https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Same Streamflow, Different Water Stories: The Hidden Impacts of Streamflow-Only Calibration in Distributed Hydrological Modeling

Nicolás A. Vásquez[1,*], Pablo A. Mendoza[1,2], Wouter Knoben[3], Martyn Clark[3], Tricia Stadnyk[3], and Naoki Mizukami[4]

[1]Department of Civil Engineering, Universidad de Chile, Santiago, Chile
[2]Advanced Mining Technology Center (AMTC), Universidad de Chile, Santiago, Chile
[3]Schulich School of Engineering, University of Calgary, Calgary, Alberta, Canada
[4]NSF National Center for Atmospheric Research (NCAR), Boulder, Colorado, United States
[*]now at Schulich School of Engineering, University of Calgary, Calgary, Alberta, Canada

Correspondence: Pablo A. Mendoza (pamendoz@uchile.cl)

Abstract. Distributed hydrological models enable the characterization of spatial heterogeneities in states and fluxes, including streamflow at inner points of a basin. Despite the growing number of remotely sensed observations, calibrating the model parameters using only streamflow observed at the catchment outlet remains a popular practice. In this paper, we examine how streamflow-only calibration impacts the average seasonality and spatial patterns of simulated evapotranspiration (ET), soil moisture (SM), land surface temperature (LST), and fractional snow-covered area (fSCA). To this end, we conduct calibration experiments with the Variable Infiltration Capacity (VIC) model in six basins located in Chile, using (i) different streamflow-based objective functions, and (ii) regularizing parameters associated with different physical processes. For the latter step, we develop and test a novel spatial regularization strategy based on principal component analysis of physiographic attributes associated with the modeling units contained within each basin. Our results suggest that these decisions may have large effects on the spatial representation of ET, SM_1 (i.e., SM from the first soil layer in VIC), LST, and fSCA, without degrading the performance of streamflow simulations. The average streamflow seasonality can be simulated reasonably well, with large biases in ET, fSCA, SM_1, and LST (in that order). In particular, different calibration configurations can yield the same annual cycle of streamflow through very different ET seasonalities, affecting the catchment-scale seasonal water balance. Additional calibration experiments incorporating ET and SM_1 besides streamflow reaffirm tradeoffs in the fidelity of different simulated variables. Overall, the results presented here reinforce the benefits of including spatial patterns of hydrological variables in the calibration of distributed hydrological models and highlight the need to verify the seasonality of other simulated variables besides streamflow.

1 Introduction

Spatially distributed hydrological models are essential tools for simulating the heterogeneity of variables involved in the terrestrial water cycle (Reed et al., 2004), enabling streamflow predictions at ungauged inner river sections (Ul Hassan et al., 2024). Such information can be used for a myriad of applications that include, for example, the allocation of surface water

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

rights, drought monitoring, short to long-term streamflow forecasting, or water resources planning under changing climatic conditions. Despite their widespread use, the implementation of distributed hydrological models is challenged by two well-recognized problems: (1) equifinality arising from multiple combinations of internal states and fluxes across the modeling domain, which may yield very similar streamflow results at the catchment outlet for very different reasons (e.g., Beven, 1993; Khatami et al., 2019; Soltani et al., 2021a; Demirel et al., 2024); and (2) the large number of parameter values – resulting from the increased number of modeling units – that need to be estimated, thereby exacerbating overparameterization and identifiability issues (e.g., Pokhrel et al., 2008; de Lavenne et al., 2019).

The increasing number of remotely sensed observational datasets (Lettenmaier et al., 2015; McCabe et al., 2017; Sheffield et al., 2018; Wagner et al., 2025) has brought new opportunities to tackle the equifinality problem in distributed hydrological modeling by incorporating variables beyond streamflow (Q) in the calibration and evaluation steps. Such variables include evapotranspiration (ET; e.g., Demirel et al., 2018b; Dembélé et al., 2020b), soil moisture (SM; e.g., Tong et al., 2021; Bajracharya et al., 2023), land surface temperature (LST; e.g., Zink et al., 2018), water storage variations (WS; e.g., Werth and Güntner, 2010; Mostafaie et al., 2018; Trautmann et al., 2023), and fractional snow-covered area (fSCA; e.g., Parajka and Blöschl, 2008; Duethmann et al., 2014; Bennett et al., 2019; Tong et al., 2021; Tang et al., 2023). In particular, this information can be used to formulate summary performance metrics to assess the ability to replicate observed spatial patterns (e.g., Koch et al., 2017; Demirel et al., 2018b; Koch et al., 2018; Tong et al., 2021; Demirel et al., 2024), which can be combined with streamflow-based performance metrics into a single objective function (e.g., Demirel et al., 2018b; Dembélé et al., 2020a), or used within a multi-objective search framework (e.g., Demirel et al., 2024) to constrain the parameter space and achieve more realistic simulations.

Despite the proliferation of satellite remote sensing products and new multivariate parameter estimation strategies (e.g., Rakovec et al., 2016; López López et al., 2017; Stisen et al., 2018; Dembélé et al., 2020b; Soltani et al., 2021a; Meyer Oliveira et al., 2021; Pool et al., 2024; Tiwari et al., 2024; Ul Hassan et al., 2024), distributed hydrological models are usually calibrated using streamflow as the only target hydrological variable (e.g., Shafii and Tolson, 2015; Melsen et al., 2016; Mizukami et al., 2017; Beck et al., 2020; Aguayo et al., 2021; Wang et al., 2022; Kurugama et al., 2026). This emphasis stems not only from the need to support water resource management applications but also to achieve a basic understanding of catchment-scale water balances (Kirchner, 2006; Troch et al., 2013; Mendoza et al., 2016; Akbar et al., 2020). Hence, several streamflow-based multi-objective calibration approaches have been proposed and tested with distributed hydrological models for simultaneously matching different parts of the hydrograph (e.g., Pokhrel and Gupta, 2010; Westerberg et al., 2011; Garcia et al., 2017; McInerney et al., 2017; Fowler et al., 2018a, b; Koppa et al., 2019; Todorović et al., 2022; Rakovec et al., 2019; Casper et al., 2023).

On the other hand, the high dimensionality of the parameter optimization problem in distributed hydrological modeling is typically addressed through regularization (Pokhrel and Gupta, 2010; Samaniego et al., 2010; de Lavenne et al., 2019), a family of mathematical techniques that introduce one or more constraints to the parameter search process. Parameter regularization usually involves the use of super-parameters or transfer functions applied under the assumption that the spatial variability of climate and/or geomorphological attributes informs the spatial structure of model parameters (e.g., Mizukami et al., 2019;

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Beck et al., 2020). However, establishing these relationships – especially for "free" (i.e., nonphysical) parameters – is often challenging. Moreover, equifinality (Beven and Binley, 1992; Beven, 2006; Khatami et al., 2019) hinders the possibility of achieving credible simulations, although this issue can be partially alleviated through multivariate model evaluation (e.g., Demirel et al., 2018b; Dembélé et al., 2020a; Shah et al., 2021; Alfieri et al., 2022; Yáñez-Morroni et al., 2023) and parameter uncertainty estimates (Beven, 2001; Koppa et al., 2019).

Crucially, the way parameter regularization techniques are implemented plays a key role in shaping simulated spatial patterns and overall model performance (Refsgaard and Knudsen, 1996; Reed et al., 2004; Samaniego et al., 2010; Rakovec et al., 2016; Dembélé et al., 2020a; Demirel et al., 2018b), whereas the choice of calibration metric has direct implications for parameter estimates and simulated hydrographs (Kollat et al., 2012; Garcia et al., 2017; Pool et al., 2018; Fowler et al., 2018b). While the effects of these methodological decisions have typically been evaluated using summary metrics (Clark et al., 2021), their influence on simulated annual cycles of key water balance components (i.e., evapotranspiration and runoff) and simulated spatial patterns has received very little attention. Instead, previous studies have mostly focused on the effects of adding remotely sensed variables to calibration schemes on the overall performance (i.e., Q-based metrics and simulated patterns) of distributed hydrological models. Therefore, we seek to answer the following research questions:

1. To what extent do the spatial regularization of different soil parameters and the choice of streamflow-based objective functions affect model performance in terms of Q simulations and the spatial patterns of LST, ET, fSCA, and SM
2. What are the tradeoffs between accurately replicating Q annual cycles and effectively simulating the seasonal patterns of LST, ET, fSCA, and SM? How can these trade-offs be overcome?

Here, we hypothesize that distributing soil parameters through a regularization strategy that is spatially coherent with grid cell attributes can improve the model's capability to replicate spatial patterns and the seasonality of other variables aside from streamflow. To test this idea, we calibrate the Variable Infiltration Capacity model (VIC; Liang et al., 1994) in six basins with different hydrological regimes in central Chile. Specifically, we use (i) Principal Component Analysis (PCA) to derive an *a priori* spatial distribution and define a regularization strategy for a given soil parameter, and (ii) different streamflow-based objective functions to obtain parameter fields by calibrating (super) parameters. To quantify the ability to replicate observed spatial patterns in ET, LST, SM_1, and fSCA, we used biased and unbiased performance metrics. Importantly, we examine the implications of only-Q calibration approaches on seasonal water balances by looking at simulated annual cycles, focusing on Q and ET.

2 Study domain

We consider six basins located in Continental Chile that span three different hydrological regimes: snowmelt-driven, mixed, and rainfall-driven (Figure 1). The basins, from north to south, are Cochiguaz River at El Peñón (675 km^2, Figure 1a), Choapa River at Cuncumén (1132 km^2, Figure 1b), Claro River at El Valle (349 km^2, Figure 1c), Colorado River at the confluence with Palos River (878 km^2, Figure 1d), Cautín River at Rari-Ruca (1306 km^2, Figure 1e) and Futa River at Tres Chiflones (517

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

km^2, Figure 1f). These basins have a low degree of human intervention and were selected based on the following criteria: (a) a near-natural flow regime defined as a maximum threshold value of 5% for the relationship between the annual volume of water assigned for permanent consumptive use and the mean annual flow, (b) absence of large reservoirs within each catchment, and (c) a small (<2%) glacierized area (Alvarez-Garreton et al., 2018).

Five out of the six basins are located on the western slopes of the Andes Cordillera, with the Futa River basin being the exception. Figure 1 shows the annual climatological cycles for precipitation (P), temperature (T), and streamflow (Q). The Cochiguaz and Choapa River basins have a snowmelt-driven regime, receiving precipitation mainly during Fall (MAM) and Winter (JJA), when water is primarily stored as snow and released during the austral spring (SON) and summer (DJF) seasons. The Claro and Colorado River basins have a mixed regime, with higher streamflow values during the winter and spring/summer seasons, respectively. The two southernmost basins, Cautín and Futa, have maximum streamflow values during winter (rainfall-driven), although a slight influence of snowmelt is detected in the Cautín River during the spring season.

3 Datasets

3.1 Meteorological and streamflow data

Daily precipitation (P) and maximum (T_{max}) and minimum (T_{min}) air temperatures are obtained from the CR2MET v2.0 dataset (Boisier et al., 2018; DGA, 2022; Boisier, 2023), which covers continental Chile with a horizontal resolution of 0.05° x 0.05° for the period 1979-2020. CR2MET precipitation estimates are obtained using a two-step approach consisting of (i) computing the probability of precipitation at each grid cell through logistic regression models, and (ii) calculating daily precipitation amounts using multiple linear regression models. All models use ERA5 reanalysis output (Hersbach et al., 2020) and geomorphological attributes as predictors, and daily precipitation from meteorological stations as predictands (note that, in the case of precipitation occurrence, the vector with predictands only contains zeros and ones). For T_{max} and T_{min}, land surface temperature from MODIS is also included as a predictor.

To obtain sub-daily meteorological time series, CR2MET daily precipitation and temperature are disaggregated into hourly time steps using the sub-daily distribution provided by ERA5-Land (Muñoz Sabater, 2019), which is bias corrected to match the CR2MET daily values. Relative humidity (RH), wind speed (W), atmospheric pressure (AP), and incoming shortwave radiation (K_{in}) are derived for the same horizontal resolution grid by spatially interpolating ERA5-Land outputs. ERA5-Land wind speed is bias-corrected using wind simulations from WRF at a 1-km horizontal resolution (Geophysics Department and Ministerio de Energía, 2018). Incoming longwave radiation (L_{in}) was computed using the parameterization proposed by Iziomon et al. (2003) using the bias-corrected hourly temperature (Tas). To conduct hydrological model simulations, all the meteorological variables are grouped to obtain 3-hourly time steps.

Finally, daily streamflow data are obtained from stations maintained by the Chilean Water Directorate (DGA, in Spanish), which are also available in the CAMELS-CL dataset (Alvarez-Garreton et al., 2018).

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

3.2 Remote sensing products

We use remotely sensed fractional snow-covered area, actual evapotranspiration, land surface temperature, and soil moisture for hydrological model evaluation (Figure 2). The fractional snow-covered area is derived from MODIS products (MOD10/MYD10; Hall and Riggs, 2016). To obtain a unique time series, daily MOD10 and MYD10 estimates are averaged at each MODIS grid cell. If only one product is available for a specific day and grid cell, we use that value to estimate fSCA. All gaps (i.e., days and grid cells unavailable) are filled at the original horizontal resolution (500 m) using the methodology proposed by Cornwell et al. (2016) (see details in section S.1, Supporting Information). Actual evapotranspiration is obtained from the MOD16 product (Mu et al., 2011) using 8-day estimates with a 1-km horizontal resolution. Land surface temperature for each day and for each grid cell is obtained as the average between the MOD11 and MYD11 products (Wan, 2014) at a 1-km horizontal resolution. If one of the two LST estimates is unavailable, the day is considered missing. For the assessment of soil moisture simulations, we use the ESA-CCI product (Dorigo et al., 2017), available at a 0.25° horizontal resolution.

