Journal of Hydrology X 18 (2023) 100147

Contents lists available at ScienceDirect

Journal of Hydrology X

journal homepage: www.sciencedirect.com/journal/journal-of-hydrology-x

Research papers

Improvement of low flows simulation in the SASER hydrological modeling chain

Omar Cenobio-Cruz [a], Pere Quintana-Seguí [a,*], Anaïs Barella-Ortiz [a], Ane Zabaleta [b], Luis Garrote [c], Roger Clavera-Gispert [a], Florence Habets [d], Santiago Beguería [e]

[a] Observatori de l'Ebre, Universitat Ramon Llull – CSIC, Roquetes, Spain
[b] Hydro-Environmental Processes Group, Science and Technology Faculty, University of the Basque Country UPV/EHU, 48940 Leioa, Basque Country, Spain
[c] Department of Civil Engineering: Hydraulics, Energy and Environment, Universidad Politécnica de Madrid, 28040 Madrid, Spain
[d] Laboratoire de Géologie de l'Ecole Normale Supérieure, CNRS UMR 8538, Institut Pierre Simon Laplace, PSL University, Paris, France
[e] Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas (EEAD–CSIC), Zaragoza, Spain

ARTICLE INFO

Keywords:
Hydrology
Land-Surface Model
Distributed modeling
Low Flows
Parameter Regionalization
Genetic Algorithm

ABSTRACT

The physically-based, spatially-distributed hydrometeorological model SASER, which is based on the SURFEX LSM, is used to model the hydrological cycle in several domains in Spain and southern France. In this study, the modeled streamflows are validated in a domain centered on the Pyrenees mountain range and which includes all the surrounding river basins, including the Ebro and the Adour-Garonne, with a spatial resolution of 2.5 km. Low flows were found to be poorly simulated by the model. We present an improvement of the SASER modeling chain, which introduces a conceptual reservoir, to enhance the representation of the slow component (drainage) in the hydrological response. The reservoir introduces two new empirical parameters. First, the parameters of the conceptual reservoir model were determined on a catchment-by-catchment basis, calibrating against daily observed data from 53 hydrological stations representing near-natural conditions (local calibration). The results show, on the median value, an improvement (ΔKGE of 0.11) with respect to the reference simulation. Furthermore, the relative bias of two low-flow indices were calculated and reported a clear improvement. Secondly, a regionalization approach was used, which links physiographic information with reservoir parameters through linear equations. A genetic algorithm was used to optimize the equation coefficients through the median daily KGE. Cross-validation was used to test the regionalization approach. The median KGE improved from 0.60 (default simulation) to 0.67 (ΔKGE = 0.07) after regionalization and execution of the routing scheme, and 79 % of independent catchments showed improvement. The model with regionalized parameters had a performance, in KGE terms, very close to that of the model with locally calibrated parameters. The key benefit if the regionalization is that allow us to determine the new empirical parameter of the conceptual reservoir in basins where calibration is not possible (ungauged or human-influenced basins).

1. Introduction

Water is an essential, but limited and variable, resource. However, society is also exposed to hydrological hazards, such as floods and droughts, which also can be influenced by human activities (Van Loon et al., 2016). In this context, our societies have adapted to this situation by controlling these variations through the development of hydraulic infrastructure to manage these extremes. Dams, for example, allow water storage during the wet or melting (in snow-dominated areas) season to be released during the high-demand dry period.

In some regions of the world (e.g. southern Europe and West Africa), droughts have experienced a trend toward more intense and longer effects in the last decades (Seneviratne et al., 2012; Prudhomme et al., 2014; Zhao & Dai, 2015). Additionally, climate change and growing water demand exert additional stress on water resources systems (Bates et al., 2008; Wanders & Wada, 2015). This highlights the importance of water resources management in areas where water availability is already highly variable and limited, at seasonal and annual scales (e.g., the Iberian Peninsula). Furthermore, it reinforces the necessity to apply methodologies to reproduce more appropriately the hydrological

:* Corresponding author.
E-mail address: pquintana@obsebre.es (P. Quintana-Seguí).

https://doi.org/10.1016/j.hydroa.2022.100147
Received 11 July 2022; Received in revised form 21 September 2022; Accepted 7 December 2022
Available online 15 December 2022

response of the basins, especially during low flow periods, to provide decision-makers with tools to efficiently manage water resources.

Droughts and low flows are characteristics of the natural water cycle (Van Loon, 2015). The former are usually defined by a water deficit in relation to a long-term average value and depending on which variable presents a deficit, drought is categorized in different types (Mishra & Singh, 2010): meteorological drought, related to a precipitation deficit; agricultural drought, due to a soil moisture deficit; and hydrological drought, which is related to a low streamflow condition. Each of them characterized by different indices (Keyantash & Dracup, 2002; Mishra & Singh, 2010). Specifically, hydrological drought can be defined by the Standardized Flow Index (SFI, Vidal et al., 2010); however, must not to be confused with low flow conditions, which are usual during the dry season every year (Smakhtin, 2001). Therefore, hydrological drought is a more general phenomenon which is characterized by more factors than by just low flows (Van Loon, 2015). Droughts have severe impacts on water availability to sustain ecosystem and societal requirements (Sheffield et al., 2012; Stahl et al., 2016). To study drought impacts, and to improve water resource management, it is necessary to improve our knowledge of low flows. However, modeling low flows through hydrological models is still a challenge (Smakhtin, 2001; Staudinger et al., 2011).

Land-surface models (LSMs) have proven very useful for studying the hydrological cycle, including droughts (Lehner et al., 2006; Vidal et al., 2010; Prudhomme et al., 2011; Van Loon et al., 2012; Mo & Lettenmaier, 2014; Xia et al., 2014; Gaona et al., 2022) and seasonal low flows (Gudmundsson et al., 2012; Quintana-Seguí et al., 2020). Being mostly physically-based models, they help understand the underlying physical processes. However, the hydrological response in these models can be potentially improved, especially the representation of the slow component of the streamflow, which in many cases, is constrained by a limited description or even the absence of groundwater modeling (Stahl et al., 2011; Gudmundsson et al., 2012), among other processes, such as lateral subsurface flows.

Low-flow periods have not been well represented in large-scale models mainly due to a too fast response between precipitation and runoff (Van Loon et al., 2012; Barella-Ortiz & Quintana-Seguí, 2019; Quintana-Seguí et al., 2020). This represents a disadvantage in areas where streamflow is dominated by slower processes, such as groundwater discharge (i.e., from aquifers), mostly during the dry season (van Loon et al., 2012). These limitations are even more important in mountainous areas, where groundwater is poorly known (Somers and McKenzie, 2020). Therefore, it is necessary to improve the simulation of processes that influence the flow's slow component and sustain the summer flows. This improvement can be done through (i) physical (groundwater models, improved lateral flows, etc.), or by (ii) conceptual approaches.

The improvement through physical models is a complex task due to the complexity and the high uncertainties involved, also demands good knowledge of the geological structures (Habets et al., 2008; Vergnes et al., 2012). For example, Sutanudjaja et al. (2011) and Tian et al. (2012) coupled groundwater models to offline LSM models, which do not allow feedback between groundwater storage and soil moisture. On the contrary, York et al. (2002), Maxwell & Miller (2004), and Vergnes et al. (2012; 2020) used coupling schemes where this kind of feedback was considered. Another physically-based approach, instead of coupling two models, is by modification of the model itself to consider the groundwater effects, such as the work done by Miguez-Macho et al. (2007). In this case, a two-way exchange between groundwater and rivers is allowed, together with the exchange between the vadose and the saturated zones.

Conceptual approaches use simple mathematical equations to describe hydrologic processes (Liu et al., 2017), it should be stressed that implementing these is not as complex as the options described in the previous paragraph. Conceptual models depend on parameters to be calibrated, which do not correspond to a physical meaning or quantity. Several studies have tested this approach. For example, Artinyan et al. (2008) and Getirana et al. (2014) added two additional reservoirs to SURFEX LSM to evaluate the water budget. Lafaysse et al. (2011) added a reservoir to represent the effect of aquifers in mountain areas, that was extended in the plain over hard rock aquifer (Le Moigne et al., 2020) within the SIM (SAFRAN - ISBA - MODCOU) model. Gascoin et al. (2009) implemented an additional storage reservoir to consider the deep groundwater flow in France. Huang et al. (2019) added an additional layer to the DBH (Distributed Biosphere Hydrological) model to connect the soil layers with a groundwater reservoir. Guimberteau et al. (2014) compared a conceptual soil hydrology scheme against a physical approach using the ORCHIDEE (ORganising Carbon and Hydrology In Dynamic EcosystEm) model over the Amazonian basin, and reported improvements in the simulated water budget with small differences between them.

The main objective of this study is to improve low flows simulation in the hydrometeorological model SASER (SAFRAN – SURFEX – EauDyssée - RAPID) through the implementation of a well-known method: a conceptual reservoir (conceptual approach), which improves the representation of the slow component of the streamflow. To this end, to determine the values of the new empirical parameters introduced by the reservoir, we compare two methods: (i) a catchment-by-catchment calibration approach (local calibration), which can only be used in near-natural basins where observational data are available, and (ii) a regionalization approach, which links the values of the parameters to physiographic and climate variables. This second approach allows determining the parameter values of the conceptual reservoir in human-influenced basins where observed low flows represent the water management effects of dams (which are not simulated) instead of the natural processes (which are expected to be simulated).

2. Study area

Our study area is the Pyrenees and all surrounding basins that drain the Pyrenees. This covers the Ebro basin to the south, some basins that flow to the Bay of Biscay to the west, the Catalan and some Languedocian basins to the east, and the Adour-Garonne basins to the north.

The Pyrenees are located on the isthmus of the Iberian Peninsula, between the Atlantic Ocean and the Mediterranean Sea, with a length of more than 400 km in the E-W direction and a maximum width of 150 km in its central part. The topography of the study area is very heterogeneous, as it includes the Pyrenees mountain range, which reaches over 3000 m. at the highest points, and flat areas of the surrounding valleys, such as the Ebro Valley. The climate is predominantly influenced by Mediterranean features in its eastern side and Atlantic influences in its western side, with an alpine climate in the highest areas. Furthermore, the topography results in large spatial precipitation and temperature variability.

In the Adour-Garonne river basin, the southeastern part is dominated by the Mediterranean climate, whereas the western part is influenced by Atlantic Ocean conditions. Precipitation varies on average from 600 mm in the middle part of the basin to 2000 mm in the south part and Atlantic coast. The precipitation decreases with both the topography and the distance to the Atlantic. Seasonally, precipitation has two maxima, one in winter and a second in spring.

Similarly, on the spanish side of the Pyrenees precipitation decreases from west to east and from north to south. Annual precipitation varies from 100 mm in the central Ebro Valley to more than 2000 mm in the highest areas. The precipitation regime is characterized by high interannual variability (López & Justribó, 2010), particularly in Mediterranean areas, most of the annual precipitation occurs during spring and autumn. Although in some regions the maximum precipitation occurs during the cold season (in the Atlantic areas). In the Pyrenees and in the central Ebro Valley the summers are mainly dry (López-Moreno et al., 2008; 2011).

The Pyrenees, considered as natural water towers for its surrounding

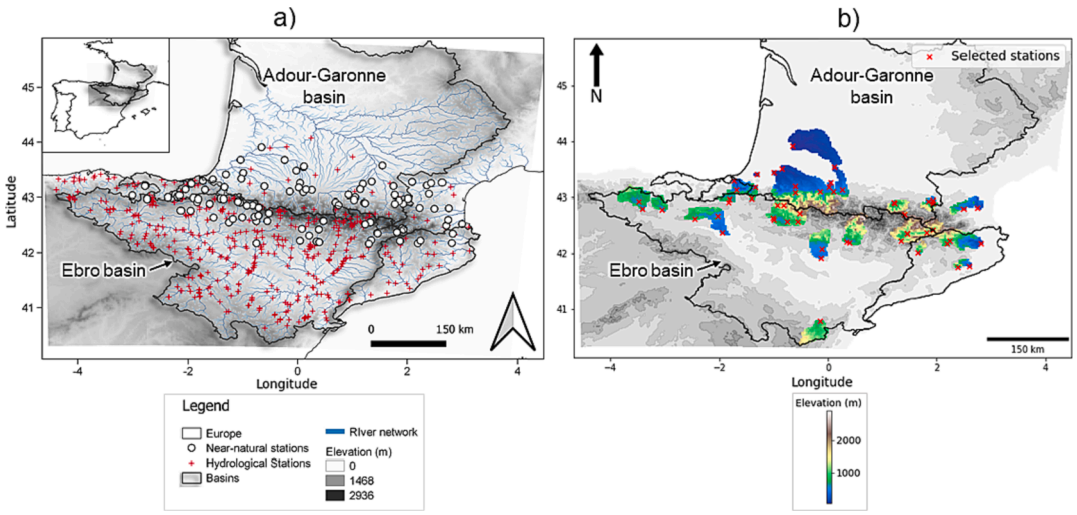

Fig. 1. (a) Study area and location of the streamflow stations; stations in black circles are defined as near-natural. The river network is depicted in blue. (b) Selected outlet gauging stations of near-natural basins.

basins, provide the water that satisfies the downstream demands for human and environmental needs (Immerzeel et al., 2019). The main water uses in the Adour-Garonne basin are agricultural and industrial. In the Ebro, water demands by the agricultural sector represent 92 % of the total water volume of the basin, this being the main water use in the basin (http://www.chebro.es/guest/uso-del-agua). This agricultural development was made possible by the construction of several dams that regulate the river flows and store water for dry periods. The large number of dams has caused major alterations in river regimes and reduced the magnitude of floods. Nevertheless, low-flows were also affected, being reduced by an order of magnitude, which is undesirable for ecological purposes (environmental flows) (Batalla et al., 2004). Additionally, the ever-expanding human activity (mainly irrigation) has increased the pressure on the water resources of the basin.

3. Data and models

In this section, we describe the observational streamflow database used it, together with the model and physiographic database.

3.1. Streamflow data

The observational streamflow database used was gathered by the EFA210/16 PIRAGUA project (Zabaleta et al., 2022). It consists of daily streamflow records from the different river basin authorities that manage the water in the study area. For our analysis, we considered a final database comprising 392 gauging stations that encompasses the temporal period between September 1979 and August 2014 (35 years), of these, 104 were selected as natural and near-natural gauging stations analyzing the data and metadata, as indicate the Fig. 1a with black circles. The selection was carried out following the criteria below:

1. Only stations with at least 20 years of data within the analysis period were considered.
2. Data gaps of less than 10 % of the total record length (these gaps were not filled) were allowed.
3. Flows must have a natural or near-natural hydrological regime, i.e. the stations must not be downstream of important human-influenced areas (e.g., reservoirs, irrigation areas).

