



## Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <http://oatao.univ-toulouse.fr/>  
Eprints ID: 8101

**To link to this article:** <http://www.hydrol-earth-syst-sci.net/13/503/2009/>

**To cite this version:**

Castaings, W. and Dartus, D. and Le Dimet, F.-X. and Saulnier, G.-M. *Sensitivity analysis and parameter estimation for distributed hydrological modeling: potential of variational methods*. (2009) *Hydrology and Earth System Sciences*, vol. 13 . pp. 503-517. ISSN 1027-5606

Any correspondence concerning this service should be sent to the repository administrator:  
[staff-oatao@inp-toulouse.fr](mailto:staff-oatao@inp-toulouse.fr)

# Sensitivity analysis and parameter estimation for distributed hydrological modeling: potential of variational methods

W. Castaings<sup>1,2,4</sup>, D. Dartus<sup>1</sup>, F.-X. Le Dimet<sup>2</sup>, and G.-M. Saulnier<sup>3</sup>

<sup>1</sup>IMFT UMR 5502 (CNRS, INP, UPS), Université de Toulouse, 31400 Toulouse Cedex, France

<sup>2</sup>LJK UMR 5224 (CNRS, INPG, UJF, INRIA), Grenoble Universités, 38041 Grenoble Cedex 9, France

<sup>3</sup>EDYTEM, UMR 5204, Université de Savoie, 73376 Le Bourget du Lac Cedex, France

<sup>4</sup>European Commission, Directorate-General Joint Research Centre, Institute for the Protection and Security of the Citizen, Econometrics and Applied Statistics Unit, T.P. 361, 21020 Ispra (VA), Italy

*Correspondence to:* W. Castaings  
(william.castaings@imft.fr)

**Abstract.** Variational methods are widely used for the analysis and control of computationally intensive spatially distributed systems. In particular, the adjoint state method enables a very efficient calculation of the derivatives of an objective function (response function to be analysed or cost function to be optimised) with respect to model inputs.

In this contribution, it is shown that the potential of variational methods for distributed catchment scale hydrology should be considered. A distributed flash flood model, coupling kinematic wave overland flow and Green Ampt infiltration, is applied to a small catchment of the Thoré basin and used as a relatively simple (synthetic observations) but didactic application case.

It is shown that forward and adjoint sensitivity analysis provide a local but extensive insight on the relation between the assigned model parameters and the simulated hydrological response. Spatially distributed parameter sensitivities can be obtained for a very modest calculation effort ( $\sim 6$  times the computing time of a single model run) and the singular value decomposition (SVD) of the Jacobian matrix provides an interesting perspective for the analysis of the rainfall-runoff relation.

For the estimation of model parameters, adjoint-based derivatives were found exceedingly efficient in driving a bound-constrained quasi-Newton algorithm. The reference parameter set is retrieved independently from the optimization initial condition when the very common dimension reduction strategy (i.e. scalar multipliers) is adopted.

Furthermore, the sensitivity analysis results suggest that most of the variability in this high-dimensional parameter space can be captured with a few orthogonal directions. A parametrization based on the SVD leading singular vectors was found very promising but should be combined with another regularization strategy in order to prevent overfitting.

---

## 1 Introduction

The distributed modelling of catchment hydrology is now recognised as a valuable approach in order to understand, reproduce and predict the behavior of hydrological systems. However, distributed hydrological models remain a simplified and imperfect representation of the physical processes using uncertain observation data for the estimation of the model inputs to be prescribed (parameters, initial condition and rainfall forcing). Analyzing and reducing uncertainty is therefore essential but issues such as sensitivity and uncertainty analysis, parameter estimation and state-updating are challenging given the dimensionality of the system. Although the approach adopted in this paper is not restricted to a specific type of model input, the focus will be on model parameters to be assigned on the basis of indirect observations (i.e. model calibration).

For both sensitivity analysis and parameter estimation, many of the approaches which are now adopted for distributed hydrological models were originally developed for parsimonious hydrological models (e.g. simplified bucket models, models based on the concept of hydrological similarity). These models are very often characterised by a very low dimensionality, a limited computational cost and no sizeable constraints on model parameters.

Experiences in the calibration of such models revealed that the corresponding response surface often contains several regions of attraction, discontinuous derivatives and other geometrical properties compromising the use of local methods, especially those using derivative information (Ibbitt and O'Donnell, 1971; Johnston and Pilgrim, 1976; Duan et al., 1992).

Therefore, many recent applications and methodological developments for model calibration involve a stochastic exploration of the parameter space using computationally intensive Monte Carlo methods and/or evolutionary algorithms. These techniques enable the global optimization of single or multiple objectives and very often characterise the uncertainty affecting model parameters (Beven and Binley, 1992; Duan et al., 1992; Kuczera and Parent, 1998; Yapo et al., 1997; Vrugt et al., 2003a,b).

Differential sensitivity analysis (McCuen, 1973a,b), which was very often the only approach computationally affordable, is now gradually replaced by assessments carried out in the statistical framework. The Regional Sensitivity Analysis (RSA) of Hornberger and Spear (1981) inspired numerous applications and developments for the analysis of hydrological systems including the contribution of Beven and Binley (1992). The combination with recursive estimation techniques (Vrugt et al., 2002; Wagener et al., 2003) or the extension to multiple objectives (Bastidas et al., 1999) can provide an interesting insight into the behaviour of hydrological models. The use of variance decomposition approaches which are based on unambiguous importance measures (Cukier et al., 1978; Sobol', 1993; Homma and Saltelli, 1996) is now emerging in the hydrological community (Tang et al., 2007a,b; Yatheendradas et al., 2008; Van Werkhoven et al., 2008a). Using these global sensitivity analysis techniques, it is possible to assess how uncertainty in the model outputs can be apportioned to different sources of uncertainty in the model inputs (Saltelli et al., 2000).

When parameter estimation and sensitivity analysis are carried out in the statistical framework, it is necessary to sample the space of uncertain inputs. However, in distributed hydrological models, parameters are discretized according to the spatial discretization of the model state variables. Approaches developed for parsimonious hydrological models are frequently transferred to distributed hydrological models by means of an empirical dimension reduction of the parameter space. For parameter estimation, scalar multipliers are used in order to adjust spatially distributed parameters featuring a variability which is fixed using prior information (Refsgaard, 1997; Madsen, 2003). The same strategy can be adopted for probabilistic sensitivity analysis (Yatheendradas et al., 2008). In some cases, spatially distributed importance measures are estimated for a very coarse grid resolution or few zones of constant value (Hall et al., 2005; Tang et al., 2007a; Van Werkhoven et al., 2008b). Sampling based approaches to sensitivity analysis and parameter estimation enable an exploration of the parameter space (essential for

non-linear systems) but can be limited in handling distributed parameter systems (i.e. curse of dimensionality).

In the deterministic framework, both sensitivity analysis and parameter estimation can be addressed using variational methods. The adjoint state method enables an efficient calculation of the derivatives of an objective function with respect to all model inputs. This technique is particularly suited when the dimension of the response function to be analysed or cost function to be optimized is small when compared to the number of inputs to be prescribed (Lions, 1968; Cacuci, 1981a; Le Dimet and Talagrand, 1986). It has contributed to numerous applications related to the analysis and forecasting of meteorological and oceanographic systems (Le Dimet and Talagrand, 1986; Hall and Cacuci, 1983; Navon, 1998; Ghil and Malanotte-Rizzoli, 1991; Bennett, 1992). Early applications of the adjoint state method to hydrological systems have been carried out in groundwater hydrology (Chavent, 1974; Carrera and Neuman, 1986; Sun and Yeh, 1990). The resolution of inverse problems (parameter, state and boundary condition estimation), local sensitivity analysis, where also addressed in this framework in land surface hydrology (Mahfouf, 1991; Callies et al., 1998; Calvet et al., 1998; Bouyssel et al., 1999; Margulis and Entekhabi, 2001; Reichle et al., 2001), vadose zone hydrology (Ngnepieba et al., 2002), river and floodplain hydraulics (Piasiecki and Katopodes, 1997; Belanger and Vincent, 2005; Honnorat et al., 2006) and catchment hydrology (White et al., 2003; Seo et al., 2003a).

The previously mentioned applications involve non-linear models and the underlying inverse problems (i.e. parameter estimation and state updating) are ill-posed. For example, equifinality is inherent in the estimation of a distributed hydraulic conductivity in groundwater hydrology or in the estimation of an initial state for the atmosphere in meteorology. The estimation of model inputs require an appropriate combination of prior information (e.g. derived from land cover and soil type) and observations of the model diagnostic variables (e.g. streamflow observations). The variational framework is suitable for the combination of the different sources of information (including statistical information) through the resolution of a regularized inverse problem.

The use of parsimonious parametrizations (i.e. scalar multipliers or parameter zonation) can be seen as a regularization approach leading to well-posed inverse problems. Using Tikhonov regularization (Tikhonov and Arsenin, 1977), prior values are specified through a penalization term in the objective function. Therefore, deviations from prior values are allowed only when strongly supported by the calibration data. Whatever the strategy adopted for the regularization of the inverse problem, the adjoint state method enable a precise and efficient estimation of the gradient driving efficient optimization algorithms. Adjoint-based sensitivities can also provide an extensive insight on the way the values specified at the different grid elements influence the simulated hydrological response. The assessment is only local (i.e. outcomes

valid for a specific point of the parameter space) but provide tremendous information which would require a prohibitive computational cost if it was to be obtained using sampling based approaches.

In rainfall-runoff modelling, deterministic sensitivity analysis and gradient-based parameter estimation have been used in the past and the contributions of McCuen (1973a,b) and Gupta and Sorooshian (1985) are directly in line with the current research endeavor. The explicit but piecewise differentiation (i.e. analytic derivatives rather than classical finite difference approximation) carried out by McCuen (1973a) and Gupta and Sorooshian (1985) corresponds to the strategy adopted for the forward mode of algorithmic differentiation (Rall, 1981). By making the computational cost independent from the dimension of the input space the adjoint state method (implemented with the reverse mode of algorithmic differentiation) represents a significant improvement for the analysis and control of spatially distributed hydrological systems. The contributions of White et al. (2003) and Seo et al. (2003a,b) address the use of this approach for parameter estimation and state updating.