To resolve the mismatch between the horizontal grid of remote sensing products and the spatial discretization of the hydrological model (0.05° x 0.05°, same as in the CR2MET product), ET, LST, and fSCA are upscaled using spatial averages, whereas SM is downscaled to the closest CR2MET grid cell (e.g., dos Santos Araujo et al., 2024). Empty values were not filled for ET, LST, and SM, as opposed to fSCA.

3.3 Ancillary data

Because this study considers a distributed hydrological model that requires *a priori* parameter fields, we derive these using grid cell attributes obtained from different sources. We use the SoilGrids dataset (Poggio et al., 2021) to derive clay and sand content, and mean bulk density for the first 2 m soil depth (using vertical weighted average), all available at a 250-m horizontal resolution. The elevation is estimated from the Shuttle Radar Topography Mission (SRTM; Farr et al., 2007), while the aspect is derived from the SRTM elevation raster at the original horizontal resolution (30 m). All datasets are upscaled to match the hydrological model grid using spatial averages.

4 Approach

Figure 2 illustrates the main steps of our experimental setup. First, we obtain and process the hydrometeorological datasets to force and evaluate hydrological model simulations (step 1, sections 3 and 4.1); secondly, we implement a spatial parameter regularization strategy (step 2, section 4.2). Finally, we conduct a suite of model calibration experiments to assess how the choice of calibration metric and the choice of regularized parameter affect (1) streamflow performance metrics, (2) the simulated spatial patterns of ET, SM_1, LST, and fSCA, and (3) the simulated annual cycles of these variables (step 3, section 4.3).

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

4.1 Hydrological Modeling

We use the VIC model (Liang et al., 1994) to simulate state variables and fluxes at a 0.05°x 0.05° horizontal resolution. VIC is a semi-distributed, physically based hydrological model that solves energy and mass balance equations. Precipitation can be partitioned into snowfall and rainfall, which serve as inputs for canopy storage. The maximum amount of water intercepted by the canopy is estimated using the Leaf Area Index (LAI; Dickinson, 1984). The snowpack is represented by two layers, where the top layer is used for energy balance computations (Andreadis and Lettenmaier, 2006). The soil column is vertically discretized into three layers, with the top and bottom layers controlling infiltration and baseflow generation, respectively. To calculate infiltration, VIC uses the Xinanjiang formulation (Zhao et al., 1980), assuming that infiltration capacity varies within an area (Wood et al., 1992). Excess runoff is generated in those areas where precipitation exceeds the available moisture storage of the first soil layer. VIC assumes that gravity drives drainage, using the formulation proposed by Brooks and Corey (1964). In this regard, water enters the cell only from the atmosphere; thus, VIC does not consider lateral fluxes among grid cells. Baseflow is computed using a formulation proposed by Franchini and Pacciani (1991). For more details on the VIC model, readers are referred to Liang et al. (1994).

Sub-grid horizontal heterogeneity in VIC can be represented through different land cover types. We use the International Geosphere-Biosphere Program (IGBP) classification for the year 2010 from the MCD12Q1 v006 land cover product (Sulla-Menashe and Friedl, 2018) to represent all land cover types spanning at least 2% of each grid cell area. Mean monthly LAI values for these land cover types are derived from the MOD15A2 product. Soil bulk density is estimated as the vertical average of values retrieved from the SoilGrids product (Poggio et al., 2021) across the top 2 m soil depth. We consider vertical heterogeneity by distributing precipitation and air temperature along 200-m elevation bands, following recommendations from previous studies (Murillo et al., 2022; Givovich et al., 2025).

We use the vector-based river routing model mizuRoute (Mizukami et al., 2016) to convert simulated runoff into streamflow. mizuRoute first performs hillslope routing using a gamma-distribution-based unit-hydrograph to delay VIC runoff, and then routes the delayed runoff for each river reach defined by the river network topology. For river routing, we used the diffusive wave scheme, following the setup recommended by Cortés-Salazar et al. (2023). Manning's roughness coefficients and channel widths are spatially distributed within each basin through regression equations that use river reach attributes as predictors (Niño, 2002; Mendoza et al., 2012). Note that none of routing parameters are included in subsequent calibration experiments (described in Section 4.3).

4.2 Parameter regularization

To assess the effects of distributing soil parameters in space, we implement a spatial parameter regularization technique based on a PCA over a suite of grid cell attributes within each basin, including bulk density, clay and sand content, elevation, and slope (e.g., Samaniego et al., 2010; Mizukami et al., 2017; Beck et al., 2020). The underlying motivation is to extract the main patterns explaining the spatial distribution of physiographic attributes and use the dominant signal to obtain VIC parameter fields. To this end, we sort the grid cell attributes within a matrix (step 2 in Figure 2), with grid cells as rows and attributes as columns,

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

and perform PCA to obtain a matrix with the same dimensions, where the first principal component (PC1; or first column of the new matrix) explains most of the variance among the attributes (See Figure S.1). Then, we use the PC1 map as a predictor to obtain *a priori* spatially distributed VIC parameters (θ_{VIC}), using three super parameters (Pokhrel et al., 2008) highlighted in red in Figure 2: (i) an additive term (θ_γ), (ii) a factor (θ_α), and (iii) an exponent controlling the linearity of the relationship (θ_β).

4.3 Model calibration experiments

We calibrated the VIC parameters identified as the most sensitive by Sepúlveda et al. (2022), including (i) b_{infilt}, which controls the infiltration process; (ii) Ds and Ws, which control the linearity of the curve used to calculate the water leaving the third (deepest) soil layer; (iii) Ds_{max}, which is the maximum baseflow rate; (iv) the saturated hydraulic conductivity (K_{sat}), which explains the water exchange between the soil layers; (v) the soil layers' depths; and (vi) maximum snow albedo and (vii) its temporal decay rate. The benchmark calibration considers spatially constant VIC parameters (e.g., Melsen et al., 2016), whereas the calibration experiments consider the spatial regularization of individual soil parameters (b_{infilt}, Ds, Ds_{max}, K_{sat} and depth of soil layers), while maintaining the rest spatially constant. With these experiments, we seek to understand the benefits of spatially distributing a parameter associated with a specific process. An additional calibration experiment considers the vertical distribution of hydraulic conductivity, for which we apply the same equation of Figure 2 (step 2; Pokhrel et al., 2008) though, instead of using the PC1, we use the vertical distribution of the bulk density obtained from the SOILGRIDS database – which contains estimation of the bulk density at different depths – to define an *a priori* vertical distribution. Hence, and for this case, K_{sat} is horizontally constant among grid cells, while adopting different values in each soil layer.

Therefore, we conduct nine (9) Q-only (number of parameter configurations) $\times$ 5 (number of objective functions) = 45 model calibrations for each basin, and – to assess the effects of adding one additional target variable – two model calibrations that consider spatially constant parameters and incorporate ET (Q+ET) and soil moisture in the first layer (Q+SM_1) in the optimization process.

4.3.1 Streamflow-based objective functions

The calibration experiments aim to maximize five streamflow-based objective functions by running the Dynamically Dimensioned Search (DDS; Tolson and Shoemaker, 2007) algorithm, implemented within the OSTRICH software (Matott, 2017) with a maximum number of 2,000 iterations. All calibration metrics consider daily time series of simulated and observed streamflow. First, we use the Kling-Gupta efficiency (KGE; Gupta et al., 2009; Kling et al., 2012), which seeks to minimize the Euclidean distance between performance metrics related to volume (β), variability (γ), and timing (r), and their optimal values (which are equal to 1):

$$OF_1 = KGE = 1 - ED = 1 - \sqrt{(1-\beta)^2 + (1-\gamma)^2 + (1-r)^2} \quad (1)$$

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Where $\beta = \mu_s/\mu_o$ is the ratio between simulated (s) and observed (o) average values (μ); $\gamma = (\sigma_s/\mu_s)/(\sigma_o/\mu_o)$, where σ represents the standard deviation; and r is the Pearson correlation coefficient between observed and simulated daily values. The second objective function is the Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970):

$$OF_2 = NSE = 1 - \frac{\sum_{t=1}^{T}(Q_s^t - Q_o^t)^2}{\sum_{t=1}^{T}(Q_o^t - \overline{Q}_o)^2} \tag{2}$$

Where Q_s and Q_o are the simulated and observed daily streamflow, respectively; t indicates the time step, and $\overline{Q}$ is the climatological mean daily streamflow. Since both KGE and NSE are influenced by high flows (Clark et al., 2021), we also used the composite metric (Eq. 3) proposed by Garcia et al. (2017), which incorporates the KGE computed with $1/Q$ to give more weight to low flows:

$$OF_3 = \frac{KGE(Q) + KGE(1/Q)}{2} \tag{3}$$

The fourth calibration metric is the Nash-Sutcliffe efficiency computed from simulated and observed daily flow duration curves (FDCs) (e.g., Nijzink et al., 2016). The last OF (Eq. 4) is based on an Euclidean distance approach (e.g., Schoups et al., 2005) that combines the daily KGE with biases in five hydrological signatures (HS_i Yilmaz et al., 2008) that focus on different parts of the flow duration curve (e.g., Westerberg et al., 2011; Shafii and Tolson, 2015):

$$OF_5 = 1 - \sqrt{[1 - KGE(Q)]^2 + \sum_{i=1}^{N=5}(HS_i^{opt} - HSi)^2} \tag{4}$$

Where HS_i^{opt} is the optimal value for the i^{th} hydrological signature. The signatures considered in Eq. 4 evaluate biases for runoff ratio (Eq. A1), high flow volumes (Eq. A2), low flow volumes (Eq. A3), the mid-segment slope of the flow duration curve (Eq. A4) and the median of daily flows (Eq. A5). The formulation of each HS can be found in the Appendix.

4.3.2 Spatial pattern performance metrics

We use the spatial pattern efficiency metric (Eq. 5) proposed by Dembélé et al. (2020a) to assess the quality of ET, SM, LST, and fSCA simulations:

$$E_{SP} = 1 - \sqrt{(1 - r_s)^2 + (1 - \gamma)^2 + (1 - \alpha)^2} \tag{5}$$

Where r_s is the Spearman rank correlation coefficient between the simulated and reference values, γ is the ratio of the coefficients of variation (as in KGE; Eq. 1), $\alpha = 1 - RMSE(Z_{X_s}, Z_{X_o})$, with Z being the time series with standardized values for the variable X (note that we follow the same notation as in Dembélé et al. (2020b)) and RMSE is the Root Mean

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Square Error. The standardization aims to avoid a direct contrast between the model and reference values, since remotely sensed products have biases that could affect the metric. However, we use bias-accounting metrics, such as KGE and RMSE, to assess the model's ability to replicate raw remotely sensed estimates.

Following Dembélé et al. (2020a), we use two approaches to contrast spatially distributed simulations against observations (options 1 and 2 in step 3, Figure 2). Option 1 compares simulated and reference maps of a variable X at each time step, resulting in a time series of performance measures ($OF_X(t)$) that can be used to compute a summary metric OF_X^{time} by temporally averaging all values. Alternatively, option 2 compares simulated and reference time series in each grid cell to obtain a map of performance metrics ($OF_X(i)$) that can be used to calculate a summary performance measure OF_X^{space} by spatially averaging all values.

For soil moisture, we use the correlation coefficient (Eq. 6) between simulated (θ_s) and reference(θ_{ref}) values, following Tong et al. (2021).

$$O_{SM} = \frac{\sum_{t=1}^{T}\left[(\theta_s^t - \overline{\theta}_s)(\theta_{ref}^t - \overline{\theta}_{ref})\right]}{\sqrt{\sum_{t=1}^{T}\left[(\theta_s^t - \overline{\theta}_s)^2(\theta_{ref}^t - \overline{\theta}_{ref})^2\right]}} \tag{6}$$

Where t represents the time step, and overlines average values. We only consider the soil moisture simulated in the first soil layer since the ESA-CCI soil moisture product accounts for the first 0.5–2 cm. To overcome mismatches between reference and SM_1 due to different soil depths and properties, we applied mean–standard deviation matching (Draper et al., 2009; López López et al., 2017):

$$\theta_s^* = \frac{\sigma_{\theta_{ref}}}{\sigma_{\theta_{sim}}} \cdot \left(\theta_s - \overline{\theta}_s\right) + \overline{\theta}_{ref} \tag{7}$$

Where $\sigma_{\theta_{ref}}$ and σ_{θ_s} are the standard deviations of the reference and simulated soil moisture, respectively.

We assess the quality of fSCA simulations using the RMSE and the metric O_{SC} proposed by Tong et al. (2021):

$$O_{SC} = 1 - (S_O + S_U) \tag{8}$$

where S_O (S_U) represents the fraction of simulated days with overestimated (underestimated) fSCA. To avoid noisy values, a threshold of 1% is applied; hence, simulated and reference grid cell fSCA values below 0.01 are considered to have no snow cover.

To assess the model's capability to replicate historically observed annual cycles of ET and SM, we normalize ET and SM annual values as $X' = (X - X_{min})/(X_{max} - X_{min})$. We use the notation SM_1, SM_2, and SM_3 to refer to the moisture content in soil layers 1, 2, and 3 of the VIC model, respectively. Finally, we focus on metrics computed for Q, ET, SM, LST, and fSCA for a calibration period spanning 01/April/2005-31/March/2018, using five years (01/April/2000-31/March/2005) for warm-up.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

4.3.3 Q+ET and Q+SM$_1$ additional calibrations

To test whether the explicit incorporation of ET or SM$_1$ in the calibration objective function improves model fidelity, we conduct two additional calibration experiments using spatially constant parameters to minimize an Euclidean-based objective function:

$$OF = \sqrt{[1 - KGE(Q)]^2 + [1 - E_{SP}(X)]^2} \tag{9}$$

Where X refers to ET or SM$_1$. For the Q+ET and Q+SM$_1$ experiments, we use the DDS optimization algorithm with a total budget of 2,000 simulations, considering the same warm-up and calibration periods as the Q-only calibrations.

