Journal of Hydrology: Regional Studies 37 (2021) 100932

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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

Comparing model complexity for glacio-hydrological simulation in the data-scarce Peruvian Andes

Randy Muñoz [a,*], Christian Huggel [a], Fabian Drenkhan [a,b,c], Marc Vis [a], Daniel Viviroli [a]

[a] Department of Geography, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland
[b] Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, United Kingdom
[c] Departamento de Humanidades, Pontificia Universidad Católica del Perú, Lima, Peru

A R T I C L E I N F O

Keywords:
Glacio-hydrological simulation
Tropical Andes
Shaman model
Scarce data
Model complexity
Multi-model comparison

A B S T R A C T

Study region: Glaciated headwaters of the Vilcanota-Urubamba river basin, Southern Peru
Study focus: A pivotal question is if robust hydrological simulation of streamflow in data-scarce and glaciated catchments can be achieved using parsimonious or more complex models. Therefore, a multi-model assessment of three glacio-hydrological models of different complexity was conducted thoroughly analyzing model performance, flow signatures and runoff components.
New hydrological insights for the region: In data-scarce catchments, such as in the tropical Andes, parsimonious glacio-hydrological models can provide more robust results than complex models. While the overall performance of all models was reasonably good (R^2: 0.65–0.70, Nash-Sutcliffe: 0.65–0.73, Nash-Sutcliffe-ln: 0.73–0.78), with increasing data scarcity more complex models involve higher uncertainties. Furthermore, complex models require substantial understanding of the underpinning hydrological processes and a comprehensive calibration strategy to avoid apparently high model performance driven by inadequate assumptions. Based on these insights we present a framework for robust glacio-hydrological simulation under data scarcity. This stepwise approach includes, among others, a multi-model focus with a comprehensive assessment of flow signatures and runoff components. Future modeling needs to be further supported by alternative data collection strategies to substantially improve knowledge and process understanding. Therefore, the extension of sensor and station networks combined with the integration of co-produced knowledge represents a meaningful measure to robust decision-making for climate change adaptation and water management under high uncertainty.

1. Introduction

Hydrological models are a common tool used to understand the water cycle and to assess the possible impacts of water management policies (Hrachowitz and Clark, 2017). In the last decades there has been a trend to develop models of higher complexity (Perrin et al., 2001) as our understanding of hydrological processes, globally available gridded datasets and computing power have increased. Furthermore, wide availability of free software and modeling packages in combination with powerful automatic calibration tools have

* Correspondence to: Environmental and Climate: Impacts, Risks and Adaptation (EClim), Glaciology and Geomorphodynamics Group (3G), Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland.

E-mail address: randy.munoz@geo.uzh.ch (R. Muñoz).

https://doi.org/10.1016/j.ejrh.2021.100932
Received 20 January 2021; Received in revised form 18 September 2021; Accepted 20 September 2021
Available online 1 October 2021

facilitated the use of complex models even to non-expert users.

However, multi-model comparisons have revealed that more complex models do not necessarily provide more consistent results than parsimonious models (Caldwell et al., 2015; Jiang et al., 2007; Perrin et al., 2001; Tegegne et al., 2017). An inadequate model complexity could lead to model over-parametrization (Perrin et al., 2001), and in particular a higher complexity could involve a more difficult parameter estimation, higher uncertainty, and higher data requirement (Li et al., 2015).

Most of multi-model comparisons were carried out by using common measures of model performance, such as the Nash-Sutcliffe efficiency or the coefficient of determination. Nevertheless, these indicators might be affected by the limitations of a proper assessment of the model performance (Krause et al., 2005). Models that perform similarly during the calibration process might provide distinct simulations under different conditions (Jiang et al., 2007), e.g. driven by climate change, often related to the model structure and calibration strategy (Caldwell et al., 2015). A good model curve-fitting may not capture the underpinning dominant hydrological processes (cf. Hrachowitz and Clark, 2017) and could thus neglect high uncertainties. A more appropriate calibration strategy should therefore consider different aspects of the fit between measured and simulated runoff (Vansteenkiste et al., 2014) by using multiple variables, model states, and hydrological signatures to comprehensively understand the respective hydrographs (Hrachowitz and Clark, 2017).

In the tropical Andes, a considerable amount of research efforts have increased the understanding of the dominant hydrological processes under present and future conditions, and possible impacts (e.g. Andres et al., 2014; Buytaert et al., 2006; Drenkhan et al., 2019; Frans et al., 2015; Guido et al., 2016; Morán-Tejeda et al., 2018; Ochoa-Tocachi et al., 2016; Rodríguez-Morales et al., 2019). Strong precipitation seasonality is one of the most relevant characteristics in the region where almost 80% of total annual precipitation falls between October and April (c.f. Baraer et al., 2012), while during the dry season from June to August glacier runoff contributes a major part of total discharge in some catchments, buffering water requirement for different water demands (c.f. Buytaert et al., 2017; Condom et al., 2012; Drenkhan et al., 2019; Frans et al., 2015). Compared to mid-latitudinal glaciers, tropical glaciers involve particular features and processes, such as when and how melt occurs, which has been described for the Cordillera Blanca (e.g. Kaser and Georges, 1997; Vuille et al., 2008) and Cordillera Vilcanota in Peru (e.g. Hastenrath, 1978; Suarez et al., 2015; Yarleque et al., 2018) and Cordillera Real in Bolivia (e.g. Sicart et al., 2011; Wagnon et al., 1999). Tropical Andean glaciers are exposed to a low thermal seasonality which leads to continuous ablation and a constant 0 °C level during the whole year (Kaser, 2001). Particularly in outer tropical environments such as Peru, precipitation and humidity play a major role controlling sublimation and glacier melt (Kronenberg et al., 2019; Sicart et al., 2008). In consequence, the widely used temperature-index models (e.g. Carenzo et al., 2016; Hock, 2003) have limitations when estimating glacier melt (Sicart et al., 2011), while more appropriate sophisticated energy balance models (e.g. Fyffe et al., 2014; Ragettli and Pellicciotti, 2012) with high data requirement (Juen et al., 2007) are of limited usefulness when hampered by data scarcity, and hence have only been applied so far in a few highly instrumented catchments.

Most of the peer-reviewed glacio-hydrological simulations on a local scale have been conducted in Asia, Europe, and North America (c.f. Tiel et al., 2020). Furthermore, model comparisons with focus on data-scarce regions are not common (e.g. Bai et al., 2017; Ragettli et al., 2014; Tegegne et al., 2017). Such a context of limited data availability, regional experience and process understanding challenges the implementation of glacio-hydrological models in the tropical Andes. There most glacio-hydrological simulations have been conducted in Peru (e.g. Condom et al., 2012), and some in Bolivia (e.g. Soruco et al., 2015) and Ecuador (e.g. Pouget et al., 2017). Lumped models are the most commonly used ranging from simple linear equations (e.g. Mark and Seltzer, 2003) to more complex estimations that consider detailed hydroclimatic variables like glacier mass balance or sublimation (e.g. Soruco et al., 2015). No lumped model seems to be preferred for simulations. Semi-distributed models are also widely applied in the region mostly using the WEAP (Condom et al., 2012; Yates et al., 2005) and RS Minerve (García Hernández et al., 2016) software. WEAP allows simulation under different model structures, while RS Minerve simulates runoff by using several conceptual models of different complexity such as HBV (Bergström, 1976) and GR4J (Perrin et al., 2003). Lumped models are particularly suitable in the data-scarce tropical Andes as they require a limited data input while semi-distributed models additionally provide information about the water cycle such as detailed runoff components.

Despite these efforts there is only a limited number and scope of hydrological modeling research available for glacierized tropical Andean catchments due to limited access, complex hydroclimatic processes, social tensions, and low institutional capacities (Drenkhan et al., 2015; Schoolmeester et al., 2018; Vuille et al., 2018). To our knowledge, there is no peer-reviewed study available for the tropical Andes implementing fully distributed models. Furthermore, glacio-hydrological multi-model assessments with strong emphasis on comprehensive model result evaluation are missing.

Limited knowledge of model performance in the region is a critical shortcome with serious implications for research. Inappropriate model selection could lead to the production of unusable model outputs or misleading interpretation of results and might thus be jeopardizing meaningful decision-making support.

In this study we focus on the headwater area of the Vilcanota-Urubamba basin (VUB) in southern Peru which remains highly understudied despite holding the second largest tropical glacier extent worldwide (~142 km^2) including one of the highest peaks in Peru (Ausangate, 6374 m a.s.l.) and the largest tropical ice cap Quelccaya at 5300 m a.s.l. (Drenkhan et al., 2018; Yarleque et al., 2018). Together with the Zongo glacier in the nearby Cordillera Real, Bolivia, Quelccaya is among the most studied glaciers in the outer tropics with important insights for tropical glacier melt processes (e.g. Guido et al., 2016; Hastenrath, 1978; Réveillet et al., 2015; Sicart et al., 2011; Yarleque et al., 2018). The upper VUB is known for its high cultural importance for surrounding communities (c.f. Vuille et al., 2018) providing water for 674 km^2 of irrigated agricultural land, about 830,000 people, and more than 290 MW of installed hydropower capacity (Drenkhan et al., 2019). This essential water supply for local livelihoods and regional economy is jeopardized by climate change. Recent research suggests that almost half of glacier areas might vanish in the VUB under an RCP2.6 scenario or most parts disappear under an RCP8.5 scenario by the end of the 21st century (Drenkhan et al., 2018; Schauwecker et al., 2017).

Accordingly, this paper aims to identify if parsimonious or more complex glacio-hydrological models can perform robust simulations in areas with a tropical climate and important glacier contribution and scarce data. Therefore, we compare three glacio-hydrological models of different complexity and assess them comprehensively by evaluating the commonly used measures of model performance, associated flow signatures and runoff components.

2. Study site and regional setting

The study is carried out in the Sibinacocha and Phinaya catchments, both located in the headwater of the VUB, in the Cusco region, Peru (Fig. 1). The Sibinacocha (Phinaya) catchment ranges from 4900 to 5975 (4709 to 6004) m a.s.l. with a current glacier extent of 10.9 km^2 (19.6 km^2) starting at about 5020 (4948) m a.s.l. (INAIGEM, 2018). A loss of about 36% (30%) in glacier area has been observed for the Sibinacocha (Phinaya) catchment between 1986 (from our baseline, see Section 3.1 Models and data source) and 2016 (from the glacier inventory from INAIGEM, 2018), consistent with regional glacier studies (e.g. Schoolmeester et al., 2018; Vuille et al., 2018; Yarleque et al., 2018). Below about 5000 m a.s.l., bare soils, overburden materials, and moraines act as an important temporary water store particularly during the dry season (Baraer et al., 2015; Glas et al., 2018). These paraglacial landscapes are sparsely populated by dry puna grasslands whose hydrological role has barely been investigated (Ochoa-Tocachi et al., 2016). An important feature in the proglacial valleys represent glacier-fed high-Andean wetlands (locally called 'bofedales') which can provide a major contribution to base flow (Buytaert and Beven, 2011).

The main sources of moisture in the region are the tropical Atlantic and the Amazon basin (Garreaud et al., 2003; Neukom et al., 2015). Total annual precipitation is about 540 mm, of which about 60% occurs during austral summer (December-February). As in other tropical regions, surface temperature shows low annual variability, ranging from 0° to 3°C at monthly scale at the Sibinacocha reservoir outlet.

The VUB is part of the outer tropics where glacier ablation occurs during the entire year, and accumulation is limited to the wet season (Kaser, 2001). Such a typical 'behavior' of tropical glaciers is related to the absence of thermal seasonality combined with the pronounced wet and dry seasons. Therefore, temperature indices are poor indicators for glacier melt while, in turn, precipitation and humidity play a major role in melting processes (Kronenberg et al., 2019; Sicart et al., 2008). These features of the Peruvian outer tropics have been widely explored in previous studies, mainly for the Quelccaya ice cap at 5300 m a.s.l. (see Buytaert et al., 2017; Hastenrath, 1978; Yarleque et al., 2018) and at the Ausangate peak at 6374 m a.s.l., one of the highest peaks in southern Peru (see Hanshaw and Bookhagen, 2014).

