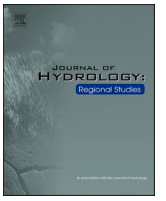
Journal of Hydrology: Regional Studies 47 (2023) 101381

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh



Construction of a daily streamflow dataset for Peru using a similarity-based regionalization approach and a hybrid hydrological modeling framework

Harold Llauca *, Karen Leon, Waldo Lavado-Casimiro

National Service of Meteorology and Hydrology of Peru, Lima 15072, Peru

ARTICLE INFO

Keywords:
Peru
Hydrological regionalization
National hydrological modeling
PISCO dataset

ABSTRACT

Study region: A total of 11,913 sub-catchments in Peru and transboundary catchments with neighboring countries in South America.
Study focus: This paper aims to develop a national hydrological model using physiographic and climatic characteristics to identify donor and receptor sub-catchments (sub-zones). Therefore, we use the hydrometeorological PISCO dataset (0.1° x 0.1°) to drive a sub-catchment conceptual rainfall-runoff (ARNO/VIC) model and a river-routing (RAPID) model in thousands of river reaches. We identify 43 hydrological zones (with 122 sub-zones) to run the hybrid hydrological modeling framework (ARNO/VIC+RAPID) with previously calibrated and validated parameters with 43 fluviometric stations for 1981–2020. Simulated flow series show a higher performance at daily scale (KGE $\geq$ 0.75, NSEsqrt $\geq$ 0.65, MARE $\leq$ 1, and $-25\% \leq$ PBIAS $\leq 25\%$) for catchments located at the Pacific coast and the Andes-Amazon transition, and good representation (R$\geq$0.75) of seasonal and interannual variability.
New hydrological insights for the region: Increasing hydrological hazards such as floods highlight the importance of a systematic hydrological analysis and modeling at national level in gauged and ungauged catchments in Peru. This study introduces a new hydrological dataset of simulated daily flow series. The results are helpful for short-term flood risk scenario simulations and long-term water resource planning as the outcomes can provide valuable information for hydrologists and water resource managers in Peruvian regions with limited or no access to in-situ networks.

1. Introduction

Hydrological models represent an important alternative to provide spatially continuous water discharge simulations across large-scale domains (Zink et al., 2017) and provide valuable information on hydrological risk even in areas with limited ground observation data (Llauca et al., 2021a). Global hydrological models (GHMs) have been used in a wide range of applications such as the global reconstruction of naturalized river flow (Lin et al., 2019), flood forecasting (Lavers et al., 2019), climate change assessment (Hattermann et al., 2017), global water economy (Dolan et al., 2021) and ecosystem services (Janse et al., 2019). The implementation of GHMs has increased significatively since the explosion of global data availability from satellites in the last decades (Tang et al., 2009). Sood and Smakhtin (2015) suggest that GHMs should not aim to replace more locally focused modeling efforts where finer spatial

* Corresponding author.
E-mail address: hllauca@senamhi.gob.pe (H. Llauca).

https://doi.org/10.1016/j.ejrh.2023.101381
Received 22 June 2022; Received in revised form 6 March 2023; Accepted 6 April 2023
Available online 12 April 2023
2214-5818/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

resolutions are needed. In that sense, national hydrological models (NHMs) are becoming an increasingly popular tool for providing information not only in gauged locations but also all ungauged river reaches. For instance, NHMs are currently applied for seasonal water forecasts in Sweden (Girons Lopez et al., 2021), climate change impacts on peak river flows in Great Britain (Kay et al., 2021), and assessing the risk of floods and droughts in New Zealand (McMillan et al., 2016). In Peru, the increasing hydrological risks (Huggel et al., 2015) and water use put pressure on water resources (Drenkhan et al., 2015) - combined with the larger country size (1,285,220 km^2), low density of fluviometric stations, and short flow data records (Llauca et al., 2021b) - have prompted hydrologists and decision makers to extend efforts for more systematic and nationwide water resource analyses.

National scale modeling requires dealing with its own challenges. McMillan et al. (2016) mention that national model input data will depend on the country size, climate variability, and catchment characteristics. In a data scarcity context, many river reaches are poorly gauged or lack any observations and continuous flow data records are often economically and logistically unfeasible (Vereecken et al., 2008).

Many of NHMs applications use physically-based and high-resolution hydrological models (e.g., WEP-CN, VIC, MIKE SHE, WRF-Hydro) that require a large number of inputs at huge computational costs (Rodriguez and Tomasella, 2016). Contrariwise, some studies in Peru (Lavado Casimiro et al., 2011; Llauca et al., 2021b; Saavedra et al., 2021) have shown that conceptual hydrological models such as the Framework for Understanding Structural Errors (Clark et al., 2008) and the Génie Rural rainfall-runoff models (Mouelhi et al., 2006) are also capable to adequately represent hydrological processes in catchments with different characteristic. The application of conceptual rainfall-runoff models in large-scale domains needs to be linked to river-routing models such as RAPID (David et al., 2011b), HYPERstream (Piccolroaz et al., 2016), mizuRoute (Mizukami et al., 2016) and others, to simultaneously compute water discharges in thousands of river reaches as is shown in Lin et al. (2019). In that sense, a hybrid modeling framework combining rainfall-runoff and river-routing models represents a low-resource alternative for operational national daily streamflow simulation in data-scarce countries such as Peru.

Parameter selection for a national model requires procedures that can be applied to many catchments and ungauged catchments. The central premise is searching for similarities between catchments to identify homogeneous hydrological areas (Song and James, 1992). A first way to deal with the extrapolation of hydrological information from a gauged catchment (donor) to an ungauged one (receptor) is by employing the spatial proximity criteria (Pagliero et al., 2014) where model parameters are transferred from a neighbor gauged catchment. Recent studies have introduced other methods for hydrological regionalization in large-scale hydrological models. For instance, (Beck et al., 2020) use the dissimilarity concept for transferring model parameters from donor to receptor

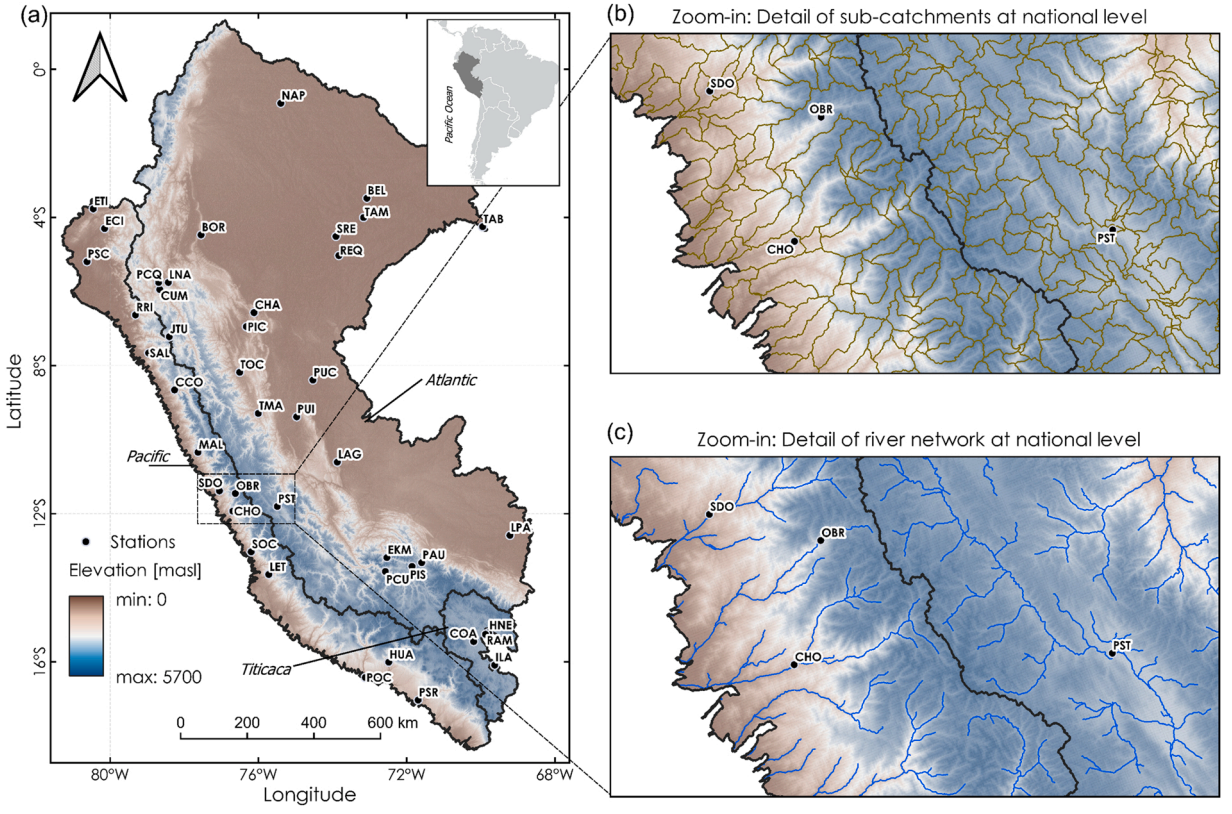


Fig. 1. (a) Study domain boundaries that consider all catchments in the Pacific, Atlantic, and Titicaca slopes in Peru, and transboundary catchments with neighbors' countries. Fluviometric stations are shown on the map as black dots. Zoom-in panels show (b) detail of sub-catchments employed for rainfall-runoff modeling, and **(c)** the river network used for river routing modeling.

catchments in a global model. Narbondo et al. (2020) apply a physical-based approach using the Richards-Baker Flashiness index to predict runoff in ungauged catchments. Rau et al. (2019) use the empirical relationship between geomorphological characteristics and conceptual model parameters. Pagliero et al. (2019) incorporate a partial least squares regression and clustering analysis for hydrological regionalization. Kratzert et al. (2019) apply the Entity-Aware Long Short-Term Memory (EA LSTM) network for regional rainfall-runoff modeling and Bock et al. (2016) use the sensitivity analysis of the rainfall-runoff and runoff variability indices to identify national hydrological calibration zones.

This study aims to develop a national hydrological model to investigate the following research questions: (a) how can similarity-based regionalization help to implement hydrological models in a data-scarce country?; (b) how can available national datasets be applied to simulate daily flow series in thousands of river reaches simultaneously?; (c) how does the national hydrological model perform in Peruvian catchments with different characteristics? For this purpose, we use cluster and dissimilarity approaches to identify pairs of donor and receptor sub-catchments across entire Peru. Furthermore, gridded meteorological datasets are used to drive a hybrid hydrological framework that combines a conceptual rainfall-runoff model and a river-routing model. Finally, we introduced a new hydrological dataset of daily flow series in 11913 river reaches in Peru. This new dataset is helpful for short-term flood risk scenario simulations and long-term water resource planning as the outcomes can provide valuable information for hydrologists and water resource managers in Peruvian regions with limited or no access to in-situ networks.

2. Data and methods

2.1. Study domain

Peru is located in Northwest South America. The Andes mountain range creates a complex topography and determines hydroclimatic variability along three main drainages: the Pacific, Atlantic, and Titicaca (Fig. 1). The Peruvian National Water Agency reports that the Pacific, Titicaca and Amazon drainages represent, respectively, 21.7%, 3.8% and 74.5% of the Peruvian territory. However, the accessibility of freshwater resources in Peru is the converse of population density; 88% of the population lives along the Pacific coast, around Titicaca and in the Andean zones of the Amazon basin, where only 2% of the freshwater resources are available (Lavado Casimiro et al., 2012).

Peruvian hydroclimatology is influenced by the disruption of the large-scale circulation patterns caused by the Andes cordillera, the contrasting oceanic boundary conditions and the landmass distribution (Garreaud et al., 2009). Atmospheric moisture from the Atlantic ascends the Andean slope and leads to high orographic rainfall rates over the eastern divide (Amazon) (Lavado Casimiro et al., 2011). Contrary, the weak rainfall along the Pacific coast (western divide) is related to the large-scale mid-tropospheric subsidence over the southeastern subtropical Pacific Ocean, enhanced by the coastal upwelling of cold water (Lavado Casimiro et al., 2012). Rainfall is highly variable in both space and time, arid conditions with low rainfall rates characterize coastal areas on the Pacific slopes ($<\sim$150 mm/year) while semi-arid conditions ($<\sim$400 mm/year) prevail in the western flanks of the Andes (Rau et al., 2019). The Atlantic and Titicaca divides are characterized by humid and semi-arid conditions, respectively, with high rainfall rates over the eastern Andean slopes ($\sim$1100 mm/year), at the Andes–Amazon transition ($\sim$3200 mm/year), and in the Amazon basin ($\sim$2550 mm/year) (Aybar et al., 2020).

According to Lavado-Casimiro (2013), the rivers in the Pacific divide are characterized by steep slopes, they are not very long, and their rainfall and discharge peaks occur simultaneously. The rivers in Amazon divide are characterized by steep slopes in the Andes and near-zero slopes in the Amazon lowlands. The rainfall peaks occur two months before the discharge peaks in the Amazon plains and simultaneously in the Andes. The Titicaca endorheic drainage is characterized by mean slopes, and because it is in the highlands, the average response between rainfall and discharge peaks is one month.