For calibration purposes, from the selection of natural stations previously done, we selected outlet stations with series starting in 1979 and with at least 30 years of records. Fig. 1b shows the final selection, which resulted in 31 outlet stations (in total 53 sub-catchments, if we consider the nested catchments).

3.2. Description of SASER model

SASER (SAfran-Surfex-Eaudysee-Rapid) is a distributed and physically-based modeling chain consisting of a meteorological forcing, a LSM, and a routing scheme.

Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige (SAFRAN) (Durand et al., 1993) is a meteorological analysis system that produces the meteorological gridded forcing dataset using an optimal interpolation algorithm which exploits the output of a meteorological model as first guess, and in-situ observations. SAFRAN provides hourly meteorological data which is then ingested by the LSM. The SAFRAN dataset used here is PIRAGUA_atmos_analysis (Quintana-Seguí & Le Cointe, 2022), which was created within the EFA210/16 PIRAGUA project and corresponds to a union of the French (Quintana-Seguí et al., 2008; Vidal et al., 2010) and the Spanish (Quintana-Seguí et al., 2016; 2017) implementations of SAFRAN. It has a temporal resolution of one hour and a spatial resolution of 2.5 km.

The LSM used by SASER is SURFEX (Surface Externalisée, in French) (Masson et al., 2013) version 8.1. It uses the Interaction Soil-Biospère-Atmosphère (ISBA) scheme (Habets et al., 2008) to simulate natural surfaces, which describes the vertical processes in the soil column and generates a runoff and a drainage that flows through the hydrosystem using Eaudysée (Saleh et al., 2011; Vergnes & Habets, 2018) and the Routing Application for Parallel Computation of Discharge (RAPID) scheme (David et al., 2011).

We used the SURFEX simulations of the PIRAGUA_hydro_analysis dataset (Beguería et al. 2022), which were performed using the same grid as the PIRAGUA_atmos_analysis. Eaudysée and RAPID were run at a resolution of 1 km, using the HydroSheds database (Lehner & Grill, 2013) to describe the stream network.

Some limitations of the current SASER implementations are that: (i) There is no lateral flow between SURFEX grid cells; (ii) there is no bidirectional interaction between the river and the alluvial aquifer; and (iii) groundwater processes are not simulated. This translates into a fast reaction between runoff and precipitation. Therefore, a better representation of groundwater storage is necessary to account for the slow processes (groundwater and snow melting).

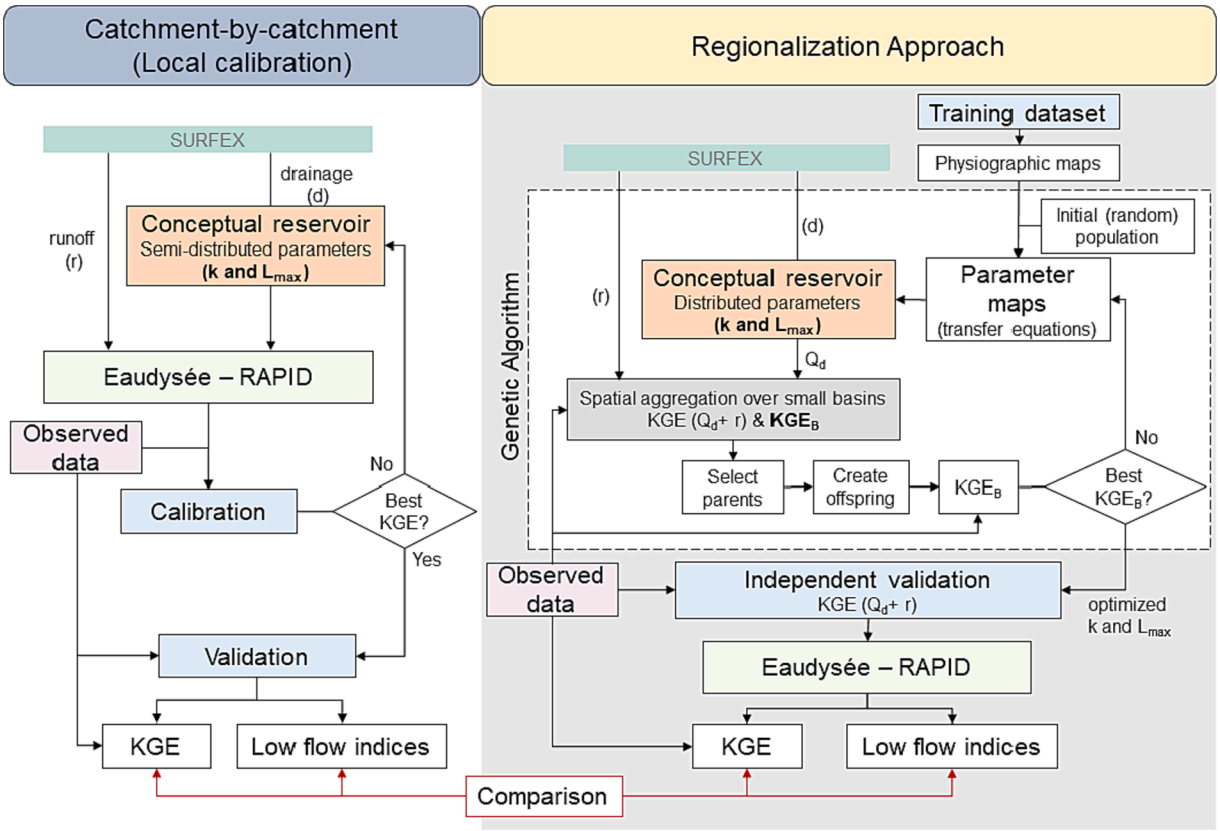

Fig. 2. General framework used in our analysis. The left panel show steps used during the local calibration and the right panel detailed the steps of regionalization approach.

3.3. Physiography

In this study, we used the SURFEX's default physiographic database ECOCLIMAP II (Faroux et al., 2013). It has a spatial resolution of 1 km and includes an ecosystem classification as well as a consistent set of land surface parameters.

In contrast to commonly used land cover products like Corine Land Cover and Global Land Cover, ECOCLIMAP-II has a new division of the existing classes with a better regional character obtained from the climatic environment (latitude, proximity to the sea, topography). The land cover parameters provided in this dataset include root depth, minimal stomatal resistance, albedo, and Leaf Area Index (LAI). The temporal variables are represented using a climatology (i.e. a mean annual cycle).

Additionally, we used the Copernicus NDVI product, which was obtained from (https://land.copernicus.eu/global/products/ndvi, last access 2022/06/07).

3.4. SIMPA model (reference)

SIMPA ("Integrated Precipitation-Streamflow Modelling System", in its original acronym in Spanish) is a conceptual and semi-distributed hydrological model developed in the Centre for Public Works Studies and Experimentation (CEDEX) in Spain (Estrela & Quintas, 1996). SIMPA simulates the natural water balance and provides information about the main hydrological variables (precipitation, evapotranspiration, streamflow) at a monthly time step. SIMPA is used by the Spanish authorities for water resources evaluation.

4. Methods

This section presents a detailed description of the methods implemented to improve low flow simulation. Fig. 2 shows the schematic flowchart with the steps of the two methodologies used, that are described below.

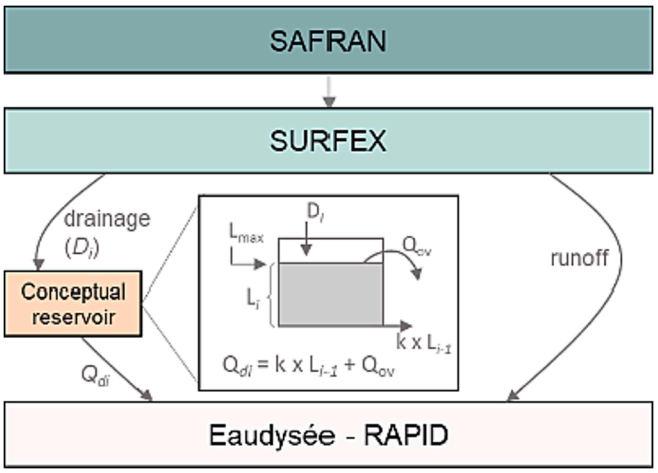

Fig. 3. Schematic diagram of the conceptual reservoir implementation, in the SASER modeling chain.

4.1. Conceptual reservoir

The implementation of a conceptual reservoir is based on the formulation from rainfall-runoff models, like ARNO (Todini, 1996) or TOPMODEL (Beven & Kirkby, 1979). The major difference being that the reservoir here presented, is implemented as an external module in the LSM model, as a postprocessing of the drainage, to account for a better representation of slow component of simulated streamflow.

We introduced a conceptual reservoir at grid point level, to improve the slow component of the streamflow. We did not modify the partitioning between surface runoff and drainage done by SURFEX LSM. The reservoir purpose is to modulate the drainage simulated by SURFEX (to sustaining it during the dry period), before being fed to the river routing component Eaudysée-RAPID (Fig. 3). Moreover, surface runoff is not modified, to avoid an excessive role of empirical parameters on the hydrological response of the model.

The input to the reservoir is the drainage generated by SURFEX. The reservoir has two outflows. The first mimics a baseflow and is the main contributing term sustaining flow during the dry period. The second occurs when the reservoir exceeds the maximum threshold, so there is no time lag when drainage is simulated during the wet season (when the reservoir is full). The total reservoir outflow is calculated following equations (1) to (3).

$$L_i = \tau(D_i + (1/\tau - k) \times L_{i-1} - Q_{ov} \quad (1)$$

$$Q_{di} = k \times L_{i-1} + Q_{ov} \quad (2)$$

And

$$Q_{ov} = max(0, L_{i-1} - L_{max})/\tau \quad (3)$$

Where D_i [mm day^{-1}] is the drainage; L_i [mm], the water content of the reservoir; k [day^{-1}] and L_{max} [mm] are empirical parameters, which correspond to the depletion coefficient and the size of the reservoir, respectively. Q_{di} [mm day^{-1}] is the total reservoir outflow, Q_{ov} [mm day^{-1}] is the reservoir outflow and τ is a constant with a value of 1 day. The subindices *i* and (*i-1*) represent the current time step and the previous step, respectively. Finally, Q_{di} is added to SURFEX's runoff and sent to Eaudysée-RAPID to compute daily streamflow.

4.2. Model performance evaluation

In different steps (calibration, validation, and regionalization), we need to evaluate the performance of the simulations. We chose the Kling-Gupta Efficiency (KGE) (Gupta et al., 2009) as our objective function:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}, \quad (4)$$

where *r* is the Pearson's correlation coefficient, α is the bias component and β represent the ratio of discharge variance:

$$\alpha = \frac{m_s}{m_o} \text{ and } \beta = \frac{\sigma_s}{\sigma_o} \quad (5)$$

m and *σ* represent the mean and standard deviation, respectively. Similarly, subscripts *s* and *o* represent simulated and observed discharge, respectively.

The KGE over untransformed discharge puts more weight on high flows (Garcia et al., 2017) and since our analysis is focused on low flows, we used a root square transformation, KGE ($Q^{1/2}$), which allows balancing the weight on low and high flow without losing the physical meaning (Santos et al., 2018).

For performance metrics is important have a benchmark to determine when the model performance is strength or not (Clark et al., 2021). The more traditional Nash-Suttclife (NSE) criterion (Nash & Sutcliffe, 1970) uses the average of the observations as a benchmark, this means that NSE > 0 if the model performed better than the benchmark. Knoben et al. (2019), demonstrated that using the same reasoning (NSE > 0) in the KGE criteria is not consistent. They showed that KGE values greater than −0.41 indicate an improvement over the mean flow benchmark. Therefore, we used a KGE value of −0.41 as the baseline value.

Only within the context of the genetic algorithm, which must evaluate the goodness of fit of each member of the population at each step, we used a transformation of the KGE:

$$KGE_B = \frac{KGE}{2 - KGE} \quad (6)$$

This transformation avoids the skewed distribution of efficiencies for large samples (Mathevet et al., 2006).

4.3. Reservoir parameter calibration procedure

The size of the reservoir (L_{max}) and the depletion coefficient (k) have to be calibrated. For calibration and validation, we use a classical split-sample procedure. Therefore, we split the entire record into two halves. The calibration period spans from 01/09/1979 to 31/08/1997 and the validation period from 01/09/1997 to 31/08/2014. Both parameters were calibrated at the sub-catchment scale on a catchment-by-catchment basis against locally observed streamflow data, which involves that all the grid points belonging to the sub-catchment had the same values of the parameters. Hereafter, we will refer to this step as local calibration.

The grid resolution and the distribution of hydrological stations allowed us to use a nested approach to calibrate the parameters where streamflow data were available. Thus the parameters were calibrated first on the upstream sub-catchments and then progressively towards the outlet.

We set the valid range for both parameters using similar criteria defined by Artinyan et al. (2008). The accumulated streamflow of the dry period from July to September should be close to the average reservoir level (L_{max} parameter). For each sub-catchment we obtained the total runoff volume for the dry period of the driest year of the calibration period (Q_{dry}, in mm), and L_{max} was estimated as $12Q_{dry} \geq L_{max} \geq Q_{dry}$. According to Artinyan et al. (2008), twelve times the volume of the driest months seems a reasonable upper bound for this parameter. The limits of the depletion coefficient (k) were calculated considering the length of the dry period, thus the reservoir has drainage releases during the length of the dry period.

We create a parameter space, for simulations purposes, where the range between the extreme values of each parameter was discretized into 12 values. A total of 144 simulations were carried out and the performance of each simulation was evaluated for each sub-catchment using the KGE ($Q^{1/2}$) (see section 4.2). The best simulation for each sub-catchment was chosen, and the parameter set associated with each of them was saved.

4.4. Regionalization of the parameters

When observational data are not available, or they do not have good quality, or they include processes that are not simulated, such as human processes (dams and irrigation), local calibration is unfeasible. To overcome this limitation, we used the regionalization approach presented by Beck et al. (2020), which allows setting the values of the reservoir parameters all over the domain, going beyond the near-natural basins used in the calibration procedure.