Phinaya (4910 m a.s.l.) is the only settlement in the study area and has about 340 inhabitants (INEI, 2017). Due to the harsh climatic conditions, agriculture is not possible and the only economic activity is livestock breeding, such as alpacas and sheep, and tourism to some extent. Local springs are the only source of potable water for both households and livestock. In 1996 the hydropower

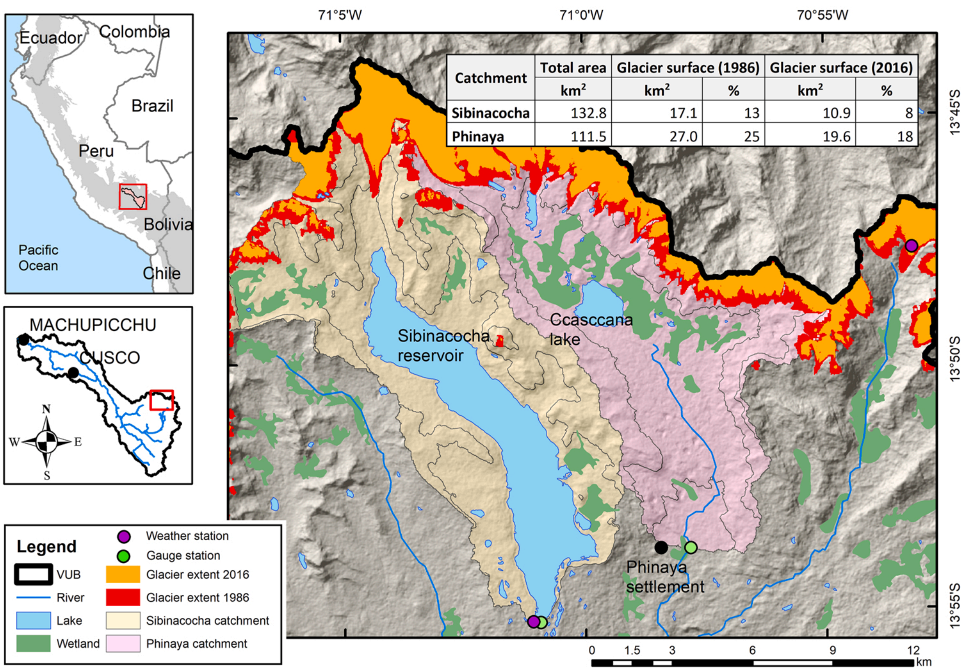


Fig. 1. Map of the study site. Grey lines within the catchments delineate the elevation bands used for hydrological modeling. Glacier extent was derived from satellite image classification for 1986 (proper analysis) and 2016 (INAIGEM, 2018). Wetlands were delimited from the Peruvian national vegetation cover map (MINAM, 2015).

company EGEMSA dammed the existing natural lake of Sibinacocha to regulate up to 110 million m^3 of the natural water body extent with the purpose to provide water for energy production at the Machu Picchu hydropower plant (about 230 km downstream of Sibinacocha).

3. Data and methods

3.1. Models and data source

Three conceptual models were implemented and compared (see Fig. 2), in order of increasing complexity: 1) the lumped Simple Hydrological Model for the Andes – Shaman model (presented in this study in Appendix A), 2) the semi-distributed Hydrologiska Byråns Vattenbalansavdelning - HBV-light model (Seibert and Vis, 2012) and, 3) the semi-distributed Routing System MINERVE - RS Minerve (García Hernández et al., 2016). A conceptual scheme of each model is shown in Appendix A. Both RS Minerve and HBV-light have been selected because they are open-access, widely documented and increasingly be used for assessments in the Peruvian Andes e.g. by the Peruvian Water Authority (ANA) and the National Meteorology and Hydrology Service of Peru (SENAMHI) and beyond. Both models allow for including glacier and non-glacier routines, require few input data, and have been implemented to study a large variety of topics including climate change impacts and water management (e.g. Astorayme et al., 2015; López López et al., 2018; Lujano Laura et al., 2016; Nauditt et al., 2017; Ochoa-Sánchez et al., 2019; Sucozhañay and Célleri, 2018). In addition, the Shaman model has been developed to model a full water balance accounting for, among others, a parsimonious glacier routine with focus on the modeling of tropical glacier melt, a subsurface reservoir, and sectorial water demand including back flows. The model structure, setup, and calibration method were designed to fit with the main features of the tropical Andean catchments like the strong precipitation seasonality, year-round but still seasonal glacier melt, and data scarcity (by using limited input data). The simulations of all three models were implemented on a monthly time step from 1981 to 2010. For both semi-distributed models (HBV-light and RS Minerve), the catchments were divided into 200 m elevation bands starting at 4800 m a.s.l.

The Shaman model includes nine parameters with one module for glacier surfaces and one for non-glacier surfaces. Glacier contribution to river streamflow (four parameters) is estimated as a function of the total glacier surface, the accumulation area ratio (AAR), and melting factors which in turn are estimated through a sinusoidal function to provide a seasonal variability. Non-glacier contribution (5 parameters) is based on Témez (1977) equations and is composed of two buckets, one for fast and one for slow runoff components. As a result, one single runoff series is estimated at the outlet of the catchment (additional details in Appendix A).

HBV-light (named hereafter as HBV) includes 23 parameters and is based on the widely applied original HBV model (Bergström, 1976) that is capable of representing diverse catchment conditions (Seibert and Vis, 2012). Glacier contribution is constrained by nine

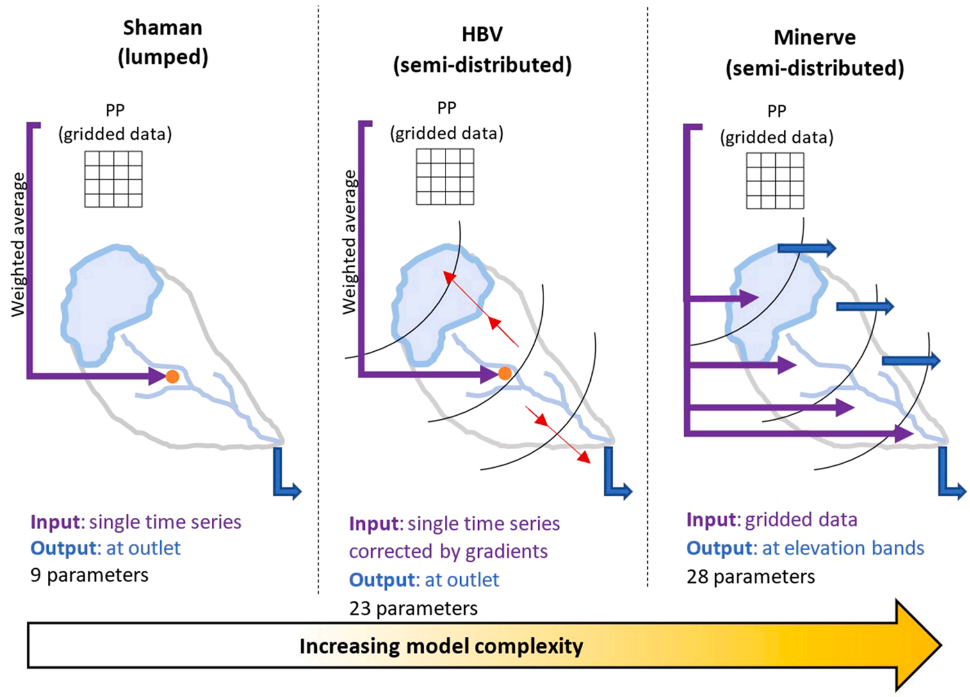


Fig. 2. Schematic representation of the three models compared in this study. PP = monthly precipitation (from gridded data). Light blue area: glacier surface. Purple arrow: input data. Blue arrow: output data. Red arrow: estimation of precipitation with gradients. Black curves: elevation bands. Orange dots: centroid of the catchment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

parameters of which six are specific to the glacier routine. Other parameters in the glacier routine, such as the temperature threshold (TT) that defines thresholds for melt and the separation of liquid/solid precipitation are shared with the routine that simulates non-glacier contributions. In particular, the SP factor allows a seasonal variability of the degree day factor (CFMAX). As a result, runoff is estimated at the outlet of the catchment.

RS Minerve allows simulating total runoff based on an ensemble of different conceptual semi-distributed models. Several models are available such as GR4J (Perrin et al., 2003), SACRAMENTO (Burnash et al., 1973), and HBV. In this study, we used the HBV model to represent the non-glacier contribution composed of fast, intermediate, and slow runoff components (14 parameters), while the Glacier-Snow Model (GSM) (García Hernández et al., 2016) was used to simulate glacier contribution to river streamflow (16 parameters). GSM computed melting from snow and ice using two different degree-day factors and a sinusoidal function that allows accounting for seasonal variations (cf. García Hernández et al. 2016). This combined approach of sinusoidal interpolation function with a seasonal minimum (austral winter) and maximum value (austral summer) is particularly suitable for modeling melt from tropical glaciers. As a result, runoff is estimated for each elevation band and the outlet of the catchment. This ensemble of the HBV and GSM models using the RS Minerve software will be named hereafter as Minerve.

The topography for both catchments Sibinacocha and Phinaya was computed using the Global Digital Elevation Model ASTER GDEM v3 at 30 m grid size available at the Earthdata NASA website (https://search.earthdata.nasa.gov/). The catchment boundaries were then constructed applying the Arc Hydro Tools removing artifacts manually.

Monthly precipitation data were retrieved from the Peruvian Interpolated data of SENAMHI's Climatological and hydrological Observations (PISCO). PISCO v2.1 is a gridded dataset at 10 km spatial resolution that combines weather station, reanalysis, and satellite data from 1981 to 2016 (Aybar Camacho et al., 2017). Monthly temperature data was taken from two local weather stations operated by SENAMHI: Sibinacocha (2016 – 2019) and Quisoquipina (2013–2019) (Fig. 1). Missing data (partially between 1981 and 2013, and 2016) was imputed with nearest neighbor linear interpolation from the PISCO v1.1 temperature dataset. A quantile-based bias correction was applied to the PISCO v1.1 since we observed higher temperature values than the local weather stations (around +6 °C). Potential evapotranspiration (PET) was estimated with the Turc (1961) method which uses mean monthly temperature (retrieved from PISCO) and a global radiation dataset. The latter was taken from the Surface Meteorological and Solar Energy (SSE) web portal (Stackhouse, 2016). For lake evaporation, a linear relationship between monthly air temperature and evaporation measurements or 1996 was calculated at Sibinacocha station ($R^2 = 0.73$). This relationship was then used to estimate lake evaporation for the entire study period at both Sibinacocha and Ccasccana lakes (Fig. 1). Monthly streamflow measurements were obtained from the hydropower company EGEMSA for Sibinacocha outflow (1981–1996) and from ANA for the Phinaya bridge gauge (1981–1996).

Although the data sources were the same, each model required a particular input data arrangement. The Shaman model requires one single input data series that represents the whole catchment. The precipitation series was estimated as a weighted area average, while PET was computed by using the temperature at the average elevation from non-glacier areas only. HBV also required one single input data series. The precipitation and PET series were the same as those used in the Shaman model. For temperature, we used the data from the Quisoquipina weather station. Additionally, the model used gradients to estimate the input data for each elevation band. In Minerve, the full PISCO v2.1 precipitation dataset and both weather stations for temperature were included. The PET values were obtained from the Minerve dataset that also uses the Turc (1961) method.

Three glacier outlines were delineated from 1986, 1994, and 2004, which represented, as far as possible, the mid-year of the calibration, validation, and simulation periods (see Section 3.2 Calibration strategy). Free multi-spectral optical satellite images were used from Landsat 5 (1986, 1994) and Landsat 7 (2004) missions, downloaded from the USGS Earth Explorer. Since 2003 Landsat 7 has scan-line corrector failures, however, due to the small extension of the study area the small affected areas were easily delineated by applying multi-temporal assessment. Images with cloud cover up to 10% were selected within the dry season (June and July) to avoid temporal snow cover leading to confusion of glacier discrimination. Glacier outlines were obtained through a semi-automatic approach based on the Normalized Difference Snow Index (NDSI) (Hall and Riggs, 2011). NDSI thresholds were iteratively assessed, ranging between 0.55 and 0.65. Glacier fragments were filtered using a minimum surface threshold of $\geq 5000\ m^2$ (similar to the Peruvian Glacier Inventory from ANA (2014)) and then manually edited to correct residual and confusion pixels (such as water bodies or cast shadows).