The objective of this paper is to demonstrate the potential of variational methods, briefly presented in Sect. 2, for the analysis of distributed rainfall-runoff models. A very simple application case is adopted for this prospective study. Other investigations involving more complex configurations and other model structures were carried out and will be reported in due course. A distributed flash flood model described in Sect. 3 was applied to a small catchment of the Thoré basin. The authors seek to illustrate what can be learned or corroborated using forward and adjoint sensitivity analysis for understanding the mapping between the model parameters and the simulated hydrological response. Using synthetic observations, the ability of efficient adjoint-based optimization to estimate reliable values for the model parameters is investigated. The results of sensitivity analysis and parameter estimation experiments are provided in Sects. 4 and 5 and the paper is concluded by a discussion of the main outcomes.

## 2 Variational methods for sensitivity analysis and parameter estimation

Variational methods provide a deterministic framework for the analysis and control of physical systems. The mathematical formalism, based on functional analysis and differential calculus, was largely expanded by related scientific disciplines such as optimal control and optimization. Adopting the vocabulary of optimal control theory, model parameters, initial and boundary conditions are referred as control variables living in a so-called control space.

This differential approach enables the exact calculation of the derivatives of a function of the model prognostic variables with respect to all control variables. Using the adjoint state method, the computational cost can be indepen-

dent from the dimension of the control space. It is precisely this feature which makes the approach very attractive for the analysis and control of spatially distributed systems.

Although any control variable can be analysed, the focus of this paper is parametric uncertainty. The use of derivatives for local sensitivity analysis and parameter estimation is addressed in Sects. 4 and 5. In this section, the calculation strategy is briefly described using a simplified mathematical formalism and the practical implementation of the approach is discussed.

For a didactic presentation of the approach, let us consider that the behavior of the system between times  $t_0$  and  $t_f$  is described by a generic model of the form:

$$\begin{cases} \frac{\partial \mathbf{x}}{\partial t} = M(\mathbf{x}, \boldsymbol{\alpha}) \\ \mathbf{x}(t_0) = \mathbf{0} \end{cases} \quad (1)$$

where  $\mathbf{x}$  is the state variable of dimension  $N_s$ ,  $M$  a nonlinear operator (after space discretization) and  $\boldsymbol{\alpha}$  a vector of parameters of dimension  $N_p$ . When the model parameters are fixed to  $\boldsymbol{\alpha} = \bar{\boldsymbol{\alpha}}$ ,  $\bar{\mathbf{x}}$  the corresponding nominal value for the state variable  $\mathbf{x}$  is obtained by solving the system given by Eq. (1).

In order to analyse or control the behaviour of the system, let us define a generic objective functional:

$$J(\mathbf{x}, \boldsymbol{\alpha}) = \int_{t_0}^{t_f} \phi(t; \mathbf{x}, \boldsymbol{\alpha}) dt \quad (2)$$

where  $\phi$  is a nonlinear function of the state variables and model parameters. The objective function  $J$  can represent a specific aspect of the system behaviour (i.e. response function) or quantify the misfit between the model diagnostic variables and the available observations (i.e. cost function).

The gradient of the functional  $J$  (real valued scalar function) with respect to  $\boldsymbol{\alpha}$  at the point  $\bar{\boldsymbol{\alpha}}$  is given by

$$\nabla_{\boldsymbol{\alpha}} J(\bar{\mathbf{x}}, \bar{\boldsymbol{\alpha}}) = \left( \frac{\partial J}{\partial \alpha_1}, \dots, \frac{\partial J}{\partial \alpha_{N_p}} \right)_{\bar{\boldsymbol{\alpha}}} \quad (3)$$

The components of this vector quantify the rate of change of  $J$  along the vectors of the standard basis in the parameter space. After the application of a normalization procedure (discussed in Sect. 4), importance measures can be estimated in order to compare the relative influence of the various model parameters on the response of interest. When  $J$  is a cost function to be optimized,  $\nabla_{\boldsymbol{\alpha}} J$  can drive very efficient gradient-based optimization algorithms (e.g. quasi-newton) for the estimation of model parameters.

### 2.1 Problems underlying the approximation of derivatives

Differential sensitivity analysis and gradient-based parameter estimation both rely on an efficient and accurate evaluation of partial derivatives. The most common technique

for the evaluation of the gradient components consists in repeated model evaluations. For example, the first order finite difference approximation for the  $i$ -th component is given by

$$\left[ \frac{\partial J}{\partial \alpha_i} \right]_{\bar{\alpha}} \approx \frac{J(\bar{\alpha}_1, \dots, \bar{\alpha}_i + \varepsilon, \dots, \bar{\alpha}_{N_p}) - J(\bar{\alpha})}{\varepsilon} \quad (4)$$

where  $\varepsilon$  refers to a perturbation applied to the nominal value of  $\alpha_i$ . Using this approach, the model can be considered as a black box and the practical implementation is straightforward. However, the precision and efficiency of this technique are very limited. The accuracy of the approximation crucially depends on choice of the step size  $\varepsilon$ . In practice, there is a very difficult compromise to be found. Large values of  $\varepsilon$  lead to truncation error, small values may give rise to cancellation/round-off error. Moreover, for the approximation of all gradient components (i.e.  $\partial J / \partial \alpha_i$   $i=1, \dots, N_p$ ), perturbations should be applied along all vectors of the canonical basis of  $\mathbb{R}^{N_p}$ . Using the first order finite difference approximation given by Eq. 4,  $N_p + 1$  model evaluations are necessary. This number is of course larger for higher order approximations. Therefore, the overall computational cost is at least linearly related to the dimension of the parameter space.

In order to avoid the use of a perturbation parameter  $\varepsilon$ , derivatives can be calculated analytically. A very general and comprehensive mathematical formalism for differential sensitivity analysis was proposed by Cacuci (1981a,b). It is based on the concept of Gâteaux derivative, a generalisation of the concept of directional derivative in differential calculus.

## 2.2 Forward sensitivity analysis

The derivative of the objective function  $J$  at the point  $\bar{\alpha}$  in the direction  $\hat{\alpha}$  is given by:

$$\hat{J}(\bar{\alpha}, \hat{\alpha}) = \int_{t_0}^{t_f} \left( \left[ \frac{\partial \phi}{\partial \mathbf{x}} \right]_{\bar{\alpha}} \hat{\mathbf{x}} + \left[ \frac{\partial \phi}{\partial \alpha} \right]_{\bar{\alpha}} \hat{\alpha} \right) dt \quad (5)$$

where  $\hat{\mathbf{x}}$  refers to variations on the state variable  $\mathbf{x}$  resulting from perturbations on the parameters  $\alpha$  in the direction  $\hat{\alpha}$ . It is important to note that  $\hat{J}(\bar{\alpha}, \hat{\alpha}) = \langle \nabla_{\alpha} J, \hat{\alpha} \rangle$  ( $\langle \cdot, \cdot \rangle$  standing for the scalar product). Given that  $\mathbf{x}$  is governed by Eq. (1), it is necessary to derive this system in order to estimate  $\hat{\mathbf{x}}$ . Therefore,  $\hat{\mathbf{x}}$  is solution of the following system:

$$\begin{cases} \frac{\partial \hat{\mathbf{x}}}{\partial t} - \left[ \frac{\partial M}{\partial \mathbf{x}} \right]_{\bar{\alpha}} \hat{\mathbf{x}} = \left[ \frac{\partial M}{\partial \alpha} \right]_{\bar{\alpha}} \hat{\alpha} \\ \hat{\mathbf{x}}(t_0) = 0 \end{cases} \quad (6)$$

where  $[\partial M / \partial \mathbf{x}]$  represents the Jacobian of the model with respect to the state variables and  $[\partial M / \partial \alpha]$  the Jacobian of the model with respect to the model parameters.

The system given by Eq. (6) is the so-called tangent linear model (TLM). For a given perturbation  $\hat{\alpha}$ ,  $\hat{\mathbf{x}}$  is obtained by the resolution of the TLM. The resulting variation of the objective function (e.g.  $\hat{J}(\bar{\alpha}, \hat{\alpha})$ ) can be calculated from Eq. (5).

However, in order to obtain all gradient components, the operation should be repeated for  $\hat{\alpha}$  corresponding to the different vectors of the canonical basis of  $\mathbb{R}^{N_p}$ . This means that the resolution of the TLM has to be carried out for each direction  $\hat{\alpha}$ . Therefore, the precision problem (i.e. approximation error) is addressed but the overall computational cost is still dependent on the dimension of the parameter space. This difficulty can be overcome by using the adjoint state method.

## 2.3 Adjoint sensitivity analysis

The linearity of  $\hat{J}(\bar{\alpha}, \hat{\alpha})$  with respect to  $\hat{\alpha}$  is produced using the introduction of an auxiliary variable  $\mathbf{p}$  (of dimension  $N_s$ ). It can be shown (Cacuci, 1981a; Le Dimet and Talagrand, 1986) that if  $\mathbf{p}$  is governed by the following system

$$\begin{cases} \frac{\partial \mathbf{p}}{\partial t} + \left[ \frac{\partial M}{\partial \mathbf{x}} \right]_{\bar{\alpha}}^T \mathbf{p} = \left[ \frac{\partial \phi}{\partial \mathbf{x}} \right]_{\bar{\alpha}} \\ \mathbf{p}(t_f) = 0 \end{cases} \quad (7)$$

the gradient is given by

$$\nabla_{\alpha} J(\bar{\mathbf{x}}, \bar{\alpha}) = \int_{t_0}^{t_f} \left( \left[ \frac{\partial \phi}{\partial \alpha} \right]_{\bar{\alpha}} - \left[ \frac{\partial M}{\partial \alpha} \right]_{\bar{\alpha}}^T \mathbf{p} \right) dt \quad (8)$$

where  $[ \ ]^T$  stands for the transpose.