The study domain includes transboundary catchments shared with Ecuador, Colombia, and Brazil, with an approximate total drainage area of 1,480,620 km^2 (Fig. 1). Given that the success of a similarity-based regionalization approach depends on the use of a large and highly diverse set of catchments (Beck et al., 2020), the study domain was divided into 11,913 sub-catchments with a mean size of 120 km^2, following the same river network as the GEOGloWS ECMWF streamflow services (Sanchez Lozano et al., 2021). This spatial discretization considers a unique river stream per sub-catchment to compute water discharge. All catchments in the study domain - and its sub-catchments - were previously classified as gauged (73%) and ungauged (27%), considering the 43 fluviometric stations displayed in Fig. 1a.

2.2. Data collection

2.2.1. The PISCO dataset

The Peruvian Interpolated data of the SENAMHI's Climatological and hydrological Observations (PISCO) is a hydrometeorological dataset developed for the Peruvian territory, including transboundary catchments. It contains gridded sub-products of precipitation (P), air temperature (TA), and potential evapotranspiration (PET), with a 0.1° spatial resolution at daily and monthly time steps. Recently, Llauca et al. (2021b) incorporated a new sub-product of monthly streamflow simulations at the national level (PISCO_HyM_GR2M) in a vector river flowline format. The PISCO dataset has been widely used in recent studies to assess the hydrological processes in Peruvian catchments (Asurza-Véliz and Lavado-Casimiro, 2020; Fernandez-Palomino et al., 2021; Saavedra et al., 2021).

The PISCO$_{P,TA,PET}$ sub-products are available from 1981 to 2016 (stable versions). However, an unstable version only for PISCO$_P$ is available from 1981 to the present, and it is daily updated for SENAMHI's operational purposes. PISCO$_P$ (Aybar et al., 2020) is generated by using geostatistical and deterministic methods that include three precipitation sources: (a) the quality-controlled

national rain gauge dataset, (b) radar-gauge merged precipitation climatologies, and (c) the Climate Hazards Group Infrared Precipitation (CHIRP) estimates. Similarly, $PISCO_{TA}$ is obtained from: a) time series of maximum and minimum air temperature data, (b) a soil temperature product from the MODIS sensor (Moderate Resolution Imaging Spectroradiometer), and (c) geographic predictors (e.g., elevation, longitude, latitude and Topographic Dissection Index). $PISCO_{PET}$ is generated from the previous $PISCO_{TA}$ data following the methodology proposed by Hargreaves and Samani (1985). The PISCO dataset is freely available on the IRI Data Library website: http://iridl.ldeo.columbia.edu/SOURCES/. SENAMHI/. HSR/. PISCO.

In this study, and for operational purposes, the mean-areal values of P and PET were calculated for each sub-catchment from 1 January 1981–31 March 2020. The $PISCO_P$ unstable version was selected to extend the data baseline beyond 2016. At the same time, we used only daily climatological values of $PISCO_{PET}$ due to the lack of an unstable version similar to $PISCO_P$.

2.2.2. Discharge data

The observed daily flow series of 43 fluviometric stations were selected from the common period of 1 January 1981–31 March 2020. A total of 79% of the stations belong to the National Service of Meteorology and Hydrology of Peru (SENAMHI, https://www.gob.pe/senamhi). Precisely, in the Amazon region (Atlantic slope), most of the stations are monitored by the SENAMHI and by the French Institute for Sustainable Development (IRD) in the framework of the HYBAM Project (https://hybam.obs-mip.fr/). The detail of selected stations is summarized in Table 1, and their spatial distribution is shown in Fig. 1a. The selection processes consider: (a) daily flows series with quality control, (b) fluviometric stations with at least 40% (15 years) of record length for period 1981–2020; (c) streamflow gauges with stage-discharge rating curves with at least 10 streamflow measurements per year and distributed during dry and wet periods; and (d) stations strategically located in river streams with major drainage areas that are used for hydrologic

Table 1
Fluviometric stations selected for hydrological modeling at the national level. Coverage of flow data records [%] is considered from 1 January 1981–31 March 2020.

Slope	Station	Abrev.	Latitude [°]	Longitude [°]	Catchment	Total drainage area [km^2]	Coverage [%]
Pacific	El Tigre	ETI	-3.77	-80.46	Tumbes	4781.57	97.70
	El Ciruelo	ECI	-4.30	-80.15	Chira	7014.84	98.00
	Pte. Sanchez Cerro	PSC	-5.19	-80.62	Piura	7519.99	49.00
	Racarumi	RRI	-6.63	-79.32	Lambayeque	2439.92	99.80
	Salinar	SAL	-7.66	-78.96	Chicama	3660.22	92.50
	Condorcerro	CCO	-8.66	-78.26	Santa	10,382.31	98.70
	Malvados	MAL	-10.34	-77.63	Fortaleza	1393.97	31.00
	Santo Domingo	SDO	-11.38	-77.05	Chancay-Huaral	1850.99	66.20
	Obrajillo	OBR	-11.45	-76.62	Chillón	786.00	59.00
	Chosica	CHO	-11.93	-76.69	Rímac	2307.22	55.00
	Socsi	SOC	-13.03	-76.20	Cañete	5926.31	93.30
	Letrayoc	LET	-13.64	-75.72	Pisco	3089.68	89.10
	Huatiapa	HUA	-16.01	-72.48	Camaná	13,220.59	45.40
	Pte. Ocoña	POC	-16.42	-73.12	Ocoña	15,369.27	34.10
	Pte. Santa Rosa	PSR	-17.03	-71.69	Tambo	13,083.31	92.50
Titicaca	Pte. Huancané	HNE	-15.22	-69.79	Huancané	3388.79	79.90
	Pte. Ramis	RAM	-15.26	-69.87	Intercuenca Ramis	14,856.33	72.90
	Pte. Unocolla	COA	-15.45	-70.19	Coata	4448.09	74.90
	Pte. Ilave	ILA	-16.09	-69.63	Ilave	7623.49	72.70
Atlantic	Napo	NAP	-0.92	-75.40	Napo	27,351.88	40.30
	Bellavista	BEL	-3.48	-73.07	Napo	99,872.13	71.30
	Tamshiyacu	TAM	-4.00	-73.16	Amazonas	653,396.03	90.70
	Tabatinga	TAB	-4.25	-69.95	Amazonas	808,803.30	87.50
	Borja	BOR	-4.47	-77.55	Marañon	114,302.21	86.80
	San Regis	SRE	-4.51	-73.91	Marañon	293,078.73	53.50
	Requena	REQ	-5.03	-73.83	Ucayali	347,564.48	59.10
	Pte. Corral Quemado	PCQ	-5.76	-78.69	Marañon	8650.17	14.40
	Los Naranjos	LNA	-5.76	-78.43	Marañon	5793.89	18.40
	Cumba	CUM	-5.94	-78.66	Marañon	35,672.21	14.20
	Chazuta	CHA	-6.57	-76.12	Huallaga	68,708.21	41.90
	Picota	PIC	-6.95	-76.33	Huallaga	56,844.64	42.70
	Jesús Tunel	JTU	-7.22	-78.40	Crisnejas	910.21	98.20
	Pte. Tocache	TOC	-8.18	-76.51	Huallaga	23,173.35	58.90
	Pucallpa	PUC	-8.39	-74.53	Ucayali	260,243.86	43.20
	Tingo Maria	TMA	-9.29	-76.00	Huallaga	12,341.11	52.20
	Puerto Inca	PUI	-9.38	-74.97	Pachitea	22,325.21	43.60
	Lagarto	LAG	-10.61	-73.87	Ucayali	19,1192.56	23.60
	Puente Stuart	PST	-11.80	-75.49	Mantaro	9222.46	69.50
	La Pastora	LPA	-12.58	-69.21	Madre de Dios	56,242.24	24.30
	Egemsa km 105	EKM	-13.18	-72.53	Urubamba	9653.03	86.20
	Paucartambo	PAU	-13.32	-71.59	Urubamba	2157.18	29.40
	Pisac	PIS	-13.42	-71.86	Vilcanota	7036.46	67.30
	Puente Cunyac	PCU	-13.56	-72.57	Apurimac	24,656.23	26.90

calibration. In strategic gauges with record lengths less than 40% (MAL, POC, PCQ, LNA, CUM, LAG, LPA, PAU and PCU), a continuous 4-year record period within 1981–2020 was required to be considered for hydrologic model calibration.

2.3. Similarity-based regionalization approach

2.3.1. Predictor maps

We used an a priori defined similarity criterion incorporating four climatic (P: precipitation; TA: air temperature; PET: potential evapotranspiration; AI: aridity index) and three physiographic (SLO: surface slope; CLAY: soil clay content; fTC: fraction of tree cover) characteristics which can be found in Table 2. According to Beck et al. (2015) and Barbarossa et al. (2018), these characteristics are strongly related to streamflow signatures. In this study, drainage area was not considered due to the high level of catchment discretization for semi-distributed modeling, and with sub-catchments of similar sizes near 120 km^2. Also, the snow cover fraction was not taken into account as a predictor due to most Peruvian catchments being dominated by pluvial contribution (Aybar et al., 2020; Fernandez-Palomino et al., 2021). Maps of mean annual P, TA, and PET values were calculated for the period 1981–2010. AI was calculated from P and PET. In case of fTC, SLO, and CLAY, mean annual maps were computed from available data during 1981–2010. Finally, the mean areal values of each predictor were calculated for all sub-catchments (gauged and ungauged) across the study domain.

2.3.2. Cluster analysis and dissimilarity assessment

Hydrological models often rely on regionalization approaches to transfer information from gauged (donor) to ungauged (receptor) catchments. According to Beck et al., (2016, 2020), benefits of the similarity-based approach include: (a) its relative ease of implementation; (b) retainment of model parameter interaction because the entire parameter set is transferred and (c) the possibility of spatial variability in model parameters according to landscape characteristics, even in ungauged regions.

This work uses a similarity-based approach (Fig. 2) since a previous study has found it outperformed well in the Pacific slope (Asurza-Véliz and Lavado-Casimiro, 2020) showing that it is possible to transfer calibrated parameter sets based on explicit consideration of climatic and physiographic similarity.

Before pairing donor-receptor sub-catchments, we apply a gauged sub-catchment classification procedure. As the main idea was to obtain more than one set of model parameters per catchment during model calibration and increase the number of donor candidates during parameter transfer, the final number of parameter sets by each gauged (calibrated) catchment will be equal to clusters in it. For this, we conduct a hierarchical cluster analysis using the predictor maps as descriptors. To avoid increasing parameter uncertainties during calibration, we adopted a maximum number of clusters per catchment of three.

Seven climatic and physiographic descriptors at the sub-catchment level were computed from predictor maps (see Table 2). The dissimilarity between donors and receptors was computed using the following equation proposed by Beck et al. (2016):

$$S_{i,j} = \sum_{p=1}^{7} \frac{\left|Z_{p,i} - Z_{p,j}\right|}{IQR_p}; \quad (1)$$

where S is the dissimilarity (−), Z are the values of the respective descriptor, p denotes the descriptor, and i and j mean, respectively, the donor and the receptor sub-catchment in question, IQR is the interquartile range of the descriptor and represents the spatial variability in all descriptor. The division by IQR in equation [1] was necessary to equalize the data variability of descriptors. From equation [1] it follows that a similar sub-catchment yields an S value close to zero.

Then, dissimilarity values compose a distance matrix (3367 receptors x 8546 donors) which characterizes the variability of the sub-catchments in the region of interest. We also incorporate the spatial proximity criterion selecting the most similar donor (with the lowest S value) from no more than 300 km distance. This was considered to avoid sub-catchment donors from too long distances and different drainage regions.

Table 2
The climatic and physiographic characteristics selected to quantify the similarity between catchments and their sub-catchments.

Variable	Units	Variable	Data source	Resolution
P	mm/year	Mean annual precipitation	$PISCO_P$ gridded data, mean of years from 1981 to 2010 (Aybar et al., 2020)	0.1°
PET	mm/year	Mean annual potential evapotranspiration	$PISCO_{PET}$ (1981–2010) calculated from $PISCO_{TA}$ using the equation proposed by (Hargreaves and Samani, 1985).	0.1°
AI	-	Aridity index	Calculated as: AI=PET/P, where P is the mean annual precipitation and PET the mean annual evapotranspiration. Values were truncated with an upper limit of 10 to avoid extremely high values on the coast.	0.1°
TA	°C	Mean air temperature	$PISCO_{TA}$ v1.1 gridded data, mean of years from 1981 to 2010.	0.1°
fTC	%	Fraction of forest cover	Landsat-based forest cover for the year 2000 (Hansen et al., 2013)	30 m
SLO	°	Surface slope	HydroSHED SRTM (Lehner et al., 2008)	90 m
CLAY	g/Kg	Soil clay content	SoilGrids 2020 version (ISRIC), mean over all layers (0–100) (Poggio et al., 2021).	250 m

