



Enhanced flood forecasting through ensemble data assimilation and joint state-parameter estimation

Item Type	Article
Authors	Ziliani, Matteo G.;Ghostine, Rabih;Ait-El-Fquih, Boujemaa;McCabe, Matthew;Hoteit, Ibrahim
Citation	Ziliani, M. G., Ghostine, R., Ait-El-Fquih, B., McCabe, M. F., & Hoteit, I. (2019). Enhanced flood forecasting through ensemble data assimilation and joint state-parameter estimation. Journal of Hydrology, 577, 123924. doi:10.1016/j.jhydrol.2019.123924
Eprint version	Post-print
DOI	10.1016/j.jhydrol.2019.123924
Publisher	Elsevier BV
Journal	Journal of Hydrology
Rights	NOTICE: this is the author's version of a work that was accepted for publication in Journal of Hydrology. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Journal of Hydrology, [[Volume], [Issue], (2019-07-12)] DOI: 10.1016/j.jhydrol.2019.123924 . © 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2024-03-13 07:03:38
Link to Item	http://hdl.handle.net/10754/656186

Enhanced flood forecasting through ensemble data assimilation and joint state-parameter estimation

Matteo G. Ziliani^a, Rabih Ghostine^b, Boujemaa Ait-El-Fquih^a, Matthew F. McCabe^{a,c},
Ibrahim Hoteit^{a,b,*}

^a*Physical Science and Engineering (PSE), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

^b*Computer, Electrical and Mathematical Science and Engineering Division (CEMSE), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

^c*Biological and Environmental Science and Engineering (BESE), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

Abstract

Accurate water level forecasts during flood events are crucial to mitigate the loss of human lives and economic damages. However, the accuracy of flood models can be affected by various factors, including the complexity of the terrain geometry and bathymetry, imperfect physics as well as uncertainties in the inflows and parameters. This paper describes a practical implementation of an ensemble Kalman filter (EnKF) based data assimilation system that is aimed towards enhancing the forecasting skill of flood models. The system was implemented and tested with a real world dam break flood, based on the experimentally scaled Toce River valley flood that occurred on July 8th, 1996. Water depth data are available for assimilation from a network of 21 sensors distributed across the domain. Our results demonstrate that assimilating data into the flood model significantly improves the model prediction by up to 90% after assimilation and 60% during forecasting. Assimilating the data more frequently significantly enhances the system performances. Estimating the two-dimensional Manning coefficient together with the model's dynamical variables (water depth and velocities) further improves the model prediction skill. Overall, our results suggest that assimilating data into the flood model, while jointly inferring the

*Corresponding author

Email address: Ibrahim.Hoteit@kaust.edu.sa (Ibrahim Hoteit)

URL: <http://assimilation.kaust.edu.sa/> (Ibrahim Hoteit)

state and (poorly known) parameters, using an EnKF may provide an efficient framework for developing an operational flood forecasting system.

Keywords: flood modeling, shallow water equations, data assimilation, ensemble Kalman filter, state-parameter estimation.

1. Introduction

With continuous growth in global population, the density of human settlements occupying areas that are vulnerable to flooding and other natural hazards is increasing. As a consequence, the human and structural cost of natural disasters has grown considerably over time. Flooding is the most common natural hazard that affects humans and properties around the world ([Few, 2003](#)), often causing colossal human and economic losses. Anticipating floods before they occur could help minimize losses through the implementation of appropriate protection, provision and rescue plans. Because of the complexity of the problem, and the number of different situations that need to be considered, the most suitable approach for generic flood prediction is often based on numerical simulations that require solution of the shallow water equations (SWE) ([Constantin and Escher, 1998](#)).

Numerical flood models based on the two-dimensional SWE are widely used to study flood events, or analyze the effects of floods in control projects. Despite many significant improvements over the last decades, flood forecasting models are still far from being perfect. They suffer from scale-related problems, with some of their parameters not being directly measurable ([Blöschl and Sivapalan, 1995](#)). These models are also subject to various sources of uncertainties. Apart from the uncertainty due to randomness in the natural process that cannot be avoided, there are also unresolved dynamical processes and uncertainties in inputs (e.g. bathymetry and boundary conditions). One way to mitigate for these uncertainties that has proven to be efficient in meteorological, oceanographic and other hydrological applications ([Hoteit et al., 2008](#); [Lermusiaux, 2006](#); [Liu and Gupta, 2007](#)), is to constrain flood forecasting models with available data. Measurements of

flood, including from satellite and in-situ measurements of water levels, are now more commonly available in near- to real-time (Zhang et al., 2011; García-Pintado et al., 2013) than ever before. These data can be used operationally to improve the model behavior and reduce uncertainties in prediction (Ghil, 1989; Holland and Malanotte-Rizzoli, 1989; Edwards et al., 2015). This process, referred to generically as data assimilation (DA), sequentially nudges the model predictions towards incoming observations to keep the model as close as possible to the real trajectory of the simulated system (Hoteit et al., 2002).

The Kalman Filter (KF) is a popular data assimilation method that was designed for linear systems with Gaussian uncertainties (Kalman et al., 1960). It recursively computes the best (minimum-variance) linear estimate of the state of the system and its error covariance matrix based on available observation. To deal with the non-linearities and large-dimensions of geophysical models, ensemble Kalman filters (EnKFs) were introduced as an alternative, providing an effective Monte Carlo framework to propagate the error covariance matrix forward in time while avoiding the need for the complex model linearization step (Evensen, 2003). The main idea behind the EnKF is to perform ensemble simulations of model runs to compute statistical estimates of the state and error covariance. Several studies have demonstrated that the EnKF presents the advantage of being portable and straightforward to implement as well as being remarkably robust and performant, even when implemented with small ensembles (Anderson and Anderson, 1999; Hamill and Snyder, 2000; Hoteit et al., 2015).

The EnKF operates in cycles of two steps: (i) a forecast step in which the ensemble members are integrated with the model to the time of the next available observation, and (ii) an analysis step to update the forecasted members with the data based on the KF analysis step, which provides the analysis ensemble from which to start the next assimilation cycle (Hoteit et al., 2002; Evensen, 2003). Importantly, this ensemble framework is flexible enough to account for model errors through random perturbations of the ensemble

members ([Hoteit et al., 2008](#)).

The EnKF is basically formulated under a linear Gaussian framework, which is actually not necessarily realizable in practice, and as such the convergence of this filter (to the true posterior distribution) is not guaranteed ([Ait-El-Fquih et al., 2016](#)). The performances of EnKFs are further often affected by the approximative nature of their ensemble covariance matrices, as they are usually implemented with limited ensembles and poorly known model errors ([Song et al., 2010](#)). This can lead to the filter inbreeding problem, with the ensemble variance being increasingly under-estimated over time ([Hendricks Franssen and Kinzelbach, 2008](#)). Covariance inflation ([Anderson and Anderson, 1999](#)) and localization ([Hamill et al., 2001](#)) are popular techniques to deal with these problems. A more straightforward approach to enhance the filter performances is to further quantify and reduce the uncertainties in the model's poorly known parameters ([Elsheikh et al., 2013](#); [Gharamti and Hoteit, 2014](#); [Gharamti et al., 2015](#); [Ait-El-Fquih et al., 2016](#)). This is known as the state-parameter filtering problem and is often tackled using the so-called Joint-EnKF ([Chen and Zhang, 2006](#); [Hendricks Franssen and Kinzelbach, 2008](#); [Gharamti et al., 2015](#)). The Joint-EnKF concurrently estimates the state of the system and the model's unknown parameters by concatenating both into one state vector to be estimated.

Modeling of flood events are usually considered as boundary value problem, with boundary conditions often described by a time changing water level (hydrograph), guiding the overflow across the domain ([Tirupathi et al., 2016](#)). The ability to update the model state through time with real time data is changing our vision of flood prediction and the way we forecast the flood evolution in time and space. In the last decade, several applications of data assimilation for flood forecasting have been presented. For instance, [Tossavainen et al. \(2008\)](#) simulated the river flow using a 2D finite-volume shallow water model, assimilating GPS equipped drifter data deployed as passive lagrangian tracers using an EnKF. [Strub et al. \(2009\)](#) latter compared EnKF and an offline optimization technique for forecasting river hydraulics. Spatially distributed water level data have also

77 been extracted from satellite images and incorporated into flood models using variational
78 DA methods (4D-VAR), demonstrating the relevance of the method to improve the consis-
79 tency between the model and the data (e.g. [Lai and Monnier 2009](#); [Hostache et al. 2010](#)).
80 Despite providing high spatial coverage, the application of satellite data to flood modeling
81 problems suffers from poor resolution and large errors ([Gichamo et al., 2012](#); [Wu et al.,](#)
82 [2014](#)).

83 Despite the widespread applications of flood modeling and forecasting methods in
84 hydrology, very few studies considered the joint state parameter estimation problem.
85 [Wilson et al. \(2010\)](#) implemented the EnKF to estimate the rear-shore bathymetry with the
86 SHORECIRC hydrodynamic model. [Wilson and Özkan-Haller \(2012\)](#) and [Landon et al.](#)
87 [\(2014\)](#) respectively assimilated depth-average and drifted-based velocity measurements
88 to estimate the bathymetry in the Kootenai River. [Mayo et al. \(2014\)](#) implemented a deter-
89 ministic EnKF to estimate the Manning roughness value in a simple configuration of the
90 ADCIRC ocean circulation model and more recently, [Siripatana et al. \(2017\)](#) demonstrated
91 the relevance of parameters estimation with the EnKF against a full Markov chain Monte
92 Carlo (MCMC) inversion. The aforementioned studies considered the parameter estima-
93 tion problem in only one dimension. Here we investigate the efficiency of (i) the EnKF
94 for estimating time-varying water height and flow directions, and (ii) the Joint-EnKF for
95 simultaneously estimating the two-dimensional Manning’s n coefficients, which define
96 the bottom roughness in a 2-D SWE model, together with the flow height and directions.
97 A detailed evaluation of the filters’ performances is reported based on realistic experiment
98 reproducing the flooding event of the Toce river valley system (Italy) that occurred on July
99 8th, 1996 ([Prestininzi, 2008](#)) and assimilating real measurements that were collected as
100 part of this experiment. We assess the sensitivity of the system to the ensemble size and
101 the frequency of the observations, with particular focus on the relevance of the Joint-EnKF
102 for enhancing the prediction skill of the flood modeling system.

103 The paper is organized as follows. Section 2 describes the flood model. The Joint-

EnKF is presented in Section 3. Section 4 provides an overview of the experimental setup and Section 5 presents the assimilation results. A summary of the main findings and conclusions is given in Section 6.

2. The Flood model

2.1. Governing equations

The model solves the Shallow Water Equations (SWE):

$$\frac{\partial U}{\partial t} + \nabla F(U) = S, \quad (1)$$

where, t is the time, U and $F(U) = (E(U), G(U))$ respectively denote the vectors of conserved flow variables and the flux, and S is the source vector. Neglecting Coriolis effects, kinetic and turbulence viscous terms and wind stresses, S only includes the bed slope source S_0 and friction source S_f . These vectors are expressed as

$$U = [h, q_x, q_y]^T, \quad (2)$$

$$E = [q_x, uq_x + gh^2/2, uq_y]^T, \quad (3)$$

$$G = [q_y, vq_x, vq_y + gh^2/2], \quad (4)$$

and

$$S = S_0 + S_f = (0, -gh \frac{\partial z}{\partial x}, -gh \frac{\partial z}{\partial y})^T + (0, -gh \frac{n^2 u \sqrt{u^2 + v^2}}{h^{4/3}}, -gh \frac{n^2 v \sqrt{u^2 + v^2}}{h^{4/3}})^T, \quad (5)$$

where h is the water depth, u and v are the velocities and q_x and q_y are the unit-width discharges in x - and y -directions (q_x equals to hu and q_y equals to hv), respectively, g is the gravitational acceleration, z is the bottom elevation, and n is the Manning roughness coefficient. A variety of numerical techniques have been proposed for solving the SWE

model (Eq. (1)), including finite element (Strang and Fix, 1973) and finite volume methods (Eymard et al., 2000), etc. In this work, we implement the finite volume method on an unstructured triangular grid using triangular domain discretisation.

