**Hydrology and
Earth System
Sciences**

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, URA1875, Toulouse, 42, Avenue Gaspard Coriolis, 31057 Toulouse Cedex 01, France
[2]INPT, CNRS, IMFT, Allée Camille Soula, 31400, Toulouse, France
[3]SCHAPI, Toulouse, 42, Avenue Gaspard Coriolis, 31057 Toulouse Cedex 01, France
[4]CERFACS, Toulouse, 42, Avenue Gaspard Coriolis, 31057 Toulouse Cedex 01, France

Abstract. The present study describes the assimilation of river water level observations and the resulting improvement in flood forecasting. The Kalman Filter algorithm was built on top of a one-dimensional hydraulic model which describes the Saint-Venant equations. The assimilation algorithm folds in two steps: the first one was based on the assumption that the upstream flow can be adjusted using a three-parameter correction; the second one consisted of directly correcting the hydraulic state. This procedure was applied using a four-day sliding window over the flood event. The background error covariances for water level and discharge were represented with anisotropic correlation functions where the correlation length upstream of the observation points is larger than the correlation length downstream of the observation points. This approach was motivated by the implementation of a Kalman Filter algorithm on top of a diffusive flood wave propagation model. The study was 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 during the flood event leads to a significant improvement in the water level and discharge in both analysis and forecast modes.

1 Introduction

River stream flow forecasting is a critical issue for the security of people and infrastructures, the function of power plants, and water resources management. Many efforts have contributed to the development of open channel flow modeling, based on mass, momentum and energy conservation equations (Chow, 1959; Hervouet, 2003). Still, uncertainties in these models are such that river stream flow modeling remains strenuous work. Major uncertainties come from the model itself as the physics of the system are simplified and discretized, as well as from hydrological boundary conditions (upstream flow or lateral inflow), meteorological boundary conditions (precipitation, pressure and wind) and from hydrological initial conditions. Hydraulic models also rely on various parameterizations expressed as numerical parameters (stability conditions for the numerical scheme), geometric parameters (cross sections, gates and weir dimensions) and hydraulic parameters (flood plain storage, friction, discharge). Calibrating a hydraulic model often means adjusting Strickler coefficients, discharge coefficients at cross or lateral devices, seepage values or cross sectional geometry. The calibration of these parameters has been widely investigated (Durand et al., 2008; Geese et al., 2011; Malaterre et al., 2010) by focusing either on calibration algorithms, sensitivity indications, or optimization of the observation network.

Both parameter calibration and physical field description can be formulated as inverse problems (Tarantola, 1987). The formulation of inverse problems in hydrology fits into a wider mathematical framework presented by Maclaughlin and Townley (1996). Data assimilation combines model simulation and observational information of the system in order to provide a better estimate of it (Ide et al., 1997; Bouttier 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 (GODAE, 2009) 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., 2010a,b) and hydraulic problems with two main objectives: optimizing model parameters and improving stream flow simulation and forecasting. Existing literature proposes several methods based on a minimization technique (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 physical fields (Jean-Baptiste et al., 2010).

The present study describes the assimilation of river water level observations and the resulting improvement in flood forecasting. The data assimilation algorithm was built on top of a one-dimensional hydraulic model describing the Saint-Venant[1] equations. The assimilation algorithm folds in two steps: the first one was based on the assumption that the upstream flow can be adjusted using a three-parameter correction, the second one consisted of directly correcting the hydraulic state. The variables to correct are gathered in the control vector; the control vector is indeed different in the two previously described assimilation steps. For both steps, a Kalman Filter algorithm was applied. In order to decrease the cost of the data assimilation algorithm, the background error covariance matrix for the second step was not propagated by the dynamics of the hydraulic model. The impact of the analysis and propagation steps in the Kalman Filter in this matrix was emulated: anisotropic correlation functions were used to represent the spatial error correlations for water level and discharge. This choice results from the implementation of the Kalman Filter algorithm on a simplified hydraulic model (representing the diffusive flood wave propagation equations). This demonstrates that the analysis and dynamics of the physics turns a Gaussian correlation function into an anisotropic function at the observation point. The data assimilation study with MASCARET was performed on the Adour (France) and the Marne Vallage (France) catchments. The improvement in river water level predictions, using data assimilation, in analysis and forecast modes are presented within this paper.

The outline of the paper is as follows: Sect. 2 describes the assimilation system, paying particular attention to the choice of the control vector for the data assimilation algorithm. Two approaches to data assimilation were implemented: the correction of the hydraulic state and the control of the upstream flow. The modeling of the background covariances matrix and the parameterization used to control the upstream flow are highlighted in this section. Section 3 provides the theoretical framework explaining the choice of anisotropic correlation functions for the spatial error correlations in the background error covariance matrix **B**. In Sect. 4, the improvements in the river flood simulations and forecasting are presented. The evaluation of the statistics describing the difference between the simulation results and the observations in re-analysis or forecast modes is used to illustrate the assimilation scheme performance. A summary and a discussion are finally given in Sect. 5.

2   Context and implementation of the data assimilation

2.1   Modeling of the physics

MASCARET is a one-dimensional free surface hydraulic model developed by EDF[2] and CETMEF[3], based on the Saint-Venant equations (Goutal and Maurel, 2002). MASCARET is widely used for modeling flood events, submersion waves resulting from the failure of hydraulic infrastructures, regulation of river infrastructures, and canal waves propagation.

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

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

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

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

2.2   The data assimilation method

The Kalman Filter (KF) approach (Gelb, 1974; Todling and Cohn, 1994; Talagrand, 1997) identifies the optimal estimate of the true value of an unknown variable x. This estimate is optimal when its variance is at a minimum, meaning, for Gaussian cases, that its probability density function is dense around the mean. Suppose that x is the control vector which can include the hydraulic 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 KF algorithm is the analysis vector x^a. The a priori knowledge of the system is the background vector x^b and the observation vector is y^o. The background, observation and analysis error covariances are respectively

---

[1]Shallow water equations

[2]Electricité de France

[3]Centre d'Etudes Techniques Maritimes Et Fluviales

gathered in the matrices $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{A}$. Assuming that the background, the observation and the analysis are unbiased, the analysis at time i can be formulated as a correction to the background state defined as:

$$x_i^{\mathrm{a}} = x_i^{\mathrm{b}} + \mathbf{K}_i \left[y_i^{\mathrm{o}} - H_i(x_i^{\mathrm{b}}) \right], \tag{2}$$

where $\mathbf{K}_i$ is the gain matrix, d_i is the innovation vector

$$d_i = y_i^{\mathrm{o}} - H_i(x_i^{\mathrm{b}}), \tag{3}$$

and $y_i = H_i(x_i)$ is the model equivalent of the observations, generated by the observation operator H_i.

The KF analysis is optimal when the variance of its error is at a minimum. Minimizing the variance of the error analysis comes down to minimizing the trace of the analysis error covariance matrix which leads to the formulation of the gain matrix (Bouttier and Courtier, 1999):

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

In this formulation, $\mathbf{H}_i$ is the Jacobian matrix of H_i in the vicinity of the background state x_i^{b} which can be written as:

$$\mathbf{H}_i = \frac{\partial y_i}{\partial x_i} = \frac{\partial H_i(x_i)}{\partial x_i}. \tag{5}$$

The analysis error covariance matrix reads

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

The analysis at time i is propagated in time by a dynamic model which defines the background at time $i+1$:

$$x_{i+1}^{\mathrm{b}} = M_{i,i+1}(x_i^{\mathrm{a}}), \tag{7}$$

where $M_{i,i+1}$ represents the model propagation between i and $i+1$.

The analysis error covariance matrix at time i is propagated in time by the dynamic model in order to define the background error covariance matrix at time $i+1$ (when the model is assumed to be perfect):

$$\mathbf{B}_{i+1} = \mathbf{M}_{i,i+1} \, \mathbf{A}_i \, \mathbf{M}_{i,i+1}^T, \tag{8}$$

where $\mathbf{M}_{i,i+1}$ is the tangent linear approximation of $M_{i,i+1}$ in the vicinity of x_i^{b}.

2.3 Implementation of the assimilation scheme

The water levels simulated with MASCARET (or any hydraulic model), may be significantly different from the observed water levels. A two step data assimilation algorithm was implemented to account for some of the uncertainties in the hydraulic model (inputs and outputs). The first step was based on the assumption that the error in the simulated water level was mainly due to an imperfect description of the upstream flows. The second step consisted of dynamically correcting the water level and the discharge states for the entire catchment (discretized in m cells) when observations were available. The two-step data assimilation procedure over the time window $[0, T_r]$ is described in Fig. 1.

2.3.1 Correction of the upstream flow

The first data assimilation approach is based on the assumption that a considerable part of the error in the simulated water level can be attributed to uncertainty in the upstream boundary condition (usually deduced from water level observations through a calibration procedure). This first step is described in the top part of Fig. 1. In order to control this uncertainty with a data assimilation procedure, a large control vector which contains the discharge boundary conditions at each time step of the simulation period, should be introduced. This would result in a computationally expensive data assimilation procedure, especially for the computation of the Jacobian of the observation operator since the relation between the control space and the observation space is given by an integration of the numerical model. In order to compute the Jacobian of the observation operator, the numerical model should be partially differentiated with respect to each element of the control vector, hence the size of the control vector should be reduced.

For this reason, the upstream flow forcing f was corrected through a three-parameter linear transformation over a time window (assimilation window $[0, T_r]$):

$$\widetilde{f}(t) = af(t-c) + b. \tag{9}$$

This parametric correction enables a simple and physical control of the time series: homothetic vertical transformation (a), shift in amplitude (b) and shift in time (c). For instance, at the upstream stations, water levels are usually observed and a rating curve (extrapolated for high discharge values) is used to describe the discharge time series used by the hydraulic model. The parameters a, b, c allow for a correction of the uncertainty related to the use of this rating curve. Additionally, the parametric correction allows for the estimation of an unknown intermediate input flow, accounting for influents that are not modeled in the hydraulic network.

For this approach, the control vector is composed of the coeficients a, b, c for each of the S upstream stations $x_i = (a_1, b_1, c_1, \cdots, a_s, b_s, c_s, \cdots, a_S, b_S, c_S)$. The characteristics of this data assimilation approach are:

- The control parameters are assumed to be constant over the time window over which the data assimilation is performed. Since no models for the temporal evolution of the parameters are described, $M_{i,i+1} = I$ in Eqs. (7) and Eq. (8). For this reason, the indice i is dropped in the following of the Sect. 2.3.1.

- The background values for the control parameters are $x^{\mathrm{b}} = (a_1^{\mathrm{b}}, b_1^{\mathrm{b}}, c_1^{\mathrm{b}}, \cdots, a_s^{\mathrm{b}}, b_s^{\mathrm{b}}, c_s^{\mathrm{b}}, \cdots, a_S^{\mathrm{b}}, b_S^{\mathrm{b}}, c_S^{\mathrm{b}}) = (1, 0, 0, \cdots, 1, 0, 0, \cdots, 1, 0, 0)$.

