Signatures-and-sensitivity-based multi-criteria variational calibration for distributed hydrological modeling applied to Mediterranean floods

Ngo Nghi Truyen Huynh, Pierre-André Garambois, Pierre Javelle, François Colleoni

Signatures-and-sensitivity-based multi-criteria variational calibration for distributed hydrological modeling applied to Mediterranean floods

Ngo Nghi Truyen Huynh[a], Pierre-André Garambois[a], Pierre Javelle[a] and François Colleoni[a]

[a]*INRAE, Aix Marseille Université, RECOVER, 3275 Route Cézanne, Aix-en-Provence, 13182, France*

ARTICLE INFO

Keywords:
Hydrological modeling
Hydrological signatures
Variational data assimilation
Multi-objective optimization
Pareto-optimal solution
Sensitivity analysis

ABSTRACT

Classical calibration methods in hydrology are commonly performed with a single cost function computed on long time series. Even though the hydrological model has acceptable scores in NSE and KGE, unbalancing problems can still arise between overall score and the model performance for flood events, and particularly flash floods. Enhancing multi-criteria calibration methods with multi-scale signatures to improve distributed flood modeling remains a challenge. In this study, the potential of hydrological signatures computed continuously and at the scale of flood events on long time series, is employed within various multi-criteria calibration approaches to attain a more efficient hydrological model. This work presents an improved and original signature-based calibration approach, implemented in the variational data assimilation algorithm of SMASH (*Spatially distributed Modelling and ASsimilation for Hydrology*) platform, applied over 141 catchments mostly located in the French Mediterranean region. Several signatures, especially flood event signatures are firstly computed, relying on a proposed automatic hydrograph segmentation algorithm. Suitable signatures for constraining the model are selected based on their global sensitivity analysis to model parameters. Several multi-criteria calibration strategies with the selected signatures are eventually performed, including a multi-objective optimization approach, and a single-objective optimization approach, that transforms the multi-criteria problem into a single-objective function. Note that in the first approach, the proposed technique based on a simple additive weighting method is used to select an optimal solution obtained from a set of non-inferior solutions. The suggested methods show that, for a global calibration, the average relative error in simulating the peak flow has been dropped from about 0.27 to 0.01-0.08 and from about 0.30 to 0.18-0.21 with various multi-criteria optimization strategies, respectively in calibration and temporal validation. For a distributed calibration, while the average NSE (resp. KGE) still slightly decreases from 0.78 (resp. 0.86) to 0.75 (resp. 0.81) in calibration, the quality of simulated peak flow has been enhanced about 1.5 times in average. In particular, the NSE (resp. KGE) calculated solely on 111 flood events which are picked from 23 downstream gauges has been improved from 0.80 (resp. 0.71) up to 0.83 (resp. 0.78) in median. These results have demonstrated the robustness and delicacy of the model constrained by the signatures for enhancing flash flood forecasting systems.

1. Introduction

The calibration of hydrological model parameters is a crucial issue in view to improve model performance in simulating catchment response but is a tough problem because of modeling-data errors and relatively sparse observability of hydrological responses. Whatever their status and complexity, hydrological models are conventionally calibrated with a discharge time series at the outlet of a catchment, representing integrative information of upstream hydrological processes. These results in ill-posed calibration inverse problems with non unique solutions, which has been called equifinality issue in hydrological modeling by K. Beven (cf. Beven (2011) and references therein) after Bertalanffy (1968) in general system theory. This poses a major challenge, especially in the context of climate change and in Mediterranean areas, where an intensification of the frequency of extreme precipitation is possible (Tramblay and Somot, 2018; Tramblay, Neppel, Carreau and Najib, 2013; Pujol, Neppel and Sabatier, 2007). To produce more accurate flood forecasts in terms of location, magnitude and timing

of hydrological responses, improving model-data fusion and exploitation of the most information available in calibration is of prior importance, which can be studied by taking advantage of hydrological signatures in calibration (e.g. Roux, Labat, Garambois, Maubourguet, Chorda and Dartus (2011); Shafii and Tolson (2015); Oliveira, Fleischmann and Paiva (2021); Mostafaie, Forootan, Safari and Schumacher (2018); Sahraei, Asadzadeh and Unduche (2020); Wu, Chen, Ye, Guo, Meng and Zhang (2021)).

In hydrology, most calibration approaches attempt to optimize input parameters of a model such that they result in a minimal misfit between simulated and observed discharge. Nevertheless, because no single metric can exhaustively represent this misfit, the calibration of a hydrological model is "inherently multi-objective" as remarked by Gupta, Sorooshian and Yapo (1998). Several performance metrics have been proposed over the past decades in the literature for hydrological modeling. The classical quadratic Nash-Sutcliffe efficiency (NSE) Nash and Sutcliffe (1970) (cf. Appendix A.1) has been used for long time. The Kling–Gupta (KGE) (cf. Appendix A.2) proposed in Gupta, Kling, Yilmaz and Martinez (2009) and based on a decomposition of the NSE has also become widely used. Other metrics, in form of signature measures (see review and in McMillan (2021)),

✉ ngo-nghi-truyen.huynh@inrae.fr (N.N.T. Huynh);
pierre-andre.Garambois@inrae.fr (P. Garambois); pierre.javelle@inrae.fr
(P. Javelle); francois.colleoni@inrae.fr (F. Colleoni)
ORCID(s): 0000-0001-5078-3865 (N.N.T. Huynh)

have been proposed in the literature for model evaluation (e.g. Yilmaz, Gupta and Wagener (2008)) and used in model optimization (e.g. Roux et al. (2011); Shafii and Tolson (2015); Mostafaie et al. (2018); Sahraei et al. (2020); Wu et al. (2021) and references therein). Moreover, hydrological signatures are a useful tool to effectively evaluate models and diagnose the role of their components in explaining the discrepancy between the simulated and observed behavior (Gupta et al., 2009), especially when combined with global sensitivity analysis (Horner, 2020). There is still a need for automated methods capable of computing signatures on observed and modeled hydrological responses, at multiple time scales with the underlying difficulty of flood events segmentation, and perform global sensitivity analysis of simulated signatures to model parameters (Horner, 2020).

Considering multiple metrics, including multi-scale signatures, in a hydrological calibration problem pertains to multi-criteria optimization problems. Basically, there are three categories for solving a multi-criteria problem in various domains: i. obtain a full non-inferior solution set (Pareto-front) by solving the multi-objective optimization problem, which is the mainstream approach (Khorram, Khaledian and Khaledyan, 2014; Abbass, Sarker and Newton, 2001; Goel, Vaidyanathan, Haftka, Shyy, Queipo and Tucker, 2007; Tavakkoli-Moghaddam, Azarkish and Sadeghnejad-Barkousaraie, 2011; Padhye and Deb, 2011; Yeh and Chuang, 2011; Das, Nikum, Krishnan and Pratihar, 2020; Torres-Treviño, Reyes-Valdes, López and Praga-Alejo, 2011), ii. transform multi-criteria problem into a single-objective optimization problem (Ross, Abbey, Bouffard and Jos, 2015; El-Ghandour and Elbeltagi, 2014; Veluscek, Kalganova, Broomhead and Grichnik, 2015; Masuda, Harada and Kurihara, 2012), iii. select an unique solution after obtaining the Pareto optimal solution set by adding some constraints based on specific preferences (Chibeles-Martins, Pinto-Varela, Barbosa-Póvoa and Novais, 2016; Wu, Liu and Lur, 2015). In hydrology, state-of-the-art multi-criteria-optimization-using-signatures-liked calibration strategies are roughly divided into three approaches:

1. *The signatures are intervened in the calibration algorithm*. For instance, Wu et al. (2021); Sahraei et al. (2020) among others proposed multi-criteria approaches to improve the optimal value of the calibration metric (NSE). This approach is more advanced in term of algorithm[1], comparing to traditional sampling methods as Shuffled Complex Evolution (SCE-UA) (Duan, Gupta and Sorooshian, 1993), by using additional criterion based on some hydrological signatures such as: the peak flow, high-low and low-flow, etc.. These signatures-based metrics enable to sort and rank efficiently the parameter sets for the optimization problem.

2. *Transform the signatures-based multi-criteria problem into a single-objective function*. For example, Mizukami, Rakovec, Newman, Clark, Wood, Gupta and Kumar (2019); Roux et al. (2011) tried to enhance the calibration metric by emphasizing specific hydrological signatures. A typical and classical approach of this idea is to utilize a calibration metric (cost function) based on the decomposition of NSE (for example: KGE in Appendix A.2). Moreover, we can combine NSE-liked cost functions and multi-signatures for the optimization problem. For example, the cost function introduced by Roux et al. (2011), in the context of "event-based hydrological modeling", to account for the shape of flash flood hydrographs and especially the timing and maximum peak flow. This method enables to calibrate the weight for each objective function (that represents the important of each objective function in the optimization problem), however, it can not generate a set of alternative solutions based on multi-objective functions.

3. *Obtain a full estimated non-inferior solution set*. Typically, Yapo, Gupta and Sorooshian (1998); Guo, Zhou, Lu, Zou, Zhang and Bi (2014); Oliveira et al. (2021) can generate a set of non-dominated solutions based on multi-objective functions but selection of an optimal solution from the non-dominated set has not received attention. Even the study in Mostafaie et al. (2018) proposed a simple technique to select the optimal solution from Pareto front by ranking the parameter sets with their distance to the reference solution, this can only be applicable when assuming that the objective functions have equal roles in the optimization problem.

While the potential of signatures for constraining hydrological models has been employed in McMillan (2021); Westerberg and McMillan (2015), and the interest of flood event signatures dedicated to multi-criteria calibration has been conveyed in Roux et al. (2011) for an event model, most of published paper on the signatures-based multi-criteria and multi-objective optimization has not yet been fully and rigorously achieved. One of the difficulties of this problem is the automatic computation of signatures, especially for flood event signatures, that requires a segmentation method (e.g. Astagneau, Bourgin, Andréassian and Perrin (2021)) to capture some hydrological information in flash flood events. In this study, we focus on multi-criteria calibration metrics with single-objective function and with multi-objective functions for uniform parameters, then a multi-criteria calibration with single-objective function for distributed parameters of a distributed model aiming at flood modeling. Note that multi-objective optimization is one of several approaches for multi-criteria calibration method, which attempts to simultaneously minimize multi-objective functions to obtain not only one but a set of optimal solutions. This study proposes an original calibration approach combining automated segmentation of flood events and signature computation within

---

[1]Mathematically, the calibration metric is unchanged (the optimization problem is always based on NSE). The signatures-based metrics are used in the algorithm to help optimize the calibration metric (assuming that the studied signatures and the discharge are closely correlated).

a variational data assimilation (VDA) algorithm from Jay-Allemand, Javelle, Gejadze, Arnaud, Malaterre, Fine and Organde (2020) enabling high dimensional spatially distributed calibration, now with multi-criteria metrics adapted to floods. Classical global calibration algorithm have also been upgraded that way. These upgrades, including new cost functions and adjoint model update, have been implemented into the SMASH platform, which Fortran source code has recently been interfaced in Python (Jay-Allemand, Colleoni, Garambois, Javelle and Julie, 2022).

The remaining sections of this paper are organized as follows: in section 2, we give a brief representation of the data and conceptual hydrological model. Next, we describe in Section 3 our methodology for computing some essential signatures and then our signatures-based calibration methods on the multi-criteria and multi-objective optimization. In Section 4, we showcase some results on the signatures, calibration results as well as give some discussions and finally, Section 5 concludes on this work.

2. Data and model

2.1. Data

The relatively large dataset from Jay-Allemand (2020) that is used in this study contains time series of hydro-meteorological variables and time invariant catchment attributes for four high rainfall-flow areas in France, identified as study areas of the PICS research project[2]. It encompasses 141 catchments including 23 drainage basins (downstream gauges), which are mostly located in the French Mediterranean region (Fig. 1). This is a subset of a larger dataset of 4,190 French catchments from INRAE-HYCAR research unit (Brigode, Génot, Lobligeois and Delaigue, 2020; Delaigue, Génot, Lebecherel, Brigode and Bourgin, 2020). Concerning the hydrological model inputs, observation data covers a period about 13 years of hourly distributed discharge and rainfall from 2006 to 2019. Discharge data are collected by the French Ministry of Environment covering the period of the forcing data and have been extracted from the (Hydro) platform[3]. The rainfall grids are the radar observation reanalysis ANTILOPE J+1 provided by Météo-France at a grid resolution of 1 km² (Champeaux, Dupuy, Laurantin, Soulan, Tabary and Soubeyroux, 2009). The potential evapotranspiration (PET) is obtained by applying a simple formula (Oudin, Hervieu, Michel, Perrin, Andréassian, Anctil and Loumagne, 2005) to SAFRAN (Quintana-Seguí, Le Moigne, Durand, Martin, Habets, Baillon, Canellas, Franchisteguy and Morel, 2008) temperature grids at 8 km resolution an empirically desaggregated at hourly time step and 1 km spatial resolution, i.e. at the same spatio-temporal resolution than rainfall. Note that observation data, rainfall grids and discharge time series, over the selected catchments have few missing data as detailed in Table 1, so that it can be neglected when performing the computations

[2]https://pics.ifsttar.fr
[3]http://www.hydro.eaufrance.fr/

and analysis in this study. Catchment information concerning river name, surface, code, number of upstream gauges and missing rates in downstream gauges are mentioned in Table 1. Raster maps, at 1 km resolution, of upstream drained area and D8 flow directions have been obtained by processing fine DEM provided by IGN (Institut Geographique National).
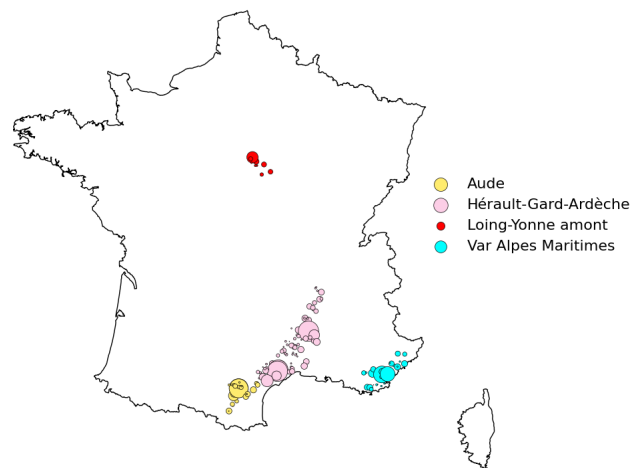

Figure 1: Spatial distribution of 141 catchments of the PICS dataset on the map of France. Four studied areas are represented by different colors. The size of each marker illustrates how large the catchment is.