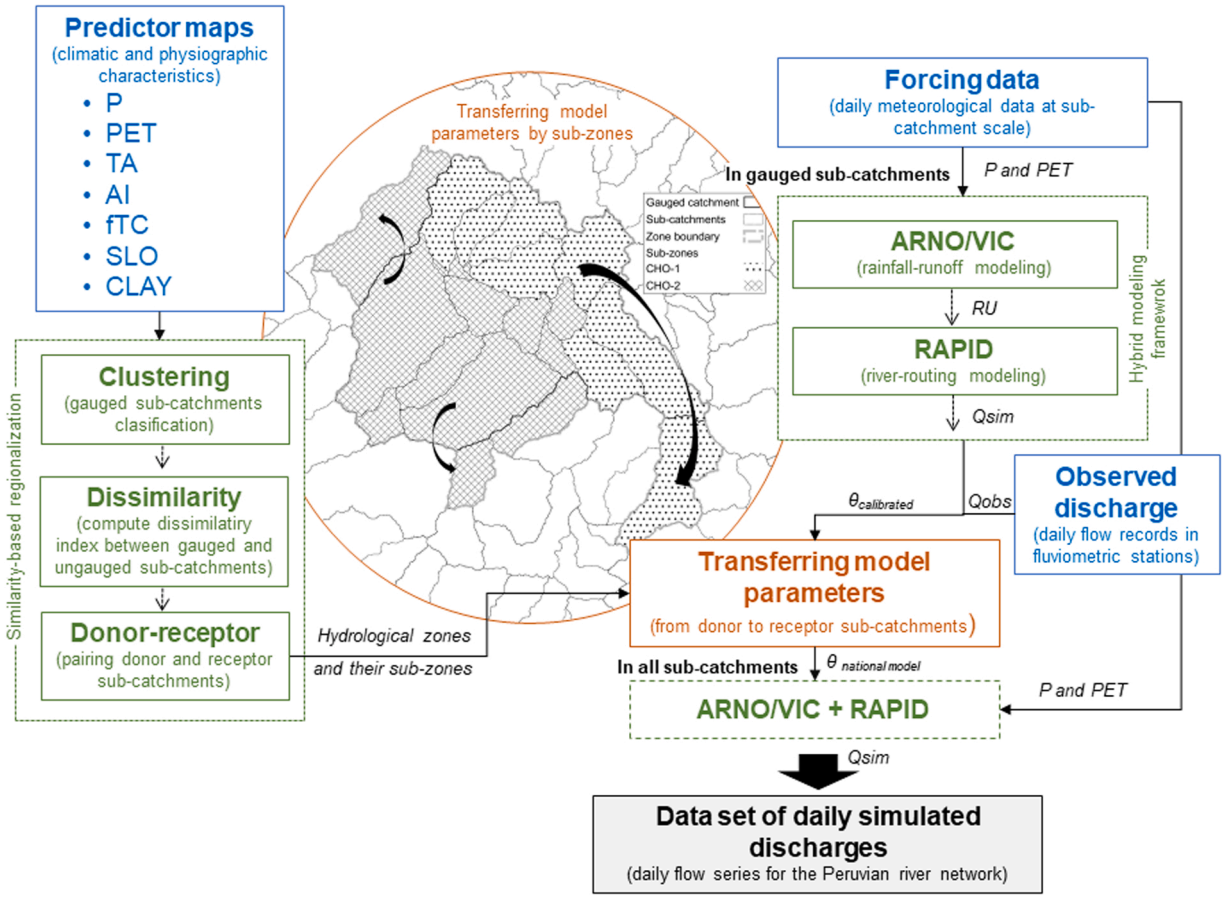


Fig. 2. Framework for the sub-catchment national simulation of daily flow series applying a similarity-based regionalization to identify hydrological zones and sub-zones, and a hybrid modeling framework for parameter calibration and streamflow simulation across the whole Peruvian river network. P: precipitation; PET: potential evapotranspiration; TA: air temperature; AI: aridity index; SLO: surface slope; CLAY: soil clay content; fTC: fraction of tree cover; RU: runoff; Qsim: simulated streamflow; Qobs: observed streamflow; θ: model parameters.

2.4. National hydrological modeling

2.4.1. Hybrid modeling framework

We conduct a hybrid framework to develop the sub-catchment national hydrological model (Fig. 2). First, meteorological data (P and PET) is used to drive a conceptual rainfall-runoff model which estimates runoff volumes (RU) at each sub-catchment. Then, rainfall-runoff outputs are inputs for a river routing modeling to generate daily streamflow (Q) across the river network. Details of both models are described as follows:

2.4.1.1. Rainfall-runoff model. The Variable Infiltration Capacity (ARNO/VIC) model (Liang et al., 1994; Todini, 1996) was selected because of its flexibility and computational efficiency. Despite several simplifications, the model is designed to provide a relatively complete representation of the dominant hydrologic fluxes (Addor and Melsen, 2019; Lane et al., 2019). Here we select de ARNO/VIC model incorporated in the Framework for Understanding Structural Errors (FUSE) developed by Clark et al. (2008), using the 'fuse' R package (Vitolo et al., 2016). The FUSE framework has 79 unique model structures by combining components of existing hydrological models. The ARNO/VIC structure has a single-state variable for the upper soil layer (S_1). The lower soil layer (S_2) can be defined by a single nonlinear baseflow reservoir of fixed size. Evaporation (e) is modeled in both soil layers and computed based on the relative root fractions in each soil layer. Percolation (q_{12}) takes available water from wilting point (θ_{wilt}) to saturation (θ_{wilt}). The surface runoff (q_{sx}) is conceptualized using the ARNO/VIC parameterization (upper zone control), and a gamma distribution is used for the time delay runoff. Readers can find more details about the model equations in Clark et al. (2008).

2.4.1.2. River routing model. The Routing Application for Parallel Computation of Discharge (RAPID) - which uses a matrix-based version of the Muskingum routing scheme - was applied to compute water discharge in thousands of river reaches simultaneously (David et al., 2011b). RAPID is an open-source model available at https://github.com/c-h-david/rapid (David et al., 2011a; Follum et al., 2017; Lin et al., 2015; Salas et al., 2018; Tavakoly et al., 2017). The Muskingum method can be rearranged by introducing a

transboundary matrix $T = (I\text{-}C_1 \cdot N)$ as described by David et al. (2015) in the RAPID model:

$$[I - C_1 \bullet (N - T)] \bullet Q(t + \Delta t) = b(t) + T \bullet C_1 \bullet Q(t + \Delta t); \quad (2)$$

where t is time, Δt is the river routing time step, I is the identity matrix, N is the river network matrix, Q is a vector of outflows from each reach, and C_1 is a diagonal parameter matrix for a given reach j.

We use the connectivity information from the GEOGloWS ECMWF streamflow services (Sanchez Lozano et al., 2021) to create the network connectivity matrix N. Sub-catchment runoff is connected to the routing scheme at the upstream node of each river reach and flow output from the model represents flow at the downstream node.Table 3.

2.4.2. Model calibration and validation

We first execute the hybrid framework only for gauged sub-catchments to find optimal model parameters. The model was calibrated and validated for the period with simultaneously observed streamflow and input data. We choose 60% of available records for calibration and 40% for validation periods. The calibration strategy was the same as that used by Llauca et al. (2021b). We consider a unique calibration step for the Pacific and Titicaca drainages, and a stepwise calibration approach (seven steps) for the Atlantic slope due to the number of fluviometric stations and complexity of river network on the Amazon basin. The Shuffled Complex Evolutionary (SCE-UA) algorithm (Duan et al., 1993) was applied to optimize the model parameters. An artificial year (1980) of a warm-up period was used before computing performance metrics. This year was created using P and PET daily climatologies to avoid discarding model simulations since 1 January 1981. The Kling-Gupta efficiency (KGE) criterion (Gupta et al., 2009) evaluated over daily streamflow time series was used as objective function. Streamflow simulations were also assessed through the Nash-Sutcliffe squared efficiency (NSE_{sqrt}) criterion (Chiew et al., 1993), the Mean Absolute Relative Error (MARE), and the Percent Bias (BIAS) to take into account different hydrograph aspects. The same selected statistical metrics (summarized in Table 4) were computed for calibration, validation, and the total simulation period.

2.4.3. Leave-one-out cross-validation

The similarity-based hydrological regionalization approach is used in this study to transfer model parameters from gauged to ungauged catchments. Thus, the uncertainties surrounding the regionalization procedure should be evaluated by applying a cross-validation procedure to assess the regionalization skill (Wang et al., 2021). Leave-one-out cross-validation (LOOCV) was conducted to assess the quality of transferring parameters. For this, each catchment is in turn considered as pseudo-ungauged while the regionalization approach is applied to estimate its flows (Parajka et al., 2005). Then, the remainder gauged catchments are candidates of donors and regionalization procedure was applied to transfer model parameters from the catchments with the least dissimilarity index. Finally, the simulated streamflow at the pseudo-ungauged catchment is then compared with observation and statistical metrics (KGE, NSE_{sqrt}, MARE and BIAS) are computed. LOOCV was only applied to 15 catchments in the Pacific drainage (see Table 1) and statistical metrics obtained with LOOCV was compared with the independent model calibration (for total period).

2.4.4. Streamflow simulation across the country

First, we transfer optimal parameters from donor to receptor sub-catchments considering a unique set of model parameters for each hydrological sub-zone. Then, a new model run was computed for all catchments and the respective gauged and ungauged sub-catchments to simulate daily flow series (1981–2020) across the national river network (Fig. 2). Also, monthly and yearly flow series were computed to assess the seasonal (September - August) and interannual (1990–2019) representation of flows in fluviometric stations. For this, we compute the Pearson correlation coefficient (R) and the Root Mean Square Error (RMSE) between observed and simulated time series. Fig. 3.

A large body of studies assesses hydrological signatures to characterize streamflow time series and to predict some hydrological processes in ungauged catchments (Addor and Melsen, 2019; McMillan et al., 2017, 2016; Tyralis et al., 2021). We compute nine hydrological signatures (summarized in Table 5) to assess model predictions. Some of the selected indices are hydrological signatures introduced by Addor et al. (2018). We choose the 5% flow quantile, mean daily discharge, 95% flow quantile, baseflow index, slope of the flow duration curve, average duration of high-flow events, frequency of high-flow events, average duration of low-flow events, and

Table 3
ARNO/VIC model parameters and defined upper and lower bounds.

Parameter	Description	Units	Lower bound	Upper bound
MAXWATER1	Depth of upper soil layer	mm	25	500
MAXWATER2	Depth of lower soil layer	mm	50	5000
FRACTEN	Fraction total storage in tension storage	-	0.05	0.95
RTFRAC1	Fraction of roots in the upper layer	-	0.05	0.95
PERCRTE	Percolation rate	mm/day	0.01	1000
PERCEXP	Percolation exponent	-	1	20
BASERTE	Baseflow rate	mm/day	0.001	1000
QB_POWR	Baseflow exponent	-	1	10
AXV_BEXP	ARNO/VIC *b* exponent	-	0.001	3
TIMEDELAY	Time delay runoff	days	0.01	7

Source:Adapted from (Lane et al. (2019).

Table 4
Statistical metrics and their corresponding equations used for evaluating the hydrological performance of the ARNO/VIC+RAPID model.

Statistical Metric	Equation	Units	Min, Max, Optimal	Emphasis
Kling-Gupta Efficiency (KGE)	$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2}\, r = \frac{\sum_{i=1}^{n}[(X_i - \overline{X})(O_i - \overline{O})]}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2}}\, \alpha = \frac{\sigma_X}{\sigma_O}; \beta = \frac{\mu_X}{\mu_O}$	-	-∞ ,1,1	High flows Medium flows (Mizukami et al., 2016)
Nash-Sutcliffe squared (NSEsqrt)	$NSE_{sqrt} = 1 - \frac{\sum_{i=1}^{n}\left(\sqrt{O_i} - \sqrt{X_i}\right)^2}{\sum_{i=1}^{n}\left(\sqrt{O_i} - \sqrt{\overline{O}}\right)^2}$	-	-∞ ,1,1	General flows (Chiew et al., 1993)
Mean Absolute Relative Error (MARE)	$MARE = \frac{1}{n}\sum_{i=1}^{n}\frac{\lvert X_i - O_i\rvert}{O_i}$	-	0,+ ∞,0	Relative error Low flows (Ferreira et al., 2020)
Percent Bias (PBIAS)	$PBIAS = 100\frac{\sum_{i=1}^{n}(O_i - X_i)}{\sum_{i=1}^{n}O_i}$	%	-∞ ,+ ∞,0	Average tendency of the simulated data (Ferreira et al., 2020)

Note: n, number of samples; Oi, observed streamflow; Xi, simulated streamflow.

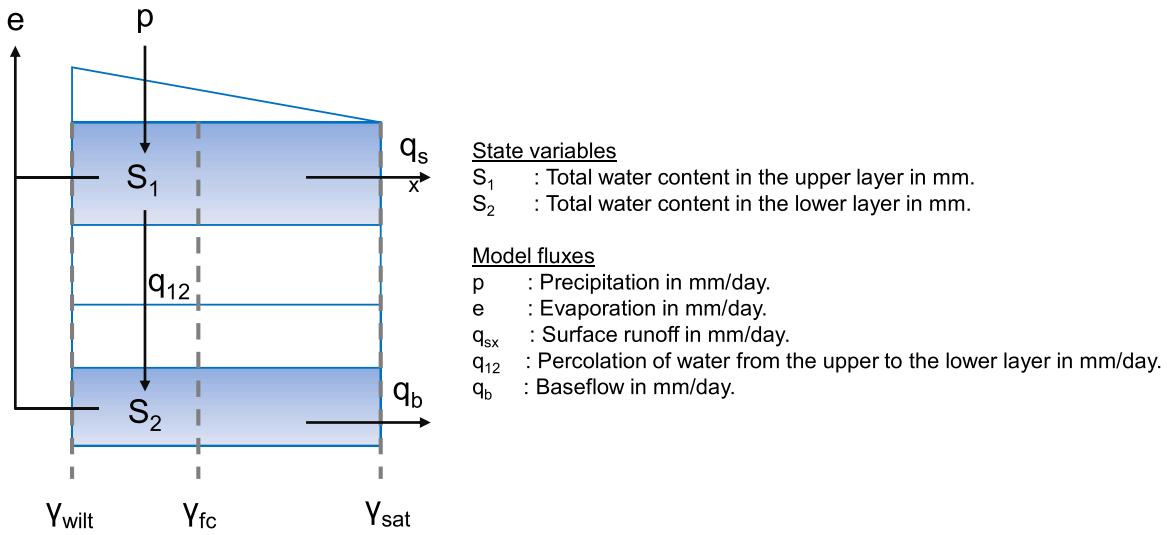


Fig. 3. Structure of the ARNO/VIC model with a single-state variable for the upper soil layer (S1) and a single lower soil layer (S2) defined by a single nonlinear baseflow reservoir of fixed size.
Adapted from Clark et al. (2008).

