On the potential of variational calibration for a fully distributed hydrological model: application on a Mediterranean catchment

Maxime Jay-Allemand[1,2], Pierre Javelle[1], Igor Gejadze[2], Patrick Arnaud[1], Pierre-Olivier Malaterre[2], Jean-Alain Fine[3], and Didier Organde[3]

[1]IRSTEA, 3275 Route de Cézanne, 13182 Aix-en-Provence, France
[2]IRSTEA, 361 rue Jean-François Breton, 34196 Montpellier, France
[3]Hydris-Hydrologie, Parc Scientifique Agropolis II, 2196 Boulevard de la Lironde, 34980 Montferrier sur Lez, France

Correspondence: Maxime Jay-Allemand (maxime.jay-allemand@irstea.fr)

Abstract.

Flash flood alerts in metropolitan France are provided by SCHAPI (Service Central Hydrométéorologique et d'Appui à la Prévision des Inondations) through the Vigicrues Flash service, which is designed to work in ungauged catchments. The AIGA method implemented in Vigicrues Flash is designed for flood forecasting on small- and medium-scale watersheds. It is based on a distributed hydrological model accounting for spatial variability of the rainfall and the catchment properties, based on the radar rainfall observation inputs. Calibration of distributed parameters describing these properties with high resolution is difficult, both technically (in terms of the estimation method), and because of the identifiability issues. Indeed, the number of parameters to be calibrated is much greater than the number of spatial locations where the discharge observations are usually available. However, the flood propagation is a dynamic process, so observations have also a temporal dimension. This must be larger enough to comprise a representative set of events. In order to fully benefit from using the AIGA method, we consider its hydrological model (GRD) in combination with the variational estimation (data assimilation) method. In this method, the optimal set of parameters is found by minimizing the objective function which includes the misfit between the observed and predicted values and some additional constraints. The minimization process requires the gradient of the cost function with respect to all control parameters, which is efficiently computed using the adjoint model. The variational estimation method is scalable, fast converging, and offers a convenient framework for introducing additional constraints relevant to hydrology. It can be used both for calibrating the parameters and estimating the initial state of the hydrological system for short range forecasting (in a manner used in weather forecasting). The study area is the Gardon d'Anduze watershed where four gauging stations are available. In numerical experiments, the benefits of using the distributed against the uniform calibration are analysed in terms of the model predictive performance. Distributed calibration shows encouraging results with better model prediction at gauged and ungauged locations.

1 Introduction

Distributed hydrological models have been introduced to take advantage of high-resolution (both in time and space) data available nowadays, from weather radars and other remote sensing tools. These models come with distributed parameters to better describe both catchment and rainfall spatial characteristics that significantly impact the dis-
5 charge modelling at low scales (Merz and Blöschl, 2009).

In the literature, the distinction between 'conceptual' and 'physically-based' models is often made, even if sometimes the limit is fuzzy. Physically-based models, such as the SHE model (Abbott et al., 1986), tend to describe local and spatial processes as precisely as possible. They involve many physical parameters, such as soil depths, porosities, conductivities, wetting front suctions or roughness coefficients. However, their complexity and the fact
10 that the reality is too heterogeneous to be completely described make their application difficult for an operational use (Beven, 1989).

For this reason, conceptual models (lumped, semi-distributed or fully-distributed) are often preferred in hydrology. Lumped models consider the watershed as a system which has no spatial dimension. They come with several operators describing the spatial-average of local processes. In distributed conceptual models, the catchment is divided into sub-
15 units, either cells displayed on a regular grid (fully-distributed models) in order to take advantage of other gridded input data or sub-catchments (semi-distributed models) in order to respect the hydrological boundaries. Distributed conceptual models can be seen as a good compromise between the detailed representation of the physically-based models and the efficiency of the lumped models. Nowadays, many run operationally in real time, over large areas, for forecasting purpose. Among them, we can cite the CREST model in the United-States (Wang et al., 2011) or the
20 G2G model in the United-Kingdom (Bell et al., 2007). Being conceptual (and not physical), their parameters must be calibrated. But due to the great number of cells (or sub-catchments), these models are largely over-parametrized. This rises both technical and scientific issues, such as parameters equifinality and uncertainty (Beven, 1993).

Methods of data assimilation (DA) have been engaged for several decades in geosciences, including meteorology, oceanography, river hydraulics and hydrology. These methods are used for estimating the driving conditions, param-
25 eters and states of a dynamical model describing the evolution of natural phenomena. The estimates are conditioned on observations (usually incomplete) of a prototype system. Some early applications of DA in hydrology are described in the review paper of (McLaughlin, 1995). It seems that the Kalman filtering has been recently the most popular DA method in hydrology (Sun et al., 2016). For instance, in (Quesney et al., 2000) the Extended Kalman Filter is applied with a lumped conceptual rainfall-runoff model to estimate the soil moisture by assimilating the SAR (syn-
30 thetic aperture radar) data. In (Munier et al., 2014), the standard Kalman Filter is applied with the semi-distributed conceptual model TGR, where the discharge observations are assimilated to adjust the initial model states. It has been shown that the predictive performance depends on the degree of 'spatialization' of the watershed and on the number of gauging stations engaged. In (Sun et al., 2015), the Extended Kalman filter is used with the distributed SWAT model to improve flood prediction on the upstream Senegal river catchment. In this work, given the large

number of state variables, only the spatially-averaged low-resolution updates are estimated. This shows that for DA involving distributed models, scalable methods must be used. The choice of DA methods is, therefore, limited to the variational estimation and the Ensemble Kalman Filter.

In variational estimation, one looks for the minimum of the cost-function using a gradient-based iterative process.

5 The cost-function itself represents the maximum a posteriori (MAP) estimator, which turns into the standard 4D-Var cost-function (Rabier and Courtier, 1992) under the Gaussian assumption. The key component of the method is the adjoint model, which allows the accurate gradient of the cost-function to be computed in a single adjoint run. Then, different minimization methods can be applied. For example, in weather/ocean forecasting, where the models involved are computationally very expensive, the Gauss-Newton method (e.g. 'incremental approach') is used. This

10 method leads to a nearest local minimum in the vicinity of the prior guess. This could be a serious problem if the posterior distribution is multimodal.

In hydrology, the variational estimation method as described above (i.e. including the adjoint model) has not been reported so far. However, similar algorithms have been used. For example, (Abbaris et al., 2014) explored the variational estimation algorithm involving the lumped conceptual HBV model in operational condition. It has been

15 used to update the soil moisture and the states of the routing tank reservoirs on some events. Due to a small number of variables, the gradient is computed by a finite-difference scheme, i.e. without the adjoint model. It has been shown that DA helps to improve peak flow prediction, however the correct choice of the assimilation period and the forecast horizon is vital. In (Thirel et al., 2010), the cost-function is minimized iteratively using the BLUE formulation, which is equivalent to the 'algebraic' form of the Gauss-Newton method. Again, the gradient is computed using the finite-

20 differences and the system matrix (Hessian) is explicitly formed and inverted. Here, DA is implemented involving the SIM model. It has been shown that the improved estimate of the moisture of the soil layers leads to a significantly better discharge simulation.

Calibration of the model parameters is a special case of data assimilation. In hydrology, calibration is very common since the parameters of conceptual models must be somehow defined. Certain past attempts with the local search

25 methods were not always successful and several authors have reported that these methods fail to deliver the global optimal solution (Moradkhani and Sorooshian, 2009), (Abbaspour et al., 2007). However, as we said already, for high-dimensional problem the choice of feasible DA methods is very limited. Let us also note that for high-dimensional, but relatively inexpensive models (such as those common in hydrology), the gradient-enhanced global search minimization methods can be considered (Laurent et al., 2019).

30 An other issue in hydrology concerns flash floods forecasting. This task is very challenging since the conditions leading to these potentially devastating events are still difficult to anticipate (Borga et al., 2010) (Braud et al., 2016). The HyMeX program (Hydrological Cycle in the Mediterranean Experiment) offers a good opportunity to conduct multi-disciplinary studies on this subject (Drobinski et al., 2014). Indeed, the Mediterranean region is particularly affected by flash floods, especially in autumn, when the warm and moistured flux coming from the sea meets inland

35 colder conditions and/or horographic forcing.