2.2. Forward hydrological model

SMASH is a computational software framework, wrapped in Python (Jay-Allemand et al., 2022), and dedicated to spatially distributed hydrological modeling including variational data assimilation (VDA) and calibration method (adjoint) adapted to high dimensional problems (cf. Jay-Allemand et al. (2020); Colleoni, Garambois, Javelle, Jay-Allemand and Arnaud (2022) and references therein) . It is designed to simulate discharge hydrographs at any spatial location within a basin and reproduce the hydrological response of contrasted catchments.

Let $\Omega \subset \mathbb{R}^2$ be a 2D spatial domain (catchment) and $t > 0$ be the physical time. A regular lattice $\mathcal{R}_\Omega$ covers Ω and $D(x)$ is the drainage plan obtained from terrain elevation processing. The number of active cells within a catchment Ω is denoted N_x.

The hydrological model is a dynamic operator $\mathcal{M}$ mapping observed input fields of rainfall and evapotranspiration $\boldsymbol{P}(x, t')$, $\boldsymbol{E}(x, t')$, $\forall (x, t') \in \Omega \times [0, t]$ onto discharge field $Q(x, t)$ such that:

$$Q(x, t) = \mathcal{M}\left[\boldsymbol{P}(x, t'), \boldsymbol{E}(x, t'), \boldsymbol{h}(x, 0), \boldsymbol{\theta}(x), t\right],$$
$$\forall x \in \Omega, t' \in [0, t] \quad (1)$$

with $\boldsymbol{h}(x, t)$ the N_s-dimensional vector of model states 2D fields and $\boldsymbol{\theta}$ the N_p-dimensional vector of model parameters 2D fields. In the following, $\boldsymbol{\theta}$ is also called control vector in optimization context.

Code	River name	Surface (km^2)	Missing rates (%)	Total upstream gauges
H3201010	Le Loing	2302	0.14 (3.68)	8
V3524010	La Cance	381	0.14 (4.31)	3
V3744010	Le Doux	621	0.14 (4.02)	2
V4154010	L'Eyrieux	649	0.14 (7.38)	3
V5064010	L'Ardèche	2264	0.14 (4.22)	9
V5474015	La Cèze	1112	0.14 (3.76)	6
V7164015	Le Gardon	1093	0.14 (16.62)	10
Y1232010	L'Aude	1828	0.14 (3.74)	11
Y1364010	Le Fresquel	935	0.14 (3.74)	4
Y1415020	L'Orbiel	242	0.14 (3.74)	2
Y1564010	L'Orbieu	589	0.14 (3.77)	3
Y1605050	La Cesse	251	0.14 (4.64)	1
Y2332015	L'Hérault	2208	0.14 (7.22)	12
Y2584010	L'Orb	1336	0.14 (4.04)	11
Y3204040	Le Lez	168	0.14 (15.55)	3
Y3444020	Le Vidourle	503	0.14 (7.97)	4
Y3534010	Le Vistre	496	0.14 (4.42)	1
Y4624010	Le Gapeau	535	0.14 (3.79)	6
Y5312010	L'Argens	2512	0.14 (5.08)	10
Y5444010	La Giscle	201	0.14 (9.96)	2
Y5534030	La Siagne	492	0.14 (5.30)	5
Y5615030	Le Loup	289	0.14 (3.79)	1
Y6434010	L'Estéron	442	0.14 (7.70)	1

Table 1
General information about 23 downstream gauges of the PICS data. Code, river name, surface, missing rate of rainfall (resp. discharge) in downstream gauge during the period 2006-2019, and number of upstream are represented by the columns from left to right.

In this study, a parsimonious 6 parameters model structure from Colleoni et al. (2022) is used (Fig. 2). For a given cell i of coordinates $x \in \Omega$, in the proposed model S6, four reservoirs $\mathcal{I}$, $\mathcal{P}$, $\mathcal{T}_r$ and $\mathcal{T}_l$ of respective capacity c_i, c_p, c_{tr} and c_{tl}, are considered for simulating respectively the interception, the production of runoff and its transfer within a cell. Their state vector is denoted:

$$\boldsymbol{h}(x,t) \equiv \left(\boldsymbol{h}_i(x,t), \boldsymbol{h}_p(x,t), \boldsymbol{h}_{tr}(x,t), \boldsymbol{h}_r(x,t), \boldsymbol{h}_{tl}(x,t)\right)^T .$$

The parameter vector of SMASH model structure S6 is:

$$\boldsymbol{\theta}(x) \equiv \left(c_i(x), c_p(x), c_{tr}(x), c_r(x), ml(x), c_{tl}(x)\right)^T .$$

Hence the size of state vector is $N_s \times N_x$ and the size of parameter vector that is optimized in the following is $N_p \times N_x$. Considering tens of cells or more over a simulated catchments domain Ω, the calibration of $\boldsymbol{\theta}$ is a high dimensional inverse problem. All details related to hydrological model operator and model description are fully explained in Colleoni et al. (2022).

2.3. Calibration algorithms

Given simulated and observed discharge at gauged cells $x_k \in \Omega, k \in 1, .., Ng$, respectively denoted $Q_k(t)$ and $Q_k^*(t)$, we define the objective function as:

$$J(\boldsymbol{\theta}) = J_{obs}(\boldsymbol{\theta}) + \alpha J_{reg}(\boldsymbol{\theta}) \quad (2)$$

where the observation cost function is:

$$J_{obs} = \frac{1}{N_g} \sum_{k=1}^{N_g} j_k^*$$

measuring the misfit, via several adapted metrics detailed later, between simulated and observed discharge. In this study, $N_g = 1$, that is for single gauge optimization. Note that simulated discharge:

$$Q_k(t) = \mathcal{M}\left[\boldsymbol{P}(x, t'), \boldsymbol{E}(x, t'), \boldsymbol{h}(x, 0), \boldsymbol{\theta}(x)\right],$$
$$\forall x \in \Omega_k, t' \in [0, t]$$

with $\Omega_k \subset \Omega$ denoting the spatial domain including all upstream cells of a gauge at x_k, depends on the control vector $\boldsymbol{\theta}$ via the hydrological model $\mathcal{M}$ (Eq. 1). The second term in Eq. (2) is weighted by α and set as a classical Thikhonov regularization:

$$J_{reg} = \left\| \boldsymbol{B}^{-1/2}(\boldsymbol{\theta} - \boldsymbol{\theta}^*) \right\|_{L^2}^2$$

with $\boldsymbol{B}$ the background error covariance, and $\boldsymbol{\theta}^*$ the first guess/background on $\boldsymbol{\theta}$. We set $\alpha = 10^{-4}$ for the spatially distributed optimizations presented in this study, $\alpha = 0$ otherwise if $\boldsymbol{\theta} \equiv \bar{\boldsymbol{\theta}}$, and $\boldsymbol{B}$ is simply defined from $\sigma_{\boldsymbol{\theta}}$ the vector of mean deviations of $\boldsymbol{\theta}$, as done in Jay-Allemand et al. (2020). Furthermore, an additional constrain is introduced for parameter optimization such that:

$$\boldsymbol{\theta}_{min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{max} \quad (3)$$

Given the hydrological model (Eq. 1) and the constrain (Eq. 3), an optimal estimate $\hat{\boldsymbol{\theta}}$ of model parameter set is obtained from the condition:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}_{min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{max}}{\arg \min} J(\boldsymbol{\theta}) \quad (4)$$

This inverse problem 4 is tackled with different global optimization algorithms considering a spatially uniform control, that is low dimensional optimization problems. For instance, optimization algorithms such as: Step-By-Step (SBS) (steepest descent algorithm summarized in Edijatno (1991)), Nelder–Mead and Genetic Algorithms (GA) can be applied in this scenario. Next, a spatially distributed control vector is sought with a VDA algorithm (Jay-Allemand et al., 2020) adapted to such high dimensional hydrological optimization problems. Considering a spatially distributed control vector $\boldsymbol{\theta}(x)$, its optimization is performed with the L-BFGS-B algorithm (limited-memory Broyden–Fletcher–Goldfarb–Shanno bound-constrained (Zhu, Byrd, Lu and Nocedal, 1997)) adapted to high dimension. This algorithm requires the gradient of the cost function with respect to the sought parameters $\nabla_{\boldsymbol{\theta}} J$, that is obtained by solving the adjoint model. This numerical adjoint model has been generated with the automatic differentiation engine TAPENADE (Hascoet and Pascual, 2013) applied to the SMASH source code including the new models structures
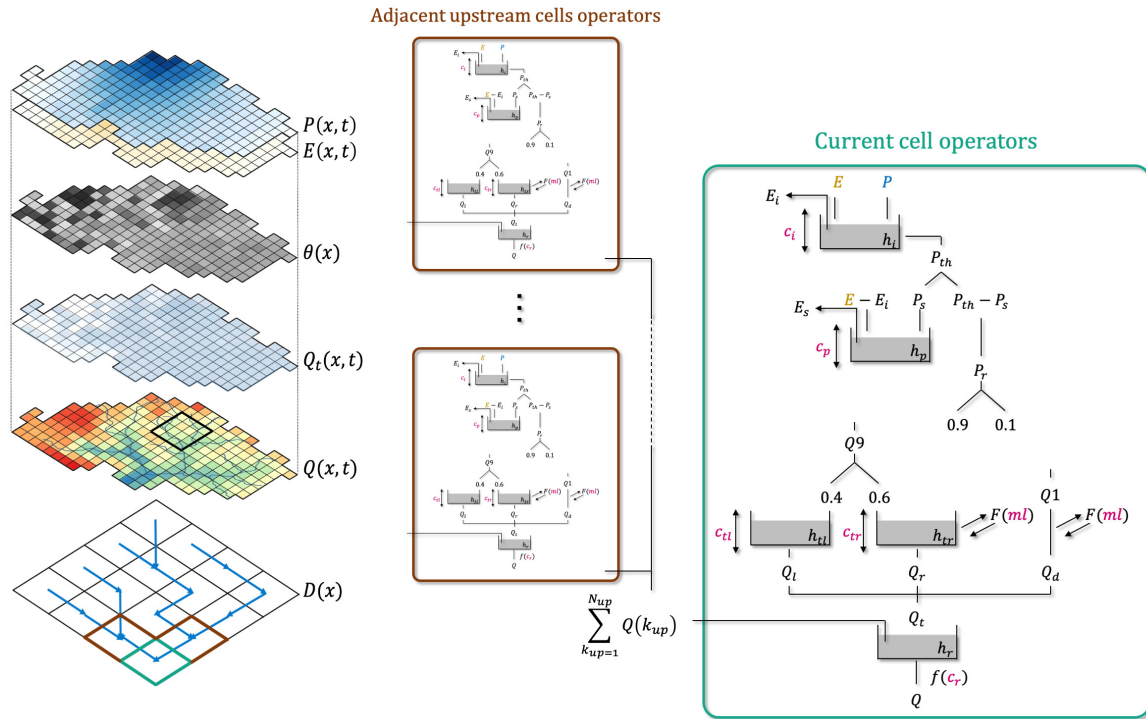
Figure 2: Distributed hydrological modeling with SMASH platform. Left: model fields from top to bottom: meteorological inputs, param eters, internal and output flux maps. Right: pixel scale and pixel-to pixel flow operators of SMASH model structure S6 studied. *Source:* Colleoni et al. (2022).

and validated with standard gradient test. The background value θ^*, used as a starting point for the optimization problem and in the regularization term, is set as in Jay-Allemand et al. (2020), i.e. as $\bar{\theta}$, a spatially uniform global optimum determined with a simple global-minimization algorithm from a uniform first guess $\bar{\theta}^*$, the steepest descent method summarized in Edijatno (1991). Given mildly non linear hydrological models as those considered in this study, this calibration approach is pertinent and sensitivity to priors is limited as shown in Jay-Allemand et al. (2020). In case of multi-objective optimization with a spatially uniform control, i.e. low dimensional inverse problems, a genetic algorithm is used.

