



# JOURNAL OF NATURAL RESOURCES AND DEVELOPMENT

## Research article

# Combining satellite-based rainfall data with rainfall-runoff modelling to simulate low flows in a Southern Andean catchment

Hann, Hamish<sup>1</sup>; Nauditt, Alexandra<sup>1</sup>; Zambrano-Bigiarini, Mauricio<sup>2</sup>; Thurner, Joschka<sup>1</sup>; McNamara, Ian<sup>1</sup>; Ribbe, Lars<sup>1</sup>

<sup>1</sup> Institute for Technology and Resources Management in the Tropics and Subtropics (ITT), Cologne Technical University, 50679 Cologne, Germany

<sup>2</sup> Department of Civil Engineering, Universidad de La Frontera, Temuco, Chile

Email: alexandra.nauditt@th-koeln.de

## Article history

Received 13/03/2020

Accepted 20/05/2020

Published 02/02/2021

## Keywords

Andean catchments

satellite-based rainfall estimates

HBV-light

SWAT

Sobol sensitivity analysis

hydroPSO

drought management

## Abstract

Since 2010, Chile has been affected by a long-term drought, often referred to as the Mega-Drought. To adapt to future drought events, we need reliable hydro-climatic information based on robust observations and models. However, reliable hydro-climatic data needed for hydrological modelling of remote Andean catchments are rarely available. Our objective was therefore to identify the most suitable combination of different rainfall estimates with two rainfall-runoff models of varying complexity to reliably simulate low flows related to drought.

We forced the widely used SWAT and HBV-light models with four satellite-based rainfall estimates (SREs) (CHIRPSv2, MSWEPv1.2, MSWEPv2.0, TMPA 3B42v7) and inverse distance weighted (IDW) precipitation observations for the main headwater catchment of the Imperial River basin (Araucanía, 38°N). Sobol global sensitivity analysis and calibration were carried out using the hydroPSO R package, using two objective functions focused on the performance of low flows (logNSE and low flows KGE).

Despite the differing amount and temporal distribution of water, for all four SREs, the Sobol sensitivity analysis resulted in robust identification of key model parameters for all SWAT optimisation runs, with most sensitivity for parameters related to deep groundwater recharge, lateral flow travel time and the SCS runoff curve number. For the HBV-light model, only parameters related to snow-melt and deep groundwater recharge were robustly identified for the wetter SRE products. For both models and the two objective functions, the best efficiencies were obtained with the wetter SREs (MSWEP v1.2 and CHIRPSv2) as well as with IDW.

We found that SREs combined with rainfall-runoff modelling provide a valuable instrument to simulate daily discharges for the Upper Imperial River Basin. The resulting models are hence suitable to project low flows and water availability; information that is requested by local decision makers needed to design irrigation systems for the agricultural downstream area and the secure water supply of settlements in the context of future droughts.

## 1 Introduction

In recent years, droughts and water scarcity have affected many regions worldwide (Spinoni et al., 2019). To cope with future droughts at the catchment scale, a deep understanding of local low flow behaviour in dependence of climate variability is needed. However, in the Andean region of Chile, hydro-climatic data are scarce. For the humid and traditionally less drought-prone Central-South, no hydrological studies are available that analyse daily and seasonal water availability and its dependence on climate variability and unknown precipitation inputs from the Andes (Zambrano-Bigiarini et al., 2017). Therefore, robust methods to quantify and project water availability, especially during droughts, need to be developed. Aside from statistical time series analyses, the most common method to achieve this is to use a hydrological model (Schuol et al., 2008; Knoche et al., 2014). Many hydrological rainfall-runoff models of varying complexity are available, ranging from lumped conceptual models to fully distributed raster-based models (Vansteenkiste et al., 2014a; Arnold et al., 2012; Abera et al., 2017; van Der Knijff et al., 2010; Paniconi & Putti, 2015; Devia et al., 2015), each with a different emphasis on flow mechanisms and structural design, which are then implemented in varying process-related algorithms of the model structure (Wei & Chih-Chiang, 2016).

Uncertainties related to model structure and input data used to calibrate a model are a constant topic of debate amongst scientists and engineers (e.g., Uhlenbrook et al., 1999; Refsgaard et al., 2006; Parasuraman & Elshorbagy 2008). Assessing parameter identifiability and input data related uncertainties is particularly relevant when simulating extreme low flows, as parameter behaviour tends to adapt to smaller precipitation inputs and measurement errors are more likely to occur when discharges are low (or very high) (Vansteenkiste et al., 2014a). Thus, to generate locally suitable parameter behaviour, it is becoming common practice to calibrate and compare models with differing structures, before selecting one to be implemented for decision making in a region (te Linde et al., 2008; Golmohammadi et al., 2014; Kaleris et al., 2016; Wei & Chih-Chiang, 2016).

The performance of the models is largely dependent on the quality and availability of the input data. Spatially distributed rainfall is the variable of greatest uncertainty in mountainous regions (Qi et al., 2016; Zambrano-Bigiarini et al., 2017; Baez-Villanueva et al., 2018), where in-situ climate monitoring is sparse and spatial variability is high. More recently, the implementation of satellite-based rainfall estimates (SREs) for hydrological and climatological studies has gained prominence in data-scarce regions (e.g., Dile & Srinivasan, 2014; Simons et al., 2016; Worqlul et al., 2017). The high spatial and temporal resolution of SREs have resulted in them becoming viable datasets for driving hydrological estimates in the absence of reliable ground-based measurements, which are sparse in remote Andean headwater catchments at high elevations.

Such is the case with the Imperial River Basin in southern Chile, which is the study region used as pilot catchment for this study. Although it is a rather wet region with average precipitation of 1200 mm per year, it has been affected by severe droughts in the last decade (Garreaud et al., 2017). To adapt to these recently recurrent summer droughts, there are plans to develop irrigation systems (CNR, 2017) to secure water supply for agriculture. McNamara et al. (2020) used a WEAP model to estimate water availability under potential drought events and irrigation infrastructure in a low-to-medium elevation agricultural sub-catchment of the basin. However, little information exists on the local hydrological and drought characteristics of the headwater catchments, and there have been no hydrological studies to quantify the seasonal availability of water in the entire basin, or its dependence on climate variability and unknown precipitation inputs in the Andes. While the performance of the most relevant SREs have been evaluated against existing rain gauge data within Chile (Zambrano-Bigiarini et al., 2017; Baez-Villanueva et al., 2018), the overall objective of this study is to identify the most robust combination of different satellite-based datasets with two rainfall-runoff models of varying complexity to reliably simulate low flows.

The specific objectives are first to quantify the role of SREs in the performance of the hydrological models HBV-light and SWAT 2.0. Second, to verify a robust model structure and parameter sets to represent hydrological processes in the Imperial River basin using a Sobol sensitivity analysis, and finally, to assess biases in simulated medium and high flows when the hydrological models are calibrated to best represent low flows.

The publicly available input data (Section 3) and open-source modelling tools (Section 4) used in this study will be valuable tools to simulate hydrological drought scenarios and seasonal and long-term low-flow forecasting to support the decision-making process of water users in the Araucanía and similar data-scarce regions.

## 2 Study Area

Our study region is the Upper Cautín (catchment area of 1257 km<sup>2</sup> at Rari-Ruca Station), a sub-catchment of the Imperial River basin, located in the Mediterranean climate of the Araucanía Region of Chile, between 37°51 and 38°56 latitude, with snow influence during spring and summer. The main tributaries of the Imperial are the Cautín, Cholchol, and Quepe rivers, which converge approximately 40 km east of the coast to form the Imperial River, totalling a drainage area of 12,600 km<sup>2</sup>.

To assess hydrological processes, we chose a headwater catchment of the Cautín river (Figure 1), referred to hereafter as the Upper Cautín River Basin (UCRB). The Cautín River originates from precipitation and snowmelt on the western slopes of the southern Andes, at altitudes reaching 3,079 m above sea level (asl), flowing to a westerly direction to drain into the Rari-Ruca streamflow