In this context, this paper presents the following novel developments. First, the GRD conceptual distributed model used in the so-called "Vigicrues Flash" French warning system (Javelle et al., 2016) is upgraded into a continuous fully distributed model, by introducing a 'cell-to-cell' routing scheme. Second, this new model is calibrated using a variational estimation algorithm, including the adjoint technique for computing the gradient. This algorithm is

5 upgraded to include the inequality constraints and scaling. The predictive performance of the calibrated model is evaluated by cross validation, using the Gardon d'Anduze watershed, in the French Mediterranean region.

The paper is organised as follows. In Section 2.1 the modified GRD model is described. In Section 2.2 we present the variational estimation algorithm adapted for the parameter calibration purpose. The testing benchmark is described in Section 2.3, and the testing methodology in Section 2.4. The results are presented in Section 3, followed by the

10 discussion and conclusions sections.

2 Methodology and data

2.1 Distributed (continuous) rainfall-runoff model GRD

The GRD model (i.e. GR 'Distributed') is a conceptual distributed hydrological model (Javelle et al., 2010) (Arnaud et al., 2011) (Javelle et al., 2014). It belongs to the 'GR' (Génie Rural) family, which includes several other bucket

15 style models, lumped or semi-distributed, developed in the last 20 years (Perrin et al., 2003), (Mouelhi et al., 2006), (Lobligeois et al., 2014), (Ficchì et al., 2016), (Santos et al., 2018), (Riboust et al., 2019).

Since march 2017, the GRD model runs operationally into the national French flash flood warning system called 'Vigicrues Flash'. As described by (Javelle et al., 2016), for this operational application, GRD runs on an 'event-based' mode, with a very simple transfer function. Furthermore, its parameters are uniform over supposed large

20 homogeneous areas (nine classes in France).

The present paper presents a new version of the GRD model: it is continuous, and cells are connected by the means of a 'cell-to-cell' routing scheme. From an operational point of view, these improvements have the following advantages: 1) the model does not require anymore to be initialised before each event, and 2) forecasts can be issued at any pixel, and not at predefined outlets as it was the case in the previous version.

25 The inputs are the radar precipitation estimates provided by Météo-France and the spatial potential evapotranspiration computed from the Oudin formula, based on the temperature (Oudin et al., 2005). The output is the discharge calculated at at any node (pixel) of the routing scheme. In the present case study, the model runs at an hourly time step on a regular $1\ km^2$ grid.

For the convenience of the reader, the elements of the model are briefly described below.

30 In each pixel, the model contains two reservoirs: a production store p and a transfer store t. Then the runoff generated in each pixel is routed from pixel to pixel (figure 1). Thus, only 3 parameters need to be defined in each cell (pixel): the capacity of the production reservoir c_p, the capacity of the transfer reservoir c_t and the local routing

velocity v.

The water balance function :

In each pixel, a water balance function determines the effective rainfall, i.e. the amount of rainfall that will produce
5   a runoff, noted P_r. This operation is carried out following several steps.

First, net rainfall P_n and net potential evapotranspiration E_n are defined from the following equations:

$$if\, P \geq E,\ P_n = P - E\ ,\ and\ E_n = 0 \tag{1}$$

$$if\, P < E,\ E_n = E - P\ and\ Pn = 0 \tag{2}$$

10   Then, the production store is filled by P_p, a part of P_n, representing the part of rainfall infiltrating in the soil.
In the same manner, the production store is emptied by E_p, a part of E_n, representing the actual evaporation. The
variation of the level h_p into the store is driven by the following differential equation (Edijatno, 1991):

$$dh_p = \left[1 - \left(\frac{h_p}{c_p} \right)^2 \right] dP_n - \frac{h_p}{c_p} \left(2 - \frac{h_p}{c_p} \right) dE_n \tag{3}$$

Assuming a stepwise approximation of input variables $P(t)$ and $E(t)$, equation (3) can be integrated over one time
15   step Δt to obtain the amount P_p filling the store and the amount E_p evapored from the store.

$$P_p = c_p \left(1 - \left(\frac{h_p}{c_p} \right)^2 \right) \frac{tanh\left(\frac{P_n}{c_p} \right)}{1 + \left(\frac{h_p}{c_p} \right) tanh\left(\frac{P_n}{c_p} \right)}, \tag{4}$$

$$E_p = h_p \left(2 - \frac{h_p}{c_p} \right) \frac{tanh\left(\frac{E_n}{c_p} \right)}{1 + \left(1 - \frac{h_p}{c_p} \right) tanh\left(\frac{E_n}{c_p} \right)}. \tag{5}$$

It should be noted that with this discret formulation, h_p is the level of the store at the begining of Δt, P_p and E_p
are the volume of water gained or lost by the store, over Δt. At the end of Δt, h_p will be updated by adding $P(t)$
20   and removing $E(t)$, before moving to the next time step.

Finally, P_r, is the remaining rainfall, i.e. the part of rainfall that is not entering into the production store. It is
noted:

$$P_r = P_n - P_p \tag{6}$$

One can see, that the state of the production reservoir h_p plays the role of the humidity state of the soil. An
25   empty store ($h_p = 0$) means that the soil is completely dry: no runoff and no evapotranspiration are produced, and

all the rainfall is absorbed ($E_p = 0$, $P_p = P_n$, $P_r = 0$). On the contrary, a full store ($h_p = c_p$) means that the soil is completely saturated: the evapotranspiration is maximal, all the rainfall contribute to the runoff and the soil reaches its maximal absorption capacity ($E_p = E_n$, $P_p = 0$, $P_r = P_n$)

5   *The transfer function (within a pixel) :*

P_r fills the second store of the model, a transfer store t with a c_t capacity. The outflow from the transfer store gives the elementary flow componant q (more precisely, q is a volume emitted by the reservoir during the time period Δt). This transformation is modeled by a conservative operator which is derived from the differential equation describing the evolution of the state h_t of the transfer reservoir and the mass conservation condition.

10    $$\frac{dh_t}{dt} + c_t h_t^\alpha = P_r, \tag{7}$$

It has been noticed (Michel C., 1989) that equation (7) correctly replicates the flooding and drying processes for $\alpha = 5$. This is an empirical knowledge which has no physical proof. Assuming P_n is the impulse function, equation (7) is integrated over one time step Δt to obtain the expression for q:

$$q = h_t - (h_t^{-4} + c_t^{-4})^{-0.25} \tag{8}$$

15   More details about the production and the transfer reservoirs can be found in (Perrin et al., 2003) and (Michel C., 1989).

The routing function (pixel-to-pixel) :

The run-off is modelled by a third operator (routing model), propagating the flow through the basin. This model
20   is build on top of a digital elevation model, which defines the runoff directions between the routing nodes. Presently, the routing nodes are placed at the center of the corresponding transfer reservoir cells. For the sake of simplicity we describe the routing model in the one-dimensional setting. The runoff from node $i-1$ to node i is delayed by a time

$$\tau_i = \frac{d_i}{v_i}, \ i = 1, \ldots, N, \tag{9}$$

where d_i and v_i are, respectively, the distance and the routing velocity between these nodes.
25   In the simplest implementation, the node output discharge (more precisely, the mass over the time step Δt) is given as

$$Q_i(t) = q_i(t) + Q_{i-1}(t - \tau_i(v_i)), \ i = 1, \ldots, N. \tag{10}$$

Since no explicit model for Q is provided, Q is not explicitly differentiable with respect to v. That is why the above formulation is not suitable for variational data assimilation, which requires the gradient of the cost function has to

be computed. In order to achieve the differentiability we represent the second term in equation (10) in the integral form as follows

$$Q_{i-1}(t - \tau_i(v_i)) = \int\limits_{t'=-\infty}^{t} Q_{i-1}(t')\,\delta(t' - \tau_i(v_i))\,dt'. \tag{11}$$

Next, instead of δ-function we use the unscaled Gaussian function, i.e.

$$5 \quad Q_{i-1}(t - \tau_i(v_i)) = \int\limits_{t'=-\infty}^{t} Q_{i-1}(t')\,\omega(t' - \frac{d_i}{v_i}, \sigma)\,dt', \tag{12}$$

where

$$\omega(t, \sigma) = exp(-\frac{t^2}{2\sigma^2}). \tag{13}$$

It is easy to see that function (12) explicitly depends on v_i via ω, therefore the gradient of Q_{i-1} with respect to v_i can be computed. Assuming $Q(t)$ is a constant during a time step period Δt, equation (12) can be written in the

10   discrete form as follows:

$$Q_{i-1}(t - \frac{d_i}{v_i}) = \sum_{k=0}^{K} \bar{\beta}_{i,k}\, Q_{i-1}(t - k\Delta t), \tag{14}$$

where

$$\bar{\beta}_{i,k} = \frac{\beta_{i,k}}{\sum_{k=1}^{K} \beta_{i,k}},$$

and

$$15 \quad \beta_{i,k} = w(t - \frac{d_i}{v_i} - k\Delta t, \sigma),\ i = 1, \ldots, N,\ k = 1, \ldots, K.$$

For the given estimate of routing velocities v_i, the coefficients $\beta_{i,k}$ does not change with time and, therefore, can be pre-computed and saved in memory. In order to avoid instability the spread parameter $\sigma = 0.5$ is used in computations. In terms of using the exponential weights the presented routing model resembles the Lag and Route (LR) model described in (Laganier et al., 2014) and (Tramblay et al., 2010). However, the Gaussian function

20   represents the hydraulic response function in a more realistic way (without the initial shock). Besides, unlike our routing model, the mentioned models employ the direct 'cell-to-output' routing, in which case the discharge at the 'ungauged' parts of the basin cannot be estimated.