2.2. Numerical implementation

Finite volume framework

Within an arbitrary cell, the definite integral form of Eq. (1) is

$$\int_{\Omega} \frac{\partial U}{\partial t} d\Omega + \int_{\Omega} \nabla F(U) d\Omega = \int_{\Omega} S d\Omega. \quad (6)$$

Using the divergence theorem, Eq. (6) can be rewritten as

$$\int_{\Omega} \frac{\partial U}{\partial t} d\Omega + \oint_{\Gamma} F(U) \cdot m d\Gamma = \int_{\Omega} (S_o + S_f) d\Omega, \quad (7)$$

where Ω and Γ denote the control volume of the system and the boundary of the volume, which in two-dimension is the control area and edges of the cell. m is the outward unit vector normal to the prescribed boundary with the components of $(m_x$ and $m_y)$. $F(U) \cdot m$ is the flux vector normal to the boundary. In Eq. (7), the first term involving time differentiation is always computed with a finite difference method, so after a time interval of ∂t , U at each cell is updated to the next time step using

$$U^{t+1} = U^t + \frac{\Delta t}{\Omega} \left(\int_{\Omega} (S_o + S_f) d\Omega - \oint_{\Gamma} F(U) \cdot m d\Gamma \right). \quad (8)$$

The integral of $F(U) \cdot m$ around the cell is evaluated with an explicit Euler method as

$$\oint_{\Gamma} F(U) \cdot m d\Gamma = \sum_{j=1}^{m_b} F_j^*(U_L, U_R) \cdot m_j l_j, \quad (9)$$

in which, j and l are the index and area (length for 2D) of the faces of the cell, respectively, and m_b is the number of faces of the cell. U_L and U_R are the reconstructed values of U

on the right and left sides of the face, respectively. To implement a stable second-order accurate scheme, the MUSCL linear interpolation approach (Van Leer, 1979) is applied to reconstruct U_L and U_R . The Harten, Lax and van Leer approximate Riemann solver with the contact wave restored (HLLC) (Toro, 2009) is employed to calculate the flux $F_j^*(U_L, U_R) \cdot m_j$ at the j th face.

The slope source term S_0 is transformed into a sum of fluxes through all faces of the cell as proposed by (Audusse et al., 2004). Finally, the friction terms are solved by a semi-implicit scheme (Begnudelli and Sanders, 2006), in order to avoid numerical instabilities in case of occurrence of rapid changes in flow depth or velocity. More details about the scheme and its validation can be found in Hou et al. (2015).

3. State and parameters estimation with the EnKF

The flood data assimilation problem can be described by a state-space model of the form:

$$\mathbf{x}_k = M(\mathbf{x}_{k-1}, \theta_{k-1}) + \eta_{k-1}, \quad (10)$$

$$\mathbf{y}_k = \mathbf{H}_k(\mathbf{x}_k) + \varepsilon_k, \quad (11)$$

where $\mathbf{x}_k \in \mathbb{R}^{N_x}$ denotes the system state (of dimension N_x) composed of the water depth and water velocity in x and y directions evaluated at every point of the domain, $\mathbf{y}_k \in \mathbb{R}^{N_y}$ the flood observations at time t_k (of dimension N_y), and $\theta \in \mathbb{R}^{N_\theta}$ the parameter vector (of dimension N_θ). M is a nonlinear operator integrating the system state from time t_k to t_{k+1} using the SWE, and \mathbf{H}_k the observational operator relating the state to the observations (the bold font reflects the linear form of the observational operator in the present study). The model process noise, $\eta = \{\eta_k\}_{k \in \mathbb{N}}$, and the observation process noise, $\varepsilon = \{\varepsilon_k\}_{k \in \mathbb{N}}$, are assumed to be statistically independent, jointly independent and independent from \mathbf{x}_0 and θ . Both η_k and ε_k are assumed to follow Gaussian distributions with zero mean and

covariances \mathbf{Q}_k and \mathbf{R}_k , respectively. We emphasize that the EnKF scheme is still valid in the perfect model case (i.e., $\mathbf{Q}_k = 0$), which is actually usually considered in the literature.

The joint state-parameter filtering problem consists of estimating the augmented vector, $\mathbf{z}_k \stackrel{\text{def}}{=} [\mathbf{x}_k^T, \theta_k^T]^T$, at each time t_k , given the previous observations $\mathbf{y}_{0:k} \stackrel{\text{def}}{=} \{\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$. One common approach to address this problem is to minimize the mean square error (MSE), which leads to the posterior mean (PM) (Ait-El-Fquih and Hoteit, 2016):

$$\mathbb{E}_{p(\mathbf{x}_k|\mathbf{y}_{0:k})}[\mathbf{x}_k] = \int \mathbf{x}_k p(\mathbf{x}_k, \theta | \mathbf{y}_{0:k}) d\mathbf{x}_k d\theta, \quad (12)$$

$$\mathbb{E}_{p(\theta|\mathbf{y}_{0:k})}[\theta] = \int \theta p(\mathbf{x}_k, \theta | \mathbf{y}_{0:k}) d\mathbf{x}_k d\theta. \quad (13)$$

Because of the nonlinear character of the system, an analytical computation of Eqs. (12) and (13) is not feasible. The Joint-EnKF was proposed as an efficient method to compute approximate solutions for Eqs. (12) and (13) with a reasonable computational cost (Ait-El-Fquih et al., 2016). The state-parameter estimation problem is first converted into a classical state-space system on which the EnKF is directly applied as follows

$$\begin{cases} \mathbf{z}_k = \tilde{M}(\mathbf{z}_{k-1}) + \tilde{\eta}_{k-1}, \\ \mathbf{y}_k = \tilde{\mathbf{H}}_k \mathbf{z}_k + \varepsilon_k, \end{cases} \quad (14)$$

where $\tilde{M}(\mathbf{z}_{k-1}) = \begin{pmatrix} M(\mathbf{x}_{k-1}) \\ \theta_{k-1} \end{pmatrix}$, $\tilde{\eta}_{k-1} = \begin{pmatrix} \eta_{k-1} \\ \mathbf{0} \end{pmatrix}$ and $\tilde{\mathbf{H}}_k = [\mathbf{H}_k \quad \mathbf{0}]$, where $\mathbf{0}$ is a zero matrix with appropriate dimensions. Starting from an analysis ensemble of N_e members at time t_{k-1} , $\{\mathbf{x}_{k-1}^{a,i}, \theta_{k-1}^{a,i}\}_{i=1}^{N_e}$, the EnKF forecast and analysis steps can be summarized as follows.

Forecast step

Starting from an analysis ensemble and assuming static dynamics for the parameter vector, i.e. $\theta_k^{f,i} = \theta_{k-1}^{a,i}$, the state analysis members, $\mathbf{x}_{k-1}^{a,i}$, are integrated with (10) to compute the

175 state forecast members

$$\mathbf{x}_k^{f,(i)} = M(\mathbf{x}_{k-1}^{a,(i)}, \theta_{k-1}^{a,(i)}) + \eta_{k-1}^{(i)}; \quad \eta_{k-1}^{(i)} \sim N(\mathbf{0}, \mathbf{Q}_{k-1}). \quad (15)$$

176 The state and parameter forecasts are then the sample mean of the $x_k^{f,i}$, and $\theta_k^{f,i}$, respec-
177 tively, and the associated error covariance matrices are their sample covariances.

178 *Analysis step*

179 When a new observation becomes available, the $\mathbf{x}_k^{f,i}$ and $\theta_k^{f,i}$ are updated following the
180 KF analysis step:

$$\mathbf{x}_k^{a,i} = \mathbf{x}_k^{f,i} + \mathbf{P}_{x_k^f, y_k^f} \mathbf{P}_{y_k^f}^{-1} (\mathbf{y}_k - \mathbf{y}_k^{f,i}), \quad (16)$$

181

$$\theta_k^{a,i} = \theta_k^{f,i} + \mathbf{P}_{\theta_k^f, y_k^f} \mathbf{P}_{y_k^f}^{-1} (\mathbf{y}_k - \mathbf{y}_k^{f,i}), \quad (17)$$

182 where $\mathbf{y}_k^{f,i} = \mathbf{H}_k \mathbf{x}_k^{f,i} + \varepsilon_k^i$; $\varepsilon_k^i \sim N(\mathbf{0}, \mathbf{R}_k)$. The (cross-)covariance matrices are evaluated
183 from the ensembles as

$$\mathbf{P}_{x_k^f, y_k^f} = (N_e - 1)^{-1} \mathbf{W}_{x_k^f} \mathbf{W}_{y_k^f}^T, \quad (18)$$

184

$$\mathbf{P}_{y_k^f} = (N_e - 1)^{-1} \mathbf{W}_{y_k^f} \mathbf{W}_{y_k^f}^T, \quad (19)$$

185

$$\mathbf{P}_{\theta_{k-1}^f, y_k^f} = (N_e - 1)^{-1} \mathbf{W}_{\theta_{k-1}^f} \mathbf{W}_{y_k^f}^T, \quad (20)$$

186 where for a given ensemble $\{\mathbf{u}^i\}_{i=1}^{N_e}$, $\mathbf{W}_{\mathbf{u}}$ denotes the perturbation matrix (Evensen, 2003)
187 (eq. 49-51). The state and parameter analysis and their error covariance matrices are
188 then the analysis ensembles mean and covariances $\mathbf{P}_{x_k^a} = (N_e - 1)^{-1} \mathbf{W}_{x_k^a} \mathbf{W}_{x_k^a}^T$ and $\mathbf{P}_{\theta_k^a} =$
189 $(N_e - 1)^{-1} \mathbf{W}_{\theta_k^a} \mathbf{W}_{\theta_k^a}^T$, respectively. The schematic representation of an assimilation cycle of
190 the (Joint)-EnKF is depicted in Figure 1.

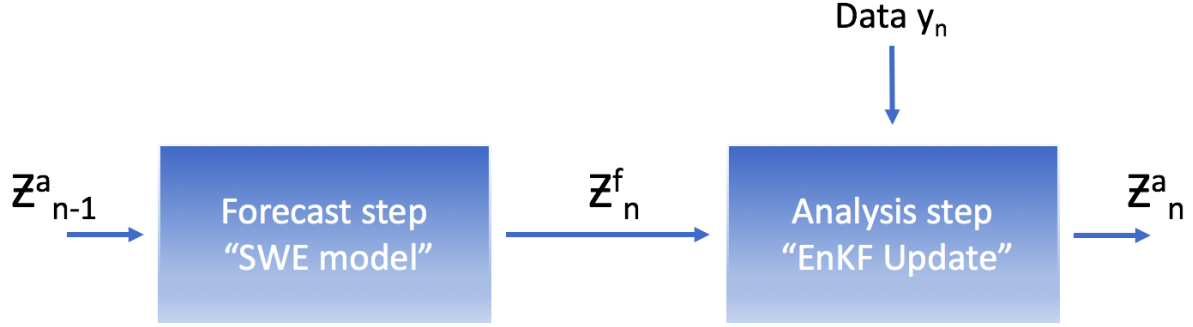


Figure 1: A schematic illustration of an assimilation cycle of the (Joint)-EnKF. State vector Z could be X = (water depth, velocities) for EnKF or (X, θ) = Manning coefficient for Joint-EnKF.

Covariance inflation (Anderson and Anderson, 1999) and localization (Hamill et al., 2001) techniques are implemented to mitigate the limitations of the EnKF with small ensembles and neglected model errors (Song et al., 2010) after every forecast step. The sample forecast error covariance matrix is inflated by a prescribed factor $\lambda^2 \equiv (1 + \delta)^2$ for a positive scalar δ (Anderson and Anderson, 1999; Hoteit et al., 2003). Localization is used in order to alleviate the problems of rank deficiency and spuriously long range correlations between distant state variables (Hamill et al., 2001). We adopt the standard local analysis (LA) technique where the state space is divided into a set of disjoint local analysis domains, and the system state in a local domain is updated using only nearby observations within a preset distance.

4. Experimental Setup

4.1. Model domain and configuration

We consider a real-scale wave propagation event, driven by torrential and fluvial flooding along an experimentally scaled river valley. An experimental model of a river located in the Occidental Alps (Italy) within a narrow flood plain was built at the National Institution for Electrical Energy (ENEL) in Milano, Italy, under the European Union CADAM (Concerted Action on Dam Break Modelling) project to investigate flow methods and their use for simulation and prediction of dam failures effects (Soares-Frazão and Testa, 1999).

A 5 km length of the Toce river valley was scaled down 100 times to reduce the dimension to approximately 50×11 m. The model is made of concrete and covers nearly the whole breadth of the valley by including the small and nearly not visible thalweg (line connecting the lowest points of successive cross-sections along the course of the river valley). Based on the model's basement, the Manning-Strickler coefficient has been set experimentally to $0.0162 \text{ s}/\sqrt[3]{m}$ (Soares-Frazão and Testa, 1999), as recommended by specialists at ENEL. Furthermore, the Digital Terrain Model (DTM) of the valley is scaled down to a resolution of 0.0025 m^2 covering the 2D domain (Fig. 3).

