Hydrol. Earth Syst. Sci. Discuss., 7, 9067–9121, 2010
www.hydrol-earth-syst-sci-discuss.net/7/9067/2010/
doi:10.5194/hessd-7-9067-2010

Hydrology and Earth System Sciences Discussions

Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting

S. Ricci[1], A. Piacentini[1], O. Thual[1,2], E. Le Pape[3], and G. Jonville[4]

[1]URA CERFACS/CNRS, URA 1875, Toulouse, France
[2]INPT, CNRS, IMFT, Toulouse, France
[3]SCHAPI, Toulouse, France
[4]CERFACS, Toulouse, France

Received: 28 September 2010 – Accepted: 5 November 2010 – Published: 30 November 2010

Correspondence to: S. Ricci (ricci@cerfacs.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The present study describes the assimilation of river water level observations and the resulting improvement of the river flood forecast. The BLUE algorithm was built on top of the one-dimensional hydraulics model MASCARET. The assimilation algorithm folds in two steps: the first one is based on the assumption that the upstream flow can be adjusted using a three-parameter correction, the second one consists in directly correcting the hydraulic state. This procedure is applied on a four-day sliding window over the whole flood event. The background error covariances for water level and discharge are represented with asymmetric correlation functions where the upstream correlation length is bigger than the downstream correlation length. This approach is motivated by the implementation of a Kalman Filter algorithm on top of an advection-diffusion toy model. The assimilation study with MASCARET is carried out on the Adour and the Marne Vallage (France) catchments. The correction of the upstream flow as well as the control of the hydraulic state along the flood event leads to a significant improvement of the water level and discharge in analysis and forecast modes.

1 Introduction

River streamflow forecast is a challenging issue for the security of the persons and the infrastructures, the exploitation of power plants and the management of water ressources. Many efforts have been made on the development of open channel flow modeling, based on mass, momentum and energy conservation equations (Chow, 1959; Hervouet, 2003). Still, the uncertainties on these models are such that river streamflow modeling remains a streneous work. Major uncertainties come from the model itself as the physics are simplified and then discretized, but also from hydrological boundary conditions (upstream flow or lateral additional discharge), meteorological boundary conditions (precipitation, pressure and wind) and from hydrological initial conditions. Hydraulics models also rely on various parameterization expressed as numerical parameters (stability coefficient for the numerical scheme), geometry parameters (cross sections, gates and weir dimensions) and hydraulic parameters (flood plain storage, friction, discharge). Calibrating a hydraulics model often means adjusting Strickler coefficients, discharge coefficients at cross or lateral devices, seepage values or cross section geometry. The calibration of these parameters has been widely investigated (Beven and Freer, 2001; Malaterre et al., 2010) focusing either on calibration algorithms, sensitivity indications or optimality of the observations network.

Both parameter calibration and physical field description can be formulated as inverse problems (Tarantola, 1987). The formulation of the inverse problem in hydrology fits into a wider mathematical framework presented by Maclaughlin and Townley (1996). Data assimilation combines numerical and observational information on a system in order to provide a better description of it (Ide et al., 1997; Boutier and Courtier, 1999; Kalnay, 2003). The benefit of data assimilation has already been greatly demonstrated in meteorology (Parrish and Derber, 1992; Rabier et al., 2000) and oceanography (Brasseur and Verron, 2006) over the past decades, especially for providing initial conditions for numerical forecast. Data assimilation is now being applied with increasing frequency to hydrology (Thirel et al., 2010, Part I; Thirel et al., 2010, Part II,) and hydraulics problems with two main objectives: optimizing model parameters and improving streamflow simulation and forecast. The literature proposes several methods based on minimization techniques approaches (Atanov et al., 1999; Das et al., 2004; Honnorat et al., 2007; Bessières et al., 2007). The filtering approach, e.g. Kalman filter or Monte Carlo algorithms, also enables the estimation of roughness coefficients (Sau et al., 2010; Pappenberger et al., 2005) and the correction of the physical fields (Jean-Baptiste et al., 2010).

The present study describes the assimilation of river water level observations and the resulting improvement of the river flood forecast. The data assimilation algorithm is built on top of the one-dimensional hydraulics model MASCARET. Given the relatively small dimension of our problem, a filtering technique was applied (further referenced as BLUE). This study focuses on the modeling of the background error covariances in the context of data assimilation for mono-dimensional hydraulics. Asymmetric correlation functions are used to represent the spatial error correlations for water level and discharge. This choice is motivated by an example on a simplified model. For that purpose, a Kalman Filter algorithm is implemented on top of an advection-diffusion toy model. This exercice shows that the analysis and the dynamics of the physics modify a gaussian correlation function into an asymmetric function at the observation point. The assimilation study with MASCARET is performed on the Adour (France) and the Marne Vallage (France) catchments. The improvement of the water level, using data assimilation, in analysis and forecast modes are shown. Most illustrations in this paper present the results on the Adour catchment.

The outline of the paper is as follows. Section 2 describes the assimilation system, paying particular attention to the choice of the control vector for the BLUE algorithm. Two approaches are implemented: the correction of the hydraulic state and the control of the upstream flow. The modeling of the background covariances matrix and the parametrization for the control of the upstream flow are highlighted. Section 3 gives the theoretical frame for the choice of asymmetric correlation functions for the spatial error correlations in the background error covariance matrix **B**. For the MASCARET application, the correlation length scale is estimated following the evolution of a perturbation on the initial state. In Sect. 4, the improvement of the river flood simulation and forecast is presented. The evaluation of usual hydraulics criteria such as precision, in re-analysis or forecast mode, illustrates the assimilation scheme properties. A summary and a discussion are finally given in Sect. 5.

HESSD
7, 9067–9121, 2010

2 Context and implementation of the data assimilation

2.1 Modeling of the physics

MASCARET is a one-dimensional free surface hydraulics model developed by EDF and CETMEF[1], based on the Saint-Venant equations (Goutal and Maurel, 2002). MASCARET is widely used for the modeling of river flood, submersion waves resulting from hydraulic infrastructure breaking, river control and canal waves propagation.

The conservative form of the mono-dimensional Saint Venant equations reads:

$$\frac{\partial S}{\partial t}+\frac{\partial Q}{\partial x}=q_a, \quad \frac{\partial Q}{\partial t}+\frac{\partial}{\partial x}(Q^2/S)+gS\frac{\partial Z}{\partial x}=-\frac{gQ^2}{SK_s^2R_H^{4/3}}. \tag{1}$$

In this formulation the river section S is expressed in m^2 and is, at each location x, a function of the water height $h=Z(x,t)-Z_{\text{bottom}}(x,t)$ where $Z(x,t)$ is the free surface height in m. The discharge in $m^3 s^{-1}$ is denoted by $Q(x,t)$, $q_a(x,t)$ is the lateral lineic discharge, K_s is the Strickler coefficient, R_H is the hydraulic radius and g is the gravity.

The non stationnary mode of MASCARET is used in this study. Significant uncertainties on the input parameters of MASCARET, such as the Strickler coefficient or the upstream flow and lateral discharge, result in errors on the simulated water level and discharge. The aim of our data assimilation approach is to reduce the uncertainties of either the inputs or the outputs of the simulation.

2.2 The assimilation method BLUE

The Best Linear Unbiased Estimator (BLUE) approach (Gelb, 1974; Talagrand, 1997) identifies the optimal estimator of the true value of an unknown variable x. This estimator is optimal as its variance is minimum, meaning, for gaussian cases, that its probability density function is dense around its mean. $\boldsymbol{x}$ is the control vector and can

[1] Centre d'Etudes Techniques Maritimes Et Fluviales

HESSD
7, 9067–9121, 2010

stand for the state variables (water level and discharge for MASCARET), the model parameters (Strickler coefficients), the boundary conditions (upstream flows) or the initial condition (initial water level and discharge), or a mix of these. The solution of the BLUE algorithm is the analysis vector $\boldsymbol{x}^{\mathrm{a}}$. The a priori knowledge of the system is the background vector $\boldsymbol{x}^{\mathrm{b}}$ and the observation vector is $\boldsymbol{y}^{\mathrm{o}}$. The background, observation and analysis error covariances are respectively gathered in the matrices **B**, **R** and **A**. Assuming that the background, the observation and the analysis are unbiased, the analysis can be formulated as a correction to the background state defined as:

$$\boldsymbol{x}^{\mathrm{a}} = \boldsymbol{x}^{\mathrm{b}} + \mathbf{K}\boldsymbol{d}, \tag{2}$$

where **K** is the gain matrix and $\boldsymbol{d}$ is the innovation vector

$$\boldsymbol{d} = \boldsymbol{y}^{\mathrm{o}} - H(\boldsymbol{x}^{\mathrm{b}}), \tag{3}$$

and $\boldsymbol{y} = H(\boldsymbol{x})$ is the model equivalent of the observations through the observation operator H.

The BLUE analysis is optimal as the variance of its error is minimum. Minimizing the variance of the analysis comes down to minimizing the trace of the analysis error covariance matrix and leads to the formulation of the gain matrix (Boutier and Courtier, 1999):

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{T}(\mathbf{H}\mathbf{B}\mathbf{H}^{T} + \mathbf{R})^{-1}. \tag{4}$$

In this formulation, **H** is the Jacobian matrix of H around the background state $\boldsymbol{x}^{\mathrm{b}}$ that can be written as:

$$\mathbf{H} = \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \frac{\partial H(\boldsymbol{x})}{\partial \boldsymbol{x}}. \tag{5}$$

The analysis error covariance matrix reads

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}. \tag{6}$$

Real-time forecast systems for meteorology and oceanography usually rely on the cycling of such algorithm, though the solution is often identified through a minimisation procedure rather than the direct computation of Eq. (2).

HESSD

7, 9067–9121, 2010

Data assimilation for flood forecasting

S. Ricci et al.

The analysis at time $i-1$ is propagated in time by the dynamical model $M_{i-1,i}$ to define the background at time i (Eq. 7). It is then corrected to provide the analysis $\boldsymbol{x}_i^{\mathrm{a}}$ at time i given by Eq. (8).

$$\boldsymbol{x}_i^{\mathrm{b}} = M_{i-1,i}(\boldsymbol{x}_{i-1}^{\mathrm{a}}), \tag{7}$$

$$\boldsymbol{x}_i^{\mathrm{a}} = \boldsymbol{x}_i^{\mathrm{b}} + \mathbf{K}_i \left[\boldsymbol{y}_i^{\mathrm{o}} - H_i(\boldsymbol{x}_i^{\mathrm{b}}) \right], \tag{8}$$

with

$$\mathbf{K}_i = \mathbf{B}_i \mathbf{H}_i^T (\mathbf{H}_i \mathbf{B}_i \mathbf{H}_i^T + \mathbf{R}_i)^{-1}, \tag{9}$$

where $M_{i-1,i}$ represents the model propagation between $i-1$ and i of the physics in Eq. (1), $\mathbf{B}_i$, $\mathbf{R}_i$ are the background and observation error covariance matrices at time i and H_i is the observation operator at time i and $\mathbf{H}_i$ is its linear approximation at time i.

The analysis error covariance matrix $\mathbf{A}_{i-1}$ is computed from Eq. (10) at each assimilation time. The analysis error covariance matrix at time $i-1$ is propagated in time by the dynamical model to define the background error covariance matrix at time i as written in Eq. (11), where $M_{i-1,i}$ is the tangent linear approximation of $M_{i-1,i}$.

$$\mathbf{A}_{i-1} = (\mathbf{I} - \mathbf{K}_{i-1} \mathbf{H}_{i-1}) \mathbf{B}_{i-1} \tag{10}$$

$$\mathbf{B}_i = \mathbf{M}_{i-1,i} \; \mathbf{A}_{i-1} \; \mathbf{M}_{i-1,i}^T. \tag{11}$$

Equations (7–10) are the Kalman Filter equations (Todling and Cohn, 1994) with no error model. If the gain matrix $\mathbf{K}_i$ is kept constant over time and $\mathbf{M}_{i-1,i} = \mathbf{I}$ is assumed, then they come down to the BLUE equations applied in our study.

2.3 Implementation of the assimilation scheme

The simulated water levels with MASCARET (or any hydraulics model), can be significantly different from the observed ones. The simulation of past flood events with

MASCARET reveals the model difficulties to catch the rise in the water level and to often underestimate the flood peak for moderate events and overestimate the flood peak for important flood events. Two data assimilation schemes were implemented on top of the MASCARET model.

2.3.1 Correction of the hydraulic state

The first data assimilation approach consists in dynamicall correcting the water level and the discharge states for the entire catchment (discretized in m cells) when observations are available. In this case, the control vector is composed of the discretized water level and discharge states $\boldsymbol{x} = \left(Z_{x_1},\cdots,Z_{x_m},Q_{x_1},\cdots,Q_{x_m}\right) = (\mathbf{Z},\mathbf{Q})$.