Table 5
Hydrological signatures selected in this study and computed from 1 September 1990–31 August 2019.

Attribute	Description	Units
Mean daily discharge	Mean daily discharge.	mm/day
5% flow quantile	5% flow quantile (low flow).	mm/day
95% flow quantile	95% flow quantile (high flow).	mm/day
Baseflow index	Baseflow as a proportion of total flow, calculated as 7-day mean annual low flow divided by mean flow.	-
Slope of the flow duration curve	Slope between 33 and 66 percentiles of flow duration curve calculated on daily data.	-
Average duration of high-flow events	Number of consecutive days > 9 times the median daily flow.	days
Frequency of high-flow events	Frequency of high-flow days (>9 times the median daily flow).	days/year
Average duration of low-flow events	Number of consecutive days < 0.2 times the mean daily flow.	days
Frequency of low-flow events	Frequency of high-flow days (<0.2 times the mean daily flow).	days/year

frequency of low-flow events. We calculate model errors in four observed data coverage intervals (0–25%, 25–50%, 50–75%, 75–100%) to compare predicted and observed signatures in different fluviometric stations across the country. The median of errors was computed, and standard error was calculated in each interval by dividing the standard deviation by the squared root of the sampling data. Error bars were built using the median values + /- the standard error.

3. Results

3.1. Selection of donor and receptor sub-catchments

The left side of Fig. 4 displays the spatial variability of four climatic (P, TA, PET, AI) and three physiographic (SLO, fTC, CLAY) characteristics. High mean annual rates of P and PET can be found in the Atlantic and Titicaca slopes, while the Pacific coast is characterized by very high AI values (with low P and high PET). TA is lower over the Andes and higher in the northeastern Amazon basin and seems to be inversely related to SLO values. CLAY tends to be relatively homogeneous across the study domain, with maximum values in the Amazon basin. fTC is also considerably higher in the Atlantic divide than in the Pacific and Titicaca drainages.

Overall, 73% of total sub-catchments are donors while the remaining 27% are receptors. It means that just a few parts of national model parameters and streamflow simulation are generated by the hydrological regionalization procedure. The Pacific slope has the major percentage of receptor sub-catchments per divide (38%), followed by Titicaca (29%) and Atlantic (16%) slopes.

The colored areas on the right side of Fig. 4 correspond to 43 gauged catchments in the study domain. The zoom-in panel shows details of hydrological zones and sub-zones for the Central Pacific slope. In this case, gauged catchments (colored polygons) were classified into one (OBR-1), two (CHO-1 and CHO-2) or three (SDO-1, SDO-2, SDO-3) clusters. Then, model parameters are transferred from donors (colored areas) to receptors (white areas) at sub-catchment level. Sub-catchments labeled with the same station name (e.g. CHO) belong to the same hydrological zone while the number next to the names distinguishes the sub-zone (e.g. CHO-1 or CHO-2). Furthermore, we identify a total of 43 hydrological zones and 122 sub-zones.

3.2. Evaluation of the national model performance

We conduct the daily national model for period 1981–2020. Fig. 5 depicts the spatial distribution of four metrics. In terms of KGE

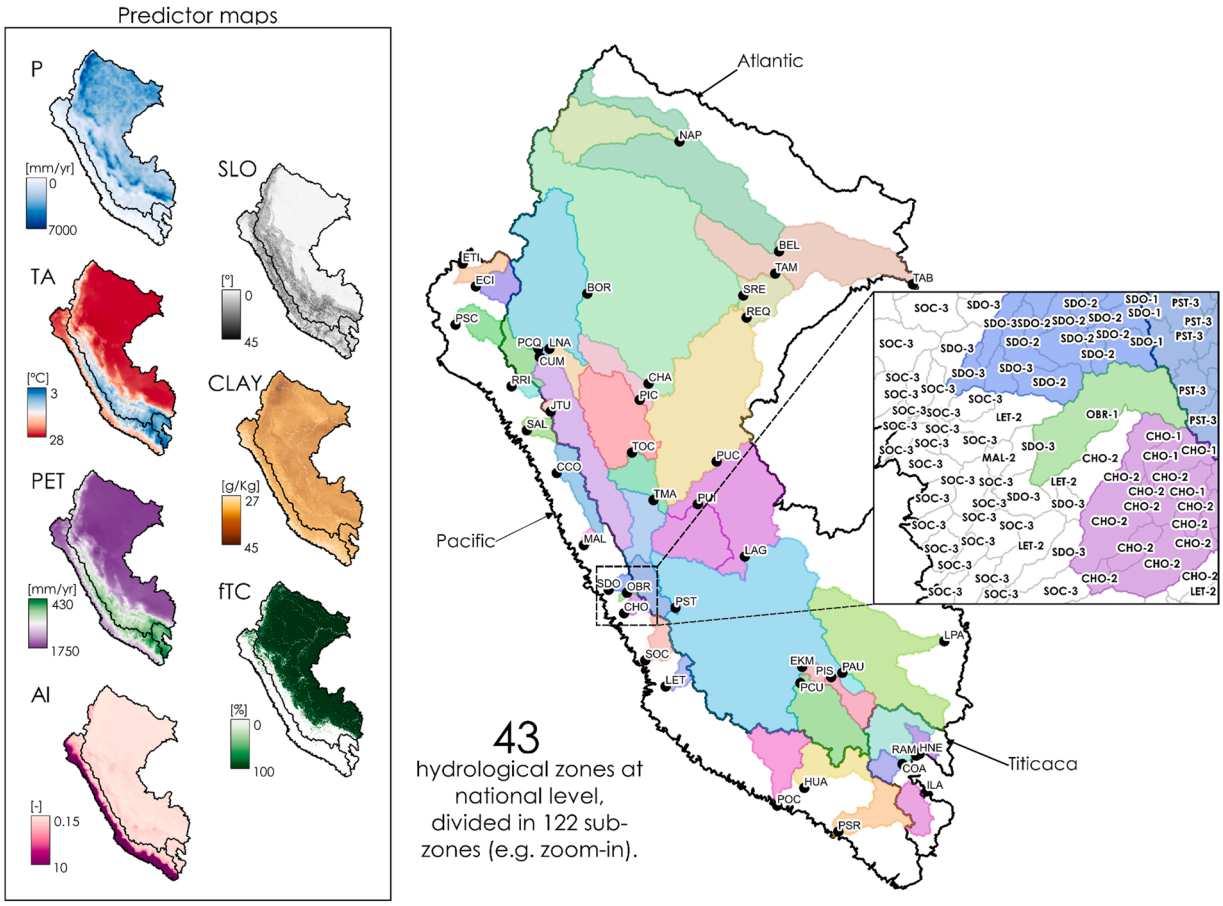


Fig. 4. On the left side: climatic (P: precipitation; TA: air temperature; PET: potential evapotranspiration; AI: aridity index) and physiographic (SLO: surface slope; CLAY: soil clay content; fTC: fraction of tree cover) predictor maps. On the right side: colored areas represent gauged catchments and sub-catchments (donors of model parameters), and white areas correspond to ungauged sub-catchments (receptors of model parameters). Points in black correspond to fluviometric stations. The zoom-in panel shows details of hydrological zones (e.g. SDO) and their respective sub-zones (e.g., SDO-1, SDO-2 and SDO-3) for three catchments in the Central Pacific slope.

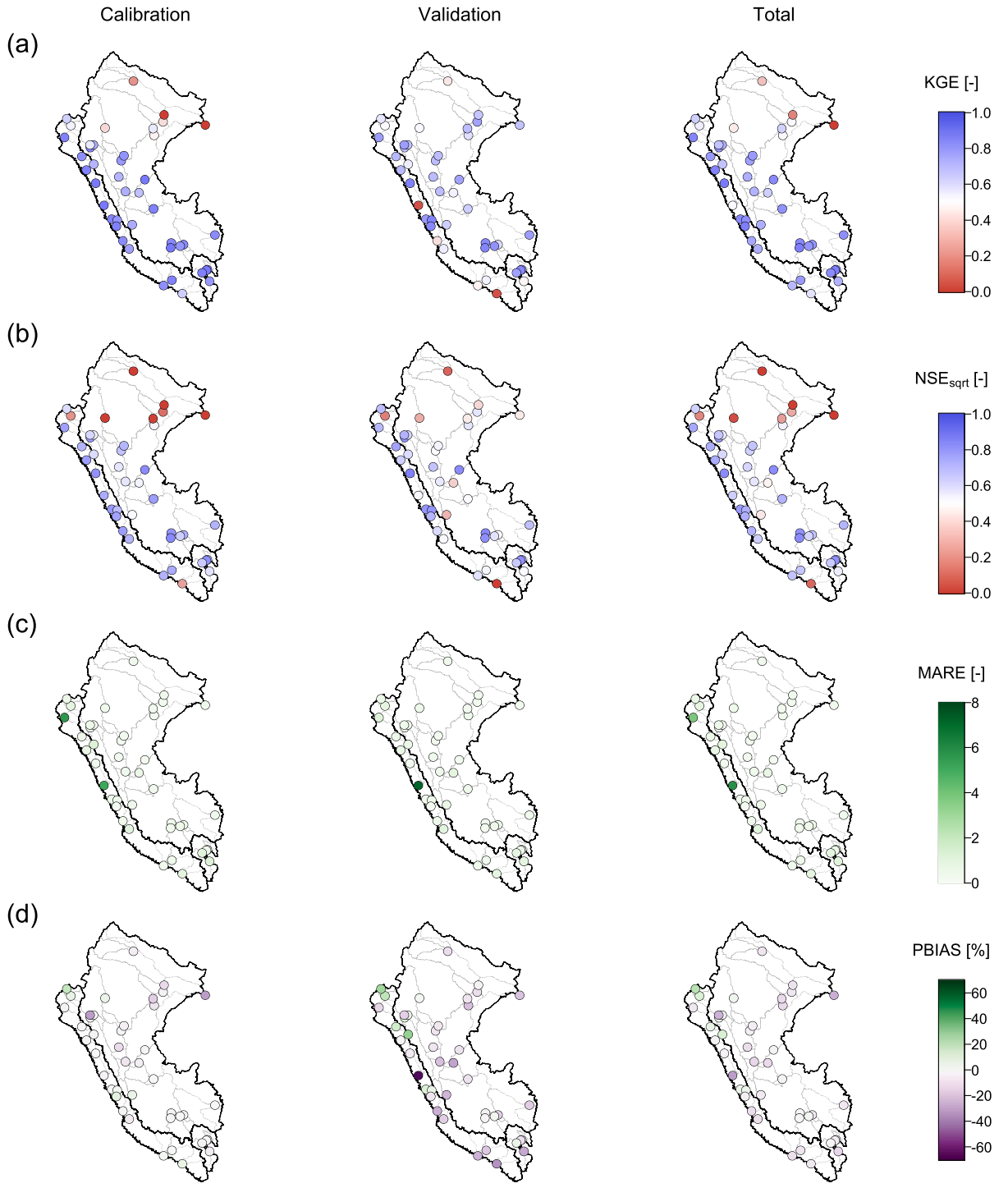


Fig. 5. (a-d) Spatial distribution of metrics for the national hydrological model performance assessment during calibration, validation, and total period (1981–2020). Each metric is focused on a specific aspect of hydrographs (KGE: middle-high flows; NSE_{sqrt}: general flows; MARE and PBIAS: model error and bias). Colored points represent gauges' performance.

and NSE_{sqrt} (Fig. 5a-b), the model performs well (KGE $\geq$ 0.75 and $NSE_{sqrt} \geq 0.65$) during calibration and validation periods for stations located in the Pacific, Titicaca, and the Andes-Amazon transition. Fluviometric stations in the Amazon lowlands (northeast) present low performance ($\leq$0.5 in both metrics). We note that in this group of gauges, KGE and NSE_{sqrt} values slightly increase from calibration to validation period. In terms of MARE and PBIAS (Fig. 5c-d), good model performance (MARE $\leq$ 1 and $-25\% \leq$ PBIAS $\leq 25\%$) can be

found for both periods for most stations. Only two gauges present higher MARE values (~8) and for the Pacific drainage, negative values of PBIAS increase from calibration to validation.

Overall, for the total period, the national hydrological model performs well with high KGE and NSE_{sqrt} values for 71% and 68% of all stations, respectively. The spatial behavior of both indices shows better model performances in the Pacific and Titicaca drainages than in the Atlantic slope (low performance prevails in the Amazon lowlands). In this study, high KGE and NSE_{sqrt} values show a good representation of high-middle and general flows, respectively, for most gauges. MARE is a good index for low data values considering that it is less sensitive to high flow values (Ferreira et al., 2020). That means that 95% of stations with lower relative error values also confirm a good representation of low flows for total period. The model performance shows lower negative PBIAS, indicating a slight tendency for model overestimation (Table 4), except for the northern Pacific coast where the model generally underestimates water discharges.