3. Methodology

A novel calibration strategy is proposed and illustrated in Fig. 3. It is based on the combination of hydrological signatures and their sensitivity analysis along with the above optimization algorithms. The computations of the signatures are firstly performed to quantify their sensitivities to model parameters following Horner (2020). These computations are covered by a whole-period-based studying (continuous signatures) and an event-based analyzing, that tries to capture the most important events to study (flood event signatures). This step gives us a more meaningful understanding of the parametric sensitivity, not only of the discharge but also of the signatures and other factors that need to be accounted as part of our minimization criterion.

Along with an analysis of signatures error, the signatures sensitivity can be computed based on Sobol indices to help decide which signatures can be selected to perform a multi-criteria optimization. Finally, we conduct a multi-criteria optimization with single-objective function or with multi-objective function using appropriate hydrological signatures to improve the simulation performance.

The numerical algorithms proposed here are implemented in Python, on top of SMASH Fortran platform that is interfaced in Python (Jay-Allemand et al., 2022) making accessible its forward-inverse algorithms (forward hydrological models, SBS and VDA calibration algorithms) and internal variables.

3.1. Signatures computation

Several signatures describing and quantifying properties of discharge time series are introduced in view to analyze and calibrate hydrological models (Appendix B). Signatures are denoted S_i, $i \in 1..N_{crit}$, with N_{crit} being the number of different signature types considered. These signatures permit to describe various aspects of the rainfall-runoff behavior such as: flow distribution (based for instance on flow percentiles), flow dynamics (Le Mesnil, 2021), flow separation (Nathan and McMahon, 1990; Lyne and Hollick, 1979)), flow timing, etc.. A so-called continuous signature is a signature that can be computed on the whole studied period. Flood event signatures on the other hand focus on the behavior of the high-flows that are observed in the flood events (Fig. 4). In
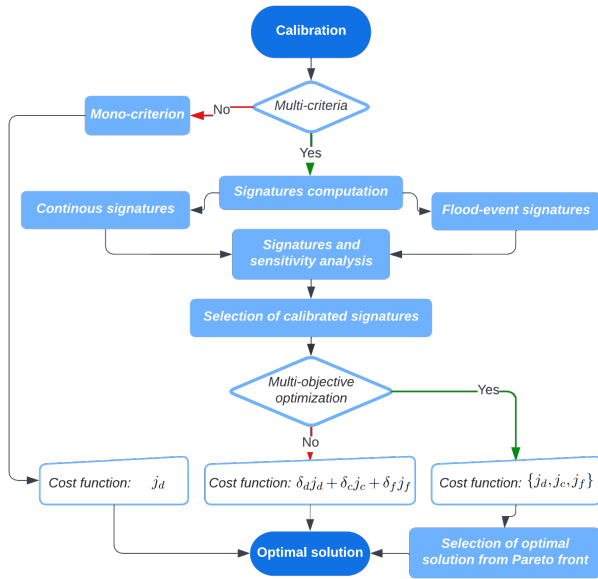
Figure 3: Illustration of a calibration process considering classical hydrological cost functions with/without continuous and/or events signatures. The notations in the cost function will be explained in 3.3.

this way, event segmentation algorithm is crucially needed before computing the flood event signatures.

A segmentation algorithm aims to capture important events occurring over a studied period for a given catchment. We propose here such an algorithm for detecting flood events with the aid of the rainfall gradient and rainfall energy (Algorithm 1). Let us consider the precipitation and the discharge at each catchment. Firstly, we detect the discharge peaks achieved with a simple peak detection algorithm, where several parameters can be used to impose a minimum peak height (mph) or a minimum distance between two successive peak (mpd), etc. (Duarte and Watanabe, 2021). For instance, we consider those exceed the 0.999-quantile of the discharge as important events (mph criterion) w.r.t. two successive events are distinguished by at least 12h (mpd criterion). Subsequently, we need to determine a starting and ending date for each event. The starting date of event is considered as the moment it starts raining dramatically sometime since 72h before the peak of discharge. For calculating this, we need to compute the gradient of the rainfall and then we choose the peaks of rainfall gradient which exceed its 0.8-quantile. These peaks correspond to the moments when we have a sharp increase in rainfall. However, this criterion is not enough to choose a good moment for starting date. In addition to gradient criterion, we need an additional criterion called "energy criterion" that takes into account the "rainfall energy". The rainfall energy is computed as the sum of squares of the rainfall observed in 12h counted from 1h before the peak of rainfall gradient. So the starting date is the first moment when the rainfall energy exceeds 0.2 of the maximal rainfall energy observed in 72h before the peak of discharge (w.r.t. gradient criterion). Finally, we aim to find

the ending date by using baseflow separation. Effectively, we compute the difference between the discharge and its baseflow from the peak of discharge until the end of researched period (lasts 10 days from starting date of event). Then, the ending date is the moment that the difference between the discharge and its baseflow is minimal in 24 hours counted from 1 hour before this moment. Note that the above values are adapted to the basins and flood scales studied.

Algorithm 1 Hydrograph segmentation algorithm based on "rainfall gradient criterion" and "rainfall energy criterion"

For each catchment, considering 2 time series (T, Q) and (T, P) where:

$T = (t_1, ..., t_n)$ is time (by hour),
$Q = (q_1, ..., q_n)$ is the discharge and
$P = (p_1, ..., p_n)$ is the rainfall.

1. Detecting peaks that exceed the 0.999-quantile of the discharge, that can be considered as important events: $E = (t_i)_{1 \leq i \leq n}$ s.t. $q_i > Quant_{0.999}(Q)$

2. For each event $t_j \in E$:

   (a) Determining a starting date based on the "rainfall gradient criterion" and the "rainfall energy criterion":

      i. Selecting rainfalls gradient those exceed its 0.8-quantile, considered as the "rainfall events":
      $RE = (t_k)_{t_k \in (t_j - 72, t_j)}$
      s.t. $\nabla P(t_k) > Quant_{0.8}(\nabla P([t_j - 72, t_j]))$

      ii. Defining the rainfall energy function:
      $f(t_x) = ||(p_x - 1, ..., p_x + 11)||_2$
      then the starting date is the first moment the rainfall energy exceeds 0.2 of the maximal rainfall energy:
      $sd = \min(t_s)_{t_s \in RE}$
      s.t. $f(s) > 0.2||(f(t_j - 72), ..., f(t_j))||_\infty$

   (b) Determining an ending date based on discharge baseflow $Qb = Baseflow(Q)$:
   $ed = \arg\min_{t_e} \sum_{t=t_e-1}^{t_e+23} |(Q - Qb)(t)|$
   s.t. $t_j \leq t_e \leq sd + 10 \times 24$

Remark. If there exists $m + 1$ ($m > 0$) consecutive events $(sd_u, ed_u), ..., (sd_{u+m}, ed_{u+m})$ occurring "nearly simultaneously", that means all of these events occur in no more than 10 days: $ed_{u+m} < sd_u + 10 \times 24$, then we merge these $m + 1$ events into a single event (sd_u, ed_{u+m}).

3.2. Sobol indices

To perform a calibration process with hydrological signatures, it is important to investigate the sensitivity of simulated signatures to the model parameters, to guide the potential selection of the signature(s) which should be used to calibrate the model. The sensitivity analysis enables us to examine how the variation of a given output/signature can be apportioned to a variation in model inputs (Saltelli, 2002). In other words, this can help detecting the possibility for
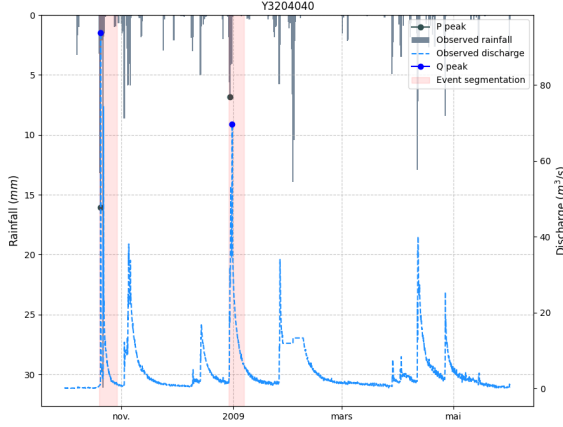
Figure 4: Example of flood events detected from hydrograph with the segmentation algorithm.

moving the model parameters to constrain some important features of the hydrological model via the signature(s). If some signatures are not sensitive to the model parameters, then it might be useless to optimize an objective function based on these signatures.

Let us consider the m-parameters set $\theta = (x_1, ..., x_m)$. Then the simulated value of a signature S_i is represented as $S_i^s = f_i(\theta)$. We are interested in Sobol indices called, first-order, second-order, and so forth. The first- (depending on x_j), second- (depending on x_j and x_k) Sobol indices of S_i^s are respectively defined as follows:

$$s_i^{1j} = \frac{\mathbb{V}[\mathbb{E}[S_i^s|x_j]]}{\mathbb{V}[S_i^s]} \text{ and } s_i^{2jk} = \frac{\mathbb{V}[\mathbb{E}[S_i^s|x_j,x_k]]}{\mathbb{V}[S_i^s]}.$$

Azzini, Mara and Rosati (2021) proposed a method to estimate the first- and total-order variance-based sensitivity indices (also called Sobol indices) on parameter sets of Monte-Carlo simulations via Saltelli generator (Saltelli, 2002). For the most part, estimating the second and higher order variance-based sensitivity indices is complicated to achieve and inconsistent in the context of high dimensional parameter space with such algorithms that require a large amount of Monte-Carlo simulations. To analyze the signatures sensitivity with the 6-parameters model, we are thus interested in the estimation of first order Sobol indices which is more reliable in the present context.

3.3. Signatures-based multi-criteria optimization

Let us consider a classical objective function j_d, which is the dominant criterion (or the most constrained criterion) in case of multi-criteria optimization, an objective function j_c combining continuous-signatures-based cost functions, and j_f combining flood-event-signatures-based cost functions. Then, the cost function to minimize J can be defined as:

$$J \equiv \begin{cases} \delta_d j_d + \delta_c j_c + \delta_f j_f \text{ for single-objective optimization,} \\ (j_d, j_c, j_f) \text{ for multi-objective optimization} \end{cases}$$

where $\delta_d, \delta_c, \delta_f$ are the corresponding optimization weights in the first case. Keep in mind that we take into account the use of signatures in both cases but the first case is a single-objective optimization while the second is a multi-objective optimization.

For each signature S_i, denote by S_i^o and S_i^s the observation and simulation respectively. The set of continuous and flood event signatures denoted N_c and N_f respectively. Then, the components j_d, j_c and j_f can be defined as follows:

- $j_d \equiv 1 - NSE$ or $1 - KGE_{(\alpha,\beta,\gamma)}$ with varying weights α, β, γ (see Appendix A.2). This metric j_d is considered as a constraining objective function for selecting an optimal solution from non-inferior solutions in case of multi-objective optimization (see Appendix C.3).

- $j_c \equiv \begin{cases} \sum_{S_i \in N_c} \sigma_{S_i} j_c^{S_i}, \text{ for single-objective} \\ \qquad\qquad \text{multi-criteria optimization;} \\ (\{j_c^{S_i}\}_{S_i \in N_c}), \text{ for multi-objective optimization} \end{cases}$

  where $j_c^{S_i} \equiv \left| \frac{S_i^s}{S_i^o} - 1 \right|$ is the objective function based on continuous signature S_i and σ_{S_i} is the corresponding optimization weight of S_i in case of single-objective function.

- $j_f \equiv \begin{cases} \sum_{S_i \in N_f} \sigma_{S_i} j_f^{S_i}, \text{ for single-objective} \\ \qquad\qquad \text{multi-criteria optimization;} \\ (\{j_f^{S_i}\}_{S_i \in N_f}), \text{ for multi-objective optimization.} \end{cases}$

  In this case, and in the context of global optimization in time, $j_f^{S_i} \equiv \frac{1}{N_E} \sum_{e=1}^{N_E} \left| \frac{S_{i,e}^s}{S_{i,e}^o} - 1 \right|$ defines the scalar objective function related to flood signature $S_i \in N_f$ over the N_E events selected with the segmentation method described in 3.1. Otherwise, to perform a season-based optimization on flood event signatures, we can compute for the events occurring in the selected season. For example, for a Spring-based optimization:

$$j_{f,spring}^{S_i} \equiv \frac{1}{\dim \mathcal{SP}} \sum_{e \in \mathcal{SP}} \left| \frac{S_{i,e}^s}{S_{i,e}^o} - 1 \right|$$

s.t. $\forall e \in \mathcal{SP} \subset \{1, ..., N_E\}, S_{i,e}$ occurs in Spring.