The background state is given by a previous integration of the model, it is the simulated water level and discharge state denoted $\left(\mathbf{Z}^b,\mathbf{Q}^b\right)$. The size of the control and the background vectors is $n = 2\,\mathrm{m}$. The observation vector contains water level at observation times and at selected locations on the hydraulic network. It is a vector of size p where p is the number of observations. The observation operator sums up to a selection matrix $p \times n$, denoted by $\mathbf{H}_{\mathrm{sel}}$ that can be written as

$$\mathbf{H}_{\mathrm{sel}} = \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}$$

where the non zero values correspond to the location of the observation point on the state vector $\boldsymbol{x}$. When the observation points do not correspond to positions on the grid, the selection matrix $\mathbf{H}_{\mathrm{sel}}$ then represents an interpolation operation or a neighbouring operation.

HESSD

7, 9067–9121, 2010

Data assimilation for flood forecasting

S. Ricci et al.

The observation covariance matrix $\mathbf{R}$ is a $p \times p$ matrix. Its diagonal terms are the observation error variances at the observation points and the off-diagonal terms are the covariances between the observation errors at different observation points. The background covariance matrix is a $n \times n$ symmetric positive-definite matrix that can be represented by blocks:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{Z,Z} & \mathbf{B}_{Z,Q}^{T} \\ \mathbf{B}_{Z,Q} & \mathbf{B}_{Q,Q} \end{pmatrix}.$$

The $n \times n$ diagonal blocks $\mathbf{B}_{Z,Z}$ and $\mathbf{B}_{Q,Q}$ represent the statistics for the errors ϵ_Z of the water level and ϵ_Q of the discharge, respectively. Its diagonals represent the variance of the background error for the water level and the discharge respectively whereas the extra diagonal terms of these blocks are the covariances between the error on the water level or the discharge at different locations on the grid. These covariances are commonly defined as *univariate* as opposed to the *multivariate* covariances in the extra-diagonal blocks $\mathbf{B}_{Z,Q}$ and $\mathbf{B}_{Z,Q}^{T}$ that represent the covariances between the errors on the water level and the errors on the discharge.

The innovation vector $\boldsymbol{d}$ (Eq. 3) expresses the difference between the simulated and the observed water level at observation points x_{obs} and times. At these locations, since the observation operator is a simple selection operator in our case, the innovation is weighed by the matrix product $\mathbf{H}_{\text{sel}}^{T}(\mathbf{H}_{\text{sel}}\mathbf{B}\mathbf{H}_{\text{sel}}^{T}+\mathbf{R})^{-1}$ in Eq. (4):

$$\widetilde{\delta Z} = \mathbf{H}_{\text{sel}}^{T}(\mathbf{H}_{\text{sel}}\mathbf{B}\mathbf{H}_{\text{sel}}^{T}+\mathbf{R})^{-1}\boldsymbol{d} \tag{12}$$

where $\widetilde{\delta Z}$ is the water level correction vector at the observation points that reads

$$\widetilde{\delta Z} = (\widetilde{\delta Z}_1, \cdots, \widetilde{\delta Z}_{\text{l}}, \cdots, \widetilde{\delta Z}_{\text{p}}) \tag{13}$$

with $l \in \{1, \cdots, p\}$.

Water level correction δZ for the whole domain results from the multiplication of $\widetilde{\delta Z}$ by $\mathbf{B}_{Z,Z}$. The water level variances translate the uncertainties on the simulated water level. An asymmetric correlation function ρ is used to describe the spatial error

HESSD
7, 9067–9121, 2010
Data assimilation for flood forecasting
S. Ricci et al.

correlations on δZ. The correlation function at an observation point such that $\boldsymbol{x}_{\text{obs}} = 250$ is presented on Fig. 1.

The correlations depend on the curvilinear distance between two points and the correlation length was computed with a systematic procedure of propagation of a local perturbation (see Sect. 3).

The discharge correction vector at the observation point $\widetilde{\delta \boldsymbol{Q}} = (\widetilde{\delta Q}_1, \cdots, \widetilde{\delta Q}_l, \cdots)$ with $l \in \{1, \cdots, p\}$, is deduced from $\widetilde{\boldsymbol{\delta Z}}$ at the observation points using the proportionality relation:

$$\widetilde{\delta Q}_{\mathrm{l}} = \frac{Q_{\mathrm{l}}^{\mathrm{b}}}{Z_{\mathrm{l}}^{\mathrm{b}}} \widetilde{\delta Z}_{\mathrm{l}}, \tag{14}$$

where $Q_{\mathrm{l}}^{\mathrm{b}}$ and $Z_{\mathrm{l}}^{\mathrm{b}}$ are the background values for the water level and discharge at the obervation points. As for the water level, the discharge correction $\delta \boldsymbol{Q}$ for the whole domain results from the multiplication of $\widetilde{\delta \boldsymbol{Q}}$ by $\mathbf{B}_{\mathrm{Q,Q}}$. The correlation function and length for $\delta \boldsymbol{Q}$ are those used for $\boldsymbol{\delta Z}$. Complementary work on the modeling of $(\widetilde{\delta \boldsymbol{Z}}, \widetilde{\delta \boldsymbol{Q}})$ error covariances modeling was initiated using a physical calibration procedure at each observation point such as

$$\widetilde{\delta Q}_{\mathrm{l}} = a(\widetilde{\delta Z}_{\mathrm{l}})^{\mathrm{r}} + b \tag{15}$$

with $l \in [1, p]$. Still, because of the tide influence at some observation points, the identification of a satisfactory approximation of these relations valid for both high tide and low tides was not always possible.

2.3.2 Correction of the upstream flow

The second data approach is based on the assumption that the error on the simulated water level is mostly due to an imperfect description of the upstream flows. These upstream flows, prescribed to the hydraulics model as boundary conditions, are usually deduced from water level observations through a calibration procedure. A considerable

HESSD
7, 9067–9121, 2010

part of the error on the simulated water level can be attributed to the uncertainty on these upstream free extremities. In order to control the uncertainty on this upstream flow, a data assimilation procedure was set up. For this approach, the control vector should contain the discharge boundary conditions at the upstream stations over the simulation period. The relation between the control space and the observation space is given by the integration of the numerical model, this operation being non linear. Still, the correction of the upstream flow for each model time step would involve a large control vector and a heavy assimilation procedure, especially for the computation of the Jacobian of the observation operator.

For that reason, the upstream flow forcing f is corrected through three linear transformations over a time window (assimilation window):

$$\widetilde{f}(t) = af(t-c) + b. \tag{16}$$

The characteristics of this data assimilation approach are:

- The background values for the control parameters are $\boldsymbol{x}^{\mathrm{b}} = (a^{\mathrm{b}}, b^{\mathrm{b}}, c^{\mathrm{b}}) = (1,0,0)$.

- The size of the background error covariance matrix is $(3 \times s)^2$ where s is the number of upstream stations. The error on the background parameters $(a^{\mathrm{b}}, b^{\mathrm{b}}, c^{\mathrm{b}})$ are assumed to be uncorrelated and the variances are estimated statically to represent the variability of the upstream flow.

- The relation between the control space and the observation space is non linear as it implies the integration of the numerical model. The observation operator H_{up} consists in the composition of two operations. The costly one is the integration of the hydraulics model given the upstream flow conditions over the assimilation window. The second one is the selection of the computed water level at the observation points and at the observation times.

- $H_{\mathrm{up}}(\boldsymbol{x}^{\mathrm{b}})$ stands for the water level at the observation points and times computed by MASCARET using the background parameters $(a^{\mathrm{b}}, b^{\mathrm{b}}, c^{\mathrm{b}})$.

HESSD
7, 9067–9121, 2010

- The Jacobian $\mathbf{H}_{\text{up}}$ of H_{up} is the tangent linear of the hydraulics model computed around $\boldsymbol{x}^{\text{b}}$, composed with a selection process at the observation points and times.

The Jacobian matrix $\mathbf{H}_{\text{up}}$ can be approximated around the background $\boldsymbol{x}^{\text{b}}$ as follow:

$$H_{\text{up}}(\boldsymbol{x}^{\text{b}}+\Delta\boldsymbol{x}) \approx H_{\text{up}}(\boldsymbol{x}^{\text{b}})+\mathbf{H}_{\text{up}}|_{\text{b}}\Delta\boldsymbol{x}, \tag{17}$$

where $\mathbf{H}_{\text{up}}|_{\text{b}}$ is discretized using an uncentered finite difference scheme:

$$\mathbf{H}_{\text{up},ij}|_{\text{b}} = \frac{\partial \boldsymbol{y}_i}{\partial \boldsymbol{x}_j} = \frac{\partial H_{\text{up},i}(\boldsymbol{x}^{\text{b}})}{\partial \boldsymbol{x}_j} \approx \frac{H_{\text{up},i}(\boldsymbol{x}^{\text{b}}+\Delta\boldsymbol{x})-H_{\text{up},i}(\boldsymbol{x}^{\text{b}})}{\Delta\boldsymbol{x}_j} = \frac{\Delta y_i}{\Delta x_j}. \tag{18}$$

In our study, Δy_i is the modification of water level at the observation station i resulting from a modification Δx_j of the control j-th variable (a, b or c). The approximation of the operator observation Jacobian is a major hypothesis. The computation of $\mathbf{H}_{\text{up}}$ requires an additional integration of the hydraulics model for each control parameter. An efficient computation of the operator $\mathbf{H}_{\text{up}}$ in the case of a larger control space was implemented by Thirel et al. (2010).

The small size of the control vector, as well as of the observation vector, enables the use of the BLUE algorithm involving matrix operations. Still, the algorithm relies on the hypothesis that the observation operator is linear on the $[\boldsymbol{x}^{\text{b}}, \boldsymbol{x}^{\text{a}}]$ interval. The linearity of the hydraulics model response to a perturbation of the control parameters $(a^{\text{b}}, b^{\text{b}}, c^{\text{b}})$ was investigated. It was shown that the relation between an upstream flow perturbation (of the form Eq. 16) and the hydraulic state response can be reasonably approximated by a linear function in the vicinity of $\boldsymbol{x}^{\text{b}}$.

This implementation allows for an improvement of the simulated water level on the assimilation window but also for an improvement of the forecast since, in forecast mode, the upstream flows is kept constant to the last analysed value.

2.3.3 Cycling of the analysis

The two previously described assimilation approaches are sequentially applied on a period covering a flood event.

The assimilation is performed over a four-day sliding window, also refered to as a cycle, with three days of re-analysis and one day of forecast. The sliding window is shifted every hour and a new assimilation is performed. The forecasted state at one hour is stored and used as the initial state for a following cycle. For the first three days of the event, the simulation starts from a standard state for water level and discharge.

The implementation of the two-step assimilation procedure is represented on Fig. 2. Over the four-day assimilation window, a free run integration of the model is achieved (black curve). The upstream flow correction – correction of the parameters (a,b,c) – is computed using observations from the second and the third days (blue dots). The observation from the first day are not used as the model is potentially not adjusted yet. The resulting analysed parameters are used to correct the upstream flows over the first, second and third days. The updated upstream flows are used for a new integration of the model (starting from the beginning of the four-day window) providing a new integration. This integration (green curve) is intermediate and it describes the background state for the hydraulic state correction procedure. Along the third day of the integration, at each observation time, the water level is adjusted. This correction is instantaneous. The model is integrated starting from the corrected state at an observation time to the next observation time, leading to a discontinuous description of the hydraulic state (discontinuous red curve). In our study, the observation time step is equal to the model time step so that the resulting integration is no more discontinuous than any model integration.

For each cycle, beyond the last observation time, the upstream flows are kept constant and the initial condition for the forecast is given by the last analysed water level and discharge states.

HESSD
7, 9067–9121, 2010

The data assimilation algorithm is implemented using the PALM (Parallel Assimilation with a Lot of Modularity, Lagarde, 2000; Lagarde et al., 2001) dynamic coupler developed at CERFACS. This software was originally developed for the implementation of data assimilation in oceanography for the use of the MERCATOR project. PALM allows the coupling of independent components with a great modularity in the data exchanges and treatment as well as and easy parallelization environment for the application (Fouilloux and Piacentini, 1999; Buis et al., 2006).

3 Modeling of B

3.1 Assumptions for the simplification of the hydraulics model

The modeling of the background covariance matrix in the context of the correction of the hydraulic state of MASCARET with data assimilation requires special attention. The modeling of the univariate covariance function in **B**, i.e. the spatial correlation for the water level on one side and on the discharge on the other side, is investigated here.

Section 3.3 justifies the choice of asymmetric correlation functions at the observation points as opposed to regular gaussian functions. The justification of this approach is made with a simple advection-diffusion model on top of which it was achievable to implement a Kalman Filter. Contrary to the BLUE algorithm, the Kalman Filter algorithm explicitly evolves the background error covariances along the cycling of the analysis. It was shown (see Appendix 3.3) that, starting from gaussian correlation functions in the initial **B**, the Kalman Filter analysis leads to asymmetric covariance functions. Specifying such correlation functions with the BLUE algorithm is then equivalent to using a Kalman Filter algorithm. This result was used for the data assimilation hydraulic state correction procedure on MASCARET on top of which it was too costly to implement a Kalman Filter. The correlation length l for the correlation function was then estimated assuming the correlation function is gaussian (see Appendix B). Finally, the correlation function at the observation point is defined locally as a gaussian of correlation

HESSD

7, 9067–9121, 2010

Data assimilation for flood forecasting

S. Ricci et al.

length l upstream the observation point and as a gaussian of correlation length $l/10$ downstream the observation point. The resulting function is local and asymmetric.