5 Results

5.1 Effects on streamflow performance metrics

First, we examine the impact of different calibration objective functions and regularized parameters on the performance of streamflow simulations across basins (Figure 3). In general, the values of the objective function are similar for a given combination of calibration metric and basin, regardless of the parameter that is regularized, suggesting a high degree of compensation among parameters. In the Choapa and Cautín River basins, lower OF values are achieved for K_{sat} parameter (and similarly for $Depth_2$ in the Futa River basin) when the OF is $0.5 \cdot KGE(Q) + 0.5 \cdot KGE(1/Q)$. Interestingly, the components of the objective functions can vary considerably, especially $KGE(1/Q)$, which gives more weight to low flows. In particular, the values of this metric are generally negative or close to zero in the Claro (mixed regime) and Futa (rainfall-driven) River basins. On the other hand, the $KGE(1/Q)$ values are stable for all configurations in the snowmelt-driven basins (Cochiguaz and Choapa), regardless of the inclusion of $KGE(1/Q)$ in the calibration process.

5.2 Effects on simulated spatial patterns

Figure 4 illustrates the impacts of calibration OF and the choice of the regularized parameter on simulated ET at the Colorado River basin with respect to the benchmark calibration. Here, we examine parameters controlling infiltration (b_{infilt}), soil depth ($Depth_2$), vertical moisture distribution (Vertical K_{sat}), and baseflow (Ds_{max}) (the remaining variables, regularized parameters, and basins are included in the Supporting Information, Figures S.2-S.25). The results show that some combinations of OF and the parameter regularization strategy may enhance the performance of ET simulations compared to the benchmark calibration. In some grid cells, absolute changes in ET performance can be larger than 1 E_{SP} units, with positive (e.g., for $KGE(Q)$ and Ds_{max}) or negative (e.g., $KGE(Q)$ & $KGE(1/Q)$ and b_{infilt}) effects.

Figure 4 also shows that, for a given OF, the degree of improvement can vary depending on the regularized parameter. For example, when using $KGE(Q)$ and $KGE(Q)$ & $NSE(FDC(Q))$ as OFs, regularizing K_{sat} and Ds_{max} (b_{infilt} and $Depth_2$) yields improved (declined) performance in ET simulations. For $NSE(Q)$, regularizing any parameter yields a slight increase

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

in E_{SP}(ET), whereas the opposite occurs with $KGE(Q)$ & $KGE(1/Q)$. For the OF combining $KGE(Q)$ and hydrological signatures, mixed results are obtained, with $\Delta E_{SP} > 0$ (blue) for lower grid cells, and $\Delta E_{SP} < 0$ (red) for high altitude grid cells. Similar ET performance results are obtained for the remaining basins, although an overall improvement in ET performance is achieved when regularizing parameters for the Claro and Futa River basins (Figures S.3 and S.7 in the Supporting Information). Notably, parameter fields that yield different performance in ET provide similar streamflow performance metrics (Fig. 3).

5.3 Simulated annual cycles

Now, we examine how different combinations of calibration metric and parameter regularization strategy affect simulated annual cycles of streamflow, normalized catchment-scale ET and SM_1, land surface temperature, and fractional snow-covered area. Figure 5 shows results for the parameter sets that maximize $OF = f(KGE(Q), KGE(1/Q))$ at the Choapa (snowmelt-driven), Colorado (mixed regime), and Cautín (rainfall-driven) River basins (results for the remaining basins are included in Figure S.32). Although streamflow is underestimated during some spring and summer months in the Choapa River basin, all calibration experiments yield a reasonable representation of the Q seasonality in the three catchments, with little difference between configurations. Similar results are obtained for LST, with an overall underestimation at the Choapa River basin. However, ET seasonalities can change drastically depending on the model configuration, as opposed to Q, SM_1, LST, and fSCA. In the Choapa River basin, regularizing different parameters yields different biases, with the largest underestimations for the spatially constant and $Depth_2$ cases during Spring. In the Colorado River basin, none of the model configurations provide a reasonable representation of ET seasonality, whereas mixed results are obtained in the Cautín River basin, with both good (e.g., b_{infilt}, K_{sat}, and Ds_{max}) and poor (e.g., spatially constant and $Depth_2$) representations. In summary, regularizing different parameters may shift simulated annual cycles and, consequently, reduce biases, without impacting the Q performance.

Figure 6 illustrates how the choice of streamflow-based OF affects simulated annual cycles of Q, ET, SM_1, LST, and fSCA in the Cautín River basin. In general, different OFs yield similar streamflow responses, except (i) when the model parameters are spatially constant and calibrated using $KGE(Q)$ (black line), which yields an overestimation in Q, and (ii) when $Depth_2$ is regularized and the OF is $KGE(Q)$ and $NSE(FDC(Q))$, which yields an underestimation of Q.

The results for LST show little disagreement among OFs and model configurations, with a reasonable representation of seasonality. For SM_1, the results reveal slight differences among configurations, with greater biases with respect to the reference (white dots) when K_{sat} is regularized. The seasonality of fSCA is well simulated with all combinations of OF and regularization strategy, though the latter decision introduces large discrepancies in biases. For example, when b_{infilt} is regularized and the calibration metric is $KGE(Q)$ ($KGE(Q)$ & $NSE(FDC(Q))$), fSCA is underestimated (overestimated).

In particular, the effects of the regularization strategy on simulated ET annual cycles can vary considerably depending on the streamflow performance metric used for calibration. For example, when $KGE(Q)$ or $KGE(Q)$ & $KGE(1/Q)$ are used as OFs, shifted ET seasonalities are obtained if b_{infilt} and $Depth_2$ are regularized, or if model parameters are spatially constant. Conversely, when $KGE(Q)$ & $NSE(FDC(Q))$ is maximized in the calibration process, all ET seasonalities collapse into

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

the reference values, regardless of the parameter that is regularized. A similar behavior is obtained for the rest of the basins (Figures S26-S32).

5.4 Relation between Q and other variables' performance metrics

Does the improvement in streamflow-based OFs relate to improvements in other Q metrics and simulated variables? To seek answers, we compute the Spearman's rank correlation coefficient between the OF values obtained in N = 2,000 iterations of the optimization process, and other performance metrics for Q, ET, SM_1, LST, and fSCA. Figure 7 displays the results for calibrations conducted at the Choapa, Colorado, and Cautín River basins using $OF = KGE(Q)$ & $KGE(1/Q)$ as OF. As expected, high correlations are obtained between streamflow-based metrics and OFs. However, negative correlations are obtained for ET, especially at the Choapa and Colorado River basins. For the case of SM_1, some parameter configurations yield positive correlations with OF values. For LST and fSCA, we obtained poor correlations with OF, which can be explained by the relatively good performance obtained for these variables, regardless of the value of OF.

5.5 Effects of adding ET or SM_1 on simulated annual cycles

Figure 8 displays the average annual cycles obtained after calibrating the model parameters with objective functions that consider (i) only streamflow, (ii) Q and ET, and (iii) Q and SM_1. The results show that streamflow performance decreases considerably in all basins when ET is added to the OF (first row in Fig. 8), whereas the annual cycle of ET improves drastically (second row in Fig. 8) for all the basins (see also Fig. S.34). The effect of calibrating Q and ET only affects SM_1 for the Cautín and Claro River Basins during the months September to January. Alternatively, calibrating Q + SM_1 improves the seasonality of ET at Choapa and, to a lesser degree, in the Cautín and Futa River basins, although important biases remain in Cautín and Futa (for example, the month of maximum ET remains shifted two months with respect to the reference ET). ET in the Colorado River Basin is not greatly modified if SM_1 is included in the objective function.

6 Discussion

6.1 Effects on spatial patterns

The effects of streamflow-based OF selection on parameter values and, therefore, the simulation of different parts of the hydrograph and other hydrological variables besides Q have been widely discussed in the literature (e.g., Gupta et al., 1998; Merz et al., 2011; Garcia et al., 2017; Westerberg et al., 2011; Fowler et al., 2018b; Khatami et al., 2019). However, the impacts of different parameter regularization techniques or the choice of regularized parameters on simulated spatial patterns are less documented. Our results show that regularizing different parameters may impact simulated spatial patterns differently, regardless of the calibration metric used, which aligns well with previous work (Samaniego et al., 2010; Demirel et al., 2018a). Moreover, we find that the simulation of spatial patterns can be improved without losing streamflow performance, in agreement with Demirel et al. (2024). Nevertheless, we did not find a unique combination of OF and parameter regularization strategy

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

that improves simulated spatial patterns for all variables and basins compared to the benchmark calibration. In this regard, it would be useful to characterize how regularizing different parameters affects the capacity of models to replicate observed spatial patterns (Demirel et al., 2018a; Saavedra et al., 2022).

6.2 Biases in simulated annual cycles

In addition to the evident effects of different parameter configurations on simulated spatial patterns of ET, SM_1, LST, and fSCA, large biases in catchment-scale simulated annual cycles may be obtained for these variables, even if realistic streamflow seasonalities are achieved. We obtain smaller biases in the average seasonality of LST and SM_1 compared to fSCA and ET. For fSCA, similar annual cycles are obtained with all the regularization strategies tested here, although with discrepancies among them during winter months with respect to the reference. For ET, and despite a reasonable representation of streamflow seasonality and improvements in spatial pattern efficiency metrics, different parameter configurations can shift ET simulated annual cycles, directly affecting seasonal water balances. Hence, our results suggest that calibrations based on streamflow data only do not necessarily yield reasonable ET simulations and, without verification, important biases in ET can remain. In this regard, our results differ from Rakovec et al. (2016), who calibrated the mHM model for Europe using only streamflow data, finding that ET was reasonably simulated.

To explore possible reasons that could explain differences in ET seasonalities, we examine simulated annual cycles in catchment-scale normalized soil moisture for the second ($SM_2^{'}$) and third ($SM_3^{'}$) soil layers, along with their combinations $SM_{2+3}^{'}$ and $SM_{1+2+3}^{'}$ (i.e., total water content in the soil column, Figure 9). In VIC, the roots of different land cover types can reach different soil layers (specified by the user), and ET is computed based on the water availability in the layers where roots are present. Thus, analyzing different average simulated soil moisture seasonalities is key to explaining when water is available for ET.

Similarly to ET, SM_1 seasonality can also be shifted with respect to the reference depending on the regularized parameter (first row in Figure 9). Interestingly, little differences in SM seasonalities arise from the model configurations in the Colorado River basin, producing the same annual cycles for ET (Figure 5). For Choapa and Cautín, simulated annual ET cycles when distributing b_{infilt} and Ds_{max} depart from the remaining configurations (Figure 5), which can be explained by similar differences in simulated SM_2 and SM_3 (Figure 9). These basins (Choapa and Cautín) are covered by forests (particularly the Cautín River Basin), and their roots have access to the three soil layers. In Choapa (a snow-driven arid basin), roots reach only the second soil layer for some land cover types (Fig. S.36). Further, discrepancies in the simulated annual cycles of ET with spatially constant parameters and $Depth_2$ at the Cautín basin are also explained by the simulated behavior in SM_2 and SM_3. Thus, similar Q but different ET seasonalities can be achieved by compensations among simulated fluxes and state variables that produce a similar streamflow response, but with discrepancies in ET, leading to an extreme case of flux equifinality (Khatami et al., 2019), stressing the need to evaluate simulated annual cycles of the target variables affecting the water mass balance.

The results presented here reveal trade-offs between the aim to achieve accurate streamflow simulations through parameter calibration and the aim to replicate other hydrological variables. Specifically, we obtained that different parameter configura-

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

tions and objective functions may yield similar streamflow performance, but for very different reasons in terms of simulated spatial patterns of other hydrological variables (Kirchner, 2006), which aligns well with previous work incorporating spatial pattern efficiency metrics in the calibration of distributed hydrological models (e.g., López López et al., 2017; Demirel et al., 2018b; Zink et al., 2018; Dembélé et al., 2020b; Tong et al., 2021; Demirel et al., 2024). However, our results also show that, despite possible improvements in spatial efficiencies of different hydrological variables (e.g., ET in Fig. 4), large biases may remain in basin-scale annual cycles. In this regard, the examination of simulated seasonalities was crucial to detect deficiencies, especially for ET and fSCA (Rakovec et al., 2016), as an extension of using only performance metrics (Clark et al., 2021). Furthermore, during the optimization of the model parameters, we found that improving Q-only performance does not necessarily lead to improvements in metrics for other hydrological variables, highlighting the extent to which equifinality can affect model fidelity.

6.3 Limitations

In this study, we used a single satellite remote sensing product for SM_1, LST, and fSCA (i.e., we did not analyze uncertainties arising from the choice of the observational reference product). Further, we regularized one parameter at a time while maintaining the others spatially constant. However, two or more parameters could be simultaneously regularized (e.g., Mendoza et al., 2012, 2016), although that would increase the number of (super)parameters to be calibrated (see equation in Figure 2; Pokhrel et al., 2008). In this regard, a sensitivity analysis of VIC model parameters focused on spatial patterns would be useful. Additionally, the analyses presented here could be extended to other lumped or semi-distributed models, taking advantage of recently developed modular modeling platforms (e.g., Clark et al., 2015; Coxon et al., 2019; Knoben et al., 2019; Craig et al., 2020) to address well-known streamflow equifinality issues arising from compensatory effects in parameters, state variables, and fluxes (Khatami et al., 2019).

This study examined the effects of streamflow-only calibration of the VIC model on simulated spatial patterns and annual cycles of variables other than Q. Although we conducted additional calibration experiments by adding more variables to the OF (Q-ET and Q-SM_1), we did not conduct a systematic evaluation that includes all variables and basins, since the associated benefits have been widely discussed in the literature (e.g., López López et al., 2017; Demirel et al., 2018b; Koppa et al., 2019; Dembélé et al., 2020a; Tong et al., 2021; Shah et al., 2021; Pool et al., 2024).