Hence, the optimization problems taking into account signatures via the cost function defined in Eq. 5 can be developed as Eq. 6 for a single-objective optimization, and as Eq. 7 for a multi-objective optimization.

$$\min_{\theta \in \mathcal{O} \subset \mathbb{R}^n} \delta_d j_d(\theta) + \delta_c \sum_{S_i \in N_c} \sigma_{S_i} \left| \frac{S_i^s(\theta)}{S_i^o(\theta)} - 1 \right|$$

$$+ \delta_f \sum_{S_i \in N_f} \sigma_{S_i} \frac{1}{N_E} \sum_{e=1}^{N_E} \left| \frac{S_{i,e}^s(\theta)}{S_{i,e}^o(\theta)} - 1 \right|$$

$$\min_{\theta \in \mathcal{O} \subset \mathbb{R}^n} \left(j_d(\theta), \left\{ \left| \frac{S_i^s(\theta)}{S_i^o(\theta)} - 1 \right| \right\}_{S_i \in N_c}, \right.$$

$$\left. \left\{ \frac{1}{N_E} \sum_{e=1}^{N_E} \left| \frac{S_{i,e}^s(\theta)}{S_{i,e}^o(\theta)} - 1 \right| \right\}_{S_i \in N_f} \right)$$
(7)

While the minimization problem with single-objective function 6 is accessible for both global and distributed calibration methods, performing a multi-objective optimization as problem 7 is sophisticated for distributed calibration considering a spatially distributed control vector adapted to a high dimensional hydrological optimization problems, and requiring a lot of cost gradient information. In global calibration with multi-objective optimization approaches, a set of feasible solutions can be found instead of an unique optimal solution in single-objective optimization (Appendix C). In such a way, a so-called "Pareto front" contains non-inferior solutions (Appendix C.1) and thus a method is proposed for selecting an optimal solution from the Pareto as depicted in Appendix C.3.

According to Yapo et al. (1998), classical generating techniques to sample the entire Pareto set are inefficient with a large number of objective functions due to time consuming (it grows rapidly at an exponential rate) and they provided in the same paper an alternative method "Multi-Objective Complex Evolution" (MOCOM-UA) for implementing efficiently a global optimization algorithm in hydrology. In addition to Strength Pareto Evolutionary Algorithm (SPEA) (Zitzler and Thiele, 1999) like MOCOM-UA, there are various multi-objective optimization techniques including Multi-Objective Genetic Algorithm (MOGA) (Murata, Ishibuchi et al., 1995), Multiple Gradient Descent Algorithm (MGDA) (Désidéri, 2012; Mercier, Poirion and Désidéri, 2018), etc.. Comparing to MGDAs, MOGAs are more effective regardless of the nature of the problem functions because they do not require gradient information, but instead are based on crossover and mutation operators. "Non-dominated Sorting Genetic Algorithm II (NSGA-II) suggested by Deb, Pratap, Agarwal and Meyarivan (2002) is a well known GA for solving multi/many-objective optimization problems including fast and elitist approach (Deb et al., 2002). Namely, a fast sorting algorithm helps optimizing the computational complexity (even with a large population size) arising from the non-dominated sorting procedure in every generation. Into the bargain, NSGA-II possesses a diversity preservation property, based on a sharing function method, that prevents the loss of good solutions involved in the mating process. Recently, NSGA-II has also been implemented in a Python library named *pymoo* (Blank and Deb (2020)) that is used in the present study thanks to the Python interface of our SMASH platform.

Note that the objective functions j_c and j_f related to continuous and flood signatures have also been implemented in Fortran and numerical adjoint model re-derived as needed by

the variational calibration algorithm (cf. section 2.3). j_f can be computed thanks to a temporal mask of corresponding to flood events selected by the Python segmentation algorithm and passed to the Fortran via the wrapped interface.

The proposed numerical algorithms are implemented in Python, on top of SMASH Fortran platform that is interfaced in Python making accessible its forward-inverse algorithms (forward hydrological models, SBS and VDA calibration algorithms) (cf. Jay-Allemand et al. (2022)).

4. Results and discussion

4.1. Signatures analysis

To start with, global calibrations using SBS algorithm are conducted to analyze the relative error between simulation and observation of the signatures. Table 2 shows that some hydrological signatures with a significant simulation error such as: $Cq2$, $Cq10$, $Cq50$, $Etre$ and Eqp need to be constrained in the calibration process (a list of all studied signatures with corresponding notations is presented in Appendix B). Next, we survey the sensitivity of these signatures to the model parameters. We considered over 10,000 spatially uniform sets of the 6 model parameters, sampled using Saltelli generator (Saltelli, 2002) to estimate the frist-order Sobol indices across 141 studied catchments (Table 3). In general, continuous signatures are less sensitive to the model than flood-event signatures. Furthermore, constraining hydrological model by flood event signatures along with a classical calibration metric (for example with $1 - NSE$ or $1 - KGE$), which is basically based on continuous records of streamflow, is ideal to balance the model between the global score and the performance on flood events. We select for example the peak flow (Eqp), which is one of flood event signatures having both significant relative error and high sensitivity (the other one is $Etre$), to perform multi-critera calibration methods. Note that multi-criteria optimization methods with multi signatures are absolutely reachable but will not be shown in this study for sake of brevity and simplify results analysis.

4.2. Signatures-based multi-criteria optimization

In this section, we discuss the performance of different models with i. uniform and ii. distributed calibration methods. For uniform calibration methods, we aim to compare different calibration metrics including classical single-objective optimization (CSOO), signature-based single-objective optimization (SSOO) and signature-based multi-objective optimization (SMOO). For distributed methods, two strategies selected for comparison are CSOO and SSOO. In both scenarios, the models are calibrated on 23 downstream gauges of the PICS data on the calibration period 2006-2013. Three validation metrics are chosen to validate the simulated results:

- on 23 downstream gauges on the validation period 2012-2019 (temporal validation, "T_Val").

- on all 141 catchments on the calibration period 2006-2013 (spatial validation, "S_Val").

Signature	Calibration metric	
	j_d^{NSE}	j_d^{KGE}
Ccr	0.14[0.28, 0.38]	0.16[0.3, 0.46]
Ccrqf	0.24[0.35, 0.35]	0.26[0.4, 0.45]
Ccrbf	0.15[0.3, 0.44]	0.15[0.33, 0.54]
Ccrsc	0.23[0.4, 0.68]	0.22[0.38, 0.69]
Cq2	0.72[3.99, 21.14]	0.76[5.99, 29.98]
Cq10	0.52[2.64, 8.8]	0.52[2.87, 9.42]
Cq50	0.29[0.49, 0.85]	0.2[0.52, 0.99]
Cq90	0.21[0.37, 0.96]	0.18[0.38, 0.99]
Evqq	0.23[0.32, 0.31]	0.19[0.31, 0.37]
Evqb	0.22[0.33, 0.39]	0.22[0.33, 0.41]
Ecr	0.2[0.28, 0.26]	0.18[0.27, 0.26]
Ecrqf	0.23[0.32, 0.31]	0.19[0.31, 0.37]
Ecrbf	0.22[0.33, 0.39]	0.22[0.33, 0.41]
Ecrsc	0.12[0.19, 0.2]	0.13[0.2, 0.24]
Etre	0.48[0.96, 1.25]	0.46[0.82, 1.1]
Eqp	0.28[0.38, 0.35]	0.25[0.36, 0.41]

Table 2
Relative error between simulated and observed signatures of the same model structure calibrated either with $1 - NSE$ or $1 - KGE$ by SBS algorithm for global optimization. The values (in the form of .[.,.]) in each case represent respectively the median, mean and standard deviation of a signature over all catchments downstream outlets of the dataset.

- on all 141 catchments on the validation period 2012-2019 (spatio-temporal validation, "S-T_Val").

4.2.1. With global calibration

Firstly, the global calibration methods are performed with NSGA for 2 single-objective-functions-based approaches and 1 multi-objective-function-based approach. Table 4 displays the mean of different objective functions for calibration and validation (with 3 validation metrics), and for 3 optimization criteria (CSOO, SSOO and SMOO) with various cost functions. In CSOO, we interpret that the model calibrated with $j_d^{KGE} = 1 - KGE$ produces a better result on the peak flow ($j_f^{Q_{peak}}$) compared to the one calibrated with $j_d^{NSE} = 1 - NSE$. This explains why KGE criterion is more robust than NSE for constraining a hydrological model, since it is built on the decomposition of NSE (Gupta et al., 2009), which emphasizes relative importance of several hydrological features. Using event signatures in addition to classical continuous metrics in SSOO, we found that simulated peak flow is highly improved in terms of relative error (about 15-18 times and 1.4-1.7 times on average resp. for calibration and temporal validation) while classical calibrated metrics are significantly deteriorated (about 1.4-1.6 times and 1.4-1.5 times on average resp. for calibration and temporal validation). This may arise from unbalancing problem between global score and performance in simulating flood event signature. To circumvent this problem, the choice of optimization weights of objective functions need to be carefully studied to balance the model performance on both short and long time series. That may be a very time consuming approach with numerous simulations running in order to find appropriate optimization weights for the objective functions. Alternatively, since global calibration algorithm does not require any gradient information and is approached in lower dimensional optimization problem, this unbalancing issue can be achieved with multi-objective optimization approach. This way keeps an acceptable level of deterioration of NSE and KGE while enhancing significantly the simulation of the peak flow (Fig. 5 and 6). However, this global multi-optimization algorithm is not capable to deal with high dimensional control vectors and the uniform setup here led to unsatisfactory results in spatial and spatio-temporal validation metrics. So, it is interesting to see how a SSOO with distributed calibration, i.e.including signatures in our VDA algorithm, can improve the model without conducting many experiments for selection of optimization weights (so the objective functions have the same optimization weights as above) at the end of this section.

Interestingly, the corresponding optimal parameters obtained with various optimization strategies are presented in Fig. 7. A preliminary review points out that the distribution over studied catchments of c_r has an important difference when performing traditional calibration (CSOO) and multi-criteria calibration methods (SSOO and SMOO). We recall that c_r is the routing parameter in our conceptual design (Fig. 2), so it has a crucial role in producing the peak flow Eqp. Additionally, the sensitivity analysis in Table 3 has indicated that c_r is one of the three parameters explaining most of the sensitivity of the signature Eqp.

4.2.2. With distributed calibration

Now we employ SSOO technique for a distributed calibration using L-BFGS-B algorithm provided a first guess by SBS algorithm. In overall, all of obtained scores in Table 5 are significantly enhanced compared to the uniform calibration method, thanks to spatially distributed control vectors granting more flexibility to reproduce observed discharge. Instead of a sharp decline of $j_f^{Q_{peak}}$ as above, this relative error slightly decreases about 1.5 times (from about 0.25 down to 0.16) in calibration and from about 0.32 down to 0.28 in temporal validation, but instead, the scores (NSE and KGE) are slightly reduced in calibration and have an inappreciable deterioration in temporal validation. So in this case, we do not have the unbalancing issue between the model performances on short and long time series when employing SSOO. We observe clearly in Fig. 8 and 9 that the error of simulated pick flow is significantly reduced while the deterioration level of the scores remains tolerable, particularly in calibration and temporal validation.

Ultimately, the scoring metrics are computed on 111 flood events picked from 23 downstream gauges (by segmentation method depicted in Algorithm 1) on the calibration period. The results plotted in Fig. 10 show that, in distributed calibration, the score of constrained calibration metric is not decreased but even improved from 0.80 (resp. 0.71) up to 0.83 (resp. 0.78) in median for NSE (resp. KGE). It indicates that the optimum of the model parameters has moved to another location that produces a better performance in simulating flood events by slightly reducing the scores in simulating the low-flow.