2.2   Variational calibration algorithm

Calibrating a distributed model is often difficult due to a number of reasons. First, the total number of sought

25   parameters can be quite large (high dimensionality). This strictly limits the choice of suitable inference method-ologies. Second, there is an identifiability issue given the sparsity of observations in space. This can be (partially)
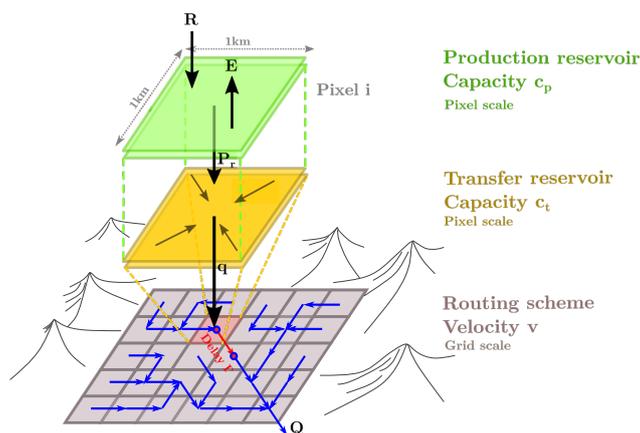
Figure 1. General outlines of the GRD model.

compensated by increasing the observation period or observation frequency to better analyze the system dynamics. However, running the model over a very long assimilation window can also be quite costly.

For distributed models the variational estimation algorithm is often a natural choice. It is perfectly scalable, i.e.

5 it works efficiently for practically unlimited size of the control vector. That is why this method (branded as 4D-Var) is commonly used in meteorology and oceanography for operational forecasting and reanalysis (Ledimet and Talagrand, 1986), (Rabier and Courtier, 1992). The method provides the exact mode of the posterior distribution by minimising the cost-function defined over the full observation window. The key element of the method is the adjoint model which provides the precise gradient of the cost-function with respect to all elements of the control vector in a

10 single run (Errico, 1997). This allows the efficient, fast converging gradient-based minimization methods to be used, such as the BFGS or Newton-type. Quite often, the need for development of the adjoint model becomes an obstacle for practical implementation of this method. Heuristic methods such as the Nelder-Mead algorithm do not require the gradient to evaluate the descent directions, but converge slowly and are not suitable for solving problems in high dimensions. The same is true as for the general purpose statistical methods such as the Markov Chain Monte

15 Carlo (e.g. Metropolis-Hastings algorithm), so for the methods specially designed for hydrology applications,such as SUFI-2 (Abbaspour et al., 2007).

Let us consider a 2D-spatial domain (basin) Ω. Let us represent the hydrological model, described in Section 2.1, as an operator A mapping the inputs $P(x,t)$ and $E(x,t)$ into the discharge $Q_k(t)$ at the observation points

8

$x_k \in \Omega, \ k = 1, N_s$:

$$Q_k(t) = A(P(x,t), E(x,t), h(x,0), p(x)), \ x \in \Omega, \ t \in (0,T), \ k = 1, N_s, \tag{15}$$

where $h(x) = (h_p(x), h_r(x))^T$ is the state vector which includes the states of all production and transfer reservoirs at time $t = 0$, and $p(x) = (c_p(x), c_r(x), v(x))^T$ is the parameter vector which includes the corresponding capacities
5 and the routing velocities at all routing nodes. If the observation period is much longer than the characteristic time of the system (which is the case for calibration/re-analysis), one can use the trivial initial state $h(x,0) = 0$, but consider the observation window $t \in (t^*, T)$, where t^* is the relaxation period. Given the observed inputs $P^*(x,t)$ and $E^*(x,t)$ and the output $Q_k^*(t)$, the calibration cost-function can be defined as follows:

$$J(p) = \int\limits_{t=t^*}^{T} \sum_{k=1}^{N_s} O^{-1/2}\left(A(P^*, E^*, 0, p) - Q_k^*\right)^2 dt + \alpha \|B^{-1/2}(p - p^*)\|_{L^2}^2, \tag{16}$$

10 where O is the observation error covariance, B is the background error covariance, p^* is a prior guess on p, which comes from special measurements, land expertise or a modeling, and α is the regularization parameter. This is more or less standard variational data assimilation (4D-Var) cost-function. The weight α is additionally introduced to mitigate the uncertainty in p^* and B.

Let us note that for the short-range forecasting (T comparable to the characteristic time of the system), the
15 parameter vector is likely to be fixed at its optimal value p_a and the initial state of reservoirs $h = h(x,0)$ will serve as a control vector. In this case, the cost-function looks as follows:

$$J(h) = \int\limits_{t=0}^{T} \sum_{k=1}^{N_s} O^{-1/2}\left(A(P^*, E^*, h, p_a) - Q_k^*\right)^2 dt + \alpha \|B^{-1/2}(h - h^*)\|_{L^2}^2, \tag{17}$$

where h^* is the background value of h. However, this paper is focused on the parameter calibration problem involving long time series of observations, thus formulation (16) is considered.

20 We use additional constraints in the form

$$p_{min} < p < p_{max}, \tag{18}$$

where p_{min} and p_{max} are the bounds which come from the empirical knowledge or physical considerations. Thus, the optimal estimate of the parameters p_a is obtained from the condition

$$p_a = \underset{p}{\arg\min} \, J(p), \tag{19}$$

25 given constraints (15) and (18).

Matrix B can be represented in the form $B = \sigma_p \cdot I \, C \, \sigma_p \cdot I$, where σ_p is the vector of mean deviations of p, C is the correlation matrix, I - the identity matrix, and '$\cdot$' stands for the elementwise (Hadamar) product. Next, the scaling

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

of parameters is introduced, such that $p = p_{min} + \tilde{p}(p_{max} - p_{min})$. Then, the penalty term in (16) takes the form

$$\alpha \|(p_{max} - p_{min}) \cdot \sigma_p^{-1} \cdot I\, C^{-1/2}(\tilde{p} - \tilde{p}^*)\|_{L^2}^2.$$

Assuming that $(p_{max} - p_{min}) \cdot \sigma_p^{-1} = const$, the cost-function (16) reads as follows:

$$J(\tilde{p}) = \int\limits_{t=t^*}^{T} \sum_{k=1}^{N_s} O^{-1/2} \left(A(P^*, E^*, 0, p) - Q_k^*\right)^2 dt + \alpha \|C^{-1/2}(\tilde{p} - \tilde{p}^*)\|_{L^2}^2, \tag{20}$$

5 given

$$p = p_{min} + \tilde{p}(p_{max} - p_{min}), \ 0 < \tilde{p} < 1. \tag{21}$$

The results presented in this paper correspond to the simplest approach to regularization: we assume that $O = I$, $C = I$, and the regularization parameter is chosen a-priori as a small value ($\alpha = 10^{-4}$) to insure well-posedness of the calibration problem. More sophisticated approaches for regularization (non-trivial correlation matrix C, a-posteriori

10 choice of α using the L-curve approach) have been tried (Jay-Allemand et al., 2018), but not presented in this paper for the sake of simplicity.