The assessment of models in which the water available for vegetation transpiration comes from different soil layers may benefit from including satellite products that provide estimates of total water storage - such as GRACE (Güntner, 2008; Werth et al., 2009; Soltani et al., 2021b) - to complement other sources of information. However, the incorporation of such observational products in our study domain is challenging because of their relatively coarser horizontal resolution. Future work could explore the potential of statistically downscaled datasets (e.g., Rakovec et al., 2016; Yin et al., 2018; Vishwakarma et al., 2021; Fatolazadeh et al., 2022; Khorrami et al., 2023; Mei et al., 2023) for the calibration and evaluation of hydrological models in mountainous catchments.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

7 Conclusions

Despite the tremendous advances in the development of satellite remote sensing products that provide information on hydrological variables in space and time, streamflow-only calibration of distributed hydrological models remains a popular practice. To identify drawbacks in simulating spatial patterns and annual cycles of variables other than streamflow, we performed several streamflow-only calibration experiments with the VIC model in six basins located in continental Chile, testing the impacts of regularizing different soil parameters and the choice of calibration objective function. Our main conclusions are as follows:

- For a given streamflow-based OF and catchment, most of the spatial regularization strategies tested in this study provide similar OF values, with varying (either positive or negative) effects on the realism of simulated spatial patterns of other variables. None of the spatial model configurations tested here was able to provide simultaneous improvements in the spatial patterns and annual cycles of the variables evaluated in all catchments.
- Improvements in spatial patterns through spatial regularization techniques do not guarantee a correct simulation of the ET and fSCA annual cycles.
- For the VIC model, ET, SM_1, LST, and fSCA simulations do not necessarily improve when Q performance improves.
- For water balance characterizations, evaluating simulated annual cycles of ET (besides Q) is required to detect model or data deficiencies and, therefore, decide on the need to incorporate these variables within the calibration process.
- In model structures where several soil layers contribute to transpiration, differences in simulated annual cycles of soil moisture – obtained from different parameter configurations – may explain discrepancies in ET annual cycles, which, in turn, may provide very similar streamflow seasonalities (an extreme case of flux equifinality).

Data availability. Daily streamflow is available at the $(CR)^2$ Climate Explorer: https://www.cr2.cl/datos-de-caudales/. CR2MET (v2.0) can be found in Boisier (2023).

Author contributions. NV and PM conceptualized the study, designed the overall approach, and wrote the article. NV conducted all the model simulations, analyzed the results, and created all the figures. NM provided support in configuring the calibration experiments and the mizuRoute model. All the authors contributed to refining the analysis and interpretation of results, and reviewing and editing the article.

Competing interests. The contact author has declared that none of the authors has any competing interests.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Acknowledgements. Powered@NLHPC: This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02). Nicolás A. Vásquez and Pablo A. Mendoza received support from the Fondecyt project No. 11200142. Nicolás Vásquez also received support from the Emerging Leaders in the Americas Program (ELAP) scholarship (Canada) and the ANID Doctorado Nacional scholarship No. 21230289 (Chile). Pablo A. Mendoza also received financial support from ANID/CIA project No. 250010. The authors thank Nicolás Cortés-Salazar for his help with the configuration of mizuRoute, and Ximena Vargas, James McPhee, Camila Álvarez-Garretón and Mauricio Zambrano for their suggestions on earlier stages of this work.

Appendix A: Hydrological signatures

The OF_5 combines the KGE(Q) with five hydrological signatures using an Euclidean distance approach (Eq. 4). The five hydrological indices (derived from Yilmaz et al., 2008), evaluate biases in mean annual runoff (HS_{RR}), high flows (HS_{FHV}), low flows (HS_{FLV}), the slope of the flow duration curve (HS_{FMS}), and median daily flows (HS_{FMM}). The formulations of each hydrological signature is as follows:

$$HS_{RR} = \frac{\sum_{t=1}^{T}(Q_s^t - Q_o^t)}{\sum_{t=1}^{T} Q_o^t} \tag{A1}$$

where T is the number of time steps, and the s and o sub-indices refer to simulated and observed values, respectively.

$$HS_{FHV} = \frac{\sum_{h=1}^{H}(Q_s^h - Q_o^h)}{\sum_{h=1}^{H} Q_o^h} \tag{A2}$$

where $h = 1, 2, ..., H$ are indices for flows with exceedance probabilities lower than 0.02. The bias in low flow volumes is computed as follows:

$$HS_{FLV} = -1 \cdot \frac{\sum_{l=1}^{L}\left[log(Q_s^l) - log(Q_s^L)\right] - \sum_{l=1}^{L}\left[log(Q_o^l) - log(Q_o^L)\right]}{\sum_{l=1}^{L}\left[log(Q_o^l) - log(Q_o^L)\right]} \tag{A3}$$

where $l = 1, 2, ..., L$ are indices for flows with exceedance probabilities between 0.7 and 1, being L the index associated with the minimum flow. The bias in the mid-segment slope of the FDC is computed as:

$$HS_{FMS} = \frac{\left[log(Q_s^{m1}) - log(Q_s^{m2})\right] - \left[log(Q_o^{m1}) - log(Q_o^{m2})\right]}{\left[log(Q_o^{m1}) - log(Q_o^{m2})\right]} \tag{A4}$$

where $m1$ and $m2$ represent the lowest and highest exceedance probabilities (0.2 and 0.7, respectively) within the mid-segment of the flow duration curve. Finally, the bias in the median of daily flows is computed as:

$$HS_{FMM} = \frac{log(Q_s^{med}) - log(Q_o^{med})}{log(Q_o^{med})} \tag{A5}$$

where med corresponds to the median value.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

References

Aguayo, R., León-Muñoz, J., Garreaud, R., and Montecinos, A.: Hydrological droughts in the southern Andes (40–45°S) from an ensemble experiment using CMIP5 and CMIP6 models, Scientific Reports, 11, 5530, https://doi.org/10.1038/s41598-021-84807-4, 2021.

Akbar, R., Short Gianotti, D. J., Salvucci, G. D., and Entekhabi, D.: Partitioning of Historical Precipitation Into Evaporation and Runoff Based on Hydrologic Dynamics Identified With Recent SMAP Satellite Measurements, Water Resources Research, 56, https://doi.org/10.1029/2020WR027307, 2020.

Alfieri, L., Avanzi, F., Delogu, F., Gabellani, S., Bruno, G., Campo, L., Libertino, A., Massari, C., Tarpanelli, A., Rains, D., Miralles, D. G., Quast, R., Vreugdenhil, M., Wu, H., and Brocca, L.: High-resolution satellite products improve hydrological modeling in northern Italy, Hydrology and Earth System Sciences, 26, 3921–3939, https://doi.org/10.5194/hess-26-3921-2022, 2022.

Alvarez-Garreton, C., Mendoza, P. A., Pablo Boisier, J., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., and Ayala, A.: The CAMELS-CL dataset: Catchment attributes and meteorology for large sample studies-Chile dataset, Hydrology and Earth System Sciences, https://doi.org/10.5194/hess-22-5817-2018, 2018.

Andreadis, K. M. and Lettenmaier, D. P.: Assimilating remotely sensed snow observations into a macroscale hydrology model, Advances in Water Resources, 29, 872–886, https://doi.org/10.1016/j.advwatres.2005.08.004, 2006.

Bajracharya, A. R., Ahmed, M. I., Stadnyk, T., and Asadzadeh, M.: Process based calibration of a continental-scale hydrological model using soil moisture and streamflow data, Journal of Hydrology: Regional Studies, 47, 101 391, https://doi.org/10.1016/j.ejrh.2023.101391, 2023.

Beck, H. E., Pan, M., Lin, P., Seibert, J., Dijk, A. I. J. M. v., and Wood, E. F.: Global Fully Distributed Parameter Regionalization Based on Observed Streamflow From 4,229 Headwater Catchments, Journal of Geophysical Research: Atmospheres, 125, e2019JD031 485, https://doi.org/10.1029/2019JD031485, 2020.

Bennett, K. E., Cherry, J. E., Balk, B., and Lindsey, S.: Using MODIS estimates of fractional snow cover area to improve streamflow forecasts in interior Alaska, Hydrology and Earth System Sciences, 23, 2439–2459, https://doi.org/10.5194/hess-23-2439-2019, 2019.

Beven, K.: Prophecy, reality and uncertainty in distributed hydrological modelling, Advances in Water Resources, https://doi.org/10.1016/0309-1708(93)90028-E, 1993.

Beven, K.: How far can we go in distributed hydrological modelling?, Hydrology and Earth System Sciences, https://doi.org/10.5194/hess-5-1-2001, 2001.

Beven, K.: A manifesto for the equifinality thesis, in: Journal of Hydrology, https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, https://doi.org/10.1002/hyp.3360060305, 1992.

Boisier, J. P.: CR2MET: A high-resolution precipitation and temperature dataset for the period 1960-2021 in continental Chile. (v2.5) [Data set], Zenodo, https://doi.org/10.5281/zenodo.7529681, 2023.

Boisier, J. P., Alvarez-Garreton, C., Cordero, R. R., Damiani, A., Gallardo, L., Garreaud, R. D., Lambert, F., Ramallo, C., Rojas, M., and Rondanelli, R.: Anthropogenic drying in central-southern Chile evidenced by long-term observations and climate model simulations, Elementa: Science of the Anthropocene, 6, https://doi.org/10.1525/elementa.328, 2018.

Brooks, R. H. and Corey, A. T.: Hydraulic properties of porous media, Hydrology Papers, Colorado State University, 3, 1964.

Casper, M. C., Salm, Z., Gronz, O., Hutengs, C., Mohajerani, H., and Vohland, M.: Calibration of Land-Use-Dependent Evaporation Parameters in Distributed Hydrological Models Using MODIS Evaporation Time Series Data, Hydrology, 10, 216, https://doi.org/10.3390/hydrology10120216, 2023.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., and Rasmussen, R. M.: A unified approach for process-based hydrologic modeling: 1. Modeling concept, Water Resources Research, https://doi.org/10.1002/2015WR017198, 2015.

Clark, M. P., Vogel, R. M., Lamontagne, J. R., Mizukami, N., Knoben, W. J., Tang, G., Gharari, S., Freer, J. E., Whitfield, P. H., Shook, K., and Papalexiou, S.: The abuse of popular performance metrics in hydrologic modeling, Water Resources Research, p. e2020WR029001, https://doi.org/10.1029/2020WR029001, 2021.

Cornwell, E., Molotch, N. P., and McPhee, J.: Spatio-temporal variability of snow water equivalent in the extra-tropical Andes Cordillera from distributed energy balance modeling and remotely sensed snow cover, Hydrology and Earth System Sciences, https://doi.org/10.5194/hess-20-411-2016, 2016.

Cortés-Salazar, N., Vásquez, N. A., Mizukami, N., Mendoza, P. A., and Vargas, X.: To what extent does river routing matter in hydrological modeling?, Hydrology and Earth System Sciences, 27, 3505–3524, https://doi.org/10.5194/hess-27-3505-2023, 2023.

Coxon, G., Freer, J., Lane, R., Dunne, T., Knoben, W. J. M., Howden, N. J. K., Quinn, N., Wagener, T., and Woods, R.: DECIPHeR v1: Dynamic fluxEs and ConnectIvity for Predictions of HydRology, Geoscientific Model Development, 12, 2285–2306, https://doi.org/10.5194/gmd-12-2285-2019, 2019.

Craig, J. R., Brown, G., Chlumsky, R., Jenkinson, R. W., Jost, G., Lee, K., Mai, J., Serrer, M., Sgro, N., Shafii, M., Snowdon, A. P., and Tolson, B. A.: Flexible watershed simulation with the Raven hydrological modelling framework, Environmental Modelling & Software, 129, 104 728, https://doi.org/10.1016/j.envsoft.2020.104728, 2020.

de Lavenne, A., Andréassian, V., Thirel, G., Ramos, M. H., and Perrin, C.: A Regularization Approach to Improve the Sequential Calibration of a Semidistributed Hydrological Model, Water Resources Research, https://doi.org/10.1029/2018WR024266, 2019.

Dembélé, M., Ceperley, N., Zwart, S. J., Salvadore, E., Mariethoz, G., and Schaefli, B.: Potential of satellite and reanalysis evaporation datasets for hydrological modelling under various model calibration strategies, Advances in Water Resources, https://doi.org/10.1016/j.advwatres.2020.103667, 2020a.

Dembélé, M., Hrachowitz, M., Savenije, H. H., Mariéthoz, G., and Schaefli, B.: Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets, Water Resources Research, https://doi.org/10.1029/2019WR026085, 2020b.

Demirel, M., Koch, J., Mendiguren, G., and Stisen, S.: Spatial Pattern Oriented Multicriteria Sensitivity Analysis of a Distributed Hydrologic Model, Water, 10, 1188, https://doi.org/10.3390/w10091188, 2018a.

Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., and Stisen, S.: Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, Hydrology and Earth System Sciences, https://doi.org/10.5194/hess-22-1299-2018, 2018b.

Demirel, M. C., Koch, J., Rakovec, O., Kumar, R., Mai, J., Müller, S., Thober, S., Samaniego, L., and Stisen, S.: Tradeoffs Between Temporal and Spatial Pattern Calibration and Their Impacts on Robustness and Transferability of Hydrologic Model Parameters to Ungauged Basins, Water Resources Research, 60, https://doi.org/10.1029/2022WR034193, 2024.

DGA: Homologación del cálculo hidrológico para la estimación de la oferta natural del agua histórica y futura en Chile., Tech. rep., SIT N° 524. Ministerio de Obras Públicas, Dirección General de Aguas, División de Estudios y Planificación, Chile. Elaborado por Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas., https://snia.mop.gob.cl/repositoriodga/handle/20.500.13000/126394, 2022.

Dickinson, R. E.: Modeling evapotranspiration for three-dimensional global climate models, in: Climate processes and climate sensitivities, edited by Hansen, J. E. and Takahashi, T., vol. 29, pp. 58–72, American Geosciences Union, https://doi.org/10.1029/GM029p0058, 1984.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185–215, https://doi.org/10.1016/J.RSE.2017.07.001, 2017.

dos Santos Araujo, D. C., Gico Lima Montenegro, S. M., Ribeiro Neto, A., and da Silva, S. F.: Evaluation of satellite-based soil moisture for agricultural drought monitoring in the Brazilian semiarid region, Remote Sensing Applications: Society and Environment, 33, 101 111, https://doi.org/10.1016/j.rsase.2023.101111, 2024.