Finally, the calibrated set of parameters of the Sibinacocha catchment were transferred to simulate total runoff for the Phinaya catchment. This is a common practice in data-scarce and ungauged catchments with shared conditions (e.g. Bárdossy, 2007; Moussa and Chahinian, 2009).

3.2. Calibration strategy

Three time periods were assessed: calibration (1981–1990), validation (1988–1996), and simulation (1998–2009). Model results from both calibration and validation were compared with measured streamflow, while for the simulation period comparison was only among model results. In each period the first two years were considered as model initialization (warm-up period) and therefore not included in the assessment. Since most of the studies in the Peruvian Andes focus on water availability, the calibration focused on a better representation of low flows than on high flows.

Parameter calibration was implemented at the Sibinacocha catchment where the longest runoff measurement is available. We took into consideration the expected behavior of the different runoff components during the hydrological year (September to August) in the Peruvian Andes. In the case of glacier contribution to total runoff, the highest absolute values should occur around November – January (austral summer – wet season) when maximum solar radiation matches with low cloudiness and thus highest glacier mass turnover (Kronenberg et al., 2016). Concerning non-glacier surfaces, it was expected that the fast runoff component plays an important role during the wet season. In the study region, high-Andean wetlands are a common feature and work as important water storage, and the intermediate flow could serve as an approximation. Since groundwater is characterized by slow flow rates it was expected that the slow runoff component remains constant during the hydrological year with higher absolute values during the wet season and an increasing relative share of overall contribution in the dry season (Baraer et al., 2015; Glas et al., 2018; Polk et al., 2017).

Parameter calibration was implemented using an objective function (OF) composed of three measures of model performance that capture average, high, and low flows. Each model has its own set of measures of model performance available for calibration. In the Shaman model, the available measures of model performance for the OF were the coefficient of determination (R^2) for average flow, the Nash-Sutcliffe efficiency (Nash) for peak flows, and the logarithmic Nash-Sutcliffe efficiency (Nash-ln) for low flows. In HBV the OF consisted of the model efficiency (Reff) for average flow, the peak flow efficiency (ReffPeak) for peak flows, and the logarithmic efficiency (LogReff) for low flows. Finally, in Minerve, the OF consisted of the Kling-Gupta efficiency (KG) for average flow, the Nash-Sutcliffe efficiency (Nash) for peak flows, and the logarithmic Nash-Sutcliffe efficiency (Nash-ln) for low flows. To facilitate the comparison between models, a common set of measures of model performance were calculated after the individual calibration process: the Spearman rank coefficient (Spearman), Nash, Nash-ln, KG, R^2, and the weighted R^2 (WR^2) with the *b* gradient of the linear correlation between observed and simulated runoff proposed by Krause et al. (2005). Also, eight commonly used flow signatures were calculated to characterize the different parts of the hydrograph (see Table 1).

Degree-day values and melting factor rates from literature (see Condom et al., 2011; Hock, 2003; Kronenberg et al., 2019)) were used to help to constrain glacier parameters. Glacier mass balance from HBV and Minerve was used to calculate glacier mass balance rates, which were compared with existing estimations from 2000 to 2009 for the same region (Dussaillant et al., 2019).

We used automatic calibration by using the available tools from each model. For that reason, no single tool was possible to use to calibrate all models. In the Shaman model we used the generalized reduced gradient (GRG) nonlinear method (Lasdon et al., 1974), in HBV the Genetic Algorithm and Powell optimization (Seibert, 2000), and in Minerve, the Shuffled Complex Evolution – University of Arizona method SCE-UA (Duan et al., 1994).

As seen for calibration, instead of using a common set of measures of model performance and calibration tool, we used the predefined options within each model. This allowed reproducing the normal conditions of model implementation, that is: usually model simulations are executed with the predefined options within the models.

4. Results

To facilitate model comparison, the results were assessed in two stages. First, we focused on the Sibinacocha catchment for which measured streamflow exist. There, we looked at the total runoff simulation and the measures of model performance (Section 4.1 Total runoff comparison), glacier runoff (Section 4.2 Glacier runoff simulation), and total runoff components (Section 4.3 Components of total runoff). In a second step, we looked at the results from the parameter transfer from Sibinacocha to the Phinaya catchment (Section 4.4 Parameter transfer to the ungauged catchment). The calibrated set of parameters for each model is indicated in Appendix A.

4.1. Total runoff comparison

Streamflow measurements at the Sibinacocha catchment allowed us to compare both measures of model performance and flow signatures (Table 2). All models provided acceptable values for measures of model performance. The Spearman coefficient suggests a better fit between measured and simulated total runoff from more complex models in both calibration and validation. The Kling Gupta coefficient (KG) ranks high for Minerve in both calibration and validation, although in calibration the highest value is for the parsimonious Shaman model. The weighted coefficient of determination (WR^2) showed the highest value in calibration (validation) for results from HBV (Shaman). Overall, Nash-ln values (logarithmic Nash-Sutcliffe efficiency) were higher than Nash values (Nash-Sutcliffe efficiency) as a result of the calibration strategy where higher importance was given to low flows. Model performance was on average lower in validation than in calibration, but Shaman shows the strongest model performance reduction, followed by HBV and Minerve. The more complex the model was, the higher the model performances that were achieved.

While measures of model performance tended to favor high model complexity, the flow signatures suggested the opposite. In Table 2, the bold numbers indicate which model provides the closest flow signature values compared to those calculated from

Table 1
Flow signatures used in this study.

Signature	Description	Reference
Qmean (Qm) (m^3/s)	Mean monthly discharge	Sawicz et al. (2011)
Runoff ratio (RR)	Ratio of mean monthly discharge to mean monthly precipitation	Sawicz et al. (2011)
Streamflow elasticity (SE)	Sensitivity of streamflow to changes in precipitation at an annual time scale	Sankarasubramanian et al. (2001)
Slope	Slope of flow duration curve between the 33rd and 66th streamflow percentiles	Sawicz et al. (2011)
Q5 (m^3/s)	5% flow quantile (high flow)	–
Q75 (m^3/s)	75% flow quantile	–
Q95 (m^3/s)	95% flow quantile (low flow)	–
Base flow index (BFI)	Ratio of mean monthly baseflow to mean monthly discharge, using a recession digital filter value of 0.985	Hosseini Duki et al. (2017)

Table 2
Measures of model performance used to assess the goodness of fit between the observed and simulated total runoff with each model at the Sibinacocha catchment. Besides, the flow signatures used to characterize the hydrograph are shown.

		Calibration (1983 – 1990)				Validation (1990 – 1996)				Simulation (2000 – 2010)		
		Qobs	Shaman	HBV	Minerve	Qobs	Shaman	HBV	Minerve	Shaman	HBV	Minerve
Measures of model performance	Spearman	–	0.85	0.91	**0.92**	–	0.82	0.87	**0.89**	–	–	–
	Nash	–	0.82	0.81	**0.83**	–	0.65	**0.73**	0.71	–	–	–
	Nash-ln	–	0.78	0.84	**0.85**	–	0.73	**0.78**	0.77	–	–	–
	KG	–	**0.91**	0.78	0.84	–	0.79	0.79	**0.80**	–	–	–
	R^2	–	0.86	**0.87**	0.72	–	0.65	**0.70**	**0.70**	–	–	–
	WR^2	–	0.07	**0.42**	0.14	–	**0.25**	0.07	0.08	–	–	–
Flow signatures	Qmean (m^3/s)	2.7	2.8	**2.8**	2.9	2.4	**2.4**	**2.3**	2.6	2.5	3.4	3.3
	Q95 (m^3/s)	0.8	**0.8**	0.7	1.2	0.2	**0.4**	0.5	0.4	0.4	0.9	0.4
	Q75 (m^3/s)	1.3	**0.98**	0.91	1.66	0.75	0.5	**0.77**	0.75	4.0	5.0	5.7
	Q5 (m^3/s)	8.4	6.9	**7.0**	6.9	7.8	**7.8**	6.5	7.1	7.5	8.9	8.5
	Slope	-2.61	-3.38	-3.61	**-2.45**	-3.29	-4.93	-3.57	**-3.13**	-5.34	-3.36	-3.92
	RR	0.81	0.84	0.83	**0.80**	0.84	**0.81**	0.78	0.82	0.81	1.12	0.80
	SE	0.91	1.12	**0.89**	0.80	0.82	1.18	0.93	**0.84**	1.16	0.81	0.84
	BFI	0.63	0.59	**0.60**	0.59	0.70	0.57	**0.58**	0.51	0.67	0.64	0.59

Bold numbers in the measures of model performance group indicate the highest value for every measure. Bold numbers in the flow signatures group indicate which model provides the closest value to flow signatures from measured runoff.

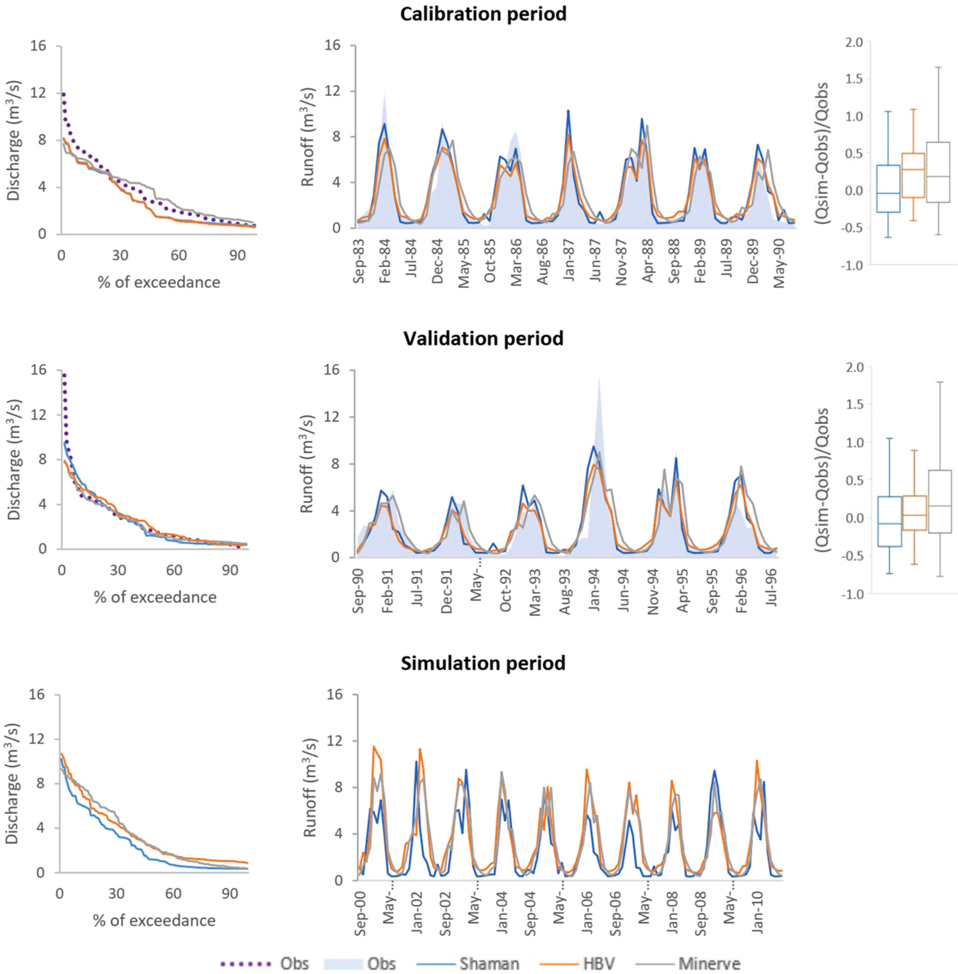
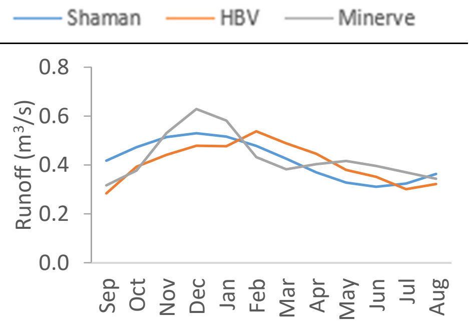
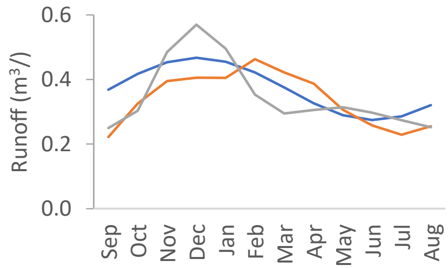


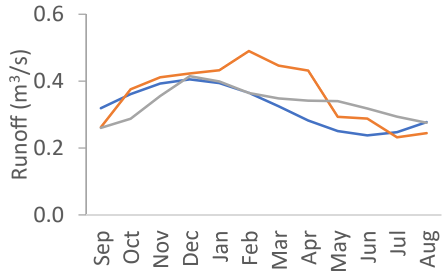
Fig. 3. Comparison between observed and simulated total runoff at Sibinacocha catchment. Left column: flow duration curves show the percentage of time in which discharges were equaled or exceeded. Middle column: historical total runoff. Right column: Box-plots of the difference between simulated and observed runoff expressed as a fraction of observed runoff.