3.3. Cross-validation in the Pacific slope

Fig. 6 displays results of the cross-validation procedure in all gauged catchments for the Pacific drainage. Panels present scatter plots between statistical metrics before and after cross-validation. Green boxes denote areas with good model performance (KGE $\geq$ 0.50, $NSE_{sqrt} \geq 0.50$, MARE ≤ 1, and $-25\% \leq$ PBIAS $\leq 25\%$). In terms of KGE, most gauges located in the central and south Pacific have good skills (points inside the green box) after LOOCV. Hence, it seems that regionalization's procedure is still a challenge in northern catchments for high-middle flows where model performance highly decreases. In terms of NSE_{sqrt}, general daily flows are good represented after transferring parameters, but northern gauges present low performances ($NSE_{sqrt} \leq 0.5$). In the case of MARE, major points have values less than 1, following the same performance as individual model calibration. Finally, gauges located in the south

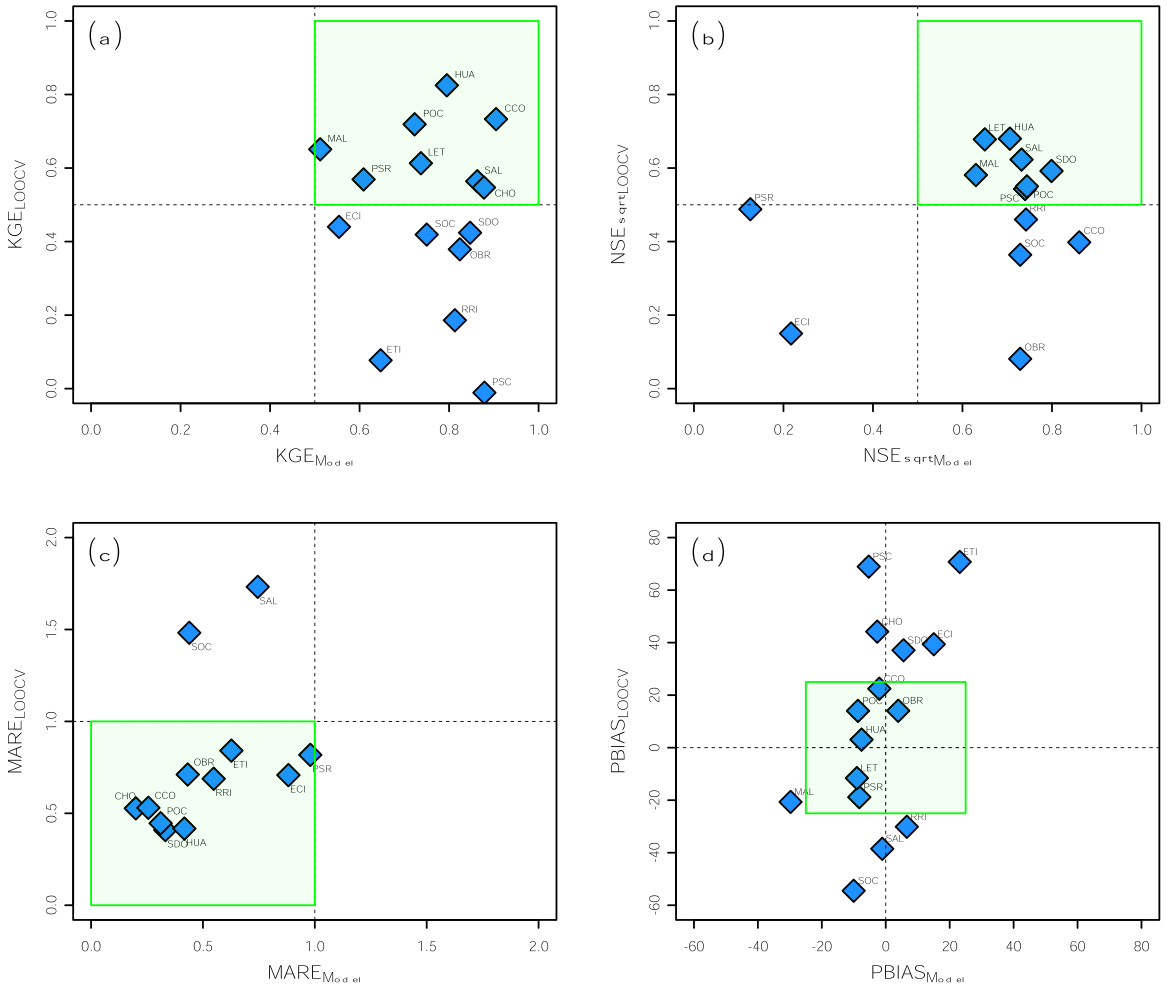


Fig. 6. (a-f) Leave-One-Out Cross Validation (LOOCV) experiment for 15 catchments in the Pacific slope. Each catchment is in turn considered pseudo-ungauged while the regionalization approach was applied to try and estimate its flows. Then, statistical metrics (KGE, NSE_{sqrt}, MARE and PBIAS) were computed (LOOCV subscript) and compared with metrics for total period evaluation after individual calibration (Model subscript). Green boxes display areas with good model performance (KGE $\geq$ 0.50, $NSE_{sqrt} \geq 0.50$, MARE ≤ 1, and $-25\% \leq$ PBIAS $\leq 25\%$).

Pacific present lower biases and we note that most catchments with negative biases in model calibration have positive biases after cross-validation.

3.4. National simulation of flow series

Fig. 7 displays simulated and observed flow time series for the six gauges with the highest record length. Comparison between

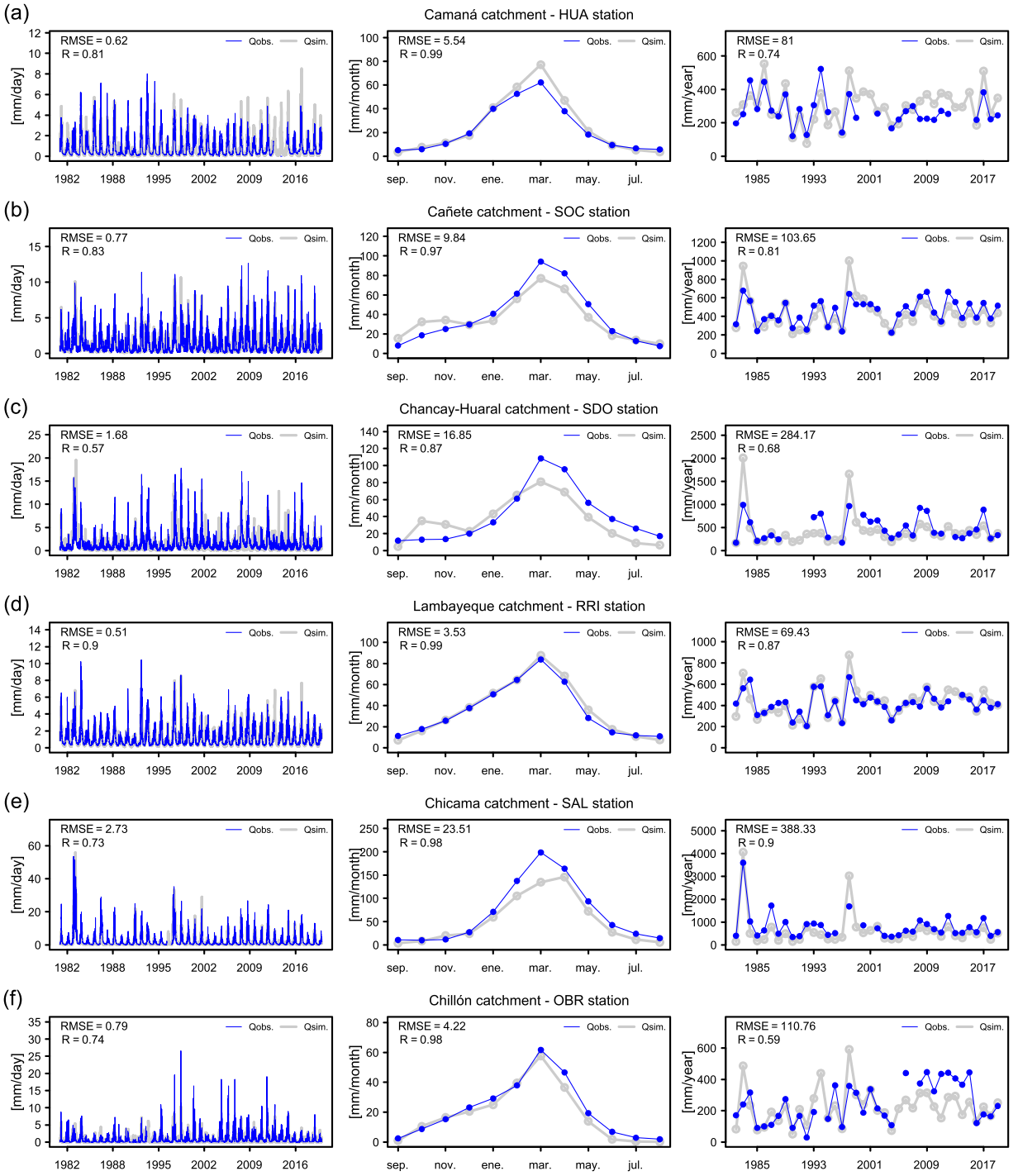


Fig. 7. (a-f) Comparison between observed (blue) and simulated (gray) flow series at daily [mm/day], monthly [mm/month] and yearly [mm/year] time scales for six fluviometric stations with the highest record length during 1981–2020.

daily, monthly (seasonal), and annual flow series show that model outputs are capable to reproduce streamflow behavior at different time scales. Simulations can represent the high variability of daily flows, including daily peak flows. Seasonal streamflow variability is well represented in many cases (with exception of SAL station), so wet (December - April) and dry seasons (June - September) are replicated by simulations. Model outputs perform well at annual scale, and simulated interannual variability follows a similar pattern to observed values; however, some extreme years (e.g. 1982/1983 and 1997/1998) are overestimated by the model.

The comparison between observed and simulated flow series in terms of R and RMSE metrics for different time scales is shown in Table 6. There are not seem to be a strong relationship between flow data coverage and R and RMSE values. For instance, streamflow gauges with low record length (e.g., CUM – 14.2%) can get higher R values than those with more observed information (e.g., PSR – 92.50%). In many gauges, simulations have a strong correlation ($R \geq 0.75$) with observed values at daily, monthly, and annual time scales. RMSE values are slightly heterogeneous in space. Higher model errors are present in gauges of the Amazon lowlands (e.g. TAM=14710.2 mm/year, BEL=974.91 mm/year, TAB=737.95 mm/year).

Fig. 8 depicts errors between predicted and observed signatures related to 5% flow quantile, mean daily discharge, 95% flow quantile, baseflow index, slope of the flow duration curve, average duration of high-flow events, frequency of high-flow events, average duration of low-flow events, and frequency of low-flow events. Points in magenta are median errors for each interval of data coverage while vertical bars correspond to + /- standard error. Positive error values correspond to overestimation of predicted signatures while negative ones mean underestimation. Median values in flow magnitude signatures (Fig. 8a-c) are the most positive and error bars increase from low (5% quantile) to high (95% quantile) flows. Error in baseflow index signature (Fig. 8d) is very low in all stations. The median of slope of flow duration curves errors (Fig. 8e) is negative and decreases when record length grows. Median

Table 6
Pearson linear correlation coefficient (R) and Root Mean Square Error (RSME) between observed and predicted flows at daily, monthly, and annual time scales, in all fluviometric stations selected in this study.