It is important to note that  $\hat{\mathbf{x}}$  and  $\hat{\alpha}$  do not appear in Eqs. (7) and (8). Therefore, once  $\mathbf{p}$  is known by integration (backward in time) of the system described by Eq. (7), all the components of the gradient  $\nabla_{\alpha} J$  needed for sensitivity analysis and parameter estimation can be calculated. In the mathematical optimization framework, the objective is to maximize/minimize the cost function  $J$  while  $\mathbf{x}$  is subject to the constraint given by Eq. 1 (i.e.  $\mathbf{x}$  should verify the governing equations). In this case, the adjoint variables correspond to the Lagrange multipliers of the constrained optimization problem. The principal difficulty resides in the derivation and transposition of complex operators.

## 2.4 Practical implementation

In principle, forward and adjoint sensitivity analysis can be performed on the continuous or discrete formulation of the model. Different implementation strategies can be adopted depending if the derivation and transposition operations are carried out on the continuous form of the direct model, on its discretized form or directly on the computer code. Algorithmic differentiation (AD) is usually preferred (see Sei and Symes, 1995 or Sirkes and Tziperman, 1997 for counterexamples) for reliable and accurate derivatives. The reader is referred to Griewank (2000) for a comprehensive description of AD.

The source code implementing the model and objective functional is a concatenated sequence of instructions. Each statement contains elementary functions which can be easily derived (i.e. local Jacobians). Algorithmic differentiation (AD) is based on a rigorous application of the chain

rule (e.g. product of local Jacobians) in the forward or reverse direction. In order to use this discrete equivalent of forward and adjoint sensitivity analysis methods, the source code of the model should be available. The availability of automatic differentiation engines (see <http://www.autodiff.org/>) provide a helpful and efficient support for the practical implementation of variational methods. It is important to emphasise that using this implementation strategy, the potential on-off switches characterising the representation of physical processes (i.e. thresholds in model formulation) are simply reported in the TLM and ADM (Zou et al., 1993). The derivative provided by AD is therefore an element of the sub-gradient.

### 3 Flash flood model

A very simple and common model structure was adopted and applied to a small catchment using synthetic observations. An event-based distributed rainfall-runoff model is applied to a small area in the upper part of the Thoré catchment (Tarn department, South West of France).

#### 3.1 Model description

The underlying physics of MARINE flash flood model (Estupina-Borrell et al., 2006) is adapted to events for which infiltration excess dominates the generation of the flood. A simplified version of the model is used for this prospective study. Rainfall abstractions are evaluated using the Green Ampt infiltration model and the resulting surface runoff is transferred using the Kinematic wave approximation (KWA). The complex geometry of the catchment is described by a structured grid in which each cell receives water from its upslope neighbors and discharge to a single downslope neighbor (steepest direction).

For a one dimensional flow of average velocity  $u$  and average depth  $h$ , the continuity equation can be expressed as:

$$\frac{\partial h}{\partial t} + \frac{\partial uh}{\partial x} = r - i \quad (9)$$

where  $r$  is the rainfall intensity and  $i$  the infiltration rate. Using the KWA approximation, which has shown the ability to represent channelized and sheet overland flow (Singh, 2001), the momentum conservation equation reduces to an equilibrium between the bed slope  $S_0$  and the friction slope  $S_f$ . The Manning equation (uniform flow on each grid cell) is used to relate the flow velocity and the flow depth:

$$u = \frac{R^{2/3} S_f^{1/2}}{n} \quad \text{with} \quad R = \frac{hw}{2h + w} \quad (10)$$

where  $R$  is the hydraulic radius,  $n$  the Manning roughness coefficient and  $w$  the elemental flow width. In this simplified version of the model, the flow width is constant (rectangular section) and given the ratio between the width (grid resolution) and the flow depth the hydraulic radius is approximated

by the water depth (i.e.  $R=h$ ). The resulting equation governing the overland flow is given by:

$$\frac{\partial h}{\partial t} + \frac{S_0^{1/2}}{n} \frac{\partial h^{5/3}}{\partial x} = r - i \quad (11)$$

In the right hand side of Eq. (11), the infiltration rate  $i(t)$  is estimated using the very common Green-Ampt infiltration model.

For an homogeneous soil column characterised by its effective hydraulic conductivity  $K$ ,  $\psi$  the soil suction at wetting front, the potential infiltration rate is given by

$$i(t) = K \left( \frac{\psi \Delta \theta}{I(t)} + 1 \right) \quad \text{with} \quad \Delta \theta = \eta(1 - \theta) \quad (12)$$

where  $\theta$  is the relative initial soil moisture (i.e.  $\theta \in [0, 1]$ ),  $\eta$  the soil porosity and  $I(t)$  the cumulative infiltration at time  $t$ . After surface ponding, the cumulative infiltration at time  $t + \Delta t$  can be calculated by the following equation

$$I_{t+\Delta t} - I_t - \psi \Delta \theta \ln \left[ \frac{I_{t+\Delta t} + \psi \Delta \theta}{I_t + \psi \Delta \theta} \right] = K \Delta t \quad (13)$$

which is solved by Newton's method.

In order to carry out the sensitivity analysis and parameter estimation experiments presented in Sects. 4 and 5, it is necessary to implement the tangent linear and adjoint models for the hydrological model presented in this section.

#### 3.2 Computer code differentiation

As emphasised in Sect. 2.4, the best representation of the operator to be derived is the associated computer code. The algorithmic differentiation of the MARINE source code (in Fortran 90) was carried out with the support of an automatic differentiation engine. The TAPENADE automatic differentiation engine (Hascoët and Pascual, 2004), a source-to-source transformation program, was adopted because of its flexibility and efficiency for both forward and reverse modes. Preliminary modifications of the source code were necessary and the code produced by TAPENADE was optimized in order to reduce the memory footprint and the computational time. This leads to a computational time for the adjoint model which is about 6 times higher than the one observed for a single model evaluation.

#### 3.3 Case study description

MARINE was applied to a very small catchment area (25 km<sup>2</sup>) from the upper part of the Thoré basin which was affected by a catastrophic flood event in November 1999. During this event, the observed cumulative rainfall was 135 mm and the maximum intensity around 75 mm/h<sup>-1</sup>. Using the observed rainfall forcing (radar data from Météo France) and prior values derived from published tables for the model parameters, synthetic observations are generated.

**Table 1.** Model parameters values derived from soil type and land cover data.

Parameter	Name	Ref. value
Effective hydraulic conductivity (mm/h)	$K$	3.
Porosity	$\eta$	0.398
Suction (mm)	$\psi$	218.5
Initial soil moisture (fraction of $\eta$ )	$\theta$	0.5
Manning roughness coefficient	$n$	0.065

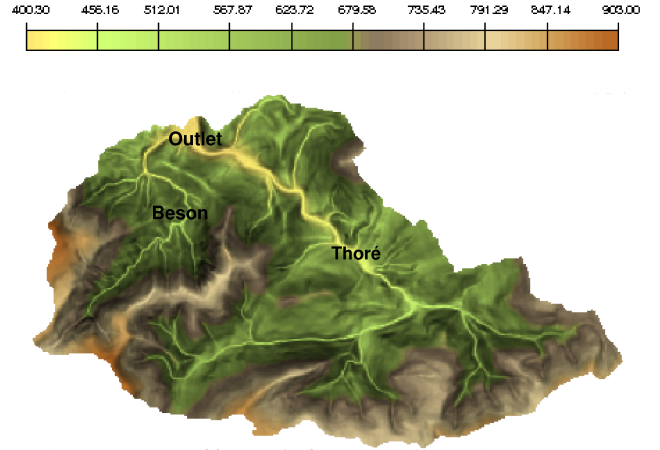
Using the values specified in Table 1 (spatially uniform values for model parameters), the resulting specific discharge is typical for Mediterranean flash flood events. The two sub-catchments contributing to the discharge in Labastide-Rouairoux (outlet situated to the north on Fig. 1) are drained by the Thoré river (eastern part) and the Beson stream (smaller sub-catchment in the western part).

#### 4 Differential sensitivity analysis

In differential sensitivity analysis, first order importance measures are calculated from the gradient when the response is scalar (e.g. peak discharge). For a vectorial response (e.g. entire flood hydrograph), the Jacobian matrix of the transformation can be evaluated. It can be computed column by column using the forward mode of AD, line by line using the reverse mode. When the dimension of the parameter space is much larger than the dimension of the response to be analysed, the adjoint technique (e.g. reverse mode of AD) is the most efficient calculation method.

However, using the rate of change along the vectors of the standard basis (components of the gradient), the parameters cannot be ranked because the nominal values might be characterised by different units and therefore different orders of magnitude. It is possible to normalize the partial derivatives with the associated nominal values for the parameter and response (e.g.  $\partial J / \partial K \cdot \bar{K} / \bar{J}$ ). In this case, the importance measure corresponds to the effect on the response from perturbing the parameter by a fixed fraction of its base value. In order to take the uncertainty underlying the different parameters into account, the associated standard deviations (or variance) can be used for the normalization. The resulting importance measure corresponds to the effect on the response from perturbing the parameter by a fixed fraction of its standard deviation (Helton, 1993). Although the approach is quite appealing, this means that derivatives are used for ranking over the space of uncertain parameters. In this paper, the choice was made to prefer a strictly local analysis in which the assessment concerns the base point where derivatives are evaluated.

In the following paragraphs, forward sensitivity analysis will be carried out to a parametrization of reduced dimensionality, adjoint sensitivity analysis will be used for the fully



**Fig. 1.** Ground elevation of the Thoré basin at Labastide-Rouairoux.

distributed parametrization. In order to facilitate the interpretation of the results, a spatially lumped rainfall forcing is used for most of the experiments presented in this section.