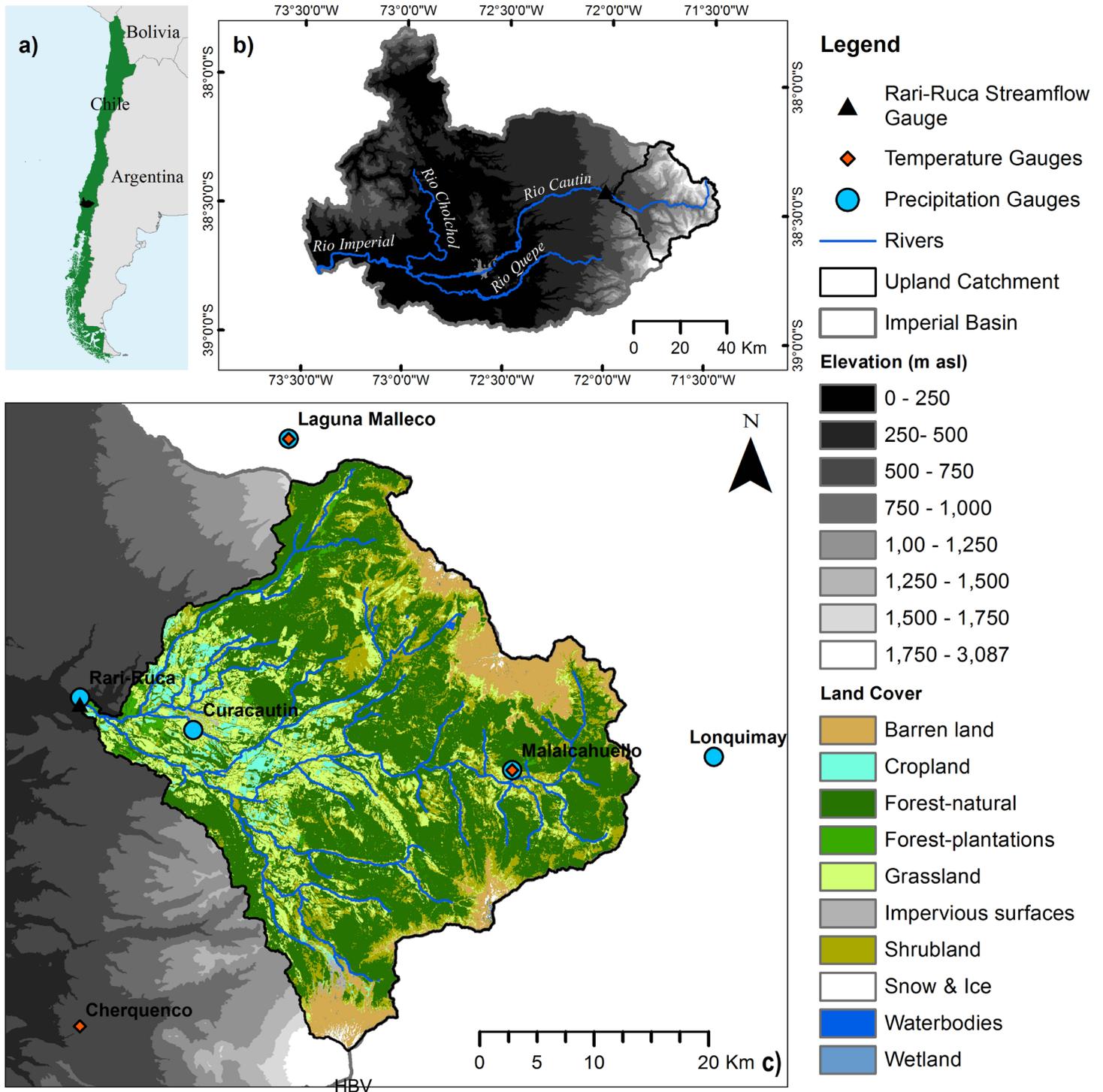


Figure 1: a) Location of the Imperial basin in Chile, b) topography and river network of the Imperial basin, c) Upper Cautín River Basin (UCRB): land cover, river network and location of hydro-climatic stations

station (415 m asl). Seasonal precipitation differs greatly in the UCRB, with high precipitation rates during winter and early spring, and very low precipitation during summer. The seasonal precipitation is reflected in the streamflow of the UCRB, which combined with snowmelt, leads to an abundance of flow throughout winter and spring and very low flows during the dry summer period.

Land cover in the UCRB is dominated by broadleaved forest in the

low to mid elevations interspersed with mixed agriculture and grassland loosely centred on the river channels. In the high elevations, land cover ranges from shrubland to barren areas to permanent snow and ice at the highest elevations. Soil in the catchment is mostly clay loam and silty clay loam (Zhao et al., 2016).

There are settlements within the catchment and much of the population relies on water pumped from local wells. The agricultural

areas are rain-fed, and irrigation is currently almost non-existent, hence the UCRB can be considered undisturbed for hydrological modelling purposes.

Climate projections for the Araucanía region foresee an increased occurrence of flood and drought events (IPCC, 2013; Prudhomme et al., 2013; Boisier et al., 2016). Due to recent drought events, an Irrigation Plan for Araucanía has been developed (CNR, 2017), which contains strategies to increase agricultural activity. A government report from 2012 predicts an increase of agricultural water demand in the region from 74.2 to 1113.2 m<sup>3</sup>/s by 2021 (Silva Rojas et al., 2012). Hence, it is extremely important to develop a water management plan to balance future water availability and demand in the region.

### 3 Data

#### 3.1 Ground-Based Rainfall

Precipitation (P) data were collected from the Dirección General de Aguas (DGA) for 28 rain gauges recording total daily P in the Imperial River basin, all located at elevations below 1,500 m asl. We selected five rain gauges to assess the spatial and temporal

distribution of P in the study catchment (1257 km<sup>2</sup>): Rari-Ruca at an elevation of 440 m asl, Curacautin upstream at 500 m asl., Laguna Malleco in the North of the catchment at 830 m asl. and Malalcahuello and Lonquimay both at 950 m asl. Maximum elevation of the catchment is 3079 m.a.s.l. Linear interpolation was used to remove data gaps of 10 days or less in the precipitation time series. A daily catchment average P dataset was then estimated using inverse distance weighting (IDW) interpolation of the observed data from these four gauges, referred to from here onwards as the IDW dataset. A theoretical 'gauge' was used for input into the models, with its location at the centre of the catchment. For the IDW dataset, the altitude of the 'gauge' was set as the mean altitude of the four P gauges (750 m asl).

#### 3.2 Satellite-Based Rainfall Estimates

There are numerous satellite-based rainfall estimation products (SREs freely accessible in near real-time to the public in a wide spectrum of data formats and spatio-temporal resolutions. Zambrano-Bigiarini et al. (2017) evaluated the performance of seven of these high-resolution SREs in a point-to-pixel evaluation along the latitudinal and longitudinal gradients of Chile. Five of the seven products performed well in the Imperial River basin, with the best results at elevation zones between 0-1,000 m asl. Based on long-

Table 1: Overview on data period, temporal and spatial resolution, coverage and source of the four SREs

SRE	Period	Temporal Resolution	Geographic Coverage	Spatial Resolution	Product description
CHIRPSv2	1981 – ongoing	Daily	Latitude: 50°S–50° N	0.05°	Funk et al., 2015
MSWEPv1.2	1979 – 2015	3-hourly	Global	0.25°	Beck et al., 2017
MSWEPv2.0	1979 – 2016	3-hourly	Global	0.1°	Beck et al., 2019
TMPA 3B42v7	1998 – ongoing	3-hourly	Latitude: 50°S – 50°N	0.25°	Huffman et al., 2007

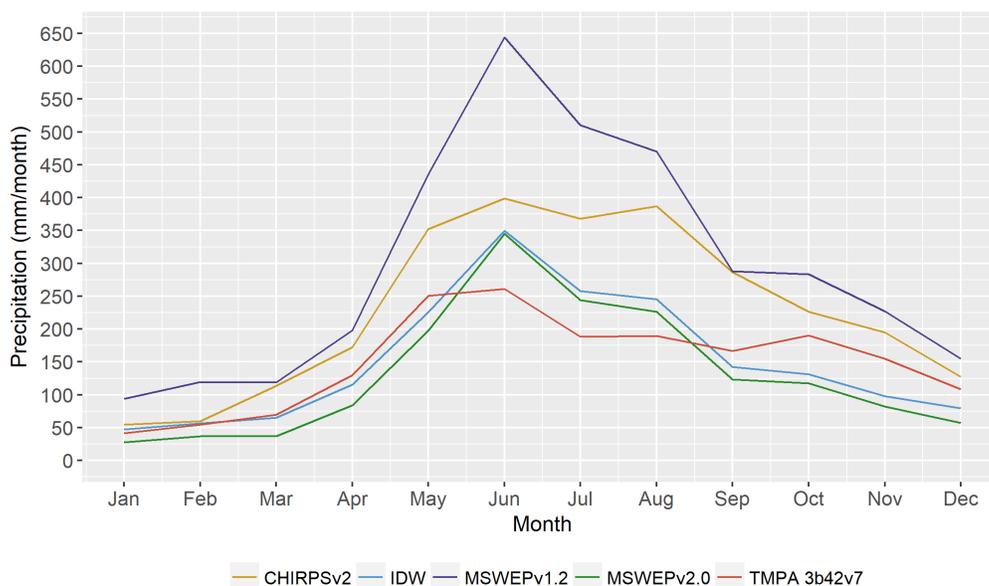


Figure 2: Mean monthly precipitation for the period 2000-2015 for the five precipitation datasets.

term data availability and whether the products are continuously updated, we selected three of these SREs for our assessment: The Climate Hazards Group InfraRed Precipitation with Stations version 2 (CHIRPSv2) (Funk et al., 2015), the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis 3B42 version 7 product (TMPA 3B42v7) (Huffman et al., 2007) and the Multi-Sourced Weighted-Ensemble Precipitation version 1.2 (MSWEPv1.2) (Beck, 2017) (Table 1). In addition, we selected MSWEPv2.0 (Beck et al., 2019), which became available after the Zambrano-Bigiarini et al. (2017) study. For each of the SREs, a daily catchment-average P dataset was calculated, using all grid cells within the catchment boundaries and setting the same location for a 'theoretical' gauge representing the precipitation inputs. The altitude of the 'gauge' was set as the mean altitude of the catchment. Figure 2 shows the long-term monthly average P for each SRE.

### 3.3 Temperature

Average daily temperature ( $T_{mean}$ ) data was non-existent at the climate gauges within the UCRB. To calculate  $T_{mean}$  we used daily catchment average  $T_{max}$  and  $T_{min}$  values calculated by applying the linear lapse rate adjustment (LLRA) method outlined by Dodson and Marks (1997).

### 3.4 Streamflow

Streamflow data (Q) for the period 1929 – 2015 were collected for the Rio Cautín at the Rari-Ruca streamflow gauge (Figure 1) by the DGA. This gauge was used as the outlet of the UCRB and the data was used for the calibration and comparison of the models.

### 3.5 Evapotranspiration

Daily potential evapotranspiration (PET) was estimated by inputting the calculated  $T_{max}$  and  $T_{min}$  data into the Hargreaves equation (Hargreaves & Samani, 1985). Long-term monthly average PET was obtained from Olivera-Guerra et al. (2014).

### 3.6 Land use/land cover

We used the 30 m resolution land use/land cover (LULC) dataset by Zhao et al. (2016). The dataset is based on Landsat 8 imagery from the years 2013-2014, supplemented with Moderate Resolution Imaging Spectroradiometer Enhanced Vegetation Index data, land cover from the Chilean Forestry Services, Google Earth imagery, and 30 m resolution Shuttle Radar Topography Mission DEM data (Zhao et al., 2016).

representing storage in the upper soil zone (SUZ) and storage in the lower soil zone (SLZ). For further details, see Uhlenbrook et al. (1999) and Seibert & Vis (2012).

### SWAT 2.0

The semi-distributed Soil and Water Assessment Tool 2012 (SWAT), developed by the United States Department of Agriculture, incorporates physically based processes related to soils and land cover and has more demanding input needs, requiring more processing power than conceptual models, but it has the potential to address changing spatial characteristics in the catchment (Neitsch et al., 2011; Arnold et al., 2012, Abbaspour et al., 2015).

### 4.2 Goodness of Fit (GOF)

We used two widely used low flow efficiency evaluation criteria: the Inverse Nash-Sutcliffe Efficiency on logarithmic-transformed flows (lnNSE) (Krause et al., 2005) and the Inverse Kling-Gupta efficiency (KGIEf) (Garcia et al., 2017) as objective functions to address low flows.