Station	Coverage [%]	Daily		Monthly		Annual	
		R	RMSE [mm/day]	R	RMSE [mm/month]	R	RMSE [mm/year]
RRI	99.80	0.83	0.77	0.97	9.84	0.81	103.65
CCO	98.70	0.90	0.51	0.99	3.53	0.87	69.43
JTU	98.20	0.74	0.79	0.98	4.22	0.59	110.76
ECI	98.00	0.57	1.68	0.87	16.85	0.68	284.17
ETI	97.70	0.73	2.73	0.98	23.51	0.90	388.33
SOC	93.30	0.81	0.62	0.99	5.54	0.74	81.00
SAL	92.50	0.87	0.76	0.99	6.05	0.94	68.70
PSR	92.50	0.63	0.39	0.98	3.22	0.69	32.29
TAM	90.70	0.60	4.29	0.58	128.51	0.18	1471.02
LET	89.10	0.75	0.98	1.00	8.63	0.69	90.63
TAB	87.50	0.49	2.85	0.98	31.53	0.29	737.95
BOR	86.80	0.45	1.82	0.96	9.05	0.38	379.11
EKM	86.20	0.89	0.49	0.97	6.90	0.63	61.36
HNE	79.90	0.84	0.42	0.98	4.21	0.75	40.90
COA	74.90	0.83	0.60	0.98	4.87	0.75	67.09
RAM	72.70	0.88	0.27	0.98	3.02	0.73	32.52
ILA	72.70	0.71	0.61	0.95	10.69	0.65	100.79
BEL	71.30	0.40	3.30	0.91	43.94	0.23	947.91
PST	69.50	0.74	0.97	0.93	11.11	0.47	173.05
PIS	67.30	0.80	0.42	0.99	3.86	0.26	74.10
SDO	66.20	0.87	0.43	0.99	2.56	0.79	49.12
REQ	59.10	0.79	1.28	0.94	18.30	0.44	227.69
OBR	59.00	0.83	0.35	0.99	3.00	0.39	60.92
TOC	58.90	0.75	1.69	0.98	13.09	0.60	267.59
CHO	55.00	0.88	0.34	0.97	4.02	0.66	53.01
SRE	53.50	0.66	1.50	0.97	24.35	0.39	364.09
TMA	52.20	0.75	1.83	0.99	18.22	0.00	260.73
PSC	49.00	0.87	1.42	0.99	4.44	0.99	85.23
HUA	45.40	0.84	0.36	0.99	3.79	0.90	37.37
PUI	43.60	0.66	5.44	0.95	58.33	0.55	643.53
PUC	43.20	0.91	0.94	0.99	9.19	0.68	129.60
PIC	42.70	0.78	1.52	0.98	13.18	0.48	173.66
CHA	41.90	0.79	1.36	0.98	12.23	0.81	117.15
NAP	40.30	0.34	3.21	0.83	62.11	0.36	472.68
POC	34.10	0.79	0.41	0.99	1.70	0.78	32.73
MAL	31.00	0.84	0.49	0.96	6.12	0.46	71.48
PAU	29.40	0.84	0.89	0.99	6.97	0.39	110.19
PCU	26.90	0.87	0.48	0.99	3.36	0.77	56.63
LPA	24.30	0.84	2.16	0.99	19.25	0.60	265.99
LAG	23.60	0.81	1.51	0.97	14.55	0.82	55.81
LNA	18.40	0.74	0.92	0.93	10.95	0.91	53.32
PCQ	14.40	0.85	3.77	0.94	65.54	0.91	626.42
CUM	14.20	0.85	0.59	0.96	8.42	0.89	46.92

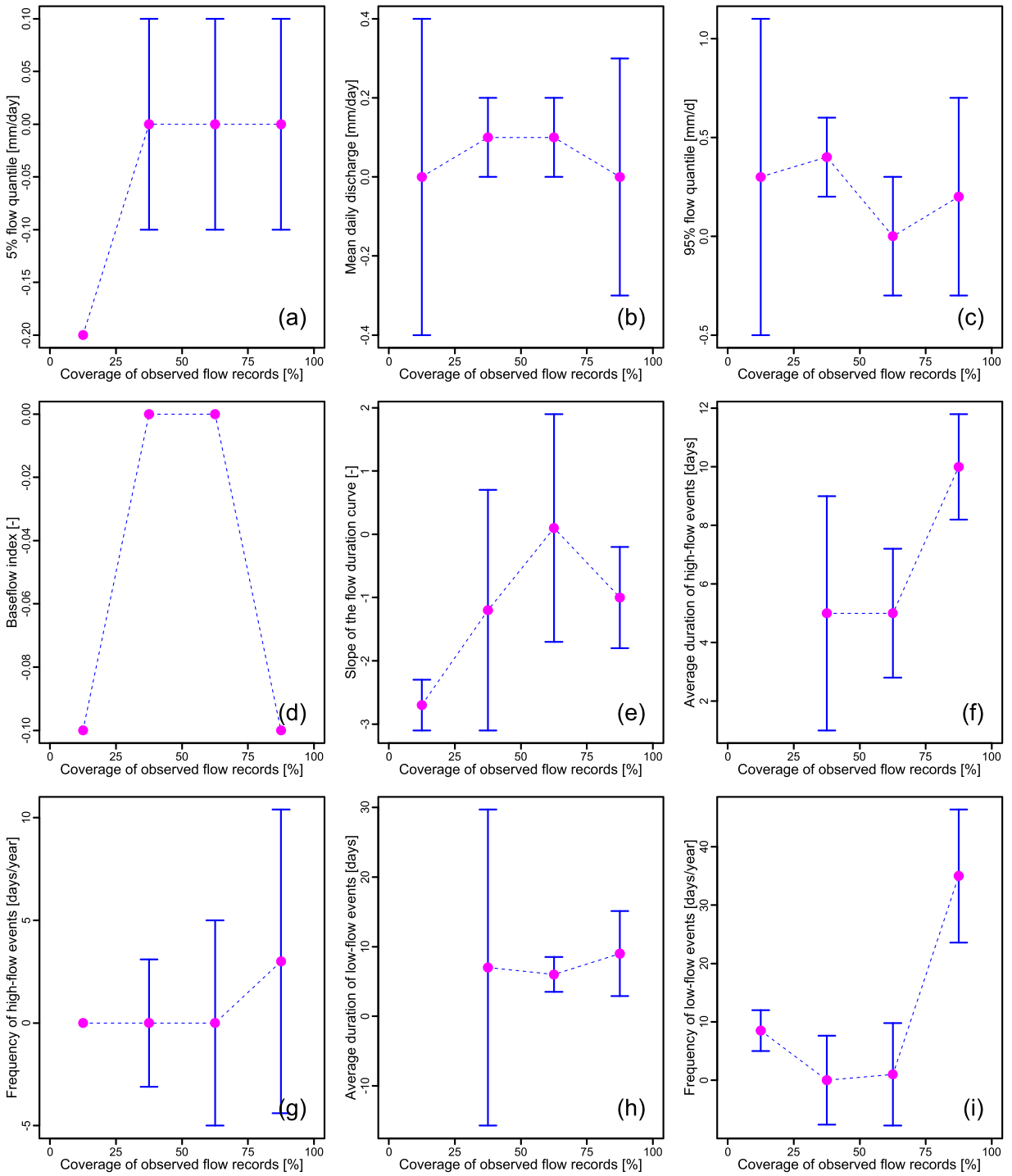


Fig. 8. (a-i) Standard error bars between nine predicted and observed hydrological signatures in gauged catchments. Standard error was computed in four intervals of observed flow record coverage (0–25%, 25–50%, 50–75%, and 75–100%) for 1981–2020. Points in magenta represent the median for each interval.

errors in duration and frequency of low and high flow events are positive (Fig. 8f-i). Errors in duration signatures tend to increase when data coverage grows while the inverse behavior occurs for frequency signatures.

Fig. 9 displays the qualitative classifications of national daily streamflow simulations in gauged catchments based on KGE (Thiemig et al., 2013) and NSE_{sqrt} (Moriasi et al., 2015) ranges. In terms of KGE ratings (Fig. 9a), with emphasis on high-middle flows (Mizukami et al., 2019), simulations are “good” (green areas) in catchments of the south, central-south, and northwest of the country. Ratings

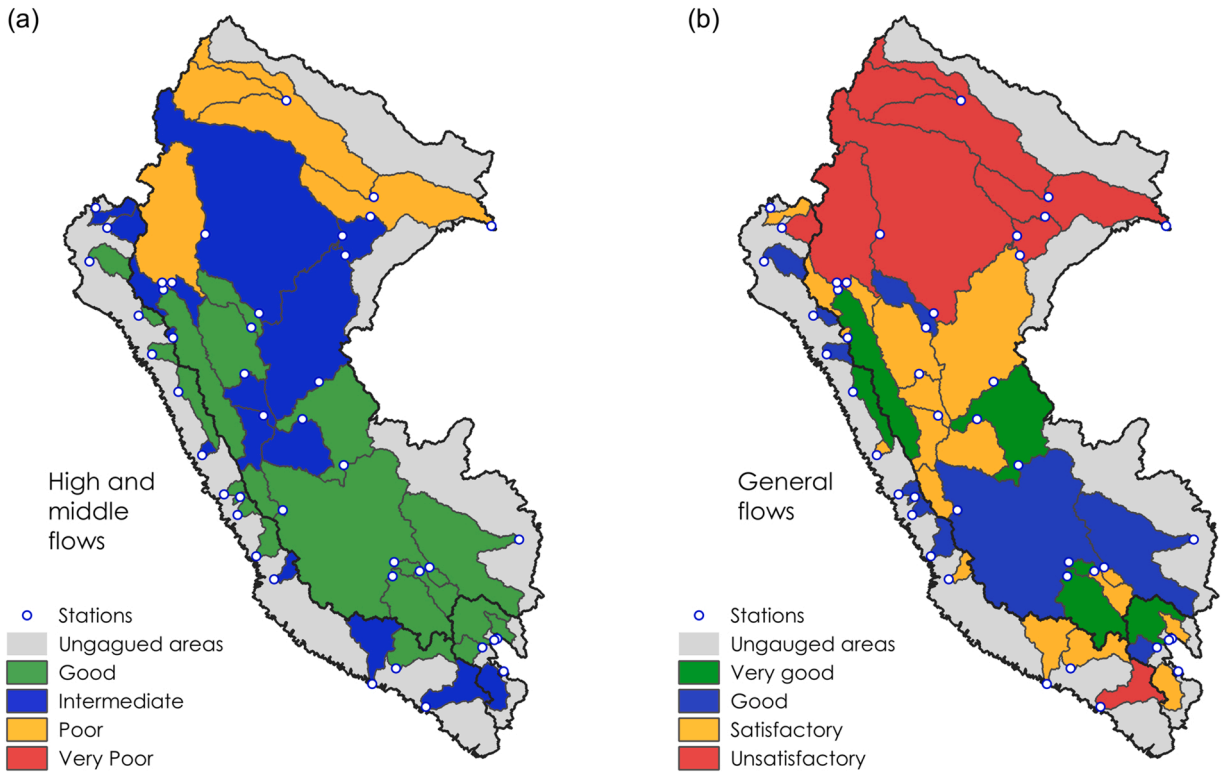


Fig. 9. Qualitative ratings of daily flow simulation on gauged catchments in the study domain based on **(a)** high and middle flows and **(b)** general flows performances in the national hydrological model.

decrease in the central-north ("intermediate", blue areas) and northeast ("poor", yellow areas). In terms of NSE_{sqrt} ratings (Fig. 9b), with emphasis on general flows (Chiew et al., 1993), simulations with "very good" (green areas) and "good" (blue areas) and "satisfactory" (yellow areas) ratings cover the majority of catchments. However, "unsatisfactory" (red areas) simulations in the country's northeast correspond to the Amazon lowlands.

This paper introduces a new hydrological dataset of simulated daily flow series in Peru for 11913 river reaches. This dataset named PISCO_HyD_ARNOVIC is available at https://www.hydroshare.org/resource/f723d6c762ca45b6936dd9489bc44842/. Flow series extraction for a specific river stream can be done by searching its COMID (users can find COMIDs at https://hllauca.github.io/map_product/Identificar_COMID.html). In an operational version (1981–present day), the PISCO_HyD_ARNOVIC dataset is now contributing to an improved understanding of past and current hydrological analysis. Furthermore, is used as crucial input for climate services such as the Monitoring and Forecasting of Potential Floods in the National Service of Meteorology and Hydrology of Peru (please visit: https://harold-llauca.shinyapps.io/sonics/).

4. Discussion

4.1. Findings in the similarity-based regionalization

Several procedures can be cited in the literature to perform stream flow prediction in an ungauged catchment. As the spatial proximity method has the optimal performance in catchments with a high density of gauges (Parajka et al., 2013), in data-scarcity domains physical similarity approach shows higher performance than other methods (Wang et al., 2021), so here we use a simple combination of both approaches (section 2.3) to classify donor and receptor sub-catchments.

The efficiency of regionalization is dependent on the regionalization method, number of catchments and catchment area. As regionalization performance tends to increase with an increasing number of sub-catchments, we use ~12 thousand sub-catchments across the study domain. Other studies have similarly found that the efficiency of regionalization increases significantly with increasing drainage area (Merz et al., 2009; Nester et al., 2011); hence, we do not consider drainage areas as regionalization descriptors because the small sub-catchments sizes (~120 km^2) could increase hydrological variability across the country. In that sense, we choose only the best global runoff predictors proposed by Beck et al., (2020, 2016, 2015), and used in a global application of the VIC+RAPID model (Lin et al., 2019).

The number of hydrological zones found here increased drastically compared with a previous regionalization study for the same domain (Llauca et al., 2021b). In this paper, hydrological zones (and sub-zone) boundaries are not continuous in space. That means

that sub-catchments from the same zone (and sub-zone) might not necessarily be proximal to neighbor catchments. Furthermore, the similarity-based regionalization approach is not restricted by gauges location as it is in a sensitivity analysis regionalization procedure (Bock et al., 2016). These differences help us to increase catchment variability representation across the country - 43 zones instead 14 presented in Llauca et al. (2021b) - especially in a context of limited or no access to in-situ networks.

We select just the best parameter set with the least dissimilarity index in a 300 km searching radius due to computational limitations and the small number of ungauged sub-catchments in the study domain (27% of the total area). However, this choice might increase regionalization uncertainties in ungauged catchments. For instance, Beck et al. (2015), proposed transferring parameters from the best ten donor candidates. In our case, cross-validation results show heterogeneous regionalization performances using the only best donor in the Pacific divide, so findings suggest that increasing the number of donor candidates could improve results. Future studies will raise the number of candidate donors and assess the impact on regionalization performance.

4.2. Limitations of flow simulation at the national level

Quantifying flows in Peru is a challenging task, given the large physiographic (Zevallos and Lavado-Casimiro, 2022), climatic (Son et al., 2020), and hydrological (Lavado Casimiro et al., 2012) heterogeneities across the territory. There are few recent large-scale hydrological modeling experiments in Peru providing simulated water discharges. Most of these have focused only on specific divides. For instance, (Asurza-Véliz and Lavado-Casimiro, 2020) and Rau et al. (2019) conduct daily and monthly models, respectively, for the Pacific drainage using regionalization techniques in a data-scarce context. Zubieta et al. (2021) and Satgé et al. (2019) assess climate change impacts and precipitation products in the Lake Titicaca catchment using GRs conceptual models. In the Peruvian Amazon Basin, local studies have performed hydrological models in the Andes-Amazon transition (Baltazar et al., 2023; Chancay and Espitia-Sarmiento, 2021) while large-sample hydrological applications have assessed the entire Amazon Basin using GHMs for detecting peak flows (Towner et al., 2019), and global datasets such as GRACE gravity satellites for runoff estimation (Chen et al., 2020). In that sense, the results present here build upon hydrological model simulations using the ARNO/VIC+RAPID modeling framework and PISCO dataset set as forcing inputs showing a spatial performance variability similar to reported in those studies.