Minimization of (20) given constraints (21) is performed by LBFGS-B (Limited memory Broyden-Fletcher-Goldfarb-Shanno Bound-constrained (Zhu et al., 1994)). The minimization process can be written in the form

15 $$\tilde{p}_{i+1} = \tilde{p}_i + \beta H^{-1}(p_i) P[J'_{\tilde{p}}(p_i)], \ i = 0, 1, \ldots, \ \tilde{p}_0 = \tilde{p}^*, \tag{22}$$

where $J'(p_i)$ and $H^{-1}(p_i)$ are the gradient (with respect to $\tilde{p}$) and the limited-memory inverse Hessian of (20) at point p_i, respectively, i is the iteration number, and P is the gradient projection operator to account for the box constraints. Let us note that $H^{-1}(p_i)$ is directly computed inside the minimisation algorithm in such a way that its norm is always bounded. This serves as an additional regularization, thus the solution p_a is always bounded, even

20 for $\alpha = 0$ in (20), i.e. even without the penalty term. The gradient $J'(p_i)$ is obtained by solving the adjoint model. This model has been generated by the Automatic Differentiation engine Tapenade (Hascoet and Pascual, 2013), then manually optimised and, finally, verified using the standard gradient test.

The background value p^* is used both as a starting point for iterations and in the penalty term. Given the fact that the information content of the test signal (rainfall) and observations (discharge) may not be sufficient to uniquely

25 resolve the distributed coefficients, evaluating an appropriate p^* becomes an important issue.

Let us consider an approximation: $c_p(x) = \bar{c}_p$, $c_r(x) = \bar{c}_r$, $v(x) = \bar{v}$, $\forall x \in \Omega$. In this case, which shall be referred below as '**uniform calibration**', the control vector $\bar{p} = (\bar{c}_p, \bar{c}_r, \bar{v})^T$ consists just of three elements. For such low-dimensional control estimating the global optimal solution, as well as the error bounds, is feasible by a variety of methods. We use a stochastic method to calibrate uniformly these parameters to ensure finding the best solution

30 in a global sense. The purpose of calibrating $c_p(x)$, $c_r(x)$, $v(x)$, $\forall x \in \Omega$ is to allow the spatial variability of these

coefficients. This shall be referred below as 'distributed calibration', the control vector $p = (c_p(x), c_r(x), v(x))^T$ consists of $3 \times N$ elements. It must be mentioned that since the model $A(\cdot)$ is nonlinear the solution obtained by (22) may not be the best possible solution in a global sense.

Finally, the stability of the calibrated set of parameters is checked by comparing the optimal solutions which are obtained by running the minimisation process with randomly perturbed p^*.

2.3 Study area

A French Mediterranean watershed, the Gardon of Anduze, has been considered for testing our model and calibration algorithm. This area has been involved in many hydrological studies made in the framework of the HYMEX (Drobinski et al., 2014), which aims to better understand the flash floods dynamic in Mediterranean areas. For instance, the FLOODSCALE project (Braud et al., 2014) enabled researchers to exploit a number of very detailed field measurements during severe storm events. Other studies involved a physically based distributed model (MARINE) in order (Roux et al., 2011), (Garambois et al., 2013), (Garambois et al., 2015), (Douinot et al., 2016), (Douinot et al., 2018).Other conceptual distributed models were also tested in this area, such as those implemented into the ATHYS platform (Bouvier and DelClaux, 1996), (Laganier et al., 2014), (Tramblay et al., 2010).

The main properties of the Gardon d'Anduze are described in (Darras, 2015). In brief, this is a steep mountainous watershed with a dense hydrographic network spreading over $540 \ km^2$ in the East part of the Cévennes mountain (France). The difference in levels between the highest elevation point and Anduze is about 800 meters and the slope reaches 50% in the upstream part. Metamorphic but fractured geological formation dominates the watershed. The ground thickness is about 30 cm in average. Water infiltrates very quickly (the saturated hydraulic conductivity is greeter than $200 \ mm.h^{-1}$) and the water circulation appends mainly underground. This area is governed by a transitional Mediterranean-Oceanic climate with warm and dry summers, alleviated by the oceanic influence, followed by recurrent short, intense but persistent heavy rainfalls in autumn and winter, known as "épisode méditerranéen", which generate flash floods. This watershed is well gauged: at least four stations with continuous data collection are operational here (see Fig.2 and Table 1). For numerical experiments, the discharge data have been extracted from the HYDRO French database (http://www.hydro.eaufrance.fr/), and the rainfall data - from the radar observations analysis provided by Météo-France for the period 2008-2015.

2.4 Testing methodology

The variational algorithm described in Sect. 2.2 is applied to the hydrological model presented in Sect. 2.1, using the Gardon d'Anduze watershed as a benchmark. The problem is considered in the rectangular spatial domain (total area $1600 \ km^2$) overlapping the watershed. The domain is covered by a uniform $1 \ km \times 1 \ km$ rectangular grid. The number of 'active' cells is about 540, so the total number of parameters to be calibrated is 3×540. Let us note that the Gardon d'Anduze watershed can be regarded as a small one. The rainfall and discharge data are available for the seven-year long period from 01/01/2008 to 01/01/2015.
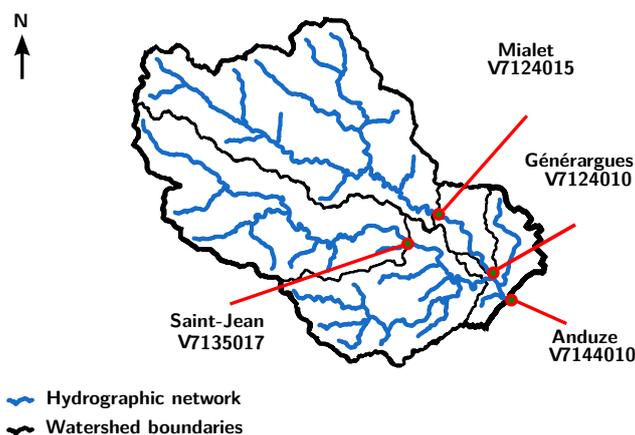
Figure 2. The Gardon watershed at Anduze: hydrographic network (blue) and gauging stations V7124015, V7124010, V7135017, V7144010 (red).

Table 1. Characteristics of the four gauging stations on the Gardon watershed.

Rivers and station names	Codes	Surfaces (km^2)
The Gardon de Mialet at Mialet	V7124015	219.7
The Gardon de Mialet at Générargues	V7124010	244.1
The Gardon de Saint-Jean at Saint-Jean-du-Gard	V7135017	157.7
The Gardon d'Anduze at Anduze	V7144010	540.8

The calibrated model validation step consists in checking the model predictive performance over the data not involved in calibration. That is, the full set of observations $Q_k^*(t)$, $k = 1, \ldots, N_s$, $t \in (0, T)$ is divided in two complementary subsets: calibration subset and validation subset. Since Q^* depends on k (defines the spatial distribution of sensors) and t we distinguish the temporal, spatial, and spatio-temporal validation. In particular, we divide the whole period in two parts: $P1$ - from 01/01/2012 to 01/01/2015, and $P2$ - from 01/01/2009 to 01/01/2012. Each period $P1$ and $P2$ can be considered as calibration or validation period. A model warm-up of one year long is performed before starting the simulations. Four gauging stations are located in the watershed (2):Anduze, Générargues, Mialet and Saint-Jean. If data from a station is used in calibration, the corresponding catchment is called the "calibration catchment", otherwise it is call the "validation catchment".

The calibration quality and the model predictive performance are measured using the Nash-Sutcliffe (NS) criteria ((Nash and Sutcliffe, 1970)). We shall refer to:

a) "**calibration**" - if the NS criteria has been computed over the calibration period for all calibration catchments;

b) "**temporal validation**" - if the NS criteria is computed over the validation period for all calibration catchments;

5 c) "**spatial validation**" - if the NS criteria is computed over the calibration period for all validation catchments;

d) "**spatio-temporal validation**" - if the NS criteria is computed over the validation period for all validation catchments.

The following numerical experiments have been performed:

1. **calibration uniform-4-sta** / **calibration distributed-4-sta** - uniform and distributed calibration, respectively,

10 using observations from all four gauging stations and two time periods $P1$ or $P2$;

2. **temporal validation uniform-4-sta** / **temporal validation distributed-4-sta** - uniform and distributed validation, respectively, using observations from all four gauging stations. The model calibrated on data from $P1$ is validated on data from $P2$, and vice versa;

3. **calibration uniform-1-sta** / **calibration distributed-1-sta** - uniform and distributed calibration, respectively,

15 using observations from one downstream gauge station (Anduze), time periods $P1$ or $P2$;

4. **spatial validation uniform-1-sta** / **spatial validation distributed-1-sta** - uniform and distributed spatial validation, respectively. The model calibrated on data from Anduze gauge station, validated on data from Générargues, Mialet and Saint-Jean gauge stations, for the same time periods $P1$ or $P2$;

5. **spatio-temporal validation uniform-1-sta** / **spatio-temporal validation distributed-1-sta** - uniform and

20 distributed spatio-temporal validation, respectively. The model calibrated on data from Anduze gauge station, validated on data from Générargues, Mialet and Saint-Jean gauge stations, but for different time periods: e.g calibrated on $P1$, validated on $P2$, and vice-versa.