#### 4.1 Numerical experiments with network/hillslopes parametrization

In this paragraph, a classical reduction of the control space is adopted. The drainage network and the hillslopes are distinguished using a threshold on the drained area. It leads to the definition of basis vectors exclusively composed of 0 and 1 on the hillslopes or drainage network. The spatially distributed parameters are expressed in this basis (e.g. rather than the canonical basis). In other words, for each parameter, a scalar multiplier is applied on the hillslopes and another in the drainage network.

##### 4.1.1 Analysis of a scalar response

The relative importance of the scalar multipliers on two aspects of the hydrological response (flood volume and flood peak) is provided by Figs. 2 and 3. The parameters  $\theta$ ,  $\psi$  and  $\eta$  appear as a product in the infiltration model (Eq. 12). Therefore, given the adopted normalization procedure, they have exactly the same influence on the model response. If a statistical normalization procedure is adopted, the ranking of those parameters would be completely driven by the associated statistical properties (variance or standard deviation).

It is important to note that for both flood volume and flood peak, all sensitivities are negative. This means that increasing the nominal value for all the parameters reduces the magnitude of the response. In fact, increasing the Green-Ampt model parameters leads to increased infiltration losses and therefore reduces the flood volume and flood peak. Increasing the friction parameters leads to a flatter flood hydrograph and therefore also reduces the flood peak.

The analysis of Fig. 2 confirms that the flood volume is mainly driven by the infiltration parameters on the hillslopes (hydraulic conductivity  $K$  and initial soil moisture  $\theta$ ). The infiltration parameters assigned on the hillslopes also have a significant influence on the flood peak but the effect of modifying the friction parameter in the drainage network is much larger (see Fig. 3).

The results presented above are strictly local and the importance measures are affected by the nominal values assigned to the different parameters. However, given the reduced computational cost of the analysis, it can be carried out at different locations in the parameter space. Additional experiments (not reported in this paper) were conducted along a transect of the parameter space ( $\theta \in [0, 1]$ ). The results show that the wetter the soil at the beginning of the event, the faster the decay of the infiltration rate to the hydraulic conductivity, and therefore the greater the relative influence of  $K$  when compared to the initial soil moisture  $\theta$ .

#### 4.1.2 Analysis of the flood hydrograph

In order to analyse the effect of parameter variations on the complete flood hydrograph, a vectorial response containing the temporal evolution (80 time steps) of the simulated discharge was considered. Given to the ratio between input and output space dimensions (i.e. 6/80), the Jacobian matrix is computed using the tangent mode of TAPENADE (i.e. forward sensitivity analysis).

Each column of the Jacobian matrix represents the variations in discharges resulting from the perturbation of one of the model parameters. After normalization, the physical interpretation of the lines and/or columns of the Jacobian matrix can provide an interesting insight (not reported in the present contribution). However, a very interesting perspective is provided by the singular value decomposition of this Jacobian matrix.

The singular value decomposition (SVD) of an  $m \times n$  matrix  $\mathbf{A}$  is a factorization of the form

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (14)$$

where  $\mathbf{S}$  is a diagonal matrix containing the singular values of  $\mathbf{A}$  in the decreasing order while  $\mathbf{U}$  and  $\mathbf{V}$  are orthogonal matrices (respectively of dimension  $m \times m$  and  $n \times n$ ). The set of entries composing the main diagonal of  $\mathbf{S}$ , denoted  $\{\sigma_1, \sigma_2, \dots, \sigma_{\min(m, n)}\}$ , is referred as the singular spectrum of  $\mathbf{A}$ . The columns of  $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$  and  $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  are the left and right singular vectors in the input and output spaces of the transformation represented by  $\mathbf{A}$ . The magnitude of the singular values in  $\mathbf{S}$  represents the importance of the corresponding singular vectors in the columns of  $\mathbf{U}$  and  $\mathbf{V}$ .

This factorization is widely used for the analysis of linear ill-posed inverse problems (Hansen, 1998). Its application to non-linear systems can serve many purposes such as sensitivity or identifiability analysis, control variables estimation

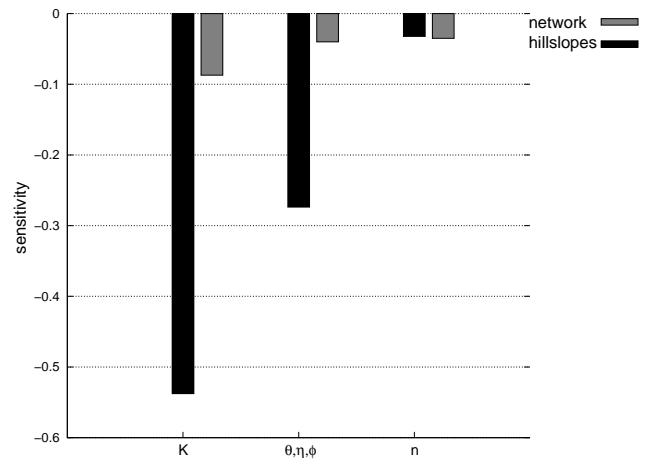


Fig. 2. Sensitivity of flood volume to model parameters for network/hillslopes parametrization.

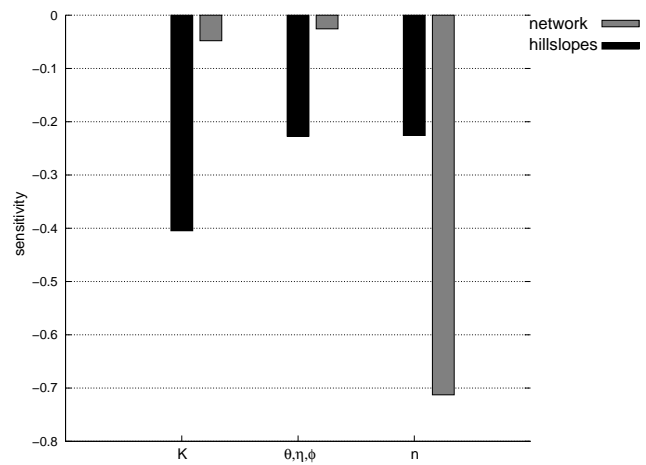


Fig. 3. Sensitivity of flood peak to model parameters for network/hillslopes parametrization.

or perturbations growth analysis in ensemble prediction (Li et al., 2005; Doherty and Randall, 2008; Clément et al., 2004; Marchand et al., 2008; Durbiano, 2001; Buizza and Palmer, 1995).

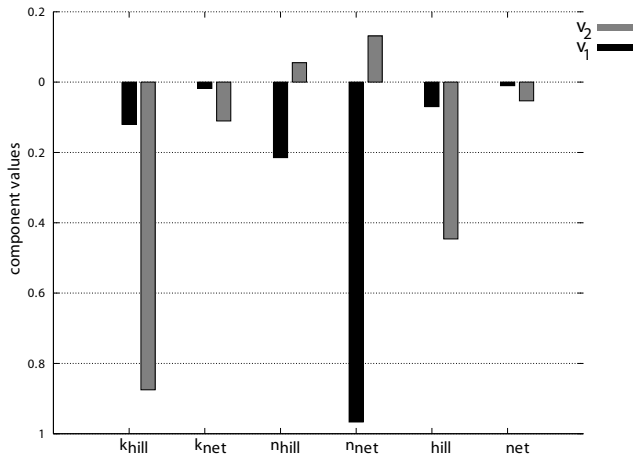
For the adopted network/hillslopes parametrization, the singular spectrum is given by Table 2 and the components of the first 2 singular vectors in the parameter space (right singular vectors) are plotted in Fig. 4. From Table 2, it can be seen that the decay of the singular values is very rapid. Most of the variability (more than 97%) is captured by the first two singular vectors. These vectors exhibit a clear distinction between the production and transfer of runoff. This could have been expected because they represent orthogonal directions in the parameter space.

Given the components of the first singular vector and the magnitude of the singular value associated, the scalar multiplier affecting the friction in the drainage network has an



**Table 2.** Singular values of the Jacobian matrix for the network/hillslopes parametrization.

Sing. values	% of variability
9.07	84.83
1.30	12.20
0.24	2.25
0.04	0.42
0.02	0.28
0.001	0.01



**Fig. 4.** Components of the first two singular vectors ( $V_1$  and  $V_2$ ) of the Jacobian in the parameter space for the network/hillslopes parametrization.

overwhelming influence on the flood hydrograph. The analysis of the second singular vector components indicates a predominance of the hillslopes infiltration parameters and a potential compensation with friction parameters.

Using the adjoint state method (reverse mode of AD) the computational cost related to the evaluation of local sensitivities is not related to the dimension of the parameter space. Therefore a similar analysis was carried out without reduction of the control space using scalar multipliers.

#### 4.2 Numerical experiments with fully distributed parameters

When no strategy is adopted for dimension reduction, the use of sampling based sensitivity analysis methods is not tractable for distributed parameter systems. In this paragraph the full potential of the adjoint method is exploited in order to analyse the effect on the hydrological response of variations on the value assigned to each element of the computational grid for the model parameters. The analysis is also carried out for a scalar response and for the entire flood hydrograph.

#### 4.2.1 Analysis of a scalar response

For a scalar response, a single integration of the adjoint model yields to sensitivity indices for all parameters at all spatial locations. In this paragraph, only the sensitivity to the flood peak is reported. Within a single river reach, it is obvious that increasing the friction will reduce the maximum discharge (i.e. negative sensitivities at all spatial locations). The situation is more complex when dealing with overland flow over the topography of a catchment.