Results in Sect. 3.3 (as well as in Appendix A and Appendix B) are based on the crude assumption that the physics of MASCARET can be approximated by the diffusive flood approximation equations. Our approach is to assume that the solution of the propagation of a given initial condition by the MASCARET equation is close to the one propagated by the diffusive flood wave approximation equations. More precisely, it is assumed in the following, that the covariance function of a signal (and then its correlation length l) propagated by MASCARET is similar to the covariance function of the same signal propagated by the diffusive flood approximation equations. The analytical solution of the MASCARET equations for a given initial condition is usually not known. On the other hand, such solution is known under the flood wave propagation assumption.

In the framework of the diffusive flood wave approximation ($S(x,t) = L h(x,t)$, where L is a constant river width), the diffusive Saint Venant equations Eq. (1) of MASCARET can be crudely approximated by

$$\frac{\partial \tilde{h}}{\partial t} + \frac{5U_n}{3}\frac{\partial \tilde{h}}{\partial x} = \kappa \frac{\partial^2 \tilde{h}}{\partial x^2}, \tag{19}$$

where $Q = hU$, $h = h_n + \tilde{h}$ and $U = U_n + \tilde{U}$, with $(\tilde{h},\tilde{U})$ small perturbations to the equilibrium (h_n,U_n) and $\kappa = \frac{U_n h_n}{2\tan\gamma}$ for a constant slope γ. The state (h_n,U_n) such that $U_n = K_s I^{1/2} h_n^{2/3}$ is a solution of the flood wave approximation equations, where $I = \sin\gamma$. The equilibrium state (h_n,U_n) for the diffusive flood wave model is chosen as a representative mean state for the following simulations with MASCARET on the each catchment. Equation (19) is a classical advection-diffusion equation where κ is the diffusion coefficient and $c = \frac{5U_n}{3}$ is the advection speed. To use this model as a support for data assimilation, an open boundary condition for Eq. (19) is imposed downstream with $\frac{\partial \tilde{h}}{\partial t}(L,t) + c\frac{\partial \tilde{h}}{\partial x}(L,t) = 0$. The upstream boundary condition is imposed by $\tilde{h}(0,t) = \tilde{q}(t)$,

HESSD
7, 9067–9121, 2010

Data assimilation for flood forecasting

S. Ricci et al.

where $\widetilde{q}$ is a random function of zero mean $<\widetilde{q}>=0$. The auto-correlation function of $\tilde{q}(t)$

$$R(\tau)=<\widetilde{q}(t)\widetilde{q}(t+\tau)>=\delta q_{\mathrm{m}}^2 \exp\left(-\frac{\tau^2}{2l_q^2}\right). \tag{20}$$

is assumed to be gaussian.

3.2 BLUE algorithm on a 1-D advection-diffusion model

To calibrate the data assimilation chain with MASCARET, data assimilation twin experiments were set up on the 1-D advection-diffusion model (Eq. 19) with $t\in[0,T]$ and $x\in[0,L]$. The 1-D domain is discretized in m cells and Eq. (19) is integrated using a Euler explicite scheme in time and an order one centered finite differences scheme in space. A reference run is integrated using the set of parameters and forcing $(c_{\mathrm{true}},\kappa_{\mathrm{true}},\widetilde{q}_{\mathrm{true}}(t))$ simulating the so called *true* water level $\widetilde{h}_{\mathrm{true}}$. The observation $\widetilde{h}_{\mathrm{obs}}(t)=\widetilde{h}_{\mathrm{true}}(t)+\epsilon_o(t)$ is calculated at the middle of the 1-D domain ($x_{\mathrm{obs}}=\frac{L}{2}$) where $\epsilon_o(t)$ is a gaussian noise defined by its standard deviation σ_o. The background trajectory $h_b(x,t)$ is integrated using a perturbed set of parameters and forcing $(c_{\mathrm{per}},\kappa_{\mathrm{per}},\widetilde{q}_{\mathrm{per}}(t))$ where $<\widetilde{q}_{\mathrm{per}}(t)\widetilde{q}_{\mathrm{per}}(t+\tau)>=\delta q_{\mathrm{m,per}}^2\exp(-\frac{\tau^2}{2l_q^2})$.

The BLUE analysis (Eq. 2) is computed at the subsequent observation times t, assimilating the single observation h_{obs} every Δt. The background is the water level $h_{\mathrm{b}}(x)$ at the analysis time t. The analysis at time t is used as the background for the assimilation at time $t+\Delta t$. The observation operator is a column matrix $\mathbf{H}=(0,\ldots,0,1,0,\ldots,0)^T$ with 1 for the x_{obs} index. The observation error matrix comes down to the scalar variance observation error σ_{o}^2. The background covariance matrix is a $m\times m$ matrix and the variances σ_{b}^2 are prescribed constant to δq_{m}^2 for the whole domain.

HESSD
7, 9067–9121, 2010

The spatial correlations in **B** are represented by the gaussian function

$$\rho(x,x') = \exp\left[-\frac{(x-x')^2}{2l_\mathrm{B}(x,x')}\right] \tag{21}$$

meaning that $\mathbf{B}(x,x') = \sigma_\mathrm{b}^2\rho(x,x')$. The length $l_\mathrm{B}(x,x')$ is a symmetric function of x and x' which represents the local correlation length for the pair (x,x'). For our study, only the correlations $\rho(x,x_\mathrm{obs})$ between the errors at x_obs and the rest of the domain are relevant. The length $l_\mathrm{B}(x,x_\mathrm{obs})$ is estimated as the variance of the impulse response of the model to a Dirac perturbation of $q(t)$. This length is first assumed to depend only on the observation point and is denoted by $l(x_\mathrm{obs})$.

Figure 3(a,b,c) shows the true $\tilde{h}_t$, the non-assimilated $\tilde{h}_\mathrm{s}$, background $\tilde{h}_\mathrm{b}$ and analysed $\tilde{h}_\mathrm{a}$ water level state over the 1-D domain at $t = T = 500\times10^3$ s when the analysis is performed every $\Delta t = 10\times10^3$ s, for different functions $l_\mathrm{B}(x,x_\mathrm{obs})$. When $l_\mathrm{B}(x,x_\mathrm{obs}) = l(x_\mathrm{obs})$ for all x (Fig. 3a), the data assimilation corrects the water level over the interval $[x_\mathrm{obs} - l(x_\mathrm{obs})/2, x_\mathrm{obs} + l(x_\mathrm{obs})/2]$. Still, the analysis (red curve) is closer to the true state (blue dotted curve) than the background (green curve) only upstream the observation point. On the contrary, when $l_\mathrm{B}(x,x_\mathrm{obs}) = l(x_\mathrm{obs})/10$ for all x, on Fig. 3b, the analysis is closer to the true state only downstream the observation point. Finally, as it appears on Fig. 3c, a better fit to the true state is obtained with an asymmetric function $\rho(x,x_\mathrm{obs})$ with

$$\begin{aligned} l_- &= l_\mathrm{B}(x,x_\mathrm{obs}) = \quad l(x_\mathrm{obs}) \quad \text{when } x < x_\mathrm{obs}, \\ l_+ &= l_\mathrm{B}(x,x_\mathrm{obs}) = l(x_\mathrm{obs})/10 \text{ when } x > x_\mathrm{obs}. \end{aligned} \tag{22}$$

The choice of an asymmetric correlation function described by Eq. (22) is quite unusual and should be treated carefully as it might lead to a non symmetric **B** matrix erroneous setting. Indeed, with an asymmetric correlation function, considering for instance two nearby observation points $x_{\mathrm{obs},1}, x_{\mathrm{obs},2}$ with $l(x_{\mathrm{obs},1}) = l(x_{\mathrm{obs},2})$ would imply that $\rho(x_{\mathrm{obs},1}, x_{\mathrm{obs},2}) \neq \rho(x_{\mathrm{obs},2}, x_{\mathrm{obs},1})$ as drawn on Fig. 4. In this case, this falls down to $\mathbf{B}(x_1,x_2) \neq \mathbf{B}(x_2,x_1)$ and **B** is not symmetric.

HESSD
7, 9067–9121, 2010

It appears that the asumption of asymmetric correlation functions must go along with the asumption of local correlation lengths so that the relation $l_B(x_1,x_2) = l_B(x_2,x_1)$ remains true and $\mathbf{B}(x_1,x_2) = \mathbf{B}(x_2,x_1)$ remains true. The definition from Eq. (22) is thus only valid if there is only one observation point or if the observation points are separated by sufficiently large distances.

It should be mentioned that in our examples the matrix **B** is not fully formulated. Only the column $\mathbf{B}(x,x_{obs})$ is used, meaning that only $\rho(x,x_{obs})$ and $\sigma_b^2(x_{obs})$ should be explicitely modeled. Since $\mathbf{B}(x_{obs},x)$ is never formulated, along with $\rho(x_{obs},x)$, there is no need to question the symmetry of the B matrix. In the case of the MASCARET application, there are several observation points but they are far enough from each other for us to assume that there is no spatial correlations between two observation points. Each observation point is treated independently of the others.

3.3 Modeling of the background error spatial correlations

The Kalman Filter algorithm is now implemented on the 1-D advection-diffusion model described by Eq. (19), using the same twin-obs framework as previously described. In this context, the background error covariance matrix is updated by the analysis and propagated in time by the tangent linear of the model (Eqs. 10–11). As a consequence, the gain matrix evolves over the assimilation cycle (Eq. 9). A detailed explanation of the evolution of the covariance function at the observation point as well as at an upstream and a downstream location are presented in Appendix along with illustrations.

Initially, the background covariance matrix is modeled by spatially constant variances and correlation length for a gaussian correlation function. After the first assimilation cycle, the error covariances are locally modified. The analysis error covariance matrix is computed from Eq. (10). The analysis error at the observation point is reduced. At upstream and downstream points, the covariance function are asymmetric. The covariance between the observation point and its neighbours is reduced since information at the observation point was brought at this location by the analysis procedure through the innovation vector. The background error covariance matrix for the next assimilation

HESSD
7, 9067–9121, 2010

cycle results from the propagation of the previous cycle analysis error covariance matrix by the tangent linear of the model $\mathbf{M}$ and its adjoint $\mathbf{M}^T$ as formulated in Eq. (11). The initially prescribed symmetric correlation function at the observation point x_2 has been modified into a local and asymmetric function. At the observation point x_2, the correlation function is asymmetric with a shorter correlation length downstream than upstream.

The application of a complete Kalman Filter on this simple advection-diffusion model enabled the understanding of the impact of the analysis and the physics on initial gaussian correlation function. It was explained (see Appendix A) how the information at the observation point leads to the reduction of the uncertainty at the observation point and then downstream this location. It was also explained how the correlation length scale is reduced downstream the observation point and how the initial gaussian correlation function evolves into an asymmetric correlation function. For the MASCARET application, the implementation of a complete Kalman Filter is not possible, mostly because of the high computing cost of the estimation and propagation of the tangent linear model (and it's adjoint). Still, the results presented here on the 1-D advection-diffusion model were used to model the correlation function for water level and discharge in the MASCARET data assimilation procedure. An approximate reduction factor of ten was taken between correlation lengths upstream and downtream the observation points, this division factor showed better results then the factor actually deduced from the Kalman Filter study on the 1-D advection-diffusion model.

In order to finalize to modeling of the background error covariance function, the value of the correlation length $l(x_{\text{obs}})$ must then be estimated. Our objective here is to determine the correlation length of the spatial correlation function for the errors on the water level and the discharge errors with MASCARET. This determination is two-fold. First a diffusion coefficient κ based on the dynamics of the diffusive flood wave approximation model (Eq. 19) is graphically estimated by studying the propagation of a perturbation of the hydraulic state. Then, this diffusion coefficient is used to formulate the spatial correlation length of the state perturbation covariance function. This estimation is used in the following to prescribe the correlation length at each observation point for the data assimilation in MASCARET as described in Eq. (22), defining the spatial correlation function along the 1-D discretized domain. The steps for the estimation of the correlation length $l(x_{\text{obs}})$ are given in Appendix B.

4 Results of the data assimilation on the Adour catchment

4.1 Description of the catchments

The data assimilation study is applied on the Adour maritime catchment area as well as on the Marne Vallage area. A description of theses catchments is given in Sect. 4.1.1 and Sect. 4.1.2, respectively.

4.1.1 Description of the adour catchment

The Adour maritime catchment area is located in Southwestern france, from the Pyrenean Piedmont to the Aquitain coast. This 16 890 km^2 drainage area estuary lies on two departements (Atlanitic Pyrénées and Landes). The Adour river rises in the Pyrenees at an altitude of 2600 m and reaches the Atlantic ocean in Bayonne 312 km further. The Adour catchment is one of the wettest in France due to strong precipitations in the upper part of the basin. The Adour catchment is divided in two regions: the mouth of the river mostly influenced by the tide and the upstream region mostly influenced by the affluents. A schematic description of the Adour catchment is shown on Fig. 7.

The Adour river has three main affluents (responsible for 65% of the total discharge in Bayonne in flood conditions). The Gaves de Pau et d'Oloron, respectivly draining 5226 km^2 and 608 km^2, are often affected by flash floods and gather into the main affluent of the catchment named Gave Réunis. The Nive drains 980 km^2 and joins up with the Adour close to Bayonne. The catchment area is limited by three thresholds controlling the tide waves propagation (at the Gave d'Oloron, Gave de Pau and Nive).