### 4.3 Sensitivity analysis

The variance-based method of Sobol (Sobol' 1993; Saltelli et al. 2010) was used for global sensitivity analysis of the hydrological models to identify the most relevant parameters to be used during calibration. Sobol is a variance-based sensitivity analysis, where two sensitivity indices are calculated for each parameter: (i) the first-order index  $S_i$ , measuring the direct contribution of parameter  $X_i$  to the total output variance; and (ii) the total-order index  $ST_i$ , representing the sum of all effects involving the parameter  $X_i$ , i.e., its direct effect plus all the interactions with other parameters. These sensitivity indices range from 0 to 1, and the difference between a parameter's first- and total-order indices represents the effects of its interactions with other parameters. Several approaches have been proposed in the literature to numerically compute the sensitivity indices  $S_i$  and  $ST_i$  (e.g., Sobol', 1993; Homma and Saltelli, 1996; Jansen, 1999), and in this study we use the most efficient formulas, as recommended by Saltelli et al. (2010).

For SWAT, 26 parameters were included in the sensitivity analysis, based on previous studies and experience, while for HBV all 12 parameters were used. Parameter ranges for HBV are those defined in Table 2, and the parameter ranges for SWAT are in Table 3.

### 4.4 HydroPSO: Calibration and validation

For all five P datasets, the 10 most relevant parameters identified with the Sobol sensitivity analysis were selected for the calibration of HBV, and for SWAT the eight most sensitive parameters were selected for calibration (Tables 4 and 5). Calibration was performed by implementing the Particle Swarm Optimisation (PSO) technique using the hydroPSO algorithm developed by Zambrano-Bigiarini and Rojas (2013). The hydroPSO algorithm, executed within the R statistical environment (R Core Team, 2016), was chosen as it was specifically designed to perform sensitivity analysis

## 4.4 Methodology

### 4.1 Rainfall-Runoff models

#### HBV-light

The Hydrologiska Byråns Vattenavdelning (HBV) model is a semi-distributed conceptual model with minimal input variables needed to establish a water balance (Seibert, 1997; Seibert & Vis, 2012). HBV-Light consists of a snow routine and two reservoirs

and model calibration on any hydrological model, making it an ideal calibration method for effective comparison between different model performances.

Both models were calibrated for the period from 2000 to 2010, with a one-year warm up period. The performance of the models was evaluated against the measured streamflow data from the Rari-Ruca gauge. We implemented hydroPSO for both GOF functions for each model, resulting in an optimal parameter set for each model and objective function (Figure 3). For SWAT we used sets of 50 particles and hydroPSO was implemented for 200 iter-

ations (10,000 model runs). The conceptual nature of HBV allows for thousands of iterations to be run without overstressing computational power, allowing us to implement sets of 50 particles for 500 iterations (25,000 model runs) within a similar time frame.

The optimised hydroPSO calibrated parameter set from each model was validated by evaluating its performance against a further five years of data (2011-2015) from the Rari-Ruca gauge, again using the KGElf and lnNSE functions as GOF performance indicators.

Table 2: Parameters used in the Sobol sensitivity analysis of HBV.

Parameter	Explanation	Unit	Default value	Lower limit	Upper limit
<i>Snow</i>					
TT	Threshold temperature	°C	0	-2.5	2.5
CFMAX	Degree-day factor	mm°C <sup>-1</sup> d <sup>-1</sup>	3	0.5	9
SFCF	Snowfall correction factor	-	1	0.2	1.2
SP	Seasonal variability in degree-Δt factor	-	1	0	1
<i>Soil and evaporation routine</i>					
FC	Maximum soil moisture	mm	200	25	1000
LP	Soil moisture threshold for reduction of evaporation	-	1	0	1
BETA	Shape coefficient	-	1	1	6
<i>Groundwater and response routine</i>					
K1	Recession coefficient	d <sup>-1</sup>	0.1	0.005	0.1
K2	Recession coefficient	d <sup>-1</sup>	0.05	0.0005	0.1
PERC	Maximal flow from upper to lower GW-box	mm d <sup>-1</sup>	1	0	12
Alpha	Non-linearity coefficient	-	0	0	1
MAXBAS	Routing, length of weighting function	d	1	1	2.5

**Table 3:** Parameters in the Sobol sensitivity analysis of the SWAT2012 model.

Parameter	Description	Unit	Default value	Lower limit	Upper limit
SFTMP	Snowfall temperature	°C	1	0	5
SNO50COV	Fraction of snow volume represented by SNOCOVMX that corresponds to 50% snow cover	-	0.5	0	0.9
ESCO	Soil evaporation compensation factor	-	0.95	0.01	1
SURLAG	Surface runoff lag time	d	4	1	12
SMTMP	Snowmelt base temperature	°C	0.5	-5	5
SMFMN	Minimum melt factor for snow	°C	4.5	1.4	6.9
TIMP	Snowpack temperature lag factor	-	1	0.01	1
SMFMX	Maximum melt factor for snow	°C	4.5	1.4	6.9
EPCO	Plant uptake compensation factor	-	1	0.01	1
ALPHA_BF	Baseflow alpha factor	d	0.048	0.01	0.99
RCHRG_DP	Deep aquifer percolation fraction	-	0.05	0	1
GWQMN	Threshold water depth in the shallow aquifer for flow	mm	0	0	5000
GW_REVAP	Groundwater "revap" coefficient	-	0.02	0	0.2
GW_DELAY	Groundwater delay time	d	31	0	500
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" to occur	mm H <sub>2</sub> O	750	1	1000
LAT_TTIME	Lateral flow travel time	d	0	0	180
OV_N	Manning's "n" value for overland flow	-	0.1	0.01	30
CANMX	Maximum canopy storage	mm H <sub>2</sub> O	0	1	10
HRU_SLP	Average slope steepness	mm <sup>-1</sup>	0.2	0	1
CN2	SCS runoff curve number	-	55	40	95
CH_K2	Effective hydraulic conductivity in main channel alluvium	mm h <sup>-1</sup>	0	0	200
CH_N2	Manning's "n" value for the main channel	-	0.014	0	0.3
SOL_AWC	Available water capacity	mm H <sub>2</sub> O mm <sup>-1</sup>	0.15	0.01	0.35
SOL_K	Saturated hydraulic conductivity	mm hr <sup>-1</sup>	50	0.001	1000
SOL_ALB	Moist soil Albedo	-	0.15	0	0.25
USLE_K	USLE equation soil erodibility factor	-	0.2	0	0.65

#### 4.5 Assessing Model Uncertainties and Flow Bias

We evaluated model uncertainties by analysing the 95% prediction uncertainty band (95PPU). The 95PPU can be used to assess the accuracy of the calibration by finding the fraction of the band that brackets the observed streamflow data, known as the P-factor (Abbaspour et al., 2007). The value for P-factor ranges from 0 to 1, where 1 indicates 100% of measured data bracketed within model prediction uncertainty. The 95PPU was calculated by setting a threshold GOF value of 0.0 and simulations which achieved a GOF above this threshold were used to calculate the 2.5 and 97.5 percentiles of the cumulative distribution of every simulated point (Abbaspour et al., 2004). The P-factors were also calculated separately for low, medium, and high flows (using 5th, 50th, and 95th percentiles, respectively) to determine differences in uncertainty in the three flow-spectra due to the use of low-flow objective functions.

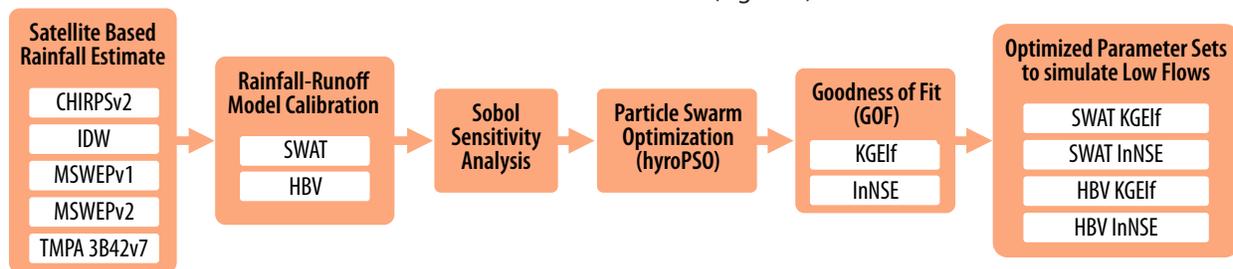


Figure 3: Overall methodology

## 5 Results

### 5.1 Sobol Sensitivity Analysis

The results of the Sobol sensitivity analysis for HBV using KGEIf as the objective function are shown in Table 4. We obtained 10 sensitive parameters to be used for calibration. For all P datasets, the most sensitive parameters were the snowmelt process related SCFC (Snow Correction Factor) and TT (Snow Melt Temperature Threshold) parameters and the groundwater parameter K2 (lower groundwater box recession threshold) followed by BETA, CFMAX, SP, PERC, LP, FC, Alpha and K1 (parameter descriptors in Table 2). Table 5 shows that the eight most sensitive parameters for SWAT are the runoff generating parameters: RCHRG\_DP (Deep aquifer percolation fraction), CN2 (SCS runoff curve number) and LAT\_TTIME (lateral flow travel time) followed by soil and vegetation related parameters SOL\_K, SOL\_AWC, CANMX and groundwater and baseflow generating parameters GW\_DELAY and ALPHA\_BF (parameter descriptions in Table 3).

### 5.2 Model calibration and validation

Figure 4 shows bar graphs of the objective function (GOF) results for calibration and validation of both model types forced with each of the P datasets. Both the HBV and SWAT models simulated streamflow reasonably well, with the MSWEPv1.2 and CHIRPSv2 forced models performing best. KGEIf obtained better results for all models (calibration and validation). Calibration with MSWEPv1.2

We also used these percentiles to produce Empirical Cumulative Density Functions (ECDFs) from hydroPSO to quantify the bias in the three flow-spectra. The ECDFs were created using a subset of the hydroPSO simulations for GOFs > 0.5 and were compared to the observed percentiles to derive percent biases (Zambrano-Bigiarini & Rojas, 2013). A positive bias indicates overestimation of flow, while a negative bias indicates underestimation of flow.