4.4.1. Predictors used for the regionalization

For the regionalization approach, we selected eight physiographic variables as predictors:

- Three of them were related to climate: ARI, aridity index; MAP, mean annual precipitation; and PET, potential evaporation. Beck et al. (2016), Nijssen et al. (2001), Singh et al. (2014), and Troch et al. (2013) demonstrated that these variables exert an important influence on the flow response in regionalization studies at global scale. The MAP predictor was transformed to square root to better fit a normal distribution.
- The NDVI, Normalized Difference Vegetation Index; and SNW, fraction of snow with respect to the total precipitation, are predictors related to land cover. The NDVI was added because the vegetation

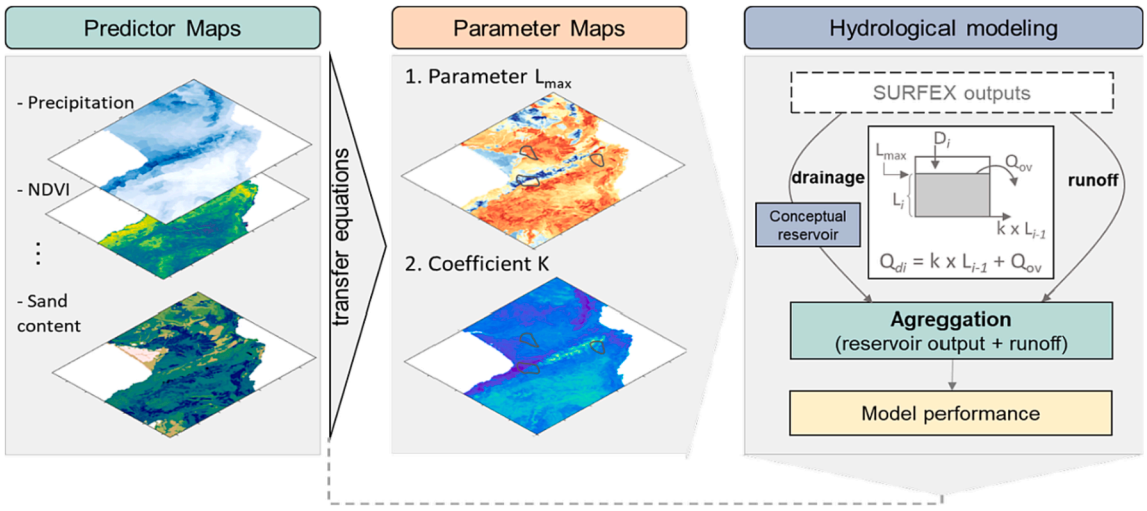

Fig. 4. Scheme of the main steps of the regionalization approach. For each time that the algorithm is run, the steps (parameter maps and hydrological modeling) are repeated in each iteration for the optimization process.

influences in the evaporation, infiltration and hydrological function of the soil, which may also affect the slow component (low flows) and runoff-rainfall conversion processes (Zhang et al., 2001; Donohue et al., 2007; Peel, 2009). The snow affects the land cover and has an important role in streamflow generation in mountain regions, and also on the slow component of the flow.

- Finally, SL, the slope; SND, soil sand content; and CLY, soil clay content, are variables related to the topography and the soil. The slope predictor was included due to the good general correlation between surface slope and soil depth (Tesfa et al., 2009), and the soil texture has a strong influence on all soil related processes, including subsurface runoff. (Price, 2011).

Most of these variables were obtained from the ECOCLIMAP II database. Although these descriptors are not directly associated with groundwater, they help determine a landscape that can be prone or not to groundwater.

4.4.2. Regionalization approach

The regionalization approach uses a genetic algorithm to optimize the coefficients of the transfer equations. These equations link the reservoir parameters (predictands) to the physiographic variables (predictors). We use the same near-natural basis that in the local calibration (all of them have an area smaller than 5,000 km^2). For the optimization, the reservoir scheme was run at daily time step and at same SURFEX spatial resolution, for the whole period (1979–2014). The SURFEX runoff and reservoir output were spatially aggregated and compared with observed runoff. This comparison is possible due to the size of the catchments that allows us to discard channel routing effects (Gericke & Smithers, 2014). The methodology is shown in Fig. 2 (right panel), and a detailed description is presented below.

Since the reservoir is implemented as an external module to post-process the SURFEX drainage before the routing step, there is no need to run SURFEX within the genetic algorithm (see Fig. 3). This is a key benefit, as SURFEX is a computationally-expensive model. Furthermore, given the size of the catchments there is no need to run the routing scheme (Eaudyssee-RAPID) at each iteration (see Fig. 4). This allows us to undertake a very high number of simulations, as we only run the very simple reservoir model, which consists of a few lines of Python code.

The transfer equations that link the predictors with the predictands are expressed as follows:

$$MP_i = w_{i1}ARI + w_{i2}MAP + w_{i3}PET + w_{i4}NDVI + w_{i5}SNW + w_{i6}SL + w_{i7}SND + w_{i8}CLY + w_{i9} \quad (7)$$

where MP_i are the model parameters (k and L_{max}) and w_i are the coefficients that will be optimized. The eight predictors chosen are (see the previous section):

- *ARI*, aridity index (P/PET);
- *MAP*, mean annual precipitation (root square transformed);
- PET, mean annual potential evaporation;
- NDVI, Mean Normalized Difference Vegetation Index;
- SNW, the fraction of snow with respect to the total precipitation;
- SL, topographic slope;
- SND, soil sand content; and
- CLY, soil clay content.

First, each predictor was interpolated to the same grid as the model uses (2.5 km of resolution). Next, as Beck et al. (2020) did, predictor values were clipped using the 99th and 1st percentiles of the area covered by the sub-catchments. Finally, to make the predictors comparable to each other, they were standardized by subtracting the mean and dividing the results by the standard deviation of the area covered by the sub-catchments.

The coefficients of the transfer equations ($9 \times 2 = 18$ w_i) were optimized using our own implementation of the ($\mu + \lambda$) genetic algorithm (Slowik & Kwasnicka, 2020). The ($\mu + \lambda$) algorithm indicates that selected parents and children together comprise the new population (offspring) for the next iteration. The algorithm starts with a random population (λ) of 32 members, all of them are evaluated using the performance score (KGE_B), and the best three members are selected and saved for the next iteration (μ, size of the parents in the population). Subsequently, through the mutation operator the offspring is created (λ/μ children are randomly created from each selected parent) following a normal distribution, within the search space. A maximum number of 100 iterations was set.

Fig. 4 represents the main steps in the optimization process. First, predictor maps are obtained depending on the area covered by the catchments used. Second, the parameter maps are calculated using transfer equations, to subsequently apply the reservoir scheme. This scheme is run using as input the daily drainage simulated by SURFEX. The reservoir output (modulated drainage) and the SURFEX runoff are spatially aggregated (all cells that comprise each catchment) and compared to observed streamflow of each catchment using the KGE score.

To avoid overfitting, and to have an indicator of uncertainty in the parameter sets, cross-validation was used. For this, the catchment set

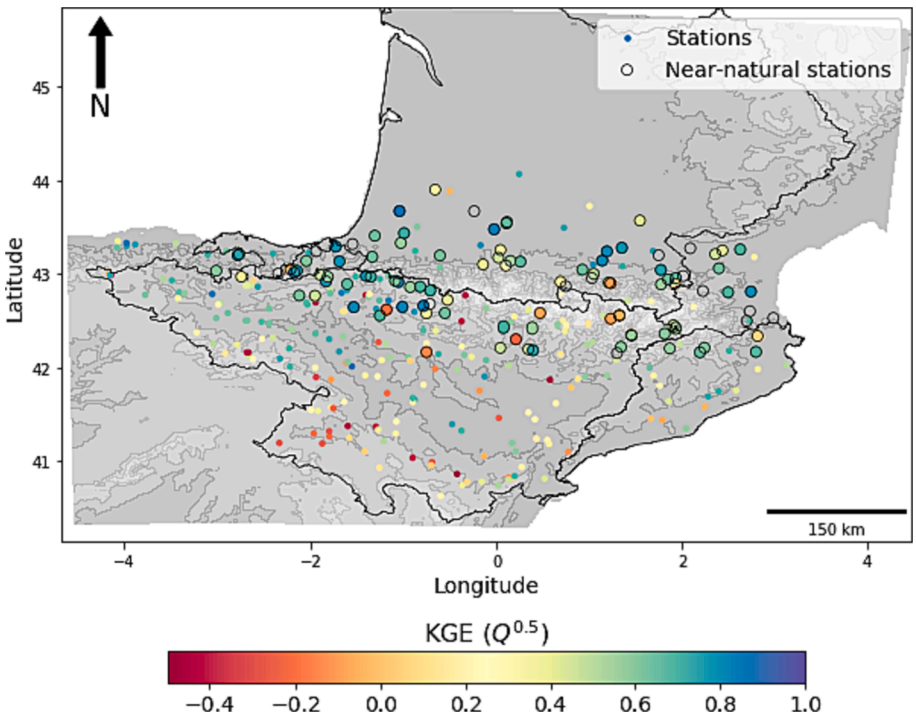

Fig. 5. KGE($Q^{1/2}$) scores between observed data and the default simulation, for the whole period (1979–2014). Larger circles with a black border indicate stations defined as natural or near-natural. The other stations are considered as influenced.

was subdivided into training (87 %) and validation (13 %) subsets. This selection was performed randomly, the catchments used for validation were used only once in each iteration, performing different experiments until all of them were used in the validation.

In total, 8 experiments were run. Each experiment consists of a training subset (27 catchments) and 4 randomly selected catchments for independent validation. In this sense, the validation subset in each experiment is different, because each catchment was used only once.

4.5. *Low flow indices*

In addition to calculating the KGE scores, and to evaluate how the conceptual reservoir implementation (local calibration and regionalization approach) impacted the simulated low flows, we calculate two low flow indices.

Different low flow indices can be estimated (Gustard et al., 1992; Smakhtin, 2001), but we focused on two common ones. First, the ratio (Q_{90}/Q_{50}), where Q_{50} is the median and the Q_{90} is the low value that is

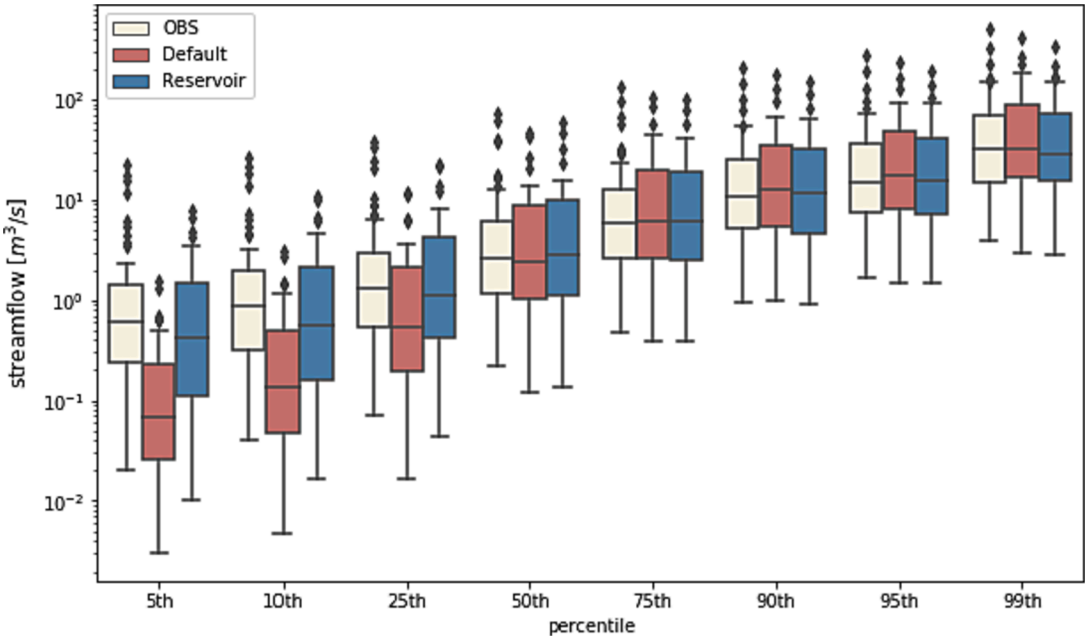

Fig. 6. Comparison of different daily runoff percentiles for observed data (OBS) and simulated streamflow (default simulation in red and simulation with reservoir scheme in blue). The horizontal line in each box represents the median, the box represents the interquartile range. The whiskers extend a maximum of 1.5 times the interquartile range.

