Dual State-Parameter Assimilation of SAR-derived Wet Surface Ratio for Improving Fluvial Flood Reanalysis

Thanh Huy Nguyen[1], Sophie Ricci[1], Andrea Piacentini[2], Christophe Fatras[3], Peter Kettig[4], Gwendoline Blanchet[4], Santiago Peña Luque[4], and Simon Baillarin[4]

[1]CECI-CERFACS CNRS UMR 5318
[2]CERFACS
[3]Collecte Localisation Satellites
[4]Centre National d'Etudes Spatiales

November 21, 2022

Abstract

Flooding is one of the most devastating natural hazards to which our society worldwide must adapt, especially as its severity and occurrence tend to increase with climate changes. This research work focuses on the assimilation of 2D flood observations derived from remote-sensing images acquired during overflowing events. To do so, the resulting binary wet/dry maps are expressed in terms of wet surface ratios (WSR) over a number of floodplain subdomains. This ratio is assimilated jointly with in-situ water-level gauge observations to improve the flow dynamics within the floodplain. An Ensemble Kalman Filter with a dual state-parameter analysis approach is implemented on top of a TELEMAC-2D hydrodynamic model. The EnKF control vector is composed of spatially-distributed friction coefficients and a corrective parameter of the inflow discharge. It is extended with the hydraulic states within the floodplain subdomains. This data assimilation strategy was validated and evaluated over a reach of the Garonne river. The observation operator associated with the WSR observations, as well as the dual state-parameter sequential correction, was first validated in the context of Observing System Simulation Experiments. It was then applied to two real flood events that occurred in 2019 and 2021. The merits of assimilating SAR-derived WSR observations, in complement to the in-situ water-level observations, are shown in the parameter and observation spaces with assessment metrics computed over the entire flood events. It is also shown that the hydraulic state correction within the dual state-parameter analysis approach significantly improves the flood dynamics, especially during the flood recession.

Posted on 21 Nov 2022 — CC-BY-NC 4 — https://doi.org/10.1002/essoar.10511831.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Dual State-Parameter Assimilation of SAR-derived Wet Surface Ratio for Improving Fluvial Flood Reanalysis

T. H. Nguyen[1], S. Ricci[1], A. Piacentini[1], C. Fatras[2], P. Kettig[3], G. Blanchet[3], S. Pena Luque[3], and S. Baillarin[3]

[1]CECI, CERFACS/CNRS UMR 5318, Toulouse, 31057, France
[2]Collecte Localisation Satellites (CLS), Ramonville Saint-Agne, 31520, France
[3]Centre National d'Etudes Spatiales (CNES), Toulouse, 31401, France

Key Points:

- An Ensemble Kalman Filter data assimilation designated to flood fluvial re-analysis, built on top of a hydrodynamic TELEMAC-2D model;
- Advanced data assimilation approaches combining in-situ and SAR-derived WSR observations are proposed and comprehensively evaluated;
- Dual state-parameter analysis treating model parameters and hydraulic states in floodplain zones provides more accurate flood simulations.

Corresponding author: T. H. Nguyen, `thnguyen@cerfacs.fr`

Abstract

Flooding is one of the most devastating natural hazards to which our society worldwide must adapt, especially as its severity and occurrence tend to increase with climate changes. This research work focuses on the assimilation of 2D flood observations derived from remote-sensing images acquired during overflowing events. To do so, the resulting binary wet/dry maps are expressed in terms of wet surface ratios (WSR) over a number of floodplain subdomains. This ratio is assimilated jointly with in-situ water-level gauge observations to improve the flow dynamics within the floodplain. An Ensemble Kalman Filter with a dual state-parameter analysis approach is implemented on top of a TELEMAC-2D hydrodynamic model. The EnKF control vector is composed of spatially-distributed friction coefficients and a corrective parameter of the inflow discharge. It is extended with the hydraulic states within the floodplain subdomains. This data assimilation strategy was validated and evaluated over a reach of the Garonne river. The observation operator associated with the WSR observations, as well as the dual state-parameter sequential correction, was first validated in the context of Observing System Simulation Experiments. It was then applied to two real flood events that occurred in 2019 and 2021. The merits of assimilating SAR-derived WSR observations, in complement to the in-situ water-level observations, are shown in the parameter and observation spaces with assessment metrics computed over the entire flood events. It is also shown that the hydraulic state correction within the dual state-parameter analysis approach significantly improves the flood dynamics, especially during the flood recession.

1 Introduction

1.1 Flood monitoring and forecasting

Flooding is one of the most common, yet most severe and costliest natural disasters worldwide. According to figures provided by the United Nations Office for Disaster Risk Reduction, flooding accounted for 43.4% of all 7,255 disaster events recorded globally between 1998 and 2017[1]. Flood forecasting systems rely on both monitoring and numerical modelling. Most modelling systems concatenate hydrologic rainfall-runoff models that represent the dynamics of the catchment with hydrodynamic models that simulate the dynamics of the river bed and the floodplain. River hydrodynamic models rely on solving the Shallow Water equations (SWE) which are depth-averaged Navier-Stokes equations. They are used to predict river water surface elevation (WSE) and discharge, thus allowing for flood risk assessment. However, these numerical models remain imperfect due to the uncertainties in the model itself and its inputs, e.g., friction and boundary conditions (BC), which translate into uncertainties in the model outputs, i.e. water level and discharge. A well-established method for reducing such uncertainties is to periodically adjust these models by assimilating various available observations. As a result, flood simulation and forecast capability have greatly improved thanks to the advances in data assimilation (DA) (Madsen & Skotner, 2005; Neal & Jeffrey, 2007; Neal et al., 2009). Continuous time-series of gauged water levels and/or discharges recorded at sparse locations have been used for model calibration and validation. DA strategies, namely EnKF, classically consist in combining these time-series measurements with numerical models to correct the hydraulic states and reduce the uncertainties in the model parameters, e.g., friction coefficients, upstream inflow (Neal & Jeffrey, 2007). EnKF relies on the stochastic computation of the forecast error covariance matrix amongst a limited number of perturbed simulations. Therefore, this approach depends strongly on the characteristics of the observing network, i.e., the density, the frequency and the statistics of errors of the observation (Mirouze et al., 2019). However, due to installation and

[1] https://www.prevention-web.net/knowledgebase

maintenance costs, limnimetric in-situ gauge stations providing water levels are only available at a few locations within a catchment (Mason et al., 2012), and they are usually installed by the river. Such a spatial scarcity is a limitation for numerical model precision in simulation and forecast, especially in the floodplain. This can be overcome by the use of other data sources such as remote sensing (RS) flood maps that, despite low revisit frequency, offer a 2D representation of the flow dynamics.

1.2 Assimilation of remote sensing flood-related data

Leveraging RS products in the context of flood risk management presents a great opportunity to improve the ability of flood monitoring and forecasting (G. Schumann et al., 2009). In the recent years, SAR systems have played a major role in operational flood management, due to its reliability to collect day-and-night observations regardless of weather conditions. Water bodies and flooded areas typically exhibit low backscatter on SAR images since most of the incident radar pulses are specularly reflected away upon arrival at the water surfaces. Therefore, the detection of flooded areas is straightforward on SAR images, with several exceptions, e.g. in urban environment, vegetated areas, or when facing variability of water roughness and speckle. Indeed, mis-detection of flooded vegetation areas (i.e. partially submerged vegetation) mainly occurs because signals cannot reach the water surfaces beneath vegetation being caught in volume scattering from the canopy, or due to multiple-bounce effects between the tree trunks and the underneath water surfaces. It could also occur in urban areas due the complexity of the landscape geometry (e.g. shadow, layover, highly reflective scatterers). Over the last decades, the literature on DA into hydrodynamic models mainly focused on the assimilation of in-situ or RS-derived WSE observations (Hostache et al., 2010), mostly because this is a state variable in any hydraulic model, thereby rendering the DA more straightforward. Such methods involve retrieving WSE from the combination of RS-derived flood extent maps with topography data. Yet, this relies on the use of precise and high-resolution Digital Terrain Models (DTM) and still requires some further research to prevent potential bias from such a usage (Cian et al., 2018). As a result, recent studies have been carried out to directly assimilate flood extent maps in hydraulic models. Flood probability maps have also been estimated by a Bayesian approach applied to SAR images, and subsequently assimilated into a particle filter-based data assimilation framework (Hostache et al., 2018; Dasgupta, Hostache, Ramsankaran, Schumann, et al., 2021; Revilla-Romero et al., 2016; Di Mauro et al., 2021). Cooper et al. (Cooper et al., 2019) proposed a new observation operator that directly uses backscatter values from SAR images as observations in order to bypass the flood edge identification or flood probability estimation processes. However, this approach has only been implemented with synthetical SAR images in the scope of a twin experiment. It relies on the hypothesis that SAR images must yield distinct distributions of wet and dry backscatter values, which may not hold for real SAR data due to aforementioned limitations.

The increasing availability of highly spatially distributed RS observations of flood extent and water levels offer new opportunities for investigation and analysis (e.g., (Bates, 2004; G. Schumann et al., 2009)). The possibility of using SAR imagery data for the validation and calibration of two-dimensional (2D) hydraulic models was first highlighted by Jung et al. (Jung et al., 2012). Since then, the increasing amount of RS data and the advances in Machine Learning algorithms dedicated to water detection have enabled a great number of research work dedicated to hydrologic and hydraulic models calibration/validation for real-time forecasting. The combination of RS data with local hydrodynamic models has thus been greatly studied in the literature as it allows to overcome the limitations of both incomplete and uncertain sources of knowledge on the river and floodplain dynamics. A comprehensive review by Grimaldi et al. (Grimaldi et al., 2016) provides an analysis on the use of coarse-, medium- and high-resolution RS observations of flood extent and water level to improve the accuracy of hydraulic models for flood forecasting. It points out that RS data should be used as a complement data source—but not as an alternative—to the in-situ data in order to calibrate, validate, and constraint the hydraulic models. This stems from their low precision and acquisition frequency (Grimaldi et al., 2016). Indeed, compared to in-situ data, RS data provide useful flood extent and flood edge information at a large coverage, usually covering the whole considered catchment, but they are much sparser in terms of frequency. In addition, uncertainty exists in flood extent mapping from RS observations, e.g. SAR images, which originates from both the input images and the classification algorithm itself. As a matter of fact, classification overall accuracy of flooded areas varies considerably and only in rare cases exceeds 90% (G. J.-P. Schumann et al., 2012). An updated review from Dasgupta et al. (Dasgupta, Hostache, Ramsankaran, Grimaldi, et al., 2021) provides the state-of-the-art on the assimilation of Earth Observation data with hydraulic models for the purpose of improved flood inundation forecasting.

1.3 Objective and Outline

As the severity and occurrence of flood events tend to intensify with climate change, the need for flood forecasting capability increases. In this regard, the Flood Detection, Alert and rapid Mapping (FloodDAM) project (Kettig et al., 2021), funded by the Space for Climate Observatory initiative, was set out to develop pre-operational numerical tools to enable quick responses in various flood-prone areas while improving the resolution, reactivity, and predictive capability. In our previous works (Nguyen et al., 2021, 2022), flood extent maps were inferred from Sentinel-1 (S1) images by a Random Forest (RF) developed in the framework of the FloodML project (Huang et al., 2020; Kettig et al., 2021). In these works, in-situ water level time series at observing stations of the river bed were assimilated in order to sequentially correct friction and inflow discharge. Accordingly, the hydrodynamic model results in re-analysis and forecast modes are improved. S1-derived flood extent were then used as independent validation observations providing valuable information, especially in the floodplain. The EnKF algorithm was favored and implemented as it allows to stochastically estimate the covariance matrices between the model inputs/parameters and its outputs, without formulating the tangent linear of the hydrodynamics model, under the assumption that the errors in the control vector are properly described by a Gaussian probability density function. Taking further advantage of S1-derived flood extents, the DA of flood extent maps, expressed in terms of wet surface ratios (WSR) computed as the ratio of wet pixels detected on S1-derived water masks, over the total number of pixels in a subdomain of the floodplain, is here investigated. This strategy aims at reducing comprehensively the uncertainties in the model parameters and forcing, and consequently improve the overall flood re-analysis and forecast capability especially in the floodplain. This article presents a DA approach to accommodate 2D WSR observations alongside with in-situ water level time-series within an EnKF framework implemented on a 2D hydrodynamics model on the Garonne river. A dual state-parameter DA strategy is implemented to reduce the uncertainties in friction coefficients, upstream forcing and hydraulic state (water level in selected floodplain subdomains). The control vector is augmented with a water level state correction that is uniform over a limited number of subdomains in the floodplain. This work is first carried out in the context of Observing System Simulation Experiment (OSSE) where observations are generated from a reference simulation with chosen settings, considered as the truth. Generated in-situ and WSR synthetical observations are then assimilated into an ensemble DA with a priori (background) settings that differ from the true settings. This strategy is common in DA studies as it allows to validate a DA algorithm and whether its analysis manages to bring the resulting control vector closer to the truth's settings, as well as the resulting model state closer to the synthetical observations than the a priori state. The DA strategy is then tested for real flood events, assimilating both water level data measured at in-situ gauge stations and WSR observations derived from S1 images. The DA results are validated with respect to independent data from Sentinel-2 (S2) optical images and high water marks (HWM) that are available and relevant. The latter is a collaborative dataset of high water marks[2] contributed by occasional observers in the floodplain is used as independent data for validation purposes.

The remainder of the paper is organized as follows. Section 2 gathers the material and data used in this study. Subsection 2.1 presents the hydrodynamic numerical solver TELEMAC-2D[3] (T2D). Its implementation on the Garonne Marmandaise catchment for the representation of flooding along with the associated sources of uncertainties are described in subsection 2.2. The reference flood events and the in-situ and remote sensing data that are used for assimilation and validation purposes are then described in subsection 2.3. The EnKF algorithm for dual state-parameter correction is presented in section 3. The description of the control vector, the forecast and analysis steps are proposed in subsection 3.1, subsection 3.2 and subsection 3.3, respectively. Section 4 provides a thorough description of the experimental settings and assessment metrics for the DA strategy. The experimental settings for DA cycling and observation errors are gathered along with the metrics that are used to assess the performance of the DA simulations with respect to assimilated and independent data. Experimental results are presented in Section 5, first in the framework of OSSE (subsection 5.1), then for real events (subsection 5.2). The merits of assimilating RS data for simulation evaluation in the river bed and the floodplain are highlighted in the control and the observation space for OSSE and real experiments, especially at the flood peak and during the flood recess. Conclusions, limitations, and perspectives are given in Section 6.