### 4.6 Overall methodology

The five P datasets were used to calibrate the two models. Then, we applied the Sobol sensitivity analysis in the hydroPSO tool to evaluate model parameter sensitivity and parameter identifiability. We then implemented a Particle Swarm Optimisation (PSO) technique to calibrate the models, using the two objective functions (GOFs) to analyse the performance of the models. The process outcome is four sets of optimised parameters for each P dataset, two for each model (Figure 3).

gave the best results when using the HBV model (InNSE = 0.89 and KGEIf = 0.93). The IDW forced models performed reasonably well considering the P gauges are all below 1000 m asl. The validation results for three of the four IDW forced models are higher than the calibration values. The MSWEPv2.0 and TMPA 3B42v7 forced models performed very poorly when forcing HBV and returned GOF function values  $\leq 0.5$  for the calibration period. Both SREs performed better when forcing SWAT, with TMPA 3B42v7 giving higher GOF values than MSWEPv2.0.

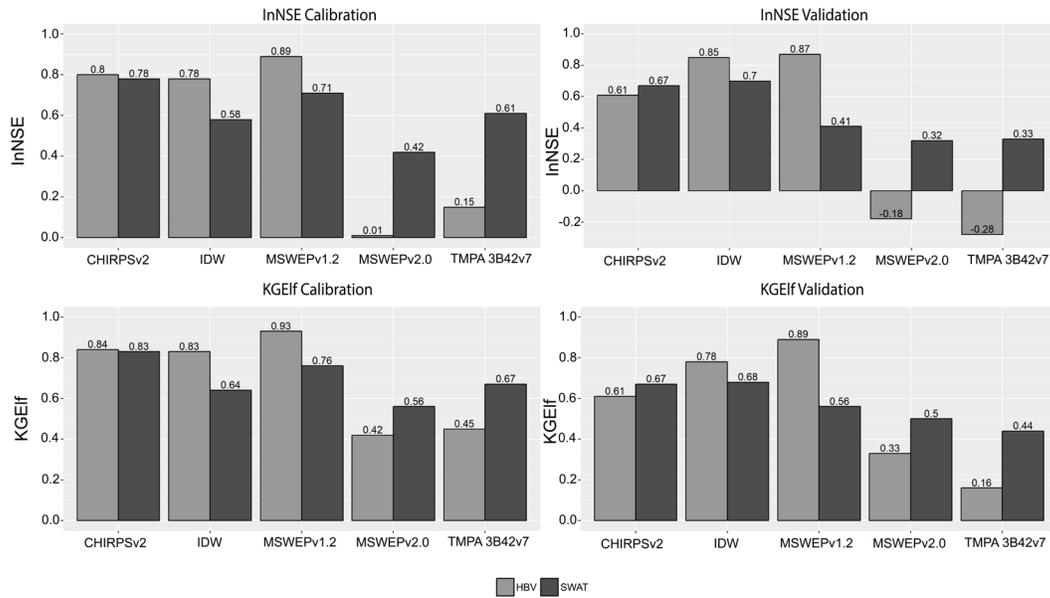
MSWEPv2.0 performed slightly better than TMPA 3B42v7 during validation. The validation GOF values are all below 0.5 except for SWAT forced with MSWEPv2.0, which gave KGEIf = 0.50 when calibrated using KGEIf as the objective function. Therefore, these models could not be validated to an acceptable level. MSWEPv1.2 performed almost as well during validation as during calibration when forcing HBV, with only a difference of 0.02 and 0.04 for InNSE and KGEIf, respectively. MSWEPv1.2 performed much better when forcing HBV than when forcing SWAT. This can be seen in the validation bar graphs (Figure 4) which show that the GOF values for the validation of HBV (InNSE = 0.87, KGEIf = 0.89) obtained almost double the values than SWAT (InNSE = 0.41, KGEIf = 0.56). CHIRPSv2 showed similar results in validation for both objective functions and models (0.61 and 0.67 respectively). Performance of the MSWEPv2.0 and TMPA 3B42v7 forced models for calibration and validation was better for KGEIf than for InNSE.

**Table 4:** Results of the Sobol sensitivity analysis of the HBV model using KGEIf. Rankings are in descending order of most sensitive parameters to least sensitive.

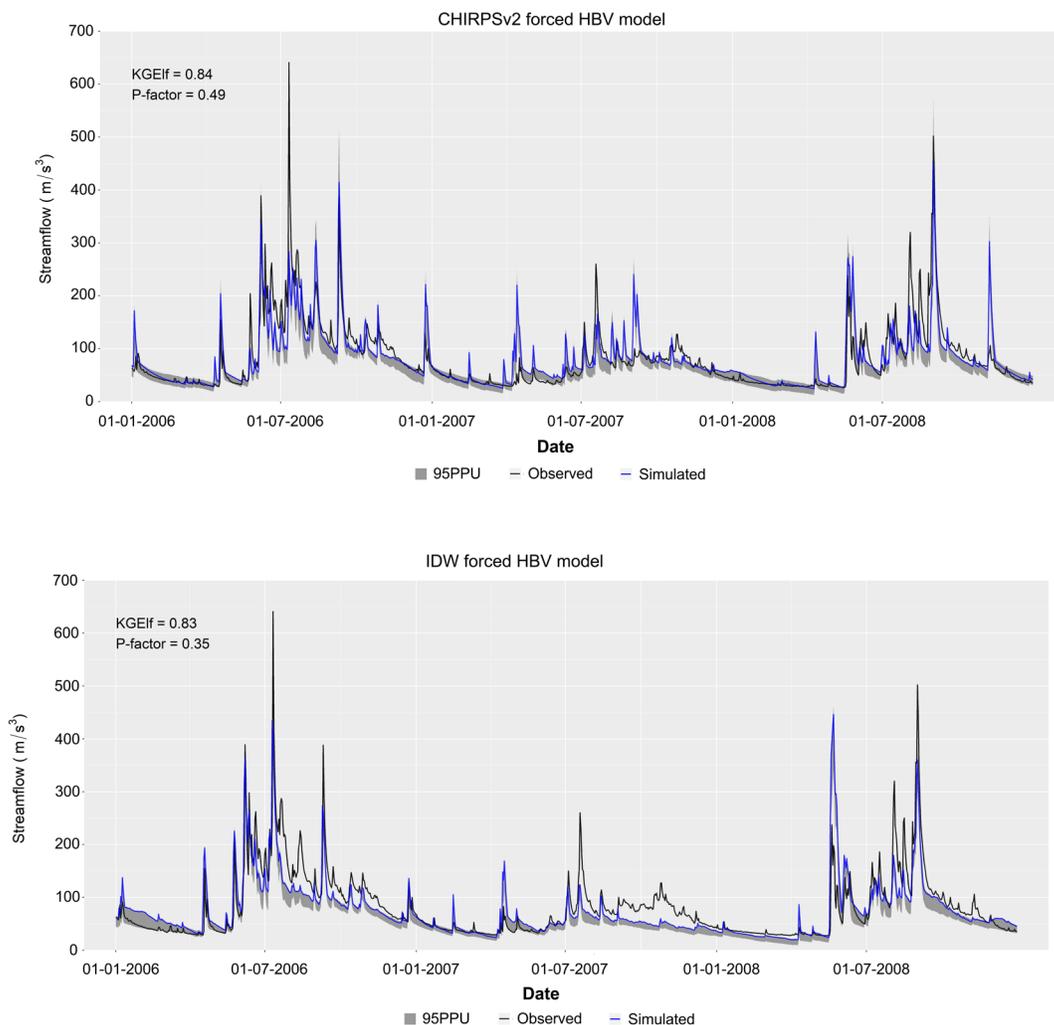
Parameter	CHIRPSv2 Ranking	IDW Ranking	MSWEPv1.2 Ranking	MSWEPv2.0 Ranking	TMPA 3B42v7 Ranking
SFCF	1	1	1	2	4
TT	2	2	2	3	2
K2	3	3	3	1	1
BETA	5	4	4	4	3
CFMAX	8	5	6	5	5
SP	6	6	7	6	6
PERC	4	7	5	8	8
LP	10	8	11	7	7
FC	11	9	10	9	9
Alpha	7	10	8	10	10
K1	9	11	9	11	11
MAXBAS	12	12	12	12	12

**Table 5:** Results of the Sobol sensitivity analysis of the SWAT model using KGEIf. Rankings are in descending order of most sensitive parameters to least sensitive.

Parameter	CHIRPSv2 Ranking	IDW Ranking	MSWEPv1.2 Ranking	MSWEPv2.0 Ranking	TMPA 3B42v7 Ranking
RCHRG_DP	2	1	1	1	1
CN2	1	2	2	3	2
LAT_TTIME	3	3	3	2	3
SOL_K	5	4	4	4	5
SOL_AWC	4	6	5	5	6
CANMX	6	5	7	6	4
GW_DELAY	8	7	8	9	7
ALPHA_BF	9	9	6	7	8
HRU_SLP	7	8	9	8	9
GW_REVAP	13	10	10	11	10
CH_K2	12	12	13	12	14
OV_N	11	13	12	14	13
SFTMP	10	15	11	15	12
GWQMN	15	11	19	10	11
ESCO	14	14	14	13	15
CH_N2	16	16	17	17	17
SMTMP	17	18	15	18	16
REVAPMN	19	19	18	20	18
SMFMN	18	20	16	21	20
EPCO	21	17	22	16	22
SMFMX	20	21	20	19	19
TIMP	22	22	21	23	21
SOL_ALB	23	23	23	22	23
SNO50COV	24	24	24	24	24
USLE_K	25	25	25	25	25
SURLAG	26	26	26	26	26



**Figure 4:** InNSE and KGEIf objective function (GOF) performance for the calibration and validation periods and the five Precipitation datasets. The numbers above each bar show the GOF value.



**Figure 5:** Hydrographs showing a three-year period of the calibration period for the combinations HBVlight and CHIRPSv2 and HBVlight and IDW. The black curve is the observed streamflow captured at Rari-Ruca, the blue curve is the optimal simulation achieved during hydroPSO calibration, and the grey curve is the 95% prediction uncertainty. The goodness of fit and P-factor values were calculated for the entire calibration period.

### 5.3 Model Uncertainties and hydrograph performance

Figure 5 displays hydrographs for the two best performing model-P data set combination. The P-factor was calculated for the entire calibration period, while the hydrographs below show a snapshot of three years of this period. The CHIRPSv2 and IDW P-factors are low considering the high KGEIf values.