The center of the valley is characterized by the presence of a large reservoir that expands across the entirety of the section and is bordered by levees, with an aperture at the river side that is always kept closed during the experiment. A pump regulates the level of the tank that is connected with the upstream boundary of the model. As soon as the water level exceeds the bottom of the initial section of the valley that is reproduced by the model, the water is released and a monitoring system estimates the discharge. The downstream end is modeled as a free fall into a tank. Critical flow occurs at the outlet, so that its influence on the upper part should be minimal (Prestininzi, 2008).

The flood event, that has been replicated, simulates a dam break occurrence where inundation overtopped the reservoir levees. The domain is initially dry (water level = 0), and the measured inflow hydrograph is reported in Fig. 2. A set of 21 water probes scattered over the domain (as shown in Fig. 3) were employed to retrieve the time evolution of the water depth. These are then assimilated in an attempt to accurately predict the flood arrival time and water level.

The SWE flood model described in Section 2 has been executed on an unstructured mesh consisting of $N_c = 10,185$ triangular cells with a resolution of 0.3 m^2 , as shown in Fig. 4. Different meshes' resolution have been tested (although not reported in the manuscript) in order to assess the performances of the SWE model and filters. In our simulations, the choice of using a course resolution mesh consisting of $N_c = 10,185$ triangular cells was a

236 good compromise between accuracy and processing time. As also mentioned in [Caleffi](#)
 237 [et al. \(2003\)](#), the use of higher resolution meshes to improve the model's forecasting skills
 238 did not lead to significant improvements in the results, but instead dramatically increased
 239 the computational cost of the simulation process.

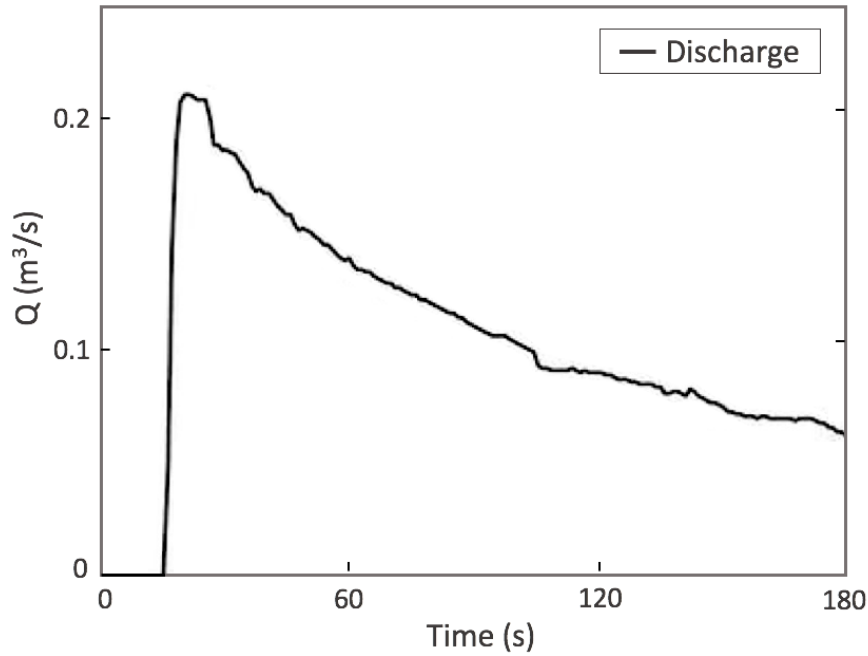


Figure 2: Hydrograph of the flash flood measured during the experiment.

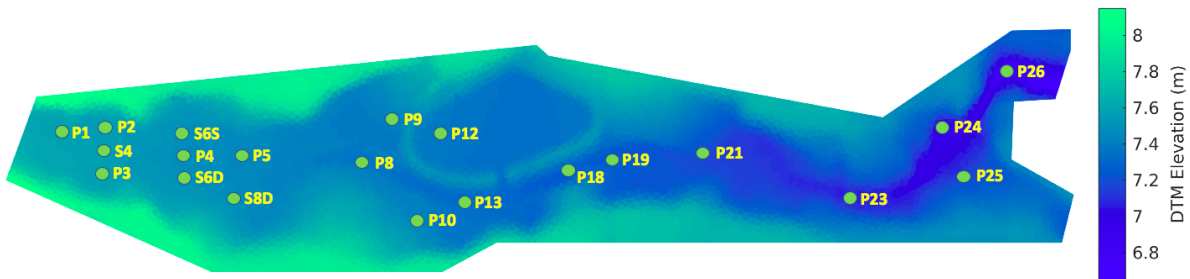


Figure 3: Digital Terrain Model (DTM) of the river valley showing the decreasing elevation from left to right side of the domain. 21 gauges are spread along the path of the river, providing water level measurements throughout the duration of the simulation.

240 The valley is very steep, with a mean slope of 2%, while some areas are characterized
 241 by abrupt drops. Due to the different slopes and shapes changes in the valley, the water
 242 flow continuously passes through critical condition. Fig. 5 depicts the propagation of the

inundation through the valley at instants 0, 20, 40 and 60 seconds, as simulated by the SWE model. Several images of the actual inundation can be also found in [Soares-Frazão and Testa \(1999\)](#) to visualize the intensity of the event. The time step is set to 10^{-4} s and the total duration of the experimental simulation is 180 s. The experiments were conducted on a 12 cores, 2.67 GHz and 48Gb RAM Dell workstation, setting the output interval to 1 s; memory requirement does not exceed 40 MB. No boundary condition needed to be applied at the lateral boundary, as the flow never reaches it.

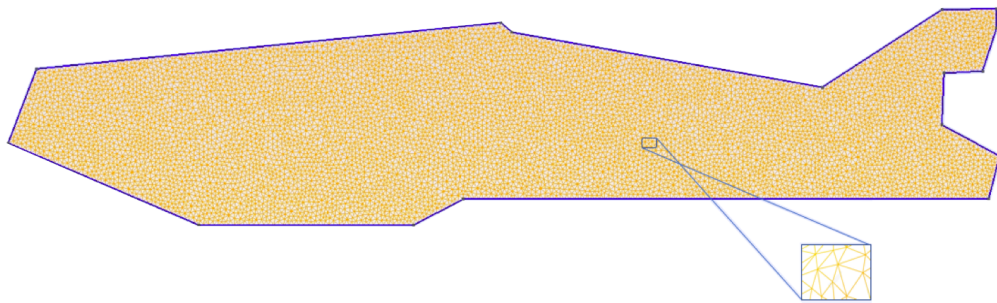


Figure 4: Unstructured mesh made by 10185 triangular cells used for the numerical simulation.

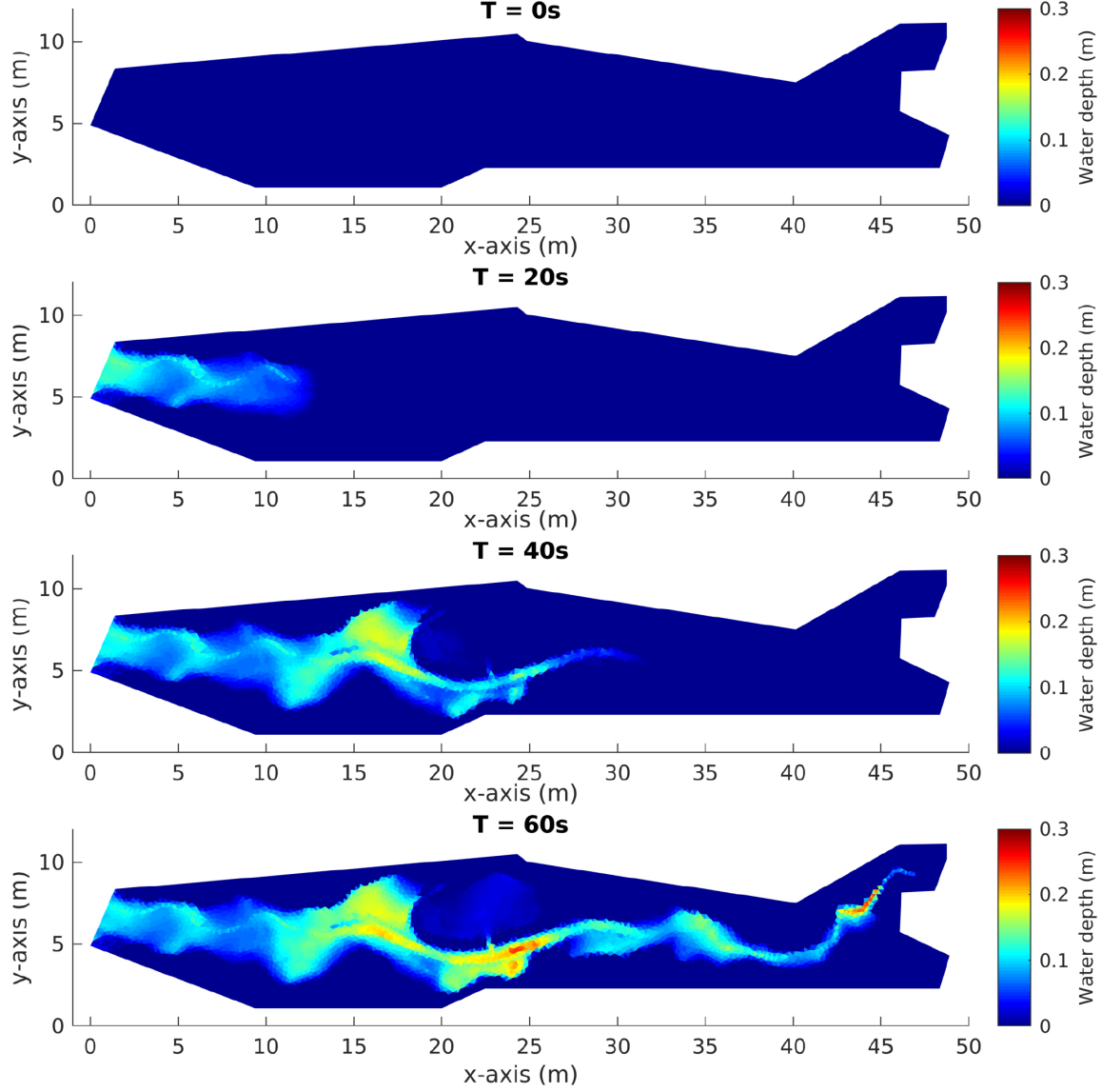


Figure 5: Initial water depth maps at initial time (dry bed, top), after 20 seconds (second from top), after 40 seconds (third from top) and after 60 seconds of simulation (bottom) using the SWE model.

4.2. Joint-EnKF implementation and assimilation experiments

The state vector is composed of $4N_c$ elements and concatenates the water depth h , the water velocities q_x and q_y in x and y directions, and the Manning value n_i at each cell of the domain. The observation vector \mathbf{y}_k is created with the measurements of water depth recorded at the 21 water probes, available every second. The measurement Gaussian errors are assumed unbiased, spatially uncorrelated, and of variance $\sigma^2 = 10^{-4} \text{ m}^2$ (i.e.

256 $R = \sigma^2 Id$).

257 The initial 2D Manning's field ensemble is generated using the public domain ANSI-C
258 code 'GCOSIM3D', developed by [Gómez-Hernández and Journel \(1993\)](#) and based on
259 the sequential Gaussian simulation algorithm. To generate the initial realizations, the
260 covariance has been set with zero mean and standard deviation equal to 20% of the
261 prescribed Manning value. The variance is properly scaled for the realizations of 2D
262 multi-Gaussian fields to fall within an appropriate range of Manning's n values (0.01 -
263 0.025). The maximum and minimum Manning's values of these realizations are 0.0187
264 and 0.0121, respectively. Examples of generated realizations of Manning's n fields are
265 shown in Fig. 6, from which the initial ensemble was selected.