The analysis of Fig. 5 reveals that positive and negative sensitivities are encountered over the catchment. Negative sensitivities are much larger than positive sensitivities in magnitude but there are locations where increasing the friction coefficient lead to a slight increase of the maximum discharge. While all sensitivities have the expected sign (i.e. negative) along the main stream (i.e. Thoré river), some positive sensitivities can counterbalance the overall effect in some concomitant sub-basins (e.g. area drained by the Beson stream and some hillslopes along the Thoré river).

Therefore, when applying a single scalar multiplier for the entire catchment, compensation effects usually occur which are very difficult to identify without such analysis. A simple corroboration can be carried out using multiple model evaluations. As an illustration, increasing the nominal by 10% for all roughness coefficients leads to  $-4.5\%$  variation on the peak discharge. This variation is  $-5.9\%$  when only the cells featuring a negative sensitivities are modified and it becomes  $+1.5\%$  when the same operation is carried out on the cells featuring positive sensitivities.

#### 4.2.2 Analysis of the flood hydrograph

When considering the entire flood hydrograph, the ratio between the input and output space dimensions is now very close to 100 (i.e.  $3 \times 2582/80$ ). Therefore, the Jacobian matrix is computed line by line using multiple integrations of the adjoint model.

In order to facilitate the physical interpretation, the SVD is performed on *sub-Jacobians*. Each *sub-Jacobian* accounts for a single parameter but for all spatial locations. For a given parameter, the analysis enables an extensive understanding for the influence of the values specified over the entire catchment.

For this specific analysis, the singular vectors in the parameter space can be mapped on the surface of the catchment. They are provided for the parameters  $n$  and  $K$  by Figs. 6 and 7. The analysis of the previously cited figures enables an extensive insight into the model behavior. The drainage network is emphasized for the friction coefficient (Fig. 6) and flow concentrations areas can be distinguished for the hydraulic conductivity (Fig. 7). Significant interactions between different regions of the catchment were already identified when analysing the variability of the sensitivity to the flood peak (see Fig. 5). A similar behaviour is

encountered when analysing the entire flood hydrograph for both friction coefficient and hydraulic conductivity.

The leading singular vector mainly corresponds to the Thoré river and drained area for both  $n$  and  $K$  (Figs. 6a and 7a). However, the interacting regions are already characterised by different signs. For the second singular vectors, the catchment regions and signs are inverted (Figs. 6b and 7b).

The interactions between the two sub-basins can be also analysed in the observation space. For the roughness coefficient  $n$ , the components of  $u_1$  and  $u_2$  (singular vectors in the observation space) are plotted together with the outlet discharge (see Fig. 8). The analysis of this figure explains that the slight disruptions of the hydrograph during both the rising limb and the recession are mainly due to the flood wave coming from the Beson stream. For the imposed spatially lumped rainfall, given that the smaller sub-basin (holding  $v_2$ ) is closer to the catchment outlet, the resulting smaller concentration time leads to a quicker response perfectly characterised by  $u_2$ . The correspondence between the singular vectors in the parameter and observation spaces is really informative and meaningful.

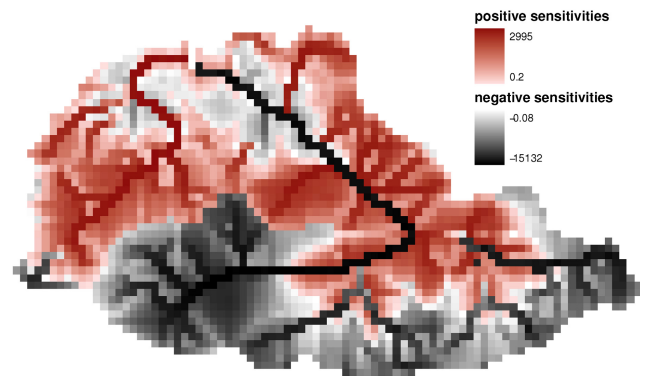
In addition, the singular spectrum for all parameters and different forcing conditions (lumped and spatially variable rainfall) was also analysed. The analysis of Fig. 9 reveals that the decay of singular values is faster for the roughness coefficient when compared to the infiltration parameters. Although the influence of friction is very important, less singular vectors are necessary to describe the sub-space producing variability in simulated discharges. This is due to the fact that this subspace is mainly restricted to the drainage network for friction parameters (see Fig. 6).

For a spatially distributed rainfall forcing the decay of singular values appears to be slower and the gap between friction and infiltration parameters is reduced. The fact that more singular vectors are necessary to describe the sub-space producing important variations of the simulated discharges is a sign of increased information content. This increase in information content was expected when comparing the results obtained with uniform rainfall forcing with those obtained for spatially variable precipitations.

It was shown in this section that the derivatives obtained with algorithmic differentiation provide a valuable introspection into the relation between the model parameters and the simulated hydrological response. The availability of accurate adjoint-based sensitivities also enables the use of efficient gradient-based optimization techniques for the estimation of model parameters.

## 5 Gradient-based parameter estimation

In this paper, a bound-constrained (inequality constraints) quasi-Newton (BFGS) optimization algorithm (Lemarchal and Panier, 2000) from the MODULOPT library was used.



**Fig. 5.** Spatial variability for the sensitivity of the flood peak to the friction coefficient.

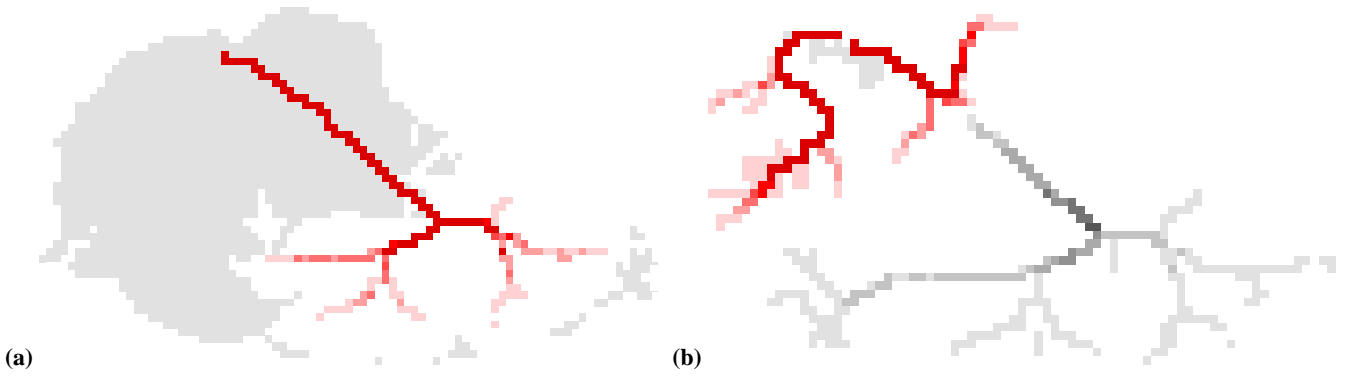
Using the adjoint sensitivities, the algorithm estimates the active set by performing a Wolfe line search along the gradient projection path.

### 5.1 Numerical experiments with network/hillslopes parametrization

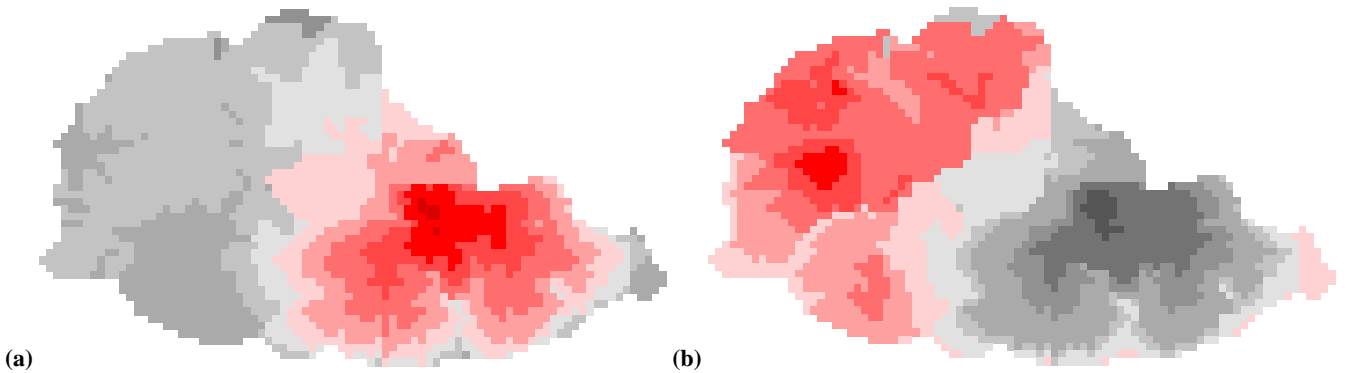
Synthetic observations are generated with the parametrization described in the previous section with  $K_{net}=4\text{ mmh}^{-1}$ ,  $K_{hill}=2\text{ mmh}^{-1}$ ,  $n_{net}=0.05$ ,  $n_{hill}=0.08$  and  $\theta=0.5$  (uniform over the catchment). The Nash criterion is used to measure the misfit between model simulations and the synthetic observations. As shown in Fig. 10, all control variables are retrieved independently from the initial parameter values (initial point for the optimization routine). The relative importance of the parameters inferred from the local sensitivity analysis results seems to be similar for the Nash efficiency over the bounded parameter space. The more sensitive is the performance measure to a parameter, the greater is the identifiability of this parameter and therefore the faster the iterates convergence to the reference value (e.g. parameters  $n_{net}$  and  $K_{hill}$ ).

It is important to emphasize that both adjoint (required to the estimation of the gradient) and direct model evaluations (required for the line search) were reported in Fig. 10. The total number of iterations (less than 50) is very small when compared to the number of model evaluations required by evolutionary algorithms like the very popular Shuffled Complex Evolution (SCE) from Duan et al. (1992). As commented by Kavetski et al. (2006), “5 min of Newton computing often replaces 24 h of SCE search and yield useful additional information”.