Table 3
Main indicators of simulated glacier runoff at Sibinacocha catchment from multiannual average from each period: calibration (1983–1990), validation (1990 – 1996), and simulation (2000 – 2010). Glacier runoff is expressed as m^3/s and as a contribution to total runoff (%). Graphs on the right panel show the multiannual average of glacier runoff simulation from each period and model.

		Shaman	HBV	Minerve	Shaman	HBV	Minerve
		m^3/s			%		
Calibration Glacier surface 1986 17 km^2	**Annual**	0.42	0.41	0.43	32	24	29
	JJA	0.33	0.32	0.37	55	39	45
	DJF	0.51	0.50	0.55	9	10	15
	Max	0.53	0.54	0.63	59	44	61
	Month	Dec	Feb	Dec	Aug	Aug	Aug
	MB	–	-0.18	-0.20	–	–	–
Validation Glacier surface 1994 15 km^2	**Annual**	0.37	0.34	0.35	31	24	26
	JJA	0.29	0.25	0.27	52	38	40
	DJF	0.45	0.42	0.47	10	10	13
	Max	0.47	0.46	0.57	62	42	62
	Month	Dec	Feb	Dec	Sep	Aug	Sep
	MB	–	-0.24	-0.34	–	–	–
Simulation Glacier surface 2004 13 km^2	**Annual**	0.32	0.36	0.33	29	18	22
	JJA	0.25	0.25	0.30	49	29	42
	DJF	0.39	0.45	0.39	8	7	7
	Max	0.41	0.49	0.41	52	30	58
	Month	Dec	Feb	Dec	Jul	Aug	Aug
	MB	–	-0.22	-0.32	–	–	–







JJA: average during the dry season June, July, August. DJF: average during the wet season December, January, and February. Month: month of maximum glacier runoff (m3/s) and maximum glacier contribution to total runoff (%). MB: glacier mass balance rate in m w.eq./year.

measured runoff for the respective calibration and validation periods. In that way, both Shaman and HBV models performed better than the more complex model Minerve. However, Minerve was the one whose performance does not largely differ between calibration and validation in terms of flow signature similarities, followed by Shaman and HBV models. Despite that, the calculated flow signatures are similar across the different models at calibration and validation periods.

The Shaman model showed higher sensitivity to changes in precipitation than the other models (the highest streamflow elasticity SE among the models). Indeed, both Shaman and HBV had similar flow signatures. Minerve showed, on the other hand, less sensibility to changes in precipitation and influence from fast flows indicated by both lower runoff ratio (RR) and slope compared to other models.

Both the flow duration curve (FDC) and hydrographs from Fig. 3 show similarities between the models during calibration and validation. Nevertheless, the small differences were amplified in the simulation period. As expected, the low flows were better estimated than the extreme peak flows due to the calibration strategy. The box-plots in the right column of Fig. 3 show the difference between the simulated (Qsim) and measured total runoff (Qobs) as a fraction of the total runoff that can be considered as error estimations (EE). Minerve showed higher EE than HBV and Shaman, represented by the largest box-plot. The magnitude of the EE from Minerve and Shaman looks the same in both calibration and validation, while EE from HBV reduces in validation. The box-plots show also a tendency to overestimate the total runoff, except for the Shaman model where over and underestimations look similar.

4.2. Glacier runoff simulation

In the absence of measured glacier runoff data, we used several assumptions to constrain the parameters of the glacier runoff simulation (see Section 3.2 Calibration strategy). Such a strategy resulted in similar glacier runoff from every model in m^3/s but a different proportion of glacier contribution to total runoff (Table 3). All models estimated glacier runoff at about 0.37 m^3/s on average between 1983 and 2010 (about 14 km^2 of glacier surface on average for this period). Estimated annual glacier runoff represents on average about 26% of total runoff which is in line with regional studies (e.g. Buytaert et al., 2017). Both Shaman and Minerve indicated that the maximum glacier contribution (m^3/s) occurs at the beginning of the summer season (December), while HBV estimated the maximum to occur at the end of the summer season (February). During the calibration period, all models simulated the maximum relative glacier contribution in August.

The hydrograph from glacier runoff from the Shaman model is smoother than the ones from HBV and Minerve (Table 3 right panel). That is because the simulations from Shaman are not based on climate variables but on the glacier surface, accumulation area ratio (AAR), and melting factors. Both Minerve and Shaman showed a decrease of the glacier runoff from the calibration to the simulation period according to changes in the glacier surface. HBV in its turn showed a decrease in glacier runoff from calibration to validation, and then a slight increase in the simulation period (Table 3 right panel), which might be related to the high streamflow elasticity (SE) of the model (Table 2).

We also calculated the glacier mass balance rate from HBV and Minerve for each period (Table 3). Low rates were calculated in the calibration period, which then increased during the validation period and decreased again in the simulation period. Despite those

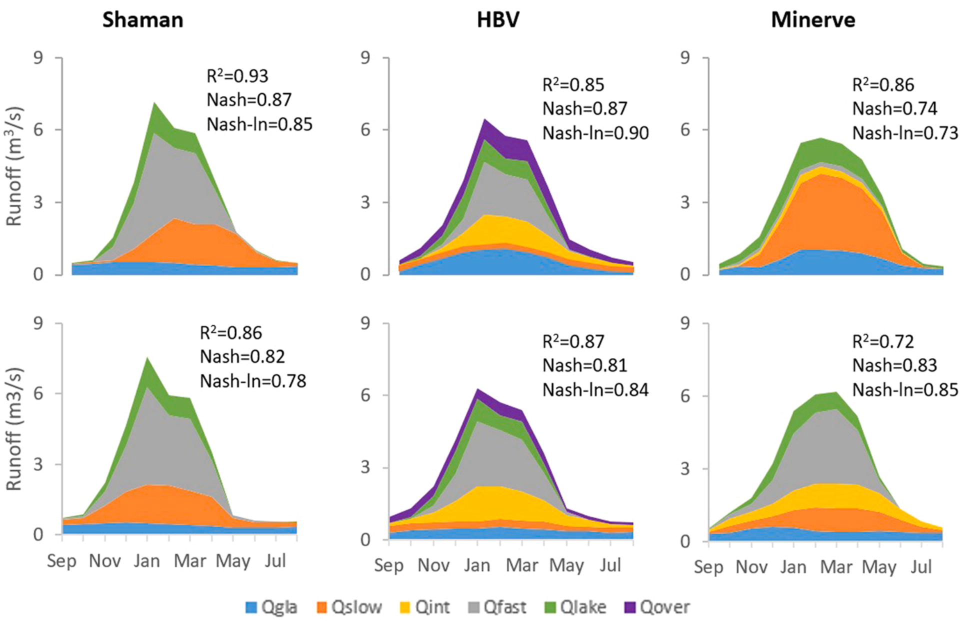


Fig. 4. Components of total runoff at Sibinacocha catchment simulated with each model from the multiannual average hydrological year. Top: results from first automatic calibration with the predefined range of parameters. Bottom: results from calibration after an exhaustive readjustment of parameter ranges and several automatic calibration runs. Qgla: glacier runoff. Qslow, Qint, Qfast: slow, intermediate, and fast flow components of total runoff. Qlake: runoff from the lake. Qover: overland flow from glacierized areas.

changes, results from the simulation period (2000–2010) are in line with the regional estimations of geodetic mass balance computed by Dussaillant et al. (2019) that ranges from − 0.35 to − 0.38 m w.eq./year for the headwater of the VUB. These results are also in line with mass balance estimations from the Cordillera Blanca (2006 – 2007) calculated as − 0.32 ± 0.4 m w.eq./year (Gurgiser et al., 2013).

4.3. Components of total runoff

Each model estimated different runoff components depending on their structure. In addition to glacier and lake runoff, the Shaman model simulated slow and fast runoff; the HBV model simulated slow, intermediate, and fast runoff from non-glacierized areas, and overland flow from glacierized areas; and Minerve simulated slow, intermediate, and fast runoff (Fig. 4).

In automatic calibration, every model had a predefined range of parameters to balance the contribution of every single runoff component. However, we observed that in most of the cases the automatic calibration tends to overestimate the slow or glacier flows to compensate for the strong precipitation seasonality. Even more, such a result provided high model performance in the very first run of automatic calibration. For non-expert users, this might indicate a good model performance despite the wrong representation of relevant hydrological processes in the study area. Fig. 4 shows both results from the first automatic calibration (top) and after a careful and exhaustive readjustment of parameter ranges and several runs of automatic calibration (bottom).

All models were able to represent the assumptions described in Section 3.2 Calibration strategy like higher fast flow contribution in the wet season and higher slow flow contribution in the dry season (bottom row Fig. 4). The slow flow from Shaman is equivalent in order of magnitude to the sum of intermediate and slow flows from HBV and Minerve. HBV indicated a higher contribution from intermediate flow in comparison with the slow flow. In turn, Minerve provided an equivalent contribution between intermediate and slow flows. In the absence of additional measurements, it is difficult to indicate which one provides a more realistic simulation. Regarding the slow flow, that approximates groundwater, HBV simulated a more stable slow flow without peaks, an expected characteristic of groundwater.

4.4. Parameter transfer to the ungauged catchment

The set of parameters calibrated from the Sibinacocha catchment were used to simulate the runoff at the Phinaya catchment. In the latter monthly streamflow measurements from December 1992 to December 1993 exist, which were compared with simulated runoff (Fig. 5). By looking at the hydrograph, results from both HBV and Shaman provided a better fit to the measured data, while Minerve overestimated the total runoff most of the time. However, the seasonality was well represented. All models provided a high coefficient of determination (R2), the highest was HBV (0.93), followed by Minerve (0.83) and then Shaman (0.82). The flow duration curve in Fig. 5 (right panel) was calculated from the full simulation at Phinaya catchment from September 1983 to August 2010. Both Shaman and HBV provided closer simulations among them mainly for high and medium flows, (<45% of exceedance), while Minerve tended to estimate higher flows for the same percent of exceedance. Regarding low flows (>65% of exceedance), all models tended to provide similar results to each other.

5. Discussion

Results outlined above show that there is no significant difference of simulation outputs from parsimonious and more complex models. This corroborates previous studies (e.g. Breuer et al., 2009; Perrin et al., 2001; Vansteenkiste et al., 2014) which indicate that more complex models do not necessarily improve the performance of hydrological simulations. Contrariwise, parsimonious glacio-hydrological models can provide more robust results than more complex models in a context of data scarcity and limited process understanding, such as in the glacierized tropical Andean catchments. If sufficient quantity and quality data is available, more complex models might provide several advantages, like estimations inside of the catchments or deeper processes understanding.