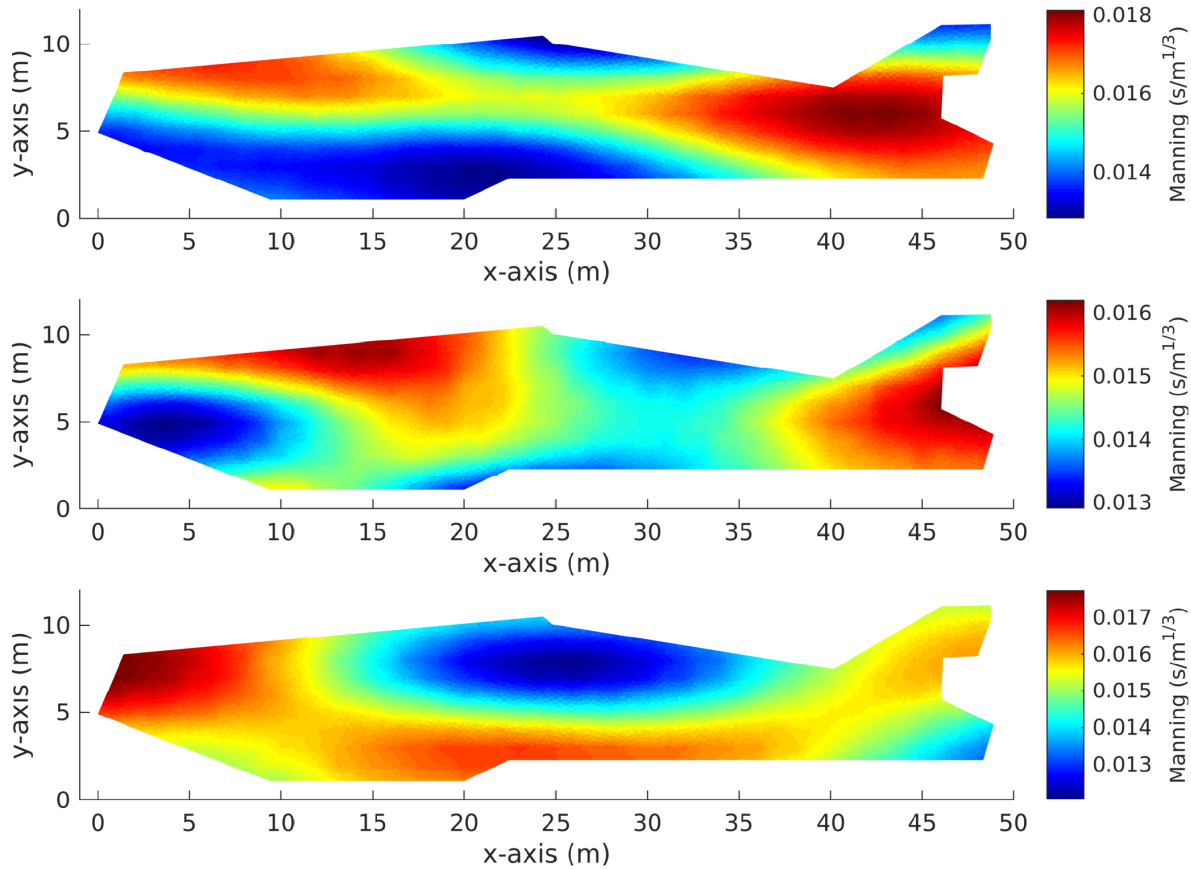


Figure 6: Three different realizations of the Manning parameter obtained using the Gaussian simulation toolbox ([Gómez-Hernández and Journel, 1993](#)).

To assess the relevance of the EnKF for enhancing the forecasting skill of the SWE model, we evaluate the filter performance and robustness under different experimental scenarios. We implement the EnKF assuming: (i) the Manning coefficient to be constant equal to $0.0162 \text{ s}/\sqrt[3]{m}$, which means that the EnKF state vector only includes water depth and velocity, and (ii) the Manning coefficient is also unknown and try to estimate its value at every point of the model mesh; we refer to this case as the Joint-EnKF. In both cases we consider a perfect model scenario ($\mathbf{Q}_k = 0$). We then conduct a number of sensitivity experiments changing the ensemble size and the assimilation period, and evaluate the performances of the filters with different combinations of localization and inflation values. The localization length is scaled according to the real river valley dimensions (5 km length). We select these scenarios to assess the performances of the standard EnKF (case (i)) and Joint-EnKF (case (ii)) under more challenging and imperfect modeling and observational conditions to demonstrate their capabilities to enhance the SWE model forecasting skills. The performance of the filters is evaluated based on the root-mean-square-errors (RMSE) of the water level, resulting from the forecast and analysis state estimates, as

$$RMSE_k = \sqrt{\frac{1}{N_y} \sum_{i=1}^{N_y} (y_k^i - \hat{y}_k^i)^2}, \quad (21)$$

where y_k^i is the i^{th} element of the observation at time k , \hat{y}_k^i its corresponding forecast (or analysis) and N_y is the total number of observations. RMSE's improvements, in terms of percentage, are also analyzed for all the gauges to evaluate the filters' results. We also examine the spatial patterns of the estimated Manning's field coefficients, and the ensemble spread as a measure of the uncertainties about the filters' estimates.

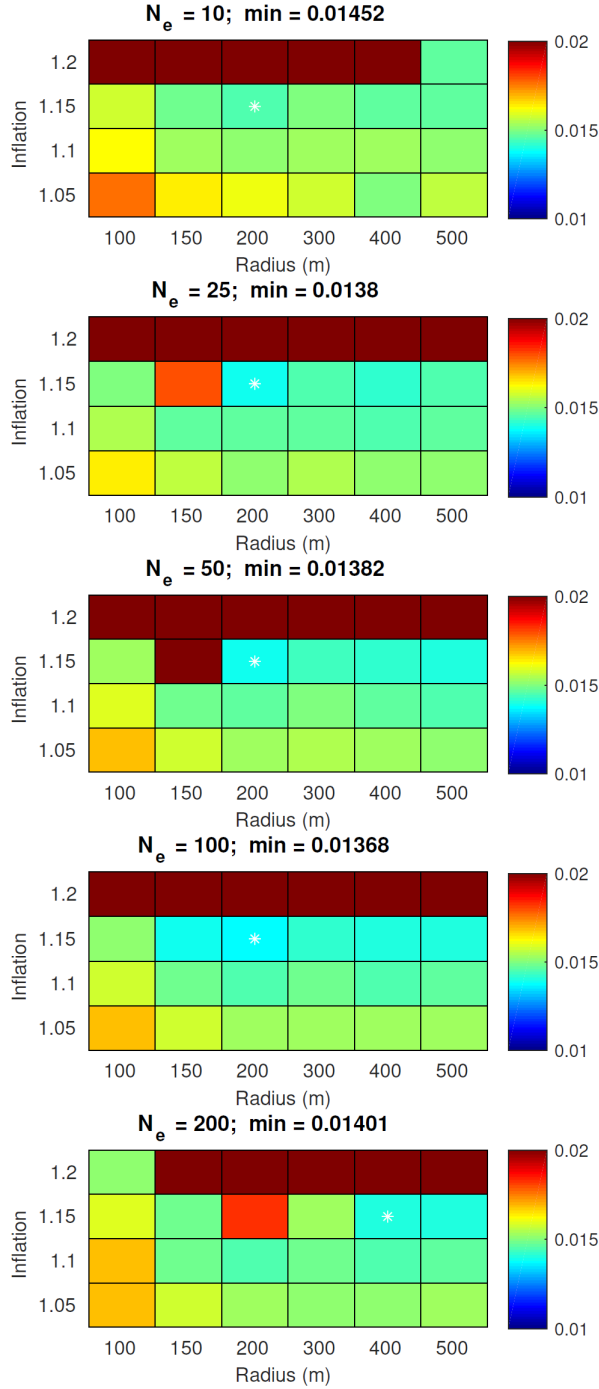
5. Assimilation results

5.1. Sensitivity to filtering parameters

We first study the sensitivity of the EnKF and Joint-EnKF to the ensemble size, N_e . In realistic and large-scale hydraulic applications, one would be restricted by computational resources to using small ensembles. We conduct the experiments testing five ensemble sizes, $N_e = 10, 25, 50, 100$ and 200 . Figs. 7 and 8 plot the RMSE of the water depth, as given in (21), averaged over the assimilation period, for the forecast and analysis as resulting from the EnKF and the Joint-EnKF, respectively. Assimilation results from different combinations of ensemble size, localization and inflation are analyzed. In particular, we consider the solution with covariance localization length scales (relative to the real field site) of 100, 150, 200, 300, 400 and 500 meters and inflation values of 1.05, 1.1, 1.15 and 1.2. The overall accuracy of the water depth estimates resulting from the EnKF and Joint-EnKF increases with increased ensemble size. EnKF and Joint-EnKF forecasts (Fig. 7) exhibit different behaviors: while the EnKF does not show a significant improvement in terms of RMSE with larger N_e (basically similar minimum RMSE values for different N_e), the Joint-EnKF estimates keep improving with increasing ensemble size. In particular, the RMSE of the EnKF solution quickly ceases to improve beyond 25 members only. This is consistent with the results of Fig. 9, which shows the RMSE of the water depth for the EnKF and Joint-EnKF, as they result from the best combination of ensemble size - localization - inflation in term of average RMSE.

Even though the EnKF solution appears to be less sensitive to the choice of the filter's parameters (Fig. 7) than the Joint-EnKF, a careful analysis of the filter solutions demonstrates the ability of the Joint-EnKF to provide better estimates (analysis and forecast) as compared with the observations. Simultaneously estimating the state and the Manning's parameters with the Joint-EnKF reduces the uncertainties in the model forecast and significantly enhances the overall performance of the assimilation system.

EnKF Forecast



Joint-EnKF Forecast

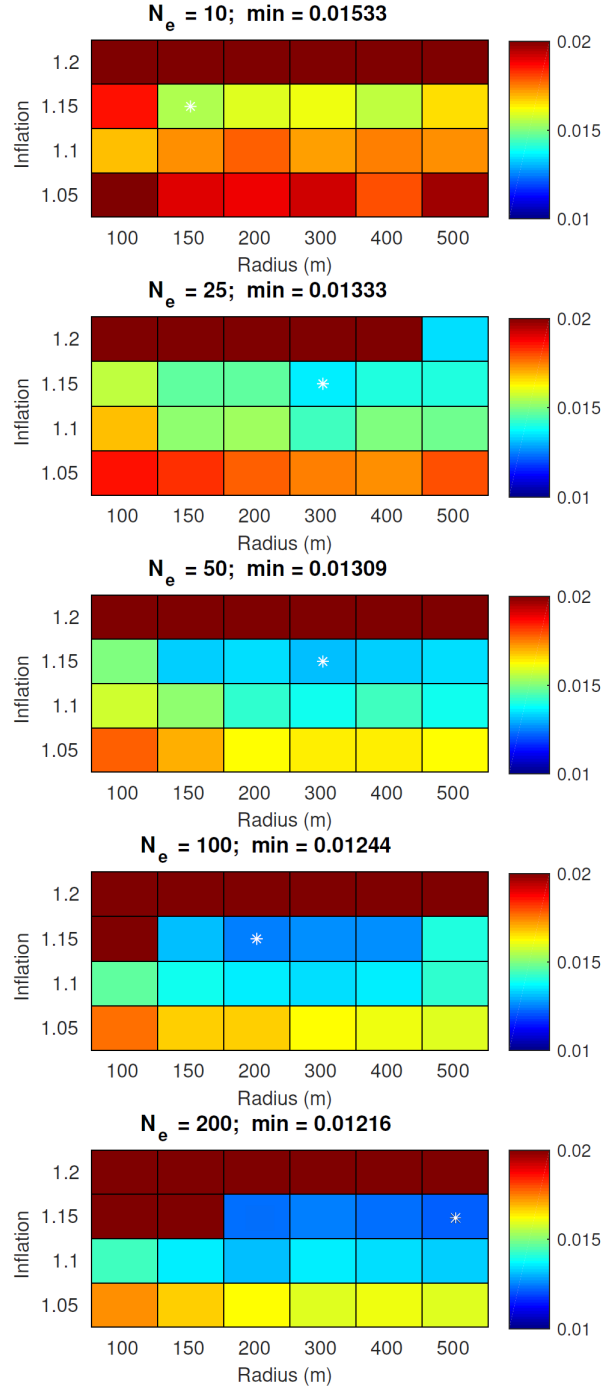
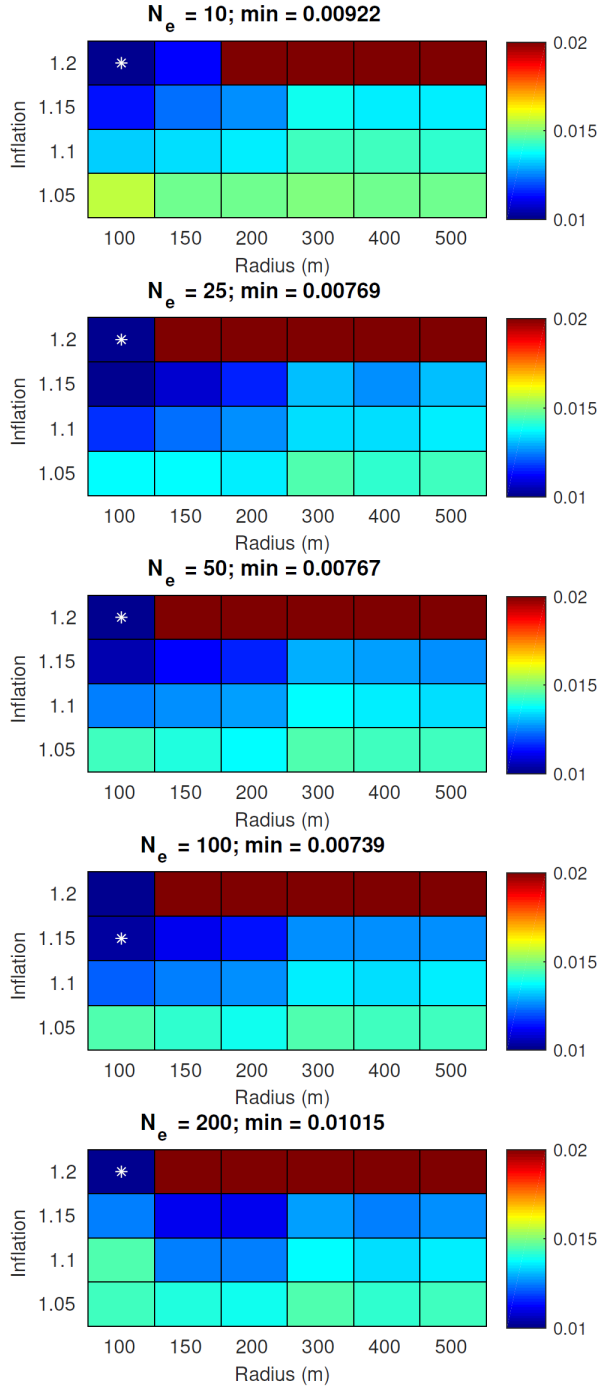


Figure 7: RMSE of the water surface elevation predictions as they result from the forecast estimates obtained by the EnKF (left column) and Joint-EnKF (right column) for different combinations of ensemble size $N_e = (10, 25, 50, 100, 200)$, localization and inflation. White stars denote the minimum RMSE value for each N_e used.