2 Methodology

2.1 Hydrodynamic model

The non-conservative form of the SWE are written in terms of water level (denoted by H [m]) and horizontal components of velocity (denoted by u and v [m.s^{-1}]). They express mass and momentum conservation averaged in the vertical dimension, assuming that (i) vertical pressure gradients are hydrostatic, (ii) horizontal pressure gradients are due to displacement of the free surface, and that (iii) horizontal length scale is significantly greater than the vertical scale. The SWE read:

$$\frac{\partial H}{\partial t} + \frac{\partial}{\partial x}(Hu) + \frac{\partial}{\partial y}(Hv) = 0 \tag{1}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -g\frac{\partial Z}{\partial x} + F_x + \frac{1}{H}div\left(H\nu_e\overrightarrow{grad}(u)\right) \tag{2}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -g\frac{\partial Z}{\partial y} + F_y + \frac{1}{H}div\left(H\nu_e\overrightarrow{grad}(v)\right) \tag{3}$$

where Z [m NGF69] is the water surface elevation and ν_e [m^2.s^{-1}] is the water diffusion coefficient. The water level $H = Z - Z_b$ is computed from Z with Z_b [m NGF69] the bottom elevation. In the following, the water surface elevation is shortened as WSE and the water level as WL. t stands for time and g is the gravitational acceleration constant. div and $\overrightarrow{grad}$ are respectively the divergence and gradient operators.

In addition, F_x and F_y [m.s^{-2}] are the horizontal components of external forces (friction, wind and atmospheric forces), defined as follows:

$$F_x = -\frac{g}{K_s^2}\frac{u\sqrt{u^2+v^2}}{H^{4/3}} - \frac{1}{\rho_w}\frac{\partial P_{atm}}{\partial x} + \frac{1}{H}\frac{\rho_{air}}{\rho_w}\,C_d\,U_{w,x}\sqrt{U_{w,x}^2+U_{w,y}^2} \tag{4}$$

$$F_y = -\frac{g}{K_s^2}\frac{v\sqrt{u^2+v^2}}{H^{4/3}} - \frac{1}{\rho_w}\frac{\partial P_{atm}}{\partial y} + \frac{1}{H}\frac{\rho_{air}}{\rho_w}\,C_d\,U_{w,y}\sqrt{U_{w,x}^2+U_{w,y}^2} \tag{5}$$

[2] https://www.reperesdecrues.developpement-durable.gouv.fr/
[3] www.opentelemac.org

where ρ_w/ρ_{air} [kg.m^{-3}] is the water/air density ratio, P_{atm} [Pa] is the atmospheric pressure, $U_{w,x}$ and $U_{w,y}$ [m.s^{-1}] are the horizontal wind velocity components, C_d [-] is the wind drag coefficient that relates the free surface wind to the shear stress, and lastly, K_s [m$^{1/3}$.s^{-1}] is the river bed and floodplain friction coefficient, using the Strickler formulation (Gauckler, 1867).

In order to solve Eq. (1)-(3), initial conditions $\{H(x,y,t=0) = H_0(x,y);\ u(x,y,t=0) = u_0(x,y);\ v(x,y,t=0) = v_0(x,y)\}$ are provided, and boundary conditions (BC) are described with a time-varying hydrogram upstream and a rating curve downstream. The Strickler coefficient is prescribed uniformly over defined subdomains, and calibrated according to the observing network. The hydrodynamic numerical model T2D is used to simulate and predict WSE and velocity from which the flood risk can be assessed. T2D solves the SWE derived from Navier-Stokes equations with an explicit first-order time integration scheme, a finite-element scheme and an iterative conjugate gradient method (Hervouet, 2007). The results are obtained at each point of the mesh mapped onto the catchment topography.

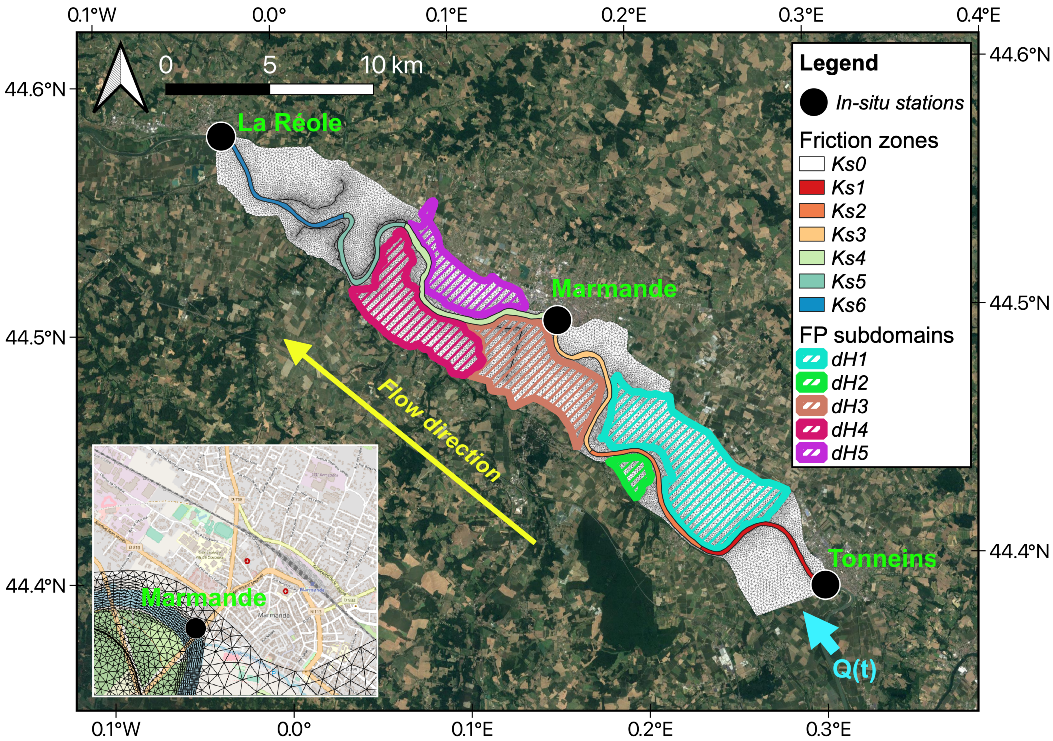


Figure 1: T2D Garonne Marmandaise domain. The VigiCrue observing stations are indicated as black circles. The different river friction zones are indicated as colored segments of the river bed. The floodplain is divided into five subdomains that are hatched in different colors. The inset figure at the bottom left corner magnifies the urban area of Marmande nearby its namesake gauging station.

2.2 Study area and description of the uncertainties

The study area is the Garonne Marmandaise catchment (southwest of France) which extends over a 50-km reach of the Garonne River between Tonneins and La Réole (Figure 1). Since the 19th century, it has been equipped with infrastructures. As such, a system of dykes and weirs had been progressively built to protect floodplains from flooding events, such as the historic flood of 1875, and to manage submersion and flood retention areas. Observing stations operated by the VigiCrue network[4] are located at Tonneins, Marmande, and La Réole (indicated as black circles in Figure 1, providing water-level measurements every 15 minutes.

A T2D model was developed and calibrated over this catchment, which was built on a mesh of 41,000 nodes using bathymetric cross-sectional profiles and topographic data (Besnard & Goutal, 2011). The topography of the catchment was generated using IGN (French National Mapping Agency) maps as well as aerial photographs for photogrammetric reconstruction (Besnard & Goutal, 2011). A local rating curve at Tonneins is used to translate the observed WL into a discharge time-series. Discharge time-series $Q(t)$ are then applied as forcing over the whole upstream interface (cyan arrow in Figure 1), including both river bed and floodplain boundary cells. This modeling strategy was implemented to allow for a cold start of the model with any inflow value. However, it prompts an over-flooding of the upstream first meander, until the water returns to the river bed. The downstream BC at La Réole is described with a local rating curve. In the following, both upstream and downstream areas are excluded from the computation of assessment metrics to limit the impact of the choice of the BC strategy and topographic errors on the results. Over the simulation domain, the friction coefficient is defined over seven zones, including six segments from K_{s_1} to K_{s_6} for the river bed and one K_{s_0} for the entire floodplain, as illustrated in Figure 1 with solid colored segments of the rived bed and white background color for the floodplain. A priori values for friction in the river bed are set from a calibration process using in-situ WL observations at Tonneins, Marmande and La Réole for selected set of past flood events, summarized by Table 1. The description of the friction coefficients is highly prone to uncertainties related to the zoning assumption, the calibration procedure, and the set of calibration events. In the following, these coefficients are considered as random variables with a gaussian Probability Density Function (PDF) with mean $\mathbf{x}_0$ and standard deviation $\sigma_{\mathbf{x}}$ estimated from the calibration process (Table 1). The a priori values are further improved with the DA strategy. The uncertainty in the upstream BC is also taken into account. Indeed, the limited number of in-situ observations yields errors in the formulation of the rating curve that is used to translate the observed WL into discharge, especially for high flow. Thus, a multiplicative factor μ on the time-dependent discharge time-series is considered as a random variable with a gaussian PDF centered at 1. Lastly, in order to account for the evapotranspiration, ground infiltration and rainfall processes that are lacking in the T2D Garonne model, a state correction is implemented in the floodplain. The floodplain is divided in five subdomains based on the description of the storage areas (Besnard & Goutal, 2011) and the dyke system of the catchment. A uniform WL correction δH_k with $k \in [1, 5]$ over each subdomain is added to the control vector. Each δH_k is considered as a random variable with a zero-mean Gaussian with a standard deviation set to 0.25 [m]. The calibrated friction coefficient values, and the default values of 1 for μ and 0 for δH_k are used as setting for the free run experiment further denoted as FR.

2.3 Flood events and Observations

Two significant flood events having occurred in December 2019 and January-February 2021 are studied in this research work. In-situ WL measured every 15 minutes at Tonneins, Marmande and La Réole are shown in Figure 2a and Figure 2b, respectively. A simulation period of 25 days was selected around the flood peak for each event in order to properly capture the flood and the recess periods. All of the time-varying plots in this article are made in local time (UTC +01:00).

[4] `https://www.vigicrues.gouv.fr/`

Table 1: Characteristics of the Gaussian PDF for friction coefficients, multiplicative coefficient for inflow and water level correction in the subdomain of the floodplain.

Variable	Unit	Calibrated/ default values $\mathbf{x}_0$	Standard deviation $\sigma_{\mathbf{x}}$	95% confidence interval
K_{s_0}	$m^{1/3}.s^{-1}$	17	0.85	17 ± 1.67
K_{s_1}	$m^{1/3}.s^{-1}$	45	2.25	45 ± 4.41
K_{s_2}, K_{s_3}	$m^{1/3}.s^{-1}$	38	1.9	38 ± 3.72
$K_{s_4}, K_{s_5}, K_{s_6}$	$m^{1/3}.s^{-1}$	40	2.0	40 ± 3.92
μ	-	1	0.06	1 ± 0.0136
δH_k ($k \in [1,5]$)	m	0	0.25	0 ± 0.0566

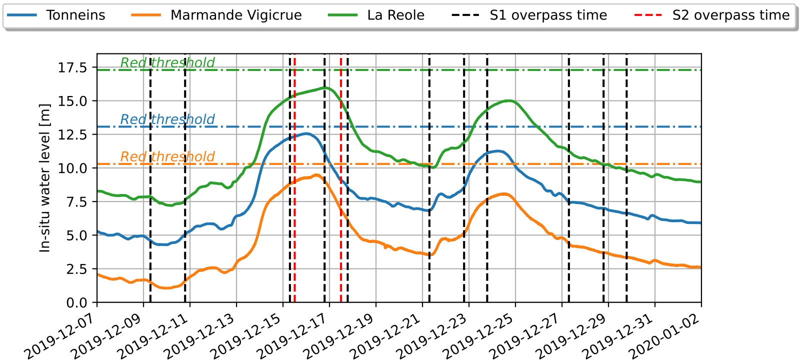
Table 2: General information on the studied flood events.

Event	First date	Last date	Nb of S1 images	Nb of usable S2 images	Nb of HWM
2019	2019-12-08	2020-01-02	11	2	120
2021	2021-01-16	2021-02-10	12	0	178

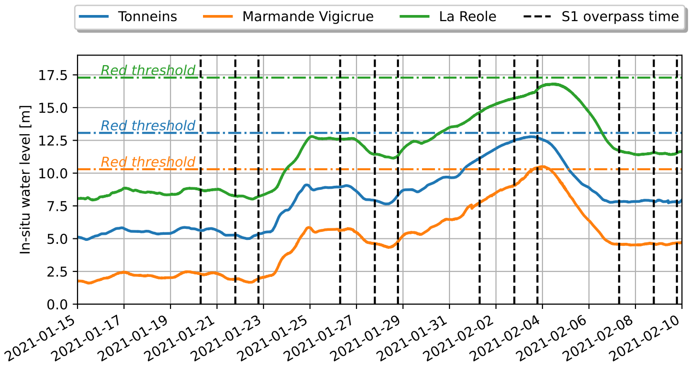
Sentinel-1 (S1) is the first satellite series of the Copernicus program (Torres et al., 2012). This SAR system works at C-band, with a central frequency of 5.405 GHz. The Interferometric Wide (IW) mode with 250-km-wide swath used in this study offers a ground resolution of approximately 20 × 22 m; this product is then resampled, reprojected and distributed at 10 × 10 m for the Ground Range Detected (GRD) products. In order to improve the revisit time, Sentinel-1 works as a constellation of two polar-orbiting identical satellites Sentinel-1A launched on 2014-04-03 and Sentinel-1B on 2016-04-26, allowing a six-day revisit time. The S1 GRD IW products are leveraged as the predominant data source to produce binary water maps using Machine Learning algorithms developed by CNES and CLS in the framework of the FloodML project (Huang et al., 2020; Kettig et al., 2021). The specifications of the flood extent mapping method applied to S1 images are detailed in (Nguyen et al., 2022).

Similarly, Sentinel-2 (S2) mission comprises a constellation of two multispectral instrument satellites, Sentinel-2A launched on 2015-06-23 and Sentinel-2B on 2017-03-07. They are placed in the same sun-synchronous orbit, phased at 180 degrees to each other. They provide 290-km swath width and a high revisit time (10 days at the equator with one satellite, and 5 days with 2 satellites under cloud-free conditions which might result in 2-3 days revisit at mid-latitudes due to orbit overlapping). In order to perform the flood extent mapping on S2 images, an extraction of features based on the Normalized Difference Vegetation Index (NDVI) (Huang et al., 2021) and the Modified Normalized Difference Water Index (MNDWI) (Xu, 2006) was carried out. They are then used as the inputs for the implemented RF classifier (Kettig et al., 2021).