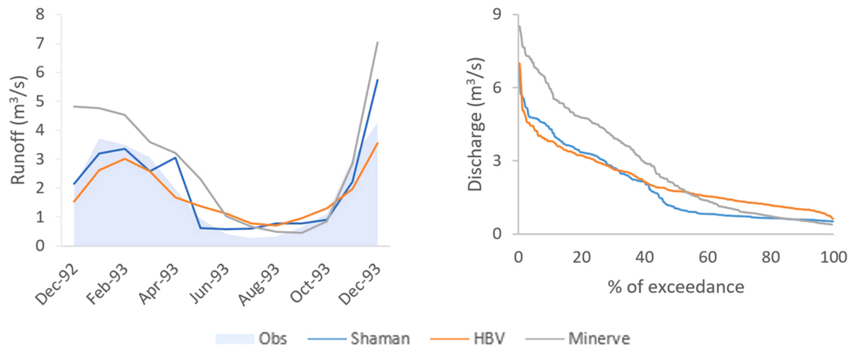


Fig. 5. Left: observed and simulated total runoff at Phinaya bridge from December 1992 to December 1993. Right: flow duration curve from the simulated total runoff from 1983 to 2010.

However, several limitations need to be considered when evaluating the presented results in this study. Firstly, the selected hydrological models depict a certain range rather than a full spectrum of model complexity. Secondly, no fully distributed model was assessed due to the inherent limitations of data input and process complexity. This might also be the main reason why no fully distributed models have been implemented yet in the region. Thirdly, the use of different OF and calibration tools influence the simulations and model results. Although some calibration tools might be more powerful than others, all model results were compared with the same measures of model performance. Lastly, another important insight is linked to the assessment of final model outputs at the catchment outlet only.

In the following sections, we elaborate and discuss in more detail the aforementioned critical points and present a modeling framework to improve the robustness of glacio-hydrological simulation in glacierized data-scarce tropical Andean catchments.

5.1. Parsimonious vs complex glacio-hydrological models

Results suggest that three models of different complex perform similar runoff simulations. The median error estimations from all models are also comparable among them (+/−15%) (see Fig. 3) although they represent only the behavior of total runoff. Small differences between model results might be related to the model structure (the inherent conceptual differences) and how the input data is used (see Section 3.1 Models and data sources). As a result, the parsimonious Shaman model can perform as well as the more complex HBV or Minerve models while providing a clear advantage of reduced computer power and data requirement.

This becomes more relevant with automatic calibration without an expert supervision and quality checking because the runoff components can be wrongly compensated to achieve good model performance (see top row in Fig. 4). This is known as the 'achieving good model performance for the wrong reasons' error and has been observed in several model applications (e.g. Finger et al., 2015; Kirchner, 2006; Konz and Seibert, 2010). Automatic calibration results are highly sensitive on the selection of the measures of model performance, the specified search scheme, and numerical measures of the goodness-of-fit (Madsen, 2000). Therefore, even when automatic calibration tools are powerful and easy to use, they need to be carefully supervised to make full benefit of the advantages such as exploring an exhaustive set of parameter options in a short time series. Fig. 4 shows that unsupervised automatic calibration from Shaman and HBV provided a distribution of runoff components than the automatic calibration from Minerve. Nonetheless, the Shuffled Complex Evolution method (SCE-UA) which is available for Minerve is well known as an efficient and robust tool for automatic calibration of hydrological models (Muttil and Jayawardena, 2008; Seong et al., 2015; Wang et al., 2010). Under the specific conditions of this study, the exhaustive calibration strategy of Minerve that requires further adjustment of the runoff components represents a disadvantage compared to the more parsimonious Shaman and HBV models.

Our assessment, under complex and data-scarce basin conditions, highlights the fact that measures based on model performance only do not provide sufficient information about the drivers behind the predominant hydrological processes. A comprehensive assessment is needed to identify or reject model results that do not meet minimum requirements (Vaché and McDonnell, 2006). Such a comprehensive assessment is directly linked to the calibration strategy. In this study, the calibration strategy used different assumptions based on the fast and slow flows to constrain the runoff components but also to reject unrealistic model results (see Section

Table 4
Compared potential benefits between the implementation of parsimonious and more complex glacio-hydrological models in the data-scarce tropical Andean catchments.

Category		Feature	Model complexity	
			Parsimonious	More complex
Input	Input data	Detailed input data at the outlet	Low	High
		Input data within the catchment	Low	High
		Distributed input data	Low	High
	Knowledge of processes	Knowledge of catchment processes	Low	High
		Knowledge of the calibration tool	Low	High
		Knowledge of model structure	Low	High
Model	Model structure	Structure complexity	Low	High
		Structure flexibility	Medium	Low
	Parameters	Number of parameters	Low	High
		Risk of parameter equifinality	Low	High
		Risk of overparameterization	Medium	High
	Calibration strategy	Complexity of the calibration tool	Medium	High
		Calibration flexibility	Low	High
Output	Output data	Output data at the outlet	Low	High
		Output data within the catchment	Low	High
	Representation of hydrological processes	Representation of internal basin processes	Low	High
		Representation of hydrological processes	Low	High
		Insights of hydrological processes	Low	High

Dark blue color indicates high potential benefits, light blue color indicates low benefits.

3.2 Calibration strategy). Vansteenkiste et al. (2014) applied a similar approach for the calibration procedure where runoff responses derived from rainfall and potential evapotranspiration were used to constrain the runoff components. Nevertheless, this approach implies some degree of subjectivity and is therefore highly dependent on the modeler's expertise.

More complex models provide several advantages such as a larger sets of calibration options (Perrin et al., 2001), spatial scenarios, specific simulation and analysis options within the catchment (Pokhrel and Gupta, 2011) and provide new insights about the hydrological processes (see Table 4). Nevertheless, in the absence of sufficient measurements and process understanding, all model results remain as unverified hypotheses (Hrachowitz and Clark, 2017). To produce robust simulations and benefit of its advantages, more complex models require sufficient data availability (of both quality and quantity), and a deep understanding of the underlying model structure and the predominant hydrological processes in the catchment.

Table 4 summarizes the potential benefits and trade-offs of the implementation of parsimonious and more complex models under data scarcity. On the one side, parsimonious models require limited input data and knowledge about the hydrological processes, with robust model outputs but of limited insight and mostly with single research goals. On the other side, more complex models require additional data and expertise to use their full potential and advantages such as providing new insights of hydrological processes. These trade-offs highlight the importance of a careful model selection depending on the research goal and how well the dominant hydrological processes are understood. Therefore, a more appropriate approach in data-scarce environments is to collect additional data, improve the calibration strategy, and implement a comprehensive model assessment of parsimonious models rather than increasing model complexity (c.f. Finger et al., 2015; Tarasova et al., 2016).

5.2. Glacier runoff simulation

In this study, glacier runoff was constrained by three factors: the change of glacier area over time, the glacier mass balance rate, and relative glacier contribution to total runoff. Accounting for changes in glacier area over time due to glacier melt is not a regular practice, instead the glacier area remains the same in hydrological simulations over large periods with the risk to overestimate glacier runoff (Tiel et al., 2020). Here, the glacier surface was delineated from satellite images using a common semi-automatic approach deploying the NDSI and manual editing with an estimated error at around ± 5% under normal conditions (Paul and Kääb, 2005). Other studies applied different approaches that take into account vertical changes in the glacier mass but require additional data such as glacier volume or snow water equivalent (e.g. Bahr et al. 2015; Huss et al., 2010). Our computed relative annual and seasonal glacier

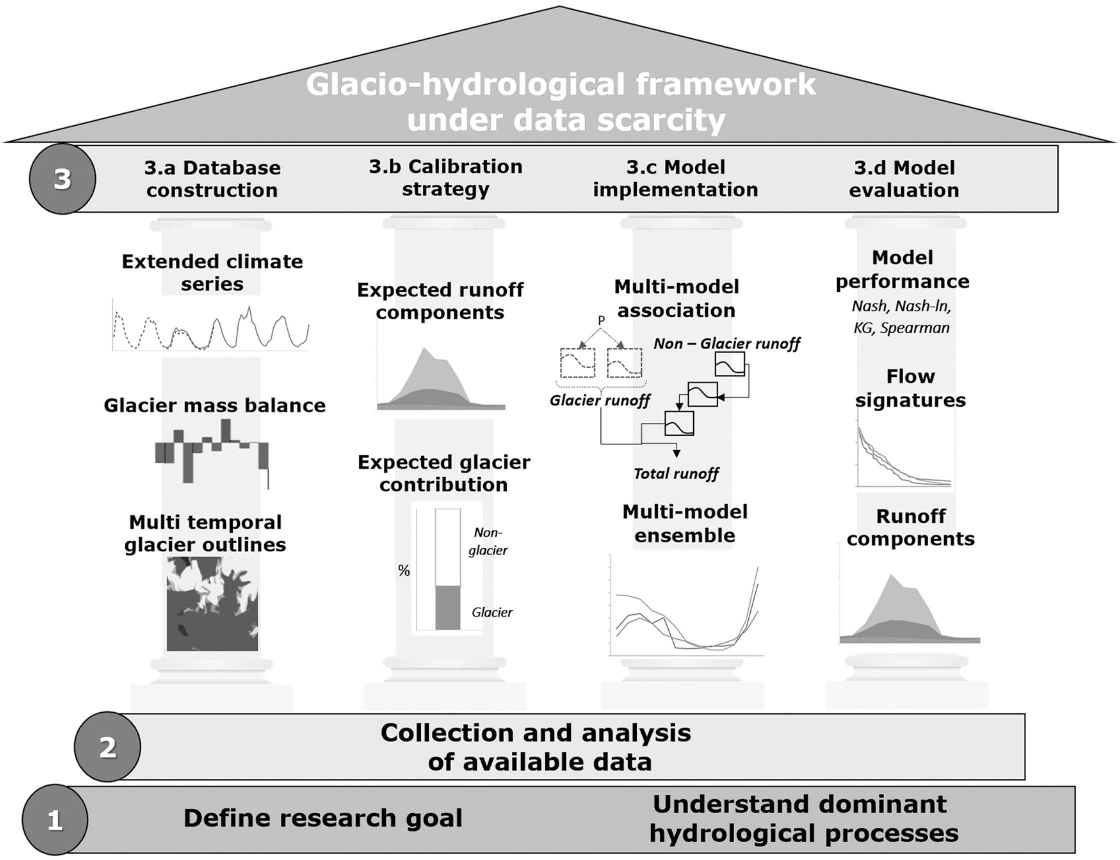


Fig. 6. Proposed framework for robust glacio-hydrological simulation under data scarcity.

contribution to streamflow are in line with existing studies (e.g. Buytaert et al., 2017). Nevertheless, without field measurements that help to constrain glacier model routines (see Konz and Seibert, 2010) results are highly dependent on user expertise and parameter selection. Therefore, we additionally used a glacier mass balance rate corresponding to available measurements which have proven to be suitable when assessing appropriate melting rates (see Section 3.2 Calibration strategy).

In tropical high-mountain conditions, our results suggest that a differentiated routine and parameters for glacier and non-glacier areas (like in Shaman and Minerve models) are more helpful than an integrated approach (like HBV) when simulating glacier runoff. The differentiated coefficients provide flexibility when trying to simulate the expected absolute and relative glacier contribution to total runoff. In any case, common degree-day models are not appropriate for simulating the melting of tropical glaciers. At sub-annual time steps, air temperature is a poor index of melt due to relative constant temperature throughout the year and low air heat content in high altitudes (Sicart et al., 2008). Therefore, adapted approaches, such as a sinusoidal interpolation function (used in Shaman and Minerve models) with a seasonal minimum (austral winter) and maximum melt factor (austral summer) seem to be more promising (García Hernández et al., 2016; Kronenberg et al., 2019; Sicart et al., 2008).

5.3. *A framework for robust glacio-hydrological simulation in data-scarce areas*

Based on the insights of our inter-model comparison for the Peruvian Andes, a framework has been developed that includes three main levels and four pillars to support robust glacio-hydrological simulation under data scarcity (Fig. 6).

In the first level, the research goal of simulations is defined and the dominant hydrological processes are identified. In this study, the focus was rather on low flows than on peak flows, while the seasonality of glacier runoff and the high contribution of slow flows were identified as key hydrological processes. The second level covers the data collection process including e.g. gridded datasets of precipitation and evapotranspiration (often not available as in-situ data), while temperature can normally be obtained from weather stations. Such a data needs to be comprehensively assessed in both quality and quantity depending on the time step and study period.