HESSD
7, 9067–9121, 2010

The upstream affluents flow in a mountainous region and their flood plain is narrow. The riverbed becomes larger as the hillslopes decrease. The river banks are partly equiped and the overflowing is stored by embankment dykes allowing the control of the Adour floods.

Meteorological data are provided by Meteo-France. They provide pressure at sea, pressure at land (Biarritz airfield), wind direction and intensity as well as water level anomaly at the coast (these are not used in the MASCARET simulation). The hydrologic (water level and corresponding discharge) data are available in real time at the SPC[2] upstream stations: Dax, Escos, Orthez and Cambo-les-bains and are used as discharge boundary conditions for the hydraulics model. The maritime boundary conditions are given by the SPC tide gauge located in the estuary. Tide forecast are given by the SHOM[3]. Additionally, tide gauges located at Lesseps, Urt and Peyrehorade stations display the water level every five minutes or hourly. These last observations are used for the data assimilation process. The water level at the river mouth varies between 0.9 m and 4.55 m on a semi-diurnal cycle coupled with a bi-weekly signal. Tide waves time propagation ranges between one hour at Lesseps and four hours at Dax for high-tide conditions.

The MASCARET model was chosen by the SCHAPI[4] to simulate the physics of the Adour catchment. A preliminary calibration procedure of several model parameters was done by the SCHAPI using data from twelve flood events of varying intensity. The geometry of the hydraulic network, the computation time step and the Strickler coefficient were ajusted so that the tide and the flood events are well represented at Urt, Dax, Lesseps and Peyrehorade. Globally, at Peyrehorade, the simulation tends to overestimate the flood peak for big flood events and underestimate the flood peak for moderate events.

[2] Service de Prévision des Crues

[3] Service Hydrographique et Océanographique de la Marine

[4] Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations

HESSD
7, 9067–9121, 2010

4.1.2 Description of the marne catchment

The Marne Vallage catchment is located East of the Paris basin. The Marne river is the main affluent of the Seine river and is 525 km long. The Marne catchment area is divided in three main regions: two separated regions upstream the Lac de Der (Marne Amont and Marne Vallage) and one region downstream the Lac de Der (Marne Moyenne). Our study focuses the Marne Vallage drainage area that lies between Condes and Chamouilley.

The Marne river has two main affluents. The Rognon is responsible for 50% of the Marne discharge. This karstic basin is caracterised by slow flood rises, long flooding period and strong sensibility to local precipitations. The Rongeant is a calcareous basin and has a strong reactivity to precipitation. The Marne Vallage catchment is divided in three regions: upstream the Rognon/Marne confluence between Condes and the confluence, upstream the Rognon/Marne confluence between Saucourt and the confluence and downstream the Rognon/Marne confluence between Chamouilley and the Rognon/Marne confluence. A schematic description of the Marne Vallage catchment is shown on Fig. 8.

The hydrologic data are provided by the Champagne-Ardenne DIREN at Condes, Saucourt, Joinville and Chamouilley. The hourly observations at Joinville and Chamouilley are used for the data assimilation procedure. There are two upstream boundary condition (Saucourt and Condes) described by hydrograms and one downstream boundary condition described by a rating curve (correlation between water level and dischage).

As for the Adour catchment, the MASCARET model was chosen by the SCHAPI to simulate the physics of the Marne Vallage catchment. A preliminary calibration procedure of several model parameters was done by the SPC Seine-aval Marne-amont using data from eight flood events of varying intensity. The geometry of the hydraulic network, the computation time step and the Strickler coefficient were ajusted so that the flood events are well represented at Joinville and Chamouilley. The water level simulated at Joinville are globally correct, though often overestimated and the flood peak is often underestimated at Chamouilley.

4.2 Data assimilation set-up

The two-step assimilation procedure described in Sect. 2.3.3 is applied on the Adour catchment and on the Marne Vallage catchment for several flood events. For the Marne catchment, illustrations are shown for an episode in April 2006 (starting 03/23/2006) with a single flood peak. For the Adour catchment, illustrations are shown for an episode in November 2002 (starting 11/02/2002) with a single flood peak as well as for an episode in November 2009 (starting 11/09/2009) with two flood peaks. Statistics are computed over five episodes between 2002 and 2004 for which the observations are hourly (for the 2009 event, the observations are available every five minutes).

For the Marne Vallage catchment, the downstream boundary condition is not considered in the control. Using Eq. (16) and considering the Marne Vallage catchment's two free extremities, the size of the control vector is equal to six. Sensitivy tests at each free extremity for the Marne Vallage catchment revealed that the tangent linear model is valid for a pertubation up to 20% on a, 6 $m^3 s^{-1}$ on b and 6 h on c. For the hydraulic state correction approach, the correlation length were set using the procedure described in Sect. 3, to 51 km and 55 km at Joinville and Chamouilley, respectively.

For the Adour catchment, the oceanic upstream water level is not considered in the control since the uncertainty on the maritime boundary condition is smaller than that of the other free extremities (no use of a rating curve). Using Eq. (16) and considering the Adour catchment's four free extremities, the size of the control vector is equal to twelve. Sensitivity tests at each free extremity for the Adour catchment revealed that the tangent linear model is valid for a pertubation up to 20% on a, 20 $m^3 s^{-1}$ on b and 6 h on c. For the hydraulic state correction approach, the correlation length were set using the procedure described in Sect. 3, to 20 km, 6 km and 34 km at Peyrehorade, Urt and Lesseps, respectively.

HESSD
7, 9067–9121, 2010

Hydraulic observations are gathered in the French *Banque Hydro* data base[5]. River gauges and tide gauges give an approximate and incomplete description of the water level in space and time. The quality of these data greatly relies on the quality of the measurement and for some of them is better in low water conditions then in flood conditions. Additionally, discharge observations are usually derived from water level measurements using a calibration procedure that could be questionned.

In our study, the observed water level reaches up to a couple of meters and the observation error standard deviations are prescribed to 0.1 m. The observation error covariances are neglected assuming that the observations stations are far enough for the spatial errors to be weakly correlated. The background error variances were chosen two to three times bigger than the observation error variances. At each observation point, only the observations above a minimal value are taken into account for the assimilation process to avoid representativeness errors. Additionaly, a threshold is applied on the misfit between the observation and the simulated water level to eliminate inconsistent observations.

4.3 Illustration of the data assimilation method

Figure 9 shows the water level over a four-day period (Day 19 to Day 22 of a flood event in November 2002) at the observation station at Peyrehorade on the Adour catchment. The integration of MASCARET (free run) starting from a previously calculated state is plotted in black and the hourly observations are plotted in blue. The difference between these two curves reaches 15% of the observation at the beginning of Day 22. The assimilation procedure is applied to improve the water level over the first three days (*re-analysis period*) as well as over a so-called *forecast* period (Day 22). The anaysis with the instantaneous correction of the water level (purple curve) shows an excellent fit to the observations over the re-analysis period but to a minor improvement over the forecats period. The model is constantly constrained to the observed state by

[5]http://www.hydro.eaufrance.fr/

HESSD
7, 9067–9121, 2010

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

the hydraulic state correction procedure from Day 19 to Day 22. Even though the analysed state is almost egal to the observed state at the beginning of Day 22, over the forecast period, the analysis remains far from the observation. This shows that the improvement of the initial conditon at Day 22 is not enough to improve the simulation over the following day. The improvement is only significant over a couple of hours and other uncertainties degrade the simulation. Effectively, the upstream flow is known up to Day 22 and is kept constant over the forecast period. The analysis after the correction of the upstream flow (green curve) shows a good fit to the observation over the past period as well as over the forecast period. The difference to the observation only reaches 9% of the observed value at Day 22.5. The upstream flow is corrected over the two-day period (Day 20 to Day 22) allowing for a better simulation of the water level over this period. Additionally, over the forecast period, the upstream flow is kept constant to the last analysed value (which is obviously better than the non-analysed one) allowing for a significant improvement of the water level over Day 22. The analysis after the two-step assimilation procedure is plotted in red and shows an improvement over the re-analysis as well as over the forecast period.

For this event, the two-step analysis is cycled every hour so that, at an observation point, the water level is forecasted over the whole flood event. Figure 10 shows the six-hour forecast for the free (black curve) and the assimilated runs (red curve) as well as the non assimilated observations (blue curve). It clearly appears that the assimilation improves the six-hour forecast. Shorter and longer ranges for the forecast were also studied leading to similar conclusions.

Similar representation is provided on Fig. 11 for the November 2009 event on the Adour catchment. Observations at the previously described stations are available every five minutes. The non assimilated observation (forecast mode) are represented by the blue curve. The two-step analysis was applied in near real time mode and led to a significant improvement of the flood forecast. At a one hour forecast range, the free run (black curve) underestimates the water level at Peyrehorade, the simulated first peak is underestimated of approximately 10%. The second flood peak is correctly represented eventhough the flood rise is too slow. The two-step data assimilation (red curve) enables the exact representation of the both flood peaks as well as a more realistic flood rise for the second peak. Same results were observed at the other observation stations.

Finally, the two-hour and the seven-hour forecasted water level on the Marne Vallage catchment, for the April 2006 event is presented on Figs. 12 and 13, respectively. The free run simulation significantly overestimates the flood peak at Joinville and Chamouilley (not shown) and the two-step data assimilation procedure allows for a good simulation of the peak at a two-hour forecast. The simulated peak at seven-hour forecast is also in better agreement with the observations than the free run, even though an overestimation of the peak remains. More analysis were carried out for flood events on the Marne Vallage catchment. Globally, the results are not as satysfing as on the Adour catchment. The main reason for this moderated performance is an incomplete and perfectible calibration of the numerical model (mostly Strickler coefficient and lateral input flow) on the Marne Vallage catchment before the assimilation procedure. It was shown on some events, that in order to improve the simulation at one observation station, the data assimilation algorithm must degradate the simulation at the other observation location. The application of the two-step data assimilation procedure enabled the detection of a model incoherence on the Marne Vallage that can not be efficiently accounted for with the present control vector. Further work towards the improvement of the calibration of the model for the Marne Vallage catchment is ongoing at the SPC Seine-aval Marne-amont and will be used with the data assimilation procedure when available.

4.4 Interpretation of the results

4.4.1 Criteria for the interpretation

The results of the assimilation approach are statistically interpreted for five flood events between 2002 and 2004 on the Adour catchment. Various criteria can be considered to estimate the impact of the analysis on the water level simulation at each observation point.

We present here the computation of the standard deviation of the difference between the Free run and the Observation at an observation point for the water level, denoted by FmO (Free run Minus Observation, in m), and the difference between the Analysis and the Observation, denoted by AmO (Analysis Minus Observation, in m). These differences are computed for a given forecast range (six-hour in our case) at each observation time (one hour or five minutes). The standard deviation is computed over time for each flood event as well as over the simulated flood events. The precision criteria Pr (in %) is also defined as

$$Pr = \frac{100}{N_{\text{obs}}} \sum_{1}^{N_{\text{obs}}} \left| \frac{h^{\text{mod}} - h^{\text{obs}}}{h^{\text{obs}}} \right|, \tag{23}$$

where N_{obs} is the number of observations over a period (24 h in re-analysis mode and six hours in forecast mode in our case) and h^{mod} is the simulated water level with or without the assimilation. When computed with the free run simulation, the precision is noted Pr^{frun} whereas when computed with the analysed water level, the precision is noted Pr^{assim}.

It must be noted that the assimilation procedure is not the same for every flood event. In some cases, because of the model instabilitiy (more precisely of its use on the Adour catchment), it was not possible to achieve the free run with MASCARET over a two-day period used as the background for the data assimilation procedure, it was then impossible to control the upstream flow. Only the instantaneous approach was possible as it constantly constrains the simulation. Whenever possible the two-step assimilation was implemented, if not, only the instantaneous hydraulic state correction was applied.

HESSD
7, 9067–9121, 2010

4.4.2 November 2002 event, adour catchment

The November 2002 flood event is first analysed in details here, then, the results are presented over the other five events.

For the 2002 event, it was first noted that the root mean square of AmO is smaller than the root mean square of FmO at any time during the re-analysis period, meaning that the assimilation procedure works correctly and brings the analysis closer to the observations than the background for each flood event. This result is confirmed by the computation of the precision over the 24 h before the last observation time over the whole flood event. The computation of this criteria is shown on Fig. 14 at Peyrehorade. For that particular case, the precision computed with the analysed simulated water level is close to 0 when the precision computed with the free run simulated water level is around 5%.

Additionally, at a six-hour forecast range, we also have rms(Am0) < rms(FmO) meaning that the correction of the initial condition and the boundary conditions (upstream flow) for the forecast leads to a significant improvement of the water level forecast. The computation of the precision over the six hour after the last observation time over the whole flood event is also presented on Fig. 14 at Peyrehorade. For that particular case, the precision computed with the analysed simulated water level is close to 2.5% when the precision computed with the free run simulated water level is around 7.5%.