The double-peak flood event in 2019 was observed by eleven S1 SAR images, provided by the constellation of S1-A and S1-B ascending and descending orbits, and two S2 images with acceptable cloud cover condition. The 2021 flood event is composed of a single peak (but with a stronger flow than that of 2019) and was observed by 12 S1 images. The flood peak was reached on 2021-02-04 and it exceeded the highest threshold level for flood risk at Marmande, set out by the French national flood forecasting center (SCHAPI) in collaboration with the departmental prefect (see Figure 2b). It should



(a) 2019 flood event



(b) 2021 flood event

Figure 2: Water level H time-series for (a) 2019 flood event, and (b) 2021 flood event, at Tonneins (blue), Marmande (orange) and La Réole (green). S1 and S2 overpass times are indicated as vertical dashed lines, respectively in black and in red. The red thresholds for the WL associated with the highest level of flooding risk at each observing stations are shown as horizontal dash-dotted lines with the same color.

be noted that for the S1 images from the ascending orbit 132, a small part of the downstream area (including La Réole) is a no-data area as it is out of range from the acquisition. As aforementioned, two S2 optical images are available for 2019 near the first flood peak at 2019-12-15 12:05 and 2019-12-17 11:54 thus providing independent data for validation, with a cloud cover percentage of 40.58% and 11.28%, respectively. Due to high cloud cover, none of the S2 images acquired during the 2021 provides reliable observations. The SAR S1 image acquired on 2021-02-02 18:55 and the derived flood extent map by FloodML for the same date are shown respectively in Figure 3a with grayscale (from dark to bright) indicating the backscatter values (from low to high), and in Figure 3b where wet pixels are indicated in white. The simulated flood extent for the free run (FR) introduced in Sect. 2.2 using the calibrated and default parameter values (Table 1) is shown in green in Figure 3b. The ratio between the number of wet pixels and the total number of pixels, named WSR, is formulated for each of the five subdomains of the floodplain indicated in hatched colored areas in Fig 1. WSR is further considered as the observation for the DA strategy. In order to account for mis-detection of wet pixels in vegetated regions, exclusion layers were identified from four land cover classes (deciduous and coniferous forests, orchards, and diffused built-up areas) of the IOTA2 land cover map (Inglada et al., 2017) produced on the French territory. These highly vegetated areas shown in red in Figure 3b were excluded from the comparison between the model flood extent and the RS-derived observation, as well as from the WSR and further assessment computation. A zoomed-in view of Figure 3c over the zones 1, 2 and 3 of the floodplain is provided in Figure 3b with the same color code. Figure 3d complements the zoomed-in view with the effective areas (color-coded according to Figure 1) for the computation of WSR, taking into account the excluded pixels (in red). Lastly, as aforementioned, there are 120 HWM observations available for the 2019 flood event whereas 178 observations were collected in the aftermath of the 2021 event (as it was more severe).

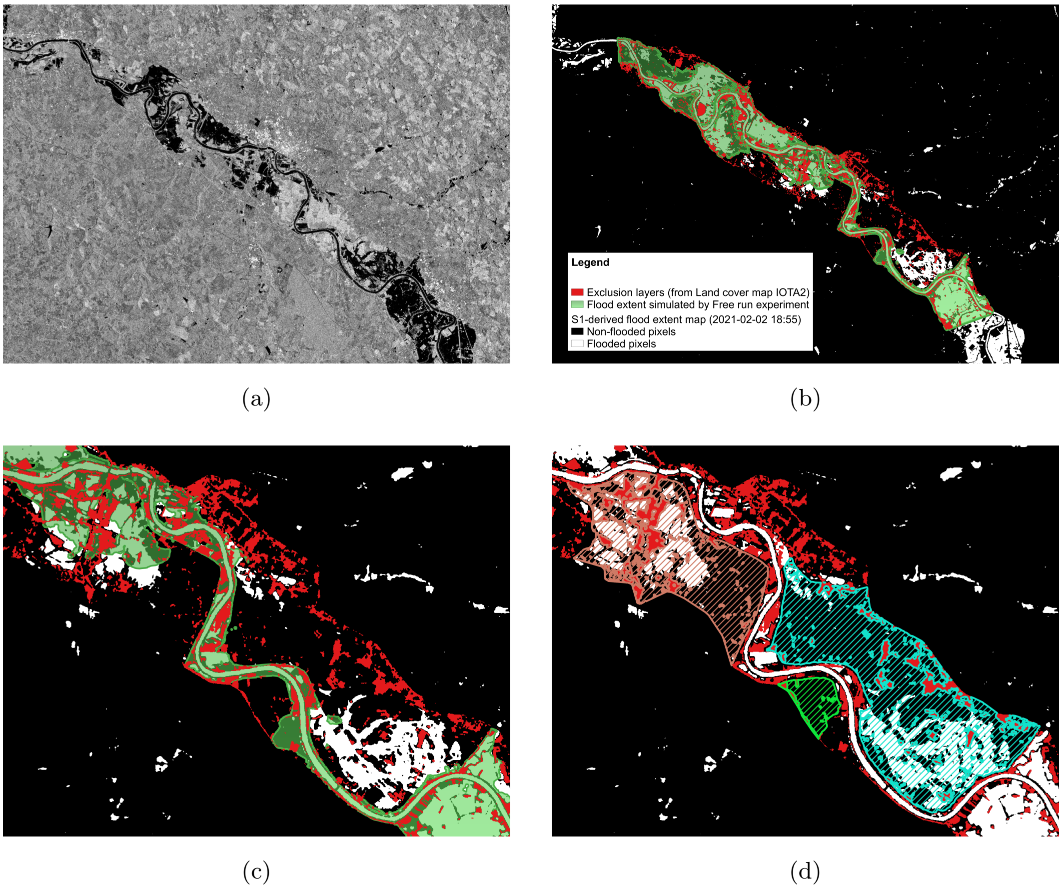


Figure 3: 2D flood extent observation derived from S1 data. (a) SAR S1 (VV polarization) image acquired on 2021-02-02 18:55:00, (b) the S1-derived flood extent binary map with red regions representing the areas to be excluded from the comparison between flood extents, green outline: the free run using calibrated and default parameter values, (c) zoomed-in view of sub-figure 3b and (d) zoomed-in view of the effective areas for the computation of WSR on the zones 1-3 taking into account exclusion layers, overlaid on the flood extent map.

3 Data Assimilation

3.1 Description of the control vector

The implemented DA algorithm consists in a cycled stochastic EnKF, where the control vector $\mathbf{x}$ is composed of the seven friction coefficients K_{s_k} with $k \in [0, 6]$, one multiplicative parameter μ to modify the time-varying upstream BC Q(t), and five state corrective variables δH_k with $k \in [1, 5]$ over the floodplain zones. Altogether these $n =$ 13 parameters are assumed to be constant over a DA cycle, yet their evolution in time is made possible by DA between cycles. The DA cycle c covers a time window, denoted by $W_c = [t_{start}, t_{end}]$ of length $T = 18$ hours over which $n_{obs,c}$ observations are assimilated. The cycling of the DA algorithms consists in sliding the time window of a period $T_{shift} = 6$ hours so that the cycles c and $c + 1$ overlap over 12 hours.

It could be argued that the DA algorithm is more a smoother than a filter as it operates over a sliding time window. Yet, as the control vector is composed of model parameters and corrections that are assumed constant over the assimilation window (as opposed to the model state), the smoothing resumes to a filtering. The EnKF algorithm relies on the propagation of N_e members with perturbed values of $\mathbf{x}$, denoted by $\mathbf{x}^i$. The forecast values of $\mathbf{x}^i$ are denoted by $\mathbf{x}_c^{f,i}$ (superscript index f stands for "forecast"), where $i \in [1, N_e]$ is the ensemble member counter.

3.2 Description of the EnKF forecast step

The EnKF forecast step consists in the propagation in time, over W_c, of the control and model state vectors. The EnKF is applied to model parameters that, by definition, do not evolve in time over the cycle c. The absence of propagation model for the control vector implies that the forecast for the control vector at cycle c should remain equal to its analysis at cycle $c - 1$. Yet, in order to avoid ensemble collapse, artificial dispersion is introduced with the addition of perturbations $\boldsymbol{\theta}$ to a global value $\mathbf{x}_{c-1}^{a,glo}$ issued from the previous cycle. For the friction coefficients K_{s_k} with $k \in [0, 6]$, and the forcing parameter μ, $\mathbf{x}_{c-1}^{a,glo}$ is chosen as the mean of the analysis from the previous cycle $\overline{\mathbf{x}_{c-1}^a}$ (superscript index a stands for "analysis" and $\overline{\bullet}$ stands for the average over the ensemble). For the floodplain state corrections δH_k with $k \in [1, 5]$, $\mathbf{x}_{c-1}^{a,glo}$ is set to 0. The forecast step thus reads:

$$\mathbf{x}_c^{f,i} = \begin{cases} \mathbf{x}_0 + \boldsymbol{\theta}_1^i & \text{if } c = 1 \\ \mathbf{x}_{c-1}^{a,glo} + \boldsymbol{\theta}_c^i & \text{if } c > 1 \end{cases} \tag{6}$$

with

$$\mathbf{x}_{c-1}^{a,glo} = \left[\overline{(K_{s_k})_{c-1}^a} \text{ with } k \in [0,6],\ \overline{\mu_{c-1}^a},\ 0 \text{ for each } \delta H_k \text{ with } k \in [1,5]\right], \tag{7}$$

and

$$\boldsymbol{\theta}_c^i \sim \mathcal{N}\left(\mathbf{0}, (\sigma_c^i)^2\right), \tag{8}$$

where

$$\sigma_c^i = \begin{cases} \sigma_{\mathbf{x}} & \text{if } c = 1, \\ \lambda\sqrt{\frac{1}{N_e-1}\sum_{i=1}^{N_e}(\mathbf{x}_{c-1}^{a,i} - \overline{\mathbf{x}_{c-1}^a})^2} + (1-\lambda)\sigma_{\mathbf{x}} & \text{if } c > 1. \end{cases} \tag{9}$$

For the first cycle, the perturbed friction, upstream forcing coefficient values and floodplain state perturbations are drawn within the PDFs described in Table 1. For the next cycles, the set of coefficients issued from the mean analysis at the previous cycle $c-1$ is further dispersed by additive perturbations $\boldsymbol{\theta}$ (Eq. (8)) drawn from the Gaussian distribution with zero mean and a standard deviation obtained from the linear combination of the standard deviation of the analysis at $c-1$ and $\sigma_{\mathbf{x}}$ described in Table 1.

The two terms are weighted by the hyperparameter λ (Eq. (9)). This technique is an advanced alternative to anomalies inflation for avoiding the well-known ensemble collapse, better suited for heterogeneous control of parameters. The combined update of the variance for the re-sampling of the parameters allows to preserve part of the information from the background statistical description that may differ amongst the parameters and over time while also inheriting analyzed variance from the previous cycle. In the following implementation, λ is set to 0.3. This tuning was chosen after the analysis of the ensemble spread in the control space along the DA cycles.

The background hydraulic state, associated with each member of the ensemble of inputs, denoted by $\mathbf{s}_c^{f,i}$, results from the integration of the hydrodynamic model $\mathcal{M}_c$: $\mathbb{R}^n \to \mathbb{R}^m$ from the control space to the model state (of dimension m) over W_c:

$$\mathbf{s}_c^{f,i} = \mathcal{M}_c(\mathbf{s}_{c-1}^{a,i}, \mathbf{x}_c^{f,i}). \tag{10}$$

The initial condition for $\mathcal{M}_c$ at t_{start} is provided by a user-defined restart file for the first cycle. For the following cycles, it takes in a full restart $\mathbf{s}_{c-1}^{a,i}$ saved from the analysis run of the previous cycle $\mathbf{s}_{c-1}^{a,i}$ at time $t_{start} + T_{shift}$. Note that in order to avoid inconsistencies between the state and the analysed set of parameters at t_{start}, a short spin-up integration is run on the 3 hours preceding t_{start}. It should be noted that the perturbations δH_k ($k \in [1,5]$) (Eq. 6) are evenly distributed on the time steps in $[t_{start} - 3h, t_{start} + T_{shift}]$ and added to the simulated WL field, while enforcing that the resulting WL at each pixel remains non-negative.

The control vector equivalent in the observation space for each member, denoted by $\mathbf{y}_c^{f,i}$, stems from:

$$\mathbf{y}_c^{f,i} = \mathcal{H}_c(\mathbf{s}_c^{f,i}) \tag{11}$$

where $\mathcal{H}_c$: $\mathbb{R}^m \to \mathbb{R}^{n_{obs}}$ is the observation operator from the model state space to the observation space (of dimension n_{obs}) that selects, extracts and eventually interpolates model outputs at times and locations of the observation vector $\mathbf{y}_c^o$ over W_c. The observation vector here gathers observations of different types (in-situ WL and WSR), at different times over W_c. The observation operator $\mathcal{H}_c$ is thus composed of two operations that are applied separately to the T2D hydraulic state. On the one hand, a selection operator that extracts the WL at time and locations of the in-situ observations. On the other hand, a flood mask generator that applies a threshold of 5 cm on the WL simulated field at S1 overpass times, in order to identify the wet/dry pixel mask, then computes WSR observations by counting the number of wet pixels in each subdomain of the floodplain. It should be noted that, in the following, the observation operator regarding the in-situ observations, also includes a bias removal step to take into account a systematic model error. Eq. (11) thus reads

$$\mathbf{y}_c^{f,i} = \mathcal{H}_c(\mathbf{s}_c^{f,i}) - \mathbf{y}^{\text{bias}} \tag{12}$$

where $\mathbf{y}^{\text{bias}}$ is an a priori knowledge of the model-observation bias. Such a bias was diagnosed and estimated during the 24-hour quasi-stationary non-overflowing period of 2021-01-15. This yields $\mathbf{y}^{\text{bias}}_{\text{Tonneins}} = 0.72$, $\mathbf{y}^{\text{bias}}_{\text{Marmande}} = 0.40$, and $\mathbf{y}^{\text{bias}}_{\text{LaRéole}} = -0.24$ m (Nguyen et al., 2021, 2022).

3.3 Description of the EnKF analysis step

The EnKF analysis step stands in the update of the control and model state vectors. When applying a stochastic EnKF (Asch et al., 2016), the observation vector $\mathbf{y}^{o,i}$ is perturbed, and an ensemble of observations $\mathbf{y}_c^{o,i}$ ($i \in [1, N_e]$) is generated:

$$\mathbf{y}_c^{o,i} = \mathbf{y}_c^o + \boldsymbol{\epsilon}_c \;\text{ with }\; \boldsymbol{\epsilon}_c \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_c). \tag{13}$$

$\mathbf{R}_c = {\sigma_{obs}}^2 \mathbf{I}_{n_{obs}}$ is the observation error covariance matrix ($\mathbf{I}_{n_{obs}}$ is the $n_{obs} \times n_{obs}$ identity matrix). $\mathbf{R}_c$ is assumed to be diagonal, of standard deviation σ_{obs} , as the observation errors are assumed to be uncorrelated, Gaussian with a standard deviation proportional to the observation $\sigma_{obs,c} = \tau \mathbf{y}_c^o$. This stochastic perturbation was not set up for the WSR observations in $\mathbf{y}_c^{o,i}$ in order to avoid values beyond the physical range $[0, 1]$ for the ratio.