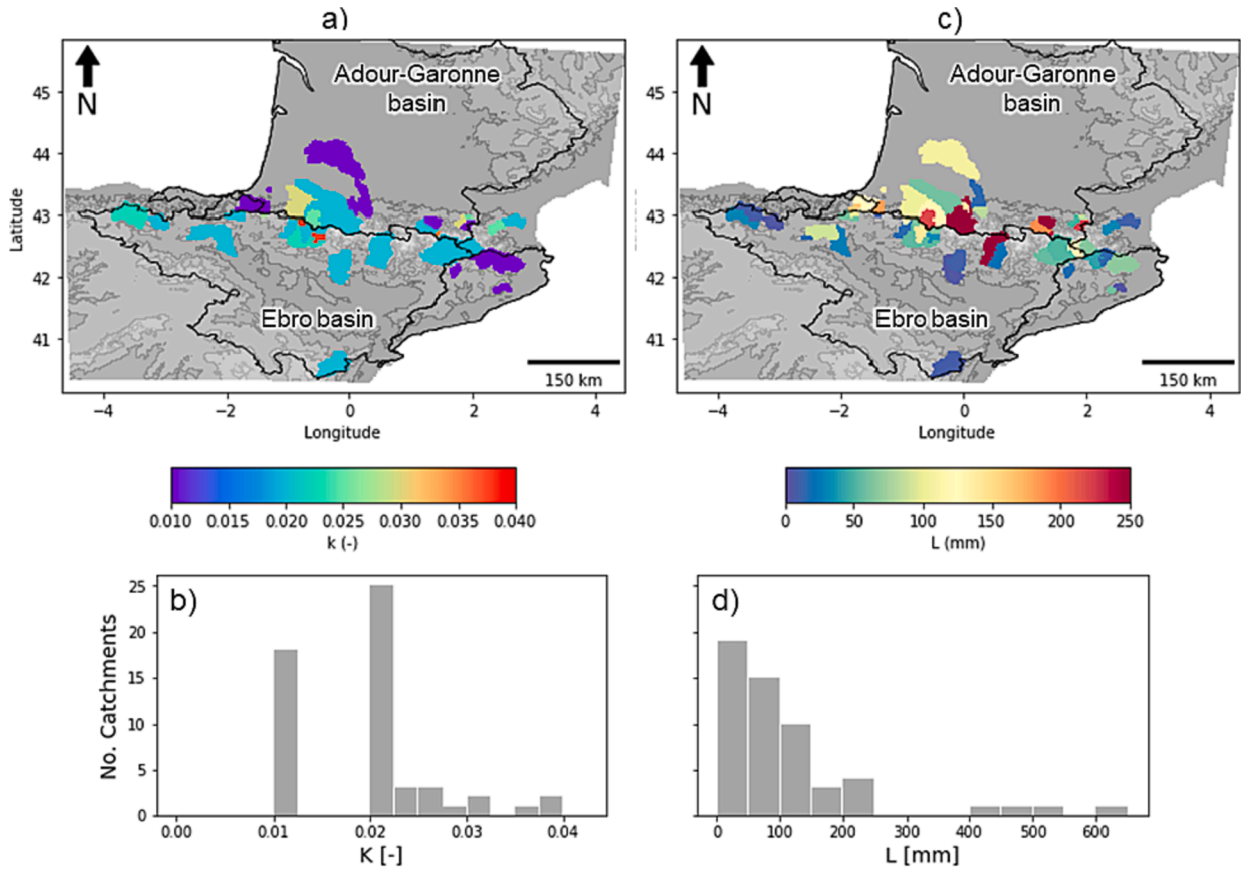

Fig. 7. Distribution of the 53 catchments used in the local calibration (catchment by catchment). Panels (a) and (c) show each catchment's calibrated values of the k [-] and L_{max} [mm] parameters, respectively. Lower panels, (b) and (c), show the histograms of each parameter.

observed 1/10th of the time, at daily time step. This index is interpreted as the proportion of streamflow originating from groundwater stores, excluding the effects of the catchment area (Smakhtin, 2001). Second, the annual minimum monthly flow with a return period of 5 years, QMNA(5), which is widely used in France for water management issues and provides information about low flow severity. To compute the QMNA(5), we calculated the 5-year return period for each sub-catchment fitting it to a log-normal distribution for low flows (Catalogne, 2012), based on a series with at least 30 years of record.

We derived the low flows indices for the summer period (July to September) for each station and computed both using the observations and the simulations. Then, we calculated the relative bias, to evaluate and compare the performance of the simulations.

5. Results

5.1. Model performance of default SASER simulation

To evaluate the performance of the SASER model (default simulation) we calculated the KGE($Q^{1/2}$) and used the 104 stations defined as natural or near-natural, which are indicated by circles with black border in Fig. 5.

Fig. 5 shows the calculated KGE scores for the entire period (1979–2014). Stations with the highest KGE values are located mainly in the Pyrenean region, where direct human influence is low. However, most of the stations in the Ebro river basin (a highly influenced basin) showed lower KGE values, especially in areas downstream of the reservoirs. This was expected, since the model must perform poorly over human influenced areas, as it only simulates natural processes.

If we focus on the near-natural basins (Fig. 1a), we see that 11 of 104 stations (uncolored circles in Fig. 4) show KGE values below the baseline value of −0.41. These lower values belong to stations located at a closer junction of small tributaries with the main channel (same SASER pixel) and due to resolution constraints (2.5 km), SASER takes the simulated streamflow to the main channel and not from the tributary. For the rest of the stations, 17 % and 42 % showed values greater than 0.7 and 0.5 respectively.

In Fig. 6, we compare the observed (OBS, in beige) and simulated (Default, in red) statistics of the streamflow for the different percentiles (near-natural basins) and the 53 catchments where the reservoir was calibrated. The SASER model (Default) simulates reasonably well high and median daily streamflow, represented by the 90th to 99th percentiles, and 50th to 75th percentiles respectively. However, low flows (25th percentile and below) are underestimated, as depicted in the red boxes in Fig. 6. For example, the median relative bias between observed discharge and default simulation for the 25th percentile is −66 %, whereas for the 95th percentile is 12 %. This shows that low flows are poorly simulated by the default SASER model.

5.2. Evaluation of the model including the reservoir with calibrated parameters

Fig. 7 shows the result of the calibration of the two model parameters L_{max} and k in the near-natural basins. They were calibrated empirically by running the reservoir scheme. The values of the L_{max} parameter ranged from a minimum of 4 mm to a maximum of 600 mm, with an average value of 108 mm, while the values of parameter k ranged from 0.01 to 0.04, with a median value of 0.018. At the end of the calibration process, we observed the following results: most Ebro catchments have the same value of parameter k, 0.02; whereas the parameter L_{max} varies over a wider range of values. These parameter values are explained by basin characteristics, such as soil type, topography, etc. For example,

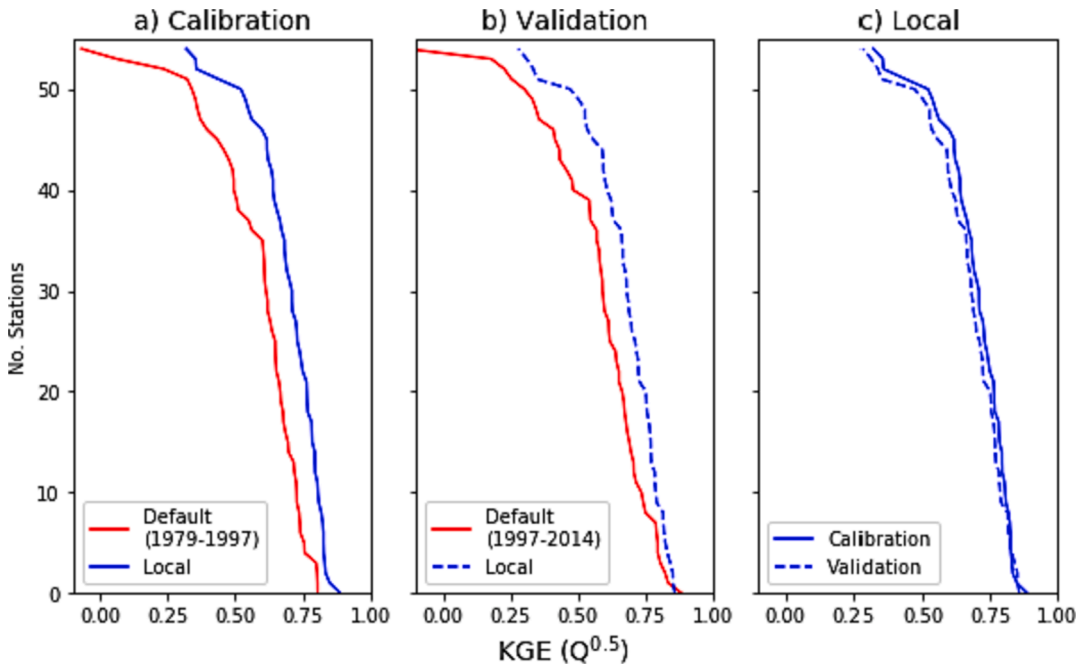

Fig. 8. Accumulated distribution of KGE scores: a) for the calibration period (1979–1997), b) for the validation period (1997–2014), and c) the comparison between KGE obtained in both periods for local calibration.

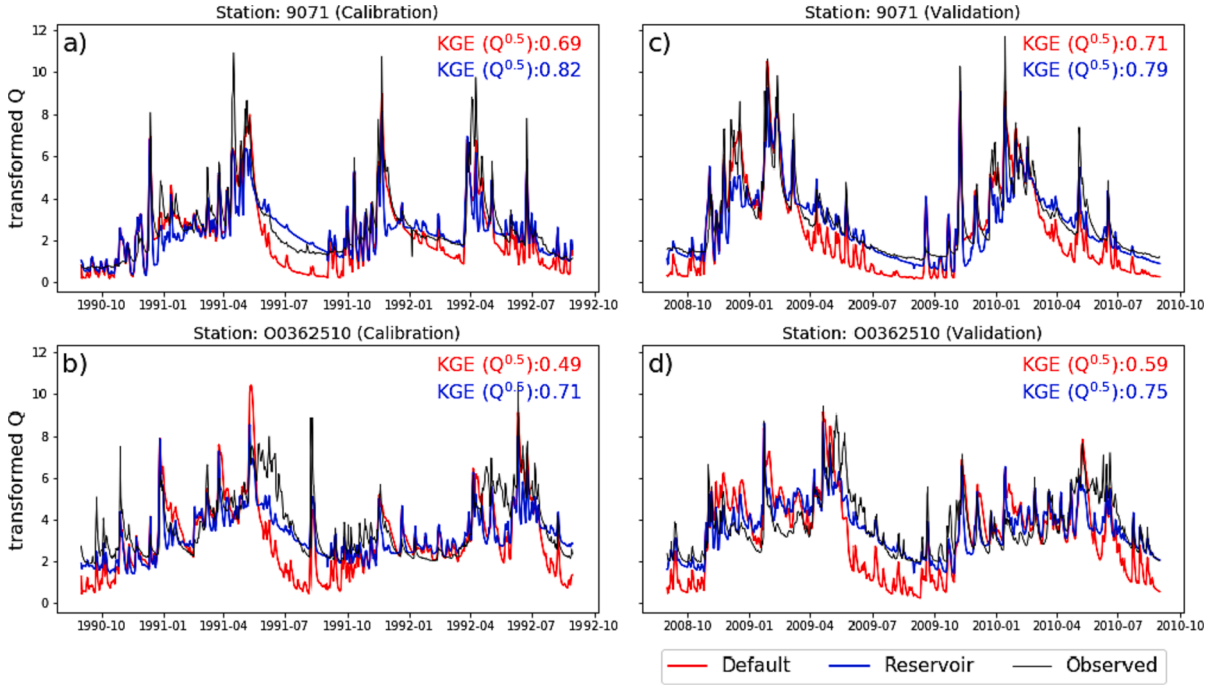

Fig. 9. Daily time series comparison of two stations, for two hydrological years of the calibration (a and b) and validation (c and d) periods, respectively. Observed streamflow (black line), simulated streamflow without reservoir scheme (red line), and simulated streamflow adding a conceptual reservoir (blue line). KGE($Q^{1/2}$) is calculated over each complete period (1979–1997 to calibration and 1997–2014 to validation, respectively).

higher L_{max} values correspond to sub-catchments with high permeability (e.g. south of France); detailed information is presented in discussion section.

Fig. 8 summarizes the resulting KGE scores and shows that adding a calibrated conceptual reservoir improves them over all catchments. On the one hand, Panel (a) shows that for the calibration period, the median KGE value of the default simulation was 0.63 and using the reservoir scheme it was 0.72, which represents a clear improvement with respect to the default simulation. For the validation period (Panel b), the default simulation has a score of 0.60 and the calibrated one of 0.71, which is a similar improvement than that from the calibration period. Panel (c) compares the difference between the scores of the calibrated model with the reservoir obtained during the calibration and validation periods. The two lines are very close, which is an indicator of robustness, the model performs similarly inside and outside the calibration period.

The resulting time series (Fig. 9) shows graphically that the

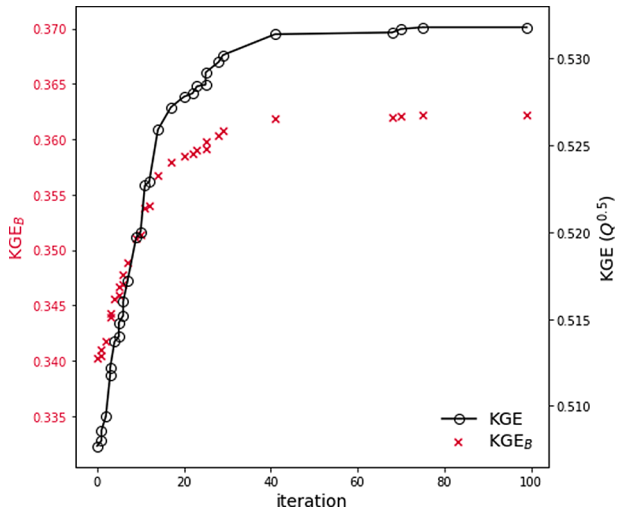

Fig. 10. Evolution of performance metric (median KGE($Q^{1/2}$) and KGE_B) in each iteration for one experiment.

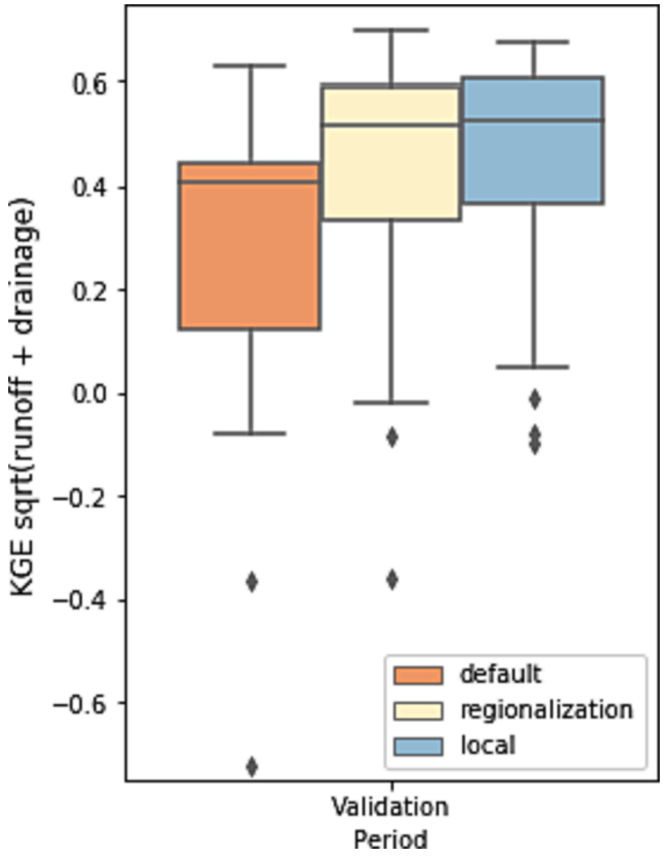

Fig. 11. Box plot of KGE scores for the validation period using the local calibration (adding conceptual reservoir), the regionalization (cross-validation) and default simulation (without conceptual reservoir).

calibrated reservoir produces time series closer to the observations. The KGE values where higher for the simulation using the reservoir scheme, for both periods. Panels a) and b) of Fig. 9 show the time series of the daily streamflow at two stations (only two years of the calibration period are depicted in order to make the plots easy to read) and panels c) and d) show the same stations for the validation period. The improved model is able to sustain the summer flows much better than the default model, without affecting the median and high flows.