4.4.3 Statistical interpretation of five events

The past (−24 h) and forecast (+6 h) precisions were computed over time for each flood event, at each observation point. Their temporal means (at Peyrehorade) are presented in Table 1.

As expected, for the free run, the mean forecast precision Pr_{+6}^{frun} is larger than the mean past precision Pr_{-24}^{frun} as the upstream flow is unknow in forecast mode (as well as other boundary conditions such as the maritim water level forcing that is forecasted by tide models at the SHOM), at Peyrehorade and also at the two other observation points (not shown).

HESSD
7, 9067–9121, 2010

With the assimilation procedure, at Peyrehorade, the mean past precision is reduced from 4.15% to 0.36% (reduction of 91%) meaning that the analysis is closer to the observation than the free run on the re-analysis period, the water level hydraulic state correction procedure plays a major role here as shown on Fig. 9 as it constantly brings the simulation towards the observation. Additionally, the mean forecast precision is reduced from 4.33% to 2.76% (reduction of 36%), the upstream flow correction is mostly responsible for this improvement (when activated) as it limits the error on the upstream forcing. Even with the assimilation procedure, the forecast precision Pr_{+6}^{assim} remains larger than the past precision Pr_{-24}^{assim} because of the uncertainties on the boundary conditions in forecast mode.

These results are similar for the observation stations at Urt and Lesseps as well as globally over the 3 stations showing that the overall effect of the assimilation on the Adour catchment improves the description of the water level. Still the Lesseps station being closer to the maritime boundary, it is significantly influenced by the tides that are not controled, as a consequence, the results of the assimilation is not as good as for the Peyrehorade and Urt stations.

The difference between the free run and the observation FmO and between the analysis and the observation AmO were computed over time respectively in re-analysis mode (−24 h) and forecast mode (+6 h) for each flood event, at each observation point. Their temporal mean and standard deviation (at Peyrehorade for the first event $A1$) are presented in Table 2.

In re-analysis mode, $\overline{\text{AmO}}_{-24}$ is significantly smaller than $\overline{\text{FmO}}_{-24}$, for event $A1$ as well as for the other four events (not shown). Additionally, as expected with the data algorithm, the standard deviation of AmO is smaller than the standard deviation of FmO meaning that the assimilation reduced the variance of the error to the observation. Same conclusions are drawn for the forecast mode, with a smaller improvement than in re-analysis mode.

HESSD
7, 9067–9121, 2010

Averaged over the flood events and the observation stations, the mean standard deviation of the difference to the observations is reduced from 0.1124 to 0.024 in re-analysis mode and from 0.16 to 0.15 in forecast mode.

5 Summary and conclusions

This paper presented the improvement of river floods forecast assimilating river water level observations. The study was carried out with the 1-D dimensional hydraulics model MASCARET on the Adour catchment and on the Marne Vallage catchment. Representative events were presented for both catchments and statistics on the results were computed over five events between 2002 and 2004 for the Adour catchment. The water level data were assimilated using a BLUE algorithm to control the upstream flow and dynamically correct the hydraulic state. Particular focus was made on the modeling of the background error covariance matrix. At first, the correlation was assumed to be gaussian and its correlation length was estimated with a graphical method based on the propagation of a perturbation from the upstream station to the observation points. Asymmetric functions were used to represent the background error spatial correlation for the water level and the discharge, respectively. The justification for this choice has been made applying a Kalman Filter algorithm on a simple advection-diffusion model. It was shown that the analysis turns a gaussian correlation function into an asymmetric correlation function where the correlation length scale is shorter downstream the observation point. This approach enabled a realistic modeling of the spatial error correlations.

The first step of the analysis was based on the assumption that the upstream flow can be adjusted using a three-parameter simple correction. These three control parameters were adjusted on a two-day time window after one day of free run. Sequentially, the second step of the assimilation consisted in correcting the hydraulic state every hour (the observation frequency) during one day. The simulation was integrated in forecast mode for one more day. This procedure was applied on a four-day sliding window over the whole flood event. It was shown that the simulation is significantly closer to the observation than the free run over the re-analysis period as well as over the forecast period. Two criteria were used to draw these conclusions: the precision and the root mean square of the difference between the simulation and the observations. For some events, the free run was not completed because of numerical reasons and only the hydraulic state correction procedure was applied. In these case, the improvement of the assimilation is less impressive in forecast mode. In fact, most of the errors on the hydraulic state are due to uncertainties on the upstream flow. The sensitivity to an initial condition for the continuation of a simulation can be totally negligible compared to the sensitivity to the upstream flow. In addition, since the upstream flow in forecast mode are kept constant to the last known value, the control of the upstream flow is predictive. On average, the correction of the hydraulic state is not as predictive and does not suffice to constrain the simulation over an interesting forecast period.

The assimilation procedure presented in this paper can be applied to other catchment areas. Still, a careful estimation of the data assimilation statistics must be carried out as they are representative of the local physics. The relation between the water level and the discharge errors should be further investigated. This aspect would be enriched by the use of local (Z,Q) calibration functions. Beyond this study, the extension of the control space can be foreseen as other sources of uncertainties or model errors (for example the simplification of the flood plain representation) result on errors on the hydraulic state. Effectively, the correction of the Strickler coefficients or of the lateral discharge could be further investigated. So far, the non linearity of the observation operator was such that the use of the BLUE algorithm did not seem promising. A closer study of the sensitivity these parameters might lead to different interpretation or, on the contrary, promote the use of other data assimilation algorithms.

HESSD
7, 9067–9121, 2010

Appendix A

Evolution of the local covariance functions with the Kalman Filter algorithm on a 1-D advection-diffusion model

The Kalman Filter algorithm is implemented on the 1-D advection-diffusion model and the covariance matrices are updated following Eqs. (11–10).

The initial background covariance matrix[6] $\mathbf{B}_{c1}$ is modeled by spatially constant variances and correlation length for a gaussian correlation function. For this analysis, $\sigma_b^2(x) = 0.25$ for all x. The covariances $\mathbf{B}_{c1}(x_1,x)$, $\mathbf{B}_{c1}(x_2,x)$ and $\mathbf{B}_{c1}(x_3,x)$ between respectively x_1,x_2,x_3 and any point x are displayed on Fig. 5a (x_2 is the observation point in our example). It should be noted that $\mathbf{B}_{c1}(x_i,x_i) = \sigma_b^2(x_i) = 0.25$ for $i \in \{1,2,3\}$. By construction, $\mathbf{B}_{c1}$ is symmetric.

After one assimilation cycle, the error covariances are locally modified, the analysis error covariances in $\mathbf{A}_{c1}$ are computed from Eq. (10) and shown on Fig. 5b for x_1, x_2 and x_3. It should be noted that, as expected, the analysis error variance σ_a^2 at the observation point x_2 is smaller than σ_b^2. At the observation point, the covariance function $\mathbf{A}_{c1}(x_2,x)$ remains symmetric. On the contrary, at the upstream point x_1 and at the downstream point x_3, the covariance function $\mathbf{A}_{c1}(x_1,x)$ and $\mathbf{A}_{c1}(x_3,x)$ are asymmetric. The covariance between x_1 and the observation point x_2, as well as between x_3 and the observation point x_2, is reduced since information at the observation point was brought at this location by the analysis procedure through the innovation vector. The covariance functions at x_1 and x_3 are symmetric around x_2. It should also be noted that the analysis covariance matrix modeled with the represented covariance function is symmetric, for example, $\mathbf{A}_{c1}(x_1,x_2) = \mathbf{A}_{c1}(x_2,x_1) = 0.0057$, these two values are represented by dots on Fig. 5b.

[6] The subscript ci denotes the number of the assimilation cycle

HESSD
7, 9067–9121, 2010

The background error covariance matrix $\mathbf{B}_{c2} = \mathbf{M}^T \mathbf{A}_{c1} \mathbf{M}$ for the next (second so far) assimilation cycle is computed from Eq. (11) meaning that the previous cycle analysis background covariance matrix is propagated by the tangent linear of the model $\mathbf{M}$ and its adjoint $\mathbf{M}^T$.

The columns of the updated $\mathbf{B}_{c2}$ for x_1, x_2 and x_3 are shown on Fig. 5c. The asymmetric covariances and correlation functions at all the upstream and downstream locations where propagated to the observation points so that, the covariance and then correlation functions are now asymmetric at the observation point. The spatial covariances in $\mathbf{B}_{c2}$ for x_{obs}, here $\mathbf{B}_{c2}(x_2, x)$, for this second assimilation are fundamentaly different from those initially prescribed for the first assimilation in $\mathbf{B}_{c1}$, here $\mathbf{B}_{c1}(x_2, x)$. Still, the symmetry property of the covariance matrix is conserved and $\mathbf{B}_{c2}(x_1, x_2) = \mathbf{B}_{c2}(x_2, x_1)$ (these values are represented by dots on Fig. 5c). The initially prescribed symmetric correlation function at the observation point x_2 has been modified into a local and asymmetric function. At the observation point x_2, the correlation function is asymmetric with a shorter correlation length downstream than upstream.

Figure 6a represents the diagonal terms of the first analysis covariance matrix $\sigma_a^2(x) = \mathbf{A}_{c1}(x, x)$ (in red dashed curves) and the second background covariance matrix $\sigma_{\mathrm{b}}^2(x) = \mathbf{B}_{c2}(x, x)$ (in black dashed curves). As expected, the analysis error variance is smaller than the initial background error variance ($\mathbf{B}_{c1}(x, x) = 0.25$) in the neighbouring of the observation point. The size of this neighbouring is prescribed by the correlation length initially prescribed in $\mathbf{B}_{c1}$. The variances initially prescribed to the spatially constant value 0.25 are now local, for example $\sigma_{\mathrm{b}}^2(x_2) = 0.0497$ and $\sigma_{\mathrm{b}}^2(x_1) = 0.2474$. Additionally, the variances in the updated $\mathbf{B}_{c2}$ (in black dashed curves) correspond to the propagation of the variances in $\mathbf{A}_{c1}$ by $\mathbf{M}$ and its adjoint $\mathbf{M}^T$, they are also local. The update of the background error covariances by the analysis and the propagation of the background error covariances matrix by the tangent linear model consists in the evolution of both the variances and the correlations. It appears that the correlation lengths tend to shorten downstream the observation point. After several iterations (7 in our example) of the Kalman Filter, the variances are globally reduced downstream the observation point as shown on Fig. 6a for $\mathbf{A}_{c7}$ and $\mathbf{B}_{c8}$. Effectively, the uncertainty at the observation is reduced by the data assimilation algorithm and then the information is propagated downstream.

The covariances $\mathbf{A}_{c7}(x_1,x)$, $\mathbf{A}_{c7}(x_2,x)$ and $\mathbf{A}_{c7}(x_3,x)$ between respectively x_1,x_2,x_3, after seven iterations of the Kalman filter, are shown on Fig. 6b in solid curves. For comparison, the covariances from cycle 1 were also plotted in dashed curves on Fig. 6b. It is worth noting that at $x_3 = 320$, the amplitude of the variance has approximately been divided by two over the seven assimilation cycles. The shape of the covariance function evolves over time, especially downstream the observation point and the correlation functions are clearly asymmetric with a shorter correlation length scale downstream than upstream the observation point. The local correlation function for $x_2 = x_{obs}$ in $\mathbf{B}_{c1}$ and $\mathbf{B}_{c8}$, respectively denoted by $\rho_{c1}(x_2,x)$ and $\rho_{c8}(x_2,x)$, are shown on Fig. 6c. The correlation length is approximatly divised by five on this plot along the 7 assimilation cycle of the Kalman Filter. This factor doesn't vary significantly when the Kalman Filter is further iterated.

Appendix B

Estimation of the local correlation length in B for the hydraulic state correction procedure

Our objective here is to determine the correlation length of the spatial correlation function for the errors on the water level and the discharge errors with MASCARET. This determination is two-fold. First a diffusion coefficient κ based on the dynamics of the diffusive flood wave approximation model (Eq. 19) is graphically estimated by studying the propagation of a perturbation of the hydraulic state. Then, this diffusion coefficient is used to formulate the spatial correlation length of the state perturbation covariance function.

HESSD
7, 9067–9121, 2010

The diffusion coefficient κ is estimated by simulating the response to an upstream perturbation along the water line. For a stationnary discharge and water level, for each reach, a small but steep perturbation is added to the up-stream flow. This perturbation is propagated and diffused over time to reach the observation points. A perturbation of the form

$$\tilde{h}(x,0) = \frac{1}{2}\mathrm{erf}\left(\frac{x}{\sqrt{2}\ l_{\mathrm{temp},0}}\right) + \frac{1}{2}, \tag{B1}$$

that sums up to a Heaviside function if $l_0 \rightarrow 0$, was added at the upstream flows $\tilde{q}(t)$. This perturbation is propagated by Eq. (19) towards the observation points where, at time t, the state is described by $\tilde{h}(x,t)$ given by:

$$\tilde{h}(x,t) = \frac{1}{2}\mathrm{erf}\left[\frac{x-ct}{\sqrt{2}\ l_{\mathrm{temp}}(t)}\right] + \frac{1}{2} \tag{B2}$$

with $l_{\mathrm{temp}}(t)^2 = l_{\mathrm{temp},0}^2 + 2\kappa t$.