The major purpose of the experiments is to compare the predictive performance of the models based on "dis-

25 tributed calibration" against those based on "uniform calibration". First, we compare the models ability to predict the discharge at observation points (experiments 1, 2); then the ability to resolve the spatial distribution of discharge, i.e. outside the observation points (experiments 3, 4) and, finally, the stability of predictions in spatio-temporal dimension (experiment 5).

3   Results

30 Fig.3 shows the results of calibration (exp.1) and temporal validation (exp.2). The left panel shows the results associated with the uniform calibration, the right panel - with the distributed calibration. When calibration is performed for the period P1, the results are validated for the period P2, and vise versa. All four gauging are involved, thus we have 4×2 calibration / validation points. The NS criteria is computed at these points and ranked

in the increasing value order. By comparing the calibration results (shown by "+") presented at these two panels one can see that the distributed calibration allows much better approximation of the observed discharge than the uniform one. This result is anticipated and simply confirms that the data assimilation procedure works correctly. We also notice that the distributed calibration is less "stable" than the uniform calibration. Indeed, performances

5 gaps between calibration and validation are larger with the distributed set of parameters. More importantly, the validation results (shown by "×") confirm that the model temporal predictive performance is noticeably better if the distributed calibration (red) is used. For convenience, the validation results are also presented in Fig.4 (left). Let us notice that the latter result has been attained despite the rainfall patterns for P1 and P2 being quite different (P1 covers 'wet years', and P2 - 'dry years').

10 Fig.4(right) shows the results of spatial validation (exp.4). Only data from the downstream gauging station (Anduze) is used for calibration (exp.3). In terms of the NS criteria the calibration results are as follows: uniform - 0.77/0.80, distributed - 0.89/0.93, for P1/P2 respectively. Data from the remaining three gauge stations are used for validation, thus we have 3×2 validation points shown in the figure by "○"(to distinguish from temporal validation results). One can see that the model spatial predictive performance is also better if the distributed calibration (red)

15 is used, with one exception. We have to notice here that the spatial predictive performance seems less significant that the temporal one. This depends, however, on the spatial variability of the test signal (rainfall). Since the basin under investigation is relatively small, this result is not surprising.

Fig.5 shows the results of spatio-temporal validation (exp.5). As before, only data from the downstream gauging station (Anduze) is used for calibration (exp.3), but data from all four gauge stations from a different time period

20 is used for validation, giving 4×2 validation points. Two of them (shown in "×")) replicate the temporal validation from Fig.4(left), other six (shown in "⊗")) stand for spatio-temporal validation. One can see that the spatio-temporal predictive performance is, again, better if the distributed calibration (red) is used (for seven out of eight points).

Fig.6 and Fig.7 represent the maps of calibrated parameters related to experiments 1 and 3, respectively. Comparing the left and right panels, which correspond to different time periods $P1$ and $P2$, one can notice a significant

25 inconsistency between the calibrated parameters for capacities C_p and C_t. This effect can be attributed to quite a different rainfall pattern over the reference periods. Besides, the successful estimation of the distributed parameters depends on the information content of the test signal. If this content is not sufficient, one could get the uniqueness issue (equifinality). As a possible remedy one may consider the idea of a pooled analysis, i.e. calibrating the model independently for different hydrological regimes (dry, medium and wet). It is interesting to see that, regarding the

30 routing velocity v, the maps are rather similar and are related to the network drainage (see Fig.2).

4 Discussion

The validation results presented above generally confirm that the distributed calibration improves the temporal, spatial, and spatio-temporal predictive performance of the GRD model as compared to the uniform calibration. For
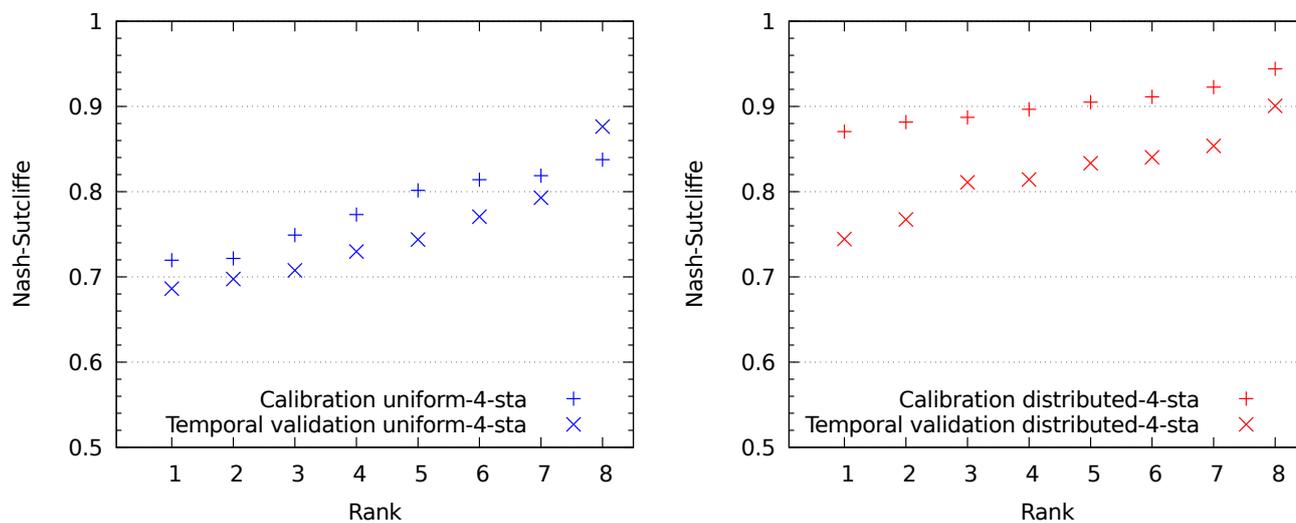
Figure 3. Distribution of the NS criteria in calibration (exp.1) and in the corresponding temporal validation (exp.2): left - **uniform-4-sta**, right - **distributed-4-sta**.
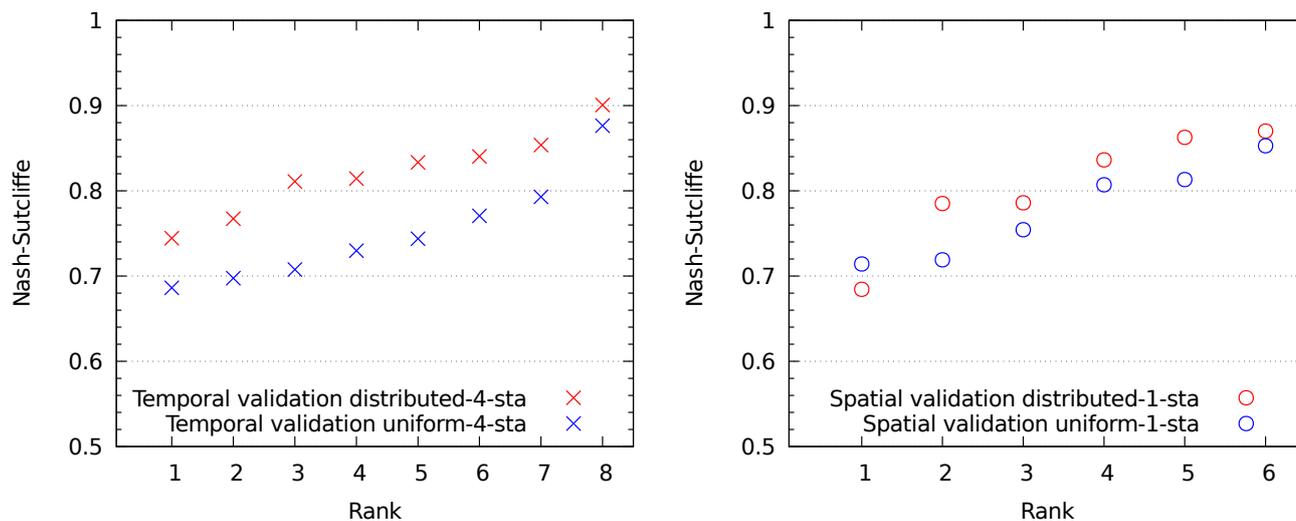