The innovation vector over W_c is the difference between the perturbed observation vector $\mathbf{y}_c^{o,i}$ and the model equivalent $\mathbf{y}_c^{f,i}$ from Eq. (11) and Eq. (13). It is weighted by the Kalman gain matrix $\mathbf{K}_c$ and then added as a correction to the background control vector $\mathbf{x}_c^{f,i}$, so that the analysis control vector $\mathbf{x}_c^{a,i}$ reads:

$$\mathbf{x}_c^{a,i} = \mathbf{x}_c^{f,i} + \mathbf{K}_c \left(\mathbf{y}_c^{o,i} - \mathbf{y}_c^{f,i}\right), \tag{14}$$

with

$$\mathbf{x}_c^{a,i} = \left[(K_{s_k})_c^{a,i} \text{ with } k \in [0,6],\ \mu_c^{a,i},\ (\delta H_k)_c^{a,i} \text{ with } k \in [1,5]\right]. \tag{15}$$

The Kalman gain reads:

$$\mathbf{K}_c = \mathbf{P}_c^{\mathbf{x},\mathbf{y}} [\mathbf{P}_c^{\mathbf{y},\mathbf{y}} + \mathbf{R}_c]^{-1}. \tag{16}$$

$\mathbf{P}_c^{\mathbf{y},\mathbf{y}}$ is the covariance matrix of the error in the background state equivalent in the observation space $\mathbf{y}_c^f$. $\mathbf{P}_c^{\mathbf{x},\mathbf{y}}$ is the covariance matrix between the error in the control vector and the error in $\mathbf{y}_c^f$. Both matrices are stochastically estimated within the ensemble:

$$\mathbf{P}_c^{\mathbf{x},\mathbf{y}} = \frac{1}{N_e} \mathbf{X}_c^T \mathbf{Y}_c \in \mathbb{R}^{n \times n_{obs}} \tag{17}$$

$$\mathbf{P}_c^{\mathbf{y},\mathbf{y}} = \frac{1}{N_e} \mathbf{Y}_c^T \mathbf{Y}_c \in \mathbb{R}^{n_{obs} \times n_{obs}} \tag{18}$$

with:

$$\mathbf{X}_c = \left[\mathbf{x}_c^{f,1} - \overline{\mathbf{x}_c^f}, \cdots, \mathbf{x}_c^{f,N_e} - \overline{\mathbf{x}_c^f}\right] \in \mathbb{R}^{n \times N_e} \tag{19}$$

$$\mathbf{Y}_c = \left[\mathbf{y}_c^{f,1} - \overline{\mathbf{y}_\mathbf{c}^\mathbf{f}}, \cdots,\ \mathbf{y}_c^{f,N_e} - \overline{\mathbf{y}_c^f}\right] \in \mathbb{R}^{n_{obs} \times N_e} \tag{20}$$

and

$$\overline{\mathbf{x}_c^f} = \frac{1}{N_e} \sum_{i=1}^{N_e} \mathbf{x}_c^{f,i} \in \mathbb{R}^n \tag{21}$$

$$\overline{\mathbf{y}_c^f} = \frac{1}{N_e} \sum_{i=1}^{N_e} \mathbf{y}_c^{f,i} \in \mathbb{R}^{n_{obs}}. \tag{22}$$

It should be noted that a localization on $\mathbf{P}_c^{\mathbf{x},\mathbf{y}}$ was implemented so that only the WSR observations are used to account for errors in the floodplain state through the estimation of $\delta H_k^{a,i}$. Consequently, the correction of the hydraulic state in the floodplain is only activated when WSR are available over the assimilation window. This prevents from equifinality issues due to the size of the ensemble. Indeed, the stochastic approximation in Eq. (17) and Eq. (18) could infer some artificial sensitivity of the hydraulic state in the floodplain with respect to the friction coefficients in the river bed.

The analyzed hydrodynamic state, associated with each analyzed control vector $\mathbf{x}_c^{a,i}$ is denoted by $\mathbf{s}_c^{a,i}$. It results from the integration of the hydrodynamic model $\mathcal{M}_c$ with the updated friction coefficients $(K_{s_k})_c^{a,i}$, the upstream forcing Q_{up} using $\mu_c^{a,i}$ and the state correction in the floodplain $\delta H_k^{a,i}$ over W_c, starting from the same initial condition as each background simulation within the ensemble. In order to preserve a smooth WL field, the mean of the analysis for $\overline{\delta H_k^a}$ computed within the ensemble is considered (Eq. (25)).

$$\mathbf{s}_c^{a,i} = \mathcal{M}_c(\mathbf{s}_{c-1}^{a,i}, \tilde{\mathbf{x}}_c^{a,i}), \tag{23}$$

with

$$\tilde{\mathbf{x}}_c^{a,i} = \left[(K_{s_k})_c^{a,i} \text{ with } k \in [0,6],\ \mu_c^{a,i},\ \overline{(\delta H_k)_c^a} \text{ with } k \in [1,5] \right], \tag{24}$$

and

$$\overline{(\delta H_k)_c^a} = \frac{1}{N_e} \sum_{i=1}^{N_e} (\delta H_k)_c^{a,i} \text{ with } k \in [1,5]. \tag{25}$$

4 Experimental settings

4.1 Specifications of Observing System Simulation Experiments

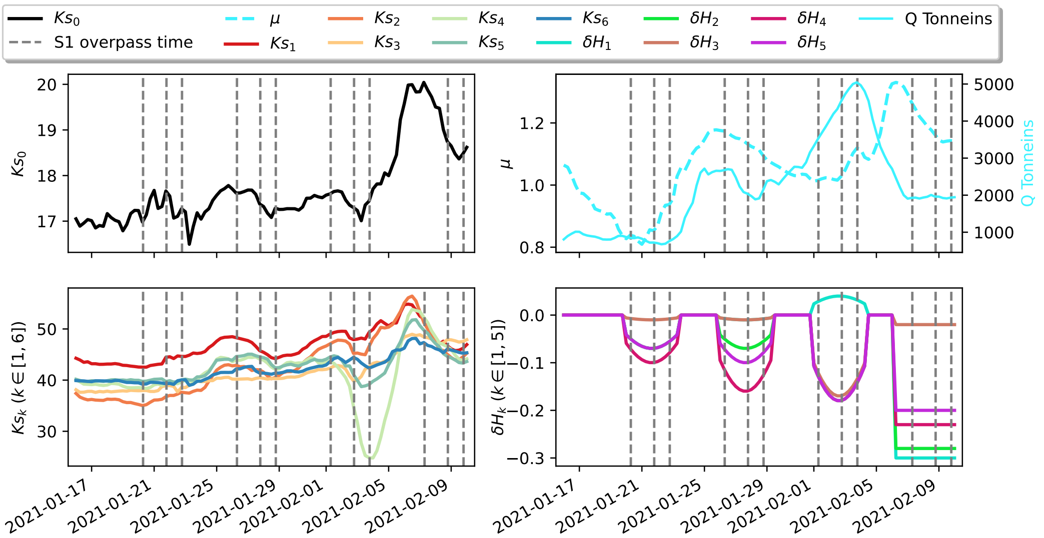


Figure 4: True values of the control vector for the reference simulation over the synthetical 2021 event in OSSE. Top left: K_{s_0}, bottom left: K_{s_k} with $k \in [1,6]$, top right: μ (left y-axis, dashed cyan curve) and $Q(t)$ (right y-axis, solid cyan curve), bottom right: δH_k with $k \in [1,5]$. These color codes are identical to those of Figure 1. The S1 overpass times are indicated as vertical black dashed lines.

The framework of an OSSE is based on a deterministic simulation with a selected set of time-varying parameters over the flood event, as shown in Figure 4. This reference simulation is further denoted as true. In the present work, the true friction parameter K_{s_k} with $k \in [0,6]$, are set from the results of a previous DA experiments on the real 2021 flood event. In Figure 4, the true friction coefficient for the floodplain K_{s_0} is plotted in black on the top left panel and the true friction coefficients K_{s_k} with $k \in [1,6]$ for the river bed are plotted on the bottom left panel. The time-series discharge for the 2021 event is used as the upstream BC for the OSSE experiment. The true multiplicative correcting factor μ for the inflow is also issued from a previous DA analysis, and is added a cosine function as perturbation. It is plotted as a dashed cyan curve on the top right panel in Figure 4 (left y-axis) along with the inflow BC at Tonneins, represented by a solid cyan curve (right y-axis). The state correction true values were set up with negative cosine curves for the three first groups of S1 observations (from the beginning of the event until the flood peak), and a constant water removal of -18 cm over the flood recession period. They are shown on the bottom right panel in Figure 4. For the sake of consistency, the color codes for K_{s_k} with $k \in [1,6]$ and for δH_k with $k \in [1,5]$ are identical to their effective areas depicted in Figure 1.

The true simulation is used to provide synthetic observations using the observation operator $\mathcal{H}_c$ from Eq. (11) applied at the in-situ and S1 observation times from the real 2021 event. This stands in the extraction of the true simulated WL values at all observation times and locations, first to generate synthetical in-situ observations, and second to extract the wet/dry pixels for WSR computation. Thus this experiment is further denoted as *synthetical 2021 event*. These synthetical in-situ and WSR observations are then assimilated in a DA experiment, with a priori settings that differ from the truth. The OSSE experiments aim at assessing the performance of the DA method involving both types of observations (in-situ and WSR), especially its capacity to retrieve the true parameters (forcing data, friction coefficients and state correction).

4.2 Experimental setup and assessment metrics

In both OSSE and real event experiments, one free run FR (without assimilation), and three DA experiments were carried out with different configurations regarding the types of observations that are assimilated and the active components of the control vector, as detailed in Table 3. Two types of observations are considered: (i) in-situ WL observations at the three VigiCrue stations Tonneins, Marmande and La Réole every 15 minutes, (ii) WSR values computed over the five floodplain zones at S1 overpass times. Then, two options of control vector are involved, one with all six friction coefficients and the inflow multiplicative coefficient, whereas the other one is extended with the water state correction in the floodplain. With these configurations, three experiments are called IDA, IWDA and IHDA. IDA experiment only assimilates in-situ WL observations (synthetical observations in the context of OSSE) and the control vector is limited to friction coefficients K_{s_k} with $k \in [0, 6]$ and the inflow multiplicative coefficient μ. IWDA experiment assimilates in-situ WL and WSR observations (synthetical in the context of OSSE) with the same control vector as IDA. IHDA has an extended control vector that also includes δH_k with $k \in [1, 5]$, it assimilates the same in-situ WL and WSR observations (synthetical in the context of OSSE) as IWDA. For the DA experiments, the proportionality coefficient used to specify the observation error τ (cf. subsection 3.3) for in-situ data is fixed to 15%, meaning that σ_{obs} amounts to 15% of the observation value, whereas the value of τ for WSR data varies from 10% to 20% depending on how early the observation time is within the 18-hour assimilation window. All DA experiments where carried out using $N_e = 75$ members. In the following, the subscript $OSSE$ is used in the experiment name to distinguish the OSSE from the real modes.

The metrics employed for 1D and 2D assessment are formulated with respect to the observations that are synthetical in the context of OSSE, or with respect to the real observations from the VigiCrue gauge stations (for the in-situ WL) and from S1/S2 images (for 2D flood extent maps and derived WSR).

Table 3: Summary of the Free Run and DA experiment settings.

Exp. name	DA	Assimilated observations	Nb of members N_e	Control variables
FR	No	-	1	-
IDA$_{(OSSE)}$/IDA	Yes	In-situ WL	75	$K_{s_{[0:6]}}, \mu$
IWDA$_{(OSSE)}$/IWDA	Yes	In-situ WL and WSR	75	$K_{s_{[0:6]}}, \mu$
IHDA$_{(OSSE)}$/IHDA	Yes	In-situ WL and WSR	75	$K_{s_{[0:6]}}, \mu, \delta H_{[1:5]}$

4.2.1 1D metrics for water level time-series assessment

The quality of the simulated WL, noted H^m, is assessed with respect to in-situ observed WL, noted H^o, computing the root-mean-square error (RMSE) between the simulated and the observed WL time-series, sampled at observation times, along the assimilation windows for the entire flood event:

$$\text{RMSE} = \sqrt{\frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} (H_i^m - H_i^o)^2} \tag{26}$$

In the case of OSSE, the RMSE is also formulated with respect to the control parameters, computing the difference between their true value and the DA analysis.

4.2.2 2D metrics for flood extent assessment

The simulated flood extent maps are generated from the T2D simulated WL 2D field, by applying a threshold of 5 cm below which the pixel is considered as dry and above which it is considered as wet. The T2D WL output field is first projected onto the regular grid of the S1 image (ground sampling distance: 10 × 10 m) so as to allow for a straightforward comparison between observed and simulated flood extent. In the case of OSSE, the water flood maps from the reference simulation are used instead of the observed flood extent maps from S1 images.

The metrics to compare the simulated and the observed flood extents are the Critical Success Index (CSI) and Cohen's kappa index (κ). CSI considers the FloodML flood extent maps as the reference observed flood maps (ground truth) based on which the T2D simulated flood extent maps are evaluated, whereas the objective of κ index is used to measure the agreement between the two flood extent estimators. The formulation of these indices relies on the count of pixels following one of four outcomes that constitute a contingency map: True Positives (TP) and True Negatives (TN), respectively, are the number of pixels correctly predicted as flooded and correctly identified as non-flooded, False Positives (FP) or *over-prediction* is the number of non-flooded pixels incorrectly predicted as flooded, and False Negatives (FN) or *under-prediction* is the number of missed flooded pixels. Based on these counts, the CSI and κ indices are computed as follows:

$$\text{CSI} = \frac{TP}{TP + FP + FN}, \tag{27}$$

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{28}$$

where p_o is the observed proportionate agreement and p_e is the probability of a random agreement, defined as follows:

$$\begin{aligned} p_o = & \quad \frac{TP + TN}{TP + FP + FN + TN}, \\ p_e = & \quad \frac{TP + FN}{TP + FP + FN + TN} \times \frac{TP + FP}{TP + FP + FN + TN}. \end{aligned}$$

These two metrics range from 0% when there is no common area (i.e. no agreement) between the simulated and the observed flood extents, and reach their highest value of 100% when the prediction provides a perfect fit to the observed flood extents. While CSI is conventionally the most widely used metric for this comparison, Cohen's kappa index provides a better overall metric with TN pixels also being taken into account. Lastly, it should be noted that the magnitude and the size of the flood (and consequently the number of pixels used for the computation) were shown by (Stephens et al., 2014) to have a significant influence on these indices; thus limiting their use for different event and different catchment comparison. This limitations has no impact here, as CSI and κ indices are used to compare different numerical experiments on a single catchment and on the same event.