EnKF Analysis



Joint-EnKF Analysis

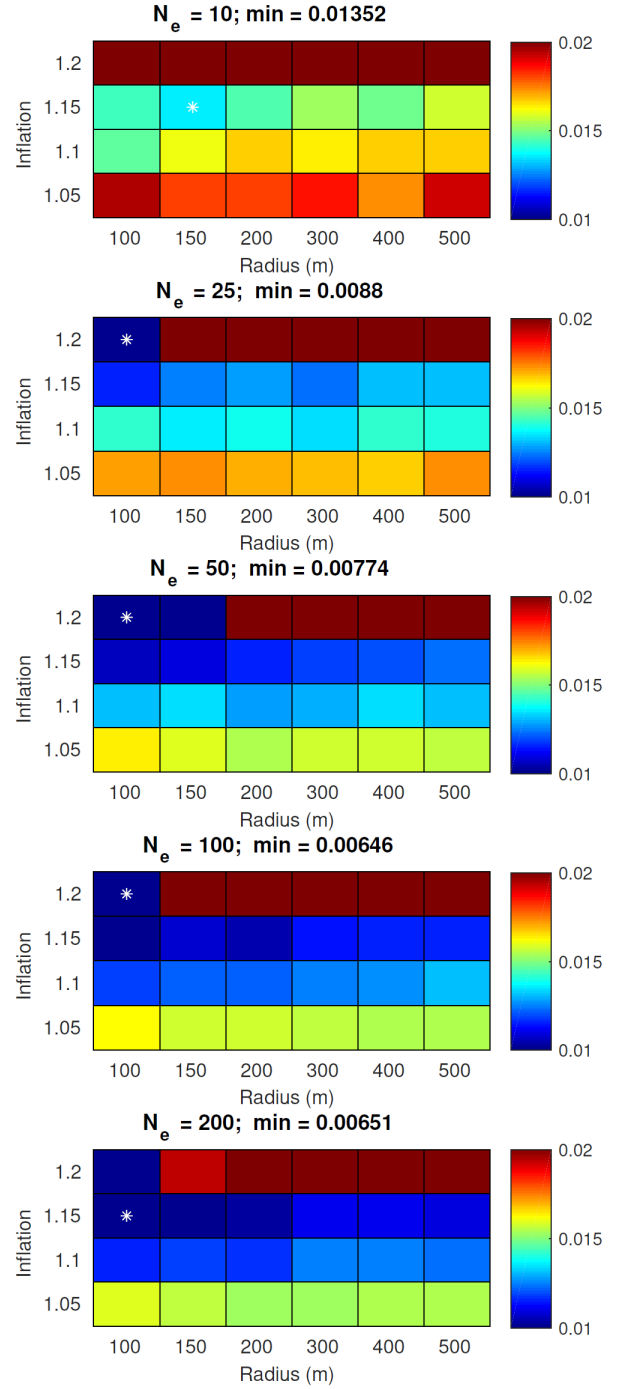


Figure 8: RMSE of the water surface elevation predictions as they result from the analysis estimates obtained by the EnKF (left column) and Joint-EnKF (right column) for different combinations of ensemble size $N_e = (10, 25, 50, 100, 200)$, localization and inflation. White stars denote the minimum RMSE value for each N_e used.

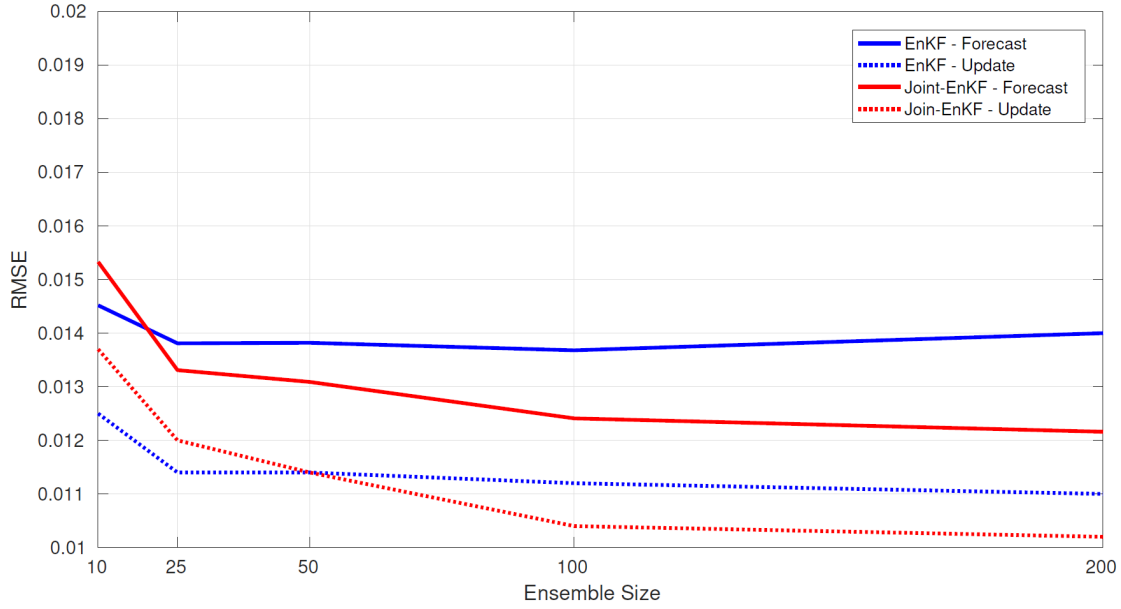


Figure 9: Comparison between EnKF (blue) and Joint-EnKF (red) RMSE trend of the water depth forecast estimate, for different values of ensemble size, considering the best estimate reached by the two filters: Joint-EnKF (localization = 500m, inflation = 1.15), EnKF (localization = 200m, inflation = 1.15)

To further analyze the solutions of the EnKF and Joint-EnKF in comparison with that of the SWE model free-run (without assimilation) and the EnKF, we compare their water surface elevation estimates against the observations for each of the 21 gauges (Figs. 10 and 11) considering the best combination, in terms of minimum RMSE, of ensemble size - localization - inflation ($N_e = 100$, radius = 200m, inflation = 1.15, for EnKF, and $N_e = 200$, radius = 500m, inflation = 1.15 for Joint-EnKF). The simulated event clearly exhibits the features of a dam break phenomena: the inflow hydrograph is based on a vertical raising stage followed by a slow falling limb. As a result, the dynamics exhibit a sharp front that occurs through the valley with high-velocity propagation and one-dimensional wave behavior in many areas of the domain. Figs 10 and 11 show the efficiency of both EnKF and Joint-EnKF to improve the behaviour of the SWE model. Although a few of the 21 gauges are not particularly well estimated by any of the simulation scenarios (SWE model, EnKF and Joint-EnKF), comparison with experimental data suggests that the overall behavior of the flooding is well reproduced.

Closer examination of the results from the upstream section toward downstream shows that the level measurements at gauges P5, P9, P12, P13, P21, P23, P24, P25 and P26 are very well predicted by both EnKF and Joint-EnKF. The Joint-EnKF provided lower RMSE values at all of the aforementioned gauges compared to the EnKF, with improvements reaching up to 90% with respect to the SWE model. The SWE model predictions at gauges P9 and P12 are tightly connected, as P9 is located in front of the portion of the levee of the reservoir where overtopping occurred in both the physical and numerical models. The model, even though it reasonably underestimates the water level at P9, does not accurately predict the levee overtopping. Indeed, the overtopping occurs 1.5 m closer to the reservoir intake, where the crest of the levee is lower. This is due to a run-up caused by a conversion of kinetic energy into potential energy, since the water level at P9 is lower than the top of the levee. However, this transfer cannot be reproduced by the SWE model (Prestininzi, 2008), and thus the volumes of water entering the reservoir are likely inaccurately filling the reservoir and overestimating P12 levels. The delay of the hydrograph is due to the different modalities and locations of the overtopping.

The water level at gauge P1 is well predicted by the Joint-EnKF, while the EnKF is less performant, providing only half of the improvement of the Joint-EnKF. EnKF and Joint-EnKF also improve the prediction of the water level at the P3, P8, P10, P18, S6S, S6D stations, by about 27 % to 57 %, on average. A bias is still clearly visible between the filters' estimates and the measurements. This suggests that the flow at these locations is also affected by the incomplete physics of the SWE model, or to uncertainties in some other parameters inputs that are not being estimated as part of the filtering process (the topography, for instance). Again, the Joint-EnKF provides better estimates than the EnKF at almost all locations. In particular, station S6S requires a more extensive analysis as it is not clear why the SWE model completely fails to simulate the water depth, both qualitatively and quantitatively, with the EnKF and Joint-EnKF only slightly improving the solution. An explanation, in this particular case, can be suggested if we consider the

physical location of the station. S6S is deployed in the shadowed area: the relief located just upstream of the gauge deflects the flow towards the center of the valley, screening the gauge from the main flood wave. Since the front of inundation here travels parallel to the cross section of the valley, the marked difference in the measured levels at the same instant reveals the deflection effect. Moreover, the surface level at S6S does not exhibit a falling limb. Instead, its tail shows a slightly raising trend and suggests the occurrence of a relatively calm and steady recirculation.

At P2, while the Joint-EnKF is able to improve the water level prediction by almost 24%, the EnKF fails and provides a solution that is even worse than the SWE model. The large, low frequency observation of water surface at gauge S4 (during the first half of the simulation period) is well predicted by both EnKF and Joint-EnKF. However, while the peak level is almost exactly predicted, the falling limb is fairly well simulated by the EnKF, but not by the Joint-EnKF, whose performance, in terms of RMSE water level time-evolution, is similar to the model. The behavior of the water depth estimates at gauge P19 is fairly well predicted only during the first part of the simulation period, while in the second half, the filters fail to accurately predict the water level. Peak level at gauges P4 is well predicted by both EnKF and the Joint-EnKF, although the latter subsequently overestimates all the rest of the hydrograph, even providing a slightly worse solution (in terms of RMSE error) than that of the SWE model. The same conclusion can be made for gauge S8D, where the main peak is fairly well simulated, but the rest of the water depth is underestimated by both EnKF and Joint-EnKF. Limited performances of the filters at P4 and S8D can be probably explained by the topography-related complexity of the flow at these specific locations, which was not fully accounted for in the prescribed model error statistics (presently through simple inflation). The EnKF and Joint -EnKF runs took about 10 and 22 hours, respectively, using a 12 cores, 2.67 GHz and 48Gb RAM workstation. The more time required by the Joint scheme is due to the larger number of variables of the joint state vector, which includes water height, velocities, and also the Manning coefficients.