The third level comprises the entire modeling process divided into four pillars. The first pillar (3.a) focuses on the database construction by extending temporally and spatially, when possible, the climate data series, the collection of glacier mass balance data or rates and the delimitation of multit-emporal glacier outlines. In this study for instance, we extended the in-situ temperature series in time with gridded data, compared our glacier mass balance data with regional estimates and delineated glacier outlines for calibration, validation and simulation periods. The second pillar (3.b) focuses on the calibration strategy by establishing assumptions about the expected runoff components and the glacier runoff contribution to total runoff (see Section 3.2 Calibration strategy). Then, the third pillar (3.c) focuses on the model implementation through the multi-model association (e.g. differentiated routines and parameters for glacier and non-glacier regions) and the multi-model ensemble (e.g. running different models of different complexities), the latter which has been successfully implemented in climate models. Finally, the last pillar (3.d) focuses on the model evaluation by using both, common model performance indicators and, additionally, evaluating e.g. flow signatures and different runoff components.

Short and limited data series do not allow the models to capture all relevant flow signatures, and models can only partially reproduce them. Therefore, the use of different models and complexities in a multi-model ensemble can provide more robust results than a single model. If available, additional data can be included in the Calibration strategy (3.b) or in the Model evaluation (3.d) as part of a multi-response or multi-data calibration (e.g. He et al., 2018; Konz and Seibert, 2010; Pellicciotti et al., 2012; Tarasova et al., 2016; Uhlenbrook and Leibundgut, 2002).

The individual steps within the framework can be considered as universally applicable to improve the quality of hydrological simulations and considered best practices in applications beyond this particular study region and purpose. Therefore, the framework is also valid in data-rich and non-tropical glacier regions. Nonetheless, a systematic integration and implementation of each step within the framework considerably improves the quality and meaningfulness of simulations in data-scarce regions.

5.4. *Implications for technical institutions and decision-makers*

The development of efficient strategies for data collection and hydrological modeling approaches that increases process understanding is of paramount importance to bridge limited in-situ data and the lack of regional studies. Modelers can benefit from a knowledge production process by implementing parsimonious models at a first stage and then increase model complexity as new data and knowledge is gathered. Therefore, the presented framework (Fig. 6) can facilitate the exploration of new models within a multi-model strategy that could fit better research goals and context.

Nevertheless, data scarcity, limited research capacities and process understanding remain a major bottleneck that both academia and (local) governments need to tackle. Future research should focus on overcoming these limitations that require, among others, additional funding, time allocation for research, capacity building and the development and use of meaningful applications. In the last years, citizen science based initiatives have increasingly been adopted in the tropical Andes and beyond, as they contribute to the flexible collection of extensive datasets, often combined with low-cost sensor technology (Buytaert et al., 2014; Njue et al., 2019). As participatory approaches, they can potentially empower e.g. local communities by engaging them with data monitoring, management, and further decision-making. Ideally, this collaborative process integrates different types of co-produced knowledge within a shared baseline since the early research planning (Muccione et al., 2019) which could then be used for hydrological modeling and further water management planning (Njue et al., 2019). Additionally, decision-making under deep uncertainty (Marchau et al., 2019) and exploratory scenario modeling (Kwakkel, 2017) approaches have an enormous potential to support policy-making for climate change and water management. Such approaches deal with high uncertainties from different sources including both natural and human factors. In the near future, glacio-hydrological modeling in the tropical Andes will most likely continue to deal with high uncertainty and limited process understanding due to the striking gap in data monitoring and modeling approaches. Model results should thus be taken as first-order indicators with limited significance and precision that need to be constantly revised and updated when new data are available. Future studies should also continue with systematic data collection in the region, exploring improvements in understanding hydrological processes, as well as, the adaptation or design of particular model structures that represent the relevant catchment behavior in the glacierized tropical Andes. The use of remote sensing data seems promising to fill data gaps but modelers should be aware of the limitations and combine the advantages of comprehensive remote sensing assessments with ground-truth validation (Sidle, 2021; Taylor et al., 2021).

6. Conclusion

This paper aims to identify if parsimonious or more complex glacio-hydrological models can perform robust simulations in tropical glaciated mountain environments with scarce data. Accordingly, three models of different complexity were compared and comprehensively assessed.

For the 2000–2010 period, total runoff was estimated as 2.5–3.3 m^3/s with an annual (maximum) glacier contribution of 18–29% (29–49%). Absolute (relative) maximum glacier contribution to runoff was calculated for December (August) as 0.41–0.49 m^3/s (30 – 58%). All model complexities allow for acceptable modeling performance (R^2 0.65–0.70, Nash-Sutcliffe 0.65–0.73, Nash-Sutcliffe-ln 0.73–0.78) with small differences related to the model structure. This confirms findings from previous studies more complex models do not necessarily improve hydrological simulation. More complex models require a more comprehensive calibration strategy and assessment to avoid simulations with apparently high model performance driven by inadequate assumptions. Besides, complex models require a better understanding of the hydrological processes that are often hampered by data scarcity, limited knowledge, and field accessibility, such as in the glacierized tropical Andes. Therefore, our study suggests that parsimonious glacio-hydrological models can provide more robust results than complex models in data-scarce catchments.

Our results imply several limitations that need to be carefully considered: 1) the selected models do not cover the full spectrum of model complexity, 2) different OF and calibration tools have been used that influence the simulation and quality of the model results, and 3) the assessment of model outputs was performed at the catchment outlet only which does not allow for further modeling insights within the catchment.

Results suggest that the use of parsimonious models combined with a careful calibration strategy, and additional data collection can provide more robust simulations rather than opt for increasing model complexity. Additional data includes, among others, extended climate data series, glacier mass balance data, multitemporal glacier outlines and single runoff measurements. Based on these insights, a framework has been developed that aims for robust glacio-hydrological simulation under data scarcity. This approach includes a multi-model focus with a comprehensive assessment of flow signatures and runoff components, and is relevant for the tropical Andes region and other regions with similar characteristics.

While such a combined model framework will most likely contribute to more robust hydrological simulations, the remaining challenges of high uncertainty, limited knowledge and capacities hamper a further improvement of model results. Academia and (local) governments should therefore continue to extend data collection, improve understanding of hydrological processes, and the adaptation or design of model structures for specific applications in glacierized tropical Andean catchments.

Funding

This work is part of a PhD project funded by the Swiss Government Excellence Scholarships for Foreign Scholars and Artists and the AguaFuturo project funded by the Swiss National Science Foundation (project no. 205121L_166272).

CRediT authorship contribution statement

Randy Muñoz: Conceptualization, Data curation, Methodology, Visualization, Software, Formal analysis, Writing – original draft. **Christian Huggel:** Conceptualization, Supervision, Writing – review & editing. **Daniel Viviroli:** Conceptualization, Methodology, Formal analysis, Supervision, Writing – review & editing. **Marc Vis**: Methodology, Writing – review & editing. **Fabian Drenkhan:** Visualization, Formal analysis, Writing – review & editing.

Declaration of Competing Interest

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

Acknowledgments

We would like to thank EGEMSA for providing the runoff and climate data from Sibinacocha, and Inés Dussaillant for sharing her detailed dataset about glacier melting rates. We also thank Rafael Muñoz-Carpena and two anonymous reviewers for their valuable comments to improve the quality of this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2021.100932.

References

ANA, 2014. In: Glaciares, Primera (Eds.), Inventario Nacional de Glaciares y Lagunas. Inventario Nacional de Glaciares y Lagunas, Huaraz, Perú https://doi.org/http://www.ana.gob.pe/media/981508/glaciares.pdf.

Andres, N., Vegas Galdos, F., Lavado, W., Zappa, M., 2014. Water resources and climate change impact modelling on a daily time scale in the Peruvian Andes. Hydrol. Sci. J. 59, 2043–2059. https://doi.org/10.1080/02626667.2013.862336.

Astorayme, M., García, J., Suarez, W., Felipe, O., Huggel, C., Molina, W., 2015. Modelización hidrológica con un enfoque semidistribuido en la cuenca del río Chillón. Perú. Rev. Peru. Geo Atmosférica RGPA 4, 109–124.

Aybar Camacho, C., Lavado, W., Huerta Julca, A., Fernández Palomino, C., Vega Jacome, F., Sabino Rojas, E., Felipe Obando, O., 2017. Uso del Producto Grillado "PISCO" de precipitación en Estudios, Investigaciones y Sistemas Operacionales de Monitoreo y Pronóstico Hidrometeorológico, in: Nota Técnica 001 SENAMHI-DHI-2017. Lima, Peru, p. 22.

Bahr, D.B., Pfeffer, W.T., Kaser, G., 2015. A review of volume-area scaling of glaciers. Rev. Geophys. 53, 95–140. https://doi.org/10.1002/2014RG000470.

Bai, P., Liu, Xiaomang, Liang, K., Liu, Xiaojie, Liu, C., 2017. A comparison of simple and complex versions of the Xinanjiang hydrological model in predicting runoff in ungauged basins. Hydrol. Res. 48, 1282–1295. https://doi.org/10.2166/nh.2016.094.

Baraer, M., Mark, B.G., McKenzie, J.M., Condom, T., Bury, J., Huh, K.-I.I., Portocarrero, C., Gómez, J., Rathay, S., 2012. Glacier recession and water resources in Peru's Cordillera Blanca. J. Glaciol. 58, 134–150. https://doi.org/10.3189/2012JoG11J186.

Baraer, M., McKenzie, J., Mark, B.G., Gordon, R., Bury, J., Condom, T., Gomez, J., Knox, S., Fortner, S.K., 2015. Contribution of groundwater to the outflow from ungauged glacierized catchments: a multi-site study in the tropical Cordillera Blanca, Peru. Hydrol. Process. 29, 2561–2581. https://doi.org/10.1002/hyp.10386.

Bárdossy, 2007. Calibration of hydrological model parameters for ungauged catchments. Hydrol. Earth Syst. Sci. 11, 703–710.

Bergström, S., 1976. Development and application of a conceptual runoff model for Scandinavian catchments. PhD Thesis. SMHI Reports RHO No. 7. Norrköping.

Breuer, L., Huisman, J.A., Willems, P., Bormann, H., Bronstert, A., Croke, B.F.W., Frede, H.-G., Gräff, T., Hubrechts, L., Jakeman, A.J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D.P., Lindström, G., Seibert, J., Sivapalan, M., Viney, N.R., 2009. Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use. Adv. Water Resour. 32, 129–146. https://doi.org/10.1016/j.advwatres.2008.10.003.

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., 1973. A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers.

Buytaert, W., Beven, K., 2011. Models as multiple working hypotheses: Hydrological simulation of tropical alpine wetlands. Hydrol. Process. 25, 1784–1799. https://doi.org/10.1002/hyp.7936.

Buytaert, W., Celleri, R., Willems, P., Bièvre, B., De, Wyseure, G., 2006. Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. J. Hydrol. 329, 413–421. https://doi.org/10.1016/j.jhydrol.2006.02.031.

Buytaert, W., Moulds, S., Acosta, L., De Bièvre, B., Olmos, C., Villacis, M., Tovar, C., Verbist, K.M.J., 2017. Glacial melt content of water use in the tropical Andes. Environ. Res. Lett. 12, 114014 https://doi.org/10.1088/1748-9326/aa926c.

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T.C., Bastiaensen, J., De BiÃvre, B., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D.M., Hergarten, C., Isaeva, A., Karpouzoglou, T., Pandeya, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Van Hecken, G., Zhumanova, M., 2014. Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. Front. Earth Sci. 2, 1–21. https://doi.org/10.3389/feart.2014.00026.

Caldwell, P.V., Kennen, J.G., Sun, G., Kiang, J.E., Butcher, J.B., Eddy, M.C., Hay, L.E., Lafontaine, J.H., Hain, E.F., Nelson, S.A.C., Mcnulty, S.G., 2015. A comparison of hydrologic models for ecological flows and water availability. Ecohydrology 8, 1525–1546. https://doi.org/10.1002/eco.1602.

Carenzo, M., Pellicciotti, F., Mabillard, J., Reid, T., Brock, B.W., 2016. An enhanced temperature index model for debris-covered glaciers accounting for thickness effect. Adv. Water Resour. 94, 457–469. https://doi.org/10.1016/J.ADVWATRES.2016.05.001.