The parameters κ and c are estimated from the numerical solution $\tilde{h}(x,t)$ of MASCARET as follow:

- $c \approx \frac{s_\mathrm{r}}{t_\mathrm{r}}$ where s_r is the curvilinear distance between the upstream station and the observation point and t_r is the time between the upstream perturbation and the arrival of the step perturbation at the observation point,

- $l_{\mathrm{temp}}(t_\mathrm{r}) \approx cT$ where T is the time between the +20% of the initial discharge and −20% of the final discharge. T is graphically estimated on the simulated discharge at the observation points,

- $\kappa = \frac{\sqrt{l_{\mathrm{temp}}(t_\mathrm{r})^2 - l_{\mathrm{temp},0}^2}}{2t_\mathrm{r}}$ with $l_{\mathrm{temp},0}^2 = 0$ for the Heaviside initial condition which is approximatively the case.

HESSD
7, 9067–9121, 2010

Using these three relations when the pertubation reaches the observation point, κ is estimated by $\frac{T^2 s_r^2}{2t_r^3}$ where s_r is known where as T and t_r are graphically estimated on the simulated discharges.

Like the initial and boundary conditions, the temporal covariance function $R(\tau)$ is also propagated by the diffusive flood wave propagation equations Eq. (19). Since the temporal covariance function $R(\tau)$ is a gaussian and using the theory of the random function diffusion, the spatial covariance function of $\widetilde{h}(x,t)$ can be approximated by a gaussian. Assuming that the spatial covariance function for the boundary condition is chosen as a gaussian and denoted by $\mathbf{B}_0$, the covariance function for the solution at time t is given by $\mathbf{B}_t = \mathbf{M}_t \mathbf{B}_0 \mathbf{M}_t^T$ where $\mathbf{M}_t$ stands for the advection-diffusion processus. Since the advection processus has no effect on the covariance function, this formulation can be written in the advected referential. In this referential, denoting by $\mathbf{L}_t$ the diffusion of a gaussian solution, then $\widetilde{\mathbf{B}_t} = \mathbf{L}_t \mathbf{L}_t^T = \mathbf{L}_{2t}$ and the spatial correlation length in $\widetilde{\mathbf{B}}_t$ decreases with the distance:

$$l^2(x) = l_0^2 + 4\kappa \frac{x}{c}. \tag{B3}$$

The correlation length $l(x)$ is locally defined for any location of the domain. For the BLUE assimilation algorithm, only the correlation length at the observation point x_{obs} is needed. The local correlation length at the observation point is then calculated using Eq. (B3). For a usual application case with MASCARET, a realistic upstream flow is prescribed from which l_0^2 can be determined. Still, when the observation point is far enough from the upstream stations $l_0^2 \ll 4\kappa \frac{x_{obs}}{c}$.

This graphical approach leads to the estimation of a local correlation length at each observation point based on the perturbation of the upstream flow at one upstream station. Since there are several upstream stations for our studies, there are several resulting signals $\widetilde{h}(x,t)$ reaching the observation point, leading to several estimations of $l(x_{obs})$. At the observation point, the spatial correlation function is approximated by a gaussian resulting from the sum of the gaussians of respective correlation length l_i and

HESSD
7, 9067–9121, 2010

amplitude a_i. The correlation length of the resulting gaussian can be approximated by:

$$l(x_{\text{obs}}) = \frac{\sum_i a_i l_i(x_{\text{obs}})}{\sum_i a_i} \tag{B4}$$

where the subscript i denotes the number of the upstream station ($i \in [1,4]$ for the Adour catchment and $i \in [1,4]$ for the Marne Vallage catchment).

Acknowledgements. The authors are grateful to N. Goutal at LNHE Department at EDF who has made the source code of the hydraulics model MASCARET available for this work. The authors would like to thank C. Ivanoff who contributed to this work during her Master intership at CERFACS and to O. Pannekoucke for valuable discussions.

The publication of this article is financed by CNRS-INSU.

References

Atanov, G., Evseeva, E., and Meselhe, E.: Estimation of roughness profile in trapezoidal open channels, Journal of Hydraulic Engineering, 125(3), 309–311, 1999. 9069

Beven, K. J. and Freer, F.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, 2001. 9069

Bessières, H., Roux, H., and Dartus, D.: Data assimilation and distributed flash flood modeling, in: Proceedings of the second Space for Hydrology Workshop, Surface Water Storage and Runoff: Modeling, In-situ data and Remote Sensing, 2007. 9069

Bouttier, F. and Courtier, P.: Data assimilation concepts and methods, ECMWF Lecture Note, 1999. 9069, 9072
Brasseur, P. and Verron. J.: The SEEK filter method for data assimilation in oceanography: a synthesis, Springer Berlin/Heidelberg, 2006. 9069
Buis, S., Piacentini, A., and Déclat, D.: PALM: A Computational framework for assembling high performance computing applications, Concurrency Computat.: Pract. Exper., 18(2), 247–262, 2006. 9080
Chow, V. T.: Open-channel hydraulics, McGraw-Hill, 1959. 9068
Das, A.: Parameter estimation for flow in open-channel networks, Journal of Irrigation and Drainage Engineering, 130(3), 160–165, 2004. 9069
Fouilloux, A. and Piacentini, A. :The PALM Project: MPMD Paradigm for an Oceanic Data Assimilation Software, Lecture Notes In Computer Science, 1685, 1423–1430, 1999. 9080
Gelb, A.: Applied optimal estimation, Cambridge Mass.: MIT Press, 1974. 9071
Goutal, N. and Maurel, F.: A finite volume solver for 1D shallow water equations applied to an actual river, Int. J. Numer. Meth. Fluids, 38(2), 1–19, 2002. 9071
Hervouet, J.-M.: Hydrodynamique des écoulements à surface libre – Modélisation numérique avec la méthode des éléments finis, Ponts et Chaussées (Presses), 2003. 9068
Honnorat, M., Marin, J., Monnier, J., and Lai, X.: Dassflow v1.0: a variational data assimilation software for 2D river flows., INRIA Technical Report, 2007. 9069
Ide, K., Courtier, P., Ghil, P., and Lorenc, A. C.: Unified notation for data assimilation: operational, sequential and variational, Journal of the Meteorological Society of Japan, 75(1B), 181–189, 1997. 9069
Jean-Baptiste, N., Doré, C., Sau, J., and Malaterre, P.-O.: Data assimilation forthe real-time update of a 1D hydrodynamic lodel, fault detection – Application to the automatic control of hydropower plants on the Rhône river, in Proceedings of SimHydro: Hydraulic modeling and uncertainty, 2010. 9069
Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003. 9069
Lagarde, T.: Nouvelle approche des méthodes d'assimilation de données: les algorithms de point selle, Ph.D. thesis, Université Paul Sabatier, 246 pp., 2000. 9080
Lagarde, T., Piacentini, A., and Thual, O.: A new representation of data assimilation methods: the PALM flow charting approach, Q. J. R. M. S., 127, 189–207, 2001. 9080

Maclaughlin, D. and Townley, L.: A reassessment of the groundwater inverse problem, Water Resour. Res., 32, 1131–1161, 1996. 9069
Malaterre, P.-O., Baume, J.-P., and Jean-Baptiste, N.: Calibration of open channel flow models: a system analysis and control engineering approach, in Proceedings of SimHydro: Hydraulic modeling and uncertainty, 2010. 9069
Pappenberger, F., Beven, K., Horrit, M., and Blazkov, S.: Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations, Journal of Hydrology, 302, 46–69, 2005. 9069
Parrish, D. F. and Derber, J. C.: The national meteorological center's spectral statistical interpolation analysis system, Mon. Weather Rev., 120, 1747–1763, 1992. 9069
Rabier, F., Jarvinen, H., Kilnder, E., Mahfouf, J.-F., and Simmons, A.: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics, Quarterly Journal of The Royal Meteorological Society, 126, 1143–1170, 2000. 9069
Sau, J., Malaterre, P.-O., and Baume, J.-P.: Sequential MonteCarlo hydraulic state estimation of an irrigation canal, Comptes Rendus de l'académie des Sciences, 338, 212–219, 2010. 9069
Talagrand, 0.: Assimilation of observations, an introduction, J. Meteorol. Soc. Jpn., 75(1B), 191–209, 1997. 9071
Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM: Society for Industrial and Applied Mathematics, 1987. 9069
Todling, R. and Cohn, S. E.: Suboptimal schemes for atmospheric data assimilation based on the Kalman filter, Mon. Weather Rev., 122, 2530–2557, 1994. 9073
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, Hydrol. Earth Syst. Sci., 14, 1623–1637, doi:10.5194/hess-14-1623-2010, 2010. 9069, 9078
Thirel, G., Martin, E., Mahfouf, J.-F., Massart, S., Ricci, S., Regimbeau, F., and Habets, F.: A past discharge assimilation system for ensemble streamflow forecasts over France – Part 2: Impact on the ensemble streamflow forecasts, Hydrol. Earth Syst. Sci., 14, 1639–1653, doi:10.5194/hess-14-1639-2010, 2010. 9069

HESSD
7, 9067–9121, 2010
Data assimilation for flood forecasting
S. Ricci et al.

Table 1. Precision for the free run Pr^{frun} and the analysis Pr^{assim} computed over 24 h before the last observation time and over six hours after the last observation time, averaged over five flood events between 2002 and 2004, at Peyrehorade, Urt and Lesseps.

Flood event	Correction	Obs. station	Pr^{frun}_{-24}(%)	Pr^{assim}_{-24}(%)	Pr^{frun}_{+6}(%)	Pr^{assim}_{+6}(%)
*A*1	Upflow,h,Q	Peyrehorade	4.86	0.39	4.98	2.21
*A*2	h,Q	Peyrehorade	4.17	0.22	5.10	3.51
*A*3	h,Q	Peyrehorade	4.84	0.31	5.44	3.91
*A*4	h,Q	Peyrehorade	4.40	0.31	3.53	2.28
*A*5	Upflow,h,Q	Peyrehorade	2.49	0.60	2.60	1.93
Mean		Peyrehorade	4.15	0.36	4.33	2.76
Mean		Urt	3.63	0.69	3.92	3.08
Mean		Lesseps	2.71	0.97	2.69	2.63
Mean		3 stations	3.49	0.67	3.64	2.82

HESSD
7, 9067–9121, 2010

Table 2. Mean and rms of FmO(Free run Minus Observation) and AmO (Analysis Minus Observation) computed over 24 h before the last observation time and over six hours after the last observation time for the November 2002 event at Peyrehorade.

$\overline{FmO}_{-24}$	$\overline{AmO}_{-24}$	$Std(FmO)_{-24}$	$Std(AmO)_{-24}$
−0.19	0.004	0.17	0.024
$\overline{FmO}_{+6}$	$\overline{AmO}_{+6}$	$Std(FmO)_{+6}$	$Std(AmO)_{+6}$
−0.20	0.008	0.20	0.18

HESSD
7, 9067–9121, 2010

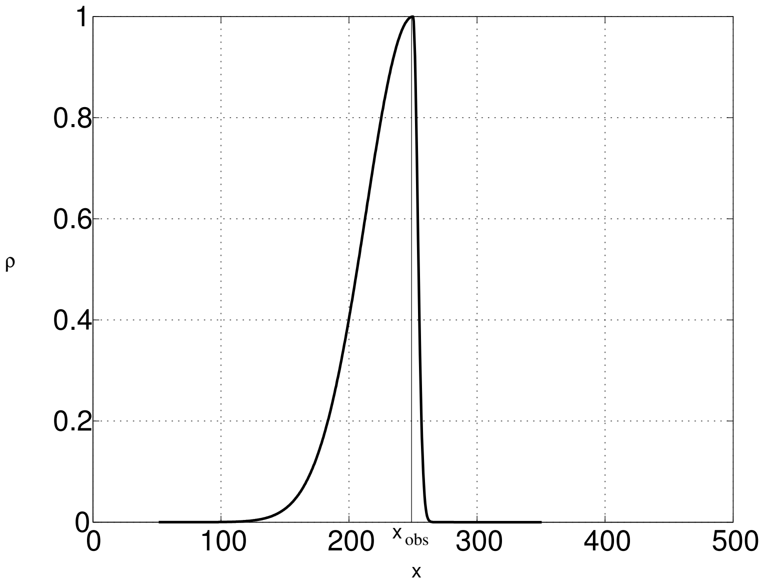

Fig. 1. Local asymmetric correlation function at observation point $x_{obs} = 250$.