EnKF Solution

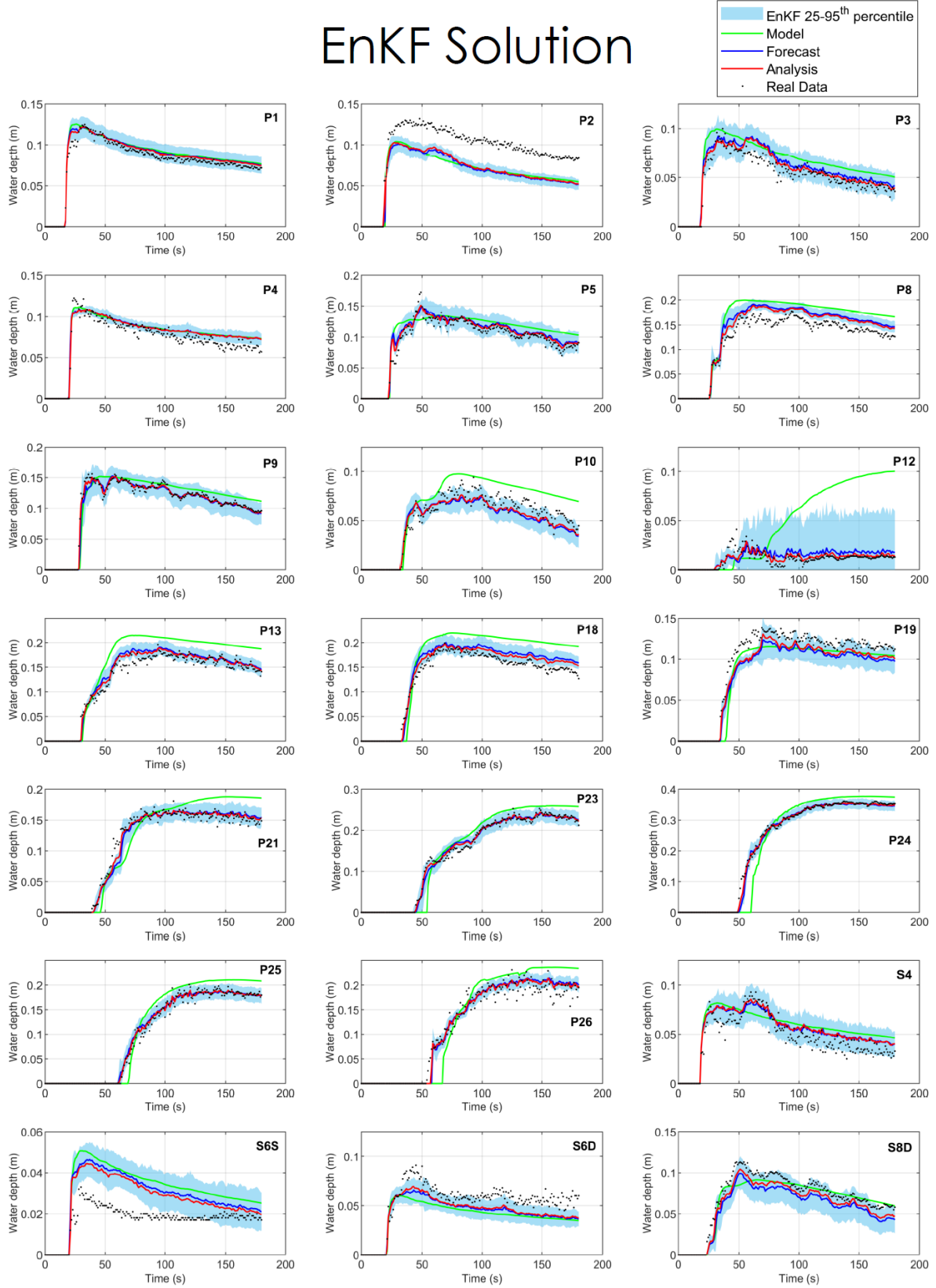


Figure 10: Comparison between experimental data (black spots), results obtained with the SWE model (green), EnKF forecast (blue) and EnKF analysis estimates (red). The shaded area represents the 25 – 95th percentile of the forecast ensemble as it results from the EnKF.

Joint-EnKF Solution

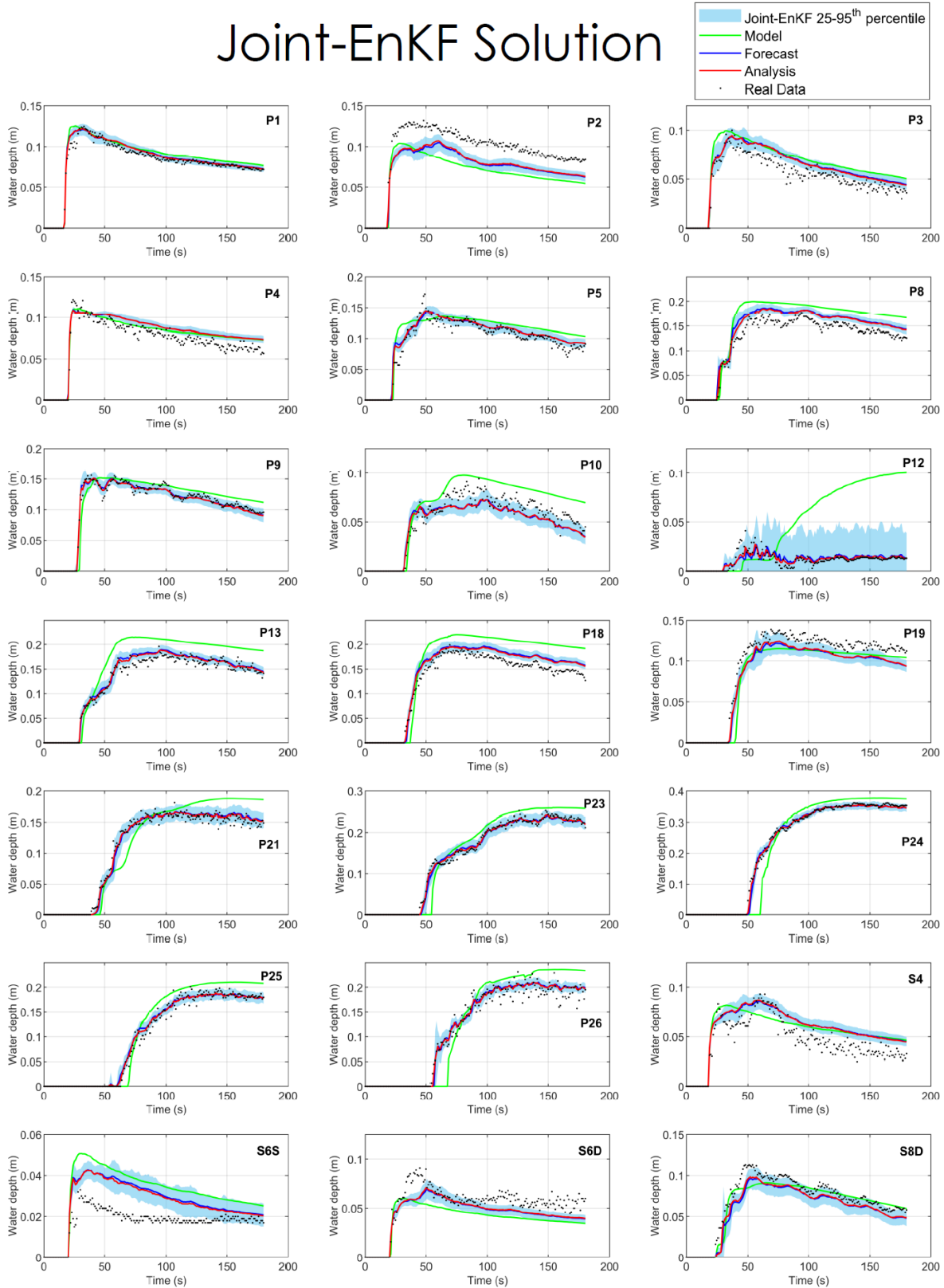


Figure 11: Comparison between experimental data (black spots), results obtained with the SWE model (green), Joint-EnKF forecast (blue) and Joint -EnKF analysis estimates (red). The shaded area represents the 25 – 95th percentile of the forecast ensemble as it results from the Joint-EnKF.

Overall, the behaviours of the SWE model and EnKFs can be summarized as follows: 1) the peak level magnitude is well estimated almost everywhere by both model and filters, though occasionally delayed and underestimated by the SWE model; 2) the solution of the SWE model is smooth compared to the EnKF and Joint-EnKF solutions, which delineate the water flow dynamic much more accurately; and 3) both the EnKF and the Joint-EnKF are able to enhance the water depth estimation at almost all the locations, improving the SWE solution by up to 90% after assimilating the data, and 60% during forecasting, on average.

For completeness, a quantitative analysis is also presented in Table 1, summarizing the RMSE values (in meters) comparing measured and predicted water depths of the SWE model (second column), Joint-EnKF (third column) and EnKF forecast estimates (fourth column). The results of the EnKFs are presented in terms of percentage improvement of the filters with respect to the SWE model prediction. The significantly better estimates of the filters compared to the free-run outlines the relevance of assimilating the gauges data into the flood model. On average, the Joint -EnKF is able to achieve better results for most of the gauges (16 out of 21) compared to the EnKF, demonstrating the relevance of estimating the parameters alongside the state. The close performances of the EnKF and Joint-EnKF are due to the robust initial Manning values, which were provided by ENEL specialists. Still, the Joint-EnKF was able to enhance the filter forecast results through improved Manning coefficients. Fig. 12 also compares the results of the SWE model, EnKF, and Joint-EnKF in terms of RMSE for every single station as they result from the best combination ensemble size - localization - inflation in term of average RMSE. Water level prediction at P12, P13, P23, P24 and P25 clearly demonstrate the ability of the two filters to provide better water depth estimates than the SWE model, reaching more than 60% improvement compared to the model free-run. The large gap between the filters and the model curves confirms the previous assertion. Conversely, the water level at stations P4 and S8D are not well predicted by either the EnKF or the Joint-EnKF, somehow leading

408 to a higher RMSE values than the SWE model.

Gauges	SWE Model	Joint -EnKF(Impr.%)	EnKF(Impr.%)
P1	0.0072	0.0048(33.61)	0.0059(18.08)
P2	0.0298	0.0227(23.756)	0.0302(-1.38)
P3	0.0142	0.0103(27.61)	0.0094(34.02)
P4	0.0085	0.0106(-24.06)	0.0087(-2.19)
P5	0.0175	0.0105(40.07)	0.0105(39.91)
P8	0.0340	0.0192(43.12)	0.0210(38.41)
P9	0.0148	0.0054(63.98)	0.0080(45.93)
P10	0.0168	0.0098(41.09)	0.0095(43.77)
P12	0.0535	0.0054(89.45)	0.0070(86.91)
P13	0.0368	0.0104(71.89)	0.0133(63.82)
P18	0.0392	0.0165(57.83)	0.0169(56.73)
P19	0.0156	0.0136(12.62)	0.0155(0.60)
P21	0.0246	0.0082(66.11)	0.0109(55.96)
P23	0.0263	0.0085(67.17)	0.0094(64.30)
P24	0.0408	0.0115(71.99)	0.0129(68.31)
P25	0.0214	0.0085(60.23)	0.0085(60.42)
P26	0.0341	0.0138(59.46)	0.0163(52.28)
S4	0.0145	0.0144(1.58)	0.0106(26.86)
S6S	0.0156	0.0111(28.52)	0.0129(17.61)
S6D	0.0170	0.0121(28.69)	0.0138(19.12)
S8D	0.0096	0.0119(-24.71)	0.0136(-42.17)

Table 1: RMSE of the forecast water depth estimate using Joint-EnKF and EnKF and improvement with respect to the SWE model. Green and red values represent respectively the best (> 60%) improvements and the negative performances of the filters for all the gauges.