Condom, T., Escobar, M., Purkey, D., Pouget, J.C., Suarez, W., Ramos, C., Apaestegui, J., Tacsi, A., Gomez, J., 2012. Simulating the implications of glaciers' retreat for water management: a case study in the Rio Santa basin, Peru. Water Int. 37, 442–459. https://doi.org/10.1080/02508060.2012.706773.

Condom, T., Escobar, M., Purkey, D., Pouget, J.C., Suarez, W., Ramos, C., Apaestegui, J., Zapata, M., Gomez, J., Vergara, W., 2011. Modelling the hydrologic role of glaciers within a Water Evaluation and Planning System (WEAP): a case study in the Rio Santa watershed (Peru). Hydrol. Earth Syst. Sci. Discuss 8, 869–916. https://doi.org/10.5194/hessd-8-869-2011.

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. https://doi.org/10.1002/wat2.1105.

Drenkhan, F., Guardamino, L., Huggel, C., Frey, H., 2018. Current and future glacier and lake assessment in the deglaciating Vilcanota-Urubamba basin, Peruvian Andes. Glob. Planet. Change 169, 105–118. https://doi.org/10.1016/j.gloplacha.2018.07.005.

Drenkhan, F., Huggel, C., Guardamino, L., Haeberli, W., 2019. Managing risks and future options from new lakes in the deglaciating Andes of Peru: The example of the Vilcanota-Urubamba basin. Sci. Total Environ. 665, 465–483. https://doi.org/10.1016/j.scitotenv.2019.02.070.

Duan, Q., Sorooshian, S., Gupta, V.K., 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. J. Hydrol. 158, 265–284. https://doi.org/10.1016/0022-1694(94)90057-4.

Dussaillant, I., Berthier, E., Brun, F., Masiokas, M., Hugonnet, R., Favier, V., Rabatel, A., Pitte, P., Ruiz, L., 2019. Two decades of glacier mass loss along the Andes. Nat. Geosci. 12, 802–808. https://doi.org/10.1038/s41561-019-0432-5.

Finger, D., Vis, M., Huss, M., Seibert, J., 2015. The value of multiple data set calibration versus model complexity for improving the performance of hydrological models in mountain catchments. Water Resour. Res. 51, 1939–1958. https://doi.org/10.1002/2014WR015712.

Frans, C., Istanbulluoglu, E., Lettenmaier, D.P., Naz, B.S., Clarke, G.K.C., Condom, T., Burns, P., Nolin, A.W., 2015. Predicting glacio-hydrologic change in the headwaters of the Zongo River, Cordillera Real, Bolivia. Water Resour. Res. 51, 9029–9052. https://doi.org/10.1002/2014WR016728.

Fyffe, C.L., Reid, T.D., Brock, B.W., Kirkbride, M.P., Diolaiuti, G., Smiraglia, C., Diotri, F., 2014. A distributed energy-balance melt model of an alpine debris-covered glacier. J. Glaciol. 60, 587–602. https://doi.org/10.3189/2014JOG13J148.

García Hernández, J., Paredes Arquiloa, J., Foehn, A., Roquier, B., 2016. Rs Minerve - Technical Manual v2.7.

Garreaud, R., Vuille, M., Clement, A.C., 2003. The climate of the Altiplano: Observed current conditions and mechanisms of past changes. In: Palaeogeography, Palaeoclimatology, Palaeoecology. Elsevier B.V, pp. 5–22. https://doi.org/10.1016/S0031-0182(03)00269-4.

Glas, R., Lautz, L., McKenzie, J., Mark, B., Baraer, M., Chavez, D., Maharaj, L., 2018. A review of the current state of knowledge of proglacial hydrogeology in the Cordillera Blanca, Peru. Wiley Interdiscip. Rev. Water 5, e1299. https://doi.org/10.1002/wat2.1299.

Guido, Z., McIntosh, J.C., Papuga, S.A., Meixner, T., 2016. Seasonal glacial meltwater contributions to surface water in the Bolivian Andes: A case study using environmental tracers. J. Hydrol. Reg. Stud. 8, 260–273. https://doi.org/10.1016/j.ejrh.2016.10.002.

Gurgiser, W., Marzeion, B., Nicholson, L., Ortner, M., Kaser, G., 2013. Modeling energy and mass balance of Shallap Glacier, Peru. Cryosph. 7, 1787–1802. https://doi.org/10.5194/tc-7-1787-2013.

Hall, D.K., Riggs, G.A., 2011. Normalized-difference snow index (NDSI). In: Encyclopedia of Snow, Ice and Glaciers. Springer, Netherlands, pp. 779–780. https://doi.org/10.1007/978-90-481-2642-2_376.

Hanshaw, M.N., Bookhagen, B., 2014. Glacial areas, lake areas, and snow lines from 1975 to 2012: Status of the cordillera vilcanota, including the Quelccaya Ice Cap, northern central Andes, Peru. Cryosphere 8, 359–376. https://doi.org/10.5194/tc-8-359-2014.
Hastenrath, S., 1978. Heat-budget measurements on the Quelccaya ice cape. Peruvian Andes. J. Glaciol. 20, 85–97.
He, Z., Vorogushyn, S., Unger-Shayesteh, K., Gafurov, A., Kalashnikova, O., Omorova, E., Merz, B., 2018. The value of hydrograph partitioning curves for calibrating hydrological models in glacierized basins. Water Resour. Res. 54, 2336–2361. https://doi.org/10.1002/2017WR021966.
Hock, R., 2003. Temperature index melt modelling in mountain areas. J. Hydrol. 282, 104–115. https://doi.org/10.1016/S0022-1694(03)00257-9.
Hosseini Duki, S.R., Seyedian, M., Rouhani, H., Farasati, M., 2017. Evaluation of base flow separation methods for determining water extraction (case study: gorganroud river basin). Water Harvest. Res. 2, 54–70. https://doi.org/10.22077/JWHR.2018.830.
Hrachowitz, M., Clark, M.P., 2017. HESS opinions: the complementary merits of competing modelling philosophies in hydrology. Hydrol. Earth Syst. Sci. 21, 3953–3973. https://doi.org/10.5194/hess-21-3953-2017.
Huss, M., Jouvet, G., Farinotti, D., Bauder, A., 2010. Future high-mountain hydrology: a new parameterization of glacier retreat. Hydrol. Earth Syst. Sci. 14, 815–829. https://doi.org/10.5194/hess-14-815-2010.
INAIGEM, 2018. Inventario Nacional de Glaciares - Las Cordilleras Glaciares del Perú, 1st ed. Huaraz, Peru.
INEI, 2017. Sistema de Consulta de Centros Poblados [WWW Document]. URL http://sige.inei.gob.pe/test/atlas/. (accessed 5.20.17).
Jiang, T., Chen, Y.D., Xu, C. yu, Chen, Xiaohong, Chen, Xi, Singh, V.P., 2007. Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China. J. Hydrol. 336, 316–333. https://doi.org/10.1016/j.jhydrol.2007.01.010.
Juen, I., Kaser, G., Georges, C., 2007. Modelling observed and future runoff from a glacierized tropical catchment (Cordillera Blanca, Perú). Glob. Planet. Change 59, 37–48. https://doi.org/10.1016/j.gloplacha.2006.11.038.
Kaser, G., 2001. Glacier-climate interaction at low latitudes. J. Glaciol. 47, 195–204. https://doi.org/10.3189/172756501781832296.
Kaser, G., Georges, C., 1997. Changes of the equilibrium-line altitude in the tropical Cordillera Blanca, Peru, 1930-50, and their spatial variations. Ann. Glaciol. 24, 344–349. https://doi.org/10.3189/s0260305500012428.
Kirchner, J.W., 2006. Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. Water Resour. Res. 42 https://doi.org/10.1029/2005WR004362.
Konz, M., Seibert, J., 2010. On the value of glacier mass balances for hydrological model calibration. J. Hydrol. 385, 238–246. https://doi.org/10.1016/j.jhydrol.2010.02.025.
Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5, 89–97. https://doi.org/10.5194/adgeo-5-89-2005.
Kronenberg, M., Rohrer, M., Huggel, C., Frey, H., Muñoz, R., 2019. Modelling melt from glaciers located in the tropical Andes. Background and practical considerations. [WWW Document]. URL https://www.proyectoglaciares.pe/materiales/modelling-melt-from-glaciers-located-in-the-tropical/.
Kronenberg, M., Schauwecker, S., Huggel, C., Salzmann, N., Drenkhan, F., Frey, H., Giraldez, C., Gurgiser, W., Kaser, G., Juen, I., Suarez, W., Hernaández, J.G., Sanmartín, J.F., Ayros, E., Perry, B., Rohrer, M., 2016. The projected precipitation reduction over the central andes may severely affect peruvian glaciers and hydropower production. Energy Procedia 97, 270–277. https://doi.org/10.1016/j.egypro.2016.10.072.
Kwakkel, J.H., 2017. The exploratory modeling workbench: an open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. Environ. Model. Softw 96, 239–250. https://doi.org/10.1016/j.envsoft.2017.06.054.
Lasdon, L.S., Fox, R.L., Ratner, M.W., 1974. Nonlinear optimization using the generalized reduced gradient method. RAIRO - Oper. Res. - Rech. Opérationnelle 8, 73–103.
Li, H., Xu, C.Y., Beldring, S., 2015. How much can we gain with increasing model complexity with the same model concepts? J. Hydrol. 527, 858–871. https://doi.org/10.1016/j.jhydrol.2015.05.044.
López López, P., Immerzeel, W.W., Rodríguez Sandoval, E.A., Sterk, G., Schellekens, J., 2018. Spatial downscaling of satellite-based precipitation and its impact on discharge simulations in the magdalena river basin in Colombia. Front. Earth Sci. 6 https://doi.org/10.3389/FEART.2018.00068.
Lujano Laura, E., Sosa Sarmiento, J.D., Lujano Laura, A., Lujano Laura, R., 2016. Modelación hidrológica semidistribuida en la región hidrográfica del Titicaca: caso de estudio cuenca del río Ramis, Perú. Rev. Investig. Altoandinas 18, 431–438. https://doi.org/10.18271/ria.2016.235.
Madsen, H., 2000. Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. J. Hydrol. 235, 276–288. https://doi.org/10.1016/S0022-1694(00)00279-1.
Marchau, V.A.W.J., Walker, W.E., Bloemen, P.J.T.M., Popper Editors, S.W. (Eds.), 2019. Decision Making under Deep Uncertainty, Decision Making under Deep Uncertainty. Springer International Publishing. https://doi.org/10.1007/978-3-030-05252-2.
Mark, B.G., Seltzer, G.O., 2003. Tropical glacier meltwater contribution to stream discharge: a case study in the Cordillera Blanca, Peru. J. Glaciol. 49, 271–281. https://doi.org/10.3189/172756503781830746.
MINAM, 2015. Mapa nacional de cobertura vegetal: memoria descriptiva., Primera Ed. ed. Lima, Perú.
Morán-Tejeda, E., Luis Ceballos, J., Peña, K., Lorenzo-Lacruz, J., López-Moreno, I.J., 2018. Recent evolution and associated hydrological dynamics of a vanishing tropical Andean glacier: Glaciar de Conejeras, Colombia. Hydrol. Earth Syst. Sci. 22, 5445–5461. https://doi.org/10.5194/hess-22-5445-2018.
Moussa, R., Chahinian, N., 2009. Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events. Hydrol. Earth Syst. Sci. 13, 519–535. https://doi.org/10.5194/hess-13-519-2009.
Muccione, V., Huggel, C., Bresch, D.N., Jurt, C., Wallimann-Helmer, I., Mehra, M.K., Pabón Caicedo, J.D., 2019. Joint knowledge production in climate change adaptation networks. Curr. Opin. Environ. Sustain. 39, 147–152. https://doi.org/10.1016/j.cosust.2019.09.011.
Muttil, N., Jayawardena, A.W., 2008. Shuffled Complex Evolution model calibrating algorithm: enhancing its robustness and efficiency. Hydrol. Process. 22, 4628–4638. https://doi.org/10.1002/HYP.7082.
Nauditt, A., Birkel, C., Soulsby, C., Ribbe, L., 2017. Conceptual modelling to assess the influence of hydro-climatic variability on runoff processes in data scarce semi-arid Andean catchments. Hydrol. Sci. J. 62, 515–532. https://doi.org/10.1080/02626667.2016.1240870.
Neukom, R., Rohrer, M., Calanca, P., Salzmann, N., Huggel, C., Acuña, D., Christie, D.A., Morales, M.S., 2015. Facing unprecedented drying of the Central Andes? Precipitation variability over the period AD 1000–2100. Environ. Res. Lett. 10, 084017 https://doi.org/10.1088/1748-9326/10/8/084017.
Njue, N., Stenfert Kroese, J., Gräf, J., Jacobs, S.R., Weeser, B., Breuer, L., Rufino, M.C., 2019. Citizen science in hydrological monitoring and ecosystem services management: state of the art and future prospects. Sci. Total Environ. 693, 133531 https://doi.org/10.1016/j.scitotenv.2019.07.337.
Ochoa-Sánchez, A., Crespo, P., Carrillo-Rojas, G., Sucozhañay, A., Célleri, R., 2019. Actual evapotranspiration in the high andean grasslands: a comparison of measurement and estimation methods. Front. Earth Sci. 7, 1–16. https://doi.org/10.3389/FEART.2019.00055.
Ochoa-Tocachi, B.F., Buytaert, W., De Bièvre, B., Célleri, R., Crespo, P., Villacís, M., Llerena, C.A., Acosta, L., Villazón, M., Guallpa, M., Gil-Ríos, J., Fuentes, P., Olaya, D., Viñas, P., Rojas, G., Arias, S., 2016. Impacts of land use on the hydrological response of tropical Andean catchments. Hydrol. Process. 30, 4074–4089. https://doi.org/10.1002/hyp.10980.
Paul, F., Kääb, A., 2005. Perspectives on the production of a glacier inventory from multispectral satellite data in Arctic Canada: Cumberland Peninsula, Baffin Island. In: Annals of Glaciology. International Glaciology Society, pp. 59–66. https://doi.org/10.3189/172756405781813087.
Pellicciotti, F., Buergi, C., Immerzeel, W.W., Konz, M., Shrestha, A.B., 2012. Challenges and uncertainties in hydrological modeling of remote hindu kush–karakoram–himalayan (HKH) basins: suggestions for calibration strategies. Mt. Res. Dev. 32, 39–50. https://doi.org/10.1659/MRD-JOURNAL-D-11-00092.1.
Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. J. Hydrol. 279, 275–289. https://doi.org/10.1016/S0022-1694(03)00225-7.
Perrin, C., Michel, C., Andréassian, V., 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. J. Hydrol. 242, 275–301. https://doi.org/10.1016/S0022-1694(00)00393-0.
Pokhrel, P., Gupta, H.V., 2011. On the ability to infer spatial catchment variability using streamflow hydrographs. Water Resour. Res. 47 https://doi.org/10.1029/2010WR009873.