HESSD
7, 9067–9121, 2010

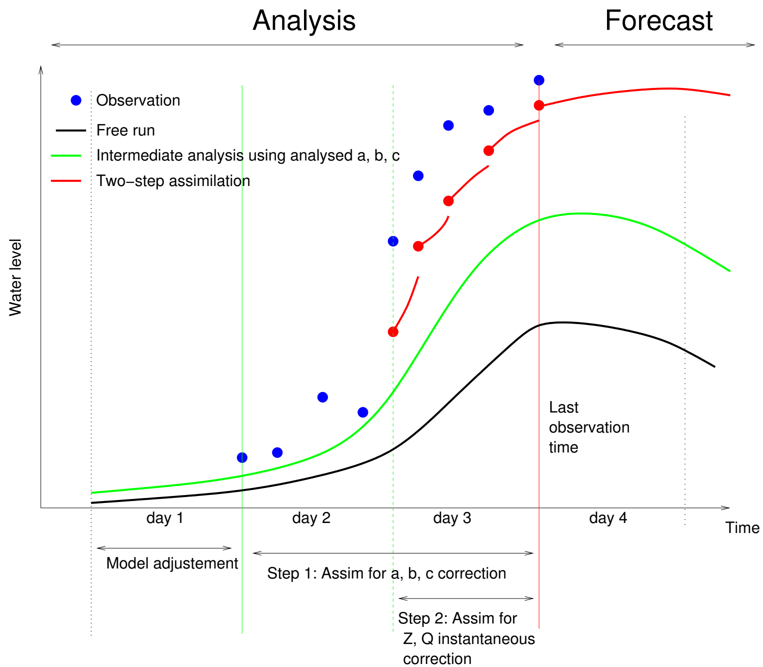

Fig. 2. Schematic representation of the two-step assimilation algorithm over one cycle.

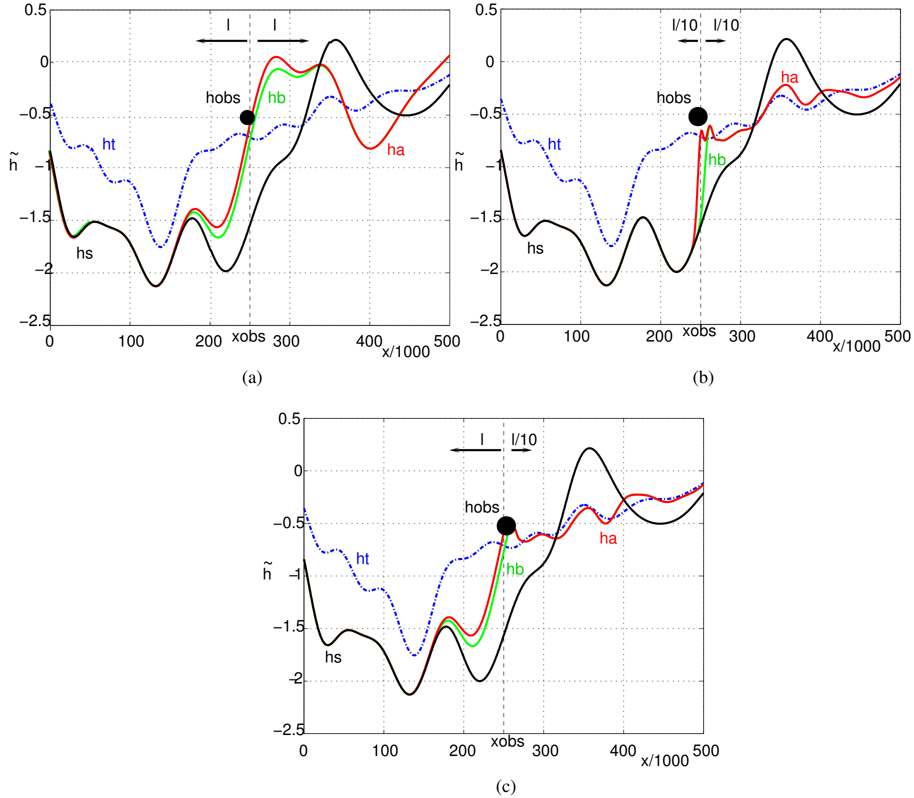

Fig. 3. Water level perturbation $\tilde{h}$ of the advection-diffusion model for the *true* simulation (dotted blue line), the *free* simulation (black solid line), the *background* simulation (green solid line) and the *analysed* simulation (red solid line) with **(a)** $l_- = l_+ = l(x_{obs})$, **(b)** $l_- = l_+ = l(x_{obs})/10$, **(c)** $l_- = l(x_{obs})$ et $l_+ = l(x_{obs})/10$.

HESSD
7, 9067–9121, 2010

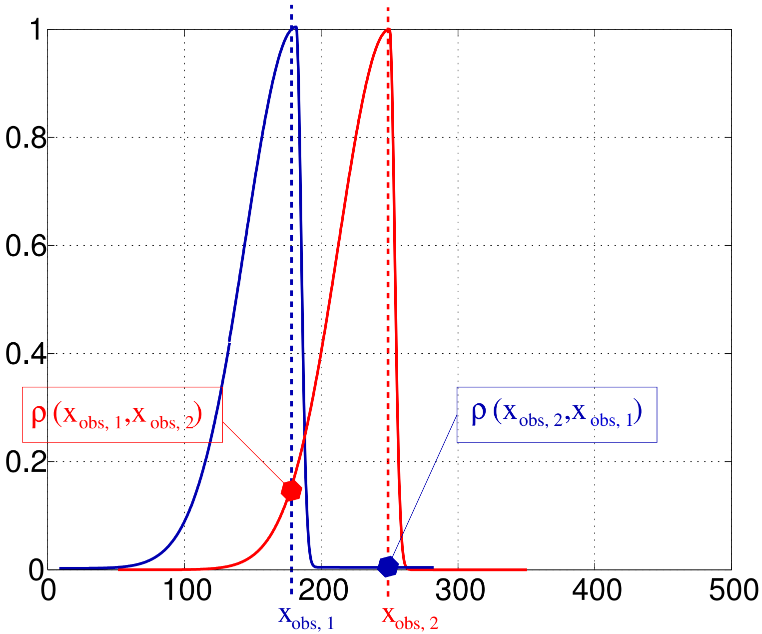

Fig. 4. Spatial correlations $\rho(x,x')$ defined by Eq. (21) at observation points $x_{\mathrm{obs},1}$ and $x_{\mathrm{obs},2}$ which could be erroneously prescribed with asymmetric correlation functions with $l_{\mathrm{B}}(x,x_{\mathrm{obs},1})$ and $l_{\mathrm{B}}(x,x_{\mathrm{obs},2})$ defined by Eq. (22) and $l(x_{\mathrm{obs},1}) = l(x_{\mathrm{obs},2})$.

HESSD
7, 9067–9121, 2010

Data assimilation for flood forecasting

S. Ricci et al.

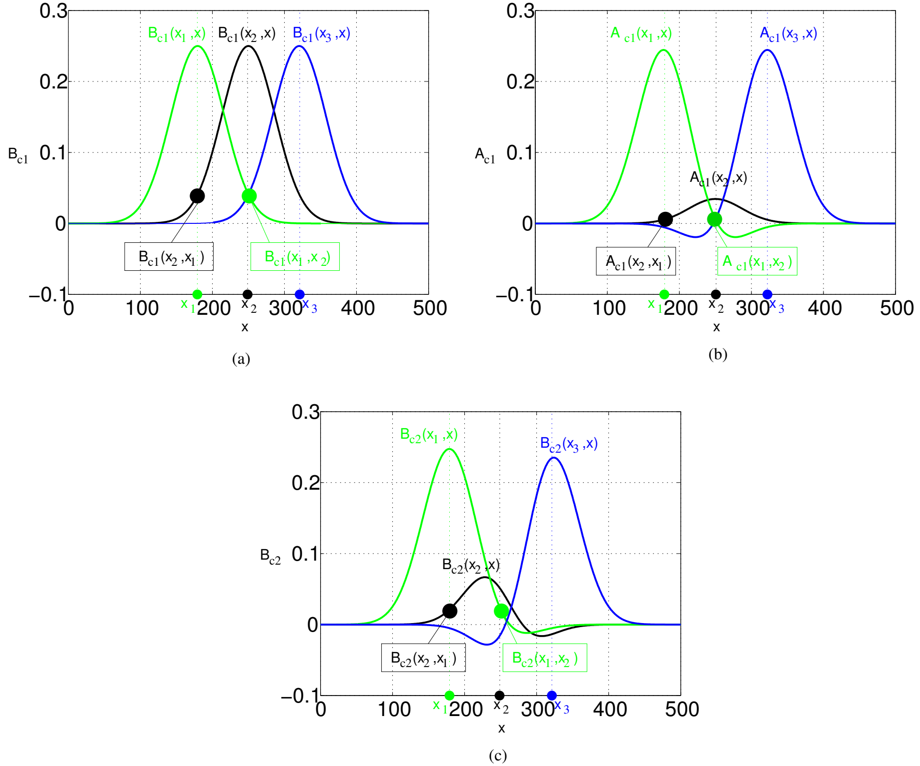

Fig. 5. **(a)** Initial background error covariances in $\mathbf{B}_{c1}$, **(b)** analysis error covariances in $\mathbf{A}_{c1}$ after one analysis and **(c)** updated covariances in $\mathbf{B}_{c2}$ for points $x_1 = 180$ (green), $x_2 = x_{obs} = 250$ (black) and $x_3 = 320$ (blue).

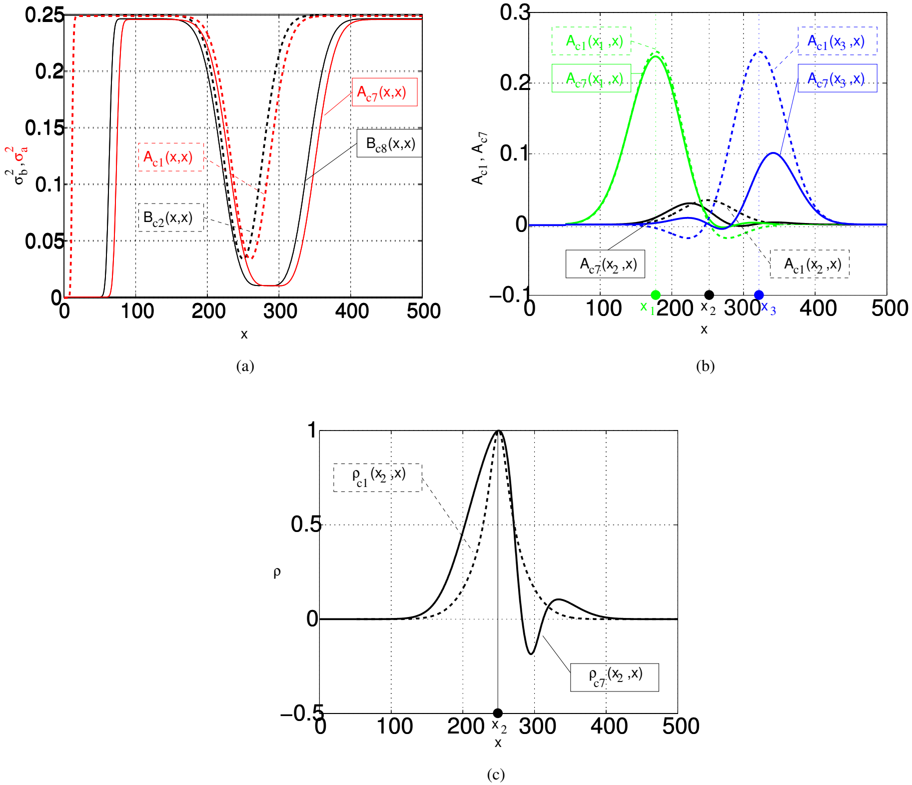

Fig. 6. **(a)** Error variance at each point for cycle one (dashed curves): analysis $\mathbf{A}_{c1}$ (red) and $\mathbf{B}_{c2}$ (black), and for cycle seven (solid curves): analysis $\mathbf{A}_{c7}$ (red) and $\mathbf{B}_{c8}$ (black). **(b)** Analysis error covariances in $\mathbf{A}_{c1}$ (dashed curves)and in $\mathbf{A}_{c7}$ (solid curves) for points $x_1 = 180$ (green), $x_2 = x_{obs} = 250$ (black) and $x_3 = 320$ (blue). **(c)** Error correlation in $\mathbf{B}_{c1}$ (dashed curve) and $\mathbf{B}_{c8}$ (solid curve) for $x_2 = x_{obs}$.

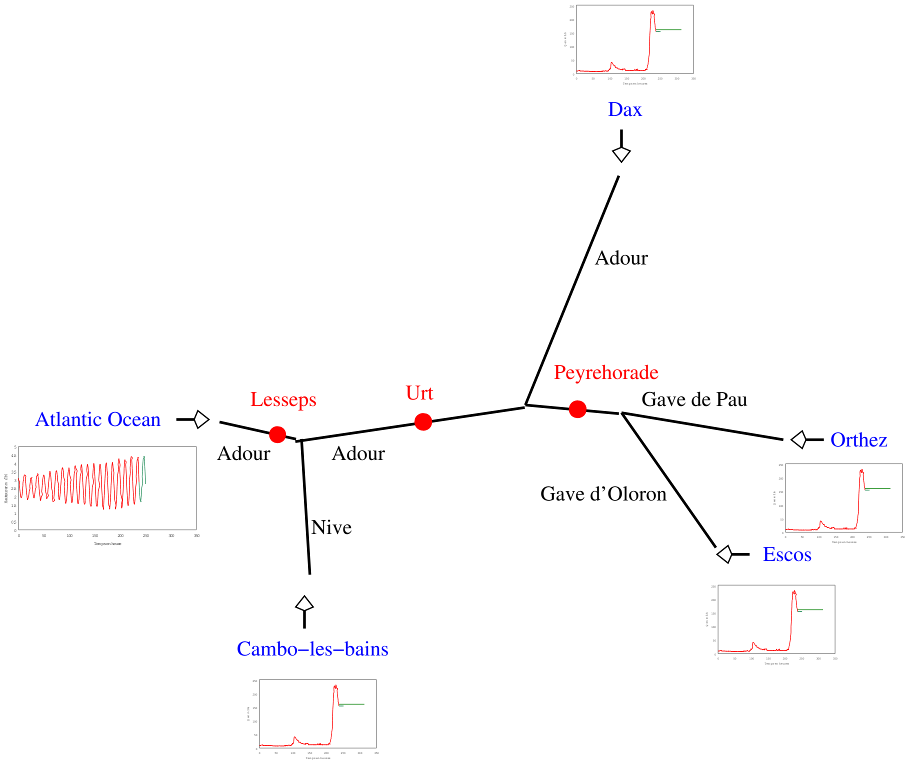

Fig. 7. The Adour catchment with the measurement stations in red and the upstream stations in blue.