The blue boxes in Fig. 6, clearly show that the lower percentiles of the simulation with the calibrated reservoir are improved compared to the default simulation. For the three percentiles related to low flows (the 5th, 10th and 25th), the values of median relative bias for default simulation were −90 %, −84 % and −66 % respectively. Whereas, the values obtained from the simulation with reservoir scheme were −37 %, −33 % and −17 % for the same percentiles. Even though the values remain slightly underestimated, it represents an improvement in low flow simulation.

5.3. *Evaluation of the regionalization approach*

As an example, Fig. 10 shows the evolution of the KGE and the KGE_B for one of the experiments. We recall that the KGE and KGE_B are related through equation (6). Each point represents the model performance after each iteration of the genetic algorithm while it searches for the best solution. For this experiment, during the first iterations, the KGE rapidly increases, and after iteration 30 the slope of the curve changes drastically to be almost flat, which indicates that the algorithm is close to a maximum (in this case, the median KGE is slightly higher than 0.53) as the new mutations do not improve the result and are not selected. Similar behavior was founded in all the experiments.

Fig. 11 shows the boxplots of the resulting KGE for the independent catchments of each experiment and compares it with the KGE obtained from the default and local calibration experiments. The median daily KGE value of the default simulation was 0.41 and the median daily KGE value of the local calibration was 0.53, while for the regionalization it was 0.52, its value is close to the obtained by the local calibration. Furthermore, we obtained improvements in 79 % of the validation catchments. This confirms the robustness of the regionalization approach to improve low flows in the catchments that were not used in the training dataset, also the KGE scores are as good as using the local calibration.

Fig. 12 shows the distribution of the KGE values for the different experiments. Starting from the left the graph shows: (i) the default simulation (brown), (ii) the local calibration (dark blue), which represents a clear improvement, (iii) the eight different regionalization experiments (light green), and (iv) three additional experiments (light brown). The eight regionalization experiments show median KGE values very close to each other, around 0.51. In Fig. 12, the 75th percentile (top edge of each box) of all experiments shows a similar value (close to 0.6), while the minimum values (lower whiskers) show greater variability, suggesting that catchments for which the KGE score was poor are more sensitive.

The three additional experiments (last three boxes in Fig. 12) where performed to see how the eight experiments can be combined to obtain a final result. In the first additional experiment, we calculate the median time series of the eight previous experiments. In the second, we run the reservoir scheme with the median value of the k and L_{max} parameter maps (median value of the parameters for each grid point) and calculate streamflow. In the last experiment, we calculated the median value of each parameter (57 mm and 0.016 for L_{max} and k, respectively) and run the reservoir scheme with these homogeneous values over the entire domain. Surprisingly, these three extra experiments showed a similar KGE median values when we compare to previous experiments, however, the last experiment (fixed median values) showed a slightly larger spread among KGE values. In any case, in the local calibration or regionalization approach, KGE values are not below −0.41, the benchmark defined previously.

An example of one year of observed and simulated (by aggregating runoff and drainage) daily streamflow for one catchment is given in Fig. 13. In this plot, we computed the median flow of the eight regionalization experiments (dark blue line) and we also plotted, in light blue, the confidence interval of the simulations.

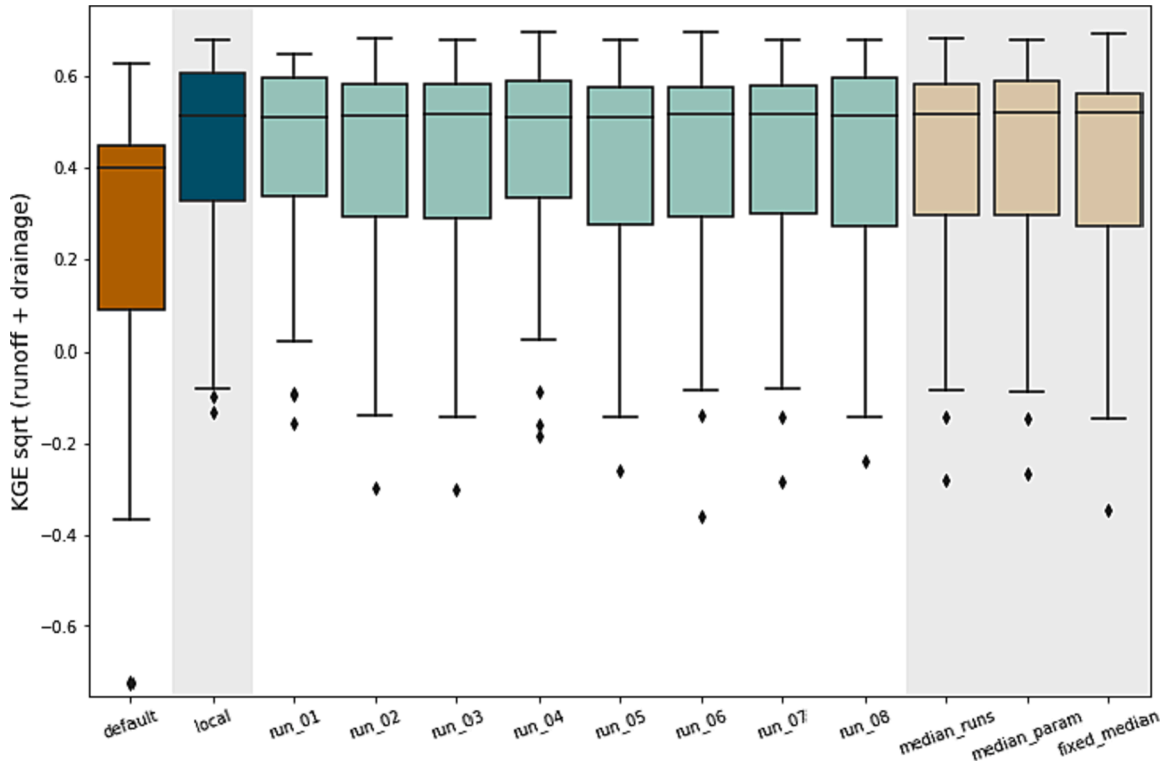

Fig. 12. Box plot of KGE for streamflow of the local calibration, using the genetic algorithm (run 01 to 08), the median of simulated streamflow (median runs), the median of eight parameter maps (median params), and simulation using fixed parameter values over the full domain. Streamflow validation period from 1997 to 2014.

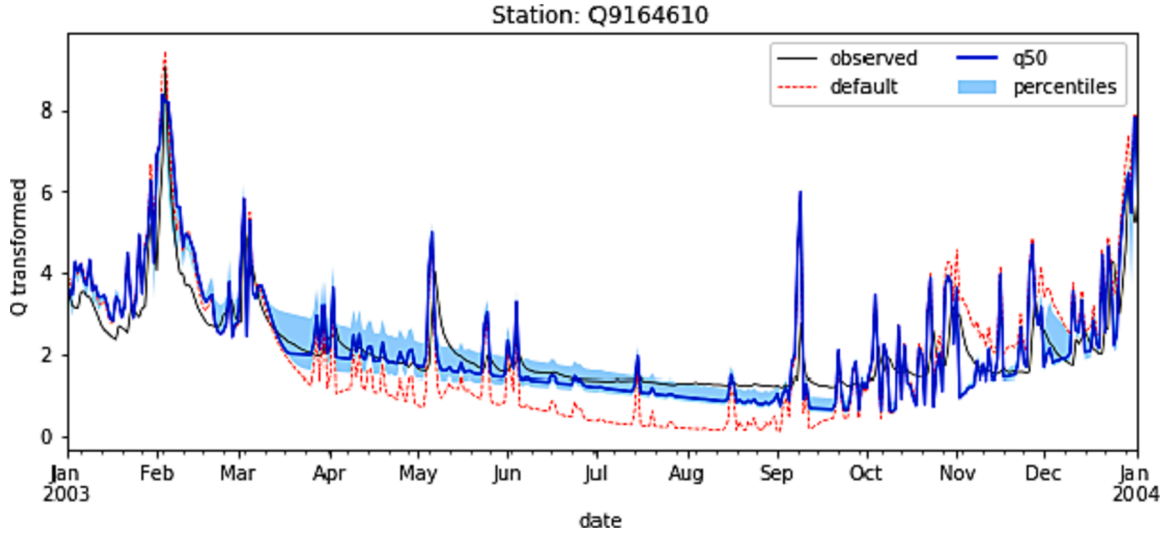

Fig. 13. Comparison of the observed (black line), simulated by SASER without adding the reservoir (red dashed line), and the median values of the eight simulations (blue line) of daily streamflow for a catchment; shaded area represent the percentile 90th and 5th of the eight simulations. Y-axis is the discharge root square.

5.4. Computing natural streamflow with the improved model.

Once the maps of the two parameters of the reservoir model were obtained using the median value of the parameter maps of the eight experiments (Fig. A1), we proceed to calculate the final streamflow using the river routing scheme.

Fig. 14 presents the streamflow simulation performance obtained using the median values of the eight regionalization experiments for the entire record (1979–2014) over the complete database. Of the total stations, 47 % showed higher KGE values than 0.5 (5 % more compared to the default simulation). Fig. 14(b) shows the improvement, in terms of ΔKGE, due to regionalization for the whole period.

For comparison purposes, we calculated the KGE values over the same periods that were used for local the local calibration (1979–1997) and validation (1997–2014). For the first period (calibration) the regionalization approach obtains a median KGE of 0.69 and for the second one (validation) it was 0.67, which is very close. Fig. 15 summarizes the performance of local calibration and regionalization approach over the calibration (solid lines) and validation (validation lines) periods on the near-natural basins. Both approaches showed an

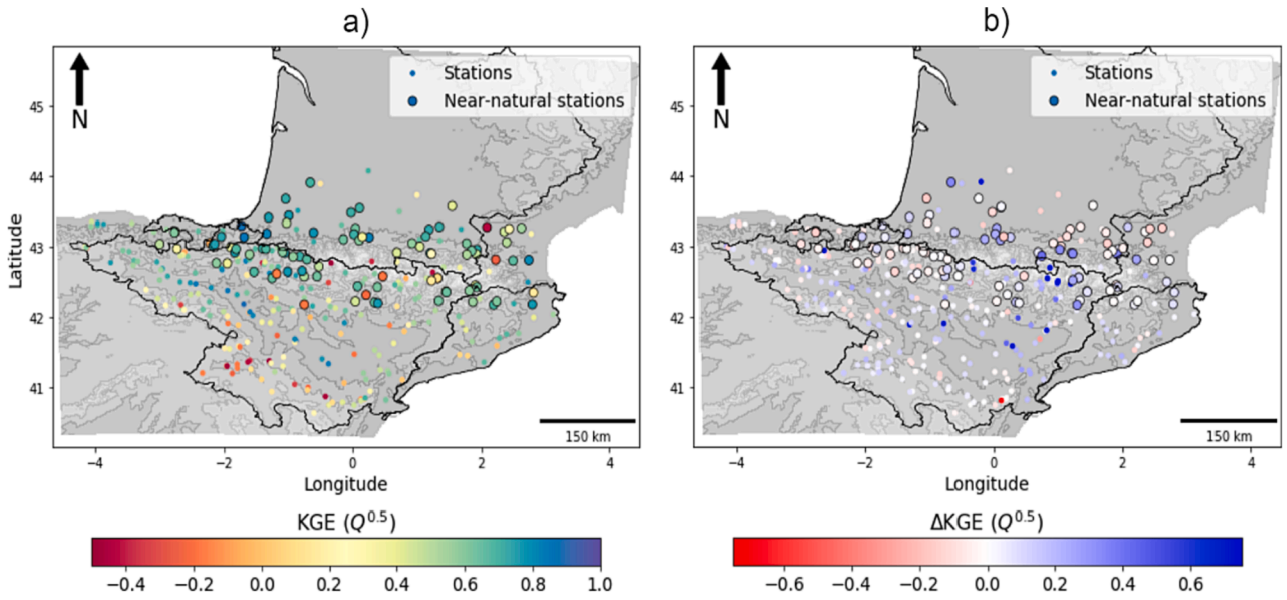

Fig. 14. (a) KGE scores obtained to calculate the streamflow with the routing scheme using the regionalization approach against the observed streamflow. (b) Improvement in KGE after regionalization (difference between Fig. 4 and Fig. 14(a)).

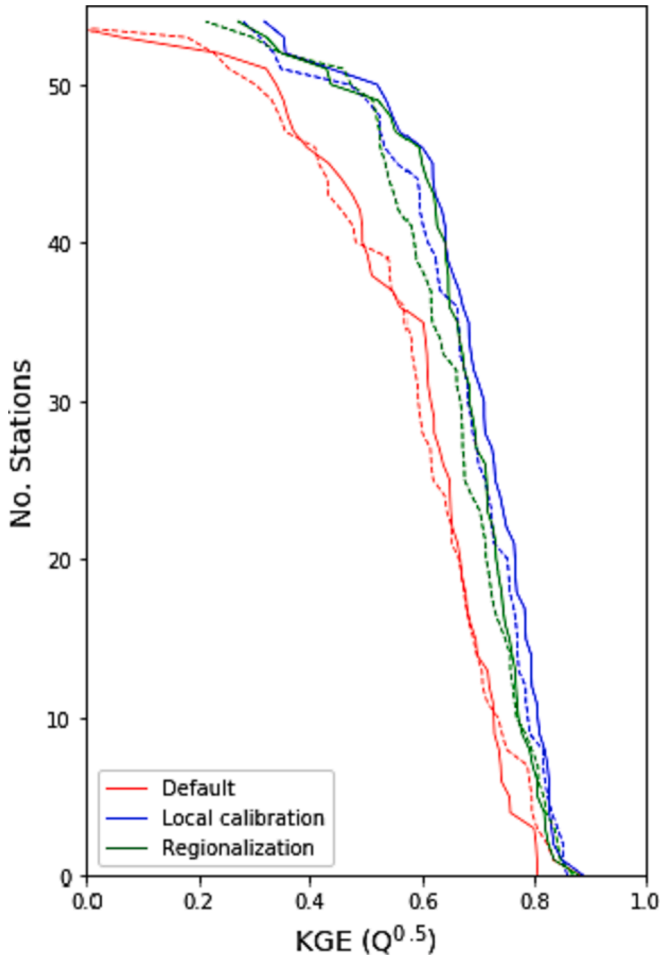

Fig. 15. Same as Fig. 7, comparing KGE from local calibration (blue) and KGE by the regionalization (green) against to default simulation (red) for calibration (solid line) and validation (dashed line) period.

improvement compared to the default simulation (without the reservoir). The best KGE values are provided by the local calibration, which we expected, even though the regionalization approach showed very close values to local calibration.