Polk, M.H., Young, K.R., Baraer, M., Mark, B.G., McKenzie, J.M., Bury, J., Carey, M., 2017. Exploring hydrologic connections between tropical mountain wetlands and glacier recession in Peru's Cordillera Blanca. Appl. Geogr. 78, 94–103. https://doi.org/10.1016/j.apgeog.2016.11.004.

Pouget, J.C., Proaño, D., Vera, A., Villacís, M., Condom, T., Escobar, M., Le Goulven, P., Calvez, R., 2017. Modélisation glacio-hydrologique et gestion des ressources en eau dans les Andes équatoriennes: l′exemple de Quito. Hydrol. Sci. J. 62, 431–446. https://doi.org/10.1080/02626667.2015.1131988.

Ragettli, S., Cortés, G., McPhee, J., Pellicciotti, F., 2014. An evaluation of approaches for modelling hydrological processes in high-elevation, glacierized Andean watersheds. Hydrol. Process 28, 5674–5695. https://doi.org/10.1002/hyp.10055.

Ragettli, S., Pellicciotti, F., 2012. Calibration of a physically based, spatially distributed hydrological model in a glacierized basin: on the use of knowledge from glaciometeorological processes to constrain model parameters. Water Resour. Res. 48 https://doi.org/10.1029/2011WR010559.

Réveillet, M., Rabatel, A., Gillet-Chaulet, F., Soruco, A., 2015. Simulations of changes to Glaciar Zongo, Bolivia (16° S), over the 21st century using a 3-D full-Stokes model and CMIP5 climate projections. Ann. Glaciol. 56, 89–97. https://doi.org/10.3189/2015AoG70A113.

Rodríguez-Morales, M., Acevedo-Novoa, D., Machado, D., Ablan, M., Dugarte, W., Dávila, F., 2019. Ecohydrology of the Venezuelan páramo: water balance of a high Andean watershed. https://doi.org/10.1080/17550874.2019.1673494 12, 573–591. https://doi.org/10.1080/17550874.2019.1673494.

Sankarasubramanian, A., Vogel, R.M., Limbrunner, J.F., 2001. Climate elasticity of streamflow in the United States. Water Resour. Res. 37, 1771–1781. https://doi.org/10.1029/2000WR900330.

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P.A., Carrillo, G., 2011. Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. Hydrol. Earth Syst. Sci. 15, 2895–2911. https://doi.org/10.5194/hess-15-2895-2011.

Schauwecker, S., Rohrer, M., Huggel, C., Endries, J., Montoya, N., Neukom, R., Perry, B., Salzmann, N., Schwarb, M., Suarez, W., 2017. The freezing level in the tropical Andes, Peru: an indicator for present and future glacier extents. J. Geophys. Res. Atmos. 122, 5172–5189. https://doi.org/10.1002/2016JD025943.

Schoolmeester, T., Johansen, K., Alfthan, B., Baker, E., Hesping, M., Verbist, K., 2018. The Andean Glacier and Water Atlas: The Impact of Glacier Retreat on Water Resources. UNESCO and GRID-Arendal.

Seibert, J., 2000. Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. Hydrol. Earth Syst. Sci. 4, 215–224. https://doi.org/10.5194/hess-4-215-2000.

Seibert, J., Vis, M.J.P., 2012. Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. Hydrol. Earth Syst. Sci. 16, 3315–3325. https://doi.org/10.5194/hess-16-3315-2012.

Seong, C., Her, Y., Benham, B.L., 2015. Automatic calibration tool for hydrologic simulation program-FORTRAN using a shuffled complex evolution algorithm. Water 7, 503–527. https://doi.org/10.3390/W7020503.

Sicart, J.E., Hock, R., Ribstein, P., Litt, M., Ramirez, E., 2011. Analysis of seasonal variations in mass balance and meltwater discharge of the tropical Zongo Glacier by application of a distributed energy balance model. J. Geophys. Res. 116, D13105 https://doi.org/10.1029/2010JD015105.

Sicart, J.E., Hock, R., Six, D., 2008. Glacier melt, air temperature, and energy balance in different climates: the Bolivian Tropics, the French Alps, and northern Sweden. J. Geophys. Res. Atmos. 113. https://doi.org/10.1029/2008JD010406.

Sidle, R.C., 2021. Strategies for smarter catchment hydrology models: incorporating scaling and better process representation. Geosci. Lett. 8, 24. https://doi.org/10.1186/s40562-021-00193-9.

Soruco, A., Vincent, C., Rabatel, A., Francou, B., Thibert, E., Sicart, J.E., Condom, T., 2015. Contribution of glacier runoff to water resources of La Paz city, Bolivia (16° S). Ann. Glaciol. 56, 147–154. https://doi.org/10.3189/2015AoG70A001.

Stackhouse, P.W., 2016. Surface Meteorology and Solar Energy (SSE) Data Release 5.1 [WWW Document]. URL https://ntrs.nasa.gov/citations/20080012141. (accessed 8.20.20).

Suarez, W., Macedo, N., Montoya, N., Arias, S., Schauwecker, S., Huggel, C., Rohrer, M., Condom, T., 2015. Balance enérgetico neto (2012-2014) y evolución temporal del nevado Quisoquipina en la región de Cusco (1990-2010). Rev. Peru. Geo-Atmosferica 4, 80–92.

Sucozhañay, A., Célleri, R., 2018. Impact of rain gauges distribution on the runoff simulation of a small mountain catchment in southern Ecuador. Water 10, 1169. https://doi.org/10.3390/w10091169.

Tarasova, L., Knoche, M., Dietrich, J., Merz, R., 2016. Effects of input discretization, model complexity, and calibration strategy on model performance in a data-scarce glacierized catchment in Central Asia. Water Resour. Res. 52, 4674–4699. https://doi.org/10.1002/2015WR018551.

Taylor, L.S., Quincey, D.J., Smith, M.W., Baumhoer, C.A., McMillan, M., Mansell, D.T., 2021. Remote sensing of the mountain cryosphere: Current capabilities and future opportunities for research. Prog. Phys. Geogr 1. https://doi.org/10.1177/03091333211023690.

Tegegne, G., Park, D.K., Kim, Y.-O.O., 2017. Comparison of hydrological models for the assessment of water resources in a data-scarce region, the Upper Blue Nile River Basin. J. Hydrol. Reg. Stud. 14, 49–66. https://doi.org/10.1016/j.ejrh.2017.10.002.

Témez, J.R., 1977. Modelo Matemático de trasformación "precipitación-escorrentía," Asociación de Investigación Industrial Eléctrica (ASINEL). Asociación de Investigación Industrial Eléctrica (ASINTEL), Madrid, España, España.

Tiel, M., Stahl, K., Freudiger, D., Seibert, J., 2020. Glacio-hydrological model calibration and evaluation. WIREs Water 7. https://doi.org/10.1002/wat2.1483.

Turc, L., 1961. Evaluation des besoins en eau d'irrigation. Formule climatique simplifiée et mise à jour. Agro 12, 13–49.

Uhlenbrook, S., Leibundgut, C., 2002. Process-oriented catchment modelling and multiple-response validation. Hydrol. Process. 16, 423–440. https://doi.org/10.1002/hyp.330.

Vaché, K.B., McDonnell, J.J., 2006. A process-based rejectionist framework for evaluating catchment runoff model structure. Water Resour. Res. 42 https://doi.org/10.1029/2005WR004247.

Vansteenkiste, T., Tavakoli, M., Van Steenbergen, N., De Smedt, F., Batelaan, O., Pereira, F., Willems, P., 2014. Intercomparison of five lumped and distributed models for catchment runoff and extreme flow simulation. J. Hydrol. 511, 335–349. https://doi.org/10.1016/j.jhydrol.2014.01.050.

Vuille, M., Carey, M., Huggel, C., Buytaert, W., Rabatel, A., Jacobsen, D., Soruco, A., Villacis, M., Yarleque, C., Elison Timm, O., Condom, T., Salzmann, N., Sicart, J.E., 2018. Rapid decline of snow and ice in the tropical Andes – Impacts, uncertainties and challenges ahead. Earth-Sci. Rev. 176, 195–213. https://doi.org/10.1016/j.earscirev.2017.09.019.

Vuille, M., Kaser, G., Juen, I., 2008. Glacier mass balance variability in the Cordillera Blanca, Peru and its relationship with climate and the large-scale circulation. Glob. Planet. Change 62, 14–28. https://doi.org/10.1016/j.gloplacha.2007.11.003.

Wagnon, P., Ribstein, P., Francou, B., Pouyaud, B., 1999. Annual cycle of energy balance of Zongo Glacier, Cordillera Real, Bolivia. J. Geophys. Res. Atmos. 104, 3907–3923. https://doi.org/10.1029/1998JD200011.

Wang, Y.-C., Yu, P.-S., Yang, T.-C., 2010. Comparison of genetic algorithms and shuffled complex evolution approach for calibrating distributed rainfall–runoff model. Hydrol. Process. 24, 1015–1026. https://doi.org/10.1002/HYP.7543.

Yarleque, C., Vuille, M., Hardy, D.R., Timm, O.E., De la Cruz, J., Ramos, H., Rabatel, A., 2018. Projections of the future disappearance of the Quelccaya Ice Cap in the Central Andes. Sci. Rep. 8, 15564. https://doi.org/10.1038/s41598-018-33698-z.

Yates, D., Sieber, J., Purkey, D., Huber-Lee, A., 2005. WEAP21 - a demand-, priority-, and preference-driven water planning model. Part 1: model characteristics. Water Int. 30, 487–500. https://doi.org/10.1080/02508060508691893.