Draper, C. S., Walker, J. P., Steinle, P. J., de Jeu, R. A., and Holmes, T. R.: An evaluation of AMSR–E derived soil moisture over Australia, Remote Sensing of Environment, 113, 703–710, https://doi.org/10.1016/j.rse.2008.11.011, 2009.

Duethmann, D., Peters, J., Blume, T., Vorogushyn, S., and Güntner, A.: The value of satellite-derived snow cover images for calibrating a hydrological model in snow-dominated catchments in Central Asia, Water Resources Research, 50, 2002–2021, https://doi.org/10.1002/2013WR014382, 2014.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, Reviews of Geophysics, 45, RG2004, https://doi.org/10.1029/2005RG000183, 2007.

Fatolazadeh, F., Eshagh, M., and Goïta, K.: New spectro-spatial downscaling approach for terrestrial and groundwater storage variations estimated by GRACE models, Journal of Hydrology, 615, 128 635, https://doi.org/10.1016/j.jhydrol.2022.128635, 2022.

Fowler, K., Coxon, G., Freer, J., Peel, M., Wagener, T., Western, A., Woods, R., and Zhang, L.: Simulating Runoff Under Changing Climatic Conditions: A Framework for Model Improvement, Water Resources Research, https://doi.org/10.1029/2018WR023989, 2018a.

Fowler, K., Peel, M., Western, A., and Zhang, L.: Improved Rainfall-Runoff Calibration for Drying Climate: Choice of Objective Function, Water Resources Research, https://doi.org/10.1029/2017WR022466, 2018b.

Franchini, M. and Pacciani, M.: Comparative analysis of several conceptual rainfall-runoff models, Journal of Hydrology, 122, 161–219, https://doi.org/10.1016/0022-1694(91)90178-K, 1991.

Garcia, F., Folton, N., and Oudin, L.: Which objective function to calibrate rainfall–runoff models for low-flow index simulations?, Hydrological Sciences Journal, 62, 1149–1166, https://doi.org/10.1080/02626667.2017.1308511, 2017.

Geophysics Department, U. d. C. and Ministerio de Energía, G. d. C.: Explorador Eólico, Tech. rep., Geophysic Deparment, Universidad de Chile, https://eolico.minenergia.cl/inicio, 2018.

Givovich, F., Mendoza, P. A., Vásquez, N. A., Murillo, O., Muñoz-Castro, E., and Ayala, A.: Representing subgrid precipitation variability in snowpack and hydrological modeling: Is adding complexity worth it?, Journal of Hydrology, 662, 134 038, https://doi.org/10.1016/j.jhydrol.2025.134038, 2025.

Güntner, A.: Improvement of Global Hydrological Models Using GRACE Data, Surveys in Geophysics, 29, 375–397, https://doi.org/10.1007/s10712-008-9038-y, 2008.

Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, Water Resources Research, https://doi.org/10.1029/97WR03495, 1998.

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

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Hall, D. K. and Riggs, G. A.: MODIS/Terra snow cover Daily L3 Global 500m SIN GRID, version 6, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD10A1.006, 2016.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Iziomon, M. G., Mayer, H., and Matzarakis, A.: Downward atmospheric longwave irradiance under clear and cloudy skies: Measurement and parameterization, Journal of Atmospheric and Solar-Terrestrial Physics, 65, 1107–1116, https://doi.org/10.1016/j.jastp.2003.07.007, 2003.

Khatami, S., Peel, M. C., Peterson, T. J., and Western, A. W.: Equifinality and Flux Mapping: A New Approach to Model Evaluation and Process Representation Under Uncertainty, Water Resources Research, https://doi.org/10.1029/2018WR023750, 2019.

Khorrami, B., Pirasteh, S., Ali, S., Sahin, O. G., and Vaheddoost, B.: Statistical downscaling of GRACE TWSA estimates to a 1-km spatial resolution for a local-scale surveillance of flooding potential, Journal of Hydrology, 624, 129 929, https://doi.org/10.1016/j.jhydrol.2023.129929, 2023.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, https://doi.org/10.1029/2005WR004362, 2006.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424-425, 264–277, https://doi.org/10.1016/J.JHYDROL.2012.01.011, 2012.

Knoben, W. J. M., Freer, J. E., Fowler, K. J. A., Peel, M. C., and Woods, R. A.: Modular Assessment of Rainfall–Runoff Models Toolbox (MARRMoT) v1.2: an open-source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous state-space formulations, Geoscientific Model Development, 12, 2463–2480, https://doi.org/10.5194/gmd-12-2463-2019, 2019.

Koch, J., Mendiguren, G., Mariethoz, G., and Stisen, S.: Spatial sensitivity analysis of simulated land surface patterns in a catchment model using a set of innovative spatial performance metrics, Journal of Hydrometeorology, https://doi.org/10.1175/JHM-D-16-0148.1, 2017.

Koch, J., Demirel, M. C., and Stisen, S.: The SPAtial EFficiency metric (SPAEF): Multiple-component evaluation of spatial patterns for optimization of hydrological models, Geoscientific Model Development, https://doi.org/10.5194/gmd-11-1873-2018, 2018.

Kollat, J. B., Reed, P. M., and Wagener, T.: When are multiobjective calibration trade-offs in hydrologic models meaningful?, Water Resources Research, 48, https://doi.org/10.1029/2011WR011534, 2012.

Koppa, A., Gebremichael, M., and Yeh, W. W.: Multivariate calibration of large scale hydrologic models: The necessity and value of a Pareto optimal approach, Advances in Water Resources, https://doi.org/10.1016/j.advwatres.2019.06.005, 2019.

Kurugama, K. M., Kazama, S., and Hiraga, Y.: Augmenting observation network design and assimilation frequency in distributed hydrological models: insights from the LISFLOOD-based hydrological data assimilation framework, Journal of Hydrology, 667, 134 853, https://doi.org/10.1016/j.jhydrol.2025.134853, 2026.

Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., and Wood, E. F.: Inroads of remote sensing into hydrologic science during the WRR era, Water Resources Research, 51, 7309–7342, https://doi.org/10.1002/2015WR017616, 2015.

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

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

López López, P., Sutanudjaja, E. H., Schellekens, J., Sterk, G., and Bierkens, M. F. P.: Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products, Hydrology and Earth System Sciences, 21, 3125–3144, https://doi.org/10.5194/hess-21-3125-2017, 2017.

Matott, L.: OSTRICH: an Optimization Software Tool, Documentation and User's Guide, Version 17.12.19., http://www.civil.uwaterloo.ca/envmodelling/Ostrich.html, 2017.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E., Franz, T. E., Shi, J., Gao, H., and Wood, E. F.: The future of Earth observation in hydrology, Hydrology and Earth System Sciences, https://doi.org/10.5194/hess-21-3879-2017, 2017.

McInerney, D., Thyer, M., Kavetski, D., Lerat, J., and Kuczera, G.: Improving probabilistic prediction of daily streamflow by identifying Pareto optimal approaches for modeling heteroscedastic residual errors, Water Resources Research, https://doi.org/10.1002/2016WR019168, 2017.

Mei, Y., Mai, J., Do, H. X., Gronewold, A., Reeves, H., Eberts, S., Niswonger, R., Regan, R. S., and Hunt, R. J.: Can Hydrological Models Benefit From Using Global Soil Moisture, Evapotranspiration, and Runoff Products as Calibration Targets?, Water Resources Research, 59, https://doi.org/10.1029/2022WR032064, 2023.

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Clark, M., and Uijlenhoet, R.: Representation of spatial and temporal variability in large-domain hydrological models: Case study for a mesoscale pre-Alpine basin, Hydrology and Earth System Sciences, 20, 2207–2226, https://doi.org/10.5194/hess-20-2207-2016, 2016.

Mendoza, P. A., McPhee, J., and Vargas, X.: Uncertainty in flood forecasting: A distributed modeling approach in a sparse data catchment, Water Resources Research, 48, https://doi.org/10.1029/2011WR011089, 2012.

Mendoza, P. A., Clark, M. P., Mizukami, N., Gutmann, E. D., Arnold, J. R., Brekke, L. D., and Rajagopalan, B.: How do hydrologic modeling decisions affect the portrayal of climate change impacts?, Hydrological Processes, https://doi.org/10.1002/hyp.10684, 2016.

Merz, R., Parajka, J., and Blöschl, G.: Time stability of catchment model parameters: Implications for climate impact analyses, Water Resources Research, https://doi.org/10.1029/2010WR009505, 2011.

Meyer Oliveira, A., Fleischmann, A., and Paiva, R.: On the contribution of remote sensing-based calibration to model hydrological and hydraulic processes in tropical regions, Journal of Hydrology, 597, 126 184, https://doi.org/10.1016/j.jhydrol.2021.126184, 2021.

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

Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resources Research, 53, 8020–8040, https://doi.org/10.1002/2017WR020401, 2017.

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

Mostafaie, A., Forootan, E., Safari, A., and Schumacher, M.: Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily GRACE data, Computational Geosciences, 22, 789–814, https://doi.org/10.1007/s10596-018-9726-8, 2018.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Mu, Q., Zhao, M., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sensing of Environment, 115, 1781–1800, https://doi.org/10.1016/j.rse.2011.02.019, 2011.

Muñoz Sabater, J.: ERA5-Land hourly data from 1981 to present, 2019.

Murillo, O., Mendoza, P. A., Vásquez, N. A., Mizukami, N., and Ayala, A.: Impacts of subgrid elevation bands on hydrological portrayals: insights from a suite of hydroclimatically diverse mountainous catchments, Earth and Space Science Open Archive, submitted to Water Resources Research, p. 31, https://doi.org/10.1002/essoar.10510847.1, 2022.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I - A discussion of principles, Journal of Hydrology, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nijzink, R. C., Samaniego, L., Mai, J., Kumar, R., Thober, S., Zink, M., Schäfer, D., Savenije, H. H. G., and Hrachowitz, M.: The importance of topography-controlled sub-grid process heterogeneity and semi-quantitative prior constraints in distributed hydrological models, Hydrology and Earth System Sciences, 20, 1151–1176, https://doi.org/10.5194/hess-20-1151-2016, 2016.

Niño, Y.: Simple Model for Downstream Variation of Median Sediment Size in Chilean Rivers, Journal of Hydraulic Engineering, 128, 934–941, https://doi.org/10.1061/(ASCE)0733-9429(2002)128:10(934), 2002.

Parajka, J. and Blöschl, G.: The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models, Journal of Hydrology, 358, 240–258, https://doi.org/10.1016/j.jhydrol.2008.06.006, 2008.

Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G. B., Kempen, B., Ribeiro, E., and Rossiter, D.: SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty, SOIL, 7, 217–240, https://doi.org/10.5194/SOIL-7-217-2021, 2021.

Pokhrel, P. and Gupta, H. V.: On the use of spatial regularization strategies to improve calibration of distributed watershed models, Water Resources Research, https://doi.org/10.1029/2009WR008066, 2010.

Pokhrel, P., Gupta, H. V., and Wagener, T.: A spatial regularization approach to parameter estimation for a distributed watershed model, Water Resources Research, https://doi.org/10.1029/2007WR006615, 2008.

Pool, S., Vis, M., and Seibert, J.: Evaluating model performance: towards a non-parametric variant of the Kling-Gupta efficiency, Hydrological Sciences Journal, https://doi.org/10.1080/02626667.2018.1552002, 2018.

Pool, S., Fowler, K., and Peel, M.: Benefit of Multivariate Model Calibration for Different Climatic Regions, Water Resources Research, 60, https://doi.org/10.1029/2023WR036364, 2024.

Rakovec, O., Kumar, R., Attinger, S., and Samaniego, L.: Improving the realism of hydrologic model functioning through multivariate parameter estimation, Water Resources Research, 52, 7779–7792, https://doi.org/10.1002/2016WR019430, 2016.

Rakovec, O., Mizukami, N., Kumar, R., Newman, A. J., Thober, S., Wood, A. W., Clark, M. P., and Samaniego, L.: Diagnostic Evaluation of Large-Domain Hydrologic Models Calibrated Across the Contiguous United States, Journal of Geophysical Research: Atmospheres, https://doi.org/10.1029/2019JD030767, 2019.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., and Seo, D. J.: Overall distributed model intercomparison project results, Journal of Hydrology, 298, 27–60, https://doi.org/10.1016/J.JHYDROL.2004.03.031, 2004.

Refsgaard, J. C. and Knudsen, J.: Operational validation and intercomparison of different types of hydrological models, Water Resources Research, 32, 2189–2202, https://doi.org/10.1029/96WR00896, 1996.

Saavedra, D., Mendoza, P. A., Addor, N., Llauca, H., and Vargas, X.: A multi-objective approach to select hydrological models and constrain structural uncertainties for climate impact assessments, Hydrological Processes, https://doi.org/10.1002/hyp.14446, 2022.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, 5523, https://doi.org/10.1029/2008WR007327, 2010.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Schoups, G., Lee Addams, C., and Gorelick, S. M.: Multi-objective calibration of a surface water-groundwater flow model in an irrigated agricultural region: Yaqui Valley, Sonora, Mexico, Hydrology and Earth System Sciences, 9, 549–568, https://doi.org/10.5194/hess-9-549-2005, 2005.

Sepúlveda, U. M., Mendoza, P. A., Mizukami, N., and Newman, A. J.: Revisiting parameter sensitivities in the variable infiltration capacity model across a hydroclimatic gradient, Hydrology and Earth System Sciences, 26, 3419–3445, https://doi.org/10.5194/hess-26-3419-2022, 2022.

Shafii, M. and Tolson, B. A.: Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives, Water Resources Research, https://doi.org/10.1002/2014WR016520, 2015.