Figure 4. Distribution of the NS criteria: left - temporal validation (exp.2), right - spatial validation (exp.4).

a chosen observation period and the associated test signal (rainfall) one can get a relatively stable set of calibrated parameters. However, for very different test signals (i.e. those coming from different populations) the calibrated sets
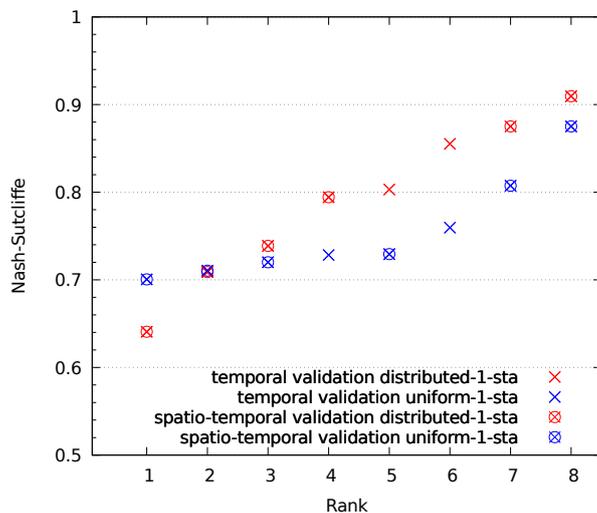
Figure 5. Distribution of the NS criteria in spatio-temporal validation (exp.5).

Table 2. Optimal uniform set of parameters for experiment uniform-4-sta.

Parameters	Period P1	Period P2
c_p (mm)	1000	1000
c_t ($mm.s^{0.25}$)	1215.7	1565.2
v ($m.s^{-1}$)	5	5

are also quite different, fully or partly. This clearly indicates a structural deficiency of the chosen model, which is not surprising since the model is conceptual.

First, the involved routing scheme may not describe properly the hydraulic properties of the basin. This poor hydraulic behaviour can be partially compensated by the others model parameters during the calibration process, which may explain the extreme (equal to the upper bound) parameter values. Second, the hydrological modelling at the cell scale is very primitive. For instance, ground water lost, lateral exchange and Horton run-off are not modelled.

The estimated velocity fields plotted in Fig.6 and Fig.7 remain similar for different experiments. Looking at the hydrographic network in Fig.2, one can see that the velocities are much higher along the main drains than on the side slopes. This result is in agreement with the true physical behaviour of the system, even though the routing scheme is conceptual. However, some unnatural phenomena can also be noticed. For example, upstream the Mialet station (north-west part of the right watershed), both the production and transfert capacities heat the upper

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

Table 3. Optimal uniform set of parameters for experiment uniform-1-sta.

Parameters	Period P1	Period P2
c_p (mm)	1000	1000
c_t $(mm.s^{0.25})$	1033.9	1625.8
v $(m.s^{-1})$	5	5

bounds and the velocities become very low along the drain. At Mialet and Générargue stations, peak discharges are often underestimated. At Anduze station, the model performance remains very good and the total discharges are compensated by the Gardon of Saint-Jean where discharges are often over-estimated during the flood (lower capacities). This bad upstream modeling is related to equifinality. To improve calibration quality, the velocities have

5 to be constrained according to the cumulative surfaces.

5 Conclusions

In this study, the distributed hydrological model GRD is generalized by using the "cell-to-cell" routing scheme. Since the model includes a high-dimensional set of distributed parameters, the variational method has been developed for its calibration. To the best of our knowledge, this is the first time when the variational estimation involving the adjoint

10 sensitivities has been applied in the field of hydrology. The main question has been whether or not one can benefit from allowing spatial variability of model parameters to improve the model's temporal, spatial and spatio-temporal predictive performance. The model and the calibration algorithm have been tested using data sets available for the Gardon d'Anduze, producing some encouraging results. For a given data record, the calibration algorithm retrieves a relatively stable set of distributed parameters which improves the model predictive performance. The temporal

15 validation has demonstrated the usefulness of the spatial calibration for discharge prediction in gauged locations, and the spatial and spatio-temporal validation - that the discharge prediction can be improved inside the watershed in ungauged locations/sub-catchments. Thus, the algorithm emulates a downscaling process which is of a great interest for modelling the hydrological processes. The utility of the fully distributed hydrological models (combined with the variational calibration) needs to be further investigated by considering a selection of different watersheds

20 and observation records. The immediate algorithmic developments include improvement of the routing scheme, implementation of constraints within the iterative regularization approach (to tackle the uniqueness/equifinality issues), and considering the pooled analysis.

Code availability. Our research code is hosted on the IRSTEA Gitlab at https://gitlab.irstea.fr/aiga/GRDv2.git.

Hydrology and
Earth System
Sciences
Discussions

EGU
Open Access

References

Abbaris, A., Dakhlaoui, H., Thiria, S., and Bargaoui, Z.: Variational data assimilation with the YAO platform for hydrological forecasting, Proceedings of the International Association of Hydrological Sciences, Volume 364, 2014, pp.3-8, 364, 3–8, https://doi.org/10.5194/piahs-364-3-2014, 2014.

5 Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., and Srinivasan, R.: Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT, Journal of Hydrology, 333, 413–430, 2007.

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'connell, P. E., and Rasmussen, J.: An introduction to the European Hydrological System—Systeme Hydrologique Europeen,"SHE", 2: Structure of a physically-based, distributed modelling system,

10 Journal of hydrology, 87, 61–77, 1986.

Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., and Javelle, P.: Sensitivity of hydrological models to uncertainty in rainfall input, Hydrological Sciences Journal–Journal des Sciences Hydrologiques, 56, 397–410, 2011.

Bell, V. A., Kay, A. L., Jones, R. G., and Moore, R. J.: Development of a high resolution grid-based river flow model for use with regional climate model output, Hydrology and Earth System Sciences Discussions, 11, 532–549, https://hal.

15 archives-ouvertes.fr/hal-00305636, 2007.

Beven, K.: Changing ideas in hydrology - The case of physically-based models, Journal of Hydrology, 105, 157–172, https://doi.org/10.1016/0022-1694(89)90101-7, 1989.

Beven, K.: Prophecy, reality and uncertainty in distributed hydrological modelling, Advances in Water Resources, 16, 41–51, 1993.

20 Borga, M., Anagnostou, E. N., Blöschl, G., and Creutin, J. D.: Flash floods: Observations and analysis of hydro-meteorological controls, Journal of Hydrology, 394, 1–3, 2010.

Bouvier, C. and DelClaux, F.: ATHYS: a hydrological environment for spatial modelling and coupling with GIS, IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences, 235, 19–28, 1996.

Braud, I., Ayral, P.-A., Bouvier, C., Branger, F., Delrieu, G., Le Coz, J., Nord, G., Vandervaere, J.-P., Anquetin, S., Adamovic,

25 M., Andrieu, J., Batiot, C., Boudevillain, B., Brunet, P., Carreau, J., Confoland, A., Didon-Lescot, J.-F., Domergue, J.-M., Douvinet, J., Dramais, G., Freydier, R., Gérard, S., Huza, J., Leblois, E., Le Bourgeois, O., Le Boursicaud, R., Marchand, P., Martin, P., Nottale, L., Patris, N., Renard, B., Seidel, J.-L., Taupin, J.-D., Vannier, O., Vincendon, B., and Wijbrans, A.: Multi-scale hydrometeorological observation and modelling for flash flood understanding, Hydrology and Earth System Sciences, 18, 3733–3761, https://doi.org/10.5194/hess-18-3733-2014, https://www.hydrol-earth-syst-sci.net/

30 18/3733/2014/, 2014.

Braud, I, Borga, M., Gourley, J., Hurlimann Ziegler, M., Zappa, M., and Gallart, F.: Flash floods, hydro-geomorphic response and risk management, Journal of hydrology, 541, 1–5, 2016.

Darras, T.: Flash flood forecasting by statistical learning, Ph.D. thesis, Université Montpellier, https://tel.archives-ouvertes.fr/tel-01816929, 2015.

35 Douinot, A., Roux, H., Garambois, P.-A., Larnier, K., Labat, D., and Dartus, D.: Accounting for rainfall systematic spatial variability in flash flood forecasting, Journal of Hydrology, 541, 359–370, 2016.