When the P-factors were calculated separately for low, medium, and high flows (Figure 6), the results differed on how well

the 95PPU band brackets the observed flow when each of the objective functions are used. Figure 6 shows the P-factors for the three flow-spectra. The P-factors for each model are much higher when only low-flows are considered and generally higher when KGEIf is used as the objective function. A similar trend can be seen in the high-flow P-factors, although they are much lower than the low-flow P-factors.

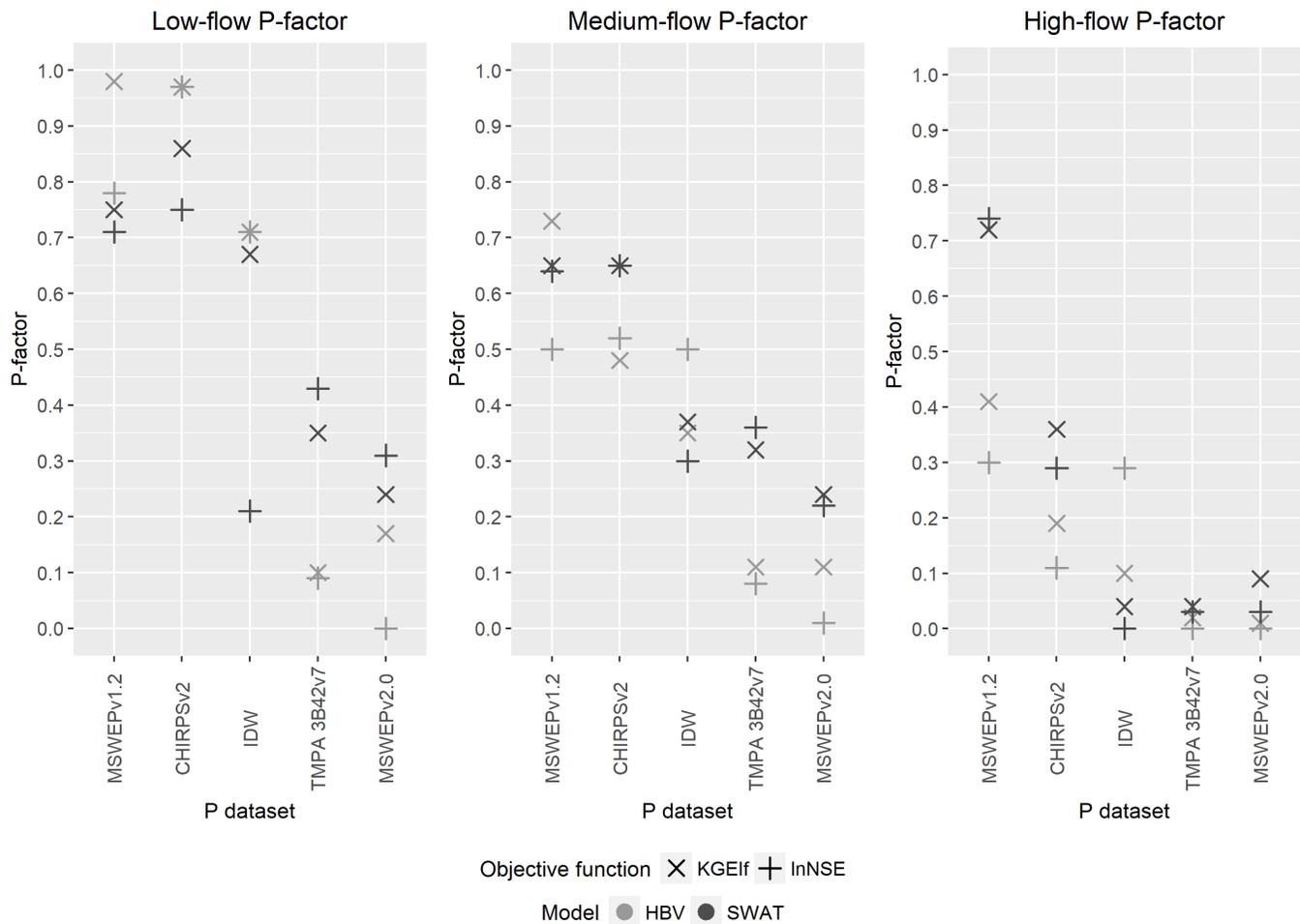


Figure 6: P-factors calculated separately for low-, medium-, and high-flows

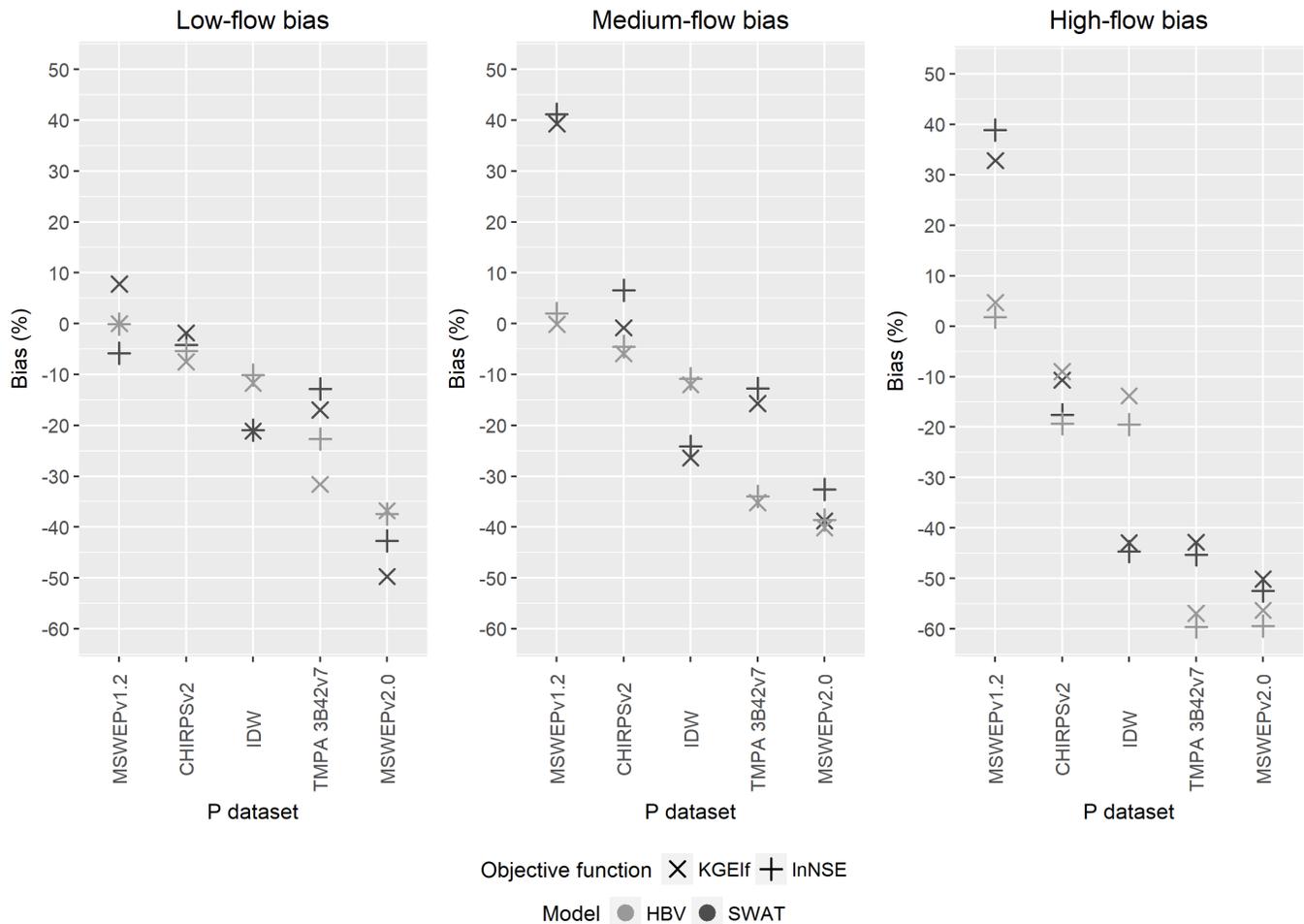
#### 5.3.1 Analysis of low-flows using ECDF bias

Figure 7 shows the percent biases, calculated from the ECDFs, for the P datasets when forcing each model and calibrated using each objective function. Overall, we observed that the models calibrated using lnNSE performed better at simulating low-flows than the models calibrated using KGEIf as the objective function. Simulated streamflow from CHIRPSv2, IDW, MSWEPv2.0, and TMPA 3B42v7 tended to underestimate low-flow periods when forcing both SWAT and HBV. MSWEPv1.2 had 0% bias when forcing HBV regardless of the objective function used during calibration and is consistent with these models achieving the highest GOF values. When forcing SWAT, MSWEPv1.2 tended to slightly overestimate low-flows when KGEIf was used (7.8%) and slightly underestimate low-

flows when lnNSE was used (-5.8%). The opposite was observed with CHIRPSv2, which tended to underestimate low-flows to a lesser extent when forcing SWAT compared to when it was used to force HBV. IDW underestimated low-flows by around 10% when forcing SWAT and around 20% when forcing HBV, regardless of the model or objective function used during calibration. MSWEPv2.0 and TMPA 3B42v7 both significantly underestimated low-flows.

#### 5.3.2 Medium and high flow simulation performance

Similar trends are observed with medium flows. Medium flows were simulated better overall with lnNSE compared to KGEIf; however, were not simulated as accurately as low flows for most models. The MSWEPv1.2 forced SWAT models overestimated medium



**Figure 7:** Flow bias for a. Low-flows, b. Medium-flows, and c. High-flows based on the 5th, 50th, and 95th percentiles of the empirical cumulative distribution function, respectively, for the simulated discharges in the Upper Cautin River Basin during the calibration period.

flows by around 40% and the newer version, MSWEPv2.0, underestimated medium flows when forcing both models by 30-40%. TMPA 3B42v7 underestimated medium flows more when forcing HBV than SWAT, which is the opposite for IDW, which underestimated medium flows more when forcing SWAT than HBV. The MSWEPv1.2 forced HBV model simulated medium flows the best. When calibrated using KGEIf, medium flows were simulated with almost no bias (-0.1%) and MSWEPv1.2 only slightly overestimated medium flows (2%) when using InNSE. The CHIRPSv2 forced SWAT model calibrated using KGEIf performed almost as well as the MSWEPv1.2 forced HBV models and only underestimated medium flows by 0.8%. When InNSE was used to calibrate the CHIRPSv2 forced SWAT model, flow was overestimated by 6.6%. High flows were better simulated using KGEIf. MSWEPv2.0, IDW and TMPA 3B42v7 both greatly underestimated high flows. MSWEPv1.2 tended to overestimate high flows only slightly when forcing HBV and greatly overestimated high flows when forcing SWAT. CHIRPSv2 simulated high flows more poorly and underestimated high flows by 10% using KGEIf and by 20% when InNSE was used as the objective function.