5 Results and Discussions

Quantitative performance assessments are carried out in the control and in the observational spaces by comparing (*i*) the parameters yielded by the different DA analysis, including a comparison to the true parameters in OSSE, (*ii*) the different analyzed WL time-series with synthetical or real in-situ observations, (*iii*) the different analyzed WSR with real or synthetical WSR observations in the floodplain, and (*iv*) the contingency maps and the overall CSI and Cohen's kappa index computed for the different analyzed flood extent maps, with respect to the synthetical or real observed flood extent maps. First, these comparisons in OSSE mode (subsection 5.1) allow to assess the benefits of assimilating spatially distributed RS-derived observations, with an augmented control vector, in order to represent the floodplain dynamics, and advocates for this strategy in real experiment mode. Then, subsection 5.2 presents all these quantitative assessments concerning the two real flood events, 2019 and 2021.

5.1 Results for OSSE experiments

5.1.1 Results in the control space for OSSE

Figure 5 shows the analyzed parameters from the different DA experiments, with blue lines for IDA, green lines for IWDA and red lines for IHDA for the synthetical 2021 flood event. The true parameter values are plotted in black and the calibrated or default values $\mathbf{x}_0$ are indicated by horizontal dashed lines. The overpass times of S1 over the 2021 event are depicted by vertical dashed lines. The analyzed values for K_{s_k} (with $k \in [0, 6]$) are shown on the left column, while that of the inflow correction μ is in the top panel of the right column. The reference and the analyzed values in IHDA experiment for δH_k^a with $k \in [1, 5]$ are shown on the other panels of the right column, respectively in black and in red (0 for the default value). The bottom right panel displays the upstream forcing for reference purposes.

For the synthetical 2021 event, it appears that all three DA analyses succeed in retrieving the true friction coefficients in the river bed, with a lesser success on the 5^{th} and 6^{th} river segments (i.e. K_{s_5} and K_{s_6}). This is most likely due to equifinality issues, as the downstream part of the flow is also influenced by the friction in the middle part of the river near Marmande (corresponding to the 3^{rd} and 4^{th} river segments controlled by K_{s_3} and K_{s_4}). Also due to the equifinality issues, the analysis for the floodplain friction K_{s_0} probably compensates for the analysis of K_{s_3} at the beginning of the event during low flow. As the water begins to occupy the floodplain, this equifinality issue lessens and the analysis on the floodplain friction becomes more efficient, and converges to the true value. It should be noted that despite these equifinality issues, all analyzed friction coefficients remain within physical ranges (both for the ones in the river bed and the one in the floodplain) and closer to the true value than to the default value, especially near the flood peak. It is also worth-noting that, as expected, the assimilation of in-situ observations at Marmande (located in the 4^{th} river segment) allows for an excellent analysis on K_{s_4} for IDA, and no additional information from the floodplain is necessary to constrain the friction in this segment. The analysis for the multiplicative factor μ is very noteworthy for all 3 DA experiments, even with a small underestimation as water recession starts. Given the localization step in the EnKF algorithm, the analysis for the state correction in the floodplain δH_k ($k \in [1, 5]$) only activates when WSR observations are present over the 18-hour assimilation window. Hence, the analysis for the WL correction (IHDA in red) is zero most of the time, including in between two S1 overpass times (in this catchment there is a minimum of 24 hours between two S1 observations from different orbits). For the assimilation windows that include WSR observations, the IHDA analysis succeeds in retrieving the values that are close to the true values for all subdomains of the floodplain and over the entire event.

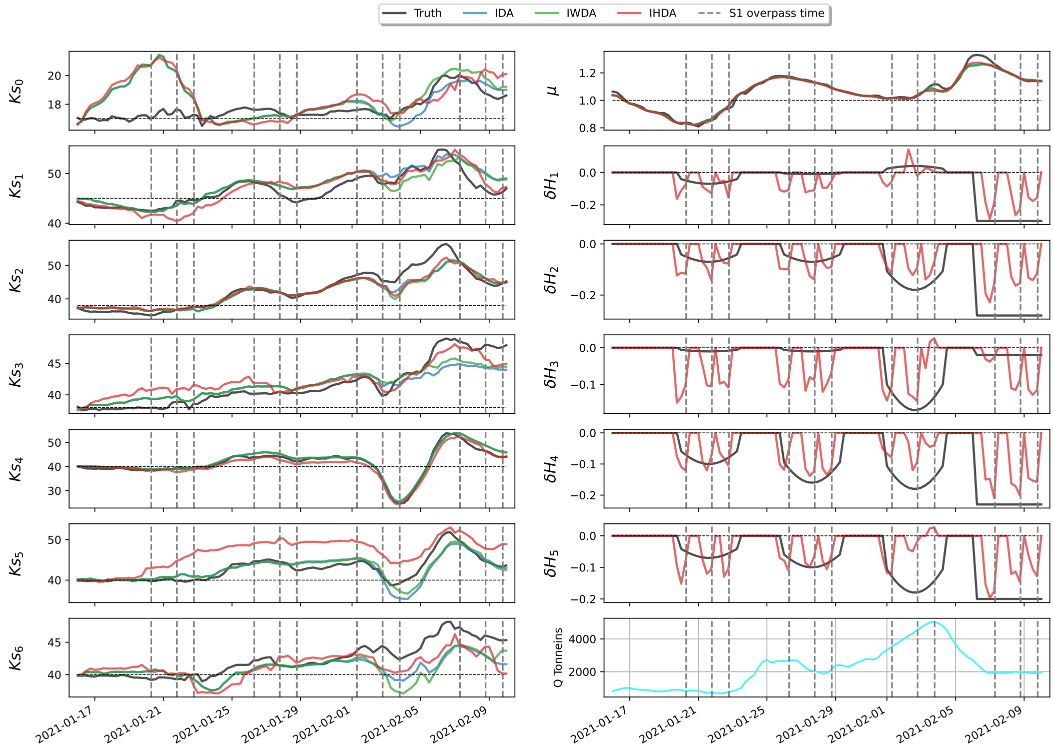


Figure 5: Analyzed values of the control vector for IDA (blue), IWDA (green), and IHDA (red) in OSSE. The default values are represented with horizontal dashed lines, whereas the S1 overpass times are shown with vertical dashed lines. Left column: friction coefficients in the floodplain K_{s_0}, and in the river bed K_{s_k} (with $k \in [1,6]$). Right column, from top to bottom: multiplicative correction to the inflow μ, state correction δH_k (with $k \in [1,5]$), and upstream forcing $Q(t)$.

5.1.2 Results in the observation space: Water levels at observing stations for OSSE

The RMSE (Eq. (26)) computed over the entire event, for the WL from FR simulation, as well as from IDA, IWDA and IHDA analyses, with respect to the reference WL at Tonneins, Marmande, and La Réole are presented in Table 4. For each observing station, the lowest RMSE values are underlined. Table 4 shows that all DA experiments succeed in significantly reducing the WL errors, compared to that of FR. The reduction in RMSE with respect to FR amounts to 79%, 89%, and 91%, respectively, at Tonneins, Marmande, and La Réole, with very close values for IDA, IWDA and IHDA. The RMSE at observing stations remains under 5.5 cm for all DA experiments. This level of precision is expected in OSSE, in coherence with the prescribed observation error. It validates the performance of the implemented EnKF solution. These results illustrate that the correction of the hydraulic state in the river bed can be properly achieved assimilating WL at observing stations only, and that the complementary assimilation of WSR is of greater interest when assessing the dynamics of the floodplain.

5.1.3 Results in the observation space: WSR in the floodplain for OSSE

The WSR in the five floodplain zones for the simulated WL in FR and the analyzed WL in the three DA experiments are compared to the WSR computed from the

Table 4: Water level RMSE w.r.t. reference water levels at VigiCrue observing stations, for 2021 synthetical event, in OSSE.

Exp. name	Root-Mean-Square Error [m] Tonneins	Marmande	La Réole
FR	0.260	0.397	0.578
IDA	0.052	0.042	0.053
IWDA	0.055	0.044	0.054
IHDA	0.052	0.045	0.050

reference simulation in Figure 6. The WSR values are shown in Figure 6a and the misfit between the reference and simulation WSR values (i.e. observed WSR - simulated WSR) are shown in 6b. The WSR for the truth are plotted in black, whereas the WSR for FR are in orange. The color code for the DA experiments is the same as in Figure 5: IDA in blue, IWDA in green, and IHDA in red. From the beginning of the event to the flood rising limb (around 2021-02-01), the impact of assimilating WSR is insignificant as the water has not overflowed to the floodplain. The WSR values in the reference and the experiment are thus null or close to zero.

Near the flood peak, FR underestimates flooding in most of the zones, with the exception of zone 5. Both IDA and IHDA present improved results with greater WSR values than FR. The merits of IHDA (red) versus IWDA (green) is clearly visible during the flood recess (after 2021-02-03) in all zones when the T2D model alone in FR fails to evacuate the water. The WSR values in IHDA are brought significantly close to the reference WSR values, while WSR values for IDA and IWDA are not better than those of FR. This illustrates how the augmented control vector with δH_k (with $k \in [1, 5]$) allows for an efficient assimilation of the information in the floodplain expressed as WSR measurements, and an associated correction of the floodplain dynamics. This shows that IHDA is the most efficient DA strategy to represent the floodplain, thus advocates for its application in real event mode.

5.1.4 2D validation with contingency maps, CSI *and* κ *indices for OSSE*

Figure 7 displays the resulting contingency maps for FR and DA experiments formulated for the T2D simulated flood extent maps with respect to those of the reference simulation in OSSE. The correctly predicted pixels are represented in light blue when non flooded, and in dark blue when flooded in the (synthetical) observations. The incorrectly predicted non-flood and flooded pixels (respectively, underprediction and overprediction) are represented in yellow and in red. Contingency maps are shown for the synthetical 2021 event at the time of the flood peak (top panel) and during water recess (bottom panel). The resulted CSI (Eq. (27)) and the κ indices (Eq. (28)) are also indicated. At the flood peak, FR significantly underestimates flooding over several subdomains of the floodplain. While the assimilation of in-situ data in IDA and the joint assimilation of WSR in IWDA bring some improvements, the most significant improvement comes from the extended control vector involving the hydraulic state associated with the assimilation of WSR in IHDA. During the water recess, IDA and IWDA fail to bring any improvement with respect to FR. Yet, the correction of the hydraulic state in the subdomains of the floodplain associated with the assimilation of WSR in IHDA leads to an effective drying of the floodplain that is in good agreement with the synthetical observation.

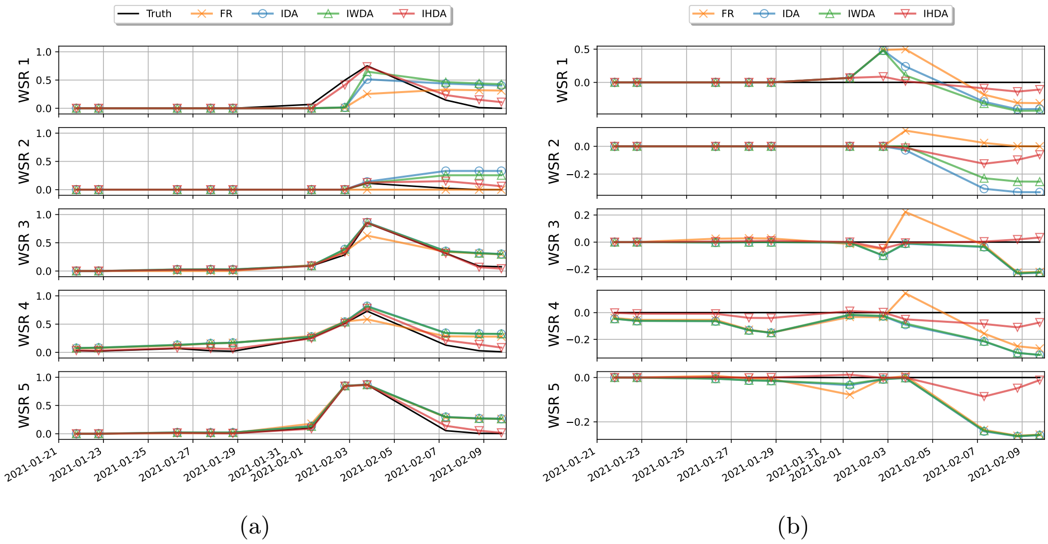


Figure 6: (a) WSR values computed in OSSE for the reference run (black), FR (orange), IDA (blue), IWDA (green), and IHDA (red) over the five subdomains of the floodplain. (b) Misfit between the reference WSR and the simulated WSR values in the five floodplain zones.

Figure 8a (respectively, Figure 8b) depicts the CSI (respectively, the κ index) yielded by FR and DA experiments at all S1 overpass times. Within the OSSE framework, all DA experiments result in flood extent maps that are in relative agreement with the reference flood maps. Indeed, IDA, IWDA and IDA allow for a significant improvement with respect to FR near the flood peak (2021-02-03 19:00). Yet, IHDA outperforms both IDA and IWDA, especially during the flood recess. IHDA leads to a CSI above 68% at every time steps (and above 88% before water recession period). During the flood recess (last three timesteps), IDA and IWDA have a CSI varying between 38-63% while IHDA has a CSI above 68% at all 3 timesteps. The results on the κ index, while also involving the TN counts (cf. subsection 4.2.1), provide the same conclusions. The analysis of the contingency maps as well as the CSI and κ indices demonstrate the merits of the assimilation of the WSR observations, together with the correction of the hydraulic state in subdomains of the floodplain. This strategy is thus applied in real event mode in the following subsection.

5.2 Results for real experiments

In this subsection, the quantitative assessments are carried out in two real flood events, 2019 and 2021.