Hydrology and
Earth System
Sciences
Discussions
Open Access
EGU

Douinot, A., Roux, H., Garambois, P.-A., and Dartus, D.: Using a multi-hypothesis framework to improve the understanding of flow dynamics during flash floods, Hydrology and Earth System Sciences, 22, 5317–5340, 2018.

Drobinski, P., Ducrocq, V., Alpert, P., Anagnostou, E., Béranger, K., Borga, M., Braud, I., Chanzy, A., Davolio, S., Delrieu, G., Estournel, C., Boubrahmi, N. F., Font, J., Grubišić, V., Gualdi, S., Homar, V., Ivančan-Picek, B., Kottmeier, C., Kotroni, V., Lagouvardos, K., Lionello, P., Llasat, M. C., Ludwig, W., Lutoff, C., Mariotti, A., Richard, E., Romero, R., Rotunno, R., Roussot, O., Ruin, I., Somot, S., Taupier-Letage, I., Tintore, J., Uijlenhoet, R., and Wernli, H.: HyMeX: A 10-Year Multidisciplinary Program on the Mediterranean Water Cycle, Bulletin of the American Meteorological Society, 95, 1063–1082, https://doi.org/10.1175/BAMS-D-12-00242.1, https://doi.org/10.1175/BAMS-D-12-00242.1, 2014.

Edijatno: Mise au point d'un modèle élémentaire pluie-débit au pas de temps journalier, Ph.D. thesis, Université Louis Pasteur, ENGEES, Cemagref Antony, France, 1991.

Errico, R. M.: What is an adjoint model?, Bulletin of the American Meteorological Society, 78, 2577–2591, 1997.

Ficchì, A., Perrin, C., and Andréassian, V.: Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events, Journal of Hydrology, 538, 454–470, https://doi.org/https://doi.org/10.1016/j.jhydrol.2016.04.016, http://www.sciencedirect.com/science/article/pii/S0022169416301974, 2016.

Garambois, P.-A., Roux, H., Larnier, K., Castaings, W., and Dartus, D.: Characterization of process-oriented hydrologic model behavior with temporal sensitivity analysis for flash floods in Mediterranean catchments, Hydrology and Earth System Sciences, 17, 2305–2322, 2013.

Garambois, P.-A., Roux, H., Larnier, K., Labat, D., and Dartus, D.: Parameter regionalization for a process-oriented distributed model dedicated to flash floods, Journal of Hydrology, 525, 383–399, 2015.

Hascoet, L. and Pascual, V.: The Tapenade Automatic Differentiation tool: principles, model, and specification, ACM Transactions on Mathematical Software (TOMS), 39, 20, 2013.

Javelle, P., Fouchier, C., Arnaud, P., and Lavabre, J.: Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations, Journal of Hydrology, 394, 267–274, 2010.

Javelle, P., Demargne, J., Defrance, D., Pansu, J., and Arnaud, P.: Evaluating flash-flood warnings at ungauged locations using post-event surveys: a case study with the AIGA warning system, Hydrological sciences journal, 59, 1390–1402, 2014.

Javelle, P., Organde, D., Demargne, J., Saint-Martin, C., de Saint-Aubin, C., Garandeau, L., and Janet, B.: Setting up a French national flash flood warning system for ungauged catchments based on the AIGA method, in: 3rd European Conference on Flood Risk Management FLOODrisk 2016, vol. 7, p. 11, 2016.

Jay-Allemand, M., Gejadze, I., Javelle, P., Organde, D., Fine, J.-A., Patrick, A., and Malaterre, P.-O.: Assimilation de données appliquée à un modèle pluie-débit distribué pour la prévision des crues, in: De la prévision des crues à la gestion de crise, Société hydaulique de France, Avignon, 2018.

Laganier, O., Ayral, P. A., Salze, D., and Sauvagnargues, S.: A coupling of hydrologic and hydraulic models appropriate for the fast floods of the Gardon River basin (France), Natural Hazards and Earth System Sciences, 14, 2899–2920, 2014.

Laurent, L., Le Riche, R., Soulier, B., and Boucard, P.-A.: An Overview of Gradient-Enhanced Metamodels with Applications, Archives of Computational Methods in Engineering, 26, 61–106, 2019.

Ledimet, F. and Talagrand, O.: Variational Algorithms For Analysis And Assimilation Of Meteorological Observations - Theoretical Aspects, Tellus Series A-Dynamic Meteorology And Oceanography, 38, 97–110, https://doi.org/10.3402/tellusa.v38i2.11706, 1986.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., and Loumagne, C.: When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events, Hydrology and Earth System Sciences, Volume 18, Issue 2, 2014, pp.575-594, 18, 575–594, https://doi.org/10.5194/hess-18-575-2014, 2014.

McLaughlin, D.: Recent developments in hydrologic data assimilation, Reviews of Geophysics, 33, 977–984, 1995.

Merz, R. and Blöschl, G.: A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria, Water Resources Research, 45, 2009.

Michel C., Maihol J.C., L. T.: Hydrologie appliquée, Tech. rep., Cemagref, 1989.

Moradkhani, H. and Sorooshian, S.: General review of rainfall-runoff modeling: model calibration, data assimilation, and uncertainty analysis, in: Hydrological modelling and the water cycle, pp. 1–24, Springer, 2009.

Mouelhi, S., Michel, C., Perrin, C., and Andréassian, V.: Linking stream flow to rainfall at the annual time step: the Manabe bucket model revisited, Journal of hydrology, 328, 283–296, 2006.

Munier, S., Litrico, X., Belaud, G., and Perrin, C.: Assimilation of discharge data into semidistributed catchment models for short-term flow forecasting: Case study of the Seine River basin, Journal of Hydrologic Engineering, 20, 05014 021, 2014.

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models. Part I. A conceptual models discussion of principles., Journal of Hydrology, 10, 282–290, 1970.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: 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, 2005.

Perrin, C., Michel, C., and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, Journal of Hydrology, Volume 279, Issue 1, p. 275-289., 279, 275–289, https://doi.org/10.1016/S0022-1694(03)00225-7, 2003.

Quesney, A., Francois, C., Ottle, C., Hegarat, S., Loumagne, C., Normand, M., et al.: Sequential assimilation of SAR/ERS data in a lumped rainfall-runoff model with an extended Kalman filter, IAHS-AISH PUBL., pp. 495–497, 2000.

Rabier, F. and Courtier, P.: Four-dimensional assimilation in the presence of baroclinic instability, Quart. J. Roy. Meteorol. Soc., 118, 649–672, 1992.

Riboust, P., Thirel, G., Moine, N. L., and Ribstein, P.: Revisiting a Simple Degree-Day Model for Integrating Satellite Data: Implementation of Swe-Sca Hystereses, Journal of Hydrology and Hydromechanics, 67, 70–81, 2019.

Roux, H., Labat, D., Garambois, P.-A., Maubourguet, M.-M., Chorda, J., and Dartus, D.: A physically-based parsimonious hydrological model for flash floods in Mediterranean catchments, Natural Hazards and Earth System Science, Volume 11, Issue 9, 2011, pp.2567-2582, 11, 2567–2582, https://doi.org/10.5194/nhess-11-2567-2011, 2011.

Santos, L., Thirel, G., and Perrin, C.: Continuous state-space representation of a bucket-type rainfall-runoff model: a case study with the GR4 model using state-space GR4 (version 1.0), Geoscientific Model Development, 11, 1591–1605, 2018.

Sun, L., Nistor, I., and Seidou, O.: Streamflow data assimilation in SWAT model using Extended Kalman Filter, Journal of Hydrology, 531, 671–684, 2015.

Sun, L., Seidou, O., Nistor, I., and Liu, K.: Review of the Kalman-type hydrological data assimilation, Hydrological Sciences Journal, 61, 2348–2366, 2016.

Thirel, G., Martin, E., Mahfouf, J. F., Massart, S., Ricci, S., and Habets, F.: A past discharges assimilation system for ensemble streamflow forecasts over France–Part 1: Description and validation of the assimilation system, Hydrology and Earth System Sciences, 14, 1623–1637, 2010.

Tramblay, Y., Bouvier, C., Crespy, A., and Marchandise, A.: Improvement of flash flood modelling using spatial patterns of rainfall: A case study in southern France, in: Sixth World FRIEND Conference, pp. 172–178, 2010.