### 5.3.3 Water Balance

The over- and under-estimations illustrated above become clear when considering how well each P dataset closes the water balance (Table 6). The average annual observed streamflow in the UCRB is 2,209.64 mm/year. For the IDW, MSWEPv2.0 and TMPA 3B42v7 forced models, the average annual P is under 2,000 mm/year. Looking first at the HBV forced models, this resulted in very small amounts of annual AET and an annual streamflow much lower than the observed value. The SWAT models forced with these P datasets produced higher values for AET. This was not at the expense of streamflow, which tended to be higher than the annual streamflow of the respective HBV model value. The MSWEPv2.0 and TMPA 3B42v7 forced SWAT models underestimate annual streamflow by at least 900 mm/year. The CHIRPSv2 and MSWEPv1.2 forced HBV models show a much higher annual P than the other three P datasets and produced much higher values for annual streamflow. The CHIRPSv2 forced HBV model was unable to close the water balance due to the low AET. The CHIRPSv2 forced SWAT model was able to close the water balance. The volume difference between simulated annual streamflow and observed annual streamflow was lowest for the MSWEPv1.2 forced HBV model calibrated using InNSE (12.74 mm/year).

**Table 6:** Average annual values (mm/year) simulated by each model forced with each precipitation dataset using lnNSE and KGEIf as the objective function during the periods 2001-2010.

	Water balance process (mm/y)	CHIRPSv2		IDW		MSWEPv1.2		MSWEPv2.0		TMPA 3B42v7	
		KGEIf	lnNSE	KGEIf	lnNSE	KGEIf	lnNSE	KGEIf	lnNSE	KGEIf	lnNSE
HBV	P	2841.31	2841.31	1839.24	1839.24	3802.55	3802.55	1607.75	1607.75	1922.95	1922.95
	Qsim	2025.64	1944.92	1960.79	1926.90	2190.04	2196.90	1193.29	1185.22	1276.33	1271.30
	AET	352.45	432.19	355.30	368.70	821.04	755.57	216.56	219.26	346.18	349.30
	Qobs – Qsim	184.00	264.71	248.85	282.74	19.59	12.74	1016.35	1024.42	933.31	938.34
SWAT	P	2741.40	2741.40	1814.40	1814.40	3545.80	3545.80	1579.20	1579.20	1807.80	1807.80
	Qsim	2068.30	2085.75	1549.59	1544.22	2330.55	2384.95	1997.54	1357.90	1598.68	1618.24
	AET	700.80	682.90	541.40	547.60	728.40	733.20	586.30	512.30	648.60	631.10
	Qobs – Qsim	141.34	123.89	660.05	665.42	-120.91	-175.31	212.10	851.74	610.96	591.40

## 6 Discussion

HydroPSO optimisation obtained very good statistical efficiencies for KGEIf and lnNSE for the MSWEPv1.2, CHIRPSv2 and IDW datasets. The hydroPSO tool was effective in terms of low computational power and time requirements and few iterations (25,000) needed to obtain a satisfactory result.

### 6.1 Model performance and role of objective function

HBV-light obtained better efficiencies compared to SWAT forced with CHIRPSv2, IDW, and MSWEPv1.2. However, HBV was unable to be satisfactorily calibrated when forced with MSWEPv2.0 and TMPA 3B42v7. This might be explained by the higher complexity of the SWAT model structure. Previous studies have focused on whether performance increases with model complexity: te Linde et al. (2008) found that the HBV model performed better than the more complex VIC distributed land-based model in the Rhine River Basin. Finger et al. (2015) showed that the HBV model obtained better results compared to the spatially distributed TOP-KAPI model, and demonstrated that model complexity does not enhance model performance if the input data does not contain the appropriate information. This suggests that if discharge is the only target variable, lumped models can be more appropriate, especially when considering the lower computation time and data requirements. However, if physical processes related to soil and LULC are targeted, a more distributed model such as SWAT is necessary (Orth et al., 2015).

The effectiveness of the models was influenced by the objective functions that we selected for calibration. When calibrating the two wetter and best performing SREs (CHIRPSv2 and MSWEPv1.2), simulated streamflow showed only minor low-flow bias regardless of the objective function used. For the remaining SREs and the IDW dataset, low flow and medium flow bias is lower when lnNSE is used. However, calibrating with lnNSE produced greater over- or under-estimation of high flows compared to KGEIf. This is due to the combined nature of the KGEIf function (Garcia et al., 2017). Garcia et al. (2017) found that while low flows were better simulated with KGE(1/Q), simulations overall were better when

combining it with KGE, due to KGE placing greater emphasis on higher flows.

We can conclude similar findings for this study when comparing KGEIf with lnNSE; KGEIf results in more accurate streamflow simulations when considering all flow-spectra biases and this is also reflected in the GOF values (Figure 4), which tended to be higher with KGEIf compared to lnNSE.

### 6.2 Model uncertainties

We evaluated model uncertainties by analysing the 95% prediction uncertainty band (95PPU) and the P-factors, finding that the P-factors for each model were much higher when only low flows were considered and generally higher when KGEIf was used as the objective function. These values are high enough (e.g., KGEIf 0.93; P-factor 0.72) to conclude that the model uncertainty is reasonably low and any low flow simulations can be considered with an adequate level of confidence. On the other hand, the high flow P-factors are much lower than the low flow P-factors and the model uncertainty is too high for most SREs. Therefore, we cannot be as confident in the simulated values for high flows. Because the focus of this study is low flows related to drought, this outcome is considered to be acceptable. The P-factors for each model show how in the search for the optimal fit, the iterative nature of hydroPSO decreases parameter uncertainty at the cost of capturing model uncertainties within the 95% prediction uncertainty band (Abbsapour et al., 2015). Due to the ease of hydroPSO to alter the desired objective function, it would be a simple exercise to include the P-factor as a weighted index in the objective function. Thus, in future studies, a balance could be found between hydrograph optimisation and accounting for model uncertainty.

### 6.3 Parameter sensitivity (Sobol)

We found only minor differences in parameter sensitivity when using different objective functions and precipitation estimates (see Table 4 and 5 for the results of Sobol with KGEIf as the objective function). The most sensitive parameters were those related to the snowmelt processes (SCFC and TT) and groundwater

processes (K2). For SWAT, the most sensitive parameters were the baseflow and groundwater related parameters RCHRG\_DP, CN2, and LAT\_TTIME. Both models also showed a lower sensitivity for soil and vegetation related parameters (FC, LP and BETA for HBV, and SOL\_K, SOL\_AWC, CANMX for SWAT) followed by further groundwater and baseflow generating parameters (PERC and Alpha for HBV, and GW\_DELAY and ALPHA\_BF for SWAT). Both models hence use the streamflow generating parameters for calibration while HBV additionally uses the temperature dependent snowmelt related parameters. This coincides with earlier HBV-Light applications in the Andes of Central Chile, that resulted in a Snow-Groundwater model governed by streamflow generating and snowmelt related parameters (Nauditt et al., 2016).

#### 6.4 Hydrological evaluation of Satellite-based Rainfall Estimates (SREs)

MSWEPv1.2 performed the best when forcing both models. HBV forced with MSWEPv1.2 performed better than SWAT forced with MSWEPv1.2, achieving the highest values for KGEIf and lnNSE. It also had the least simulation bias when looking at the three flow spectra. Both models closed the water balance, mainly due to the high P amounts available to be utilised by each model. As shown in Figure 2, annual average P estimated by MSWEPv1.2 is much higher than the other SREs. When forced with MSWEPv1.2, both SWAT and HBV took advantage of the high amounts of P to close the water balance and calibrated the groundwater parameters to encourage storage and percolation of the excess water during hydroPSO optimisation. CHIRPSv2 also performed well and performed better when forcing SWAT than HBV. This was due to the ability of SWAT to more accurately simulate the three flow spectra, leading to better simulations of the validation period. In contrast, the IDW dataset only performed well when forcing HBV and not when forcing SWAT.

In a point to pixel evaluation of different SREs (Zambrano-Bigiarini et al., 2017), TMPA 3B42v7 showed a good overall performance in the central south region of Chile, but this result was not confirmed by our hydrological evaluation of the product. The same can be said for MSWEPv2.0, which also performed poorly. Baez-Villanueva et al. (2018) compared MSWEPv2.0 to observed data in the Imperial River basin and found that it performed as well as CHIRPSv2. The discrepancies between the point-to-pixel comparisons and the findings from this study could be: i) because the SREs (including CHIRPSv2) use observed P for bias correction during the estimation process; and ii) there are no stations above 1,500 m asl in the region. Orographic precipitation occurs in the UCRB (Viale & Nuñez, 2001) and these effects are not captured at higher elevations due to the lack of observation stations. This could lead to underestimations of P in these high elevation areas, especially for the TMPA 3B42v7 and MSWEPv2.0 datasets, which appear to use observation stations for bias correction more rigorously (see Figure 2).

#### 6.5 Model Water Balance: evapotranspiration (AET) and precipitation estimates

The SWAT models produced much higher values for actual evapotranspiration (AET) than the HBV models (Table 6). The reason for this cannot be the PET estimated by each model, as the annual PET estimated by HBV is 1,347 mm/year, while the PET estimated by SWAT is approximately 820 mm/year. The only available reference for long-term annual PET in the region is a DGA (1988) publication, which calculated pan evapotranspiration to be 1,300 mm/year at Rari-Ruca station. Therefore, HBV produces more accurate values of PET than SWAT, for which the Hargreaves method is used. Olivera-Guerra et al. (2014) estimated AET in Chile using satellite imagery. They estimated that the annual AET in the UCRB is 717 mm/year. As Table 6 shows, the AET produced by the HBV model is much lower than this value for all models except when forced with moisture rich MSWEPv1.2. The AET produced by the SWAT models is closer to the AET that was estimated by Olivera-Guerra et al. (2014).