Signature	Model parameter					
	c_i	c_p	c_{tr}	c_{tl}	c_r	ml
Ccr	-0.01[-0.0, 0.01]	-0.04[-0.04, 0.01]	13.46[13.46, 0.26]	0.08[0.12, 0.1]	-0.0[-0.0, 0.0]	12.07[12.56, 2.75]
Ccrqf	0.33[0.63, 0.89]	2.72[3.44, 1.93]	36.2[36.13, 2.47]	1.65[1.91, 1.24]	2.18[2.88, 2.19]	23.49[23.94, 6.31]
Ccrbf	-0.01[-0.0, 0.01]	-0.04[-0.03, 0.01]	12.87[12.88, 0.19]	0.08[0.11, 0.09]	-0.0[-0.01, 0.01]	11.74[12.12, 2.4]
Ccrsc	-0.05[0.03, 0.31]	1.95[1.85, 1.21]	11.78[14.75, 10.6]	13.28[13.34, 3.96]	11.72[14.38, 9.46]	42.54[42.89, 13.98]
Cq2	0.12[-28.39, 202.72]	0.01[-3.78, 24.13]	10.05[-3753.53, 31924.79]	0.0[-7.08, 53.76]	0.0[-0.34, 13.31]	-0.26[-2312.18, 18970.9]
Cq10	-0.0[0.05, 0.11]	-0.0[-0.03, 0.05]	12.79[13.64, 2.53]	0.02[0.04, 0.04]	-0.0[-0.0, 0.01]	9.4[7.79, 3.24]
Cq50	-0.01[-0.02, 0.01]	-0.01[-0.01, 0.02]	12.63[12.63, 0.13]	0.07[0.11, 0.09]	0.0[0.0, 0.0]	10.93[11.13, 1.51]
Cq90	-0.02[-0.02, 0.01]	-0.06[-0.04, 0.06]	13.05[13.17, 0.51]	0.16[0.26, 0.25]	-0.02[-0.02, 0.02]	16.02[17.66, 5.87]
Evqq	0.03[0.42, 3.41]	7.28[8.16, 5.81]	30.2[30.87, 6.77]	2.75[2.89, 1.96]	1.72[2.71, 3.02]	20.5[20.12, 10.61]
Evqb	0.01[0.03, 0.13]	0.18[0.97, 1.91]	13.57[13.9, 3.73]	0.25[0.79, 1.22]	-0.01[-0.02, 0.24]	18.85[21.03, 12.04]
Ecr	0.0[0.07, 0.6]	0.87[1.88, 2.76]	17.47[17.66, 3.85]	0.5[1.07, 1.44]	-0.01[0.06, 0.53]	19.79[22.49, 13.01]
Ecrqf	0.03[0.42, 3.41]	7.28[8.16, 5.81]	30.2[30.87, 6.77]	2.75[2.89, 1.96]	1.72[2.71, 3.02]	20.5[20.12, 10.61]
Ecrbf	0.01[0.03, 0.13]	0.18[0.97, 1.91]	13.57[13.9, 3.73]	0.25[0.79, 1.22]	-0.01[-0.02, 0.24]	18.85[21.03, 12.04]
Ecrsc	0.58[1.16, 3.11]	1.4[2.34, 3.09]	16.93[21.97, 18.64]	3.64[4.45, 3.15]	28.09[32.35, 21.69]	14.09[22.13, 19.7]
Etre	-0.0[1.08, 8.07]	0.07[0.99, 3.59]	0.73[2.15, 5.65]	0.12[0.54, 1.34]	82.79[71.05, 28.65]	1.88[4.09, 5.82]
Eqp	-0.04[0.12, 1.5]	3.92[4.89, 4.46]	26.82[27.19, 6.71]	1.12[1.48, 1.33]	12.09[14.21, 9.8]	13.11[14.45, 9.0]

Table 3
First-order variance-based sensitivity indices of all studied signatures by the 6 model parameters (in %). The values (in the form of .[.,.]) in each case represent respectively the median, mean and standard deviation of the sensitivity of a signature by a parameter over all catchments.

Method	Calibration metric	$\overline{j_d^{NSE}}$				$\overline{j_d^{KGE}}$				$\overline{j_f^{Q_{peak}}}$			
		Cal	T_Val	S_Val	S-T_Val	Cal	T_Val	S_Val	S-T_Val	Cal	T_Val	S_Val	S-T_Val
CSOO	j_d^{NSE}	0.274	0.277	0.901	0.616	0.239	0.369	0.687	0.736	0.279	0.324	0.387	0.357
	j_d^{KGE}	0.352	0.330	1.048	0.795	0.183	0.323	0.665	0.721	0.267	0.280	0.379	0.344
SSOO	$j_d^{NSE} + j_f^{Q_{peak}}$	0.447	0.418	1.056	0.889	0.377	0.476	0.759	0.853	0.014	0.189	0.346	0.372
	$j_d^{KGE} + j_f^{Q_{peak}}$	0.551	0.431	1.259	0.956	0.335	0.443	0.777	0.833	0.017	0.209	0.337	0.358
SMOO	$\{j_d^{NSE}, j_f^{Q_{peak}}\}$	0.341	0.351	1.020	0.845	0.271	0.420	0.703	0.803	0.087	0.215	0.336	0.391
	$\{J_d^{KGE}, j_f^{Q_{peak}}\}$	0.456	0.409	1.163	0.821	0.243	0.368	0.683	0.724	0.048	0.182	0.316	0.389

Table 4
Calibration, temporal, spatial and spatio-temporal validation metrics for uniform model. The mean of calibration and validation cost values for different objective functions are displayed for each calibration metric.
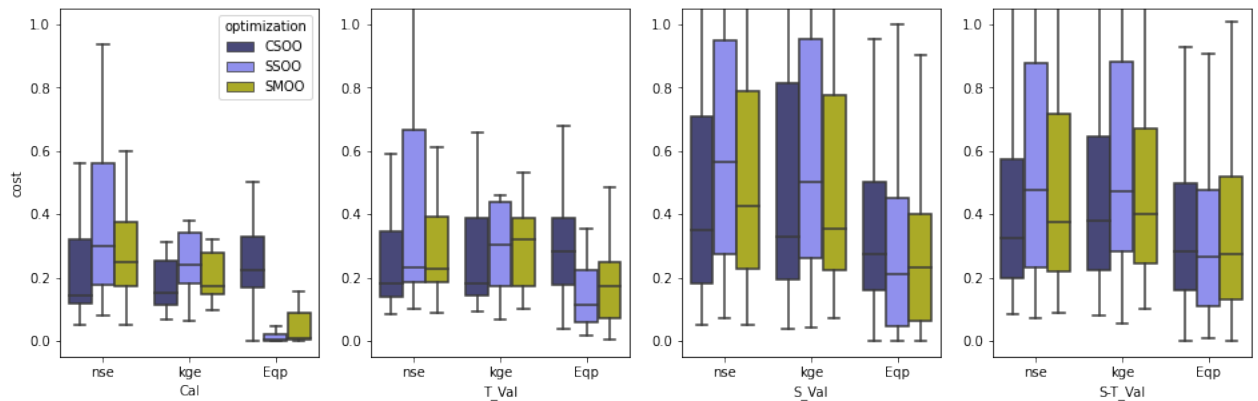

Figure 5: Comparison of calibration and validation metrics (optimal fit $cost = 0$) for three optimization approaches (CSOO, SSOO, SMOO) by constraining $1 - NSE$ in case of global calibration. From left to right: calibration, temporal validation, spatial validation and spatio-temporal validation.

Method	Calibration metric	$\overline{j_d^{NSE}}$				$\overline{j_d^{KGE}}$				$\overline{j_f^{Q_{peak}}}$			
		Cal	T_Val	S_Val	S-T_Val	Cal	T_Val	S_Val	S-T_Val	Cal	T_Val	S_Val	S-T_Val
CSOO	j_d^{NSE}	0.221	0.244	0.655	0.596	0.233	0.355	0.553	0.597	0.274	0.334	0.381	0.376
	j_d^{KGE}	0.239	0.231	0.802	0.702	0.140	0.292	0.617	0.701	0.226	0.295	0.365	0.364
SSOO	$j_d^{NSE} + j_f^{Q_{peak}}$	0.251	0.241	0.831	0.639	0.231	0.305	0.586	0.612	0.183	0.298	0.392	0.383
	$j_d^{KGE} + j_f^{Q_{peak}}$	0.297	0.245	0.964	0.671	0.190	0.300	0.617	0.647	0.152	0.271	0.376	0.387

Table 5
Calibration, temporal, spatial and spatio-temporal validation metrics for distributed model. The mean of calibration and validation cost values for different objective functions are displayed for each calibration metric.
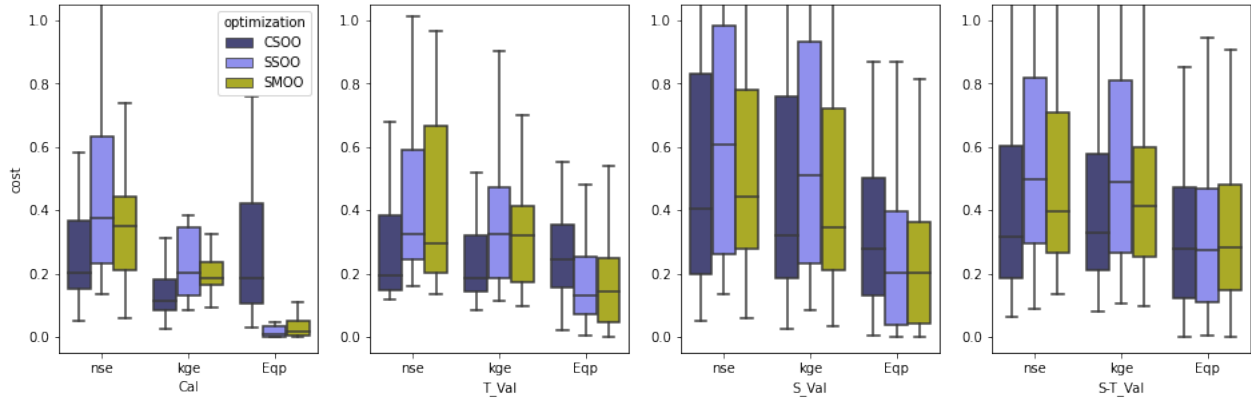
Figure 6: Comparison of calibration and validation metrics (optimal fit $cost = 0$) for three optimization approaches (CSOO, SSOO, SMOO) by constraining $1 - KGE$ in case of global calibration. From left to right: calibration, temporal validation, spatial validation and spatio-temporal validation.
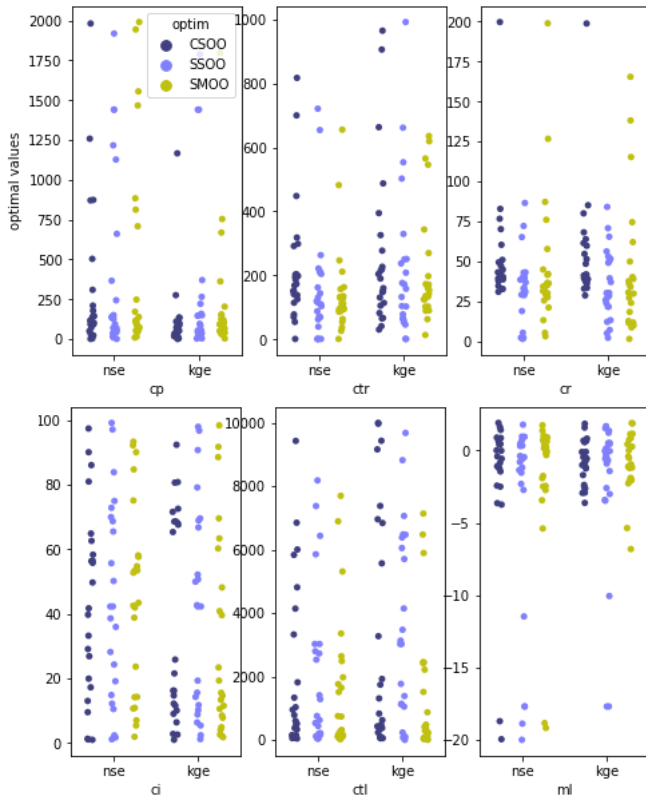

Figure 7: Uniform parameter of the SMASH 6-parameters structure. In each boxplot, the first column present the distribution of a parameter for 3 optimization strategies (CSOO, SSOO, SMOO) using j_d^{NSE}, whereas the strategies in the second column use j_d^{KGE} as the dominant (or constrained) objective function. The boundary conditions of the model parameters are given in Appendix D.

Regarding to the parameter space, Table 6 conveys some statistical analysis over studied catchments for the spatially uniform parameter sets obtained with 4 calibration metrics. Comparing to the spatially distributed optimal parameters in Table 7, we interpret that the mean of distributed parameters

over all catchments in the 4 cases (corresponding to 4 calibration metrics) is globally coherent to the distribution of the first guess. Some parameters are almost spatially uniform for example the non conservative water exchange parameter ml has a small distributed deviation in median (resp. in average) over all catchments 0.01 (resp. 0.05) (calibrated with j_d^{NSE}) compared to its distributed mean in median (resp. in average) -0.59 (resp. -4.98). Conversely, the transfer parameter c_{tl} has a great distributed deviation (in median over all catchments) 193.72 compared to its distributed average 347.87, that also has a massive difference to its distributed median 114.79. Fig. 11 illustrates the spatially distributed optimal parameters at the largest catchment Y5312010, the Argens River, for a distributed calibration with $j_d^{KGE} + j_f^{Q_{peak}}$. The boxplots analyzing the mean and standard deviation of distributed parameter over all studied catchments are displayed in Appendix E.