The meteorological data are the critical input for hydrological modeling. Singh and Reza Najafi (2020) depict that the propagation of gridded climate datasets biases has strong implications in the simulation of extreme events. Thus, the unsatisfactory modeling results found in northeastern Peru may reflect two main issues: first, the significant uncertainty of spatial rainfall distribution in the Marañón, Ucayali, and Huallaga catchments reported by Fernandez-Palomino et al. (2021) and Zubieta et al. (2017), and second, higher $PISCO_P$ (unstable version) uncertainties due to the lack of rainfall estimates in equatorial regions (Aybar et al., 2020). Furthermore, as evapotranspiration plays a crucial role in hydrological processes in the Amazon catchment (de Oliveira et al., 2021), daily PET climatologies might be increasing forcing uncertainties and water balance in the Amazon lowlands.

Overall, we demonstrate the feasibility to implement a sub-catchment hybrid hydrological modeling framework (ARNO/VIC+RAPID) for generating flow simulations in Peru, similar to national hydrological applications in other study domains using physical-based and conceptual models such as S-HYPE (Girons Lopez et al., 2021), WEP-CN (Liu et al., 2020), TopNet (McMillan et al., 2016), and HBV (Veijalainen et al., 2010). In addition, because large-scale hydrological modeling is complex in data-scarce sites, the following assumptions might be impacts in model results: a) streamflow gauges with shorter records might influence model calibration due to insufficient information available for discharge uncertainty distribution estimation (Westerberg et al., 2020); thus this work incorporate a flow data coverage assessment; b) we estimate naturalized river flows across the country even in the Pacific divide which has the largest population and major infrastructure of water regulation (Stensrud, 2019); c) we conduct a mono-objective function with emphasis on high-middle flow calibration (Mizukami et al., 2019) used in national water balance applications such as Vásquez et al. (2021); d) due to floodplain playing an essential role in the Amazon basin flow routing (Wongchuig Correa et al., 2017), uncertainties in channel geometry information could affect simulations in the Atlantic drainage more than Pacific and Titicaca slopes.

5. Conclusions

In this study, a hybrid modeling framework and a similarity-based regionalization approach were described for running a national hydrological model across an extensive river network. The methodology proposed allows us to build daily simulations of flow series in the Peruvian catchments. The national hydrological model performance is good in most parts of the country and simulations follow observed daily flow series, especially for catchments located on the Pacific coast and in the Andes-Amazon transition of Peru. Moreover, the model can represent streamflow seasonal and interannual variability and main hydrological signatures across the country.

Low model performance was observed in northeastern Peru (Amazon lowlands), where simulations tend to overestimate the observed water discharges. This might be related to forcing data uncertainties due to the low density of ground meteorological observations in the Amazon basin and biases in operational gridded meteorological products. Also, the limited number of fluviometric stations in the Amazon slope could be increasing parameter, states, and outputs uncertainties, due to error propagation from head basins to downstream in the Andes Amazon transition. Hydrological regionalization proved to be a useful tool to estimate daily streamflow series in ungauged catchments and to deal with the problem of data scarcity in Peru. Future studies will incorporate different conceptual model structures and donor candidates.

Ongoing research focuses on the implementation of a national hydrological model for Peru. The main application is the generation of a new operational hydrological dataset of daily flow series from 01 January 1981 to the present day. This dataset will also be used to include future risk scenario simulations in often poorly gauged catchments, climate change impacts on water resources and the

implementation of climate services in Peru such as the Monitoring and Forecasting of Potential Floods in the National Service of Meteorology and Hydrology of Peru.

CRediT authorship contribution statement

Harold Llauca: Conceptualization, Data curation, Methodology, Software, Validation, Visualization, Writing – original draft preparation, Writing – review & editing, **Karen Leon:** Conceptualization, Data curation, Methodology, Software. **Waldo Lavado:** Conceptualization, Methodology, Resources, Supervision.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Harold Llauca reports administrative support, article publishing charges, and writing assistance were provided by Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). Harold Llauca reports a relationship with Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) that includes: employment.

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2023.101381.

References

Addor, N., Melsen, L.A., 2019. Legacy, rather than adequacy, drives the selection of hydrological models. Water Resour. Res. 55, 378–390. https://doi.org/10.1029/2018wr022958.

Addor, N., Nearing, G., Prieto, C., Newman, A., Le Vine, N., Clark, M., 2018. Selection of hydrological signatures for large-sample hydrology. EarthArXiv. https://doi.org/10.31223/osf.io/2em53.

Asurza-Véliz, F.A., Lavado-Casimiro, W.S., 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., Fernández, C., Huerta, A., Lavado, W., Vega, F., Felipe-Obando, O., 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.

Baltazar, L.A., Viola, M.R., Rogério de Mello, C., Junqueira, R., da Silva Amorim, J., 2023. Hydrological modeling in a region with sparsely observed data in the eastern Central Andes of Peru, Amazon. J. South Am. Earth Sci. 121, 104151 https://doi.org/10.1016/j.jsames.2022.104151.

Barbarossa, V., Huijbregts, M.A.J., Beusen, A.H.W., Beck, H.E., King, H., Schipper, A.M., 2018. FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015. Sci. Data 5, 180052. https://doi.org/10.1038/sdata.2018.52.

Beck, H.E., de Roo, A., van Dijk, A.I.J.M., 2015. Global maps of streamflow characteristics based on observations from several thousand catchments. J. Hydrometeorol. 16, 1478–1501. https://doi.org/10.1175/JHM-D-14-0155.1.

Beck, H.E., van Dijk, A.I.J.M., de Roo, A., Miralles, D.G., McVicar, T.R., Schellekens, J., Bruijnzeel, L.A., 2016. Global-scale regionalization of hydrologic model parameters. Water Resour. Res. 52, 3599–3622. https://doi.org/10.1002/2015wr018247.

Beck, H.E., Pan, M., Lin, P., Seibert, J., Dijk, A.I.J.M., Wood, E.F., 2020. Global fully-distributed parameter regionalization based on observed streamflow from 4229 headwater catchments. J. Geophys. Res. D: Atmos. https://doi.org/10.1029/2019JD031485.

Bock, A.R., Hay, L.E., McCabe, G.J., Markstrom, S.L., Atkinson, R.D., 2016. Parameter regionalization of a monthly water balance model for the conterminous United States. Hydrol. Earth Syst. Sci. 20, 2861–2876. https://doi.org/10.5194/hess-20-2861-2016.

Chancay, J.E., Espitia-Sarmiento, E.F., 2021. Improving hourly precipitation estimates for flash flood modeling in data-scarce Andean-Amazon basins: An integrative framework based on machine learning and multiple remotely sensed data. Remote Sens. (Basel) 13, 4446. https://doi.org/10.3390/rs13214446.

Chen, J., Tapley, B., Rodell, M., Seo, K.-W., Wilson, C., Scanlon, B.R., Pokhrel, Y., 2020. Basin-scale river runoff estimation from GRACE gravity satellites, climate models, and in situ observations: a case study in the Amazon basin. Water Resour. Res 56. https://doi.org/10.1029/2020wr028032.

Chiew, F.H.S., Stewardson, M.J., McMahon, T.A., 1993. Comparison of six rainfall-runoff modelling approaches. J. Hydrol. 147, 1–36. https://doi.org/10.1016/0022-1694(93)90073-I.

Clark, M.P., Slater, A.G., Rupp, D.E., Woods, R.A., Vrugt, J.A., Gupta, H.V., Wagener, T., Hay, L.E., 2008. Framework for Understanding Structural Errors (FUSE): a modular framework to diagnose differences between hydrological models. Water Resour. Res. Water Sci. Appl. 44, 2135. https://doi.org/10.1029/2007WR006735.

David, C.H., Habets, F., Maidment, D.R., Yang, Z.-L., 2011a. RAPID applied to the SIM-France model. Hydrol. Process. 25, 3412–3425. https://doi.org/10.1002/hyp.8070.

David, C.H., Maidment, D.R., Niu, G.-Y., Yang, Z.-L., Habets, F., Eijkhout, V., 2011b. River network routing on the NHDplus dataset. J. Hydrometeorol. 12, 913–934. https://doi.org/10.1175/2011JHM1345.1.

David, C.H., Famiglietti, J.S., Yang, Z.-L., Eijkhout, V., 2015. Enhanced fixed-size parallel speedup with the Muskingum method using a trans-boundary approach and a large subbasins approximation. Water Resour. Res. 51, 7547–7571. https://doi.org/10.1002/2014wr016650.

Dolan, F., Lamontagne, J., Link, R., Hejazi, M., Reed, P., Edmonds, J., 2021. Evaluating the economic impact of water scarcity in a changing world. Nat. Commun. 12, 1915. https://doi.org/10.1038/s41467-021-22194-0.

Drenkhan, F., Carey, M., Huggel, C., Seidel, J., Oré, M.T., 2015. The changing water cycle: climatic and socioeconomic drivers of water-related changes in the Andes of Peru. Wiley Interdiscip. Rev.: Water 2, 715–733. https://doi.org/10.1002/wat2.1105.

Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76, 501–521. https://doi.org/10.1007/BF00939380.

Fernandez-Palomino, C.A., Hattermann, F.F., Krysanova, V., Lobanova, A., Vega-Jácome, F., Lavado, W., Santini, W., Aybar, C., Bronstert, A., 2021. A novel high-resolution gridded precipitation dataset for Peruvian and Ecuadorian watersheds – development and hydrological evaluation. J. Hydrometeorol. 1 https://doi.org/10.1175/jhm-d-20-0285.1.

Ferreira, P.M., de, L., Paz, A.R., da, Bravo, J.M., 2020. Objective functions used as performance metrics for hydrological models: state-of-the-art and critical analysis. RBRH 25. https://doi.org/10.1590/2318-0331.252020190155.

Follum, M.L., Tavakoly, A.A., Niemann, J.D., Snow, A.D., 2017. AutoRAPID: a model for prompt streamflow estimation and flood inundation mapping over regional to continental extents. J. Am. Water Resour. Assoc. 53, 280–299. https://doi.org/10.1111/1752-1688.12476.

Garreaud, R.D., Vuille, M., Compagnucci, R., Marengo, J., 2009. Present-day South American climate. Palaeogeogr. Palaeoclimatol. Palaeoecol. 281, 180–195. https://doi.org/10.1016/j.palaeo.2007.10.032.

Girons Lopez, M., Crochemore, L., Pechlivanidis, I.G., 2021. Benchmarking an operational hydrological model for providing seasonal forecasts in Sweden. Hydrol. Earth Syst. Sci. 25, 1189–1209. https://doi.org/10.5194/hess-25-1189-2021.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853. https://doi.org/10.1126/science.1244693.

Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from ambient air temperature. Chic. Appl. Eng. Agric. 1, 96–99. https://doi.org/10.13031/2013.26773.

Hattermann, F.F., Krysanova, V., Gosling, S.N., Dankers, R., Daggupati, P., Donnelly, C., Flörke, M., Huang, S., Motovilov, Y., Buda, S., Yang, T., Müller, C., Leng, G., Tang, Q., Portmann, F.T., Hagemann, S., Gerten, D., Wada, Y., Masaki, Y., Alemayehu, T., Satoh, Y., Samaniego, L., 2017. Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins. Clim. Change 141, 561–576. https://doi.org/10.1007/s10584-016-1829-4.

Huggel, C., Raissig, A., Rohrer, M., Romero, G., Diaz, A., Salzmann, N., 2015. How useful and reliable are disaster databases in the context of climate and global change? a comparative case study analysis in Peru. Nat. Hazards Earth Syst. Sci. 15, 475–485. https://doi.org/10.5167/uzh-118125.

Janse, J.H., van Dam, A.A., Hes, E.M.A., de Klein, J.J.M., Finlayson, C.M., Janssen, A.B.G., van Wijk, D., Mooij, W.M., Verhoeven, J.T.A., 2019. Towards a global model for wetlands ecosystem services. Curr. Opin. Environ. Sustain. 36, 11–19. https://doi.org/10.1016/j.cosust.2018.09.002.

Kay, A.L., Rudd, A.C., Fry, M., Nash, G., Allen, S., 2021. Climate change impacts on peak river flows: combining national-scale hydrological modelling and probabilistic projections. Clim. Risk Manag. 31, 100263 https://doi.org/10.1016/j.crm.2020.100263.

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., Nearing, G., 2019. Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. Hydrol. Earth Syst. Sci. 23, 5089–5110. https://doi.org/10.5194/hess-23-5089-2019.

Lane, R.A., Coxon, G., Freer, J.E., Wagener, T., Johnes, P.J., Bloomfield, J.P., Greene, S., Macleod, C.J.A., Reaney, S.M., 2019. Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain. Hydrol. Earth Syst. Sci. 23, 4011–4032. https://doi.org/10.5194/hess-23-4011-2019.

Lavado Casimiro, W.S., Labat, D., Guyot, J.L., Ardoin-Bardin, S., 2011. Assessment of climate change impacts on the hydrology of the Peruvian Amazon–Andes basin. Hydrol. Process. 25, 3721–3734.