The advantage of the regionalization approach is that it allows us to estimate the values of the parameters anywhere in the area of study, thus we can also calculate improved natural flows on influenced basins, where local calibration would be impossible.

5.4.1. *Comparison with a reference model*

As the observations are affected by water management, we compare our results with an independent model. In this case, we use the SIMPA model, which is used by water managers in the Ebro river basin.

We extracted the time series simulated by the default and the improved model on the outlet of the Ebro River basin, located in Tortosa. We computed the mean annual cycle of streamflow, for the period 1980 to 2006, at the monthly time step, and plotted it (Fig. 16) together with the observations (in black) and the SIMPA model (in green).

The default simulation, in red, presents the lowest flows during the summer (from July to September), as expected, whereas the peak flow occurs in January. In general, the hydrograph of the default simulation has a shape similar to that of the observations, but with a positive bias. The positive bias is expected, as in the real basin there is more evapotranspiration, due to irrigation, than in the naturalized one. In contrast, the SIMPA model (gray line) shows two peaks, one in December and the other in April, and the minimum streamflow corresponds to the summer months, but with values twice higher than the observations. The flows obtained with the improved model, setting the parameters with the regionalization approach (green line), show a clear increase of low flows, but they are still very low compared to SIMPA. It is notable that the improved model reduces the January peak, but increases the streamflow of the following months (February and March). The new hydrograph presents a double peak, which is what we expected and what can be seen in SIMPA. It is notable how the reservoir modifies the streamflow of the winter months. This could not be seen in the daily hydrographs we showed in Fig. 13, but this arises clearly in the monthly annual cycle of the outlet, which aggregates the changes throughout the basin. Nevertheless, is difficult to determine which model is closer to reality, because the observations represent the real basin, not the naturalized one, as the models do.

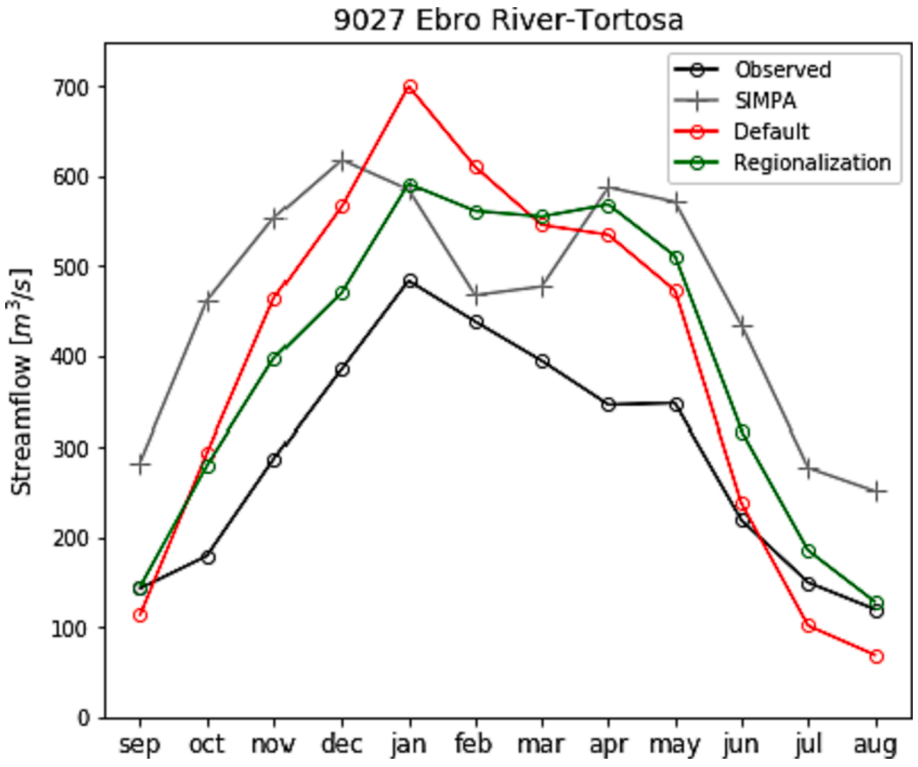

Fig. 16. Mean monthly streamflow time series at Tortosa station (main Ebro river) from observations and different models.

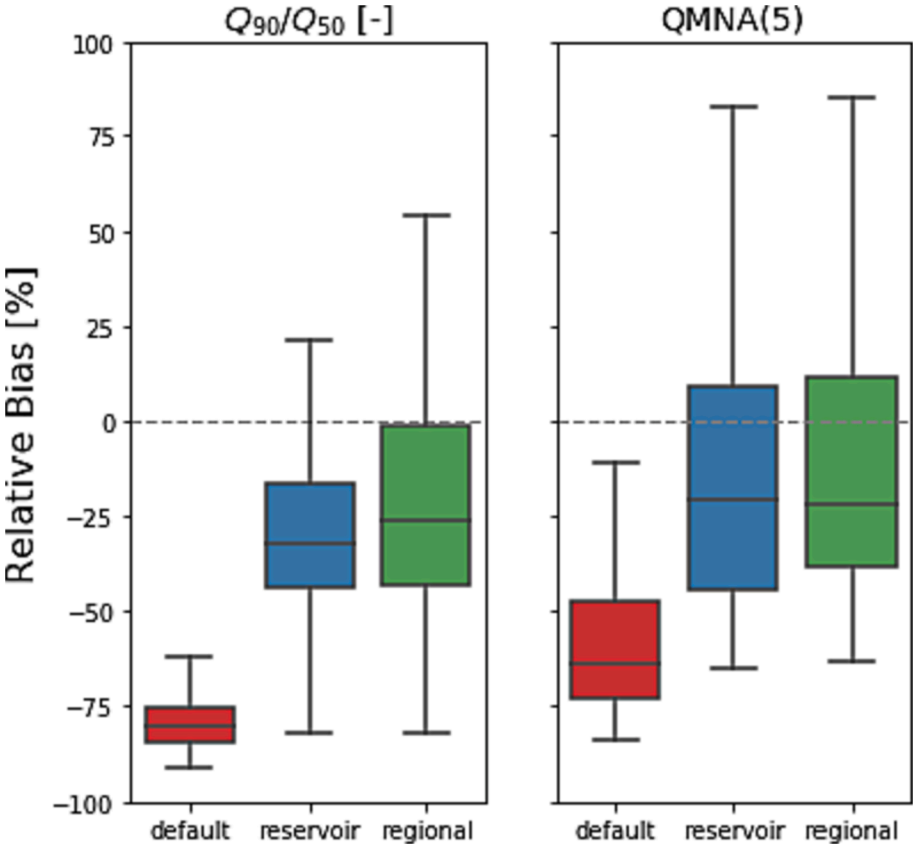

Fig. 17. Comparison of relative bias [%] for low flow indices for the default simulation and for the two approaches. The left panel shows the ratio Q_{90}/Q_{50} and the right panel the QMNA(5). The line in each box represents the median and the box the interquartile range. The whiskers extend a maximum of 1.5 times the interquartile range.

5.5. Low flow indices evaluation

In Fig. 17, we can see the improvements in the low flow indices. For the Q_{90}/Q_{50} ratio (left panel of Fig. 17), the default simulation (without reservoir) showed a median relative bias of −80 % with respect to observations, whereas the median relative bias of the simulation adding a reservoir was of –32 %. In both cases, this index is underestimated, but there is a significant gain due to the reservoir implementation. On the other hand, the QMNA(5) (right panel of Fig. 17) also showed an improvement in median relative bias, from −64 % of the default simulations to −20 % of the simulation adding a reservoir.

In the case of regionalization approach (green boxes in Fig. 17), low flow indices reported values of median relative bias of −26 % and –22 % for the Q_{90}/Q_{50} ratio and QMNA(5), respectively. These values are practically same to the local calibration; only the 75th percentile (top edge of the box) for the Q_{90}/Q_{50} shows a larger difference. In any case, although low flow indices still remain underestimated, the inclusion and calibration of a reservoir (using either approach) noticeably improves the simulated values of the low flow indices and the regionalization approach behaves very well, compared to the local calibration one.

6. Discussion

We used a conceptual approach, based on rainfall-runoff models, to improve the streamflow simulated by SASER (adding a reservoir at the grid-scale resolution). This is intended to provide a better representation of the slow component in the hydrological response, which is not well simulated by SURFEX. Results show that the additional reservoir, which is simple and easy to calibrate, has a very positive impact on the streamflow simulation. However, the calibration of the reservoir parameters is only possible in gauged natural basins, which are not numerous in the area of study.

The implementation of the conceptual reservoir has as its main objective to sustain the flows without having a significant impact on the daily high and medium flows (surface runoff). For this reason, the reservoir scheme is implemented as a postprocessing of just the drainage, which is the slow component. This redistribution of the water volume does not affect the shape of the hydrograph during peak events. During the summer months, there is a significant improvement in the low flow simulation, as shown in Fig. 16. This fact supports the effectiveness of the reservoir to improve the streamflow simulation. At the same time, at the monthly step, we see that the annual cycle has changed considerably, redistributing water from winter to summer and obtaining a more realistic double-peaked hydrograph at the outlet of the Ebro basin.

A regionalization approach was established to find the values of the parameters for all the grid points of the area of study, not only those located in near-natural basins. This approach uses physiographic and climate-related predictors, although these variables are not immediately associated with groundwater, they exert an indirect influence on the runoff response. Therefore, those predictors are acting as proxies for defining the predictands. Moreover, several previous regionalization studies have emphasized the use of those predictors (Nijssen et al., 2001; Singh et al., 2014; Beck et al., 2016; 2020).

The performance of the model with regionalized parameters is almost as good as that of the catchment-by-catchment (local calibration) approach as reported KGE values. In the local calibration approach the values of the parameters are lumped, which probably would not work well in larger basins. Thus, the regionalization approach would probably be more advantaged if tested in larger basins. Additionally, having lumped values of the parameters per basin does not play well with a model that is mostly physical and which tries to be as spatially distributed as possible.

The regionalization approach allows us to apply the reservoir everywhere in the study domain, even in heavily human-influenced areas where a local calibration approach is not feasible. Due to the lack of validation data, we have compared our simulation with a reference one (SIMPA). The results from such an experiment are less easily interpretable. Still, least it makes it evident that the differences are lower now, even if they remain very large when compare with the reference model.

The maps of the reservoir parameters produced by the regionalization approach are presented in Fig. A1. These maps cover the full domain at 2.5 km resolution and vary according to climate and physiographic information that we used as predictors. It is difficult to explain the spatial patterns due to the complex and strong connection between the different variables (e.g. climate, vegetation and soil properties) involved (Troch et al., 2013).

Despite, we can identify spatial patterns related to known hydrological processes at least in one of the parameter maps (L, reservoir size). Patterns in this parameter, associated to soil moisture content, can be explained by the land cover, precipitation and soil granulometric distribution. For example, in the region of the south-west of France where large sand deposits exist (Landes), the runoff response is quick and infiltration has high rates, thus a buffer is required to play the role of groundwater storage and sustain the flows during the dry season, and hence the values of the reservoir size are expected to be larger than in clay predominated regions. On the other hand, the patterns in the map of the k parameter are not easily interpreted. Both, soil properties and topography, influence the distribution of water storage and isolating the influence of each other is problematic (Price, 2011).

Different regionalization experiments were carried out, to account the equifinality problem (multiple optimal solutions providing reasonably similar or equal model performance values), which is one of the most important sources of uncertainty in hydrological modeling (Beven & Freer, 2001). Furthermore, in each experiment cross-validation process was used, that allows us to estimate the generalizability of the parameters and provides an indication of uncertainty. Previous experiments done (not showed here) indicated that the parameters used to initialize the genetic algorithm have a strong influence over the search direction and efficiency score.

As shown in Fig. 13 (light blue bands), time series of simulated daily streamflow provide, in general, a wide range of streamflow uncertainty bands derived from the parameter sets, especially at the end of the wet season and gradually decreasing during the dry season (July to August). Otherwise, the uncertainty bands during the winter months were of much lesser extent, except in some months (November and December), this could be explained by the fact that the reservoir configuration focuses only on low flows, however, an extensive parameter uncertainty assessment is beyond the scope of this study.

In addition to KGE scores, and to understand how well the additional reservoir simulates low flows, which is the main objective of the study, we evaluated the Q_{90}/Q_{50} and the QMNA(5) indices. The negative relative bias in both indices of default simulation suggests a fast precipitation-to-runoff reaction of the SASER model; hence, the low flows are underestimated during the dry periods. Results of our different simulations (local calibration or regionalization approach) adding the reservoir showed a considerable improvement in both indices, although they are still slightly underestimated, suggesting the suitability of the conceptual reservoir to improve the simulation of the slow component of the streamflow, even though other aspects of the model still need to be improved.

A key result of this work is that introducing a conceptual reservoir, in the SASER model, is useful even if we cannot easily calibrate the parameters in each catchment, provided that values of the parameters are reasonable. This was demonstrated by different simulations (eight experiments and the three extra experiments using lumped parameters) carried out. Even if it is difficult to know which the most appropriated parameter maps is, at least we know that all of them produce reasonable results and that all of them are considerably better than the simulation without the reservoir (default).