In the parameter estimation experiments presented in this section, a network/hillslopes parametrization was adopted in order to ensure the identifiability of model parameters. However, the sensitivity analysis results described in Sect. 4.2.2 have shown that the sub-space from the original parameter space driving the simulated discharges is spanned by the



**Fig. 6.** First singular vectors (i.e. (a)  $v_1$  and (b)  $v_2$ ) in the parameter space for the roughness coefficient  $n$  (red color ramp for positive components and gray for negative).



**Fig. 7.** First singular vectors (i.e. (a)  $v_1$  and (b)  $v_2$ ) in the parameter space for the hydraulic conductivity  $K$  (red color ramp for positive components and gray for negative).

leading singular vectors (i.e. in the parameter space) of the Jacobian matrix. Given that a small number of singular values are dominant, as illustrated in Fig. 9, most of the variability can be captured with very few orthogonal directions in the parameter space. In the following paragraph, the leading singular vectors of the Jacobian matrix are used in order to reduce the dimensionality of the parameter estimation problem.

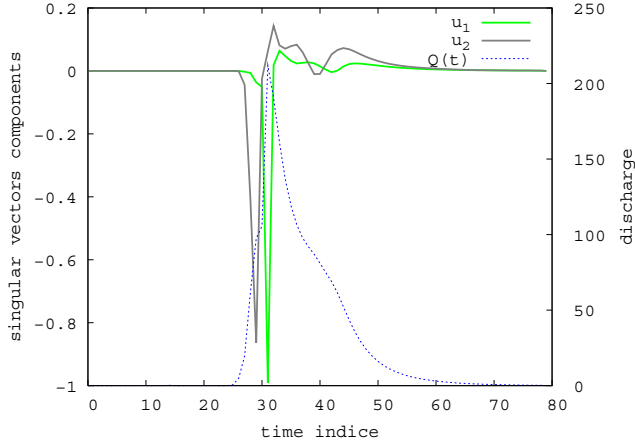
## 5.2 Numerical experiments with TSVD parametrization

Using the sensitivity analysis outcomes, it is possible to carry out parameter estimation in a reduced basis taking the observations information content into account. The spatially distributed model parameters are therefore expressed in the basis spanned by the Jacobian leading singular vectors.

The derivation was carried out at a specific point of the parameter space. However, singular vectors in the parameter space are mainly determined by the topography of the catchment and the spatial variability of rainfall. Important

variations of those orthogonal directions were not encountered when the Jacobian is evaluated at different locations in the parameter space. In order to compute the singular vectors describing the relevant sub-space for parameter estimation, the SVD was performed for the Jacobian calculated with spatially uniform rainfall forcing. A more rigorous approach would require the Jacobian to be computed with the response of the catchment for several rainfall events.

When compared to the parameter estimation experiments carried out in the previous section, a more complex virtual hydrological reality was adopted for the generation of the synthetic flood hydrograph. A relatively meaningful spatial variability was imposed for the different model parameters. The hydraulic conductivity  $K$  is decreasing linearly with the distance to the hydrographic network and the roughness coefficient  $n$  on the hillslopes is assigned according to a land-use classification derived from a SPOT satellite image. The friction coefficient is constant in the drainage network and the initial soil moisture  $\theta$  constant over the catchment. The initial condition for the optimization consists in the values specified in Table 1 uniformly applied on the catchment.



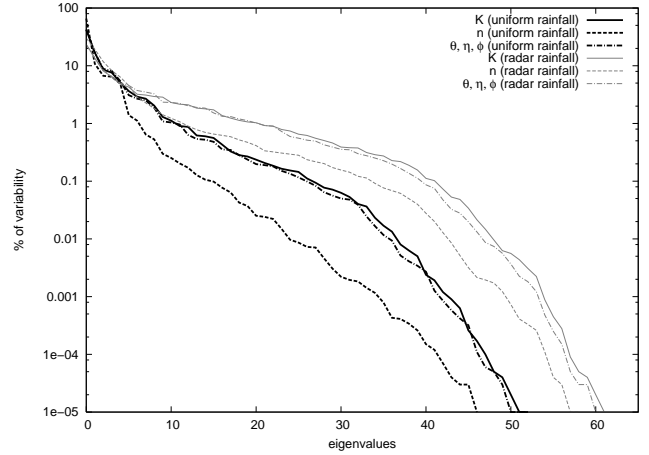
**Fig. 8.** Singular vector in the observation space for the roughness coefficient  $n$ .

The calibration problem is then tackled using parametrizations of increasing dimensionality. The simpler parametrization  $P_1$  assigns a single scalar multiplier for each parameter over the entire catchment area (i.e.  $n_K = n_n = n_\theta = 1$  in Table 3). For  $P_2$ , a scalar multiplier is applied to the hillslopes and another one to the drainage network for the hydraulic conductivity and friction coefficient (i.e.  $n_K = n_n = 2$ ).

Then, apart from the parameter  $\theta$ , the number of degrees of freedom is gradually increased by taking as a basis the singular vectors driving  $X\%$  of the variability for parameters  $K$  and  $n$  (parametrizations  $P_{SVX}$  in the table). As reported during the sensitivity analysis (Sect. 4.2.2), the number of degrees of freedom required for the roughness coefficient is much lower than that obtained for the hydraulic conductivity. The number of degrees of freedom for each parameter, the Nash performance for the estimated parameter set and the inverse of the condition number are given in Table 3. The condition number was calculated with an approximation of the hessian after the last BFGS update (i.e. at the optimum) of the quasi-newton algorithm. It is reminded that the larger the ratio  $1/\kappa(H)$ , the better is the conditioning of the optimization problem.

From the results shown in Table 3, it seems that using this description of the parameter space the number of control variables can be increased without altering the conditioning of the optimization problem. The previous statement is valid as long as the vectors describing the kernel in the parameter space (the specified degrees of freedom which do not significantly alter the hydrological response) are not introduced in the parametrization. The results obtained with  $P_{SV90}$  show that even with noise-free observations, the use of those directions for the description of the affordable sub-space lead to instability in the inverse problem.

However, it is interesting to note that with respectively 7 and 10 degrees of freedom (i.e.  $P_{SV70}$  and  $P_{SV80}$ ), the



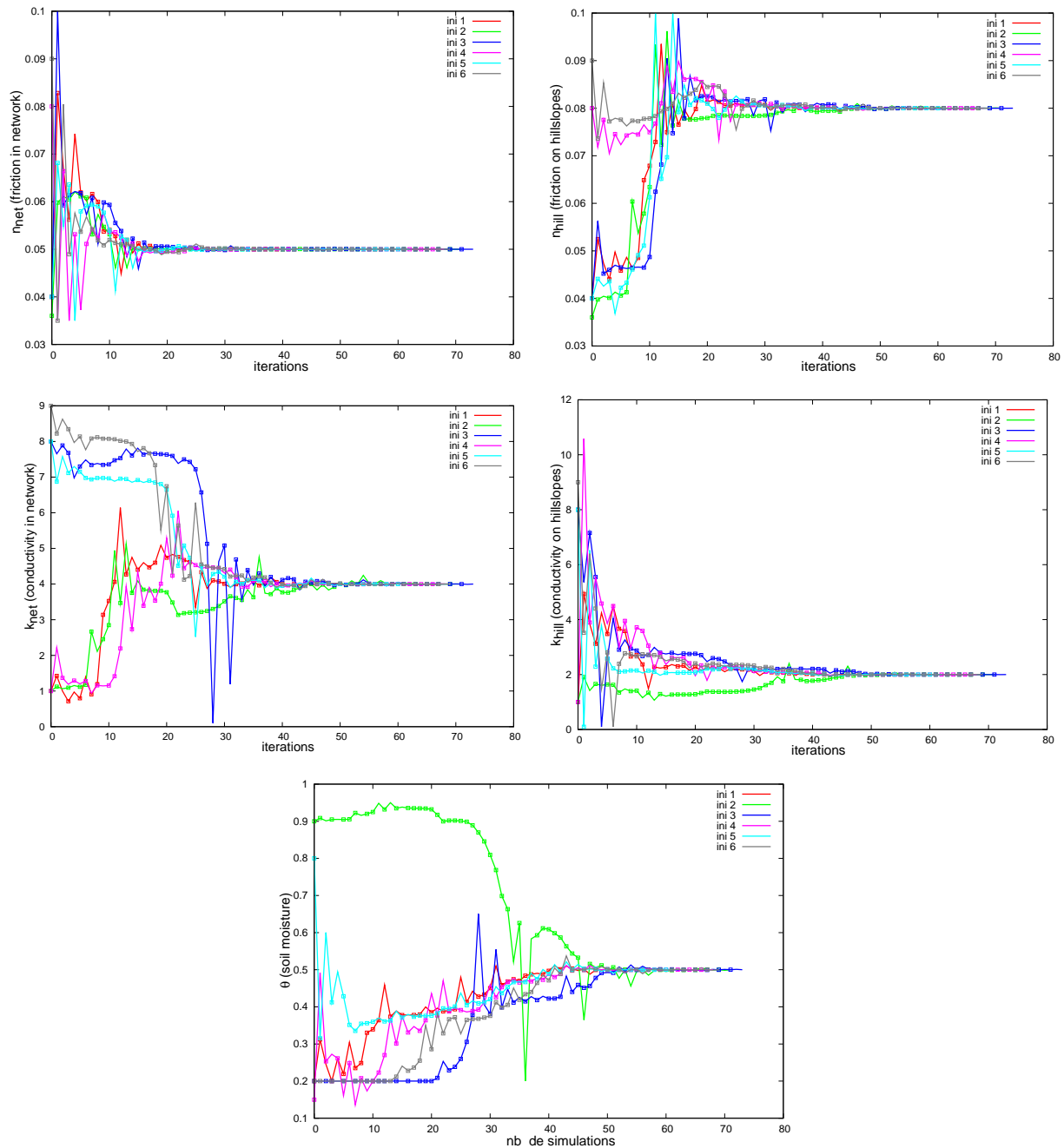
**Fig. 9.** Singular values spectrum for lumped and spatially variable rainfall.