SWAT and HBV both have a precipitation lapse rate parameter that can be included, which could allow for some of this lost P to be recovered. However, we kept the lapse rates set as the default for each model to enable a better comparison of the performance of the models and the SREs used as P inputs. The default P lapse rate for HBV (PCLAT) is 10%/100 m while the default for SWAT (PLAPS) is 0 mm/km. Tuo et al. (2016) found that calibrating the PLAPS parameter improved simulated streamflow and SWAT model performance increased by at least 9%.

The differences can also be seen in how well each P dataset was able to close the water balance. All P datasets underestimated annual streamflow when forcing both models, except for the MSWEPv1.2 forced SWAT model, which slightly overestimated annual streamflow. MSWEPv1.2 and CHIRPSv2 estimate the highest annual and mean monthly rainfall (Figure 2) so it is not surprising that these SREs performed better than the drier MSWEPv2.0 and TMPA 3B42v7. The latter also have the largest volume differences of the four datasets. Dinku et al. (2009) also evaluated TMPA 3B42v7 and found that it underestimated both the occurrence and amount of rainfall in their Columbian study catchment, which they partially attributed to orographic processes over the Andes. However, Tuo et al. (2016) considered orographic effects in their study by including elevation bands and calibrated the PLAPS parameter, yet still achieved unsatisfactory results when calibrating SWAT with TMPA 3B42v7.

## 7 Conclusion

Accurately modelling sparsely monitored headwater catchments is vital for water management in drought prone regions. We tested the combination of different SREs and modelling tools to simulate discharge from a data-scarce snowmelt driven pilot catchment in south-central Chile. By forcing two hydrological models of differing complexity with data from four SREs, we assessed if remote sensing data can be used as an alternative to ground-based observation data.

The wet MSWEPv1.2 data product performed best out of the SREs followed by CHIRPSv2, with the IDW dataset also achieving good simulations when forcing HBV. Both InNSE and KGEI performed reasonably well at simulating low-flows, as evidenced by a low bias in the MSWEPv1.2 and CHIRPSv2 simulations. As CHIRPSv2 is continuously updated and considering MSWEPv1.2 is only available until the end of 2015, we recommend using CHIRPSv2 for further studies in the Imperial River basin.

As further research steps in the region, we recommend the inclusion of the P-factor in the objective function during hydroPSO calibration. This updated method could be applied to other headwater catchments. This method could also be applied to downstream catchments, where CHIRPSv2 may perform better according to Zambrano-Bigiarini et al. (2017) and Baez-Villanueva et al.

(2018) as well as TMPA 3B42v7 and MSWEPv2.0. Furthermore, recent P products such as MSWEPv2.2 (Beck et al., 2019), the updated gridded CR2-MET, ERA5 (Hersbach et al., 2020) and RF-MEP applied over Chile (Baez-Villanueva et al., 2020) should be evaluated. Finally, the calibrated models could be forced with short and long term climate projections as well as artificial climate scenarios to simulate future drought periods.

We conclude that the methodology suggested here represents a transferable set of tools to be applied in data-scarce headwater catchments along the Chilean latitudinal gradient. Combining best performing SREs, hydrological modelling, Sobol sensitivity analysis, hydroPSO calibration and uncertainty analysis represent a valuable instrument for drought and water management in data-scarce regions. It relies on open access data, open source scripts and models that are well documented and require little computation time. Moreover, our findings provide important information on the character of low flows, valuable for drought management in the region. Our approach can be easily transferred to other study regions to allow the improved modelling of low flows, and consequently, provide valuable information and knowledge in drought management planning.

## References

- Abbaspour, K. C., Johnson, C. A., & van Genuchten, M. Th. (2004). Estimating Uncertain Flow and Transport Parameters Using a Sequential Uncertainty Fitting Procedure. *Vadose Zone Journal* 3 (4), pp. 1340–1352. doi: 10.2113/3.4.1340.
- Abbaspour, K.C., J. Yang, I. Maximov, R. Siber, K. Bogner, J. Mieleitner, J. Zobrist, & R. Srinivasan. (2007). Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology*, 333:413-430. doi: 10.1016/j.jhydrol.2006.09.014
- Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., & Kløve, B. (2015). A continental-scale hydrology and water quality model for Europe. Calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology* 524, pp. 733–752. doi: 10.1016/j.jhydrol.2015.03.027.
- Abera, W., Formetta, G., Brocca, L., & Rigon, R. (2017). Modeling the water budget of the Upper Blue Nile basin using the JGrass-NewAge model system and satellite data. *Hydrology and Earth System Sciences*, 21(6), 3145. doi:10.5194/hess-21-3145-2017
- Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R. et al. (2012). SWAT: MODEL USE, CALIBRATION, AND VALIDATION. *Transactions of the ASABE* 55, pp. 1491–1508, checked on 1/25/2017. doi: 10.13031/2013.42256
- Baez-Villanueva, O.M., Zambrano-Bigiarini, M., Ribbe, L., Nauditt, A., Giraldo-Osorio, J. D., & Thinh, N. X. (2018). Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. *Atmospheric Research*. doi: 10.1016/j.atmosres.2018.05.011.
- Baez-Villanueva, O. M., Zambrano-Bigiarini, M., Beck, H. E., McNamara, I., Ribbe, L., Nauditt, A., Birkel, C., Verbist, K., Giraldo-Osorio, J.D. & Thinh, N. X. (2020). RF-MEP: A novel Random Forest method for merging gridded precipitation products and ground-based measurements. *Remote Sensing of Environment*, 239, 111606. doi: <https://doi.org/10.1016/j.rse.2019.111606>
- Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & Roo, A. (2017). MSWEP. 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences* 21 (1), pp. 589–615. doi: 10.5194/hess-21-589-2017.
- Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I., McVicar, T. R. & Adler, R. F. (2019). MSWEP V2 global 3-hourly 0.1 precipitation: methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, 100(3), 473-500. doi: <https://doi.org/10.1175/BAMS-D-17-0138.1>
- Blandford, T.R., Humes, K.S., Harshburger, B.J., Moore, B.C., Walden, V.P., & Ye, H., (2008): Seasonal and Synoptic Variations in Near-Surface Air Temperature Lapse Rates in a Mountainous Basin. *J. Appl. Meteor. Climatol.*, 47, 249–261, <https://doi.org/10.1175/2007JAMC1565.1>
- Boisier, J. P., Rondanelli, R., Garreaud, R. D., & Muñoz, F. (2016). Anthropogenic and natural contributions to the Southeast Pacific precipitation decline and recent megadrought in central Chile. *Geophysical Research Letters* 43 (1), pp. 413–421. doi: 10.1002/2015GL067265.
- CCRR (2015). The 2010-2015 mega-drought: A lesson for the future. Center for Climate and Resilience Research (CR)2 (Report to the Nation), checked on 10/2/2017.
- CNR. (2017). Estudio Básico Diagnóstico para Desarrollar Plan de Riego en la Región de la Araucanía: Informe Final [Basic diagnostic study to develop an irrigation plan in the region of Araucanía: Final report]. National Irrigation Commission, performed by Amphos 21.
- Devia, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A review on hydrological models. *Aquatic Procedia*, 4, 1001-1007. doi: 10.1016/j.aqpro.2015.02.126.
- Dile, Y. T., & Srinivasan, R. (2014). Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. *Journal of the American Water Resources Association*, 50 (5), 1226–1241. doi: 10.1111/jawr.12182.
- Dinku, T., Ruiz, F., Connor, S. J., & Ceccato, P. (2010). Validation and Inter-comparison of Satellite Rainfall Estimates over Colombia. *Journal of Applied Meteorology and Climatology* 49 (5), pp. 1004–1014. doi: 10.1175/2009JAMC2260.1.
- Dodson, R. & Marks, D. (1997). Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Climate Research* 8, pp. 1–20, doi: 10.3354/cr008001.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S. et al. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data* 2. doi: 10.1038/sdata.2015.66.
- García, F., Folton, N., & Oudin, L. (2017). Which objective function to calibrate rainfall–runoff models for low-flow index simulations? *Hydrological Sciences Journal*, pp. 1–18. doi: 10.1080/02626667.2017.1308511.
- Garreaud, R. D., Alvarez-Garreton, C., Barichivich, J., Boisier, J. P., Duncan, C., Galleguillos, M., ... & Zambrano-Bigiarini, M. (2017). The 2010–2015 megadrought in central Chile: impacts on regional hydroclimate and vegetation. *Hydrology and Earth System Sciences*, 21(12), 6307. doi:10.5194/hess-21-6307-2017.

