888/Meteoro.tpcd.270112>.
- , K. Yang, J. Qin, X. Li, and X. Niu, 2019: A 16-year dataset (2000–2015) of high-resolution (3 h, 10 km) global surface solar radiation. *Earth Syst. Sci. Data*, **11**, 1905–1915, <https://doi.org/10.5194/essd-11-1905-2019>.
- Tarek, M., F. P. Brissette, and R. Arsenault, 2020: Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.*, **24**, 2527–2544, <https://doi.org/10.5194/hess-24-2527-2020>.
- Teuling, A. J., I. Lehner, J. W. Kirchner, and S. I. Seneviratne, 2010: Catchments as simple dynamical systems: Experience from a Swiss prealpine catchment. *Water Resour. Res.*, **46**, W10502, <https://doi.org/10.1029/2009WR008777>.
- Tobar, V., and G. Wyseure, 2018: Seasonal rainfall patterns classification, relationship to ENSO and rainfall trends in Ecuador. *Int. J. Climatol.*, **38**, 1808–1819, <https://doi.org/10.1002/joc.5297>.
- Towner, J., H. L. Cloke, E. Zsoter, Z. Flamig, J. M. Hoch, J. Bazo, E. Coughlan de Perez, and E. M. Stephens, 2019: Assessing the performance of global hydrological models for capturing peak river flows in the Amazon basin. *Hydrol. Earth Syst. Sci.*, **23**, 3057–3080, <https://doi.org/10.5194/hess-23-3057-2019>.
- van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang, 2009: Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. *Adv. Water Resour.*, **32**, 1154–1169, <https://doi.org/10.1016/j.advwatres.2009.03.002>.
- Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo, 2014: The WFDEI meteorological forcing data set: WATCH Forcing data methodology applied to ERA-Interim reanalysis data. *Water Resour. Res.*, **50**, 7505–7514, <https://doi.org/10.1002/2014WR015638>.
- Wongchuig Correa, S., R. C. D. de Paiva, J. C. Espinoza, and W. Collischonn, 2017: Multi-decadal hydrological retrospective: Case study of Amazon floods and droughts. *J. Hydrol.*, **549**, 667–684, <https://doi.org/10.1016/j.jhydrol.2017.04.019>.
- Wu, W., Y. Li, X. Luo, Y. Zhang, X. Ji, and X. Li, 2019: Performance evaluation of the CHIRPS precipitation dataset and its utility in drought monitoring over Yunnan Province, China. *Geomatics Nat. Hazards Risk*, **10**, 2145–2162, <https://doi.org/10.1080/19475705.2019.1683082>.
- Xavier, A. C., C. W. King, and B. R. Scanlon, 2016: Daily gridded meteorological variables in Brazil (1980–2013). *Int. J. Climatol.*, **36**, 2644–2659, <https://doi.org/10.1002/joc.4518>.
- , —, and B. R. Scanlon, 2017: An update of Xavier, King and Scanlon (2016) daily precipitation gridded data set for the Brazil. *Proc. XVIII Simpósio Brasileiro de Sensoriamento Remoto*, São Paulo, Brazil, INPE, 28–31, <https://proceedings.science/sbsr/papers/an-update-of-xavier-king-and-scanlon-2016-daily-precipitation-gridded-data-set-for-the-brazil>.
- Xu, X., S. K. Frey, A. Boluwade, A. R. Erler, O. Khader, D. R. Lapen, and E. Sudicky, 2019a: Evaluation of variability among different precipitation products in the northern Great Plains. *J. Hydrol. Reg. Stud.*, **24**, 100608, <https://doi.org/10.1016/j.ejrh.2019.100608>.
- Xu, Z., Z. Wu, H. He, X. Wu, J. Zhou, Y. Zhang, and X. Guo, 2019b: Evaluating the accuracy of MSWEP V2.1 and its performance for drought monitoring over mainland China. *Atmos. Res.*, **226**, 17–31, <https://doi.org/10.1016/j.atmosres.2019.04.008>.
- Xue, X., K. Zhang, Y. Hong, J. J. Gourley, W. Kellogg, R. A. McPherson, Z. Wan, and B. N. Austin, 2016: New multisite cascading calibration approach for hydrological models: Case study in the Red River Basin using the VIC model. *J. Hydrol. Eng.*, **21**, 05015019, [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001282](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001282).
- Yamazaki, D., S. Kanae, H. Kim, and T. Oki, 2011: A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resour. Res.*, **47**, 4501, <https://doi.org/10.1029/2010WR009726>.
- , and Coauthors, 2017: A high-accuracy map of global terrain elevations. *Geophys. Res. Lett.*, **44**, 5844–5853, <https://doi.org/10.1002/2017GL072874>.
- Yilmaz, K. K., H. V. Gupta, and T. Wagener, 2008: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resour. Res.*, **44**, W09417, <https://doi.org/10.1029/2007WR006716>.
- Zambrano-Bigiarini, M., A. Nauditt, C. Birkel, K. Verbist, and L. Ribbe, 2017: Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrol. Earth Syst. Sci.*, **21**, 1295–1320, <https://doi.org/10.5194/hess-21-1295-2017>.
- Zubieta, R., A. Getirana, J. C. Espinoza, and W. Lavado, 2015: Impacts of satellite-based precipitation datasets on rainfall–runoff modeling of the Western Amazon basin of Peru and Ecuador. *J. Hydrol.*, **528**, 599–612, <https://doi.org/10.1016/j.jhydrol.2015.06.064>.
- , —, J. C. Espinoza, W. Lavado-Casimiro, and L. Aragon, 2017: Hydrological modeling of the Peruvian–Ecuadorian Amazon Basin using GPM-IMERG satellite-based precipitation dataset. *Hydrol. Earth Syst. Sci.*, **21**, 3543–3555, <https://doi.org/10.5194/hess-21-3543-2017>.
- Zulkafli, Z., and Coauthors, 2014: A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean–Amazon River Basins. *J. Hydrometeorol.*, **15**, 581–592, <https://doi.org/10.1175/JHM-D-13-094.1>.