5.2.1 Results in the control space

Similarly to Figure 5, the analyzed parameters from the different DA experiments in real mode for 2019 event (respectively, 2021 event) are shown in Figure 9a (respectively, in Figure 9b), where horizontal black dashed lines stand for the default values $\mathbf{x}_0$, blue curves for IDA, green curves for IWDA, and red curves for IHDA. The analyzed values for K_{s_k} (with $k \in [0, 6]$) over the flood events are shown on the left column of each figure. The analysis for the inflow correction μ is shown in the top panel of the right column. The analyzed values for δH_k^a (with $k \in [1, 5]$) by IHDA are shown on the other

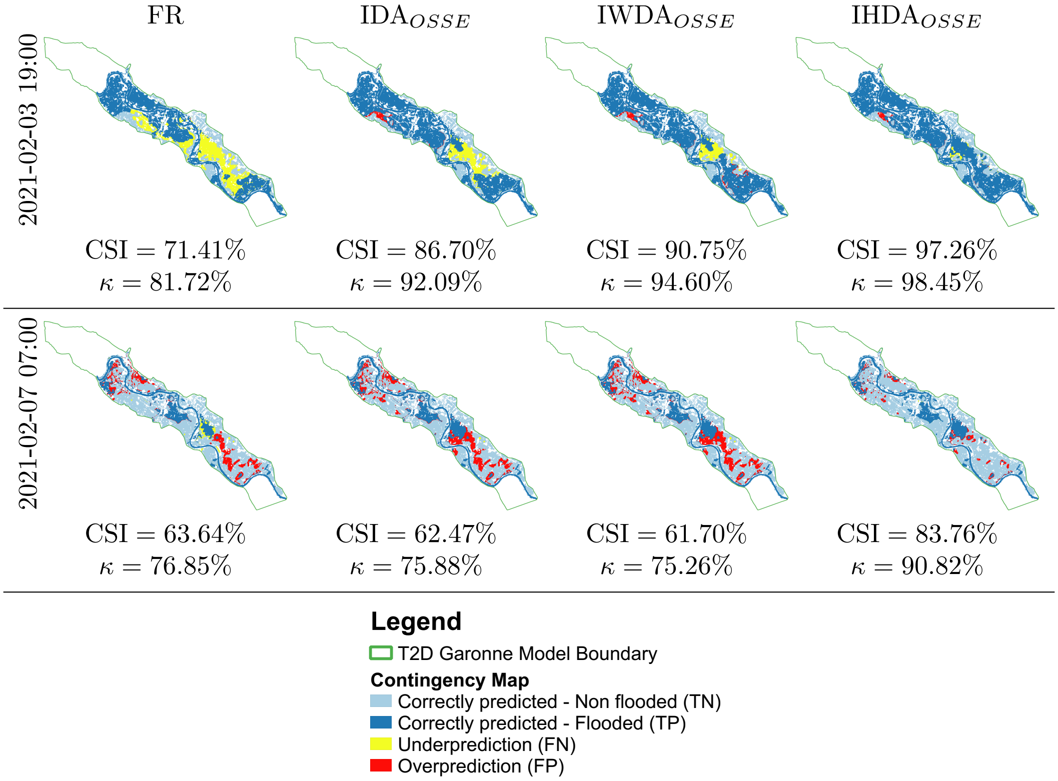
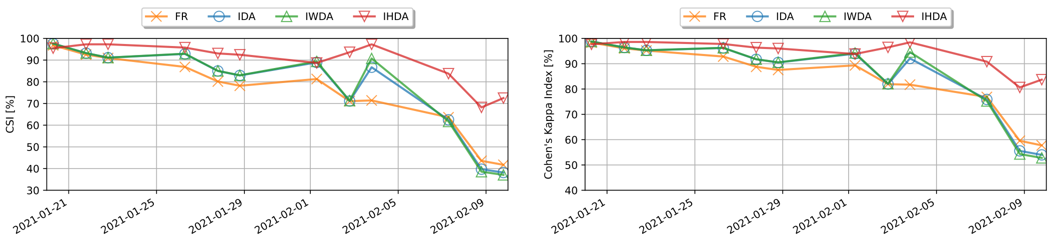


Figure 7: Contingency maps computed between simulated flood extent (from left to right: FR, IDA, IWDA and IHDA) with respect to the synthetical flood extent maps from the reference simulation in OSSE. First row: flood peak on 2021-02-03 18:48; Second row: flood recess on 2021-02-07 07:06.



(a) CSI

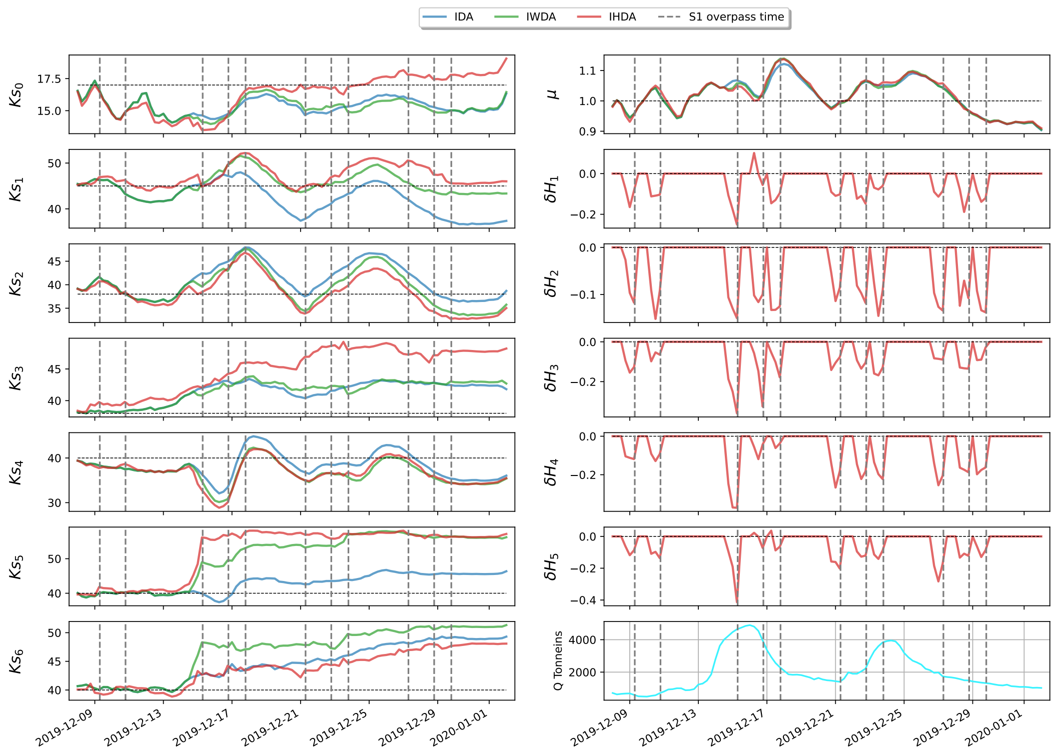
(b) Cohen's kappa index

Figure 8: (a) CSI and (b) κ indices computed for the FR and DA experiments with respect to the synthetical flood extent from the reference simulation in OSSE, at S1 overpass times.

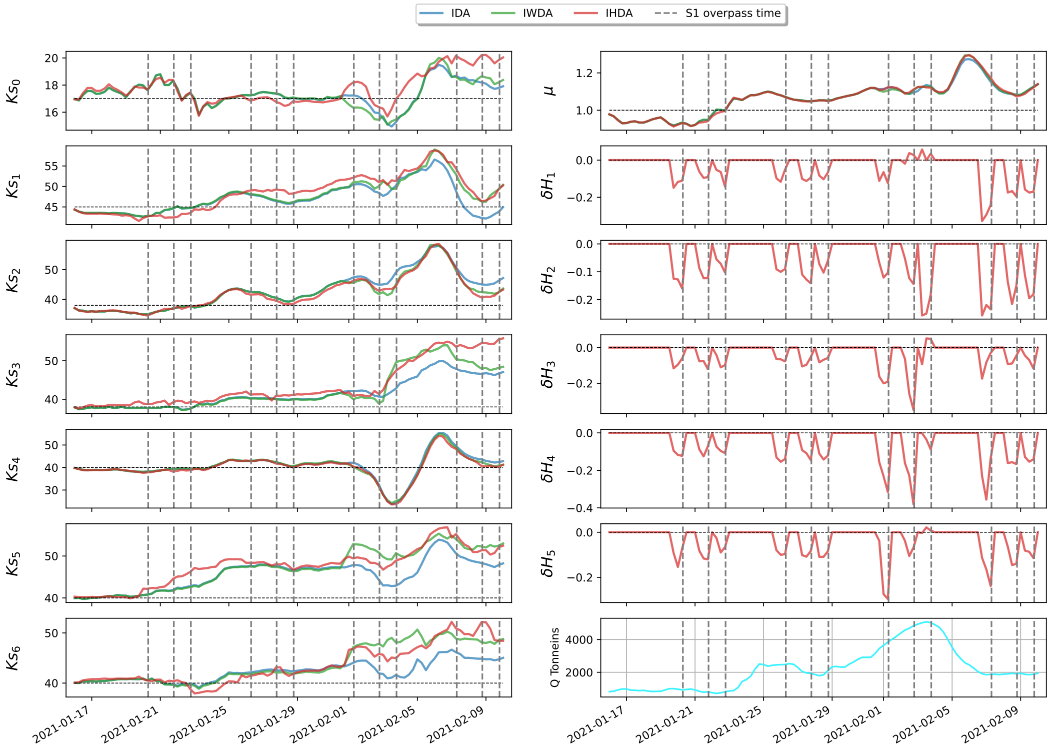
panels of the right column. The bottom panel of the right column displays the upstream forcing for reference purposes.

For all DA experiments and for both 2019 and 2021 events, the analysis values for the friction coefficients in the river bed and the floodplain remain within physical ranges, including the ones in the river bed and the one in the floodplain. The increment are larger during the flood event, as the misfit between the background run and the observations increases. The analysis for IHDA are closer to that of IWDA, compared to IDA, as in both experiments the control vector is extended with the hydraulic state. The analysis is quite far from the calibrated values for the friction of the 5^{th} and 6^{th} river segments (i.e. K_{s_5} and K_{s_6}), which is most likely due to the poor quality of the model topography in the downstream part of the domain, and the large misfit between the in-situ and the simulated WLs at La Réole. As previously remarked in OSSE mode, the analysis in the 4^{th} friction segment (i.e. K_{s_4}), that includes Marmande, is similar for IDA, IWDA and IHDA, showing that the assimilation of in-situ WLs suffices to account for friction errors in this area. Over the other friction zones, IDA is often closer to the default values. The analyses on μ are similar for IDA, IWDA, and IHDA for both events. This suggests that the in-situ WLs observed at Tonneins are enough to constraint the multiplicative correction to the inflow and that the use of additional data in the floodplain is not necessary. Concerning IHDA, the mostly negative correction on all δH values increases (i.e. more water is removed in the corresponding floodplain zones) as the flood rises, especially at the flood peak and during recess in order to account for the T2D model's limitation in physical process. During the recess period, the correction of the hydraulic state contributes in evacuating the water in the floodplain.

The results of IDA, IWDA and IHDA on the 2019 event show a greater dispersion than on the 2021 event. This may be due to the more complex flood dynamic of the 2019 event with two peaks and thus results in a degraded representation between the first recess and the second flood peak. As opposed to the assessment carried out in OSSE mode (subsection 5.1.1), the evaluation of the DA experiment results in the control space does not allow to quantitatively assess which DA strategy provides the best performance due to unknown true values of controlled parameters, thus further validations in the observation space are necessary.



(a) 2019 flood event



(b) 2021 flood event

Figure 9: Analyzed values of the control vector for IDA (blue), IWDA (green), and IHDA (red), for (a) 2019 and (b) 2021 real events. The default values are represented by horizontal dashed lines, whereas the S1 overpass times are shown as vertical dashed lines. Left column: friction coefficients in the floodplain K_{s_0}, and in the river bed K_{s_k} (with $k \in [1, 6]$). Right column, from top to bottom: multiplicative correction to the inflow μ, state correction δH_k (with $k \in [1, 5]$), upstream forcing $Q(t)$.

5.2.2 Results in the observation space: Water levels at observing stations

The RMSEs computed over time for the 2019 event (respectively, 2021 event), for the WLs from the FR, IDA, IWDA and IHDA, with respect to the observed WLs at Tonneins, Marmande, and La Réole are summarized in Table 5a (respectively, in Table 5b). For each observing station, the lowest RMSE values are underlined. Table 5 shows that all DA experiments succeed in significantly reducing the WL errors compared to those of FR, even though such reductions are less significant than in OSSE (Table 4).

For the 2019 event, the reductions in RMSE with respect to FR amount to 50%, 77%, and 57%, respectively, at Tonneins, Marmande, and La Réole, with close values between IDA, IWDA, and IHDA. For the 2021 event, those reductions are 34%, 80%, and 84%, respectively, at Tonneins, Marmande, and La Réole. For both event, the RMSEs at Tonneins and Marmande remain under 8 cm for all DA experiments, whereas it is under 14 cm at La Réole. While the RMSEs at Tonneins remain similar between the two events, a trade-off between Marmande and La Réole can be remarked for the 2019 and the 2021 events. These indicates that the model struggles to represent the dynamics at La Réole, most likely due to errors in topography in the downstream part of the domain, to errors in the rating curve used as downstream BC or to the presence of non-modeled tributaries that might play a significant role for high flows.

Table 5: Water level RMSE w.r.t. in-situ water levels at VigiCrue observing stations. The lowest RMSE is underlined.