- The size of the background error covariance matrix is $(3 \times S)^2$. The errors in the background parameters $(a_1^{\mathrm{b}}, b_1^{\mathrm{b}}, c_1^{\mathrm{b}}, \cdots, a_s^{\mathrm{b}}, b_s^{\mathrm{b}}, c_s^{\mathrm{b}}, \cdots, a_S^{\mathrm{b}}, b_S^{\mathrm{b}}, c_S^{\mathrm{b}})$ are assumed
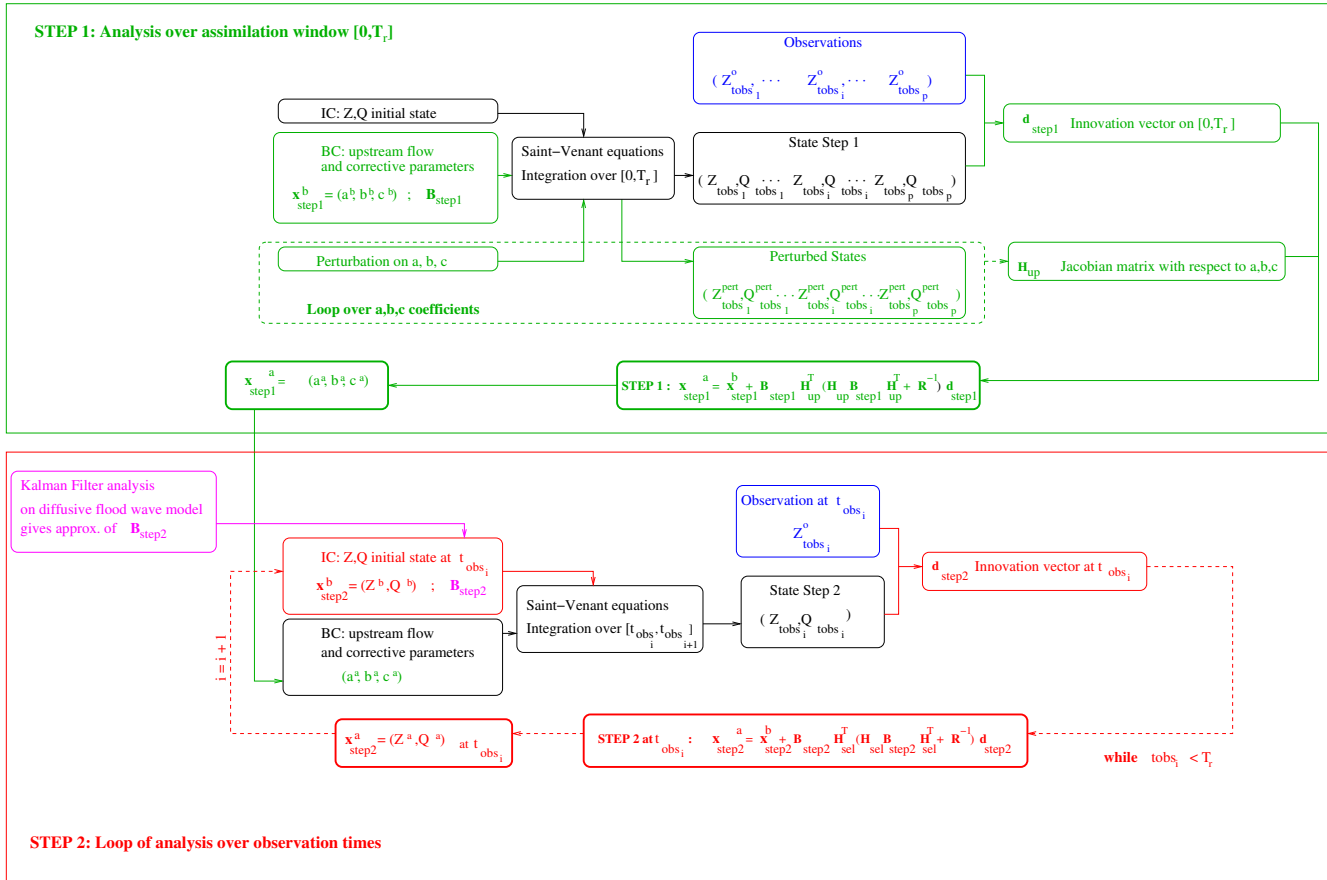
Fig. 1. Two-step data assimilation procedure flowchart on the time window $[0, T_r]$.

to be uncorrelated and the variances are estimated statistically to represent the variability of the upstream flow.

– The observation vector represents the water level at observation times at selected locations in the hydraulic network. It is a vector of size p where p is the number of observations over the data assimilation time window.

– The observation covariance matrix $\mathbf{R}$ is a $p \times p$ matrix. Its diagonal terms are the observation error variances at the observation points, which are estimated from statistics using several sets of measurements. The off-diagonal terms are covariances between the observation errors at different observation points; these correlations are assumed to be negligible since the observation points are far enough from each other.

– 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 of two operations, the more costly of which is the integration of the hydraulic model given the upstream flow conditions over the assimilation window. The second operation is the selection of the cal-

culated water level at the observation points and at the observation times.

– $H_{up}(\boldsymbol{x}^{b})$ represents the water level at the observation points and times computed by MASCARET using the background parameters $(a_1^b, b_1^b, c_1^b, \cdots, a_s, b_s, c_s, \cdots, a_S^b, b_S^b, c_S^b)$.

– The Jacobian $\mathbf{H}_{up}$ of H_{up} is the tangent linear of the hydraulic model computed in the vicinity of $\boldsymbol{x}^{b}$.

The Jacobian matrix $\mathbf{H}_{up}$ can be approximated in the vicinity of the background $\boldsymbol{x}^{b}$ as follows:

$$H_{up}(\boldsymbol{x}^{b} + \Delta\boldsymbol{x}) \approx H_{up}(\boldsymbol{x}^{b}) + \mathbf{H}_{up}|_b \Delta\boldsymbol{x}, \tag{10}$$

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

$$\mathbf{H}_{up,jk}|_b = \frac{\partial y_j}{\partial x_k} = \frac{\partial H_{up,j}(\boldsymbol{x}^{b})}{\partial x_k}$$

$$\approx \frac{H_{up,j}(\boldsymbol{x}^{b} + \Delta\boldsymbol{x}) - H_{up,j}(\boldsymbol{x}^{b})}{\Delta x_k} = \frac{\Delta y_j}{\Delta x_k}. \tag{11}$$

In this study, Δy_j is the change in water level at the observation station j resulting from a change Δx_k in the k-th control

variable (a_s, b_s or c_s) with $s \in \{1, \cdots, S\}$. The ability of the observation operator to represent, in an approximate sense, the response of the water level to changes in the control vector is a crucial tenet of this algorithm. The computation of $\mathbf{H}_{up}$ requires an additional integration of the hydraulic model for each control parameter. An efficient computation of the operator $\mathbf{H}_{up}$ in the case of a larger control space was implemented by Thirel et al. (2010a).

The small size of the control vector, as well as of the observation vector, enables the use of a KF algorithm involving matrix operations for the computation of the gain matrix. Still, the algorithm relies on the hypothesis that the observation operator can be approximated by a linear operator on the $[\mathbf{x}^b, \mathbf{x}^a]$ interval. The linearity of the hydraulic model response to a perturbation in the control parameters (a_s^b, b_s^b, c_s^b) (with $s \in \{1, \cdots, S\}$) was investigated. The difference between the right hand side and the left hand side of Eq. 10 should be quantified, and ideally should not exceed a couple of percents, in order to assess the integrity of the linearity assumption. It was found that the relation between an upstream flow perturbation (of the form Eq. 9) and the hydraulic state response is reasonably approximated by a linear function in the vicinity of $\mathbf{x}^b$.

The implementation of this algorithm allows not only for an improvement in the simulated water level within the assimilation window but also for an improvement in the forecast since, in forecast mode, the upstream flow is set equal to the last analyzed value.

2.3.2 Correction of the hydraulic state

The second data assimilation approach consists of dynamically correcting the water level and discharge states for the entire catchment (discretized in m cells) when observations are available (time i in Eqs.2 to 8). This second step is described in the bottom part of Fig. 1. The observation vector is kept the same as the one previously described but in this case evaluated at a given time t_{obs_i} instead of over a time window (in the following, the subscript t_{obs_i} is replaced by the subscript i).

Here, the control vector at time i, is composed of the discretized water level and discharge states $x = \left(Z_{x_1}, \cdots, Z_{x_m}, Q_{x_1}, \cdots, Q_{x_m} \right) = (\mathbf{Z}, \mathbf{Q})$.

The background state is given by a previous integration of the model M describing the Saint-Venant equations; it is composed of the simulated water level and discharge vectors and is denoted by $\left(\mathbf{Z}^b, \mathbf{Q}^b \right)$. The size of the control and the background vectors is n = 2m.

In Eq. (7), $M_{i,i+1}$ denotes the propagation of the hydraulic state by the non-linear equation Eq. (1). In Eq. (8), $\mathbf{M}_{i,i+1}$ denotes the tangent linear approximation of $M_{i,i+1}$. In the application with MASCARET, as the computation of $\mathbf{M}_{i,i+1}$ was too costly, the propagation of the background error covariance matrix $\mathbf{B}_i$ was not explicitly implemented; it was assumed that $\mathbf{M}_{i,i+1} = \mathbf{I}$ in Eq. (8) so that $\mathbf{B}_{i+1} = \mathbf{B}$.

A parametrization for the $\mathbf{B}$ matrix was chosen to emulate the propagation of the covariance function by the hydraulic model. This parametrization, presented in Sect. 3, results from the application of a full Kalman filter algorithm (where Eq. (8) is solved) on a simplified propagation model with a steady observation network. The propagated covariance matrix from this application was then used as the invariant $\mathbf{B}$ in the Kalman Filter algorithm for MASCARET. This algorithm will be further referred to as IKF (Invariant Kalman Filter). The background error matrix for this second step of assimilation is denoted by $\mathbf{B}_{step2}$ in Fig. 1.

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 respectively the statistics of the errors ϵ_Z in the water level and ϵ_Q in the discharge. Its diagonals represent respectively the variance of the background error in the water level and discharge whereas the extra diagonal terms of these blocks are the covariances between the error in the water level or 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 $\mathbf{B}_{Z,Q}^T$, which represent the covariances between the errors in the water level and the errors in the discharge.

The innovation vector $\mathbf{d}$ (Eq. 3) expresses the difference between the observed water level and the simulated water level at the nearest grid point. The observation operator is a selection matrix with dimensions, $p \times n$, denoted by $\mathbf{H}_{sel}$. In this study, the observation network is stationary, thus the observation operator doesn't vary over the assimilation cycles, the index i is then dropped in the following. The water level correction vector at the observation points, $\widetilde{\delta \mathbf{Z}} = (\widetilde{\delta \mathbf{Z}_1}, \cdots, \widetilde{\delta \mathbf{Z}_l}, \cdots, \widetilde{\delta \mathbf{Z}_p})$ (with $l \in \{1, \cdots, p\}$) is the product of the innovation vector and the matrix product $\mathbf{H}_{sel}^T (\mathbf{H}_{sel} \mathbf{B} \mathbf{H}_{sel}^T + \mathbf{R})^{-1}$ in Eq. (4):

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

Water level correction $\delta \mathbf{Z}$ over the entire domain results from the multiplication of $\widetilde{\delta \mathbf{Z}}$ by $\mathbf{B}_{Z,Z}$. The water level variances translate as uncertainties in the simulated water level. An anisotropic correlation function ρ was used to describe the spatial error correlations of $\delta \mathbf{Z}$ as presented in Fig. 2.