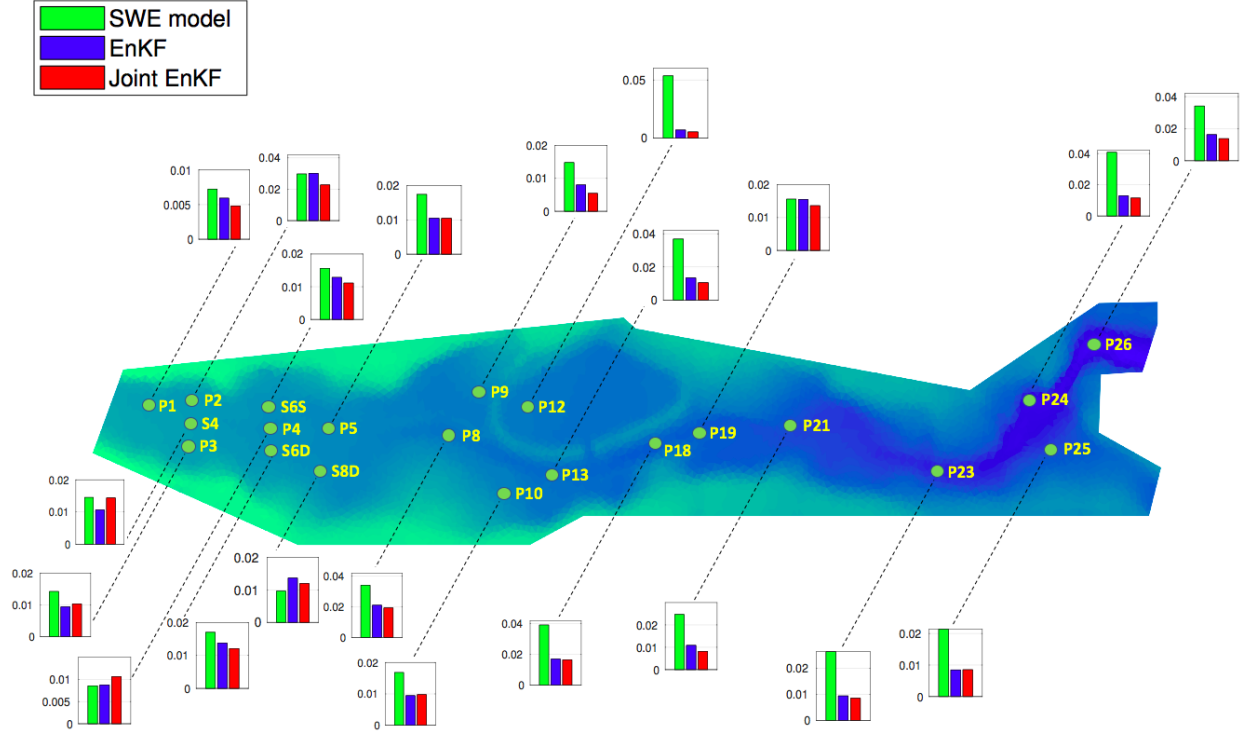


Figure 12: RMSE of the filters' forecast water depth at the 21 gauges using the SWE model (green), Joint-EnKF (red) and EnKF (blue). Data are assimilated every 1 s.

We further analyzed the spatial patterns of the Manning coefficients as estimated by the Joint-EnKF using different ensemble sizes: 10, 25, 50, 100, 200, 500, 1000 and 2000 members. As shown in Fig. 13, which plots the estimated Manning values, a larger ensemble not only improves the RMSE of the water depth estimation but also better delineates the Manning areas and most importantly, provides a spatially smoother solution. The noisy pattern resulting from an ensemble smaller than 200 members is clearly improved after using larger ensembles $N_e = 500, 1000$ and 2000. This is consistent with the results of Li et al. (2012) and Gharamti et al. (2015) who suggested that larger ensembles are needed to efficiently handle the indirect nonlinear relation between the parameters and the assimilated observations in an EnKF.

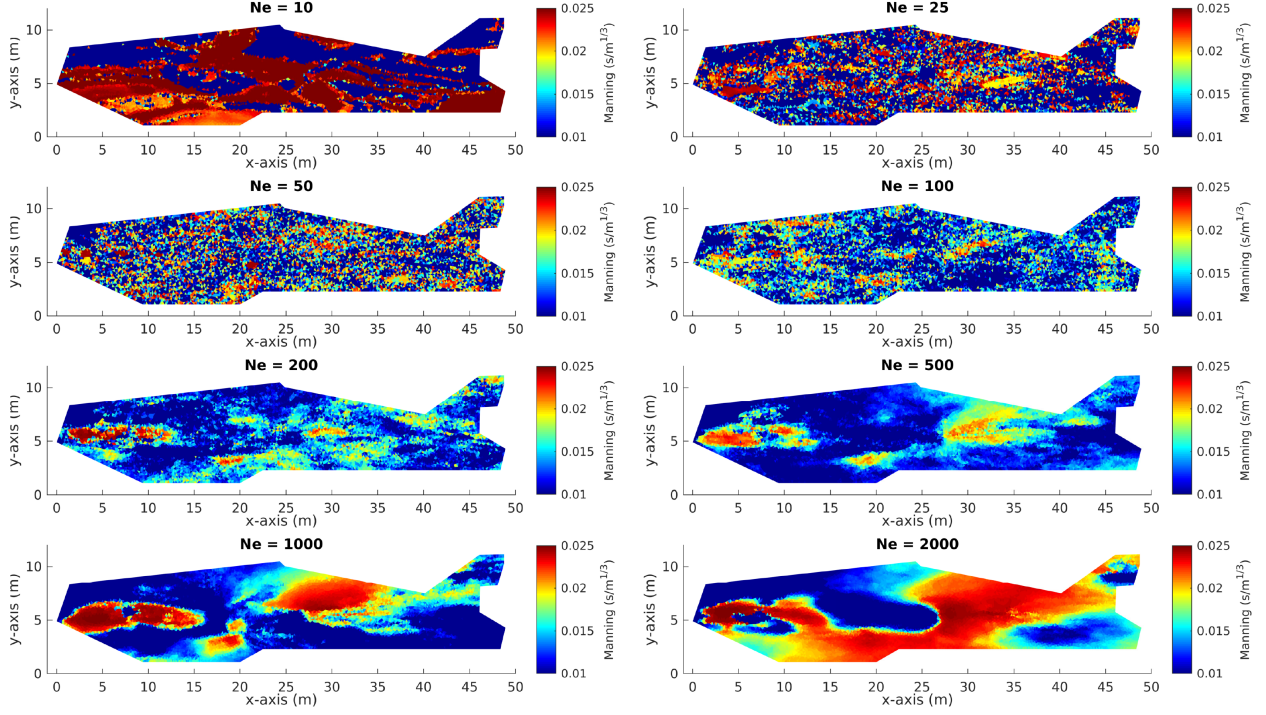


Figure 13: Spatial pattern of the Manning forecast estimates values by the Joint-EnKF for different ensemble values $N_e = 10, 25, 50, 100, 200, 500, 1000$ and 2000 . Data are assimilated every 1 s.

A further experiment has been carried out by integrating the deterministic SWE model forward with the estimated parameter field by the Joint-EnKF at the end of the assimilation experiment. We used the spatially-varying Manning resulting from the best combination of ensemble size $N_e = 200$, localization radius = 500m, and inflation = 1.15. The comparison with the original SWE model free run (without assimilation) revealed an improvement of about 10%.

Considering this same combination of filter parameters, we further estimated only the Manning coefficient with the EnKF (i.e. we included only the Manning coefficient in the filter state vector). In terms of RMSE of the water level averaged over the assimilation period, the data assimilation system that estimated only the Manning coefficient reduced the error by about 10% compared to the SWE model free run, in agreement with the other results.

5.2. Sensitivity to the frequency of observations

The sensitivity of the EnKF and Joint-EnKF to the size of assimilation interval of water depth observations was examined. Six scenarios were considered in which water depth measurements were assimilated every 1, 2, 5, 10, 15, and 20 seconds. The filters were implemented using the best combination of ensemble size - radius - inflation obtained from the first set of the experiments and the standard deviation was set to the measurement error of $10^{-2}m$. Fig. 14 plots the RMSE of the water depth forecasts for the different assimilation intervals of the observations.

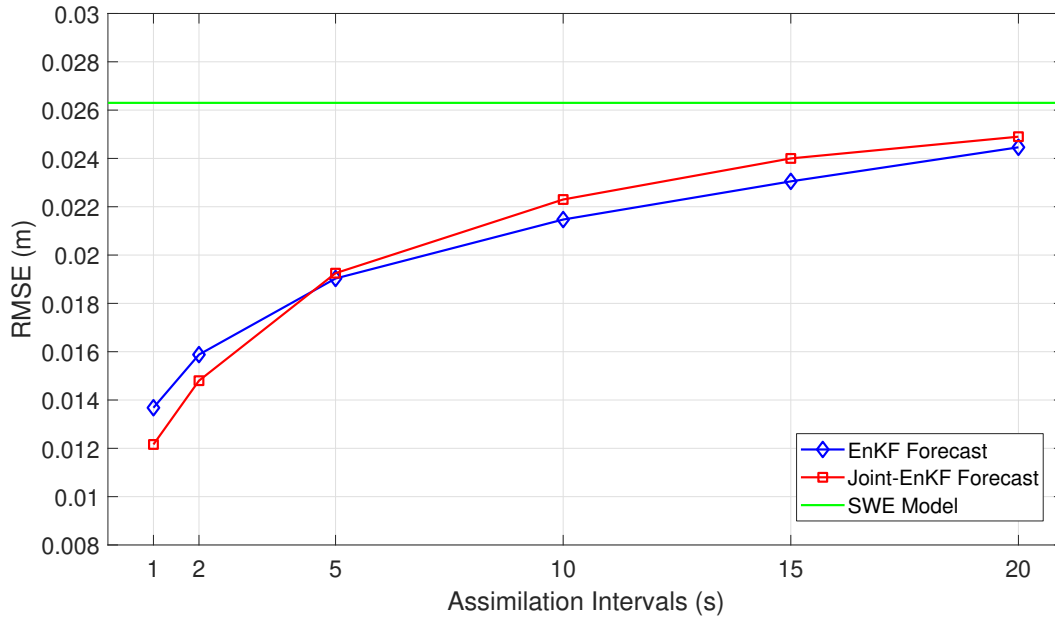


Figure 14: RMSE of the water depth using EnKF and Joint-EnKF filters for different data assimilation intervals (1, 2, 5, 10, 15 and 20 s).

Both Joint-EnKF and EnKF are able to improve the solution as long as the assimilation periods with which the data are assimilated remains small, i.e. every 1 s, with an RMSE of 0.01216 and 0.01389, respectively. The estimated errors increase as the observation interval increases, reaching an RMSE of 0.0148, 0.01925, 0.0223, 0.024 and 0.0249 m, for the Joint-EnKF, and 0.01588, 0.01903, 0.02147, 0.02305 and 0.02446 m, for the EnKF. One may notice that when the data are frequently assimilated (every 5 s and less), the Joint-EnKF only slightly outperforms the EnKF. This is because, in this case, one can rely

on enough observations to adjust the model forecast and avoid its deviation from the true trajectory even with inaccurate Manning coefficients. Overall, both filters perform better when the data are assimilated more frequently. The Joint-EnKF seems to be more sensitive to the temporal size of assimilation interval, which increases the non linearity of the estimation problem, as has been also reported in [Gharamti et al. \(2015\)](#).

6. Summary and conclusions

A reconstruction and prediction experiment, based upon an experimentally reproduced dam break event, was performed using a data assimilation framework based on a shallow water equation (SWE) model and an ensemble Kalman filter (EnKF). Numerical experiments assimilated real experimental data from the Toce physical model test used in the Concerted Action on Dambreak Modeling (CADAM) project. In this context, the physical model was a 1:100 scale model of a reach of the Toce River valley developed at the ENEL-HYDRO (National Institution for Electrical Energy). The concurrent estimation of the state of the system (composed of water height and velocity) together with the Manning coefficients across the domain was explored using a Joint-EnKF approach. Results suggest that the EnKF and Joint-EnKF were both able to improve the SWE model behavior and forecasting skills, while more accurately reconstructing the propagation of such an impulsive wave over complex topography.

Overall, the EnKF assimilation system was able to replicate the general development of the dynamic process and to reconstruct the inundation arrival time, the water surface elevation and the local peak values with significantly better accuracy than the SWE model free-run. Comparing the EnKF and the Joint-EnKF performances, we found that estimating the parameters alongside the state significantly improves the filter's results. However, large ensembles were needed in the Joint-EnKF to obtain robust estimates of the Manning field. The filter failed to improve the model behavior at two (out of 21) gauges, likely because of the unreliable ensemble statistics sampled from the model outputs at these

complex locations that are not well described by the physics of the numerical forecast model. This suggests the importance of treating the true source of errors in the model to obtain the best filters' performances. The Joint-EnKF also behaved better when the data were assimilated more frequently.

In this work, only the Manning field was estimated by the Joint-EnKF as it was the main unknown. However, this framework could be expanded to include other important parameters in the estimation process, such as the bathymetry elevation and the inflow hydrograph. The latter in particular, being a time-varying parameter, is generally poorly known and as such might need to be included in the estimation process as well. In future work, we will also explore ways to reduce the sensitivity of the Joint-EnKF to the size of the ensemble and borrow ideas from the dual state-parameters EnKF estimation framework ([Ait-El-Fquih et al., 2016](#)) to use different ensemble sizes, and even different update schemes ([Ait-El-Fquih and Hoteit, 2018](#)), for the state and parameters.

Acknowledgement

Research reported in this publication was supported by King Abdullah University of Science and Technology (KAUST).