(a) 2019 flood event

Exp. name	Root-Mean-Square Error [m] Tonneins	Marmande	La Réole
FR	0.129	0.220	0.318
IDA	0.060	0.045	0.125
IWDA	0.064	0.049	0.128
IHDA	0.064	0.051	0.138

(b) 2021 flood event

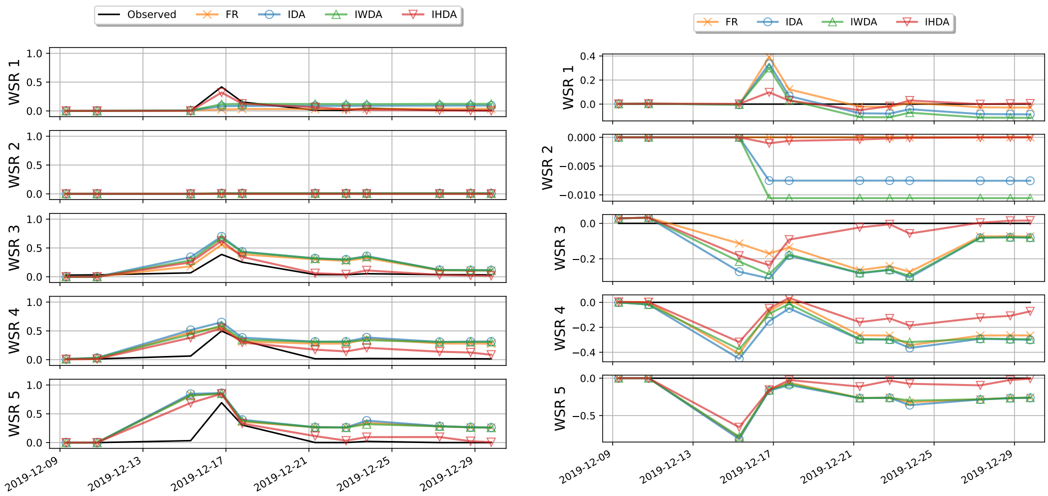
Exp. name	Root-Mean-Square Error [m] Tonneins	Marmande	La Réole
FR	0.106	0.392	0.536
IDA	0.062	0.071	0.081
IWDA	0.069	0.077	0.081
IHDA	0.065	0.073	0.079

It should be noted that the best DA strategy according to in-situ WL RMSE is IDA (although ever so slightly). The assimilation of WSR in the floodplain (in IWDA and IHDA) leads to a smaller WL improvement from FR at observing stations than IDA does. This is because the dynamics of the T2D model may be consistent with the real dynamics within the river bed, but not coherent with real dynamics in the floodplain. Indeed, while in OSSE mode (subsection 5.1), the observations in both the river bed and the floodplain were obtained from the same set of reference parameters which results in IHDA achieving the lowest RMSEs, it is highly probable that, for the real events, no set of model parameters allows to represent simultaneously a realistic and consistent dynamics in the river bed and in the floodplain. Therefore, a more complex hydrodynamic model should be considered to overcome these limitations, for instance, by considering a finer zoning of friction in the river bed and the floodplain, an addition of lateral tributaries that mainly carry a large volume of water for high flows, a more precise description of the topography in the floodplain, or an addition of physical processes in the SWE solver such as rain and evapotranspiration. A preliminary conclusion here is that the assimilation of data in the floodplain is shown to under-perform the assimilation of (in-situ) WL data in the river bed, especially when the performance is only assessed through the metrics in the river bed.

5.2.3 Results in the observation space: WSR in the floodplain

The WSR in the five floodplain subdomains for the simulated WL in FR and the analyzed WL in the three DA experiments with a threshold of 5 cm are compared to the WSR computed from the observed S1-derived flood extent maps, and shown in Figure 10 and Figure 11, respectively for the 2019 and 2021 events. For the 2019 event (respectively, 2021 event), the WSR values are shown in Figure 10a (respectively, Figure 11a) and the misfit between simulation and observation WSR values (i.e. observed WSR - simulated WSR) are shown in Figure 10b (respectively, Figure 11b). The color codes for the experiments are the same as in previous figures, i.e. FR in orange, IDA in blue, IWDA in green, and IHDA in red.

As previously noted in OSSE, the impact of assimilation WSR is not significant until the floodplain is active. In most subdomains, when the floodplain is active, the model (FR and all DA experiments) tend to overflood, especially during flood recess period. First, it should be noted that the analysis for IDA and IWDA does not bring much improvement with respect to FR in the 2019 flood event. The improvement is much more evident for IHDA, in both events, especially at the flood peak and during the recess period. For the 2019 event, IHDA brings a significant improvement for the subdomains 3, 4 and 5 as the misfits in subdomains 1 and 2 have already been small for FR (hence the contributions from IHDA are less obvious). Such an improvement over all subdomains is much more evident in the 2021 event. A significant overprediction at the timestep right before the first peak (2019-12-15 07:00) in subdomain 4 and 5 can be observed. This could stem from the characteristics of SAR backscatter which intensifies as the soil moisture increases due to rainfalls while the area has not been flooded. The correction of the hydraulic state in the floodplain for IHDA, during the recess of the first peak (between 2019-12-17 and 2019-12-21), allows for a better simulation of the second flood peak than in FR. For both events, the assimilation of WSR by IHDA with the hydraulic state corrections brings an improvement is all subdomains and the floodplain is efficiently emptied after the flood peak.



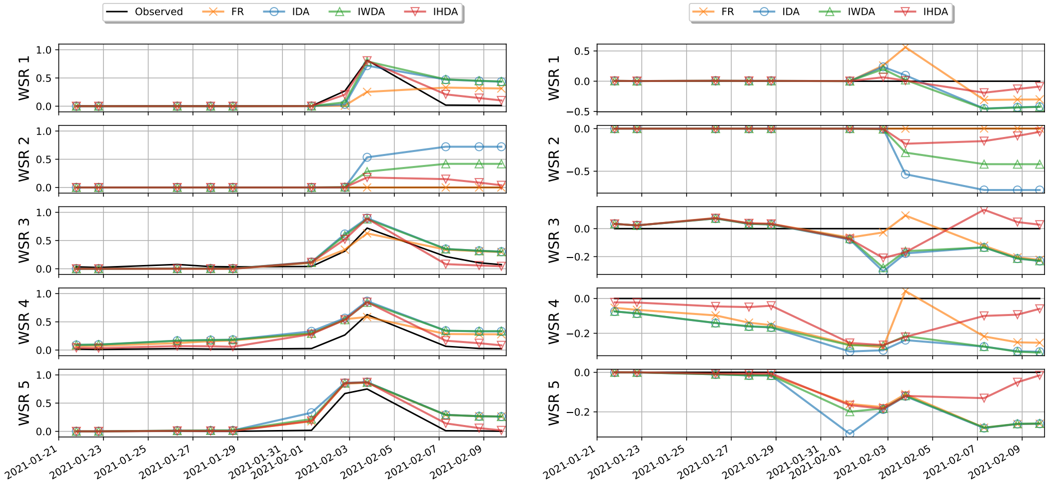
(a) WSR values

(b) Misfits

Figure 10: 2019 flood event - (a) WSR values computed for the S1-derived flood extent (black), FR (orange), IDA (blue), IWDA (green), and IHDA (red) over the five floodplain zones. (b) Misfit between the observed WSR and the simulated WSR values in the five floodplain zones.

5.2.4 2D validation with contingency maps, CSI *and* κ *indices*

Similarly to the OSSE, 2D validations are carried out by evaluating contingency maps comparing T2D water masks with S1- or S2-derived flood maps at their overpass times, and by quantitatively assessing the resulting CSI and the κ index scores. However, since the flood dynamic is quite different and even more complex in the 2019 flood event, let us start with the 2D validation on this event. Figure 12 depicts the contingency maps based on the comparison of the T2D simulated flood extent maps from FR and DA experiments with respect to those derived from S1 or S2 images during the 2019 flood event. The contingency maps are shown from top to bottom, at satellite overpass time right before the first flood peak by S2 (2019-12-15 12:00), at flood peak by S1 (2019-12-16 19:00), during the flood falling limb by S2 (2019-12-17 12:00) and by S1 (2019-12-17 19:00), and at the beginning of the second flood peak (2019-12-23 19:00) for S1. It should be stressed that, in this work, S2 imagery data are not assimilated and only are used for



(a) WSR values

(b) Misfits

Figure 11: 2021 flood event - (a) WSR values computed for the S1-derived flood extent (black), FR (orange), IDA (blue), IWDA (green), and IHDA (red) over the five floodplain zones. (b) Misfit between the observed WSR and the simulated WSR values in the five floodplain zones.

validation as independent data. The associated CSI and the κ indices are indicated on each contingency map.

For 2019 flood event, IHDA brings noticeable improvements with respect to FR, IDA and IWDA before the flood peak (first row in Figure 12), with better predictions of the flooded pixels, mostly in subdomain 1 and 3. A relatively significant overprediction on subdomain 4 and 5 from all experiments can be observed on these first-row figures. It is coherent with the remark made on the WSR validation (subsection 5.2.3). At the first flood peak observed by S1 image (second row in Figure 12), IHDA allows better predictions of the flooded pixels, mostly in subdomain 1. During the first flood recess (third and fourth row in Figure 12), the improvement brought by IHDA is not as visible as at the flood peak (second row). The added validation of the S2 image at 2019-12-17 12:00 provides an interesting remark. Indeed, the observed flood extent detected on this image is more similar to the one captured by the S1 image at 2019-12-16 19:00 (or 17 hours backward) than the one right afterward at 2019-12-17 19:00 (or 5 hours forward). Such a situation, taking into account the fact that these three images in particular were acquired in the span of 24 hours during the start of the falling limb, shows the different tendencies between the in-situ WL and the floodplain dynamics. This emphasizes the complexity of the flood dynamics in the floodplain, and advocates for the further addition of the S2-derived flood observations in the DA. Such a remark of S1 and S2 incoherence is rarely possible due to the unavailability of S2 images during a flood event because of cloud cover problem. Lastly, the fifth row of Figure 12 shows an overall improvement spread out over the five subdomains. This is also thanks to the state corrections applied at the timesteps between the two flood peaks.

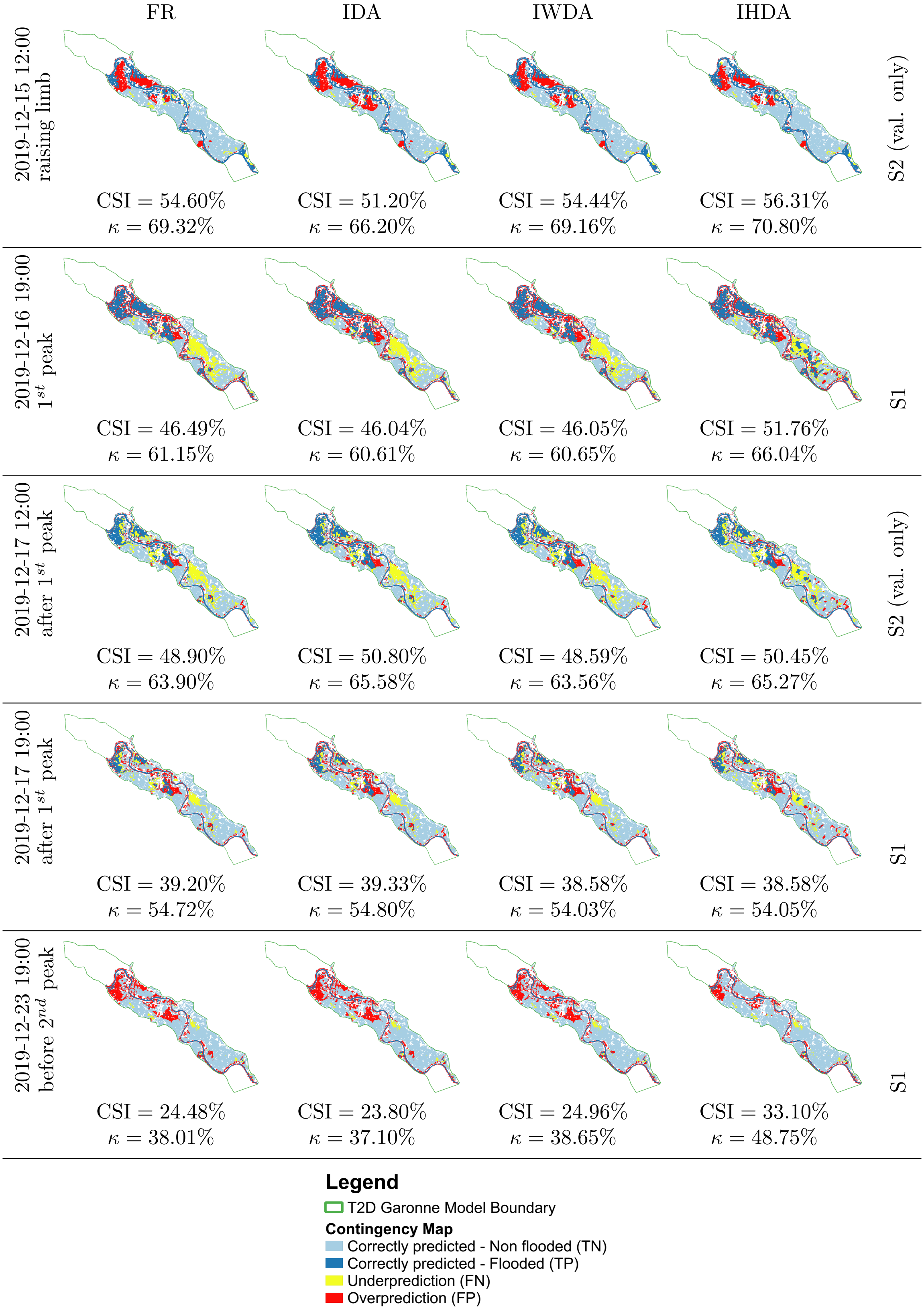


Figure 12: 2019 flood event - Contingency maps computed between simulated flood extent (from left to right: FR, IDA, IWDA and IHDA) with respect to S1-derived flood extent (row 2, 4 and 5) and S2-derived flood extent (rows 1 and 3).

Figure 13 displays the contingency map for the 2021 flood event (flood peak observed at 2021-02-03 19:00 and recess 2021-02-07 07:00), with metrics computed with respect to S1 derived flood extent as no S2 data where available. For 2021, the assimilation of WSR data brings a significant improvement at the flood peak (first row in Figure 13) in all subdomains in terms of the correctly predicted flood pixels. The recess period (second row in Figure 13) simulated by IHDA is also better than that of IDA and IWDA, yet, some residual flooded pixels remain, leading to still over-predicted areas. Figure 14 depicts the CSI (left column) and the κ index (right column) yielded by FR and DA experiments at all S1 overpass times, for 2019 event (top panels) and 2021 event (bottom panels), with the same color code used previously. These confirm the merits of the DA strategy in IHDA, especially for flood recess.