Shah, S., Duan, Z., Song, X., Li, R., Mao, H., Liu, J., Ma, T., and Wang, M.: Evaluating the added value of multi-variable calibration of SWAT with remotely sensed evapotranspiration data for improving hydrological modeling, Journal of Hydrology, 603, 127 046, https://doi.org/10.1016/j.jhydrol.2021.127046, 2021.

Sheffield, J., Wood, E. F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., and Verbist, K.: Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions, Water Resources Research, 54, 9724–9758, https://doi.org/10.1029/2017WR022437, 2018.

Soltani, M., Bjerre, E., Koch, J., and Stisen, S.: Integrating remote sensing data in optimization of a national water resources model to improve the spatial pattern performance of evapotranspiration, Journal of Hydrology, 603, 127 026, https://doi.org/10.1016/j.jhydrol.2021.127026, 2021a.

Soltani, S. S., Ataie-Ashtiani, B., and Simmons, C. T.: Review of assimilating GRACE terrestrial water storage data into hydrological models: Advances, challenges and opportunities, Earth-Science Reviews, 213, 103 487, https://doi.org/10.1016/j.earscirev.2020.103487, 2021b.

Stisen, S., Koch, J., Sonnenborg, T. O., Refsgaard, J. C., Bircher, S., Ringgaard, R., and Jensen, K. H.: Moving beyond run-off calibration—Multivariable optimization of a surface–subsurface–atmosphere model, Hydrological Processes, 32, 2654–2668, https://doi.org/10.1002/hyp.13177, 2018.

Sulla-Menashe, D. and Friedl, M. A.: User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product, Tech. rep., National Aeronautics and Space Administration, https://doi.org/10.5067/MODIS/MCD12Q1, 2018.

Tang, G., Clark, M. P., Knoben, W. J. M., Liu, H., Gharari, S., Arnal, L., Beck, H. E., Wood, A. W., Newman, A. J., and Papalexiou, S. M.: The Impact of Meteorological Forcing Uncertainty on Hydrological Modeling: A Global Analysis of Cryosphere Basins, Water Resources Research, 59, https://doi.org/10.1029/2022WR033767, 2023.

Tiwari, D., Trudel, M., and Leconte, R.: On optimization of calibrations of a distributed hydrological model with spatially distributed information on snow, Hydrology and Earth System Sciences, 28, 1127–1146, https://doi.org/10.5194/hess-28-1127-2024, 2024.

Todorović, A., Grabs, T., and Teutschbein, C.: Advancing traditional strategies for testing hydrological model fitness in a changing climate, Hydrological Sciences Journal, 67, 1790–1811, https://doi.org/10.1080/02626667.2022.2104646, 2022.

Tolson, B. A. and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, Water Resources Research, 43, 1413, https://doi.org/10.1029/2005WR004723, 2007.

Tong, R., Parajka, J., Salentinig, A., Pfeil, I., Komma, J., Szeles, B., KubáÅ, M., Valent, P., Vreugdenhil, M., Wagner, W., and Bloschl, G.: The value of ASCAT soil moisture and MODIS snow cover data for calibrating a conceptual hydrologic model, Hydrology and Earth System Sciences, 25, 1389–1410, https://doi.org/10.5194/hess-25-1389-2021, 2021.

Trautmann, T., Koirala, S., Guentner, A., Kim, H., and Jung, M.: Calibrating global hydrological models with GRACE TWS: does river storage matter?, Environmental Research Communications, 5, 081 005, https://doi.org/10.1088/2515-7620/acece5, 2023.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

Troch, P. A., Carrillo, G., Sivapalan, M., Wagener, T., and Sawicz, K.: Climate-vegetation-soil interactions and long-term hydrologic partitioning: signatures of catchment co-evolution, Hydrology and Earth System Sciences, 17, 2209–2217, https://doi.org/10.5194/hess-17-2209-2013, 2013.

Ul Hassan, Z., Jefferson, A. J., Avellaneda, P. M., and Bhaskar, A. S.: Assessment of hydrological parameter uncertainty versus climate projection spread on urban streamflow and floods, Journal of Hydrology, 638, 131 546, https://doi.org/10.1016/j.jhydrol.2024.131546, 2024.

Vishwakarma, B. D., Zhang, J., and Sneeuw, N.: Downscaling GRACE total water storage change using partial least squares regression, Scientific Data, 8, 95, https://doi.org/10.1038/s41597-021-00862-6, 2021.

Wagner, P. D., Duethmann, D., Kiesel, J., Pool, S., Hrachowitz, M., Ceola, S., Herzog, A., Houska, T., Loritz, R., Spieler, D., Staudinger, M., Tarasova, L., Thober, S., Fohrer, N., Tetzlaff, D., Wagener, T., and Guse, B.: The Unexploited Treasures of Hydrological Observations Beyond Streamflow for Catchment Modeling, WIREs Water, 12, https://doi.org/10.1002/wat2.70018, 2025.

Wan, Z.: New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product, Remote Sensing of Environment, 140, 36–45, https://doi.org/10.1016/j.rse.2013.08.027, 2014.

Wang, S., Peng, H., Hu, Q., and Jiang, M.: Analysis of runoff generation driving factors based on hydrological model and interpretable machine learning method, Journal of Hydrology: Regional Studies, 42, 101 139, https://doi.org/10.1016/j.ejrh.2022.101139, 2022.

Werth, S. and Güntner, A.: Calibration analysis for water storage variability of the global hydrological model WGHM, Hydrology and Earth System Sciences, 14, 59–78, https://doi.org/10.5194/hess-14-59-2010, 2010.

Werth, S., Güntner, A., Petrovic, S., and Schmidt, R.: Integration of GRACE mass variations into a global hydrological model, Earth and Planetary Science Letters, 277, 166–173, https://doi.org/10.1016/j.epsl.2008.10.021, 2009.

Westerberg, I. K., Guerrero, J. L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., and Xu, C. Y.: Calibration of hydrological models using flow-duration curves, Hydrology and Earth System Sciences, https://doi.org/10.5194/hess-15-2205-2011, 2011.

Wood, E. F., Lettenmaier, D. P., and Zartarian, V. G.: A land-surface hydrology parameterization with subgrid variability for general circulation models, Journal of Geophysical Research, 97, 2717, https://doi.org/10.1029/91JD01786, 1992.

Yáñez-Morroni, G., Suárez, F., Muñoz, J. F., and Lagos, M. S.: Hydrological modeling of the Silala River basin. 2. Validation of hydrological fluxes with contemporary data, WIREs Water, https://doi.org/10.1002/wat2.1696, 2023.

Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, Water Resources Research, https://doi.org/10.1029/2007WR006716, 2008.

Yin, W., Hu, L., Zhang, M., Wang, J., and Han, S.: Statistical Downscaling of GRACE-Derived Groundwater Storage Using ET Data in the North China Plain, Journal of Geophysical Research: Atmospheres, 123, 5973–5987, https://doi.org/10.1029/2017JD027468, 2018.

Zhao, R., Zhang, Y. L., Fang, X., Liu, R., and Zhang, Q. S.: The Xinanjiang Model, in: Hydrological Forecasting Proceedings Oxford Symposium, IAHS 129, pp. 351–356, 1980.

Zink, M., Mai, J., Cuntz, M., and Samaniego, L.: Conditioning a Hydrologic Model Using Patterns of Remotely Sensed Land Surface Temperature, Water Resources Research, 54, 2976–2998, https://doi.org/10.1002/2017WR021346, 2018.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

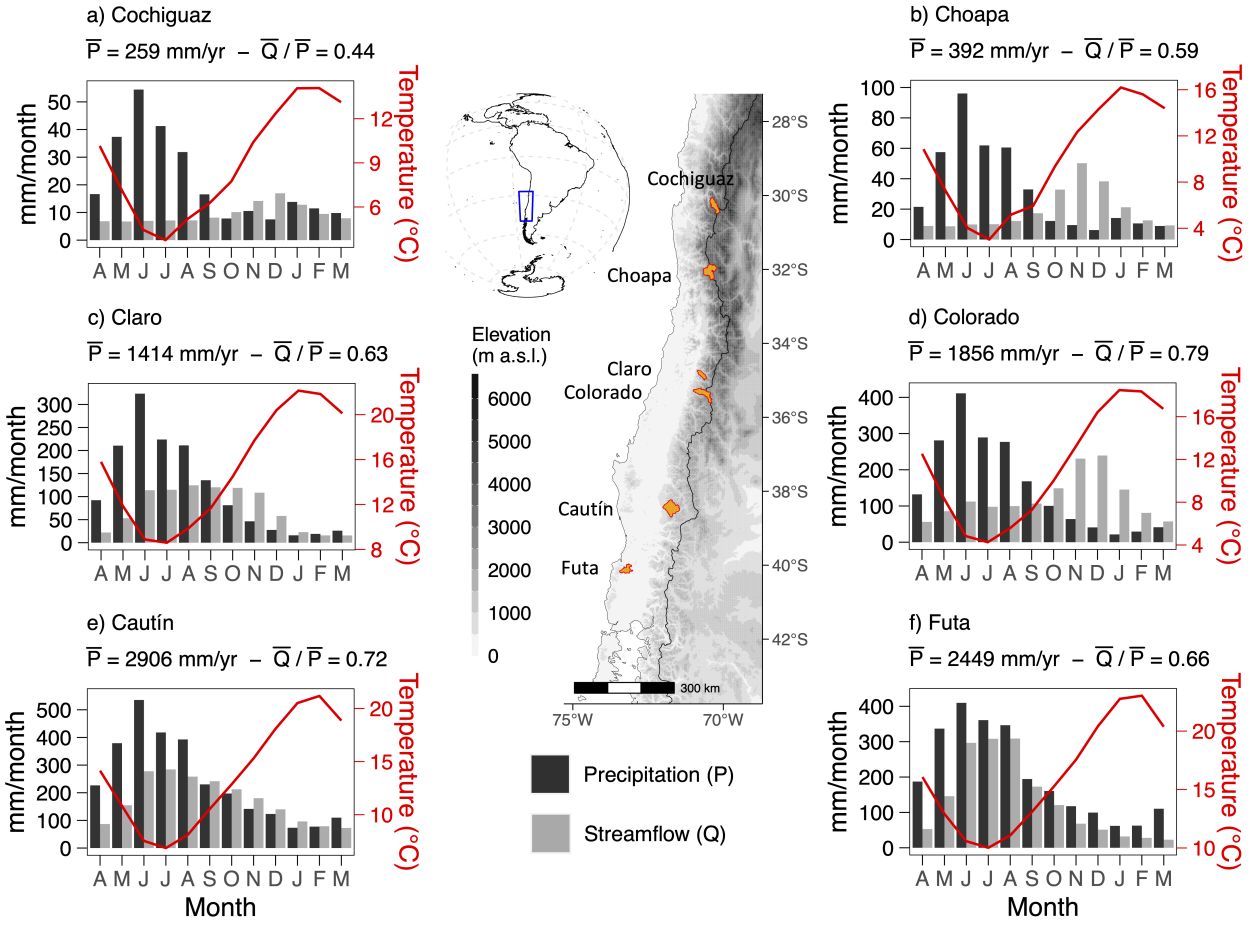

Figure 1. Location of the six case study basins considered in this study (center panel), along with the seasonal cycles of precipitation (P), streamflow (Q), and temperature (in red) for the climatological period 1980-2018. Overlines represent mean annual values. Notice that the winter season corresponds to JJA, while the summer season to DJF.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

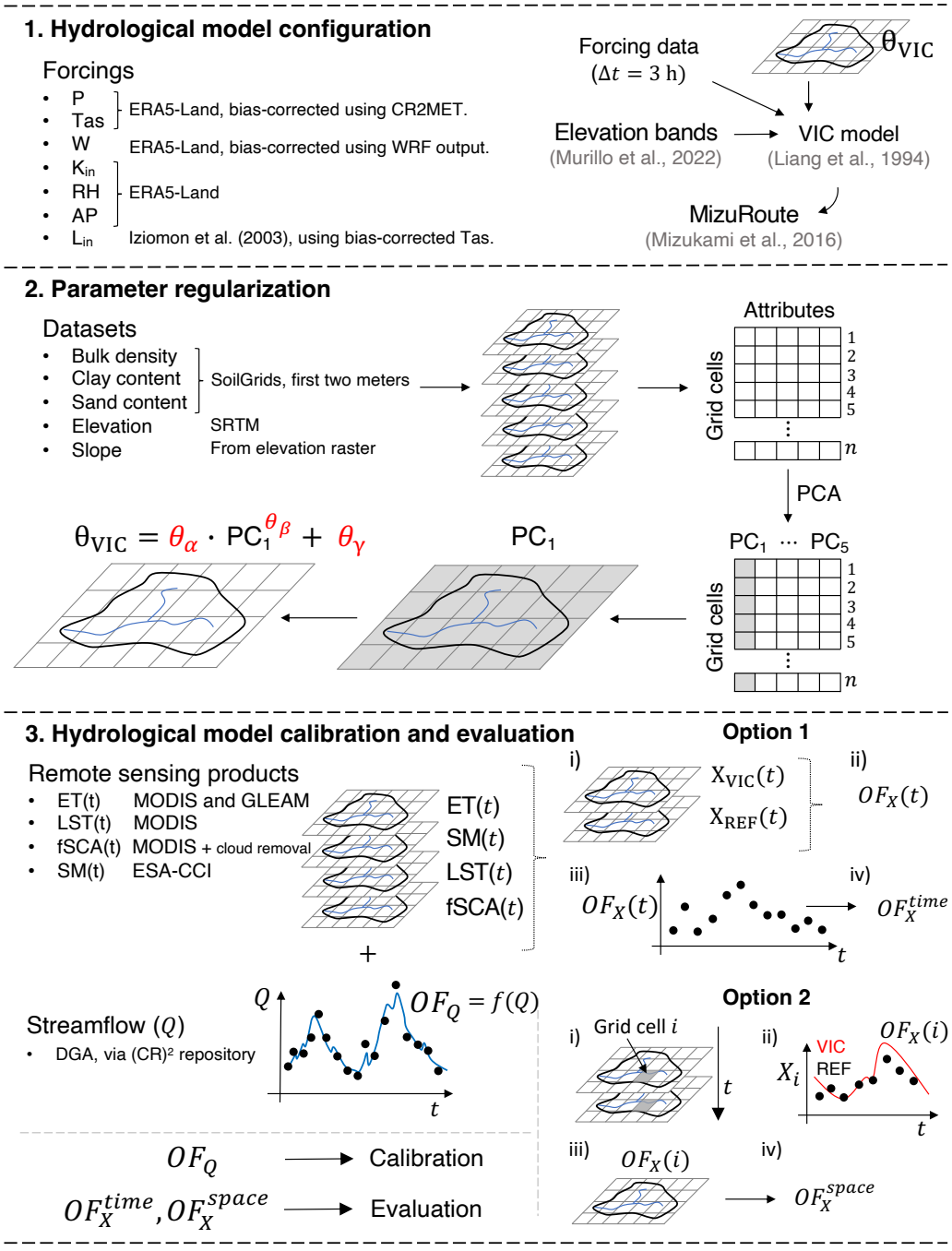

Figure 2. Schematics of the methodology used in this study. (1) Datasets and the main model configuration features used to run the VIC and mizuRoute models. (2) Approach used to obtain *a priori* soil parameter fields for each case study basin. (3) Evaluation of streamflow (Q) simulations and spatial patterns of other variables (X) by contrasting: a simulated and reference map for the variable X (e.g., ET, SM_1, LST or fSCA) for each time step t (Option 1), and simulated and reference time series at each grid cell (Option 2).