In order to keep the anayzed control vector coherent with the Saint-Venant equations, the discharge state should be corrected along with the water level state. This was done specifying multivariate error covariances in $\mathbf{B}_{Z,Q}$. The discharge correction vector at the observation point $\widetilde{\delta \mathbf{Q}} = (\widetilde{\delta \mathbf{Q}_1}, \cdots, \widetilde{\delta \mathbf{Q}_l}, \cdots)$ with $l \in \{1, \cdots, p\}$, was deduced from $\widetilde{\delta \mathbf{Z}}$ at the observation points using the local rating curve. Assuming the rating curve can be formulated as

$$Q_{rc}(Z) = \alpha Z^\gamma + \beta, \tag{13}$$

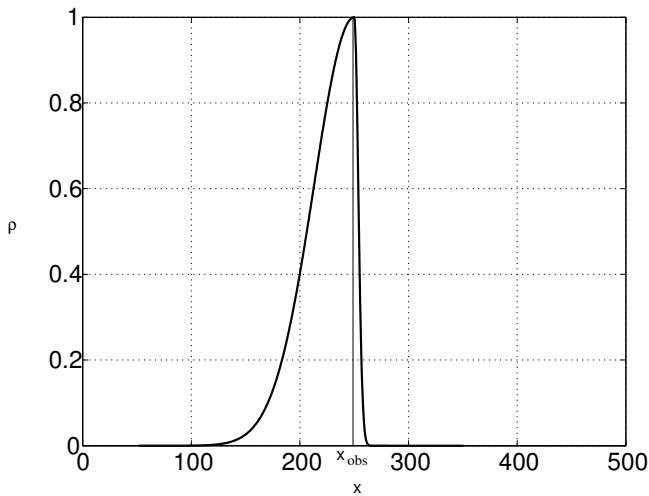**Fig. 2.** Local anisotropic correlation function at observation point x_{obs}.

the discharge correction at the observation point reads:

$$\widetilde{\delta Q_l} = Q_l^b \frac{Q_{\text{rc},l}(Z_l^b + \widetilde{\delta Z_l}) - Q_{\text{rc}}(Z_l^b)}{Q_{\text{rc}}(Z_l^b)}, \quad \text{for} \quad l \in [1, p], \quad (14)$$

where Q_l^b and Z_l^b are the background values for the water level and discharge at the observation points.

A calibration procedure was used to evaluate the coefficients α, β, γ at each observation point. However, because of the tidal influence at some observation points, the identification of a bijective function (valid for both high and low tides) was not always possible. In this case, the rating curve was crudely approximated by the identity function, leading to

$$\widetilde{\delta Q_l} = \frac{Q_l^b}{Z_l^b} \widetilde{\delta Z_l}, \quad \text{for} \quad l \in [1, p]. \quad (15)$$

As for the water level, the discharge correction δQ for the whole domain was determined from the multiplication of $\widetilde{\delta Q}$ by $\mathbf{B}_{Q,Q}$. The correlation function and length for δQ are the same as those used for δZ.

2.3.3 Cycling of the analysis

The two previously described assimilation approaches are sequentially applied over the period covering a flood event as described in Fig. 1. The assimilation is performed over a four-day sliding window, also referred to as a cycle, with three days of re-analysis and one day of forecast. The last observation time from which the forecast integration starts is the reference time T_r. The sliding window is shifted every hour and a new assimilation performed. The forecasted state at T_r is stored and used as the initial state for the 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 and forecast procedure is schematically represented in Fig. 3. 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 observations from the first day are not used as the model is potentially not adjusted yet. The resulting analyzed parameters are used to correct the upstream flows over the first, second and third days. This is the first step of the analysis. The updated upstream flows are then 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 as it describes the background state for the hydraulic state correction procedure. During the third day of the integration, at each observation time, the water level is adjusted; this correction is instantaneous and correspond to the second step of the analysis. The model is then integrated starting from the corrected state at the current observation time to the next observation time, leading to a discontinuous description of the hydraulic state (discontinuous red curve). In this study, the observation time step is equal to the model time step so that the resulting integration is no more discontinuous than any other model integration.

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

The data assimilation algorithm was 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 use with the MERCATOR project. PALM allows for the coupling of independent code components with a high level of modularity in the data exchanges and treatment while providing a straightforward parallelization environment (Fouilloux and Piacentini, 1999; Buis et al., 2006).

3 Modeling of B

As explained in Sect. 2.3.2, when the Kalman Filter algorithm is applied to correct the hydraulic state computed by MASCARET, the explicit propagation of the background error covariance matrix is not implemented because the computation of the tangent linear of the model is too costly. The covariance functions initially described in **B** are kept constant instead of being propagated by the dynamic model. For that reason, it is crucial to model covariance functions that account for some of the physics of the dynamic model, rather than Gaussian functions.

The objective of the present Section is to provide a parametrization for the **B** matrix that emulates the propagation of the covariance function by the hydraulic model. In
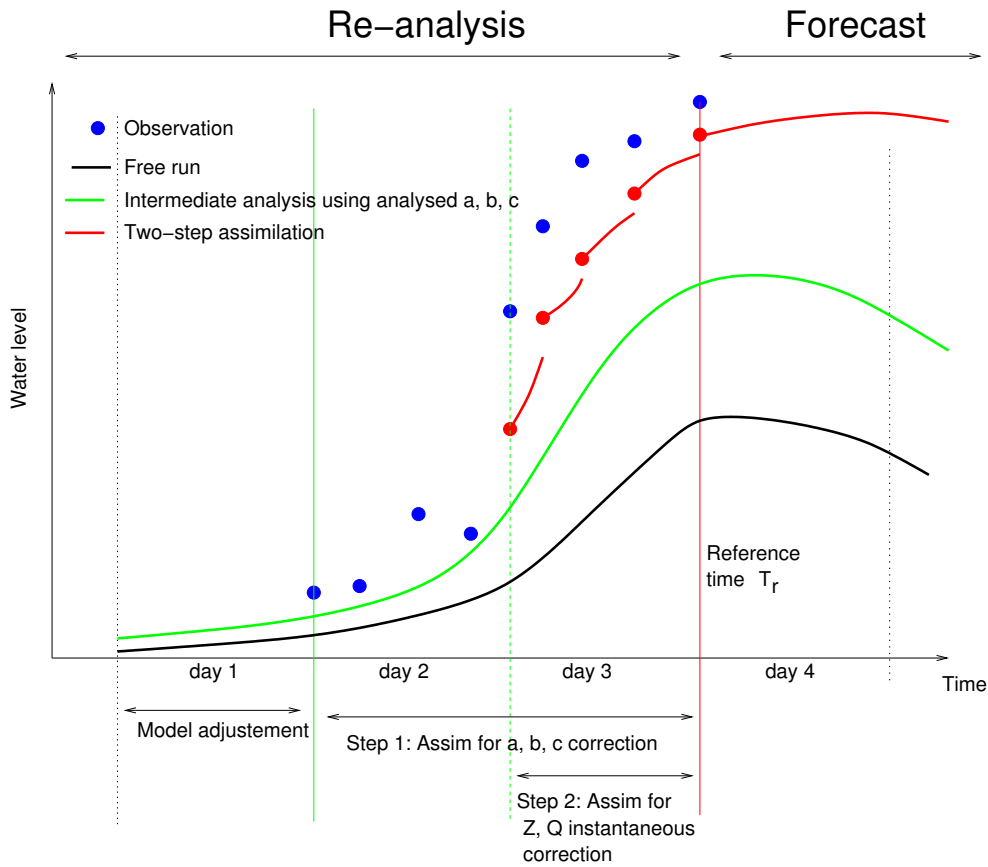
Fig. 3. Schematic representation of the two-step assimilation algorithm over one cycle.

order to find such a parametrization, an initially Gaussian covariance function is used by a full Kalman Filter algorithm applied to a simplified dynamic model. For the simplified model, which will be described in Sect. 3.1 (diffusive flood wave approximation), the analysis and propagation steps of the Kalman Filter are achieved since the computation of the tangent linear for a diffusive flood wave propagation model on a restricted spatial domain is reasonable. With this experiment, it will be demonstrated that the analysis and propagation steps of the Kalman Filter modify the covariance function at the observation point. As will be shown in Sect. 3.2, the resulting covariance function at the observation point is anisotropic, with a shorter correlation length downstream of the observation point than upstream. A validation step for this parametrization will be performed with the diffusive flood wave propagation model in Sect. 3.3. In this section, an anisotropic covariance function will be described for the IKF and the results of the assimilation will be compared to those of the IKF with a Gaussian covariance function. It will be shown that the emulation of the dynamics of the model with a parametrized covariance function for the IKF is then equivalent to the full KF.

These parametrization and validation exercises seek to justify the choice of an anisotropic covariance function at the observation points as opposed to Gaussian functions. This result will later be used to model the background error covariance function for the hydraulic state correction procedure performed on MASCARET with IKF as represented in Fig. 1.

3.1 The diffusive flood wave approximation

In this study, it is to assumed that the solution of the propagation of a given initial condition by the MASCARET equations 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 thus its correlation length l) propagated by MASCARET is similar to the covariance function of the same signal propagated by the diffusive flood wave approximation equations.

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

$$\frac{\partial \widetilde{h}}{\partial t} + \frac{5U_n}{3}\frac{\partial \widetilde{h}}{\partial x} = \kappa \frac{\partial^2 \widetilde{h}}{\partial x^2}, \tag{16}$$

where $Q = hU$, $h = h_n + \widetilde{h}$ and $U = U_n + \widetilde{U}$, with $(\widetilde{h}, \widetilde{U})$ representing 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) is 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 propagation model is chosen as a representative mean state for the following simulations with MASCARET over each catchment. Eq. (16) is a classical advection-diffusion equation where κ is the diffusion coefficient and $c = \frac{5U_n}{3}$ is the advection speed. In order to use this model as a support for data assimilation, an open boundary condition for Eq. (16) is imposed downstream with $\frac{\partial \widetilde{h}}{\partial t}(L,t) + c \frac{\partial \widetilde{h}}{\partial x}(L,t) = 0$. The upstream boundary condition is imposed by $\widetilde{h}(0,t) = \widetilde{q}(t)$, where $\widetilde{q}$ is a random Gaussian function of zero mean $<\widetilde{q}> = 0$. The auto-correlation function of $\tilde{q}(t)$

$$R(\tau) = <\widetilde{q}(t)\widetilde{q}(t+\tau)> = \delta q_{\mathrm{m}}^2 \exp\left(-\frac{\tau^2}{2l_q^2}\right) \qquad (17)$$

is assumed to have the shape of a Gaussian function. Using the theory of random function diffusion, the spatial covariance function of $\widetilde{h}(x,t)$ can be approximated by a Gaussian function.

3.2 Kalman Filter algorithm applied to the diffusive flood wave propagation model

The Kalman Filter algorithm was implemented on the 1-D diffusive flood wave propagation model described by Eq. (16), using the identical-twin experiment framework (also known as OSE[4]). The identical-twin experiment was set up with $t \in [0,T]$ and $x \in [0,L]$. The 1-D domain was discretized in m cells and Eq. (16) was integrated using an explicit Euler scheme in time and a first order centered finite difference scheme in space. A reference run was integrated using a set of parameters and forcing $(c_{\mathrm{true}}, \kappa_{\mathrm{true}}, \widetilde{q}_{\mathrm{true}}(t))$, to simulate the *true* water level $\widetilde{h}_{\mathrm{true}}$. The observation $h_{\mathrm{obs}}(t) = \widetilde{h}_{\mathrm{true}}(t) + \epsilon_o(t)$ was calculated in 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 $h_b(x,t)$ was integrated using a perturbed set of parameters and forcing $(c_{\mathrm{per}}, \kappa_{\mathrm{per}}, \widetilde{q}_{\mathrm{per}}(t))$ where $<\widetilde{q}_{\mathrm{per}}(t)\widetilde{q}_{\mathrm{per}}(t+\tau)> = \delta q_{\mathrm{m,per}}^2 \exp(-\frac{\tau^2}{2l_q^2})$.

In this context, the background error covariance matrix was updated by the analysis and propagated in time, to the next observation time, by the tangent linear of the diffusive flood wave propagation model (Eqs. 7–8). As a consequence, the gain matrix evolves over the assimilation cycles.