**Table 3.** Complexity (label and number of degrees of freedom for the parameters  $K$ ,  $n$  and  $\theta$ ), Nash efficiency (with synthetic noise free observations) and conditioning for the different parametrizations.

	$n_K$	$n_n$	$n_\theta$	Nash	$1/\kappa(H)$
$P_1$	1	1	1	0.908	0.965E-08
$P_2$	2	2	1	0.938	0.217E-11
$P_{SV70}$	4	2	1	0.968	0.889E-08
$P_{SV80}$	6	3	1	0.978	0.947E-08
$P_{SV90}$	9	5	1	0.986	0.242E-16

conditioning is even better than the one obtained with parametrization  $P_2$  (5 degrees of freedom). As emphasised by Tonkin and Doherty (2005), the subspace determined from the truncated singular value decomposition of the Jacobian (TSVD) is determined from the observations information content whereas the subspace constructed from a prior parsimony strategy is not. In the previously cited contribution, the Jacobian matrix was approximated using finite differences and used in the linearized equations of the Levenberg-Marquardt method. In order to prevent over-fitting and combine the advantages of TSVD and Tikhonov regularizations an hybrid regularization methodology was proposed.

In fact, while the truncation level of the SVD is the only mechanism for preventing over-fitting, the penalization term of the Tikhonov approach is also a way to insert prior information on the parameters. In the experiments carried out in this paper, the hydrological reality is known. The comparison with the estimated parameters have shown that the improvement of performances in terms of Nash efficiency do not necessarily come with parameters closer to the reference values.



**Fig. 10.** Convergence of the model parameters to the reference values for various initial parameter sets using quasi-newton algorithm. The displayed iterations contain both gradient calculations with the adjoint (marked with a square) and model evaluations for the line search.

## 6 Discussion and conclusions

Using a relatively simple application case, it has been shown that the potential of variational methods for distributed catchment scale hydrology should be considered. Although for this particular application many outcomes are limited to evidence retrieval, the adopted approach should be further exploited.

The fact that other techniques might offer a practical advantage (quick and easy implementation) when compared to variational methods cannot be contested. However, the practical implementation of the adjoint state method is largely facilitated by the advent of very efficient automatic differentiation engines such as the one adopted for this study. The use of this approach is at best anticipated in the development stage of hydrological models but remain affordable for existing computer models.

It is important to emphasise that a single integration of the adjoint code, encompassing the forward integration of the direct model and the backward integration of the adjoint model, yields all spatial and temporal sensitivities (Hall and Cacuci, 1983). The key advantage of this technique is that the computational cost is independent from the dimension of the control space. The results provided in this paper show that spatially distributed parameter sensitivities can be obtained for a very modest calculation effort ( $\sim 6$  times the computing time of a single model run). The analysis of an essential aspect of the simulated hydrograph such the maximum discharge has shown that the influence of the friction coefficient assigned at different spatial locations was characterised by relatively complex sensitivity patterns.

For the analysis of a vectorial response such as the flood hydrograph, the Jacobian of the transformation can be calculated using the adjoint technique. The physical interpretation of the singular vectors in the parameter and observations spaces brings out relevant features of the rainfall-runoff transformation. Furthermore, the analysis of the singular spectrum can be used to apprehend the complexity of an affordable parametrization and compare the information content of different rainfall events. Sensitivity analysis can be motivated by different goals. For understanding the model behaviour for parameter values leading to an acceptable fit to the available observations, the analysis should be carried out with posterior probability distribution functions (PDFs). Using sampling based approaches, this is rarely the case in practice, partly because posterior PDFs are often characterized by dependence which can be difficult to represent and compromise the use of many existing global sensitivity analysis techniques (see Kanso et al., 2006 for one of the few attempts). For this specific setting, local sensitivity analysis at the best estimate might prove very informative. Many of the outcomes, such as the spatially distributed importance measures, are mainly driven by the topography of the catchment and the spatial variability of the rainfall forcing rather than the specific point in the parameter space used for the analysis.

For the estimation of model parameters, even when it is obtained with an empirical dimension reduction (e.g. scalar multipliers), uni-modality or at least limited multi-modality can be achieved for many hydrological models. In this case, gradient-based optimization techniques are indisputably the most efficient (i.e. convergence to single mode or exploration of multi-modality). As underlined by Kavetski et al. (2006) convergence safeguards such as line-searches and trust-regions in modern gradient-based algorithms improve significantly the reliability of the estimates. The accuracy of the gradient can be essential and to this regard, the approach used in this paper avoid the specification of a perturbation parameter (required for finite difference approximation). The derivatives computed with the reverse mode of algorithmic differentiation (i.e. adjoint method) were found exceedingly efficient in driving a bound-constrained quasi-newton optimization

algorithms to the reference values used to generate synthetic observations.

Although the number of gradient/model evaluations is already eloquent, the authors acknowledge that a comparison with global non-smooth optimization techniques such as the Shuffle Complex Evolution from Duan et al. (1992) would strengthen the argument of the paper. However, the mathematical representation of hydrological processes in distributed models tend to produce smoother response surfaces. Although the presence of thresholds remains in the model formulations, they usually occur at the grid element level and do not produce discontinuous derivatives for the cost function. In fact, as far a discharge is concerned, the cost functions used for the calibration of model parameters involve an integration of the residuals over time for an integrated hydrological response (i.e. spatio-temporal smoothing).

As advocated by Moore and Doherty (2006a,b), empirical dimension reduction do not allow all the information to be extracted from observation data in order to reduce the predictive model error. Parameters at different locations over the surface of the catchment are not equally constrained by the observations of the hydrological response. The identifiable subspace can be described using the truncated singular value decomposition of the Jacobian matrix (TSVD). Doherty and Randall (2008) recently proposed statistics for evaluating parameter identifiability and error reduction using this factorization of the Jacobian matrix.

The experiments carried out with TSVD parametrization show that this technique represents a promising regularization strategy. However, as emphasised by Tonkin and Doherty (2005), it is essential to combine this approach with Tikhonov regularization in order to account for prior information and prevent overfitting. The appropriate calibration paradigm would therefore require a good compromise to be found between flexibility and stability. The objective would be to improve prior values rather than conducting a blind search over an arbitrarily reduced parameter space.

*Acknowledgements.* This work was supported by a bi-annual research contract between the Midi-Pyrénées région and the French government funded and managed by the French National Space Agency (CNES). The authors gratefully acknowledge the support of French National Institute for Research in Computer Science and Control (INRIA) and Toulouse Institute of Fluid Mechanics (IMFT).

Edited by: R. Moussa

## References

- Bastidas, L. A., Gupta, H., Sorooshian, S., Shuttleworth, W., and Yang, Z.: Sensitivity Analysis of a Land Surface Scheme using Multi-Criteria Methods, *J. Geophys. Res.*, 104, 19 481–19 490, 1999.