Lavado Casimiro, W.S., Ronchail, J., Labat, D., Espinoza, J.C., Guyot, J.L., 2012. Basin-scale analysis of rainfall and runoff in Peru (1969–2004): Pacific, Titicaca and Amazonas drainages. Hydrol. Sci. J. 57, 625–642. https://doi.org/10.1080/02626667.2012.672985.

Lavado-Casimiro, Felipe, 2013. ENSO impact on hydrology in Peru. Adv. Eng. Educ.

Lavers, D.A., Harrigan, S., Andersson, E., Richardson, D.S., Prudhomme, C., Pappenberger, F., 2019. A vision for improving global flood forecasting. Environ. Res. Lett. 14, 121002 https://doi.org/10.1088/1748-9326/ab52b2.

Lehner, B., Verdin, K., Jarvis, A., 2008. New global hydrography derived from spaceborne elevation data. Eos 89, 93. https://doi.org/10.1029/2008eo100001.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res 99, 14415. https://doi.org/10.1029/94jd00483.

Lin, P., Yang, Z.-L., Cai, X., David, C.H., 2015. Development and evaluation of a physically-based lake level model for water resource management: a case study for Lake Buchanan, Texas. J. Hydrol.: Reg. Stud. 4, 661–674. https://doi.org/10.1016/j.ejrh.2015.08.005.

Lin, P., Pan, M., Beck, H.E., Yang, Y., Yamazaki, D., Frasson, R., David, C.H., Durand, M., Pavelsky, T.M., Allen, G.H., Gleason, C.J., Wood, E.F., 2019. Global reconstruction of naturalized river flows at 2.94 million reaches. Water Resour. Res. 55, 6499–6516. https://doi.org/10.1029/2019WR025287.

Liu, H., Jia, Y., Niu, C., Su, H., Wang, J., Du, J., Khaki, M., Hu, P., Liu, J., 2020. Development and validation of a physically-based, national-scale hydrological model in China. J. Hydrol. 590, 125431 https://doi.org/10.1016/j.jhydrol.2020.125431.

Llauca, H., Lavado-Casimiro, W., León, K., Jimenez, J., Traverso, K., Rau, P., 2021a. Assessing near real-time satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in Peruvian Andes. Remote Sens. 13, 826. https://doi.org/10.3390/rs13040826.

Llauca, H., Lavado-Casimiro, W., Montesinos, C., Santini, W., Rau, P., 2021b. PISCO_HyM_GR2M: a model of monthly water balance in Peru (1981–2020. Water. https://doi.org/10.3390/w13081048.

McMillan, H., Westerberg, I., Branger, F., 2017. Five guidelines for selecting hydrological signatures. Hydrol. Process. 31, 4757–4761. https://doi.org/10.1002/hyp.11300.

McMillan, H.K., Booker, D.J., Cattoën, C., 2016. Validation of a national hydrological model. J. Hydrol. 541, 800–815. https://doi.org/10.1016/j.jhydrol.2016.07.043.

Merz, R., Parajka, J., Blöschl, G., 2009. Scale effects in conceptual hydrological modeling. Water Resour. Res. https://doi.org/10.1029/2009WR007872.

Mizukami, N., Clark, M.P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., Viger, R.J., Markstrom, S.L., Hay, L.E., Woods, R., Arnold, J.R., Brekke, L.D., 2016. mizuRoute version 1: a river network routing tool for a continental domain water resources applications. Geosci. Model Dev. 9, 2223–2238. https://doi.org/10.5194/gmd-9-2223-2016.

Mizukami, N., Rakovec, O., Newman, A.J., Clark, M.P., Wood, A.W., Gupta, H.V., Kumar, R., 2019. On the choice of calibration metrics for “high-flow” estimation using hydrologic models. Hydrol. Earth Syst. Sci. 23, 2601–2614. https://doi.org/10.5194/hess-23-2601-2019.

Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and water quality models: performance measures and evaluation criteria. Trans. ASABE 58, 1763–1785. https://doi.org/10.13031/trans.58.10715.

Mouelhi, S., Michel, C., Perrin, C., Andréassian, V., 2006. Stepwise development of a two-parameter monthly water balance model. J. Hydrol. 318, 200–214. https://doi.org/10.1016/j.jhydrol.2005.06.014.

Narbondo, S., Gorgoglione, A., Crisci, M., Chreties, C., 2020. Enhancing physical similarity approach to predict runoff in ungauged watersheds in sub-tropical regions. Water 12, 528. https://doi.org/10.3390/w12020528.

Nester, T., Kirnbauer, R., Gutknecht, D., Blöschl, G., 2011. Climate and catchment controls on the performance of regional flood simulations. J. Hydrol. 402, 340–356. https://doi.org/10.1016/j.jhydrol.2011.03.028.

de Oliveira, R.G., Valle Júnior, L.C.G., da Silva, J.B., Espíndola, D.A.L.F., Lopes, R.D., Nogueira, J.S., Curado, L.F.A., Rodrigues, T.R., 2021. Temporal trend changes in reference evapotranspiration contrasting different land uses in southern Amazon basin. Agric. Water Manag. 250, 106815 https://doi.org/10.1016/j.agwat.2021.106815.

Pagliero, L., Bouraoui, F., Willems, P., Diels, J., 2014. Large-scale hydrological simulations using the soil water assessment tool, protocol development, and application in the danube basin. J. Environ. Qual. 43, 145–154. https://doi.org/10.2134/jeq2011.0359.

Pagliero, L., Bouraoui, F., Diels, J., Willems, P., McIntyre, N., 2019. Investigating regionalization techniques for large-scale hydrological modelling. J. Hydrol. 570, 220–235. https://doi.org/10.1016/j.jhydrol.2018.12.071.

Parajka, J., Merz, R., Blöschl, G., 2005. A comparison of regionalisation methods for catchment model parameters. Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-9-157-2005.
Parajka, J., Viglione, A., Rogger, M., Salinas, J.L., Sivapalan, M., Blöschl, G., 2013. Comparative assessment of predictions in ungauged basins – Part 1: Runoff-hydrograph studies. Hydrol. Earth Syst. Sci. 17, 1783–1795. https://doi.org/10.5194/hess-17-1783-2013.
Piccolroaz, S., Lazzaro, M.D., Zarlenga, A., Majone, B., Bellin, A., Fiori, A., 2016. HYPERstream: a multi-scale framework for streamflow routing in large-scale hydrological model. Hydrol. Earth Syst. Sci. 20, 2047–2061. https://doi.org/10.5194/hess-20-2047-2016.
Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7, 217–240. https://doi.org/10.5194/soil-7-217-2021.
Rau, P., Bourrel, L., Labat, D., Ruelland, D., Frappart, F., Lavado, W., Dewitte, B., Felipe, O., 2019. Assessing multidecadal runoff (1970-2010) using regional hydrological modelling under data and water scarcity conditions in Peruvian Pacific catchments. Hydrol. Process. 33, 20–35. https://doi.org/10.1002/hyp.13318.
Rodriguez, D.A., Tomasella, J., 2016. On the ability of large-scale hydrological models to simulate land use and land cover change impacts in Amazonian basins. Hydrol. Sci. J. 61, 1831–1846. https://doi.org/10.1080/02626667.2015.1051979.
Saavedra, D., Mendoza, P.A., Addor, N., Llauca, H., Vargas, X., 2021. A multi-objective approach to select hydrological models and constrain structural uncertainties for climate impact assessments. Hydrol. Process. https://doi.org/10.1002/hyp.14446.
Salas, F.R., Somos-Valenzuela, M.A., Dugger, A., Maidment, D.R., Gochis, D.J., David, C.H., Yu, W., Ding, D., Clark, E.P., Noman, N., 2018. Towards real-time continental scale streamflow simulation in continuous and discrete space. J. Am. Water Resour. Assoc. 54, 7–27. https://doi.org/10.1111/1752-1688.12586.
Sanchez Lozano, J., Romero Bustamante, G., Hales, R.C., Nelson, E.J., Williams, G.P., Ames, D.P., Jones, N.L., 2021. A streamflow bias correction and performance evaluation web application for GEOGloWS ECMWF streamflow services. Hydrology 8, 71. https://doi.org/10.3390/hydrology8020071.
Satgé, F., Ruelland, D., Bonnet, M.P., 2019. … precipitation products in space and over time compared with gauge observations and snow- hydrological modelling in the Lake Titicaca …. Hydrol. Earth Syst. Sci.
Singh, H., Reza Najafi, M., 2020. Evaluation of gridded climate datasets over Canada using univariate and bivariate approaches: Implications for hydrological modelling. J. Hydrol. 584, 124673 https://doi.org/10.1016/j.jhydrol.2020.124673.
Son, R., Wang, S.-Y.S., Tseng, W.-L., Barreto Schuler, C.W., Becker, E., Yoon, J.-H., 2020. Climate diagnostics of the extreme floods in Peru during early 2017. Clim. Dyn. 54, 935–945. https://doi.org/10.1007/s00382-019-05038-y.
Song, Z., James, L.D., 1992. An objectwe test for hydrologic scale. J. Am. Water Resour. Assoc. 28, 833–844. https://doi.org/10.1111/j.1752-1688.1992.tb03185.x.
Sood, A., Smakhtin, V., 2015. Global hydrological models: a review. Hydrol. Sci. J. 60, 549–565. https://doi.org/10.1080/02626667.2014.950580.
Stensrud, A.B., 2019. The social embeddedness of hydraulic engineers in the regulation of water and infrastructure in Peru. Environ. Plan. C: Polit. Space 37, 1235–1251. https://doi.org/10.1177/2399654419866835.
Tang, Q., Gao, H., Lu, H., Lettenmaier, D.P., 2009. Remote sensing: hydrology. Prog. Phys. Geogr.: Earth Environ. 33, 490–509. https://doi.org/10.1177/0309133309346650.
Tavakoly, A.A., Snow, A.D., David, C.H., Follum, M.L., Maidment, D.R., Yang, Z.-L., 2017. Continental-scale river flow modeling of the Mississippi river basin using high-resolution NHDPlusDataset. J. Am. Water Resour. Assoc. 53, 258–279. https://doi.org/10.1111/1752-1688.12456.
Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., De Roo, A., 2013. Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. J. Hydrol. 499, 324–338. https://doi.org/10.1016/j.jhydrol.2013.07.012.
Todini, E., 1996. The ARNO rainfall—runoff model. J. Hydrol. 175, 339–382. https://doi.org/10.1016/S0022-1694(96)80016-3.
Towner, J., Cloke, H.L., Zsoter, E., Flamig, Z., Hoch, J.M., Bazo, J., Coughlan de Perez, E., Stephens, E.M., 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.
Tyralis, H., Papacharalampous, G., Langousis, A., Papalexiou, S.M., 2021. Explanation and probabilistic prediction of hydrological signatures with statistical boosting algorithms. Remote Sens. 13, 333. https://doi.org/10.3390/rs13030333.
Vásquez, N., Cepeda, J., Gómez, T., Mendoza, P.A., Lagos, M., Boisier, J.P., Álvarez-Garretón, C., Vargas, X., 2021. Catchment-Scale Natural Water Balance in Chile. In: Fernández, B., Gironás, J. (Eds.), Water Resources of Chile. Springer International Publishing, Cham, pp. 189–208. https://doi.org/10.1007/978-3-030-56901-3_9.
Veijalainen, N., Lotsari, E., Alho, P., Vehviläinen, B., Käyhkö, J., 2010. National scale assessment of climate change impacts on flooding in Finland. J. Hydrol. 391, 333–350. https://doi.org/10.1016/j.jhydrol.2010.07.035.
Vereecken, H., Huisman, J.A., Bogena, H., Vanderborght, J., Vrugt, J.A., Hopmans, J.W., 2008. On the value of soil moisture measurements in vadose zone hydrology: A review. Water Resour. Res 44. https://doi.org/10.1029/2008wr006829.
Vitolo, C., Wells, P., Dobias, M., Buytaert, W., 2016. fuse: An R package for ensemble hydrological modelling. J. Open Source Softw. 1, 52.
Wang, H., Cao, L., Feng, R., 2021. Hydrological Similarity-Based Parameter Regionalization under Different Climate and Underlying Surfaces in Ungauged Basins. Water.
Westerberg, I.K., Sikorska-Senoner, A.E., Viviroli, D., Vis, M., Seibert, J., 2020. Hydrological model calibration with uncertain discharge data. Hydrol. Sci. J. 1–16. https://doi.org/10.1080/02626667.2020.1735638.
Wongchuig Correa, S., Paiva, R.C.D., de, Espinoza, J.C., Collischonn, W., 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.
Zevallos, J., Lavado-Casimiro, W., 2022. Climate change impact on Peruvian Biomes. . Trees Livelihoods 13, 238. https://doi.org/10.3390/f13020238.
Zink, M., Kumar, R., Cuntz, M., Samaniego, L., 2017. A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty. Hydrol. Earth Syst. Sci. 21, 1769–1790. https://doi.org/10.5194/hess-21-1769-2017.
Zubieta, R., Getirana, A., Espinoza, J.C., Lavado-Casimiro, W., Aragon, L., 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.
Zubieta, R., Molina-Carpio, J., Laqui, W., Sulca, J., Ilbay, M., 2021. Comparative analysis of climate change impacts on meteorological, hydrological, and agricultural droughts in the Lake Titicaca Basin. Water 13, 175. https://doi.org/10.3390/w13020175.