- Golmohammadi, G., Prasher, S., Madani, A., & Rudra, R. (2014). Evaluating Three Hydrological Distributed Watershed Models. MIKE-SHE, APEX, SWAT. *Hydrology* 1 (1), pp. 20–39. doi: 10.3390/hydrology1010020.
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria. Implications for improving hydrological modelling. *Journal of Hydrology* 377 (1-2), pp. 80–91. doi: 10.1016/j.jhydrol.2009.08.003.
- Hargreaves, G. H. & Samani, Z. A. (1985). Reference Crop Evapotranspiration from Temperature. *Applied Engineering in Agriculture*. doi: 10.13031/2013.26773
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D. & Simmons, A. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999-2049. doi: <https://doi.org/10.1002/qj.3803>
- Homma, T. & Saltelli, A., (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety* 52, 1–17. doi: 10.1016/0951-8320(96)00002-6
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P. et al. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology* 8, pp. 38–55. doi: 10.1175/JHM560.1
- International Civil Aviation Organization (1993). *Manual of the ICAO Standard Atmosphere*. extended to 80 kilometres. 3rd ed., checked on 8/5/2017.
- IPCC (2013). *Climate Change 2013: the physical science basis*. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535pp
- IPCC (2014). *Climate Change 2014: Synthesis Report*. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R. K. and Meyer, L. A.]. IPCC, Geneva, Switzerland, 151 pp.
- Jansen, M.J.W. (1999). Analysis of variance designs for model output. *Computer Physics Communications* 117, 35–43. doi:10.1016/S0010-4655(98)00154-4.
- Kaleris, V. & Langousis, A. (2016). Comparison of two rainfall–runoff models. Effects of conceptualization on water budget components. *Hydrological Sciences Journal*, pp. 1–20. doi: 10.1080/02626667.2016.1250899.
- Knoche, M., Fischer, C., Pohl, E., Krause, P., & Merz, R. (2014). Combined uncertainty of hydrological model complexity and satellite-based forcing data evaluated in two data-scarce semi-arid catchments in Ethiopia. *Journal of Hydrology* 519, pp. 2049–2066. doi: 10.1016/j.jhydrol.2014.10.003.
- Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5, pp. 89–97. doi: 10.5194/adgeo-5-89-2005.
- McNamara, I., Nauditt, A., Zambrano-Bigiarini, M., Ribbe, L., & Hann, H. (2020). Modelling water resources for planning irrigation development in drought-prone southern Chile. *International Journal of Water Resources Development*, doi: 10.1080/07900627.2020.1768828
- Minder, J. R., Mote, P. W., & Lundquist, J. D. (2010). Surface temperature lapse rates over complex terrain. Lessons from the Cascade Mountains. *Journal of Geophysical Research* 115 (D14), F02011. doi: 10.1029/2009JD013493.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE* 50 (3), pp. 885–900. doi: 10.13031/2013.23153.
- Nachtergaele, F., van Velthuisen, H., Verelst, L., & Wiberg, D. (2012). *Harmonized World Soil Database*. version 1.2. With assistance of Niels Batjes, Koos Dijkshoorn, Vincent van Engelen, Guenther Fischer, Jones Arwyn, Luca Montanarella et al. Edited by FAO & IIASA. FAO, Rome Italy & Laxenburg, Austria, checked on 08/04/17.
- Nash, J. E. & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I -A discussion of principles, *Journal of Hydrology*, 10 (3), pp. 282-290. doi: 10.1016/0022-1694(70)90255-6
- Nauditt, A., Birkel, C., & Ribbe, L. (2016). Conceptual modelling to assess the influence of hydroclimatic variability on runoff processes in data scarce semi-arid Andean catchments. *Hydrological Sciences Journal*, DOI:10.1080/02626667.2016.1240870.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., & Williams, J.R. (2011). *Soil and Water Assessment Tool Theoretical Documentation Version 2009*. Texas (Texas Water Resources Institute Technical Report, 406), checked on 5/2/2017.
- Nyeko, M. (2015). Hydrologic Modelling of Data Scarce Basin with SWAT Model. Capabilities and Limitations. *Water Resources Management* 29 (1), pp. 81–94. doi: 10.1007/s11269-014-0828-3.
- Paniconi, C., & Putti, M. (2015). Physically based modeling in catchment hydrology at 50: Survey and outlook. *Water Resources Research*, 51(9), 7090-7129. doi: 10.1002/2015WR017780

- Parasuraman, K., & Elshorbagy, A. (2008). Toward improving the reliability of hydrologic prediction: Model structure uncertainty and its quantification using ensemble-based genetic programming framework. *Water Resources Research*, 44(12). doi:10.1029/2007WR006451
- Pilgrim, D.H., Chapman, T.G.A., & Doran, D.G., (1988). Problems of rainfall runoff modelling in arid and semiarid regions. *Hydrological Sciences Journal*, 33 (4), 379–400. doi:10.1080/02626668809491261
- Prudhomme, Christel, Giuntoli, Ignazio, Robinson, Emma L., Clark, Douglas B., Arnell, Nigel W., Dankers, Rutger et al. (2014). Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences of the United States of America* 111 (9), pp. 3262–3267. doi: 10.1073/pnas.1222473110.
- Pushpalatha, R., Perrin, C., Le Moine, N., & Andréassian, V. (2012). A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology* 420-421, pp. 171–182. doi: 10.1016/j.jhydrol.2011.11.055.
- Qi, W., Zhang, C., Fu, G., Sweetapple, C., & Zhou, H. (2016). Evaluation of global fine-resolution precipitation products and their uncertainty quantification in ensemble discharge simulations. *Hydrology and Earth System Sciences* 20 (2), pp. 903–920. doi: 10.5194/hess-20-903-2016.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Refsgaard, J. C., Van der Sluijs, J. P., Brown, J., & Van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, 29(11), 1586-1597. doi: 10.1016/j.advwatres.2005.11.013
- Rostamian, R., Jaleh, A., Afyuni, M., Mousavi, S. F., Heidarpour, M., Jalian, A., & Abbaspour, K. C. (2008). Application of a SWAT model for estimating runoff and sediment in two mountainous basins in central Iran. *Hydrological Sciences Journal* 53 (5), pp. 977–988. doi: 10.1623/hysj.53.5.977.
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S., 2010. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications* 181, 259–270. doi: 10.1016/j.cpc.2009.09.018
- Saxton, K. E. & Rawls, W. J. (2006). Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions. *Soil Science Society of America Journal* 70 (5), p. 1569. doi: 10.2136/sssaj2005.0117.
- Schuol, J., Abbaspour, K. C., Srinivasan, R., & Yang, H. (2008). Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model. *Journal of Hydrology* 352 (1-2), pp. 30–49. doi: 10.1016/j.jhydrol.2007.12.025.
- Seibert, J. (1997). Estimation of parameter uncertainty in the HBV model. *Nordic Hydrology* 28, pp. 247–262, checked on 5/9/2017.
- Seibert, J. & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences* 16 (9), pp. 3315–3325. doi: 10.5194/hess-16-3315-2012.
- Seibert, J. & Vis, M. J. P. (2016). How informative are stream level observations in different geographic regions? *Hydrol. Process.* 30, 2498–2508 (2016) doi: 10.1002/hyp.10887
- Dirección General de Aguas (DGA) (1988). Balance Hídrico de Chile. Contribution of the Chilean Committee for the International Hydrological Program (UNESCO) to the Hydric Balance of South America. Dirección General de Aguas, checked on 10/11/17.
- Silva Rojas, L., Molina Magofke, A., & Nuñez Barruel, S. (2012). Plan Regional de Infraestructura y Gestión del Recurso Hídrico al 2021. Región de La Araucanía. Ministerio de Obras Públicas.
- Simons, G., Bastiaanssen, W., Le N., Hain, C., Anderson, M., & Senay, G (2016). Integrating Global Satellite-Derived Data Products as a Pre-Analysis for Hydrological Modelling Studies. A Case Study for the Red River Basin. *Remote Sensing* 8 (4), p. 279. doi: 10.3390/rs8040279.
- Sobol', I.M., 1993. Global sensitivity indices for nonlinear mathematical models. *Mathematical Modelling and Computational Experiment* 1, 407–414. Translated from Russian: I.M. Sobol', *Matematicheskoe Modelirovanie*, 2 (1) (1990), pp. 112-118. doi: 10.1016/S0378-4754(00)00270-6
- Tallaksen, L. M. & van Lanen, H. A. J. (Eds.). *Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater*, Developments in Water Sciences 48, Elsevier B.V., Amsterdam, 2004.
- te Linde, A. H., Aerts, J. C. J. H., Hurkmans, R. T. W. L., & Eberle, M. (2008). Comparing model performance of two rainfall-runoff models in the Rhine basin using different atmospheric forcing data sets. *Hydrology and Earth System Sciences* 12 (3), pp. 943–957. doi: 10.5194/hess-12-943-2008.
- Tuo, Y., Duan, Z., Disse, M, & Chiogna, G. (2016). Evaluation of precipitation input for SWAT modeling in Alpine catchment: A case study in the Adige river basin (Italy). *The Science of the total environment* 573, pp. 66–82. doi: 10.1016/j.scitotenv.2016.08.034.

- Uhlenbrook, S., Seibert, J., Leibundgut, C., & Rodhe, A. (1999). Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrological Sciences Journal* 44, pp. 779–797. doi: 10.1080/02626669909492273
- Van Der Knijff, J. M., Younis, J., & De Roo, A. P. J. (2010). LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, 24(2), 189–212. doi: 10.1080/13658810802549154
- Vansteenkiste, T., Tavakoli, M., Van Steenberghe, N., De Smedt, F., Batelaan, O., Pereira, F., & Willems, P. (2014a). Intercomparison of five lumped and distributed models for catchment runoff and extreme flow simulation. *Journal of Hydrology*, 511, 335–349. doi: 10.1016/j.jhydrol.2014.01.050
- Vansteenkiste, T., Tavakoli, M., Ntegeka, V., De Smedt, F., Batelaan, O., Pereira, F., & Willems, P. (2014b). Intercomparison of hydrological model structures and calibration approaches in climate scenario impact projections. *Journal of Hydrology*, 519, 743–755. doi: 10.1016/j.jhydrol.2014.07.062
- Viale, M. & Nuñez, M. N. (2011). Climatology of Winter Orographic Precipitation over the Subtropical Central Andes and Associated Synoptic and Regional Characteristics. *Journal of Hydrometeorology* 12 (4), pp. 481–507. doi: 10.1175/2010JHM1284.1.
- Walsh, C. L., Blenkinsop, S., Fowler, H. J., Burton, A., Dawson, R. J., Glenis, V., Manning, L. J., Jahanshahi, G., & Kilsby, C. G. (2016). Adaptation of water resource systems to an uncertain future, *Hydrol. Earth Syst. Sci.*, 20, 1869–1884, doi: 10.5194/hess-20-1869-2016.
- Wei, Chih-Chiang (2016). Comparing single- and two-segment statistical models with a conceptual rainfall-runoff model for river streamflow prediction during typhoons. *Environmental Modelling & Software* 85, pp. 112–128. doi: 10.1016/j.envsoft.2016.08.013.
- Worqlul, A. W., Yen, H., Collick, A. S., Tilahun, S. A., Langan, S., Steenhuis, T. S. (2017). Evaluation of CFSR, TMPA 3B42 and ground-based rainfall data as input for hydrological models, in data-scarce regions. The upper Blue Nile Basin, Ethiopia. *CATENA* 152, pp. 242–251. doi: 10.1016/j.catena.2017.01.019.
- Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., & Ribbe, L. (2016). Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrology and Earth System Sciences Discuss.*, pp. 1–43. doi: 10.5194/hess-2016-453.
- Zambrano-Bigiarini, M. & Rojas, R. (2013). A model-independent Particle Swarm Optimisation software for model calibration. *Environmental Modelling & Software* 43, pp. 5–25. doi: 10.1016/j.envsoft.2013.01.004.
- Zhao, Y., Feng, D., Yu, L., Wang, X., Chen, Y., Bai, Y. et al. (2016). Detailed dynamic land cover mapping of Chile. Accuracy improvement by integrating multi-temporal data. *Remote Sensing of Environment* 183, pp. 170–185. doi: 10.1016/j.rse.2016.05.016.