5. Conclusion

In this study, we enhanced the calibration process for a conceptual distributed hydrological model SMASH, applied to Mediterranean floods, using the hydrological signatures with various multi-criteria optimization strategies. Continuous signatures and flood event signatures are firstly computed and analyzed. Subsequently, sensitivity analysis is used to help selecting appropriate signatures for constraining the model. Ultimately, signatures-based multi-criteria optimization approaches are performed and have demonstrated their robustness and reliability while improving the simulated peak flood event without excessively losing the NSE and KGE. In particular, for distributed calibration, along with preferable results on the signature, the scoring metrics (NSE and KGE) computed on flood events are even enhanced compared to the model calibrated without using the signatures. It shows that the model constrained by the signature has a better performance in simulating flood event and thus more delicate than traditional calibration approaches, specially in the context of flash flood prediction.
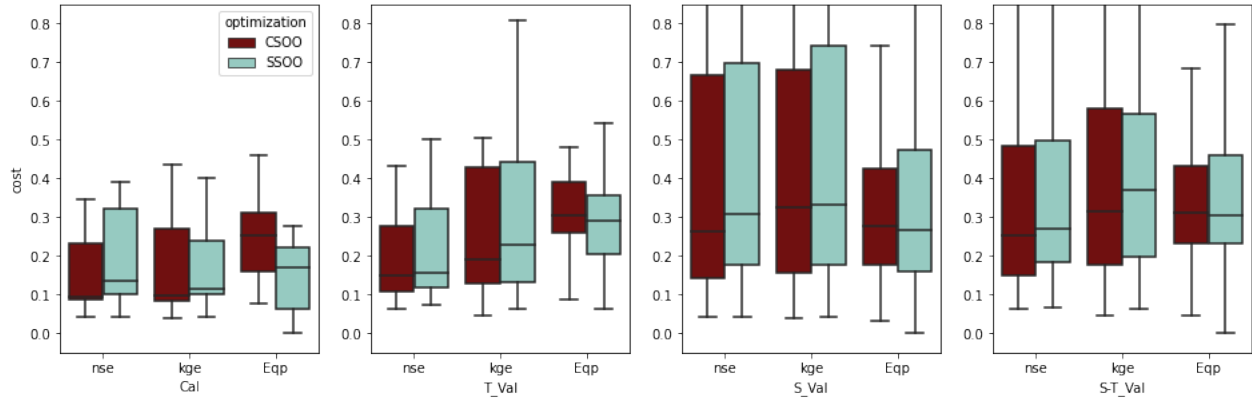
Figure 8: Comparison of calibration and validation metrics (optimal fit $cost = 0$) for two optimization approaches (CSOO, SSOO) by constraining $1 - NSE$ in case of distributed calibration. From left to right: calibration, temporal validation, spatial validation and spatio-temporal validation.
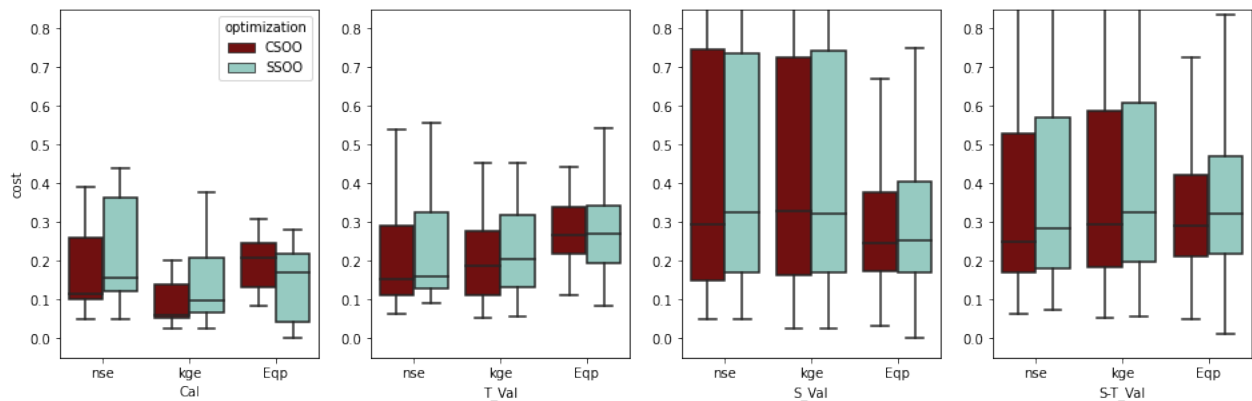

Figure 9: Comparison of calibration and validation metrics for two optimization approaches (CSOO, SSOO) by constraining $1 - KGE$ in case of distributed calibration. From left to right: calibration, temporal validation, spatial validation and spatio-temporal validation.

Interestingly, the parameter spaces in different models are also compared to help understand the optimum movement from traditional calibration approaches to signatures-based calibration approaches.

Future work will aim to (i) upgrade the variational calibration algorithm with Bayesian approach and global sensitivity metrics, including for spatially distributed controls, using adjoint model, as well as (ii) perform extended tests on an even larger sample of catchments and signatures, with complementary data, with various model structures.

Parameter	Calibration metric			
	j_d^{NSE}	$j_d^{NSE} + j_f^{Q_{peak}}$	j_d^{KGE}	$j_d^{KGE} + j_f^{Q_{peak}}$
c_i	14.71[20.3, 26.22]	16.93[20.83, 26.48]	17.6[27.15, 33.07]	17.27[30.17, 35.83]
c_p	169.99[291.17, 434.58]	146.04[310.14, 505.68]	151.87[286.79, 483.33]	141.56[289.69, 466.99]
c_{tr}	171.76[286.6, 269.5]	162.66[313.49, 304.83]	266.32[431.2, 355.04]	267.21[436.83, 360.62]
c_{tl}	347.87[812.15, 1274.12]	250.42[1366.73, 2789.22]	383.51[1413.93, 2749.35]	262.89[1337.96, 2795.16]
c_r	41.32[52.63, 34.05]	40.94[50.97, 30.97]	41.33[51.2, 30.29]	40.24[50.2, 27.58]
ml	-0.59[-4.98, 8.21]	0.0[-3.81, 7.34]	-0.0[-3.62, 7.31]	-0.0[-3.28, 6.39]

Table 6
Uniform optimal parameters calibrated by SBS algorithm with 4 calibration metrics for each catchment on its downstream gauge. The values (in the form of .[.,.]) in each case represent respectively the median, mean and standard deviation of the optimal parameter values over all catchments of the dataset.
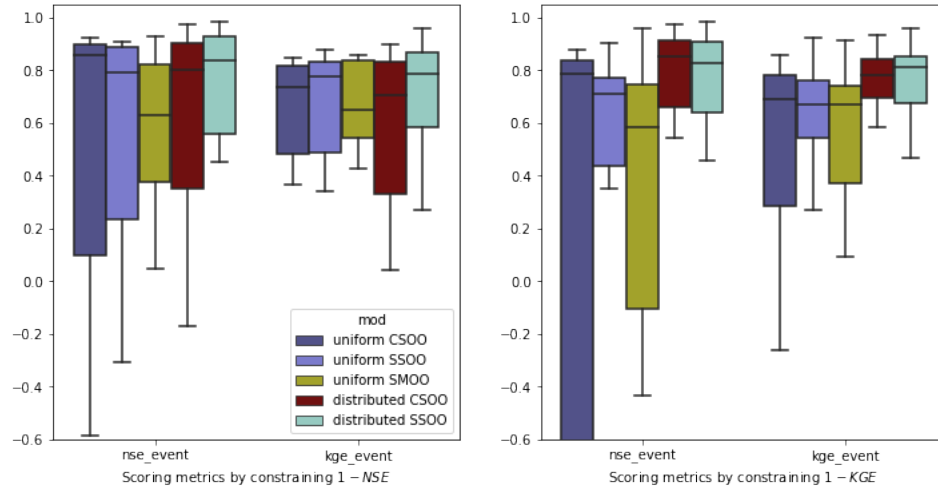
Figure 10: Comparison of scoring metrics computed on 111 events picked from 23 downstream gauges on the period 2006-2013 for all studied models, by constraining $1 - NSE$ (left) and $1 - KGE$ (right).

Parameter	Calibration metric			
	j_d^{NSE}	$j_d^{NSE} + j_f^{Q_{peak}}$	j_d^{KGE}	$j_d^{KGE} + j_f^{Q_{peak}}$
$\tilde{c}_i$	15.45[20.22, 26.34]	10.91[20.14, 26.75]	17.6[27.19, 33.16]	17.3[30.32, 36.05]
$\overline{c}_i$	14.71[20.3, 26.22]	16.93[20.83, 26.48]	17.6[27.15, 33.07]	17.27[30.17, 35.83]
σ_{c_i}	0.22[0.91, 1.46]	0.07[1.13, 4.16]	0.13[0.57, 1.28]	0.05[0.82, 3.09]
$\tilde{c}_p$	161.81[286.05, 435.48]	145.79[314.19, 518.0]	156.65[288.75, 476.53]	148.27[296.94, 485.79]
$\overline{c}_p$	169.99[291.17, 434.58]	146.04[310.14, 505.68]	151.87[286.79, 483.33]	141.56[289.69, 466.99]
σ_{c_p}	38.52[60.58, 59.64]	8.95[37.57, 44.15]	31.08[53.49, 57.35]	12.03[46.82, 102.47]
$\tilde{c}_{tr}$	174.6[287.44, 270.08]	158.48[317.5, 311.03]	266.09[429.0, 353.73]	267.12[447.27, 372.32]
$\overline{c}_{tr}$	171.76[286.6, 269.5]	162.66[313.49, 304.83]	266.32[431.2, 355.04]	267.21[436.83, 360.62]
$\sigma_{c_{tr}}$	13.88[28.84, 35.82]	3.23[25.3, 57.2]	5.68[24.28, 60.74]	1.45[24.39, 74.56]
$\tilde{c}_{tl}$	114.79[675.7, 1276.32]	127.09[1322.22, 2806.59]	180.63[1332.45, 2784.42]	146.96[1322.17, 2803.08]
$\overline{c}_{tl}$	347.87[812.15, 1274.12]	250.42[1366.73, 2789.22]	383.51[1413.93, 2749.35]	262.89[1337.96, 2795.16]
$\sigma_{c_{tl}}$	193.72[355.92, 433.62]	34.67[139.21, 252.41]	69.91[222.54, 388.17]	31.82[61.87, 81.06]
$\tilde{c}_r$	41.37[52.08, 34.42]	41.37[52.08, 34.42]	41.37[52.08, 34.42]	41.37[52.08, 34.42]
$\overline{c}_r$	41.32[52.63, 34.05]	40.94[50.97, 30.97]	41.33[51.2, 30.29]	40.24[50.2, 27.58]
σ_{c_r}	4.66[6.01, 5.04]	1.45[5.17, 10.17]	3.01[5.31, 10.29]	1.34[5.08, 13.82]
$\overline{ml}$	-0.59[-4.98, 8.21]	0.0[-3.72, 7.42]	0.0[-3.61, 7.31]	-0.0[-3.27, 6.39]
$\overline{ml}$	-0.59[-4.98, 8.21]	0.0[-3.81, 7.34]	-0.0[-3.62, 7.31]	-0.0[-3.28, 6.39]
σ_{ml}	0.01[0.05, 0.09]	0.0[0.17, 0.73]	0.02[0.05, 0.09]	0.0[0.08, 0.34]

Table 7

Analysis of spatially distributed parameter sets of the models corresponding to 4 calibration metrics. First, spatial median ($\tilde{\cdot}$), average ($\bar{\cdot}$) and standard deviation ($\sigma_{\cdot}$) for each parameter field are calculated for each catchment, then their median, mean and standard deviation over all catchments are represented in the form of $.[.,.]$.
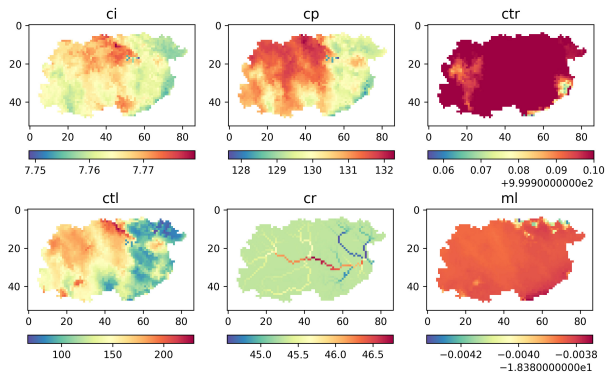

Figure 11: Spatially distributed optimal parameters at the Argens River (Y5312010), by minimizing $j_d^{KGE} + j_f^{Q_{peak}}$.

A. Classical calibration metrics in hydrology

A.1. Nash–Sutcliffe efficiency (NSE)

$$NSE = 1 - \frac{\sum_{t=0}^{T}(Q(t) - Q^*(t))^2}{\sum_{t=0}^{T}(Q^*(t) - \overline{Q^*})^2}$$

where $Q(t)$ is the simulated discharge at time t, $Q^*(t)$ is the observed discharge at time t and $\overline{Q^*}$ is the mean observed discharge.

A.2. Kling–Gupta efficiency (KGE)

$$KGE = 1 - \sqrt{\alpha(r-1)^2 + \beta(\frac{\sigma}{\sigma^*} - 1)^2 + \gamma(\frac{\mu}{\mu^*} - 1)^2}$$

Notation	Signature	Description	Formula	Unit
Ccr		Coefficient relating the amount of runoff to the amount of precipitation received	$\frac{\int^{t\in\mathbf{U}} Q(t)dt}{\int^{t\in\mathbf{U}} P(t)dt}$	-
Ccrqf	Runoff coefficients	Coefficient relating the amount of high-flow to the amount of precipitation received	$\frac{\int^{t\in\mathbf{U}} Qq(t)dt}{\int^{t\in\mathbf{U}} P(t)dt}$	-
Ccrbf		Coefficient relating the amount of low-flow to the amount of precipitation received	$\frac{\int^{t\in\mathbf{U}} Qb(t)dt}{\int^{t\in\mathbf{U}} P(t)dt}$	-
Ccrsc		Coefficient relating the amount of high-flow to the amount of runoff	$\frac{\int^{t\in\mathbf{U}} Qq(t)dt}{\int^{t\in\mathbf{U}} Q(t)dt}$	-
Cq2 Cq10 Cq50 Cq90	Flow percentiles	2%, 10%, 50% and 90%-quantiles from flow duration curve	$quantile(Q(t), 0.02)$ $quantile(Q(t), 0.1)$ $quantile(Q(t), 0.5)$ $quantile(Q(t), 0.9)$	mm

Table 8
List of all studied continuous signatures.