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

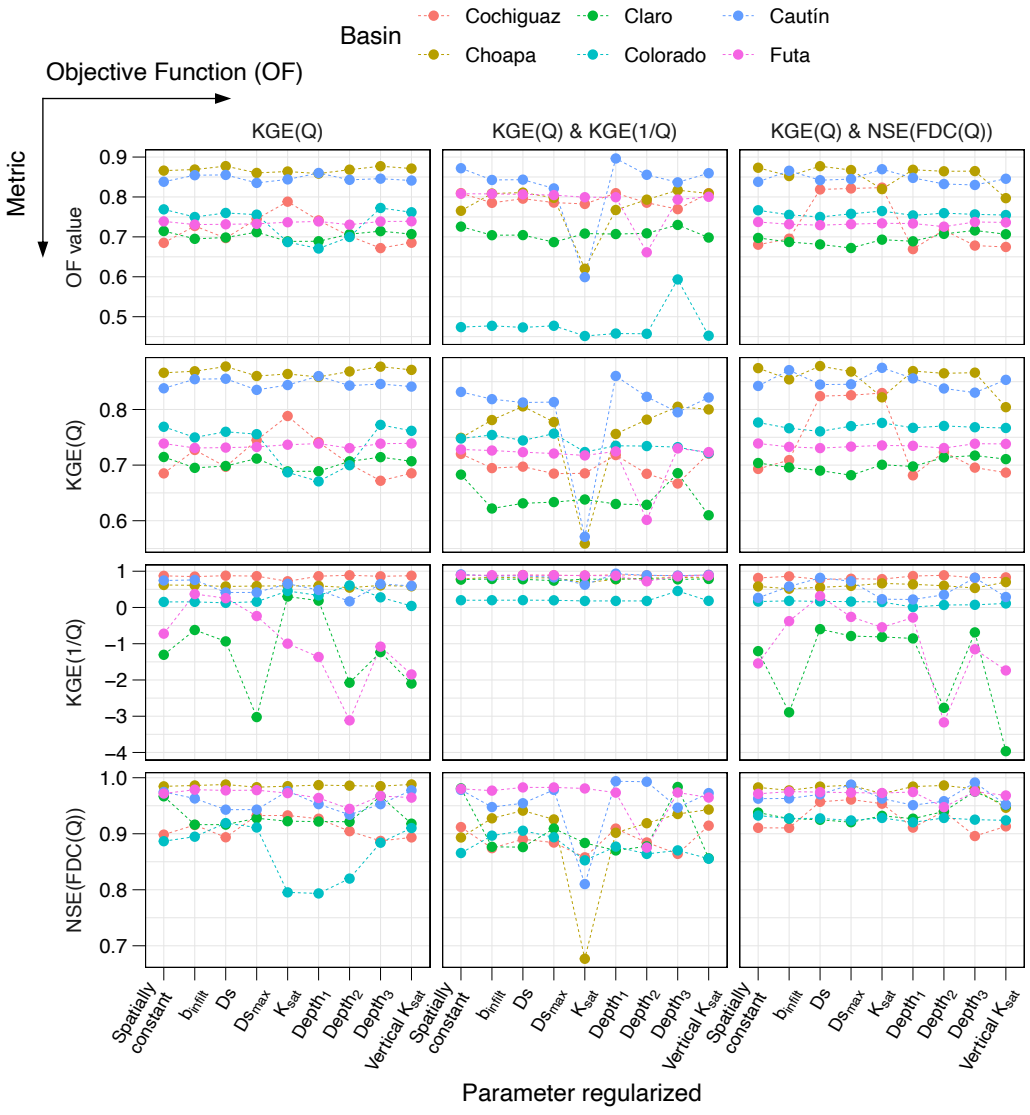

Figure 3. Best objective function value (top row) and streamflow performance metrics (remaining rows) associated with the best parameter set obtained with spatially constant parameters (i.e., benchmark calibration, $\theta_\alpha = 0$) and different regularization strategies during the calibration period (April/2005-March/2018). The columns show results for three different calibration objective functions. In each panel, the results for individual basins are displayed with different colors.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

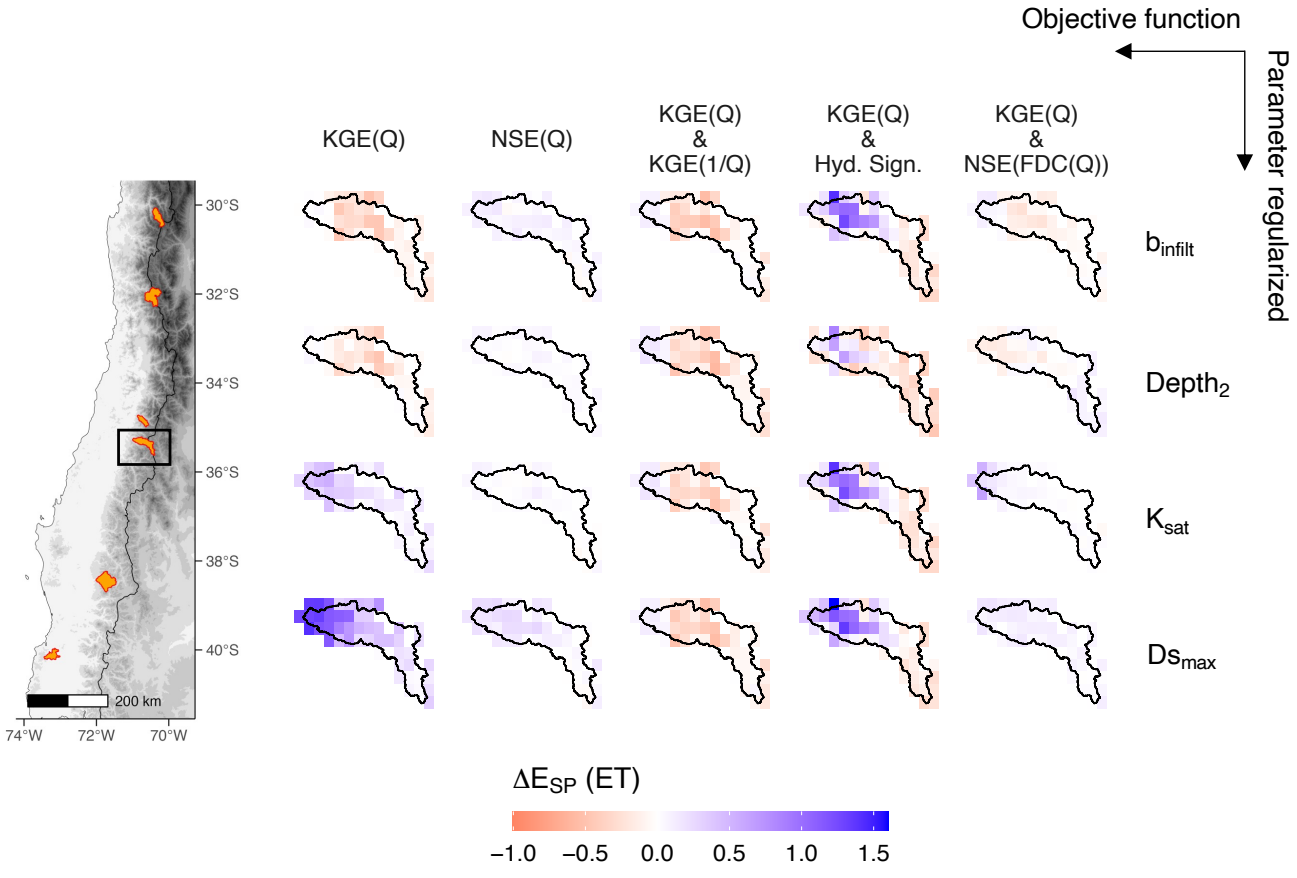

Figure 4. Change in ET performance (with respect to the benchmark calibration) for the Colorado River basin during the calibration period (Apr/2005-March/2018). Blue (red) colors represent an improvement (decline) in the performance of simulations. Note that the benchmark changes among objective functions.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

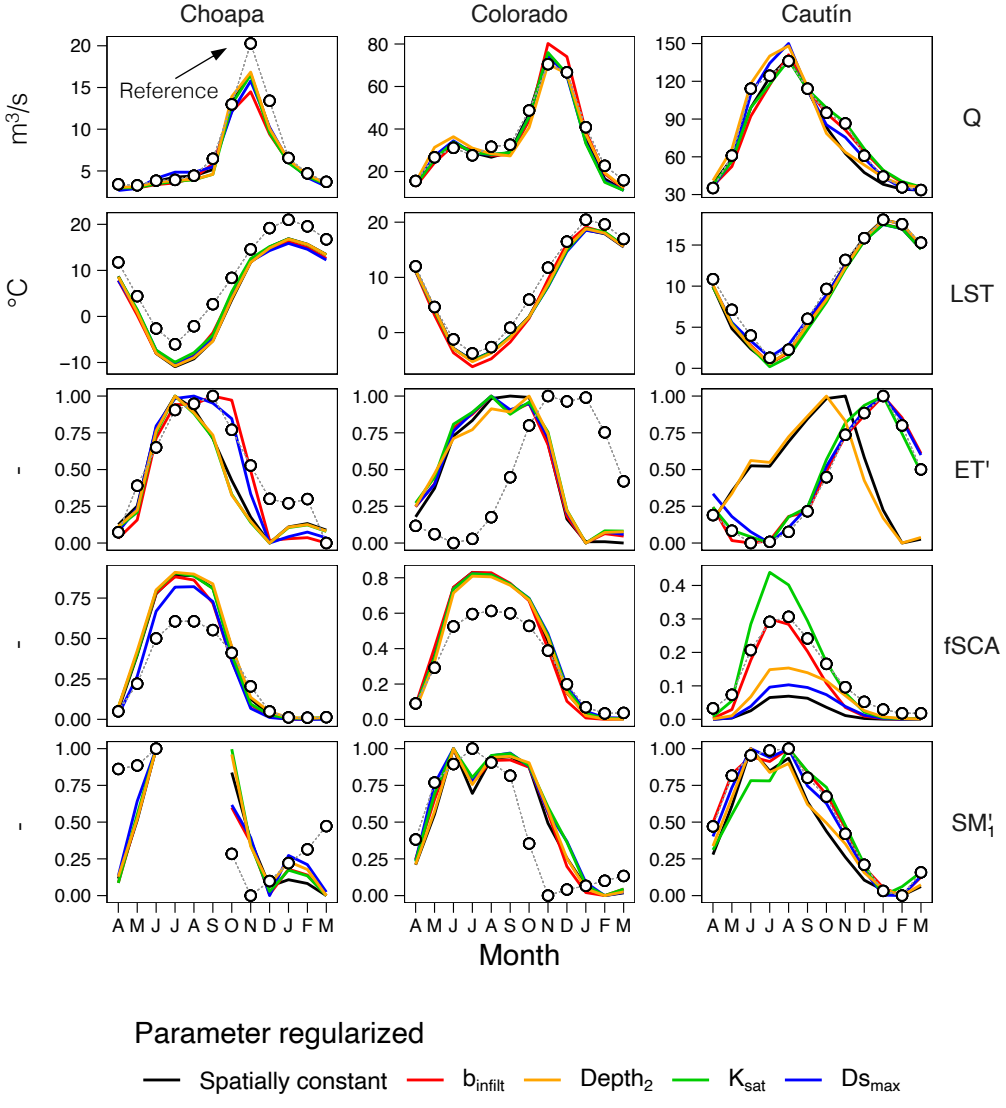

Figure 5. Catchment-scale annual cycles of streamflow (Q), normalized evapotranspiration (ET'), normalized soil moisture (SM_1'), land surface temperature (LST), and fractional snow-covered area (fSCA) in the Choapa (snowmelt-driven), Colorado (mixed regime), and Cautín (rainfall-driven) River basins (calibration period 2005-2018). Mean monthly values are computed only if at least 50 days with information are available, and the calculation considers only days with information. Note that there is not enough data to compute monthly SM_1 averages at the Choapa River basin during winter. All the results are associated with the parameter sets that maximize KGE(Q) & KGE(1/Q). Notice that winter and summer correspond to JJA and DJF, respectively. Streamflow observations and remotely sensed variables are referred to as "references" and are represented with white dots. ET and SM_1 annual values are normalized as $X' = (X - X_{min})/(X_{max} - X_{min})$.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