As described in Sect. 3.1, the initial covariance function at the observation point, for the signal $\widetilde{h}(x,t)$, is close to a Gaussian as represented in black in Fig. 4. Appendix A describes how this isotropic covariance function is modified
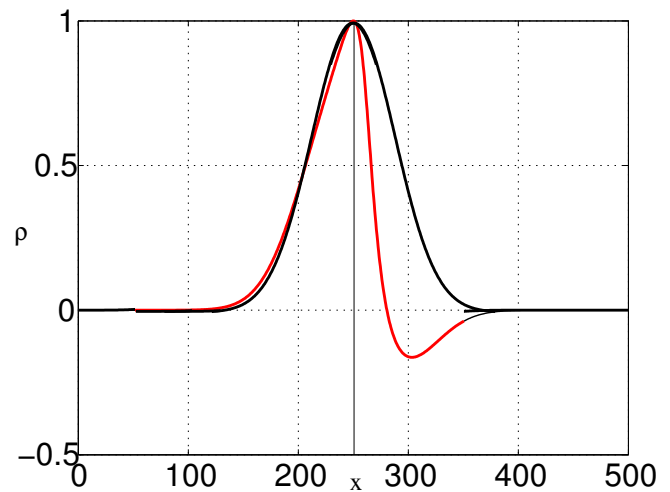
[4]Observing System Experiment

Fig. 4. Initial Gaussian background error covariance function (black curve) and anisotropic background error covariance function (red curve) from KF, at the observation point.

along the assimilation cycles by the analysis and propagation steps of the Kalman Filter algorithm. Considering a steady observation network, the shape of the covariance function converges to an anisotropic function as represented in red in Fig. 4. The covariance between the observation point and its neighbours is reduced since information at the observation point was introduced at this location by the analysis procedure through the innovation vector. The background error covariance matrix for the next assimilation 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. (8). After several KF cycles, the covariance function at the observation point is characterized by a shorter correlation length scale downstream of the observation point than upstream (see Appendix A for further details).

3.3 Invariant Kalman Filter algorithm applied to the diffusive flood wave propagation model

The covariance function computed with the Kalman Filter algorithm in Sect. 3.2 was used to parameterize the invariant covariance function of the IKF; here applied to the diffusive flood wave propagation model, for validation. The results of the IKF using this parametrisation of $\mathbf{B}$ will be compared to those of a IKF using an isotropic Gaussian function.

By default, the spatial correlations in $\mathbf{B}$ are represented by the Gaussian function

$$\rho(x,x') = \exp\left[-\frac{(x-x')^2}{2l_B(x,x')}\right] \qquad (18)$$

meaning that $\mathbf{B}(x,x') = \sigma_b^2 \rho(x,x')$. The length $l_B(x,x')$, is an isotropic function of x and x' which represents the local

correlation length for the pair (x, x'). For this study, only the correlations $\rho(x, x_{obs})$ between the errors at x_{obs} and the rest of the domain are relevant. The length $l_B(x, x_{obs})$ is assumed to depend only on the observation location and is denoted by $l(x_{obs})$.

Figure 5a, b, c shows the true $\tilde{h}_t$, the non-assimilated $\tilde{h}_s$, background $\tilde{h}_b$ and analysed $\tilde{h}_a$ water level state over the 1-D domain at $t = T = 500 \times 10^3$ s where the analysis is performed every $\Delta t = 10 \times 10^3$ s, for different functions $l_B(x, x_{obs})$. When $l_B(x, x_{obs}) = l(x_{obs})$ for all x (Fig. 5a), the data assimilation corrects the water level over the interval $[x_{obs} - l(x_{obs})/2, x_{obs} + l(x_{obs})/2]$. Still, the analysis (red curve) is closer to the true state (blue dotted curve) than the background (green curve) only upstream of the observation point. To the contrary, when $l_B(x, x_{obs}) = l(x_{obs})/10$ for all x, in Fig. 5b, the analysis is closer to the true state only downstream of the observation point. Finally, as it appears in Fig. 5c, a better fit to the true state is obtained with an anisotropic function $\rho(x, x_{obs})$ as predicted by the full Kalman Filter algorithm. An optimal value for the reduction factor of the length scale was identified by trial and error; the best results were obtained with

$$
\begin{aligned}
l_- &= l_B(x, x_{obs}) = \quad l(x_{obs}) \quad \text{when } x < x_{obs}, \\
l_+ &= l_B(x, x_{obs}) = l(x_{obs})/10 \text{ when } x > x_{obs}.
\end{aligned}
\tag{19}
$$

3.4 Parameterization of the covariance function for the MASCARET application

The application of a full Kalman Filter on a diffusive flood wave propagation model enabled the understanding of the impact of the analysis and the physics on an initial Gaussian correlation function. It was shown that the correlation length scale is reduced downstream of the observation point and that the initial Gaussian correlation function evolves into an anisotropic correlation function. These results were used to model the correlation function for the water level and discharge computed in the MASCARET data assimilation procedure. An approximate reduction factor of ten was applied between the correlation lengths upstream and downstream of the observation points.

In order to complete the modeling of the background error covariance function, the value of the correlation length $l(x_{obs})$ was then estimated. The estimation of the correlation length of the spatial correlation function for the errors in the water level and the discharge simulated with MAS-CARET occured in two steps. First a diffusion coefficient κ based on the dynamics of the diffusive flood wave approximation model (Eq. 16) was graphically estimated by studying the propagation of a perturbation of the hydraulic state. Then, this diffusion coefficient was used to formulate the spatial correlation length of the state perturbation covariance function. This procedure was used to predict the correlation length at each observation point for the data assimilation in

MASCARET. The details for the estimation of the correlation length $l(x_{obs})$ are given in Appendix B.

4 Results of the IKF analysis

4.1 Experimental procedure

4.1.1 Description of the catchments

The Adour maritime catchment area is located in Southwestern France, from the Pyrenean Piedmont to the Aquitain coast. The drainage area $(16\,890\,\text{km}^2)$ covers the departements of Atlantic Pyrenees and Landes. The Adour river rises in the Pyrenees at an altitude of 2600 m and reaches the Atlantic ocean at Bayonne 312 km further. The Adour catchment is one of the wettest in France due to heavy precipitations in the upper part of the basin. The Adour catchment is divided in two regions: the mouth of the river which is mostly influenced by the tide and the upstream region which is mostly influenced by influents. A schematic description of the Adour catchment is shown in Fig. 6. The Adour river has three main influents (responsible for 65% of the total discharge at Bayonne during flood conditions). The Gaves de Pau and d'Oloron, respectively draining $5226\,\text{km}^2$ and $608\,\text{km}^2$, are often affected by flash floods and join with the main influent of the catchment Gave Réunis. The Nive drains $980\,\text{km}^2$ and joins up with the Adour close to Bayonne.

The hydrological data at the upstream stations (Dax, Escos, Orthez and Cambo-les-bains) are provided in real time by the SPC[5]. The discharge time series are used as boundary conditions for the hydraulic model. The maritime boundary conditions are given by the SPC tide gauge located in the estuary. Tide forecasts are given by the SHOM[6]. The uncertainty in the maritime boundary condition is smaller than that in the hydrological upstream station, as a consequence, only upstream stations were controled by the data assimilation algorithm. Sensitivity tests revealed that the tangent linear model (Eq. 11) is valid for a pertubation up to 20 % in a, $6\,\text{m}^3\,\text{s}^{-1}$ in b and 6 h in c. Additionally, tide gauge observations located at Lesseps, Urt and Peyrehorade stations display the water level every five minutes or hourly. These observations were used for the data assimilation process. The correlation lengths were set, using the procedure described in Sect. 3, to 20 km, 6 km and 34 km at Peyrehorade, Urt and Lesseps, respectively.

The Marne Vallage catchment is located East of the Paris basin. The Marne river is the main influent of the Seine river and is 525 km long. This study focuses on the Marne Vallage drainage area that lies between Condes and Chamouilley. This karstic basin is characterized by slow flood rises, long flooding periods and a strong sensibility to local precipitations. The Marne river has two main influents, of which

---

[5] Service de Prévision des Crues

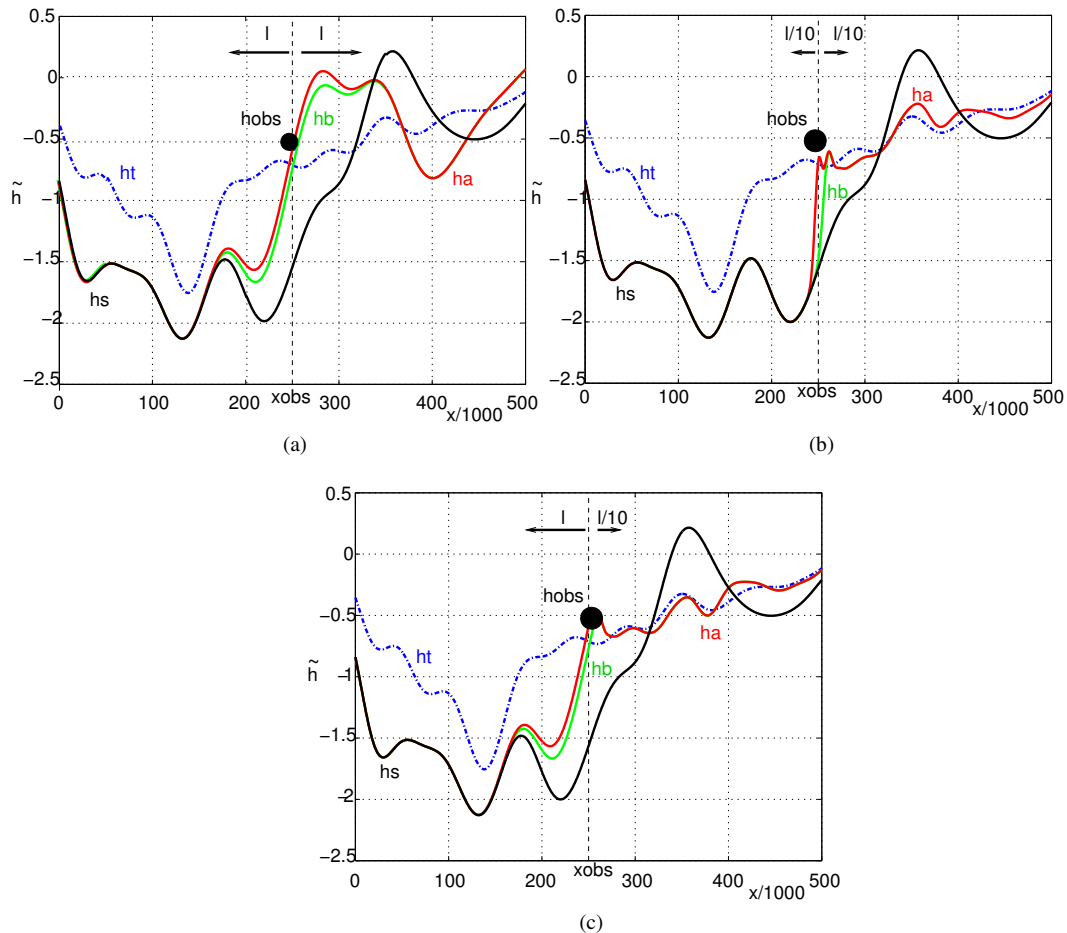[6] Service Hydrographique et Océanographique de la Marine

Fig. 5. Water level perturbation $\widetilde{h}$ 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)** $l_- = l_+ = l(x_{\mathrm{obs}})$, **(b)** $l_- = l_+ = l(x_{\mathrm{obs}})/10$, **(c)** $l_- = l(x_{\mathrm{obs}})$ and $l_+ = l(x_{\mathrm{obs}})/10$.

the Rognon is responsible for 50 % of the Marne discharge. A schematic description of the Marne Vallage catchment is shown in Fig. 7. The hydrological data at the upstream stations (Condes and Saucourt) are provided in real time by the Champagne-Ardenne DIREN. The downstream boundary condition in Chamouilley is described by a rating curve. Only the upstream stations were controlled by the data assimilation algorithm. Sensitivity tests revealed that the tangent linear model (Eq. 11) is valid for a pertubation up to 20 % in a, $6\,\mathrm{m}^3\,\mathrm{s}^{-1}$ in b and 6 h in c. The hourly tide gauge observations at Joinville and Chamouilley were used for the data assimilation procedure. The correlation lengths were set, using the procedure described in Sect. 3, to 51 km and 55 km at Joinville and Chamouilley, respectively.