Wang, J., Hong, Y., Li, L., Gourley, J. J., Khan, S. I., Yilmaz, K. K., Adler, R. F., Policelli, F. S., Habib, S., Irwn, D., et al.: The coupled routing and excess storage (CREST) distributed hydrological model, Hydrological sciences journal, 56, 84–98, 2011.

Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J.: L-BFGS-B: a limited memory FORTRAN code for solving bound constrained optimization problems, EECS Department, Northwestern University, Evanston, IL, Technical Report No. NAM-11, 1994.
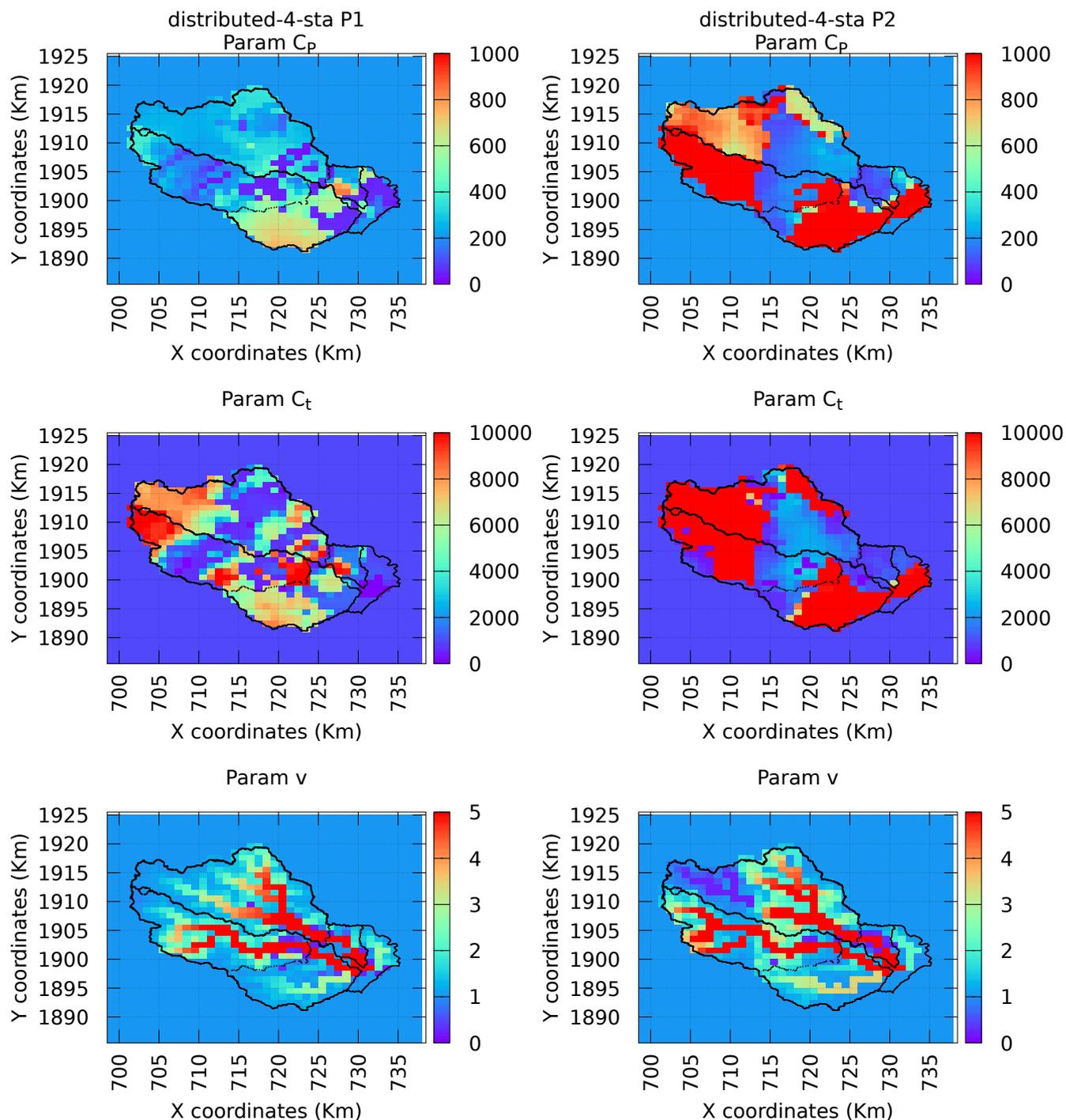
Figure 6. Maps of the calibrated coefficients (exp.1 - 4-sta): left - data from $P1$, right - data from $P2$.
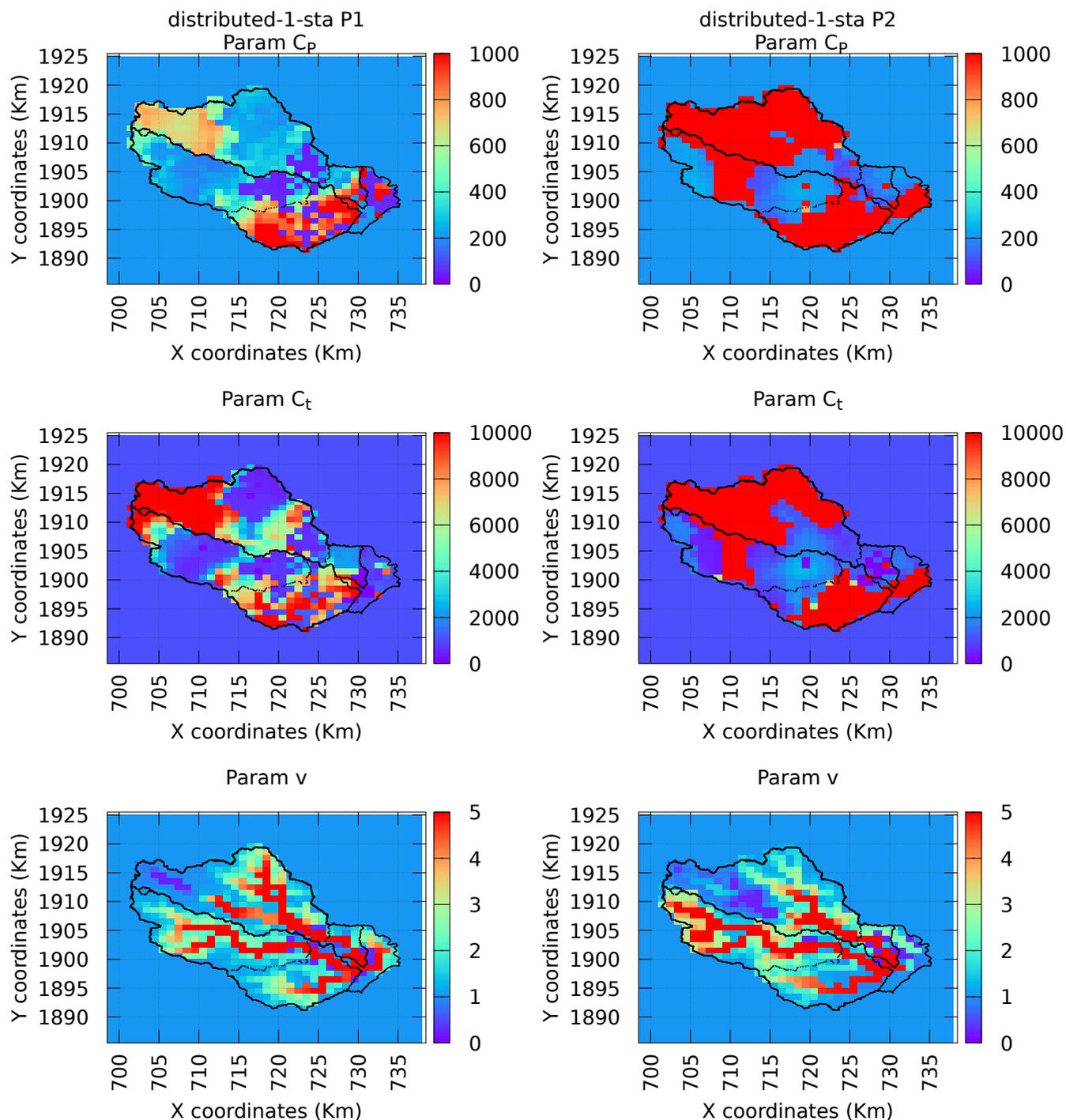
Figure 7. Maps of the calibrated coefficients (exp.3 - 1-sta): left - data from $P1$, right - data from $P2$.