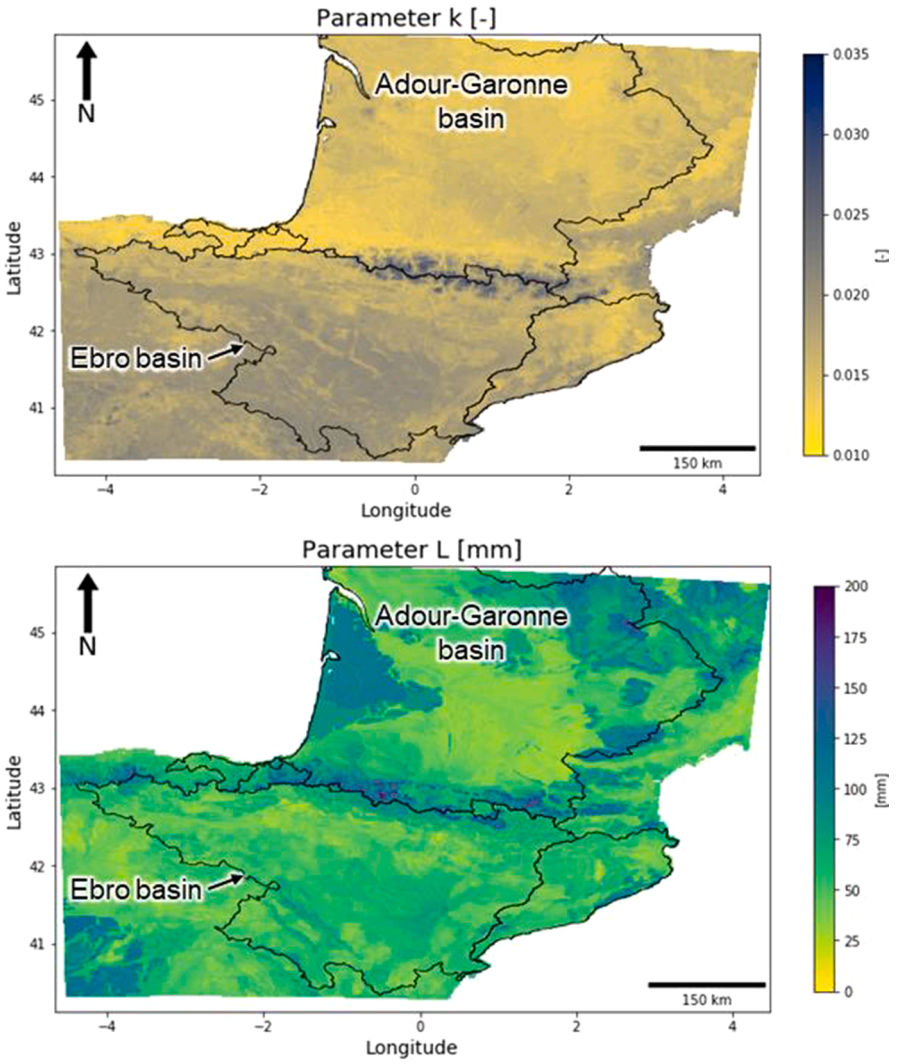

Fig. A1. Median of the eight maps of the two parameters (k and L_{max}) of the reservoir obtained applying the regionalization approach. The map set was obtained by cross-validation.

7. Conclusions

The main objective of this study is to improve the simulation of low streamflow in the hydrometeorological model SASER by using a simple conceptual reservoir scheme to postprocess the drainage and thus improve the slow component of the streamflow. The SASER simulations were made over a period of 35 years in an area that covers the Pyrenees and surrounding basins.

The default SASER model presents a strong negative bias of low flow indices. The addition of a reservoir scheme to modulate the drainage had a positive result in terms of KGE values, when the parameters were set using a local calibration method, both for the calibration and validation periods. The addition of a conceptual reservoir proved to be a simple and efficient option, with a limited number of parameters, which improves the low flows simulated by SURFEX without deteriorating high flows, as shown in the improvement of the low flow indices studied (QMNA(5) and Q_{90}/Q_{50}).

A regionalization approach, based on a genetic algorithm, was introduced to determine the values of the parameters all over the domain, including basins heavily influenced by water management, where a standard local calibration of these parameters is not possible. Results of the regionalization approach showed a clear general improvement of simulated streamflow ($\Delta KGE = 0.11$) with an improvement for 79 % of the validation catchments. Those results were almost as good as using local calibration. Both, KGE scores and the two low flow indices indicate improvements, especially for the low flow indices, although a small negative bias of the low flow indices remains.

The regionalization approach based on a genetic algorithm was possible due to the simplicity of the conceptual model and its implementation as an external module of SURFEX. Applying this approach to calibrate other empirical parameters within the SURFEX model itself would be desirable, but impractical due to computational constrains.

We established a relationship between the parameters of the conceptual reservoir with climate and physiographic variables through the genetic algorithm, which allows us to take into account the within-catchment variability all over the study area. Although this approach is not physically based, it allows linking the two new parameters with variables that are physical, which is a good compromise for a model that tries to be distributed and as physical as possible.

In conclusion, the addition of a conceptual reservoir to postprocess the SASER simulated drainage led to considerable improvement in low flow simulation. The regionalization approach allows us to apply the reservoir all over the domain including the human influenced basins, not just natural ones.

Funding

This work was partially funded by the HUMID project (CGL2017-

85687-R, AEI/FEDER, UE), the predoctoral grant PRE2018-085027 (AEI/FSE), the PIRAGUA project (EFA210/16-PIRAGUA, INTERREG V-A España-Francia-Andorra POCTEFA2014-2020) and the IDEWA project (PRIMA PCI2020-112043/AEI/ 10.13039/501100011033).

CRediT authorship contribution statement

Omar Cenobio-Cruz: Conceptualization, Formal analysis, Methodology, Visualization, Writing – original draft. **Pere Quintana-Seguí:** Formal analysis, Conceptualization, Software, Writing – review & editing, Supervision, Funding acquisition. **Anaïs Barella-Ortiz:** Software, Writing – review & editing. **Ane Zabaleta:** Resources, Data curation, Formal analysis, Writing – review & editing. **Luis Garrote:** Supervision, Writing – review & editing. **Roger Clavera-Gispert:** Writing – review & editing. **Florence Habets:** Writing – review & editing. **Santiago Beguería:** Writing – review & editing.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Appendix A. . - Parameter maps of the reservoir scheme

.

References

Artinyan, E., Habets, F., Noilhan, J., Ledoux, E., Dimitrov, D., Martin, E., Le Moigne, P., 2008. Modelling the water budget and the riverflows of the Maritsa basin in Bulgaria. Hydrol. Earth Syst. Sci. 12 (1), 21–37. https://doi.org/10.5194/hess-12-21-2008.

Barella-Ortiz, A., Quintana-Seguí, P., 2019. Evaluation of drought representation and propagation in regional climate model simulations across Spain. Hydrol. Earth Syst. Sci. 23 (12), 5111–5131. https://doi.org/10.5194/hess-23-5111-2019.

Batalla, R.J., Gómez, C.M., Kondolf, G.M., 2004. Reservoir-induced hydrological changes in the Ebro River basin (NE Spain). J. Hydrol. 290 (1–2), 117–136. https://doi.org/10.1016/j.jhydrol.2003.12.002.

Bates, B.C., Kundzewicz, Z.W., Wu, S., (2008). Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariar, Geneva, 210 pp. ISBN: 978-92-9169-123-4.

Beck, H.E., van Dijk, A.I.J.M., de Roo, A., Miralles, D.G., McVicar, T.R., Schellekens, J., Bruijnzeel, L.A., 2016. Global-scale regionalization of hydrologic model parameters. Water Resour. Res. 52 (5), 3599–3622. https://doi.org/10.1002/2015WR018247.

Beck, H.E., Pan, M., Lin, P., Seibert, J., van Dijk, A.I.J.M., Wood, E.F., 2020. Global fully distributed parameter regionalization based on observed streamflow from 4,229 headwater catchments. J. Geophys. Res. Atmos. 125 (17) https://doi.org/10.1029/2019JD031485.

Beguería, S., Palazón, L., Grusson, Y., Sánchez Pérez, J.M., Sauvage, S., Cakir, R., Quintana-Seguí, P., Barella, A., Vidal, J.P. (2022); PIRAGUA_hydro_climate ; Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas (EEAD-CSIC); https://doi.org/10.20350/digitalCSIC/14668.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249 (1–4), 11–29. https://doi.org/10.1016/S0022-1694(01)00421-8.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrol. Sci. Bull. 24 (1), 43–69. https://doi.org/10.1080/02626667909491834.

Catalogne, C., 2012. *Amélioration des méthodes de prédétermination des débits de référence d ' étiage en sites peu ou pas jaugés*. Doctoral dissertation, Doctorat Ocean Atmosphere Hydrologie. Université Joseph Fourrier, Grenoble.

Clark, M.P., Vogel, R.M., Lamontagne, J.R., Mizukami, N., Knoben, W.J.M., Tang, G., Gharari, S., Freer, J.E., Whitfield, P.H., Shook, K.R., Papalexiou, S.M., 2021. The abuse of popular performance metrics in hydrologic modeling. e2020WR029001 Water Resour. Res. 57. https://doi.org/10.1029/2020WR029001.

David, C.H., Maidment, D.R., Niu, G.Y., Yang, Z.L., Habets, F., Eijkhout, V., 2011. River network routing on the NHDPlus dataset. J. Hydrometeorol. 12 (5), 913–934. https://doi.org/10.1175/2011JHM1345.1.

Donohue, R.J., Roderick, M.L., McVicar, T.R., 2007. On the importance of including vegetation dynamics in Budyko's hydrological model. Hydrol. Earth Syst. Sci. 11 (2), 983–995. https://doi.org/10.5194/hess-11-983-2007.

Durand, Y., Brun, E., Merindol, L., Guyomarc'h, G., Lesaffre, B., Martin, E., 1993. A meteorological estimation of relevant parameters for snow models. Ann. Glaciol. 18, 65–71. https://doi.org/10.3189/s0260305500011277.

Estrela, T., Quintas, L. (1996). El sistema integrado de modelización precipitación-aportación SIMPA. Revista Digital del Cedex, (104), 43. Recuperado a partir de http://ingenieriacivil.cedex.es/index.php/ingenieria-civil/article/view/1153 (last access: 2022/06/07).

Faroux, S., Kaptué Tchuenté, A.T., Roujean, J.-L., Masson, V., Martin, E., Le Moigne, P., 2013. ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. Geosci. Model Dev 6, 563–582. https://doi.org/10.5194/gmd-6-563-2013.

Gaona, J., Quintana-seguí, P., Escorihuela, M.J., Boone, A., Llasat, M.C., 2022. Interactions between precipitation, evapotranspiration and soil moisture-based indices to characterize drought with high-resolution remote sensing and land-surface model data. Nat. Hazards Earth Syst. Sci. [PREPRINT]. https://doi.org/10.5194/nhess-2022-65.

Garcia, F., Folton, N., Oudin, L., 2017. Which objective function to calibrate rainfall–runoff models for low-flow index simulations? Hydrol. Sci. J. 62 (7), 1149–1166. https://doi.org/10.1080/02626667.2017.1308511.

Gascoin, S., Ducharne, A., Ribstein, P., Carli, M., Habets, F., 2009. Adaptation of a catchment-based land surface model to the hydrogeological setting of the Somme River basin (France). J. Hydrol. 368 (1–4), 105–116. https://doi.org/10.1016/J.JHYDROL.2009.01.039.

Gericke, O.J., Smithers, J.C., 2014. Review of methods used to estimate catchment response time for the purpose of peak discharge estimation. Hydrol. Sci. J. 59 (11), 1935–1971. https://doi.org/10.1080/02626667.2013.866712.

Getirana, A.C.V., Boone, A., Peugeot, C., 2014. Evaluating LSM-based water budgets over a West African basin assisted with a river routing scheme. J. Hydrometeorol. 15 (6), 2331–2346. https://doi.org/10.1175/JHM-D-14-0012.1.

Gudmundsson, L., Tallaksen, L.M., Stahl, K., Clark, D.B., Dumont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J., Hanasaki, N., Voss, F., Koirala, A.S., 2012. Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe. J. Hydrometeorol. 13 (2), 604–620. https://doi.org/10.1175/JHM-D-11-083.1.

Guimberteau, M., Ducharne, A., Ciais, P., Boisier, J.P., Peng, S., De Weirdt, M., Verbeeck, H., 2014. Testing conceptual and physically based soil hydrology schemes against observations for the Amazon Basin. Geosci. Model Dev. 7 (3), 1115–1136. https://doi.org/10.5194/gmd-7-1115-2014.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

Gustard, A., Bullock, A., Dixon, J. M. (1992). *Low flow estimation in the United Kingdom*. Institute of Hydrology. ISBN 0 948540 45 1.

Habets, F., Boone, A., Champeaux, J.L., Etchevers, P., Franchistéguy, L., Leblois, E., Ledoux, E., Le Moigne, P., Martin, E., Morel, S., Noilhan, J., Seguí, P.Q., Rousset-Regimbeau, F., Viennot, P., 2008. The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France. J. Geophys. Res. Atmos. 113 (6), 1–18. https://doi.org/10.1029/2007JD008548.

Huang, Z., Tang, Q., Lo, M.H., Liu, X., Lu, H., Zhang, X., Leng, G., 2019. The influence of groundwater representation on hydrological simulation and its assessment using satellite-based water storage variation. Hydrol. Process. 33 (8), 1218–1230. https://doi.org/10.1002/HYP.13393.

Immerzeel, W.W., Lutz, A.F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., Hyde, S., Brumby, S., Davies, B.J., Elmore, A.C., Emmer, A., Feng, M., Fernández, A., Haritashya, U., Kargel, J.S., Koppes, M., Kraaijenbrink, P.D.A., Kulkarni, A.V., Mayewski, P.A., Nepal, S., Pacheco, P., Painter, T.H., Pellicciotti, F., Rajaram, H., Rupper, S., Sinisalo, A., Shrestha, A.B., Viviroli, D., Wada, Y., Xiao, C., Yao, T., Baillie, J.E.M., 2019. Importance and vulnerability of the world's water towers. Nature 577 (7790), 364–369.

Keyantash, J., Dracup, J.A., 2002. The quantification of drought: an evaluation of drought indices. Bull. Am. Meteorol. Soc. 83 (8), 1167–1180 https://doi.org/10.-0477.

Knoben, W.J.M., Freer, J.E., Woods, R.A., 2019. Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. Hydrol. Earth Syst. Sci. 23 (10), 4323–4331. https://doi.org/10.5194/hess-23-4323-2019.

Lafaysse, M., Hingray, B., Etchevers, P., Martin, E., Obled, C., 2011. Influence of spatial discretization, underground water storage and glacier melt on a physically-based hydrological model of the Upper Durance River basin. J. Hydrol. 403 (1–2), 116–129. https://doi.org/10.1016/J.JHYDROL.2011.03.046.

Le Moigne, P., Besson, F., Martin, E., Boé, J., Boone, A., Decharme, B., Etchevers, P., Faroux, S., Habets, F., Lafaysse, M., Leroux, D., Rousset-Regimbeau, F., 2020. The latest improvements with SURFEX v8.0 of the Safran–Isba–Modcou hydrometeorological model for France. Geosci. Model Dev. 13, 3925–3946. https://doi.org/10.5194/gmd-13-3925-2020.