In this study, the observed water level reached a maximum of several meters and the observation error standard deviations were set to 0.1 m. The observation error covariances were neglected, assuming that the observation stations are far enough apart for the spatial errors to be weakly correlated. The background error variances were chosen to be two

to three times larger than the observation error variances. At each observation point, only the observations above a minimum value were taken into account for the assimilation process in order to avoid representativeness errors. The observed values that were found to be too far from the simulated values were not assimilated; thus was accomplished by applying a threshold to the misfit between the observed and the simulated water levels.

The MASCARET model was chosen by the SCHAPI[7] to simulate the physical processes of the study catchments. A preliminary calibration procedure of several model parameters was performed by the SCHAPI and the SPC using data from twelve flood events of varying intensity. The geometry of the hydraulic network, the computation time step and the Strickler coefficient were adjusted so that the flood events were, on average, well represented at the observation stations. Globally, at Peyrehorade (Adour), the simulation tends to overestimate the flood peak for extreme

---

[7]Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations
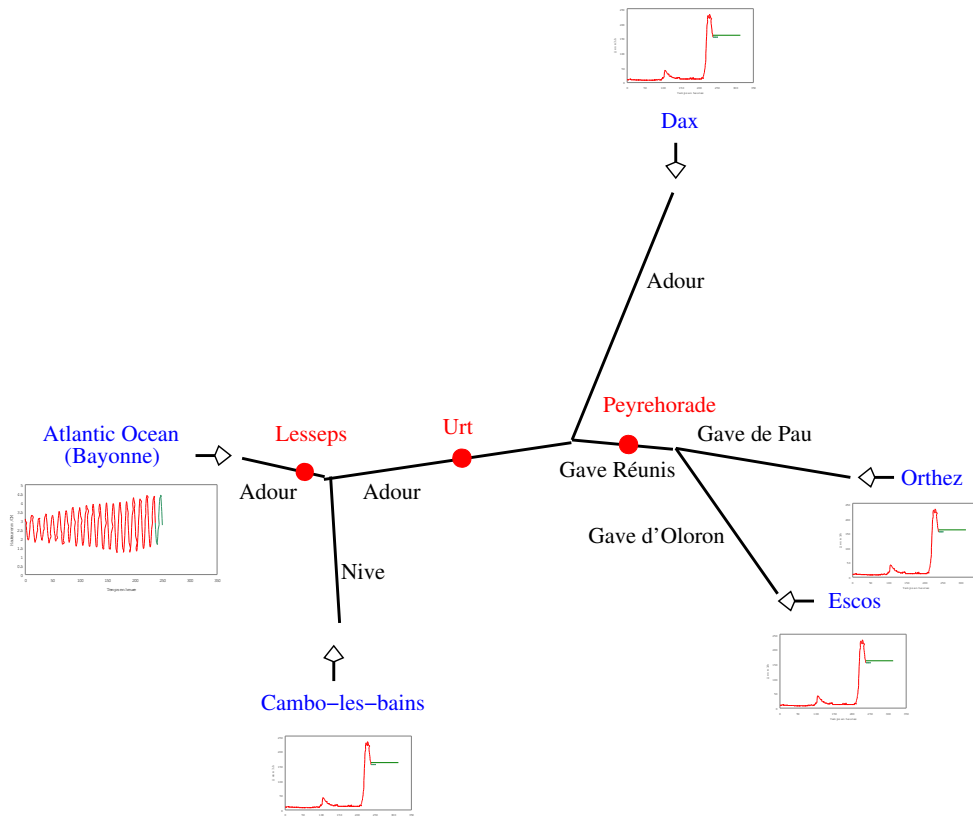
Fig. 6. The Adour catchment with the measurement stations in red and the upstream stations in blue.

flood events and underestimate the flood peak for moderate events. In the Marne catchment, a constant lateral inflow was ajusted so that the flood events were, on average, well represented at Joinville and Chamouilley. The simulated water levels at Joinville were globally correct, though often overestimated while the flood peak was often underestimated at Chamouilley.

4.1.2 Criteria for the interpretation

The baseline scenario was chosen as the simulation without assimilation (Free run). A post-treatment scenario for the Free run forecast was also explored: at the reference time T_r, the increment between the observation and the Free run is computed and added to the Free run water level. A correction that linearly decreases to zero over a 6-h forecast period is then applied. This post-treatment correction scenario (Interp run) allows for a perfect fit with the observation at T_r, for the water level at the observation locations. By construction, the Interp run merges with the Free run after 6 h of forecast.

The comparison of the Free run, Interp run and assimilation run (Assim run) with the observation was performed at the observation locations and times for the water level. The difference between the simulation and the observations is denoted by MmO (Model minus Observation) where the model run is either the Free run, the Interp run or the Assim run. MmO is computed at a given lead time in re-anaysis and forecast: 24 h before the reference time, and hourly, up to 12 h, after the reference time. The mean and standard deviation of MmO, respectively denoted by C1 and C2, are computed over the analysis cycles, at each observation station and for each flood event. The criteria C3 (in %) is defined as

$$\text{C3} = \frac{100}{N_{\text{obs}}} \sum_1^{N_{\text{obs}}} \left| \frac{h^{\text{M}} - h^{\text{obs}}}{h^{\text{obs}}} \right|, \tag{20}$$

where N_{obs} is the number of observations over a period of time and h^{M} is the simulated water level for either the Free run, the Interp run or the Assim run. C3 is a cumulative criteria as opposed to C1 and C2.

In practical terms, when the simulation is close to the observations, the criteria C1, C2 and C3 are small. C1, C2 and C3 were computed for the Free run, the Interp run and the Assim run. The percentage of reduction for each criteria was computed twice: (1) comparing the Free run and the Interp run and (2) comparing the Free run and the Assim run.
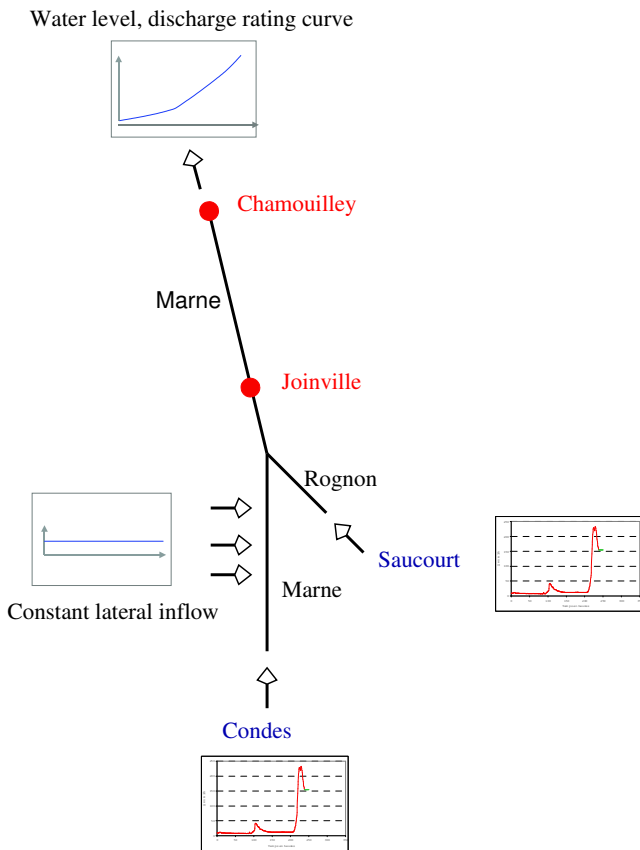
Fig. 7. The Marne Vallage catchment with the measurement stations in red and the upstream stations in blue.

4.2 Flood events simulation with data assimilation

4.2.1 Interpretation of the November 2002 event in the Adour catchment

Figure 8 shows the water level over a four-day period (Day 19 to Day 22 of a flood event starting 2 November 2011) at the observation station at Peyrehorade in the Adour catchment. The reference time for this cycle is $T_r = $ Day 22. The Free run integration of MASCARET 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 was applied to improve the water level over the first three days (re-analysis period) as well as over the forecast period (Day 22). The analysis with the instantaneous correction of the water level (green dashed curve) shows an excellent fit with the observations over the re-analysis period but leads only to a minor improvement over the forecast period. The model is constrained to the observed state by the hydraulic state correction procedure from Day 19 to Day 22. Though the analyzed state is almost equal 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 condition at the reference time $T_r = $ Day 22 is not enough to improve the simulation during the following day. The improvement is only significant over a couple of hours as the simulation is also degraded by other uncertainties. The analysis after the correction of the upstream flow (green curve) shows a good fit with the observation over the re-analysis period as well as over the forecast period. The difference with the observation only reaches 9 % of the observed value at Day 22.5. The upstream flow was 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 held equal to the last analysed value (which is better than the non-analysed one) allowing for an improvement in the water level simulation during Day 22. In summary, the water level hydraulic state correction procedure plays a major role in the re-analysis mode and the upstream flow correction plays a major role in forecast mode. The analysis after the two-step assimilation procedure is plotted in red and shows an improvement over the re-analysis period as well as during the forecast period. It should be noted that after a couple of hours of forecast, only the upstream flow correction is still effective as the green curve merges with the red curve.

For this event, the two-step analysis was cycled every hour (T_r varies from Day 18 to Day 26) so that, at an observation point, the water level is forecasted over the whole flood event. Fig. 9 shows the six-hour forecast for the Free run (black curve) and the Assim run (red curve) as well as the non assimilated observations (blue curve). For this lead time, the average C3 criteria for the flood event is 4.98 % for the Free run and 2.21 % for the Assim run, meaning that this criteria is improved by 55 % with the two-step data assimilation algorithm. For the Interp run, the average C3 criteria is equal to 3.5 % and is only improved by 29 % when compared to the Free run. At a six-hour forecast range, on average over the flood event, the assimilation procedure brings the simulation significantly closer to the observation. The accuracy of the assimilation procedure was therefore found to be better than the accuracy of the post-treatment procedure (Interp run).