List of Figures

1	A schematic illustration of an assimilation cycle of the (Joint)-EnKF. State vector Z could be $X = (\text{water depth, velocities})$ for EnKF or $(X, \theta) = \text{Manning coefficient for Joint-EnKF}$	11
2	Hydrograph of the flash flood measured during the experiment.	13
3	Digital Terrain Model (DTM) of the river valley showing the decreasing elevation from left to right side of the domain. 21 gauges are spread along the path of the river, providing water level measurements throughout the duration of the simulation.	13
4	Unstructured mesh made by 10185 triangular cells used for the numerical simulation.	14
5	Initial water depth maps at initial time (dry bed, top), after 20 seconds (second from top), after 40 seconds (third from top) and after 60 seconds of simulation (bottom) using the SWE model.	15
6	Three different realizations of the Manning parameter obtained using the Gaussian simulation toolbox (Gómez-Hernández and Journel, 1993).	16
7	RMSE of the water surface elevation predictions as they result from the forecast estimates obtained by the EnKF (left column) and Joint-EnKF (right column) for different combinations of ensemble size $N_e = (10, 25, 50, 100, 200)$, localization and inflation. White stars denote the minimum RMSE value for each N_e used.	19

509	8	RMSE of the water surface elevation predictions as they result from the analysis estimates obtained by the EnKF (left column) and Joint-EnKF (right column) for different combinations of ensemble size $N_e = (10, 25, 50, 100, 200)$, localization and inflation. White stars denote the minimum RMSE value for each N_e used.	20
510			
511			
512			
513			
514	9	Comparison between EnKF (blue) and Joint-EnKF (red) RMSE trend of the water depth forecast estimate, for different values of ensemble size, considering the best estimate reached by the two filters: Joint-EnKF (localization = 500m, inflation = 1.15), EnKF (localization = 200m, inflation = 1.15)	21
515			
516			
517			
518	10	Comparison between experimental data (black spots), results obtained with the SWE model (green), EnKF forecast (blue) and EnKF analysis estimates (red). The shaded area represents the 25 – 95 th percentile of the forecast ensemble as it results from the EnKF.	24
519			
520			
521			
522	11	Comparison between experimental data (black spots), results obtained with the SWE model (green), Joint-EnKF forecast (blue) and Joint -EnKF analysis estimates (red). The shaded area represents the 25 – 95 th percentile of the forecast ensemble as it results from the Joint-EnKF.	25
523			
524			
525			
526	12	RMSE of the filters' forecast water depth at the 21 gauges using the SWE model (green), Joint-EnKF (red) and EnKF (blue). Data are assimilated every 1 s.	28
527			
528			
529	13	Spatial pattern of the Manning forecast estimates values by the Joint-EnKF for different ensemble values $N_e = 10, 25, 50, 100, 200, 500, 1000$ and 2000. Data are assimilated every 1 s.	29
530			
531			
532	14	RMSE of the water depth using EnKF and Joint-EnKF filters for different data assimilation intervals (1, 2, 5, 10, 15 and 20 s).	30
533			

References

- Ait-El-Fquih, B., El Gharamti, M., Hoteit, I., 2016. A bayesian consistent dual ensemble kalman filter for state-parameter estimation in subsurface hydrology. *Hydrol. Earth Syst. Sci* 20, 3289–3307.
- Ait-El-Fquih, B., Hoteit, I., 2016. A variational bayesian multiple particle filtering scheme for large-dimensional systems. *IEEE Transactions on Signal Processing*, 64 (20), 5409–5422.
- Ait-El-Fquih, B., Hoteit, I., 2018. An efficient state–parameter filtering scheme combining ensemble kalman and particle filters. *Monthly Weather Review* 146 (3), 871–887.
- Anderson, J. L., Anderson, S. L., 1999. A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127 (12), 2741–2758.
- Audusse, E., Bouchut, F., Bristeau, M.-O., Klein, R., Perthame, B., 2004. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM Journal on Scientific Computing*, 25 (6), 2050–2065.
- Begnudelli, L., Sanders, B. F., 2006. Unstructured grid finite-volume algorithm for shallow-water flow and scalar transport with wetting and drying. *Journal of Hydraulic Engineering*, 132 (4), 371–384.
- Blöschl, G., Sivapalan, M., 1995. Scale issues in hydrological modelling: a review. *Hydrological processes*, 9 (3-4), 251–290.
- Caleffi, V., Valiani, A., Zanni, A., 2003. Finite volume method for simulating extreme flood events in natural channels. *J. Hydraul. Res.* 41 (2), 167–177.
- Chen, Y., Zhang, D., 2006. Data assimilation for transient flow in geologic formations via ensemble kalman filter. *Advances in Water Resources*, 29 (8), 1107–1122.

- Constantin, A., Escher, J., 1998. Wave breaking for nonlinear nonlocal shallow water equations. *Acta Mathematica*, 181 (2), 229–243.
- Edwards, C. A., Moore, A. M., Hoteit, I., Cornuelle, B. D., 2015. Regional ocean data assimilation. *Annual Review of Marine Science*, 7, 21–42.
- Elsheikh, A. H., Wheeler, M. F., Hoteit, I., 2013. An iterative stochastic ensemble method for parameter estimation of subsurface flow models. *Journal of Computational Physics*, 242, 696–714.
- Evensen, G., 2003. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53 (4), 343–367.
- Eymard, R., Gallouët, T., Herbin, R., 2000. Finite volume methods. *Handbook of Numerical Analysis*, 7, 713–1018.
- Few, R., 2003. Flooding, vulnerability and coping strategies: local responses to a global threat. *Progress in Development Studies*, 3 (1), 43–58.
- García-Pintado, J., Neal, J. C., Mason, D. C., Dance, S. L., Bates, P. D., 2013. Scheduling satellite-based sar acquisition for sequential assimilation of water level observations into flood modelling. *Journal of Hydrology*, 495, 252–266.
- Gharamti, M., Ait-El-Fquih, B., Hoteit, I., 2015. An iterative ensemble kalman filter with one-step-ahead smoothing for state-parameters estimation of contaminant transport models. *Journal of Hydrology*, 527, 442–457.
- Gharamti, M., Hoteit, I., 2014. Complex step-based low-rank extended kalman filtering for state-parameter estimation in subsurface transport models. *Journal of Hydrology*, 509, 588–600.
- Ghil, M., 1989. Meteorological data assimilation for oceanographers. Part I: Description and theoretical framework. *Dynamics of Atmospheres and Oceans*, 13 (3-4), 171–218.

- Gichamo, T. Z., Popescu, I., Jonoski, A., Solomatine, D., 2012. River cross-section extraction from the aster global dem for flood modeling. *Environmental Modelling & Software* 31, 37–46.
- Gómez-Hernández, J. J., Journel, A. G., 1993. Joint sequential simulation of multigaussian fields. In: *Quantitative Geology and Geostatistics*, Springer, pp. 85–94.
- Hamill, T. M., Snyder, C., 2000. A hybrid ensemble kalman filter–3d variational analysis scheme. *Monthly Weather Review*, 128 (8), 2905–2919.
- Hamill, T. M., Whitaker, J. S., Snyder, C., 2001. Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. *Monthly Weather Review*, 129 (11), 2776–2790.
- Hendricks Franssen, H., Kinzelbach, W., 2008. Real-time groundwater flow modeling with the ensemble kalman filter: Joint estimation of states and parameters and the filter inbreeding problem. *Water Resources Research*, 44 (9).
- Holland, W. R., Malanotte-Rizzoli, P., 1989. Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies. *Journal of Physical Oceanography*, 19 (10), 1507–1534.
- Hostache, R., Lai, X., Monnier, J., Puech, C., 2010. Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of mosel river. *Journal of Hydrology*, 390 (3), 257–268.
- Hoteit, I., Pham, D.-T., Blum, J., 2002. A simplified reduced order kalman filtering and application to altimetric data assimilation in tropical pacific. *Journal of Marine Systems*, 36 (1), 101–127.
- Hoteit, I., Pham, D.-T., Gharamti, M., Luo, X., 2015. Mitigating observation perturbation sampling errors in the stochastic enkf. *Monthly Weather Review*, 143 (7), 2918–2936.

- Hoteit, I., Pham, D.-T., Triantafyllou, G., Korres, G., 2008. A new approximate solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography. *Monthly Weather Review*, 136 (1), 317–334.
- Hoteit, I., Triantafyllou, G., Petihakis, G., Allen, J., 2003. A singular evolutive extended kalman filter to assimilate real in situ data in a 1-d marine ecosystem model. In: *Annales Geophysicae*, Vol. 21. pp. 389–397.
- Hou, J., Liang, Q., Zhang, H., Hinkelmann, R., 2015. An efficient unstructured muscl scheme for solving the 2d shallow water equations. *Environmental Modelling and Software*, 66, 131–152.
- Kalman, R. E., et al., 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82 (1), 35–45.
- Lai, X., Monnier, J., 2009. Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: Mathematical method and test case. *Journal of Hydrology*, 377 (1), 1–11.
- Landon, K. C., Wilson, G. W., Özkan-Haller, H. T., MacMahan, J. H., 2014. Bathymetry estimation using drifter-based velocity measurements on the kootenai river, idaho. *Journal of Atmospheric and Oceanic Technology*, 31 (2), 503–514.
- Lermusiaux, P. F., 2006. Uncertainty estimation and prediction for interdisciplinary ocean dynamics. *Journal of Computational Physics*, 217 (1), 176–199.
- Li, L., Zhou, H., Gómez-Hernández, J. J., Franssen, H.-J. H., 2012. Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble kalman filter. *Journal of hydrology*, 428, 152–169.
- Liu, Y., Gupta, H. V., 2007. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, 43 (7).

- Mayo, T., Butler, T., Dawson, C., Hoteit, I., 2014. Data assimilation within the advanced circulation (adcirc) modeling framework for the estimation of manning's friction coefficient. *Ocean Modelling* 76, 43–58.
- Prestininzi, P., 2008. Suitability of the diffusive model for dam break simulation: Application to a cadam experiment. *Journal of Hydrology*, 361 (1), 172–185.
- Siripatana, A., Mayo, T., Sraj, I., Knio, O., Dawson, C., Le Maitre, O., Hoteit, I., 2017. Assessing an ensemble kalman filter inference of manning's coefficient of an idealized tidal inlet against a polynomial chaos-based mcmc. *Ocean Dynamics*, 1–28.
- Soares-Frazão, S., Testa, G., 1999. Numerical model analysis: The toce river test case. In: *CADAM Proceedings, Milan Meeting*.
- Song, H., Hoteit, I., Cornuelle, B. D., Subramanian, A. C., 2010. An adaptive approach to mitigate background covariance limitations in the ensemble kalman filter. *Monthly Weather Review*, 138 (7), 2825–2845.
- Strang, G., Fix, G. J., 1973. *An analysis of the finite element method*. Vol. 212. Prentice-hall Englewood Cliffs, NJ.
- Strub, I. S., Percelay, J., Tossavainen, O.-P., Bayen, A. M., 2009. Comparison of two data assimilation algorithms for shallow water flows. *Networks and Heterogeneous Media* 4 (2), 409–430.
- Tirupathi, S., Tchrakian, T. T., Zhuk, S., McKenna, S., 2016. Shock capturing data assimilation algorithm for 1d shallow water equations. *Advances in Water Resources*, 88, 198–210.
- Toro, E. F., 2009. The hll and hllc riemann solvers. In: *Riemann Solvers and Numerical Methods for Fluid Dynamics*,. Springer, pp. 315–344.

- 653 Tossavainen, O.-P., Percelay, J., Tinka, A., Wu, Q., Bayen, A. M., 2008. Ensemble kalman
654 filter based state estimation in 2d shallow water equations using lagrangian sensing and
655 state augmentation. In: Decision and Control, 2008. CDC 2008. 47th IEEE Conference
656 on., IEEE, pp. 1783–1790.
- 657 Van Leer, B., 1979. Towards the ultimate conservative difference scheme. V. A second-order
658 sequel to Godunov's method. *Journal of Computational Physics*, 32 (1), 101–136.
- 659 Wilson, G., Özkan-Haller, H., Holman, R., 2010. Data assimilation and bathymetric inver-
660 sion in a two-dimensional horizontal surf zone model. *Journal of Geophysical Research:*
661 *Oceans*, 115 (C12).
- 662 Wilson, G., Özkan-Haller, H. T., 2012. Ensemble-based data assimilation for estimation of
663 river depths. *Journal of Atmospheric and Oceanic Technology*, 29 (10), 1558–1568.
- 664 Wu, H., Adler, R. F., Tian, Y., Huffman, G. J., Li, H., Wang, J., 2014. Real-time global flood
665 estimation using satellite-based precipitation and a coupled land surface and routing
666 model. *Water Resources Research*, 50 (3), 2693–2717.
- 667 Zhang, G., Xie, H., Duan, S., Tian, M., Yi, D., 2011. Water level variation of lake qinghai
668 from satellite and in situ measurements under climate change. *Journal of Applied*
669 *Remote Sensing*, 5 (1), 053532.