- Belanger, E. and Vincent, A.: Data assimilation (4D-VAR) to forecast flood in shallow-waters with sediment erosion, *J. Hydrol.*, 300, 114–125, 2005.
- Bennett, A.: *Inverse Methods in Physical Oceanography*, Cambridge Monographs on Mechanics and Applied Mathematics, Cambridge University Press, 1992.
- Beven, K. and Binley, A.: The future of distributed models : model calibration and predictive uncertainty, *Hydrol. Process.*, 279–298, 1992.
- Bouyssel, F., Cassé, V., and Pailleux, J.: Variational surface analysis from screen level atmospheric parameters, *Tellus A*, 51, 453–+, 1999.
- Buizza, R. and Palmer, T. N.: The singular vector structure of the global atmospheric global circulation, *J. Atmos. Sci.*, 52, 1434–1456, 1995.
- Cacuci, D.: Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach, *J. Math. Phys.*, 22(12), 2794–2802, 1981a.
- Cacuci, D.: Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses, *J. Math. Phys.*, 22(12), 2803–2812, 1981b.
- Callies, U., Rhodin, A., and Eppel, D. P.: A case study on variational soil moisture analysis from atmospheric observations, *J. Hydrol.*, 212–213, 95–108, 1998.
- Calvet, J.-C., Noilhan, J., and Bessemoulin, P.: Retrieving the Root-Zone Soil Moisture from Surface Soil Moisture or Temperature Estimates: A Feasibility Study Based on Field Measurements, *J. Appl. Meteorol.*, 37, 371–386, 1998.
- Carrera, J. and Neuman, S.: Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability and solution algorithms, *Water Resour. Res.*, 2, 211–227, 1986.
- Chavent, G.: Identification of functional parameter in partial differential equations, in: *Identification of Parameters in Distributed Systems*, edited by: Goodson R. E., and Polis, M., ASME, New York, 31–48, 1974.
- Clément, F., Khvoenkova, N., Cartalade, A., and Montarnal, P.: Analyse de sensibilité et estimation de paramètres de transport pour une équation de diffusion, approche par état adjoint, INRIA-Rocquencourt, *Projet ESTIME*, Tech. Rep. RR-5132, 2004.
- Cukier, R., Levine, H., and Shuler, K.: Nonlinear sensitivity analysis of multiparameter model systems, *J. Comput. Phys.*, 26, 1–42, 1978.
- Doherty, J. and Randall, J. H.: Two Statistics for Evaluating Parameter Identifiability and Error Reduction, *J. Hydrol.*, 366(1–4), 119–127, 2009.
- Duan, Q., Sorooshian, S., and Gupta, V.: Effective and Efficient Global Optimisation for Conceptual Rainfall-Runoff Models, *Water Resour. Res.*, 28, 1015–1031, 1992.
- Durbiano, S.: Vecteurs caractéristiques de modèles océaniques pour la réduction d'ordre en assimilation de données, Ph.D. thesis, Université Joseph Fourier (Grenoble I), 2001.
- Estupina-Borrell, V., Dartus, D., and Ababou, R.: Flash flood modeling with the MARINE hydrological distributed model, *Hydrol. Earth Syst. Sci. Discuss.*, 3, 3397–3438, 2006, <http://www.hydrol-earth-syst-sci-discuss.net/3/3397/2006/>.
- Ghil, M. and Malanotte-Rizzoli, P.: Data assimilation in meteorology and oceanography, *Adv. Geophys.*, 33, 141–226, 1991.
- Griewank, A.: Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, in: *Frontiers in Appl. Math.*, SIAM, Philadelphia, PA, Vol. 19, 369 pp., 2000.
- Gupta, V. and Sorooshian, H.: The automatic calibration of conceptual catchment models using derivative based optimization algorithms, *Water Resour. Res.*, 21, 473–485, 1985.
- Hall, C. and Cacuci, D.: Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models, *J. Atmos. Sci.*, 40, 2537–2546, 1983.
- Hall, J., Tarantola, S., Bates, P., and Horritt, M.: Distributed Sensitivity Analysis of Flood Inundation Model Calibration, *J. Hydrol. Eng.-ASCE*, 131, 117–126, 2005.
- Hansen, P.: *Rank-Deficient and Discrete Ill-Posed Problems. Numerical Aspects of Linear Inversion*. SIAM, Philadelphia, 1998.
- Hascoët, L. and Pascual, V.: TAPENADE 2.1 user's guide, Tech. Rep. RT-0300, Institut National de Recherche en Informatique et en Automatique (INRIA), 2004.
- Helton, J. C.: Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal, *Reliab. Eng. Syst. Safe.*, 42, 327–367, 1993.
- Homma, T. and Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models, *Reliab. Eng. Syst. Safe.*, 52, 1–17, 1996.
- Honorat, M., Monnier, J., Lai, X., and Le dimet, F.-X.: Variational data assimilation for 2D fluvial hydraulics simulation, in: *CMWR XVI-Computational Methods for Water Resources*, Copenhagen, 2006.
- Hornberger, G. and Spear, R.: An approach to the preliminary analysis of environmental systems, *J. Environ. Manage.t*, 12, 7–18, 1981.
- Ibbitt, R. and O'Donnell, T.: Designing conceptual catchment models for automatic fitting methods, *IAHS Publication*, 101, 462–475, 1971.
- Johnston, P. and Pilgrim, D.: Parameter optimization for watershed models, *Water Resour. Res.*, 12, 477–486, 1976.
- Kanso, A., Chebbo, G., and Tassin, B.: Application of MCMC-GSA model calibration method to urban runoff quality modeling, *Reliab. Eng. Syst. Safe.*, 91, 1398–1405, 2006.
- Kavetski, D., Kuczera, G., and Franks, S. W.: Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis, *J. Hydrol.*, 320, 187–201, 2006.
- Kuczera, G. and Parent, E.: Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm, *J. Hydrol.*, 211, 69–85, 1998.
- Le Dimet, F.-X. and Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations, *Tellus A*, 38, 97–110, 1986.
- Lemarchal, C. and Panier, E.: *Les modules M2QN1 et MQHESS*, 2000.
- Li, Z., Navon, I., and Hussaini, M. Y.: Analysis of the singular vectors of the full-physics Florida State University Global Spectral Model, *Tellus A*, 57, 560–574, 2005.
- Lions, J.: *Optimal control of systems governed by partial differential equations*, Springer-Verlag, 1968.
- Madsen, H.: Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Adv. Water Resour.*, 26, 205–216, 2003.
- Mahfouf, J.-F.: Analysis of Soil Moisture from Near-Surface Parameters: A Feasibility Study, *J. Appl. Meteorol.*, 30, 1534–1547, 1991.

- Marchand, E., Clément, F., Roberts, J. E., and Pépin, G.: Deterministic sensitivity analysis for a model for flow in porous media, *Adv. Water Resour.*, 31(8), 1025–1037, 2008.
- Margulis, S. and Entekhabi, D.: A coupled Land Surface-Boundary Layer Model and its adjoint, *J. Hydrometeorol.*, 2, 274–296, 2001.
- McCuen, R.: Component sensitivity: a tool for the analysis of complex water resources systems, *Water Resour. Res.*, 9(1), 243–247, 1973a.
- McCuen, R.: The role of sensitivity analysis in hydrologic modelling, *J. Hydrol.*, 18, 37–53, 1973b.
- Moore, C. and Doherty, J.: The cost of uniqueness in groundwater model calibration, *Adv. Water Resour.*, 29, 605–623, 2006a.
- Moore, C. and Doherty, J.: Role of the calibration process in reducing model predictive error, *Water Resour. Res.*, 41, W05020, doi:10.1029/2004WR003501, 2006b.
- Navon, I.: Practical and Theoretical Aspect of Adjoint Parameter Estimation and Identifiability in Meteorology and Oceanography, *Dynam. Atmos. Oceans*, 27, 55–79, 1998.
- Ngnepieba, P., Le Dimet, F.-X., Boukong, A., and Nguetseng, G.: Inverse problem formulation for parameters determination using the adjoint method, *ARIMA Journal – Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées*, 1, 127–157, 2002.
- Piasecki, M. and Katopodes, N.: Control of Contaminant Releases in Rivers and Estuaries. Part I: Adjoint Sensitivity Analysis, *J. Hydraul. Eng.-ASCE*, 123, 486–492, 1997.
- Rall, L. B.: Automatic Differentiation: Techniques and Applications. Lecture Notes in Computer Science 120, Springer, Berlin, 1981.
- Refsgaard, J.: Parameterisation, calibration and validation of distributed hydrological models, *J. Hydrol.*, 198, 69–97, 1997.
- Reichle, R., Entekhabi, D., and McLaughlin, D.: Downscaling of radiobrightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach, *Water Resour. Res.*, 37(9), 2353–2364, 2001.
- Saltelli, A., Chan, K., and Scott, E.: Sensitivity analysis, Wiley series in probability and statistics, Wiley, 2000.
- Sei, A. and Symes, W.: A note on consistency and adjointness for numerical schemes, Tech. rep., Department of Computational and Applied Mathematics, Rice University, Tech. Report TR95-04, 1995.
- Seo, D., Koren, V., and Cajina, N.: Real-Time Variational Assimilation of Hydrologic and Hydrometeorological Data into Operational Hydrologic Forecasting, *J. Hydrometeorol.*, 4, 627–641, 2003a.
- Seo, D., Koren, V., and Cajina, N.: Real-time assimilation of radar-based precipitation data and streamflow observations into a distributed hydrologic model, Proceedings of Symposium HS03 held during IUGG2003 at Sapporo, IAHS Publ., 282, 138–142, 2003b.
- Singh, V.: Kinematic wave in water resources: a historical perspective, *Hydrol. Process.*, 15, 671–706, 2001.
- Sirkes, Z. and Tziperman, E.: Finite difference of adjoint or adjoint of finite difference?, *Mon. Weather rev.*, 49, 5–40, 1997.
- Sobol', I.: Sensitivity analysis for non-linear mathematical models, *Mathematical Modeling & Computational Experiment*, 1, 407–414, 1993 (Engl. transl.).
- Sun, N.-Z. and Yeh, W.-G.: Coupled inverse problems in groundwater modeling I. Sensitivity and parameter identification, *Water Resour. Res.*, 26, 2507–2525, 1990.
- Tang, Y., Reed, P., van Werkhoven, K., and Wagener, T.: Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis, *Water Resour. Res.*, 43, W06415, doi:10.1029/2006WR005813, 2007a.
- Tang, Y., Reed, P., Wagener, T., and van Werkhoven, K.: Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, *Hydrol. Earth Syst. Sci.*, 11, 793–817, 2007b, <http://www.hydrol-earth-syst-sci.net/11/793/2007/>.
- Tikhonov, A. N. and Arsenin, V. Y.: Solutions of Ill-posed Problems, W. H. Winston, Washington, DC, 1977.
- Tonkin, M. J. and Doherty, J.: A hybrid regularized inversion methodology for highly parametrized environmental models, *Water Resour. Res.*, 41, W10412, doi:10.1029/2005WR003995, 2005.
- Van Werkhoven, K., Wagener, T., Reed, P., and Tang, Y.: Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, 44, W01429, doi:10.1029/2007WR0062711, 2008a.
- Van Werkhoven, K., Wagener, T., Reed, P., and Tang, Y.: Rainfall characteristics define the value of streamflow observations for distributed watershed model identification, *Geophys. Res. Lett.*, 35, L11403, doi:10.1029/2008GL034162, 2008b.
- Vrugt, J. A., Bouten, W., Gupta, H. V., and Sorooshian, S.: Toward improved identifiability of hydrologic model parameters: The information content of experimental data, *Water Resour. Res.*, 38(12), 1312, doi:10.1029/2001WR001118, 2002.
- Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, 39(8), 1214, doi:10.1029/2002WR001746, 2003a.
- Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, 39(8), 1201, doi:10.1029/2002WR001642, 2003b.
- Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., and Gupta, H. V.: Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis, *Hydrol. Process.*, 17, 455–476, 2003.
- White, L., Vieux, B., Armand, D., and Le Dimet, F.-X.: Estimation of optimal parameters for a surface hydrology model, *Adv. Water Resour.*, 26(3), 337–348, 2003.
- Yapo, P., Gupta, H., and Sorooshian, S.: Multi-objective global optimization for hydrologic models, *J. Hydrol.*, 204, 83–97, 1997.
- Yatheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M., and Stewart, A.: Understanding uncertainty in distributed flash flood forecasting for semiarid regions, *Water Resour. Res.*, 44, W05S19, doi:10.1029/2007WR005940, 2008.
- Zou, X., Navon, I. M., and Sela, J.: Variational data Assimilation with Moist Threshold Processes using the NMC Spectral Model, *Tellus A*, 45(5), 370–387, 1993.