Figure 10 displays the criteria C3 for the flood event at Peyrehorade, computed 24 h before the reference time (dashed curves) and 6 h after the reference time (solid curves), for the Free run (black curves) and the Assim run (red curves). It appears that during the re-analysis and the forecast period, the assimilation procedure brings the analysis closer to the observations than the Free run. As expected, C3 remains larger in forecast mode than in re-analysis mode because of the uncertainties in the boundary conditions at the controlled upstream stations (as well as other boundary conditions such as the maritime water level forcing as forecasted by tide models at SHOM).
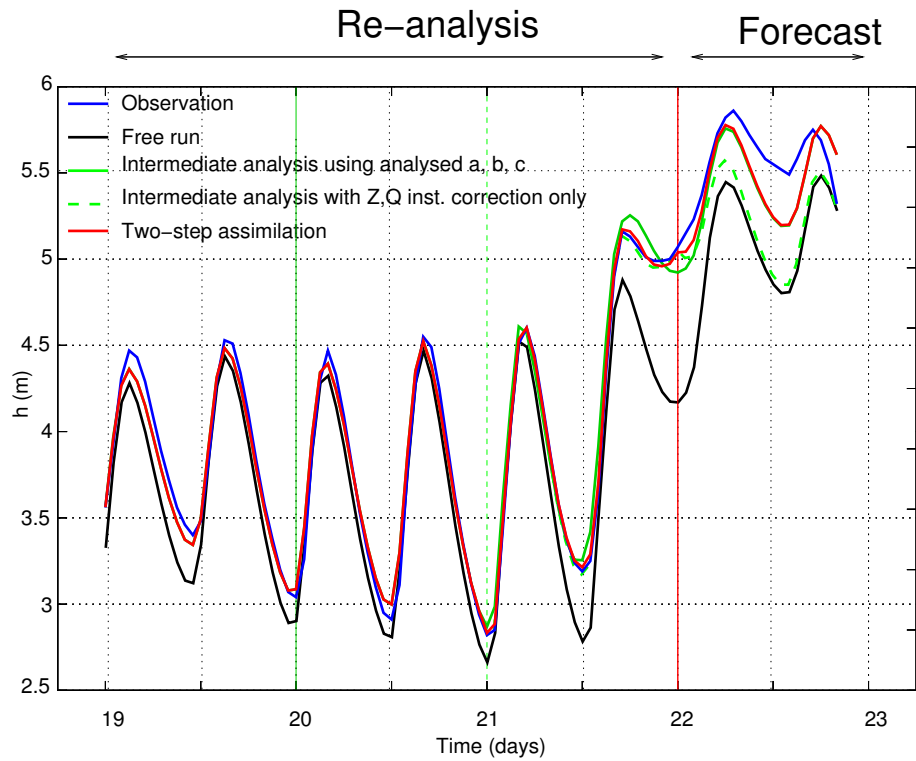
Fig. 8. November 2002 event, Adour catchment. Water level at Peyrehorade from Day 19 to Day 23 (T_r = Day 22): hourly observation in blue, Free run in black, analysis with the correction of the upstream flow only in green, analysis with the instantaneous Z, Q correction in dashed green and analysis with the two-step assimilation (Assim run) in red.
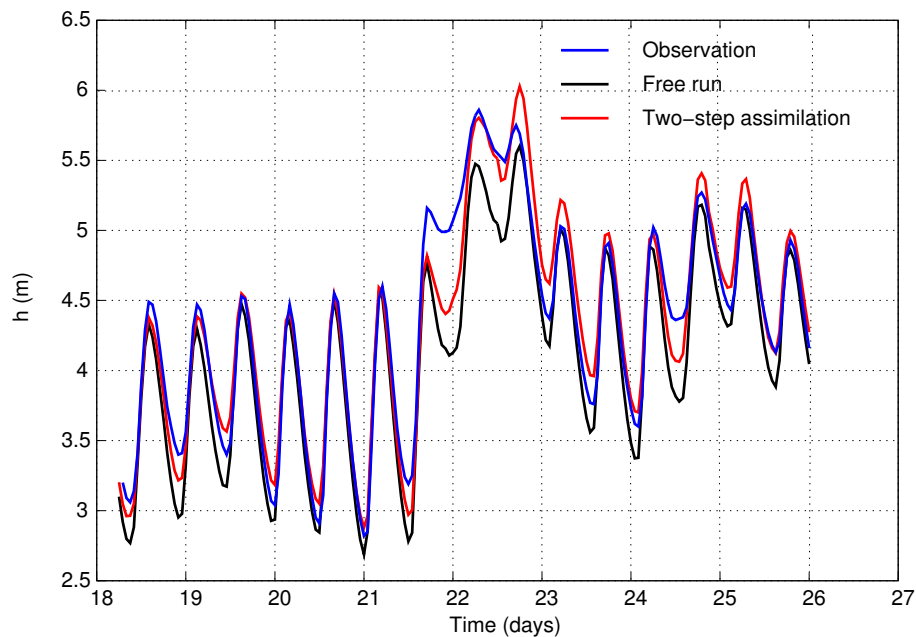

Fig. 9. November 2002 event, Adour catchment. Six-hour forecasted water level at Peyrehorade for the Free run (black curve), the observation (blue curve) and the analysis of the Assim run with the two-step assimilation (red curve).
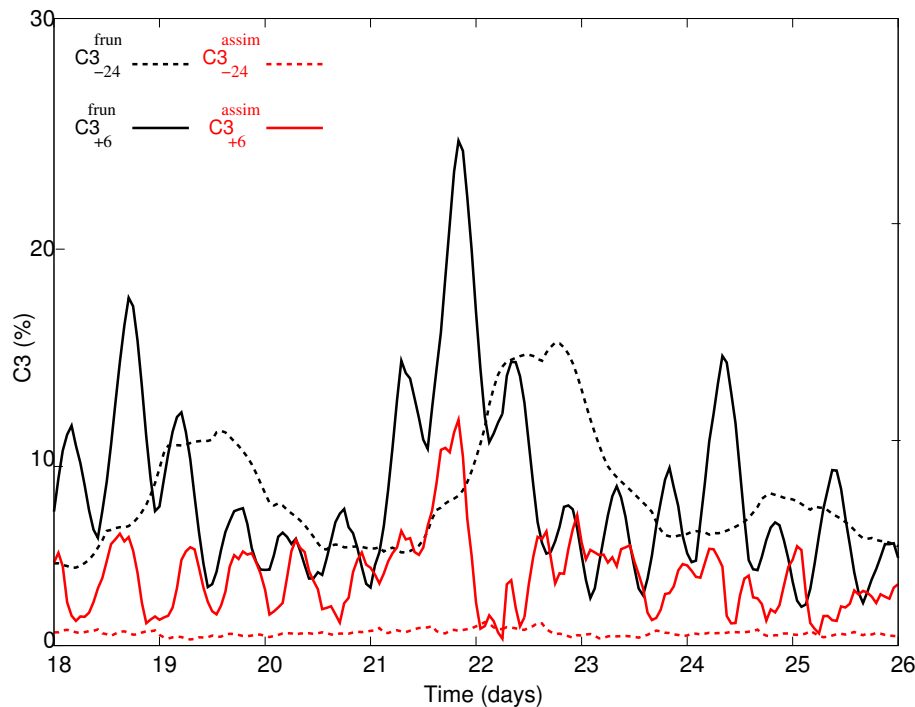
Fig. 10. November 2002 event, Adour catchment. C3 for the Free run (in black) and Assim run (in red) computed over the 24 h before T_r (dashed curves) and over the six hours after T_r (solid curves) at Peyrehorade.

4.2.2 Interpretation of the April 2006 event in the Marne catchment

The two-hour and the six-hour forecasted water levels in the Marne Vallage catchment, at Joinville, for the April 2006 event are presented in Figs. 11 and 12, respectively. The Free run significantly overestimates the flood peak at Joinville and Chamouilley (not shown) however the two-step data assimilation procedure allows for a good simulation of the peak at the two-hour forecast. The simulated peak at the six-hour forecast is also in better agreement with the observations than the Free run, though an overestimation of the peak remains. For the 2-h forecast, the average C3 (for the whole flood event) was improved by 35 % at Joinville and 32 % at Chamouilley. For the 6-h forecast, the average C3 was improved by 33 % at Joinville and 26 % at Chamouilley. During this flood event in forecast mode, the assimilation procedure brings the simulation significantly closer to the observation. It should be noted that in Figs. 11 and 12, the Assim run merges with the Free run when all the observations are under a minimim value as explained in Sect. 4.1.1.

Further analyses were carried out for flood events in the Marne Vallage catchment. Globally, the results were not as satisfying as in the Adour catchment. The main reason for this dampened performance is an incomplete calibration of the numerical model (namely for the Strickler coefficient and lateral input flow) in the Marne Vallage catchment before the assimilation procedure. It was shown during some events,

that in order to improve the simulation at one observation station, the data assimilation algorithm must degrade the simulation at the other observation location. The application of the two-step data assimilation procedure enabled the detection of a model incoherence in the Marne Vallage that could not be satisfactorily 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.

4.2.3 Statistical interpretation

Figure 13 shows that the forecasting ability of the model with data assimilation decreases with lead time. The mean reductions in criteria C1, C2 and C3 between the Free run and the Assim run (red curves) as well as between the Free run and the Interp run (black curves) were computed over seven flood events in the Adour catchment, at Peyrehorade. The dashed-dotted, dashed and solid curves represent the improvement in C1, C2 and C3 respectively. The two-step assimilation algorithm improves C1 by 72 % and C2 by 67 % at the reference time where as the Interp scenario improves C1 and C2 by 100 % as the simulated water level is literally corrected to the observed values at T_r. The improvement of C3 over the first hour of forecast with the two-step assimilation algorithm is 60 % and 70 % for the Interp scenario. Even though the Interp scenario gives the best results at short forecast range (up
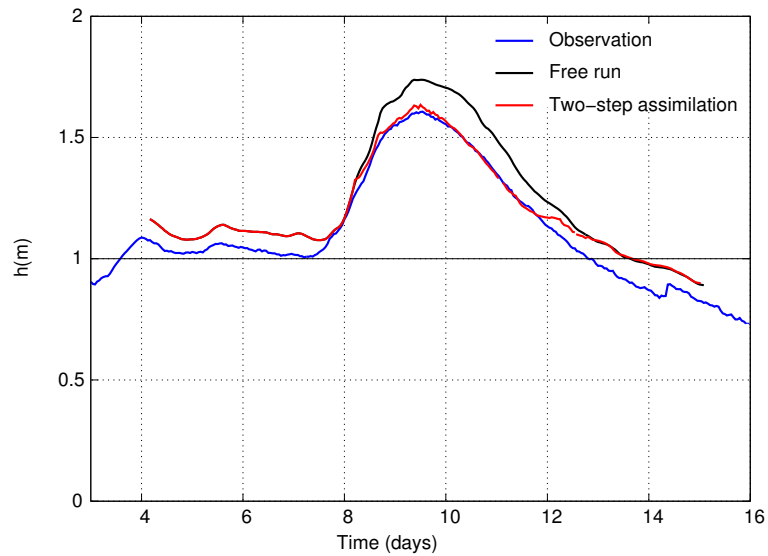
Fig. 11. April 2006 event, Marne Vallage catchment. Two-hour forecasted water level at Joinville for the Free run (black curve), the observation (blue curve) and the analysis of the Assim run with the two-step assimilation (red curve).
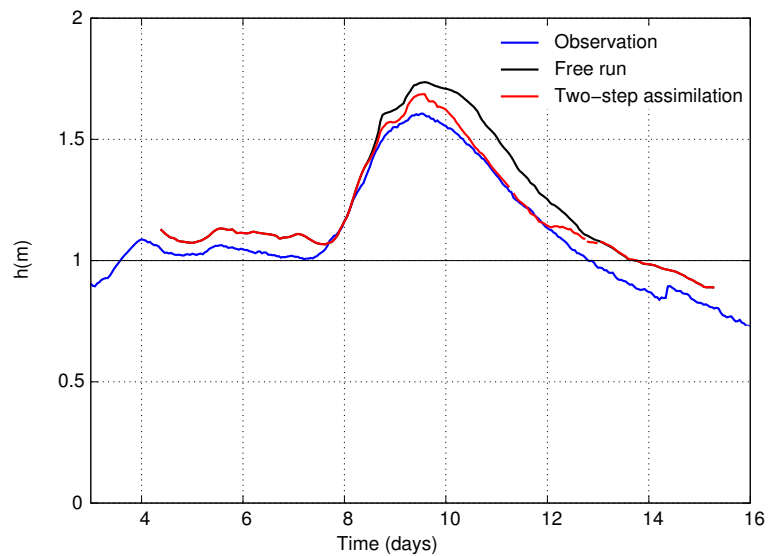

Fig. 12. April 2006 event, Marne Vallage catchment. Six-hour forecasted water level at Joinville for the Free run (black curve), the observation (blue curve) and the Assim run (red curve).

Table 1. C1, C2, C3 improvement between Free run and Assim run over the twenty four hours before the reference time. Average of seven flood events in the Adour catchment, at Peyrehorade, Urt and Lesseps.