Notation	Signature	Description	Formula	Unit
Evqq	Flood flow	Amount of quickflow in flood event	$\int^{t\in\mathbf{E}} Qq(t)dt$	mm
Evqb	Base flow	Amount of baseflow in flood event	$\int^{t\in\mathbf{E}} Qb(t)dt$	mm
Ecr		Coefficient relating the amount of runoff to the amount of precipitation received	$\frac{\int^{t\in\mathbf{E}} Q(t)dt}{\int^{t\in\mathbf{E}} P(t)dt}$	-
Ecrqf	Runoff coefficients	Coefficient relating the amount of high-flow to the amount of precipitation received	$\frac{\int^{t\in\mathbf{E}} Qq(t)dt}{\int^{t\in\mathbf{E}} P(t)dt}$	-
Ecrbf		Coefficient relating the amount of low-flow to the amount of precipitation received	$\frac{\int^{t\in\mathbf{E}} Qb(t)dt}{\int^{t\in\mathbf{E}} P(t)dt}$	-
Ecrsc		Coefficient relating the amount of high-flow to the amount of runoff	$\frac{\int^{t\in\mathbf{E}} Qq(t)dt}{\int^{t\in\mathbf{E}} Q(t)dt}$	-
Etre	Lag time	Difference time between the peak runoff and the peak rainfall	$\arg\max_{t\in\mathbf{E}} Q(t) - \arg\max_{t\in\mathbf{E}} P(t)$	h
Eqp	Peak flow	Peak runoff in flood event	$\max_{t\in\mathbf{E}} Q(t)$	mm

Table 9
List of all studied flood event signatures.

where r is the linear correlation between observations and simulations, σ and σ^* are the standard deviation in simulations and observations, respectively, μ and μ^* are the mean discharge in simulations and observations, respectively, and α, β, γ are the optimization weight parameters.

B. List of studied signatures

Denote $P(t)$ and $Q(t)$ are the rainfall and runoff at time $t \in \mathbf{U}$, where $\mathbf{U}$ is the studied period. Then $Qb(t)$ and $Qq(t)$ are the baseflow and quickflow computed using a classical technique for streamflow separation (please refer to Lyne and Hollick (1979) and Nathan and McMahon (1990) for more details).

B.1. Continuous signatures

The continuous signatures are calculated on the whole studied period as Table 8.

B.2. Flood event signatures

For an event occurring in $\mathbf{E} \subset \mathbf{U}$, the flood event signatures are calculated as Table 9.

C. Multi-objective optimization for spatially uniform model

We look into multi-objective optimization for a global calibration of spatially uniform parameters of the distributed model $\mathcal{M}$, i.e. a low dimensional control $\overline{\theta}$. The multi-objective calibration is simply defined as the optimization problem:

$$\min_{\theta\in\mathcal{O}\subset\mathbb{R}^n} (j_1(\theta), ..., j_m(\theta)) \tag{8}$$

where θ is the n-dimensional vector of model parameters in the feasible space $\mathcal{O} \subset \mathbb{R}^n$ and $j_1, ..., j_m$ are the m single-objective functions to be simultaneously minimized.

C.1. Pareto front

In single-objective optimization, the Pareto optimal solution is unique (in term of objective space) but in multi-objective problem, it common to have several solutions that can not be defined which one is the best. If the optimization problem is non-dominated, or non-inferior (each objective function is its own entity, so no individual can be better off without making at least one individual worse off), then we call that Pareto optimality, or Pareto efficiency. A Pareto front (in term of parameter space) is a set of all Pareto efficient solutions that need to be estimated. Let us consider two feasible solutions: $\theta_1, \theta_2 \in \mathcal{O}$. Then, θ_1 is said to Pareto dominate θ_2 if the following properties hold:

1. $\forall i \in \{1, ..., m\}, j_i(\theta_1) \leq j_i(\theta_2)$;
2. $\exists i \in \{1, ..., m\}, j_i(\theta_1) < j_i(\theta_2)$.

We call $\mathcal{P}$ the Pareto set representing all of Pareto solutions. By definition, a Pareto solution $\theta^* \in \mathcal{P}$ of problem 8 must fill the two following conditions:

1. $\nexists \theta' \in \mathcal{O} \setminus \mathcal{P}, \exists i \in \{1, ..., m\}, j_i(\theta') < j_i(\theta^*)$;
2. $\nexists \theta'' \in \mathcal{P}, \theta''$ dominates θ^*.

The first statement indicates that there does not exist other point in the feasible space that reduces at least one objective function while keeping others unchanged, so the Pareto set is the optimal set. The second says that, no other point exists in the Pareto set that decreases one objective function without increasing another one, so it is impossible to distinguish any solution as being better than the other in the Pareto set. Fig. 12 illustrates this for a simple problem where we have 2-objective functions j_1, j_2. The Pareto-front (in term of objective space) represents all of non-dominated optimal solutions. It implies that, it is impossible to move from any point in the feasible space and simultaneously decrease the two objective functions without violating a constraint.
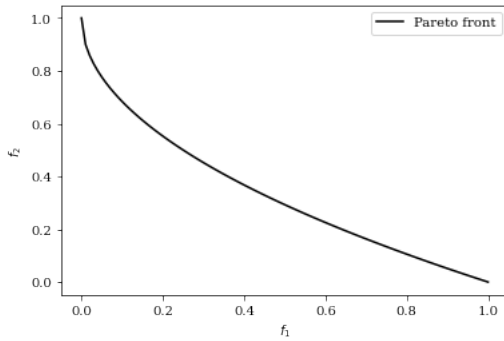

Figure 12: Illustration of Pareto-front in term of objective space.

C.2. Overview of GA

GA is a "heuristic algorithm" (or search heuristic) in optimization, inspired by the Theory of Natural Evolution, whose selection operators include "crossover" and "mutation". Basically, the process of a GA and a MOGA consist of the following 3 phases:

1. *Population initialization.* The population is randomly initialized based on the problem range and constraint. The size of the population determines also the number of solutions, called "pop-size".
2. *Parents selection (sorting).* A fitness function is defined to calculate the fitness score (also called Pareto ranking in multi-objective optimization) that determines how fit an individual is to the problem. Then, the fitness score decides the probability of selecting an individual as a parent to reproduce offspring population.
3. *Mating.* For each pair of parent to be mated, new offspring are created by exchanging the genes of parents among themselves (crossover operator). To maintain the diversity within the population and prevent premature convergence, some of the bits in the gene of certain new offspring can be flipped with a low random probability (mutation operator). Offspring are created until their pop-size is equal to the pop-size of previous generation.

C.3. Selection of an optimal solution from Pareto front

We aim to select an optimal solution that is acceptable for every objective within a constraint on principal objective function. Many strategies can be chosen to perform such a selection, based for instance on the sensitivity ratio that is the ratio of the average variabilities of a certain non-inferior solution to the corresponding value of the objective function in the Pareto front (Wang, Zhao, Wu and Wu, 2017), or the Euclidean distance from the ideal solution (Wang and Rangaiah, 2017). A simple additive weighting (SAW) method in Wang and Rangaiah (2017) can be used in our case by adding a normalization operator and assigned weightage for the objective functions.

Considering an objective matrix $(j_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$, where m is the number of non-dominated solutions, n is the number of objective functions. Then each row i represents the ith solution set of the Pareto-front and each column j represents all non-inferior solutions of the jth objective function. Denote c be the index of the classical objective function (for example $1 - NSE$ or $1 - KGE$), which is the most constrained function to find an unique optimal solution from Pareto-front. This algorithm is detailed in the following three phases:

- Objective matrix normalization $(F_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$:

$$F_{ij} = \frac{f_j^+ - f_{ij}}{f_j^+ - f_j^-}$$

where $f_j^+ = \max_{1 \leq i \leq m} f_{ij}$ and $f_j^- = \min_{1 \leq i \leq m} f_{ij}$.

- Assigning weightage for normalized objective matrix $(G_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$:

$$d = f_c^+ - f_c^-$$

$$w_j = \begin{cases} e^d, & \text{if } j = c \\ e - e^d, & \text{otherwise.} \end{cases}$$

$$G_{ij} = w_j \times F_{ij}$$

- Finding optimal solution θ:

$$\theta = (f_{k1}, ..., f_{kn})$$

where $k = \arg\max_{1 \leq i \leq m} \left(\sum_{j=1}^{n} G_{ij} \right)$.

D. Calibration bounds

The parameter vector of SMASH model structure S6 is $\theta(x) \equiv \left(c_i(x), c_p(x), c_{tr}(x), c_r(x), ml(x), c_{tl}(x) \right)^T$ and bound constrains used in optimization (Eq. 3) are set with values given in Table 10.

	c_i	c_p	c_{tr}	c_r	ml	c_{tl}
Lower boundary	1	1	1	1	-20	1
Upper boundary	100	2000	1000	200	5	10000

Table 10
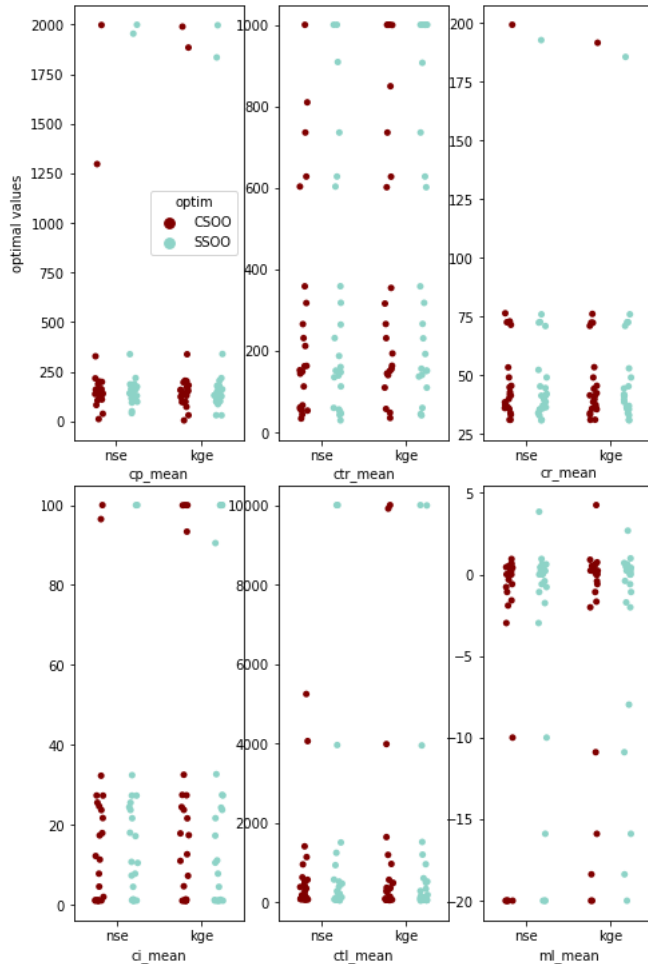Boundary conditions of SMASH 6-parameters model.

Figure 13: Distributed parameter space of the SMASH 6-parameters structure. In each boxplot, the first column present the mean of a parameter over studied catchment for 2 optimization strategies (CSOO, SSOO) using j_d^{NSE}, whereas the strategies in the second column use j_d^{KGE} as the dominant (or constrained) objective function.
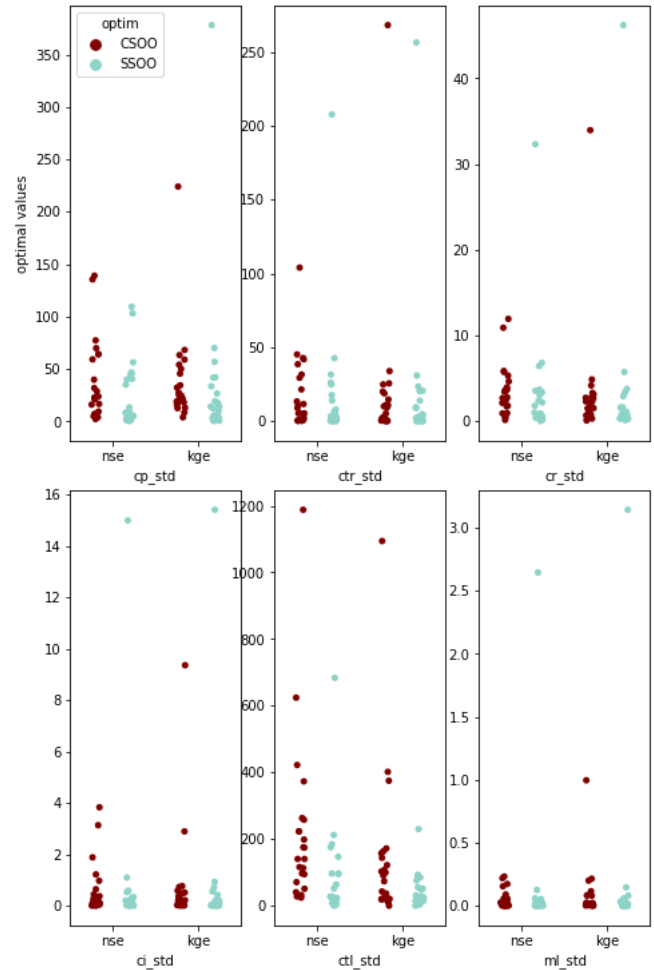

Figure 14: Standard deviation of distributed parameter space of the SMASH 6-parameters structure. In each boxplot, the first column present the std of a parameter over studied catchment for 2 optimization strategies (CSOO, SSOO) using j_d^{NSE}, whereas the strategies in the second column use j_d^{KGE} as the dominant (or constrained) objective function.