HESSD
7, 9067–9121, 2010

Data assimilation for flood forecasting

S. Ricci et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀◀ | ▶▶
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

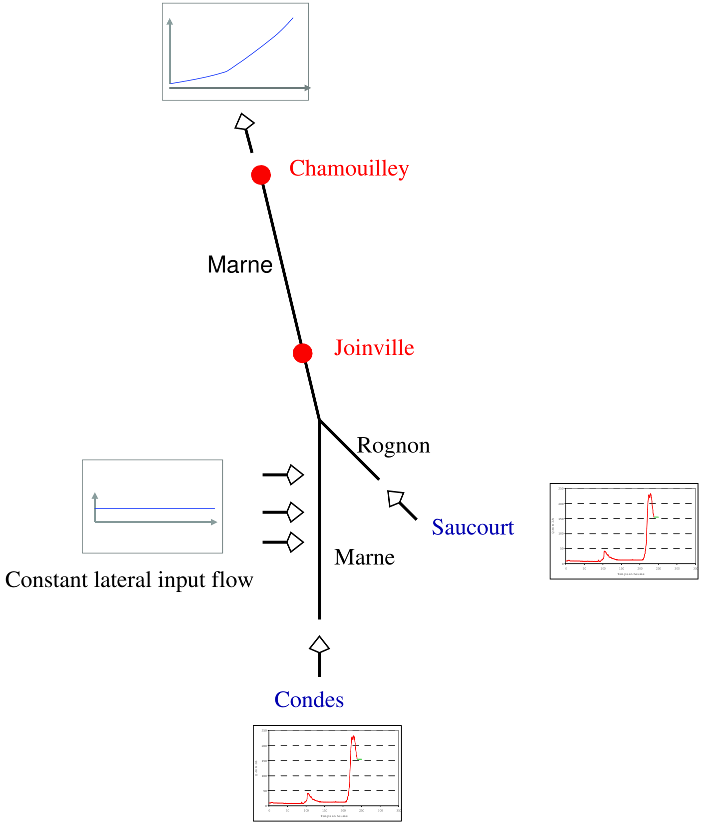

Fig. 8. The Marne Vallage catchment with the measurement stations in red and the upstream stations in blue.

HESSD
7, 9067–9121, 2010

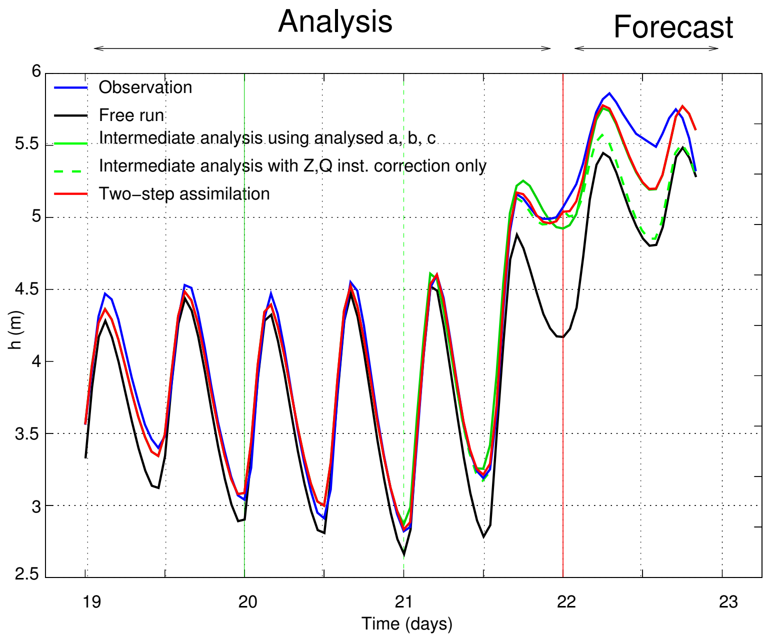

Fig. 9. November 2002 event, Adour catchment. Water level at Peyrehorade from Day 19 to Day 23: hourly observation in blue, *free run* in black, analysis with the correction of the upstream flow only in green, anaysis with the instantaneous Z, Q correction only in dashed green and analysis with the two-step assimilation in red.

HESSD
7, 9067–9121, 2010
Data assimilation for flood forecasting
S. Ricci et al.

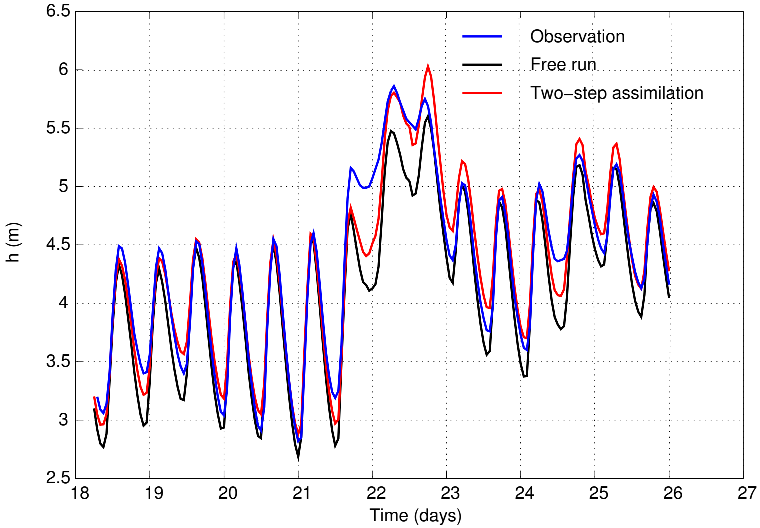

Fig. 10. November 2002 event, Adour catchment. Six-hour forecasted water level at Peyrehorade for the free run (black curve), the observation (blue curve) and the analysis with the two-step assimilation (red curve).

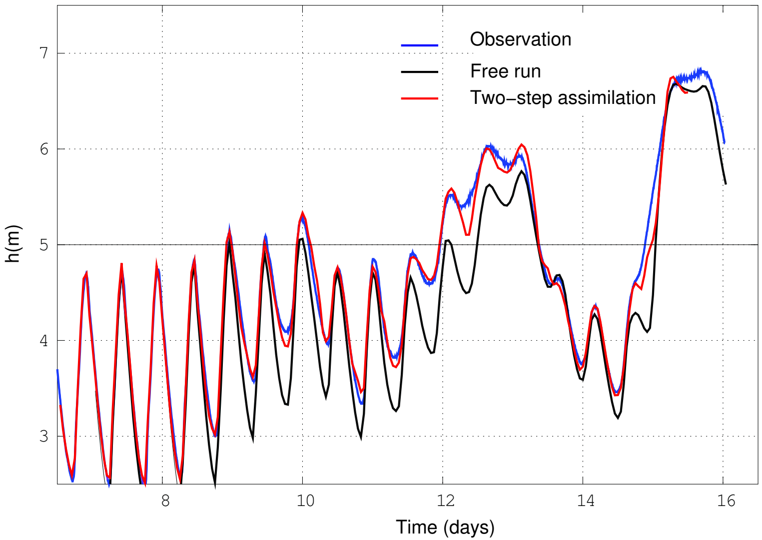

Fig. 11. November 2009 event, Adour catchment. One-hour forecasted water level at Peyrehorade for the free run (black curve), the observation (blue curve) and the analysis with the two-step assimilation (red curve).

HESSD
7, 9067–9121, 2010

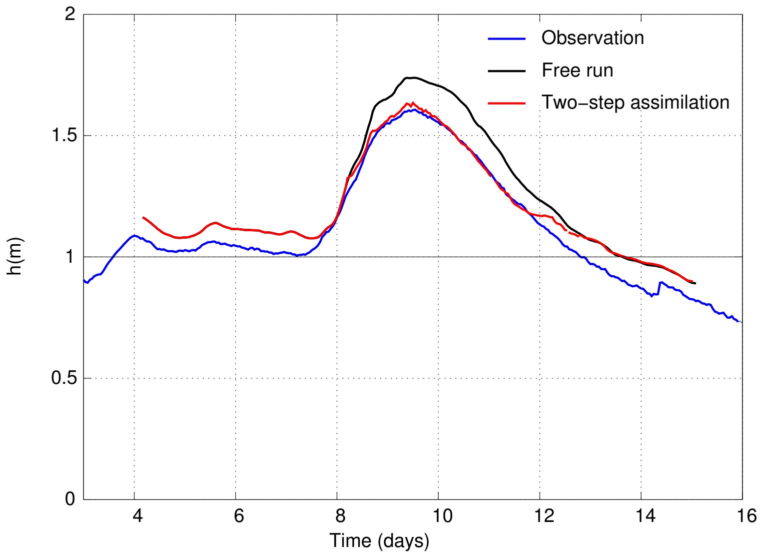

Fig. 12. April 2006 event, Marne Vallage catchment. Two-hour forecasted water level at Joinville for the free run (black curve), the observation (blue curve) and the analysis with the two-step assimilation (red curve).

HESSD

7, 9067–9121, 2010

Data assimilation for flood forecasting

S. Ricci et al.

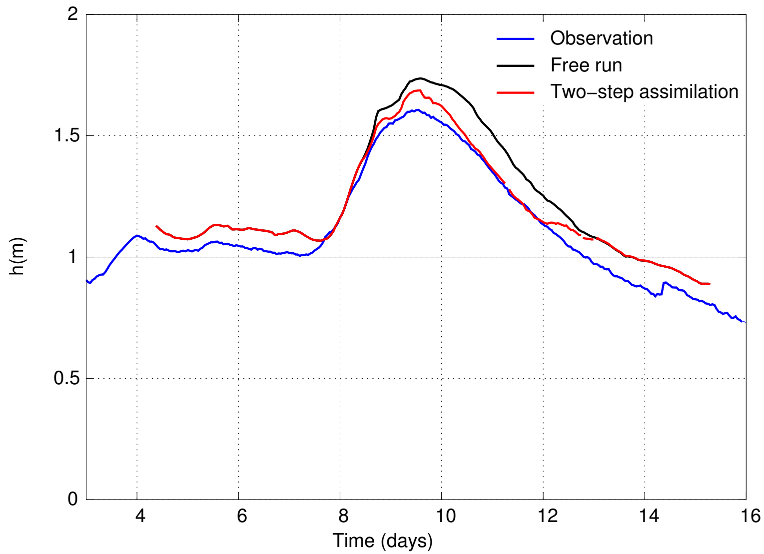

Fig. 13. April 2006 event, Marne Vallage catchment. Seven-hour forecasted water level at Joinville for the free run (black curve), the observation (blue curve) and the analysis with the two-step assimilation (red curve).

HESSD

7, 9067–9121, 2010

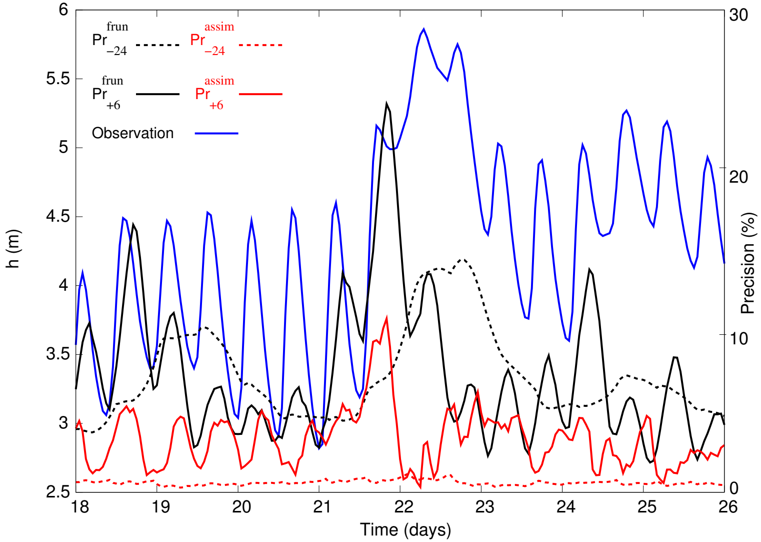

Fig. 14. November 2002 event, Adour catchment. Precision Pr for the free run (in black) and the analysis from the two-step assimilation procedure (in red) computed over 24 h before the last observation time (dashed curves) and over six hours after the last observation time (solid curves) at Peyrehorade.