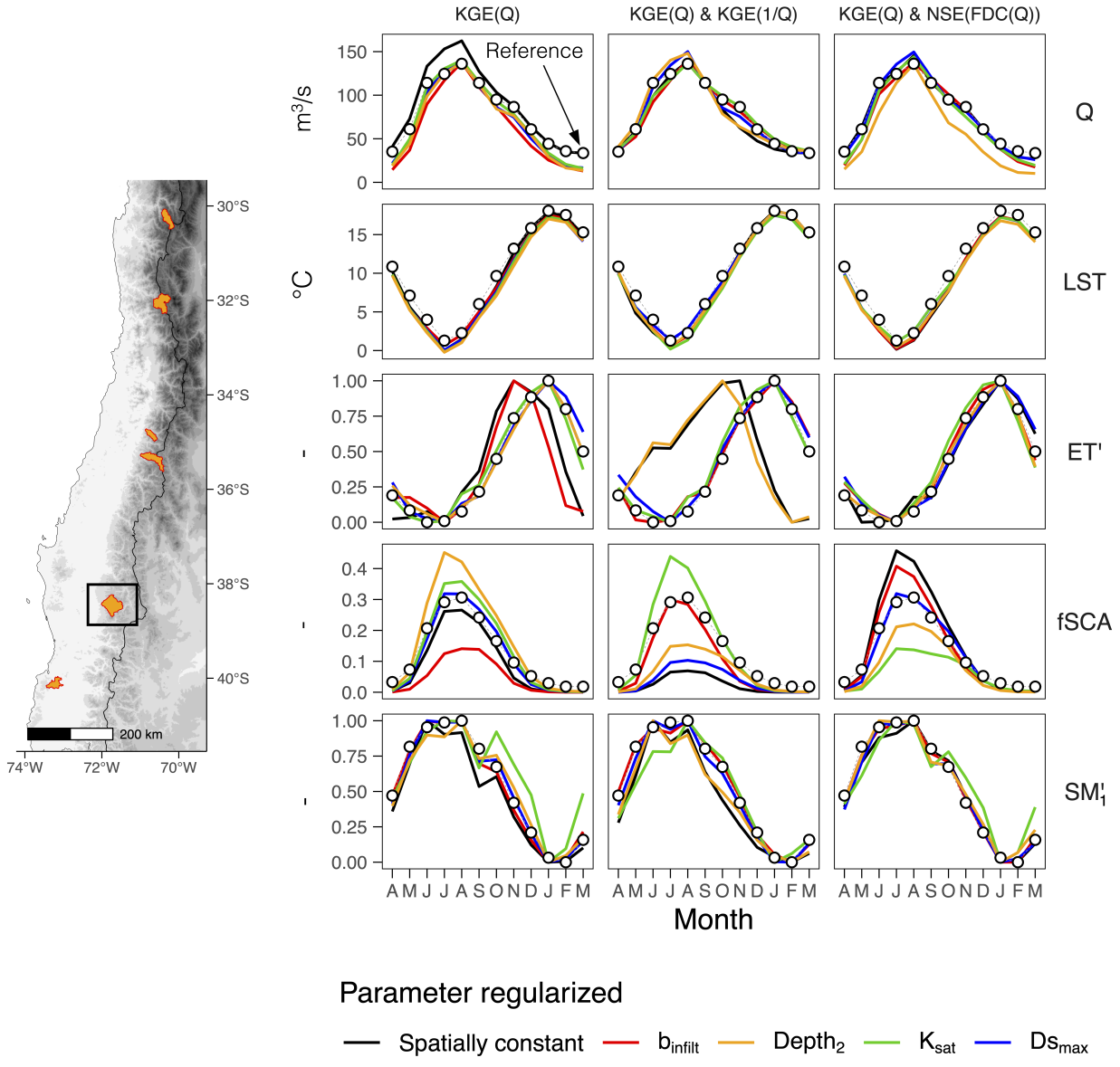

Figure 6. Impacts of the spatial regularization strategy on simulated annual cycles for a given calibration metric (columns). The variables analyzed are streamflow (Q), normalized evapotranspiration (ET'), normalized soil moisture (SM_1'), land surface temperature (LST), and fractional snow-covered area (fSCA) for the Cautín River basin (highlighted by a square on the map) during the calibration period (2005-2018). The reference datasets are shown as white dots. The normalization of ET and SM_1 is computed as $X' = (X - X_{min})/(X_{max} - X_{min})$.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

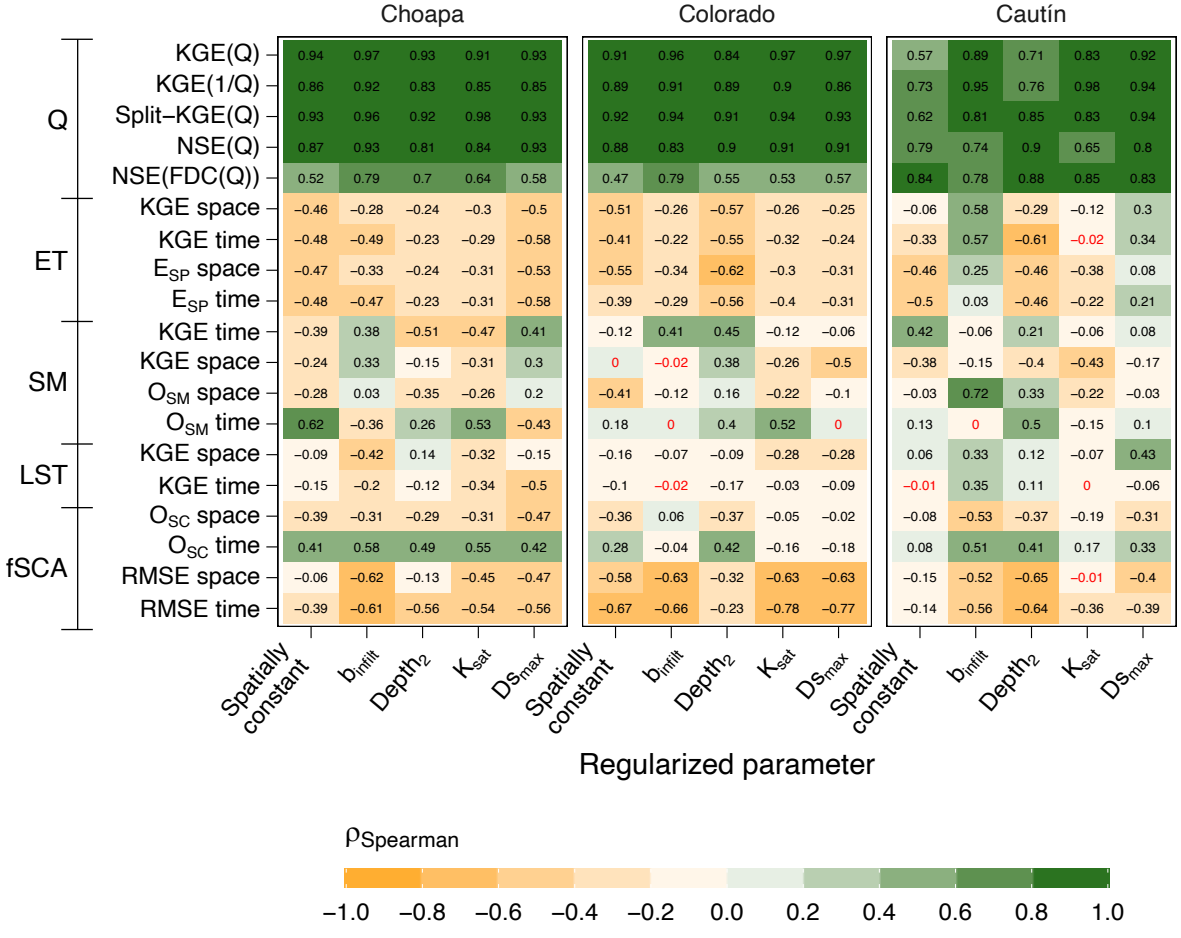

Figure 7. Spearman's rank correlation coefficient between the values of calibration metric ($OF = KGE(Q)$ & $KGE(1/Q)$) and performance measures of Q and other simulated variables during the calibration process (N = 2,000 parameter sets). Black(red) numbers represent correlations with $p_{value} < 0.05$ ($p_{value} > 0.05$). Note that lower values are better for $RSME(fSCA)$, while the Q-based metrics are positively oriented (higher, better).

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

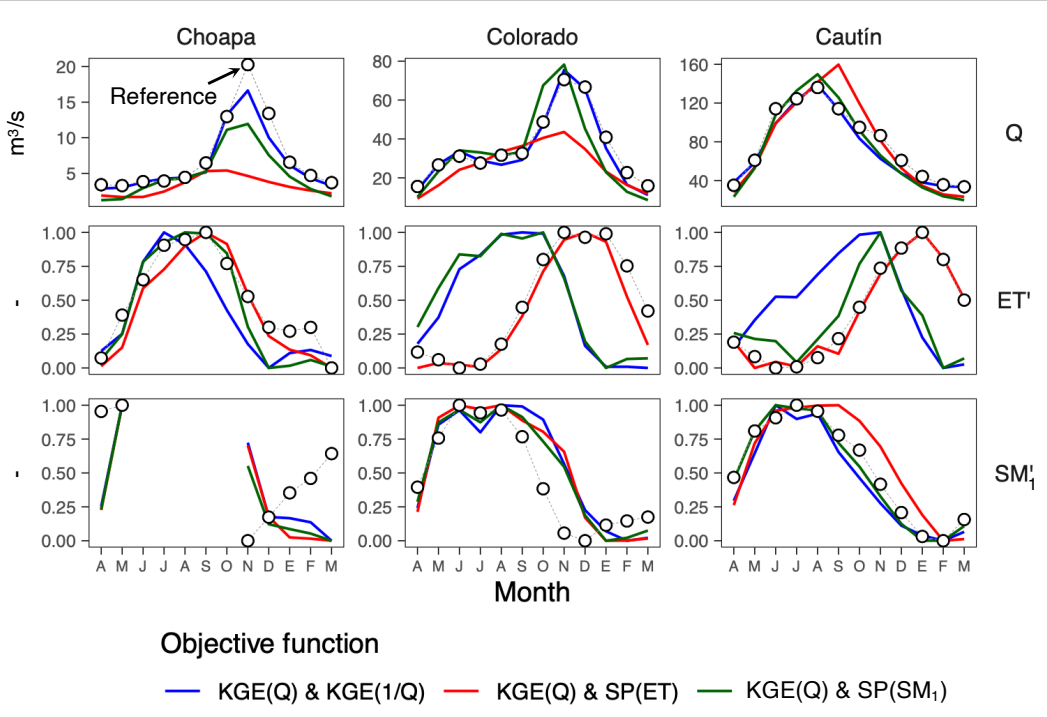

Figure 8. Catchment-scale annual cycles of streamflow (Q), normalized evapotranspiration (ET'), and normalized soil moisture (SM_1') obtained from a multivariate calibration where parameters are spatially constant. Mean monthly values are computed only if there are at least 50 days with information and consider the same days during the calibration period (2005-2018). Results correspond to the best parameter set for each objective function (colors). Notice that winter corresponds to JJA, while summer to DJF. Streamflow observations and remotely sensed variables are referred to as "reference" and symbolized with white dots. The normalization of ET and SM_1 is computed as $X' = (X - X_{min})/(X_{max} - X_{min})$.

https://doi.org/10.5194/egusphere-2026-1363
Preprint. Discussion started: 17 April 2026

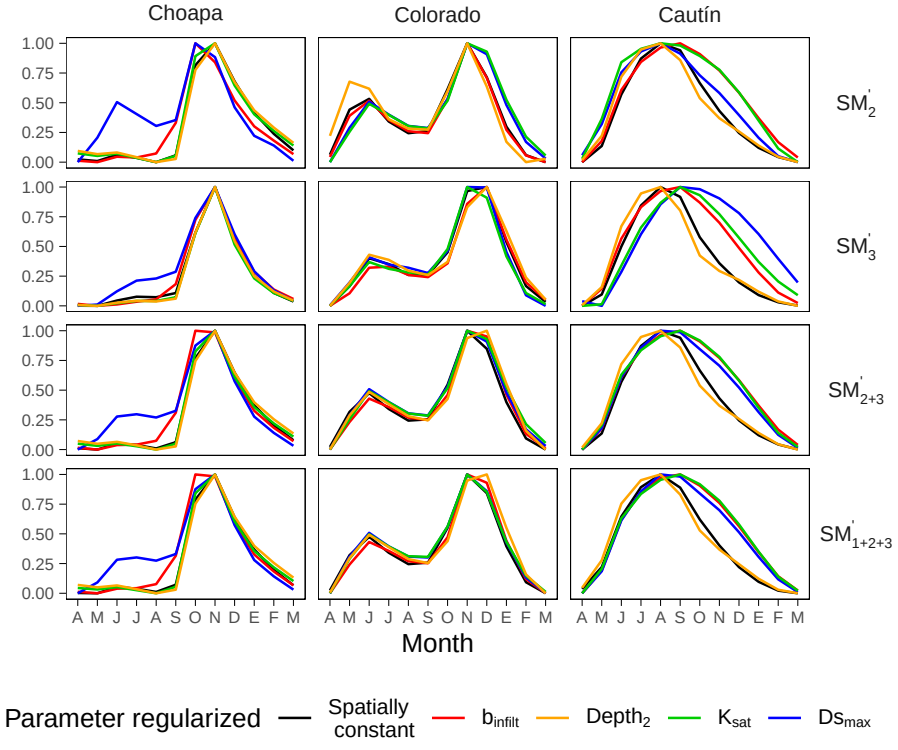

Figure 9. Catchment-scale annual cycles of normalized soil moisture (SM') at the Choapa (snow-driven), Colorado (mixed regime), and Cautín (rainfall-driven) River basins over the calibration period (April/2005-March/2018). The results are associated with the best parameter set for the objective function KGE(Q) & KGE(1/Q). Notice that winter corresponds to JJA, while summer to DJF. The normalization of SM is computed as $X' = (X - X_{min})/(X_{max} - X_{min})$.