E. Analysis of parameter space for distributed calibration

CRediT authorship contribution statement

Ngo Nghi Truyen Huynh: research plan, conceptualization, coding and numerical result, analysis, draft preparation, final redaction. **Pierre-André Garambois:** research plan, conceptualization, analysis, draft preparation, final redaction. **Pierre Javelle:** research plan, conceptualization, analysis, final redaction. **François Colleoni:** conceptualization, coding and numerical results, analysis, final redaction.

References

Abbass, H.A., Sarker, R., Newton, C., 2001. Pde: a pareto-frontier differential evolution approach for multi-objective optimization problems, in: Proceedings of the 2001 congress on evolutionary computation (IEEE Cat. No. 01TH8546), IEEE. pp. 971–978.

Astagneau, P.C., Bourgin, F., Andréassian, V., Perrin, C., 2021. When does a parsimonious model fail to simulate floods? learning from the seasonality of model bias. Hydrological Sciences Journal 66, 1288–1305. URL: https://doi.org/10.1080/02626667.2021.1923720, doi:10.1080/02626667.2021.1923720, arXiv:https://doi.org/10.1080/02626667.2021.1923720.

Azzini, I., Mara, T.A., Rosati, R., 2021. Comparison of two sets of monte carlo estimators of sobol'indices. Environmental Modelling & Software 144, 105167.

Bertalanffy, L.v., 1968. General system theory: Foundations, development, applications. G. Braziller.

Beven, K.J., 2011. Rainfall - Runoff Modelling, The Primer. John Wiley and Sons, LTD.

Blank, J., Deb, K., 2020. pymoo: Multi-objective optimization in python. IEEE Access 8, 89497–89509.

Brigode, P., Génot, B., Lobligeois, F., Delaigue, O., 2020. Summary sheets of watershed-scale hydroclimatic observed data for France. URL: https://doi.org/10.15454/UV01P1, doi:10.15454/UV01P1.

Champeaux, J.L., Dupuy, P., Laurantin, O., Soulan, I., Tabary, P., Soubeyroux, J.M., 2009. Les mesures de précipitations et l'estimation des lames d'eau à météo-france: état de l'art et perspectives. La Houille Blanche , 28–34.

Chibeles-Martins, N., Pinto-Varela, T., Barbosa-Póvoa, A.P., Novais, A.Q., 2016. A multi-objective meta-heuristic approach for the design and planning of green supply chains-mbsa. Expert Systems with Applications 47, 71–84.

Colleoni, F., Garambois, P.A., Javelle, P., Jay-Allemand, M., Arnaud, P., 2022. Adjoint-based spatially distributed calibration of a grid gr-based parsimonious hydrological model over 312 french catchments with smash platform. submitted .

Das, A.K., Nikum, A.K., Krishnan, S.V., Pratihar, D.K., 2020. Multi-objective bonobo optimizer (mobo): an intelligent heuristic for multi-criteria optimization. Knowledge and Information Systems 62, 4407–4444.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: Nsga-ii. IEEE transactions on evolutionary computation 6, 182–197.

Delaigue, O., Génot, B., Lebecherel, L., Brigode, P., Bourgin, P.Y., 2020. Database of watershed-scale hydroclimatic observations in france. URL: https://webgr.inrae.fr/base-de-donnees.

Désidéri, J.A., 2012. Multiple-gradient descent algorithm (mgda) for multiobjective optimization. Comptes Rendus Mathematique 350, 313–318.

Duan, Q., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. Journal of optimization theory and applications 76, 501–521.

Duarte, M., Watanabe, R.N., 2021. Notes on Scientific Computing for Biomechanics and Motor Control. URL: https://doi.org/10.5281/zenodo.4599319, doi:10.5281/zenodo.4599319.

Edijatno, 1991. Mise au point d'un modele elementaire pluie-debit au pas de temps journalier. Ph.D. thesis. Universite Louis Pasteur, ENGEES, Cemagref Antony.

El-Ghandour, H.A., Elbeltagi, E., 2014. Optimal groundwater management using multiobjective particle swarm with a new evolution strategy. Journal of Hydrologic Engineering 19, 1141–1149.

Goel, T., Vaidyanathan, R., Haftka, R.T., Shyy, W., Queipo, N.V., Tucker, K., 2007. Response surface approximation of pareto optimal front in multi-objective optimization. Computer methods in applied mechanics and engineering 196, 879–893.

Guo, J., Zhou, J., Lu, J., Zou, Q., Zhang, H., Bi, S., 2014. Multi-objective optimization of empirical hydrological model for streamflow prediction. Journal of Hydrology 511, 242–253.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. Journal of hydrology 377, 80–91.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research 34, 751–763.

Hascoet, L., Pascual, V., 2013. The tapenade automatic differentiation tool: Principles, model, and specification. ACM Trans. Math. Softw. 39. URL: https://doi.org/10.1145/2450153.2450158, doi:10.1145/2450153.2450158.

Horner, I., 2020. Design and Evaluation of hydrological signatures for the diagnosis and improvement of a process-based distributed hydrological model. Ph.D. thesis. Université Grenoble Alpes.

Jay-Allemand, M., 2020. Estimation variationnelle des paramètres d'un modèle hydrologique distribué. Ph.D. thesis. Aix-Marseille.

Jay-Allemand, M., Colleoni, F., Garambois, P.A., Javelle, P., Julie, D., 2022. Smash - spatially distributed modelling and assimilation for hydrology: Python wrapping towards enhanced research-to-operations transfer. IAHS URL: https://hal.archives-ouvertes.fr/hal-03683657.

Jay-Allemand, M., Javelle, P., Gejadze, I., Arnaud, P., Malaterre, P.O., Fine, J.A., Organde, D., 2020. On the potential of variational calibration for a fully distributed hydrological model: application on a mediterranean catchment. Hydrology and Earth System Sciences 24, 5519–5538.

Khorram, E., Khaledian, K., Khaledyan, M., 2014. A numerical method for constructing the pareto front of multi-objective optimization problems. Journal of Computational and Applied Mathematics 261, 158–171.

Le Mesnil, M., 2021. Signatures Hydrologiques des Bassins Karstiques. Theses. Montpellier SupAgro. URL: https://tel.archives-ouvertes.fr/tel-03578569.

Lyne, V., Hollick, M., 1979. Stochastic time-variable rainfall-runoff modelling, in: Institute of Engineers Australia National Conference, Institute of Engineers Australia Barton, Australia. pp. 89–93.

Masuda, K., Harada, K., Kurihara, K., 2012. Design of a simple evolutionary multiobjective optimization method based on the combined use of scalarization and particle swarm optimization. Electronics and Communications in Japan 95, 1–13.

McMillan, H.K., 2021. A review of hydrologic signatures and their applications. Wiley Interdisciplinary Reviews: Water 8, e1499.

Mercier, Q., Poirion, F., Désidéri, J.A., 2018. A stochastic multiple gradient descent algorithm. European Journal of Operational Research 271, 808–817.

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

Mostafaie, A., Forootan, E., Safari, A., Schumacher, M., 2018. Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily grace data. Computational Geosciences 22, 789–814.

Murata, T., Ishibuchi, H., et al., 1995. Moga: multi-objective genetic algorithms, in: IEEE international conference on evolutionary computation, IEEE Piscataway, NJ, USA. pp. 289–294.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part i—a discussion of principles. Journal of hydrology 10, 282–290.

Nathan, R.J., McMahon, T.A., 1990. Evaluation of automated techniques for base flow and recession analyses. Water resources research 26, 1465–1473.

Oliveira, A.M., Fleischmann, A., Paiva, R., 2021. On the contribution of remote sensing-based calibration to model hydrological and hydraulic processes in tropical regions. Journal of Hydrology 597, 126184.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., Loumagne, C., 2005. Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2 towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. Journal of hydrology 303, 290–306.

Padhye, N., Deb, K., 2011. Multi-objective optimisation and multi-criteria decision making in sls using evolutionary approaches. Rapid Prototyping Journal .

Pujol, N., Neppel, L., Sabatier, R., 2007. Regional tests for trend detection in maximum precipitation series in the french mediterranean region. Hydrological Sciences Journal 52, 956–973. doi:10.1623/hysj.52.5.956.

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. Journal of Applied Meteorology and Climatology 47, 92. doi:10.1175/2007JAMC1636.1.

Ross, M., Abbey, C., Bouffard, F., Jos, G., 2015. Multiobjective optimization dispatch for microgrids with a high penetration of renewable generation. IEEE Transactions on Sustainable Energy 6, 1306–1314.

Roux, H., Labat, D., Garambois, P.A., Maubourguet, M.M., Chorda, J., Dartus, D., 2011. A physically-based parsimonious hydrological model for flash floods in mediterranean catchments. Natural Hazards and Earth

System Sciences 11, 2567–2582.

Sahraei, S., Asadzadeh, M., Unduche, F., 2020. Signature-based multi-modelling and multi-objective calibration of hydrologic models: Application in flood forecasting for canadian prairies. Journal of Hydrology 588, 125095.

Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. Computer physics communications 145, 280–297.

Shafii, M., Tolson, B.A., 2015. Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives. Water Resources Research 51, 3796–3814. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014WR016520, doi:https://doi.org/10.1002/2014WR016520, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014WR016520.

Tavakkoli-Moghaddam, R., Azarkish, M., Sadeghnejad-Barkousaraie, A., 2011. A new hybrid multi-objective pareto archive pso algorithm for a bi-objective job shop scheduling problem. Expert Systems with Applications 38, 10812–10821.

Torres-Treviño, L.M., Reyes-Valdes, F.A., López, V., Praga-Alejo, R., 2011. Multi-objective optimization of a welding process by the estimation of the pareto optimal set. Expert systems with applications 38, 8045–8053.

Tramblay, Y., Neppel, L., Carreau, J., Najib, K., 2013. Non-stationary frequency analysis of heavy rainfall events in southern france. Hydrological Sciences Journal 58, 280–294. doi:10.1080/02626667.2012.754988.

Tramblay, Y., Somot, S., 2018. Future evolution of extreme precipitation in the mediterranean. Climatic Change 151, 289–302. URL: https://doi.org/10.1007/s10584-018-2300-5, doi:10.1007/s10584-018-2300-5.

Veluscek, M., Kalganova, T., Broomhead, P., Grichnik, A., 2015. Composite goal methods for transportation network optimization. Expert Systems with Applications 42, 3852–3867.

Wang, N., Zhao, W.j., Wu, N., Wu, D., 2017. Multi-objective optimization: a method for selecting the optimal solution from pareto non-inferior solutions. Expert Systems with Applications 74, 96–104.

Wang, Z., Rangaiah, G.P., 2017. Application and analysis of methods for selecting an optimal solution from the pareto-optimal front obtained by multiobjective optimization. Industrial & Engineering Chemistry Research 56, 560–574.

Westerberg, I., McMillan, H.K., 2015. Uncertainty in hydrological signatures. Hydrology and Earth System Sciences 19, 3951–3968.

Wu, H., Chen, B., Ye, X., Guo, H., Meng, X., Zhang, B., 2021. An improved calibration and uncertainty analysis approach using a multicriteria sequential algorithm for hydrological modeling. Scientific Reports 11, 1–14.

Wu, Y.K., Liu, C.C., Lur, Y.Y., 2015. Pareto-optimal solution for multiple objective linear programming problems with fuzzy goals. Fuzzy Optimization and Decision Making 14, 43–55.

Yapo, P.O., Gupta, H.V., Sorooshian, S., 1998. Multi-objective global optimization for hydrologic models. Journal of hydrology 204, 83–97.

Yeh, W.C., Chuang, M.C., 2011. Using multi-objective genetic algorithm for partner selection in green supply chain problems. Expert Systems with applications 38, 4244–4253.

Yilmaz, K.K., Gupta, H.V., Wagener, T., 2008. A process-based diagnostic approach to model evaluation: Application to the nws distributed hydrologic model. Water Resources Research 44. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007WR006716, doi:https://doi.org/10.1029/2007WR006716, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2007WR006716.

Zhu, C., Byrd, R.H., Lu, P., Nocedal, J., 1997. Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. ACM Trans. Math. Softw. 23, 550–560. URL: http://dblp.uni-trier.de/db/journals/toms/toms23.html#ZhuBLN97.

Zitzler, E., Thiele, L., 1999. Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. IEEE transactions on Evolutionary Computation 3, 257–271.