−24 h re-analysis (Adour)	Peyrehorade	Urt	Lesseps
C1 improvement (%)	72	60	54
C2 improvement (%)	67	58	50
C3 improvement (%)	80	65	54

Table 2. C1, C2, C3 improvement between Free run and Assim run over the six hours after the reference time. Average of seven flood events in the Adour catchment, at Peyrehorade, Urt and Lesseps.

+6 h forecast (Adour)	Peyrehorade	Urt	Lesseps
C1 improvement (%)	15	15	0.5
C2 improvement (%)	13	11	0
C3 improvement (%)	36	25	3

to 3-h forecast), it should be kept in mind that this scenario only corrects the water level at observation locations and thus describes a hydraulic state that is not coherent with the physical equations. In this case, water level state is discontinuous in space and discharge values are not corrected according to water level values. Therefore, the resulting hydraulic state from the Interp run can not be used as an initial condition for a forecast integration of the MASCARET model. Additionally, the improvement of C1 and C2 for the Interp run is negligible for a lead time equal 6 h and above (by construction) whereas the improvement with the assimilation approach remains between 13 % (at +6 h) and 4 % (at +12 h). For a lead time equal to 3 h and above, the improvement in C3 is significantly larger with the two-step assimilation approach than with the post-treatment approach.

The average improvement of C1, C2 and C3 was computed for each observation station in the Adour catchment (over seven events) and in the Marne catchment (over four events) showing that the overall effect of the assimilation improves the description of the water level and discharge for all the observation stations.

The improvement in C1, C2 and C3 is displayed in Table 1 for the three Adour stations (Peyrehorade, Urt and Lesseps) for the 24 h before the reference time and, in Table 2, for the 6 h after the reference time. For the three stations, the criteria were improved, with the improvement being larger in re-analysis mode than in forecast mode. The evaluation of the criteria at Urt and Peyrehorade are similar as both stations are influenced by hydrological conditions. The Lesseps station which is closer to the maritime boundary, is significantly influenced by the tides that are not controled; as a consequence, the results of the assimilation are not as good at Lesseps as at Peyrehorade and Urt.

Table 3. C3 improvement between Free run and Assim run over the twenty four hours before the reference times and the six hours after the reference time. Average of four flood events in the Marne catchment, at Joinville and Chamouilley.

C3 improvement (%)	Chamouilley	Joinville
−24 h re-analysis	41	24
+6 h forecast	24	10

The improvement in C1 is shown in Table 3 for the two Marne stations (Joinville and Chamouilley) for the 24 h before the reference time and for the 6 h after the reference time. Again, for both stations, the criteria are improved and the improvement is larger in re-analysis mode than in forecast mode. Still, the improvement at the Marne stations is smaller than that described at the Adour stations, especially at Joinville. As previously stated, the calibration of the model parameters for the Marne catchment should be revised. For example, a lateral inflow could be added to represent additional inputs from the karstic drainage area. With a more physical model for the Marne catchment, the two-step data assimilation algorithm would then lead to better results.

5 Summary and conclusions

This paper presented the improvement in river flood forecasting when assimilating water level observations. The study was carried out with the one-dimensional hydraulic model MASCARET, on the Adour and Marne Vallage catchments. Representative events were presented for both catchments and statistics of the results were computed. The water level data were assimilated using a Kalman Filter algorithm to control the upstream flow and dynamically correct the hydraulic state. The first step of the analysis was based on the assumption that the upstream flow can be adjusted using a simple three-parameter correction. These three control parameters were adjusted over a two-day time window after one day of free run. The second step of the assimilation consisted of correcting the hydraulic state every hour (the observation frequency) during one day. The simulation was then integrated in forecast mode for an additional day. With this algorithm, the background error covariance matrix is not explicitly propagated by the dynamics of the system. Still, a particular effort was made to model background error covariance functions which were coherent with the dynamics of the hydraulic model. Anisotropic functions were used to represent the background error spatial correlations for the water level and the discharge, respectively. The justification for this choice was made by applying a full Kalman Filter algorithm on a diffusive flood wave propagation model. It was shown that the analysis turns a Gaussian correlation function into an anisotropic correlation function where the correlation length
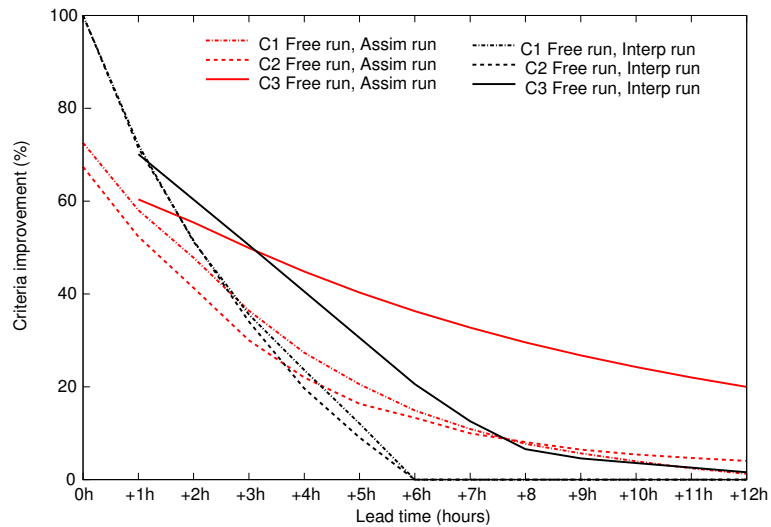
Hydrol. Earth Syst. Sci., 15, 3555–3575, 2011

www.hydrol-earth-syst-sci.net/15/3555/2011/

Fig. 13. Improvement in C1 (dashed-dotted),C2 (dashed) and C3 (solid) (in %) between Assim run and Free run (red curves) and between Interp run and Free run (black curves). Average of seven flood events, at Peyrehorade (Adour).
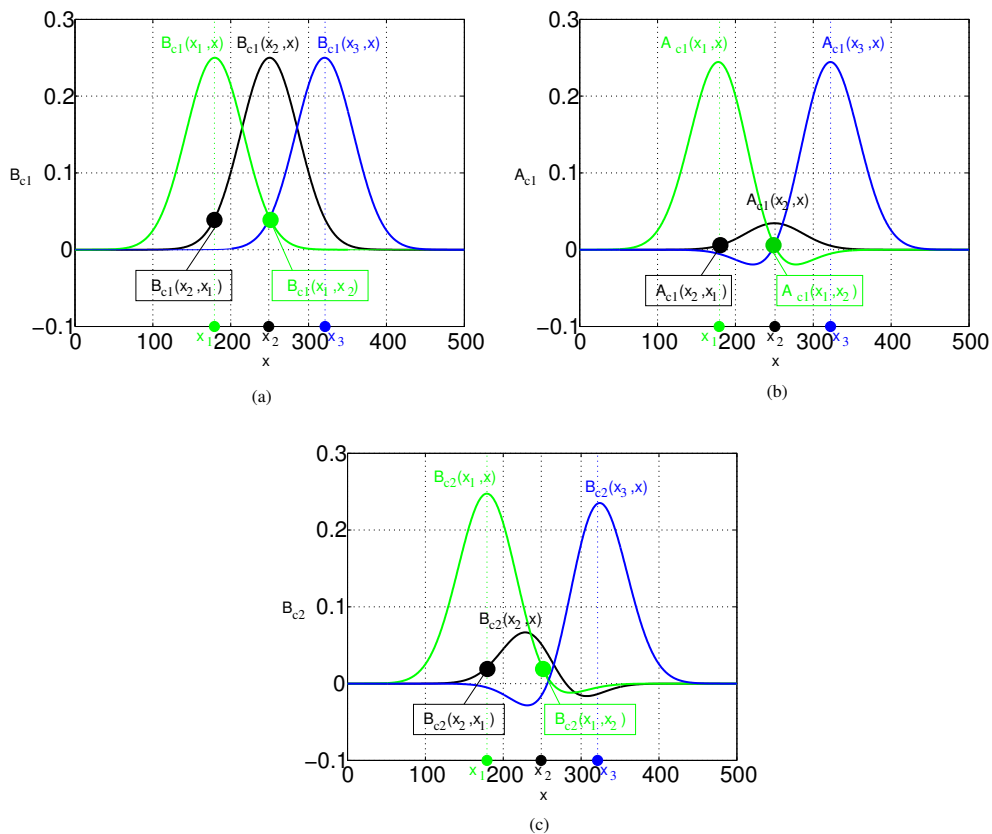

Fig. 14. (a) Initial background error covariances in $\mathbf{B}_{c1}$, **(b)** analysis error covariances in $\mathbf{A}_{c1}$ after one analysis and **(c)** updated covariances in $\mathbf{B}_{c2}$ for points $x_1 = 180$ (green), $x_2 = x_{obs} = 250$ (black) and $x_3 = 320$ (blue).
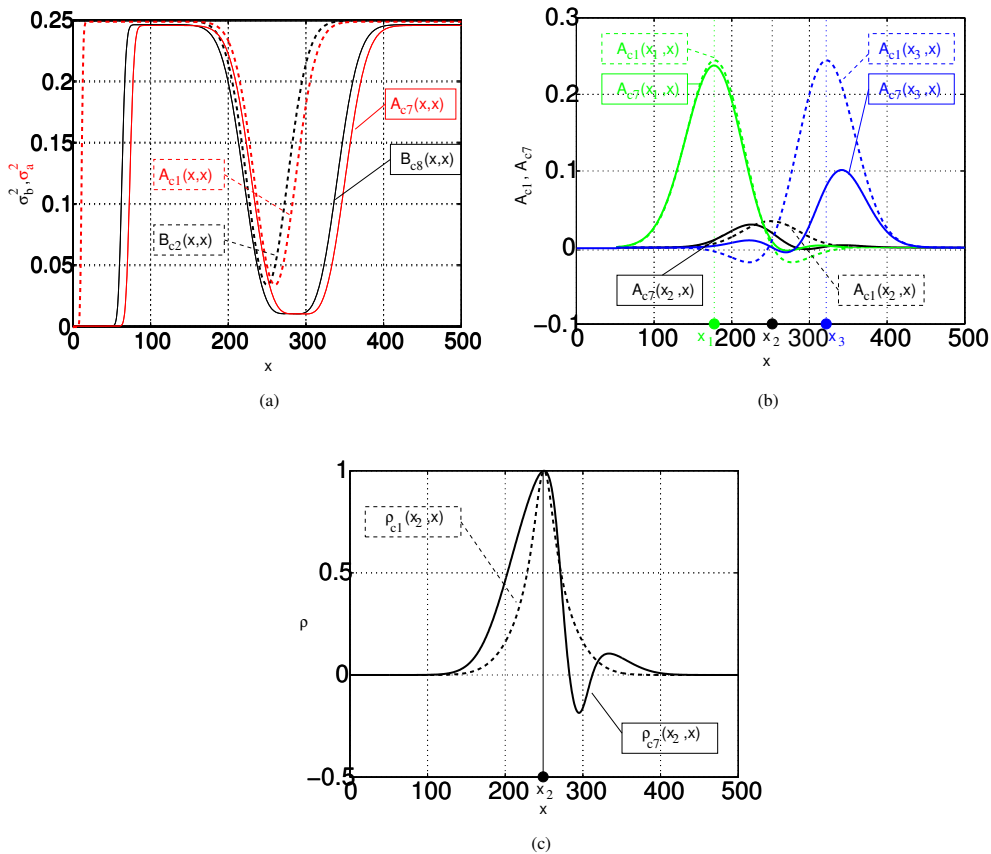
(a)

(b)

(c)

Fig. 15. (a) Error variance at each point for cycle one (dashed curves): analysis A_{c1} (red) and B_{c2} (black), and for cycle seven (solid curves): analysis A_{c7} (red) and B_{c8} (black). **(b)** Analysis error covariances in A_{c1} (dashed curves) and in 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 B_{c1} (dashed curve) and B_{c8} (solid curve) for $x_2 = x_{obs}$.

scale is shorter downstream of the observation point. This approach enabled a realistic modeling of the spatial error correlations for the data assimilation algorithm with MASCARET.