Lehner, B., Döll, P., Alcamo, J., Henrichs, T., Kaspar, F., 2006. Estimating the impact of global change on flood and drought risks in europe: a continental, integrated analysis. Clim. Change 75 (3), 273–299. https://doi.org/10.1007/S10584-006-6338-4.

Lehner, B., Grill, G., 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrol. Process. 27 (15), 2171–2186. https://doi.org/10.1002/HYP.9740.

Liu, Z., Wang, Y., Xu, Z., Duan, Q., 2017. Conceptual Hydrological Models. In: Duan, Q., Pappenberger, F., Thielen, J., Wood, A., Cloke, H.L., Schaake, J.C. (Eds.), Handbook of Hydrometeorological Ensemble Forecasting. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–23.

López, R., Justribó, C., 2010. The hydrological significance of mountains: a regional case study, the Ebro River basin, northeast Iberian Peninsula. Hydrol. Sci. J. 55 (2), 223–233. https://doi.org/10.1080/02626660903546126.

López-Moreno, J.I., Beniston, M., García-Ruiz, J.M., 2008. Environmental change and water management in the Pyrenees: Facts and future perspectives for Mediterranean mountains. Global Planet. Change 61 (3–4), 300–312. https://doi.org/10.1016/J.GLOPLACHA.2007.10.004.

López-Moreno, J.I., Vicente-Serrano, S.M., Moran-Tejeda, E., Zabalza, J., Lorenzo-Lacruz, J., García-Ruiz, J.M., 2011. Impact of climate evolution and land use changes on water yield in the ebro basin. Hydrol. Earth Syst. Sci. 15 (1), 311–322. https://doi.org/10.5194/hess-15-311-2011.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., Voldoire, A., 2013. The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geosci. Model Dev. 6 (4), 929–960. https://doi.org/10.5194/gmd-6-929-2013.

Mathevet, T., Michel, C., Andréassian, V., Perrin, C. (2006). A bounded version of the Nash–Sutcliffe criterion for better model assessment on large sets of basins. In: *Andréassian, V., Hall, A., Chahinian, N., Schaake, J. (Eds.), Large Sample Basin Experiment for Hydrological Model Parameterization: Results of the Model Parameter Experiment – MOPEX*. IAHS Publ, p. 567.

Maxwell, R.M., Miller, N.L., 2004. On the development of a coupled land surface and groundwater model. Dev. Water Sci. 55 (PART 2), 1503–1510. https://doi.org/10.1016/S0167-5648(04)80161-8.

Miguez-Macho, G., Fan, Y., Weaver, C.P., Walko, R., Robock, A., 2007. Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation. J. Geophys. Res. Atmos. 112 (13), 13108. https://doi.org/10.1029/2006JD008112.

Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. J. Hydrol. 391 (1–2), 202–216. https://doi.org/10.1016/j.jhydrol.2010.07.012.

Mo, K.C., Lettenmaier, D.P., 2014. Objective drought classification using multiple land surface models. J. Hydrometeorol. 15 (3), 990–1010. https://doi.org/10.1175/JHM-D-13-071.1.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. J. Hydrol. 10 (3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Nijssen, B., O'Donnell, G.M., Lettenmaier, D.P., Lohmann, D., Wood, E.F., 2001. Predicting the discharge of global rivers. J. Clim. 14 (15), 3307–3323. https://doi.org/10.1175/1520-0442(2001)014<3307:PTDOGR>2.0.CO;2.

Peel, M.C., 2009. Hydrology: catchment vegetation and runoff. Prog. Phys. Geogr. 33 (6), 837–844. https://doi.org/10.1177/0309133309350122.

Price, K., 2011. Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: a review. Prog. Phys. Geogr. 35 (4), 465–492. https://doi.org/10.1177/0309133311402714.

Prudhomme, C., Parry, S., Hannaford, J., Clark, D.B., Hagemann, S., Voss, F., 2011. How well do large-scale models reproduce regional hydrological extremes in Europe? J. Hydrometeorol. 12 (6), 1181–1204. https://doi.org/10.1175/2011JHM1387.1.

Prudhomme, C., Giuntoli, I., Robinson, E.L., Clark, D.B., Arnell, N.W., Dankers, R., Fekete, B.M., Franssen, W., Gerten, D., Gosling, S.N., Hagemann, S., Hannah, D.M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., Wisser, D., 2014. Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. Proc. Natl. Acad. Sci. 111 (9), 3262–3267. https://doi.org/10.1073/pnas.1222473110.

Quintana-Seguí, P., Le Cointe, P. (2022); PIRAGUA_atmos_analysis [Dataset]; Observatori de l'Ebre (URL – CSIC); http://digital.csic.es/handle/10261/271111; https://doi.org/10.20350/digitalCSIC/14665.

Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., Morel, S., 2008. Analysis of near-surface atmospheric variables: validation of the SAFRAN analysis over France. J. Appl. Meteorol. Climatol. 47 (1), 92–107. https://doi.org/10.1175/2007JAMC1636.1.

Quintana-Seguí, P., Peral, C., Turco, M., Llasat, M.C., Martin, E., 2016. Meteorological analysis systems in North-East Spain: validation of SAFRAN and SPAN. J. Environ. Inf. 27 (2), 116–130. https://doi.org/10.3808/jei.201600335.

Quintana-Seguí, P., Turco, M., Herrera, S., Miguez-Macho, G., 2017. Validation of a new SAFRAN-based gridded precipitation product for Spain and comparisons to Spain02 and ERA-Interim. Hydrol. Earth Syst. Sci. 21 (4), 2187–2201. https://doi.org/10.5194/hess-21-2187-2017.

Quintana-Seguí, P., Barella-Ortiz, A., Regueiro-Sanfiz, S., Miguez-Macho, G., 2020. The utility of land-surface model simulations to provide drought information in a water management context using global and local forcing datasets. Water Resour. Manag. 34 (7), 2135–2156. https://doi.org/10.1007/s11269-018-2160-9.

Saleh, F., Flipo, N., Habets, F., Ducharne, A., Oudin, L., Viennot, P., Ledoux, E., 2011. Modeling the impact of in-stream water level fluctuations on stream-aquifer interactions at the regional scale. J. Hydrol. 400 (3–4), 490–500. https://doi.org/10.1016/J.JHYDROL.2011.02.001.

Santos, L., Thirel, G., Perrin, C., 2018. Technical note: pitfalls in using log-transformed flows within the KGE criterion. Hydrol. Earth Syst. Sci. 22 (8), 4583–4591. https://doi.org/10.5194/hess-22-4583-2018.

Seneviratne, S., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., Zhang, X. (2012). Changes in climate extremes and their impacts on the natural physical environment. *Managing the Risk of Extreme Events and Disasters to Advance Climate Change Adaptation.*, 109–230. https://doi.org/10.1017/CBO9781139177245.006.

Sheffield, J., Wood, E.F., Roderick, M.L., 2012. Little change in global drought over the past 60 years. Nature 491 (7424), 435–438. https://doi.org/10.1038/nature11575.

Singh, R., Archfield, S.A., Wagener, T., 2014. Identifying dominant controls on hydrologic parameter transfer from gauged to ungauged catchments - A comparative hydrology approach. J. Hydrol. 517, 985–996. https://doi.org/10.1016/j.jhydrol.2014.06.030.

Slowik, A., Kwasnicka, H., 2020. Evolutionary algorithms and their applications to engineering problems. Neural Comput & Applic 32 (16), 12363–12379. https://doi.org/10.1007/s00521-020-04832-8.

Smakhtin, V.U., 2001. Low flow hydrology: a review. J. Hydrol. 240 (3–4), 147–186. https://doi.org/10.1016/S0022-1694(00)00340-1.

Somers, L.D., McKenzie, J.M., 2020. A review of groundwater in high mountain environments. Wiley Interdiscip. Rev. Water 7 (6), e1475. https://doi.org/10.1002/wat2.1475.

Stahl, K., Tallaksen, L.M., Gudmundsson, L., Christensen, J.H., 2011. Streamflow data from small basins: a challenging test to high-resolution regional climate modeling. J. Hydrometeorol. 12 (5), 900–912. https://doi.org/10.1175/2011JHM1356.1.

Stahl, K., Kohn, I., Blauhut, V., Urquijo, J., De Stefano, L., Acácio, V., Dias, S., Stagge, J. H., Tallaksen, L.M., Kampragou, E., Van Loon, A.F., Barker, L.J., Melsen, L.A., Bifulco, C., Musolino, D., de Carli, A., Massarutto, A., Assimacopoulos, D., Van Lanen, H.A.J., 2016. Impacts of European drought events: insights from an international database of text-based reports. Nat. Hazards Earth Syst. Sci. 16 (3), 801–819. https://doi.org/10.5194/nhess-16-801-2016.

Staudinger, M., Stahl, K., Seibert, J., Clark, M.P., Tallaksen, L.M., 2011. Comparison of hydrological model structures based on recession and low flow simulations. Hydrol. Earth Syst. Sci. 15 (11), 3447–3459. https://doi.org/10.5194/hess-15-3447-2011.

Sutanudjaja, E.H., Van Beek, L.P.H., De Jong, S.M., Van Geer, F.C., Bierkens, M.F.P., 2011. Large-scale groundwater modeling using global datasets: a test case for the Rhine-Meuse basin. Hydrol. Earth Syst. Sci. 15 (9), 2913–2935. https://doi.org/10.5194/HESS-15-2913-2011.

Tesfa, T.K., Tarboton, D.G., Chandler, D.G., McNamara, J.P., 2009. Modeling soil depth from topographic and land cover attributes. Water Resour. Res. 45 (10), 1–16. https://doi.org/10.1029/2008WR007474.

Tian, W., Li, X., Cheng, G.D., Wang, X.S., Hu, B.X., 2012. Coupling a groundwater model with a land surface model to improve water and energy cycle simulation. Hydrol. Earth Syst. Sci. 16 (12), 4707–4723. https://doi.org/10.5194/hess-16-4707-2012.

Todini, E., 1996. The ARNO rainfall-runoff model. J. Hydrol. 175 (1–4), 339–382. https://doi.org/10.1016/S0022-1694(96)80016-3.

Troch, P.A., Carrillo, G., Sivapalan, M., Wagener, T., Sawicz, K., 2013. Climate-vegetation-soil interactions and long-term hydrologic partitioning: signatures of catchment co-evolution. Hydrol. Earth Syst. Sci. 17 (6), 2209–2217. https://doi.org/10.5194/hess-17-2209-2013.

Van Loon, A.F., 2015. Hydrological drought explained. WIREs. Water 2 (4), 359–392. https://doi.org/10.1002/wat2.1085.

Van Loon, A.F., Van Huijgevoort, M.H.J., Van Lanen, H.A.J., 2012. Evaluation of drought propagation in an ensemble mean of large-scale hydrological models. Hydrol. Earth Syst. Sci. 16 (11), 4057–4078. https://doi.org/10.5194/hess-16-4057-2012.

Van Loon, A.F., Gleeson, T., Clark, J., Van Dijk, A.I.J.M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A.J., Tallaksen, L.M., Uijlenhoet, R., Hannah, D.M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., Van Lanen, H.A.J., 2016. Drought in the Anthropocene. Nat. Geosci. 9 (2), 89–91. https://doi.org/10.1038/ngeo2646.

Vergnes, J.P., Decharme, B., Alkama, R., Martin, E., Habets, F., Douville, H., 2012. A simple groundwater scheme for hydrological and climate applications: description and offline evaluation over france. J. Hydrometeorol. 13 (4), 1149–1171. https://doi.org/10.1175/JHM-D-11-0149.1.

Vergnes, J.-P., Habets, F., 2018. Impact of river water levels on the simulation of stream–aquifer exchanges over the Upper Rhine alluvial aquifer (France/Germany). Hydrgeol. J. 26 (7), 2443–2457. https://doi.org/10.1007/s10040-018-1788-0.

Vergnes, J.P., Roux, N., Habets, F., Ackerer, P., Amraoui, N., Besson, F., Caballero, Y., Courtois, Q., De Dreuzy, J.R., Etchevers, P., Gallois, N., Leroux, D.J., Longuevergne, L., Le Moigne, P., Morel, T., Munier, S., Regimbeau, F., Thiéry, D., Viennot, P., 2020. The AquiFR hydrometeorological modelling platform as a tool for improving groundwater resource monitoring over France: evaluation over a 60-year period. Hydrol. Earth Syst. Sci. 24 (2), 633–654. https://doi.org/10.5194/HESS-24-633-2020.

Vidal, J.-P., Martin, E., Franchistéguy, L., Habets, F., Soubeyroux, J.-M., Blanchard, M., Baillon, M., 2010. Multilevel and multiscale drought reanalysis over France with the Safran-Isba-Modcou hydrometeorological suite. Hydrol. Earth Syst. Sci. 14 (3), 459–478. https://doi.org/10.5194/hess-14-459-2010.

Wanders, N., Wada, Y., 2015. Human and climate impacts on the 21st century hydrological drought. J. Hydrol. 526, 208–220. https://doi.org/10.1016/j.jhydrol.2014.10.047.

Xia, Y., Ek, M.B., Mocko, D., Peters-Lidard, C.D., Sheffield, J., Dong, J., Wood, E.F., 2014. Uncertainties, correlations, and optimal blends of drought Indices from the NLDAS multiple land surface model ensemble. J. Hydrometeorol. 15 (4), 1636–1650. https://doi.org/10.1175/JHM-D-13-058.1.

York, J.P., Person, M., Gutowski, W.J., Winter, T.C., 2002. Putting aquifers into atmospheric simulation models: an example from the Mill Creek Watershed, Northeastern Kansas. *Adv. Water Resour.* 25 (2), 221–238. https://doi.org/10.1016/S0309-1708(01)00021-5.

Zabaleta, A., Beguería, S., Antigüedad, I., Lambán, J., Hakoun, V., Jung, M., Le Cointe, P., Caballero, Y., (2022). PIRAGUA_indicators [Dataset]. https://doi.org/10.20350/DIGITALCSIC/14658.

Zhang, L., Dawes, W.R., Walker, G.R., 2001. Response of mean annual evapotranspiration to vegetation changes at catchment scale. The research on the hydrological role of vegetation has. *Water Resour. Res.* 37 (3), 701–708. https://doi.org/10.1029/2000WR900325.

Zhao, T., Dai, A. (2015). The Magnitude and Causes of Global Drought Changes in the Twenty-First Century under a Low–Moderate Emissions Scenario, Journal of Climate, 28(11), 4490-4512. Retrieved Jun 10, 2022, from https://journals.ametsoc.org/view/journals/clim/28/11/jcli-d-14-00363.1.xml.