This procedure was applied on a four-day sliding window over the entire period of each flood event. It was shown that the simulation with data assimilation is significantly closer to the observation than the free run over the re-analysis period as well as over the forecast period. This conclusion is evidenced by the fact that the mean and standard deviation of the distance between the simulation and the observation at a given time (as well as the sum of this difference over a time period) are reduced by the data assimilation procedure. It was shown that the instantaneous correction of the hydraulic state leads to a significant improvement in re-analysis mode and for short forecast range. It was also shown that the sensitivity to an initial condition for the forecast mode is negligible compared to the sensitivity to the upstream flow, except at very short forecast range. For this reason, the upstream flow correction leads to a larger correction in forecast mode. On average, the correction of the hydraulic state is not as predictive as the upstream flow correction and is not sufficient to constrain the simulation over an interesting forecast

period. This justifies the need for the two-step data assimilation approach. This two-step procedure was applied to the Adour and Marne catchments (France) and the results were interpreted for several events in each catchment.

The assimilation procedure presented in this paper could potentially be applied to other catchment areas. Yet, 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 discharge errors should also be further investigated. This aspect would be enriched by the use of local (Z, Q) calibration functions. Additionally, the approximation of a state independent background error covariance matrix, was validated with the flood wave propagation model. Further work could be done to assess the integrity of this assumption using the Saint-venant equations model. Also, the impact of the observation frequency and the observation error statistics could be investigated: on going work suggests that when the observation frequency is low, the propagation of the covariance function is negligible and the covariance matrix remains state independent. Beyond this study, the extension of the control space could be useful as other sources of uncertainties or model errors (for

example the simplification of the flood plain representation) result in errors in the hydraulic state. For example, the correction of the Strickler coefficients or of the lateral discharge could be investigated.

Appendix A

Evolution of the local covariance functions with the Kalman Filter algorithm in a 1-D diffusive flood wave propagation model.

The Kalman Filter algorithm was implemented on a 1-D diffusive flood wave propagation model and the covariance matrices were updated following Eqs. (7–8).

The initial background covariance matrix[8] $\mathbf{B}_{c1}$ was modeled by spatially constant variances and correlation lengths 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 in Fig. 14a (x_2 is the observation point in this 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 and isotropic.

After one assimilation cycle, the error covariances are locally modified. The analysis error covariances in $\mathbf{A}_{c1}$ are computed from Eq. (6) and shown in Fig. 14b 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 isotropic. Conversely, at the upstream point x_1 and at the downstream point x_3, the covariance functions $\mathbf{A}_{c1}(x_1, x)$ and $\mathbf{A}_{c1}(x_3, x)$ are anisotropic. 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 introduced at this location by the analysis procedure through the innovation vector. The covariance functions at x_1 and x_3 are isotropic 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 in Fig. 14b.

The background error covariance matrix $\mathbf{B}_{c2} = \mathbf{M}^T \mathbf{A}_{c1} \mathbf{M}$ for the next (the second so far) assimilation cycle is computed from Eq. (8), 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 in Fig. 14c. The anisotropic covariances and correlation functions at all the upstream and downstream locations were propagated to the observation points so that the covariance and correlation functions become anisotropic 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 fundamentally different from those initially described in $\mathbf{B}_{c1}$ for the first assimilation, here $\mathbf{B}_{c1}(x_2, x)$. Still, the symmetric 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 in Fig. 14c). The initially described isotropic correlation function at the observation point x_2 has been modified into a local and anisotropic function. At the observation point x_2, the correlation function is anisotropic with a shorter correlation length downstream than upstream.

Figure 15a 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 $\sigma_b^2(x) = \mathbf{B}_{c2}(x, x)$ (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 vicinity of the observation point. The size of the area is controled by the correlation length initially described in $\mathbf{B}_{c1}$. The variances which were initially set to the spatially constant value 0.25 are now local, for example $\sigma_b^2(x_2) = 0.0497$ and $\sigma_b^2(x_1) = 0.2474$. Additionally, the variances in the updated $\mathbf{B}_{c2}$ (in black dashed curves) are also local, they correspond to the propagation of the variances in $\mathbf{A}_{c1}$ by $\mathbf{M}$ and its adjoint $\mathbf{M}^T$. 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 of the evolution of both the variances and the correlations. It appears that the correlation lengths tend to shorten downstream of the observation point. After several iterations (7 in this example) of the Kalman Filter, the variances are globally reduced downstream of the observation point as shown in Fig. 15a for $\mathbf{A}_{c7}$ and $\mathbf{B}_{c8}$. Effectively, the uncertainty at the observation point is reduced by the data assimilation algorithm, allowing the information to be 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 in Fig. 15b in solid curves. For comparison, the covariances from cycle 1 were also plotted in dashed curves in Fig. 15b. It is worth noting that at $x_3 = 320$, the amplitude of the variance has been divided by approximately two over the seven assimilation cycles. The shape of the covariance function evolves over time, especially downstream of the observation point. The correlation functions are clearly anisotropic with a shorter correlation length scale downstream than upstream of 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 in Fig. 15c. The correlation length is divided by approximately five in this plot along the 7 assimilation cycles of the Kalman Filter. This factor doesn't vary significantly when the Kalman Filter is further iterated.

---

[8]The subscript ci denotes the number of the assimilation cycles

Appendix B

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

The objective here was to determine the correlation length of the spatial correlation function for the errors in the water level and the discharge errors produced with MASCARET. This determination was two-fold. First a diffusion coefficient κ, based on the dynamics of the diffusive flood wave approximation model (Eq. 16), was graphically estimated by studying the propagation of a perturbation of the hydraulic state. Then, this diffusion coefficient was used to calculate the spatial correlation length of the state perturbation covariance function.

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

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

that sums up to a Heaviside function if $l_0 \to 0$, was added at the upstream flows $\widetilde{q}(t)$. This perturbation is propagated by Eq. (16) towards the observation points where, at time t, the state is described by $\widetilde{h}(x,t)$ given by:

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

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 $\widetilde{h}(x,t)$ of MASCARET as follows:

- $c \approx \frac{s_{\mathrm{r}}}{t_{\mathrm{r}}}$ 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 for 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 close to the conditions described in this case.

Using these three relations when the perturbation reaches the observation point, κ is estimated by $\frac{T^2 s_{\mathrm{r}}^2}{2t_{\mathrm{r}}^3}$ where s_{r} is known when as T and t_{r} are graphically estimated for the simulated discharges.

Like the initial and boundary conditions, the temporal covariance function $R(\tau)$ of $\widetilde{q}(t)$ is also propagated by the diffusive flood wave propagation equation Eq. (16). Since the temporal covariance function $R(\tau)$ is Gaussian and because the theory of random function diffusion applies, the spatial covariance function of $\widetilde{h}(x,t)$ can be approximated as Gaussian. Assuming that the spatial covariance function for the boundary condition is chosen to be 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 processes. Since advection process no effect on the covariance function, this formula can be written in the advected referential. In this referential, $\widetilde{\mathbf{B}}_t = \mathbf{L}_t \mathbf{L}_t^T = \mathbf{L}_{2t}$ where $\mathbf{L}_t$ represents the diffusion operator. 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 in the domain. For the IKF 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 an application with MASCARET, a realistic upstream flow is described from which l_0^2 could be determined or when the observation point is far enough from the upstream stations $l_0^2 \ll 4\kappa \frac{x_{\mathrm{obs}}}{c}$.

This graphical approach leads to the estimation of a local correlation length $l^2(x)$ at each observation point based on the perturbation of the upstream flow at one upstream station. Since there are several upstream stations for this study, there are several resulting signals $\widetilde{h}(x,t)$ reaching the observation point, leading to several estimations of $l(x_{\mathrm{obs}})$. At the observation point, the spatial correlation function is approximated as Gaussian, resulting from the sum of Gaussian functions of respective correlation length l_i and amplitude a_i. The correlation length of the resulting Gaussian can be approximated by:

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

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).

Edited by: J. Freer

References

Atanov, G., Evseeva, E., and Meselhe, E.: Estimation of roughness profile in trapezoidal open channels, J. Hydraul. Eng., 125, 309–311, 1999.

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.

Bouttier, F. and Courtier, P.: Data assimilation concepts and methods, ECMWF Lecture Note, 1999.

Buis, S., Piacentini, A., and Déclat, D.: PALM: A Computational framework for assembling high performance computing applications, Concurrency Computat.: Pract. Exper., 18, 247–262, 2006.

Chow, V. T.: Open-channel hydraulics, McGraw-Hill, 1959.

Das, A.: Parameter estimation for flow in open-channel networks, J. Irrig. Drain. E., 130, 160–165, 2004.

Durand M., Andreadis K. M. , Alsdorf D. E, Lettenmaier D. P., Moller D., and Wilson M.: Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model, Geophys. Res. Lett., 35, L20401, doi:10.1029/2008GL034150, 2008.

Fouilloux, A. and Piacentini, A. :The PALM Project: MPMD Paradigm for an Oceanic Data Assimilation Software, Lect. Notes Comp. Sci., 1685, 1423–1430, 1999.

Geese, A. F., Sellier M. Van Houten E., and Smart G.: Reconstruction of river bed topography from free surface data using a direct numerical approach in one-dimensional shallow water flow, Inverse Probl., 27, 1–12, 2011.

Gelb, A.: Applied optimal estimation, Cambridge Mass.: MIT Press, 1974.

Special Issue on the Revolution in Global Ocean Forecasting. GO-DAE: 10 Years of Achievement, Oceanography Society, 22, 96–109, 2009.

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, 1–19, 2002.

Hervouet, J.-M.: Hydrodynamique des écoulements à surface libre – Modélisation numérique avec la méthode des éléments finis, Ponts et Chaussées (Presses), 2003.

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.

Ide, K., Courtier, P., Ghil, P., and Lorenc, A. C.: Unified notation for data assimilation: operational, sequential and variational, J. Meteorol. Soc. JPN, 75, 181–189, 1997.

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.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003.

Lagarde, T.: Nouvelle approche des méthodes d'assimilation de données: les algorithmes de point selle, Ph.D. thesis, Université Paul Sabatier, 246 pp., 2000.

Lagarde, T., Piacentini, A., and Thual, O.: A new representation of data assimilation methods: the PALM flow charting approach, Q. J. Roy. Meteorol. Soc., 127, 189–207, 2001.

Maclaughlin, D. and Townley, L.: A reassessment of the groundwater inverse problem, Water Resour. Res., 32, 1131–1161, 1996.

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.

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, J. Hydrol., 302, 46–69, 2005.

Parrish, D. F. and Derber, J. C.: The national meteorological center's spectral statistical interpolation analysis system, Mon. Weather Rev., 120, 1747–1763, 1992.

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, Q. J. Roy. Meteorol. Soc., 126, 1143–1170, 2000.

Sau, J., Malaterre, P.-O., and Baume, J.-P.: Sequential MonteCarlo hydraulic state estimation of an irrigation canal, CR Acad. Sci., 338, 212–219, 2010.

Talagrand, 0.: Assimilation of observations, an introduction, J. Meteorol. Soc. JPN., 75, 191–209, 1997.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM: Society for Industrial and Applied Mathematics, 1987.

Todling, R. and Cohn, S. E.: Suboptimal schemes for atmospheric data assimilation based on the Kalman Filter, Mon. Weather Rev., 122, 2530–2557, 1994.

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, 2010a.

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, 2010b.