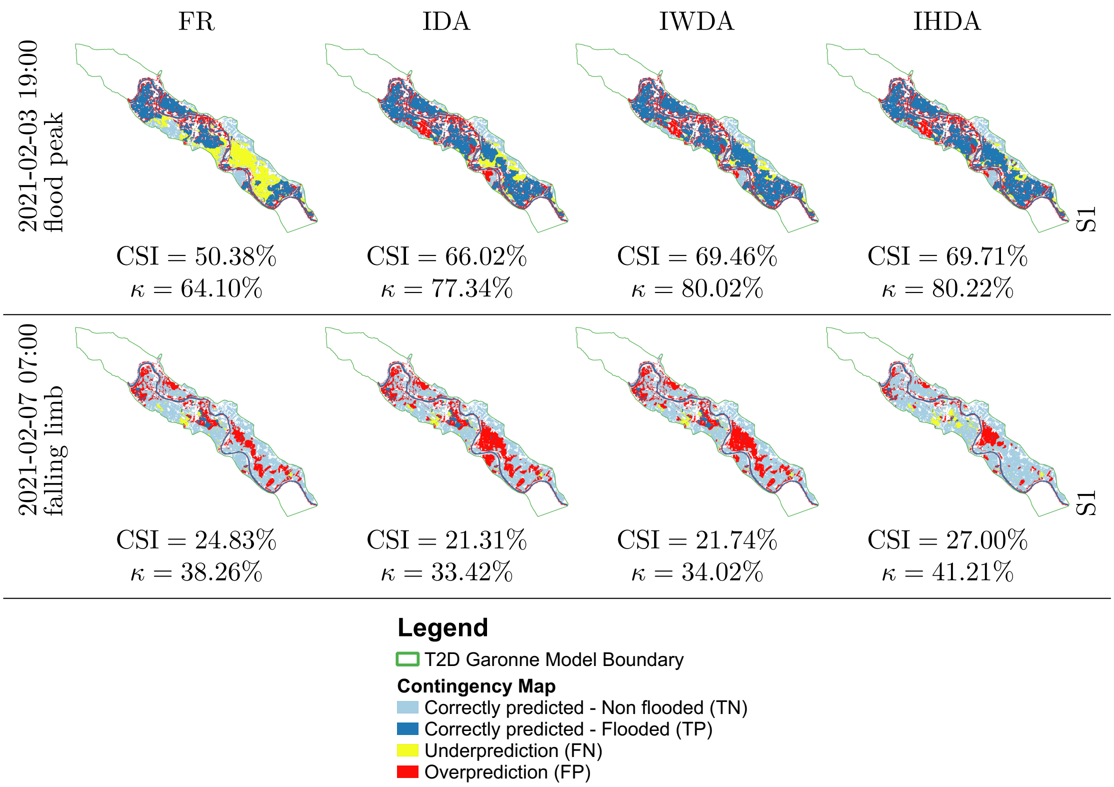
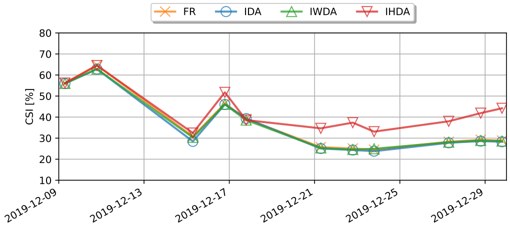


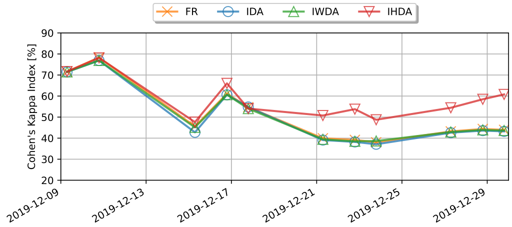
Figure 13: 2021 flood event - Contingency maps computed between simulated flood extent (from left to right: FR, IDA, IWDA and IHDA) with respect to S1-derived flood extent.

5.2.5 Post-event measure validation with HWM observations

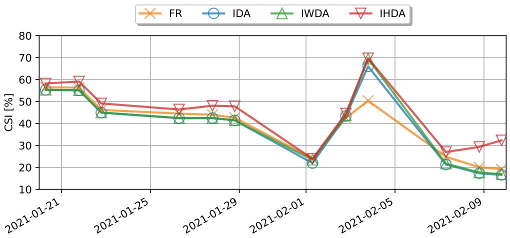
Figure 15a and 15b illustrate the validations leveraging the collective public datasets of HWM for the 2019 and the 2021 flood events. They allow us to evaluate the spatially distributed observed highest WL at various points on the river banks or within the floodplain, as opposed to the remote sensing-derived 2D flood extents that lack WL information. However, for the sake of conciseness, only the comparisons between FR and IHDA experiments are shown. First, since the flood 2019 event is of a smaller scale compared to the 2021 event, fewer HWM observations were collected. The HWM errors between the simulated WL and the observed WL are classified into four range, taking $\pm$ 1 meter as a baseline for small errors While an agreement between the errors is not available



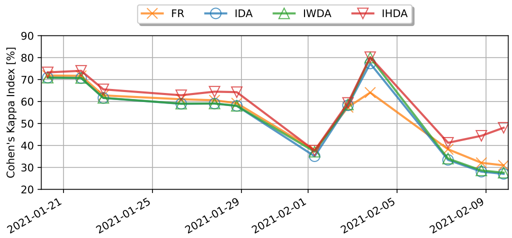
(a) CSI - 2019



(b) Cohen's kappa index - 2019



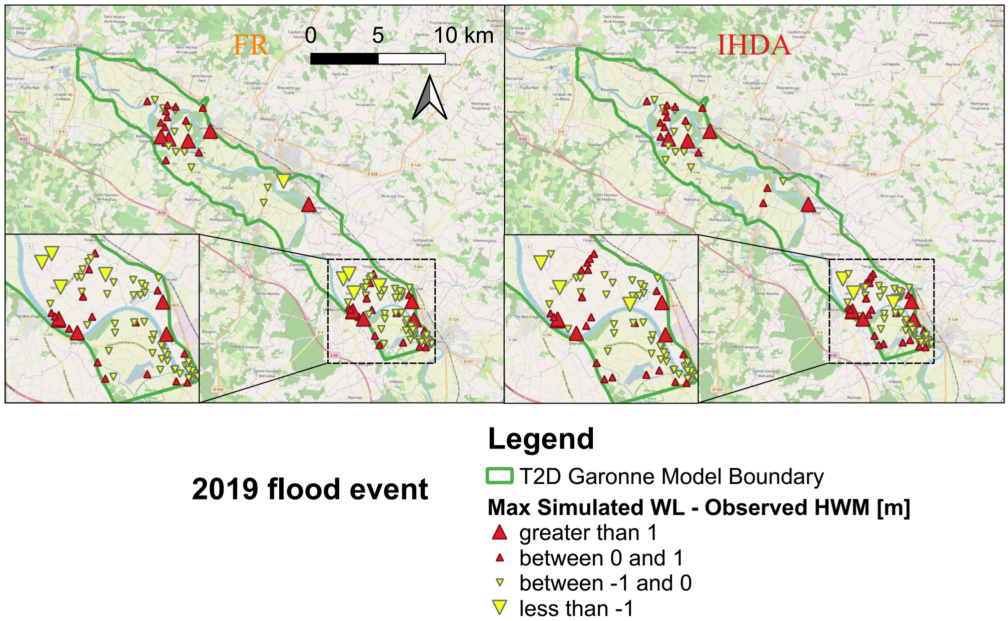
(c) CSI - 2021



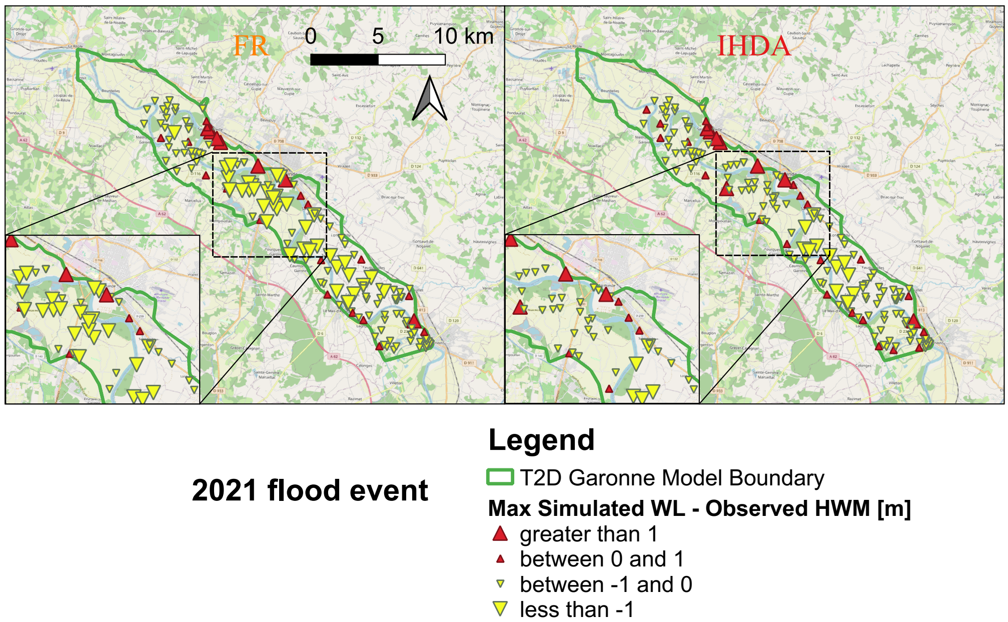
(d) Cohen's kappa index - 2021

Figure 14: CSI (left column) and κ (right column) index computed for the FR (orange), IDA (blue), IHDA (green), and IWDA (red) experiments with respect to the S1-derived observed flood extent, for the 2019 flood event (top) and 2021 flood event (bottom).

in the 2019 flood event, an unanimous improvement by IHDA compared to FR at various points can be noted, from strong to weak underprediction (visually, from big yellow triangles to small yellow ones). Similar results are found between IDA, IWDA, and IHDA. Since this validation only concerns the highest WL after an event, the relevance of IHDA demonstrated strongly over the flood recess becomes unseen.



(a) HWM validation - 2019



(b) HWM validation - 2021

Figure 15: Post-event HWM validations over the (a) 2019 and (b) 2021 flood events. A negative value indicates an underprediction (yellow triangles) by the simulation whereas a positive value indicates an overprediction (red triangles).

6 Conclusions and Perspectives

This study presents the merits of assimilating 2D flood extent observations derived from remote sensing Sentinel-1 SAR images with an Ensemble Kalman Filter implemented on the 2D hydrodynamics model TELEMAC-2D. The flood extent information is expressed in terms of Wet Surface Ratio computed over defined sensitive subdomains of the floodplain. The WSR is assimilated jointly with in-situ water level observations. The study was carried out over the Garonne Marmandaise catchment, focusing on two flooding events in 2019 and 2021. Four experiments were realized; one in free run mode and three in DA mode. The control vector gathers friction and forcing correction, and is augmented with correction of the hydraulic state in subdomains of the floodplains (IHDA experiment) that constitute the innovative strategy of this work. All of the DA experiments were implemented by a cycled EnKF with an 18-hour assimilation window sliding with 6-hour overlapping. The DA strategy was first assessed in OSSE that mimics the 2021 flood event, then applied in re-analysis mode to both real events. The simulation results were comprehensively assessed with 1D and 2D metrics with respect to assimilated data as well as with respect to independent flood extent, derived from Sentinel-2 optical imagery data or High Water Mark collective public observations when they are available.

The first DA experiment (IDA) involves only in-situ observations whereas the second one (IWDA) assimilates both in-situ observations and WSR observations derived from 2D flood extent maps. These two experiments focus on the sequential correction of friction coefficients and inflow discharge. In OSSE, they demonstrated effectiveness in retrieving the true parameters and providing relevant assessment results. The spotlight of the article is the IHDA experiment, which not only assimilates both types of observations (similar to IWDA), but also handles a dual state-parameter estimation within the EnKF, by treating inflow discharge and friction coefficients as well as the hydraulic state variable in five particular floodplain subdomains, representing evapotranspiration and/or ground infiltration processes that are unavailable in the T2D model.

We have shown that the assimilation of in-situ data in IDA improves the simulation in the river bed, yet, the dynamics in the floodplain remains incorrect with a significant underestimation of the flood (both events). Indeed, the in-situ observations located in the river bed, do not bring information on the dynamics in the floodplain. The assimilation of WSR data in the floodplain, in IWDA, brings additional improvements, that remains limited as the dynamics of the rived bed and that of the floodplain are not sensitive to model parameters that are accounted for in the control vector. The correction of the augmented control vector in IHDA allows to better represent the flood peak and to efficiently dry out the floodplain during the recess period. In OSSE mode, IHDA results in simulated WLs and WSRs that are very close to the synthetic observations, and yields better estimates of true friction and discharge parameters than IDA and IWDA. In real event mode, from FR to IHDA, the RMSE computed with respect to in-situ data in the river bed is reduced by up to 77-80% at Marmande, whereas the CSI computed with respect to remote-sensing flood extent maps is improved by up to 19.33 percentage points for the 2021 flood event (and 5.27 percentage points for the 2019 flood event). This study confirms the assertion that a densification of the observing network, especially in the floodplain, with remote sensing data and advanced DA strategy, allows to improve the representation of the dynamics of the flow in the floodplains.

This work rely on the implementation of an advanced DA strategy for TELEMAC-2D, especially the development of the observation operator dedicated to WSR, as well as the definition of the associated augmented control vector. Yet, it should be noted that the definition of the subdomains in the floodplain over which the hydraulic state is uniformly corrected, requires a deep understanding of the dynamics of the flood, and is thus not straightforward. This aspect could be further investigated, for instance based on a global sensitivity analysis with respect to the hydraulic state but also to other sources of uncertainty such as topography, especially in the downstream area. Indeed, the same dual state-parameter estimation approach could be applied to correct the bathymetry and topography provided that the size of the uncertainties is reduced, for instance working with a spatially uniform correction or a correction that is only projected onto a limited number of principal components of the errors. In this perspective, we aim to consider using high- and very-high-resolution topography as additional inputs to the model. The use of other imagery datasets (e.g. Landsat-8 and Landsat-9) can also be investigated. In the present work, the combination between remote-sensing data with regards to S1 and S2 data requires further investigation as it seems that the improvements made using S1-derived flood extent maps does not necessarily lead to an improvement with regards to S2-derived flood extents. This could stem from the differences between the S1 and S2 measurement, and the flood extent mapping algorithm. In addition, the identification of S1 or S2 exclusion maps should also be considered taking into account the limitations of each data source. Finally, an major perspective of this work stands in the potential non-gaussianity of the WSR observations. This limitation can amount to a loss of optimality of the EnKF which relies on the assumption that the observational error follows a gaussian distribution. On going work, based on a rich literature based on a change of variable to transform the non-gaussian error into gaussian errors (widely known as Gaussian anamorphosis) is on going and yield promising early results.

Acronyms

BC Boundary condition
CSI Critical Success Index
DA Data Assimilation
EnKF Ensemble Kalman Filter
FloodML Flood Machine Learning
FR Free Run
HWM High Water Marks
IDA In-situ (only) DA Experiment
IWDA In-situ and WSR DA experiment
IHDA In-situ and WSR DA experiment with extended control vector
OSSE Observing System Simulation Experiment
PDF Probability Density Function
RMSE Root-Mean-Square Error
RS Remote Sensing
SAR Synthetic Aperture Radar
SWE Shallow Water Equations
S1 Sentinel-1
S2 Sentinel-2
T2D TELEMAC-2D
WL Water Level
WSE Water Surface Elevation
WSR Wet Surface Ratio

Acknowledgments

This work was supported in part by the Centre National d'Etudes Spatiales (CNES) and in part by the Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS) within the framework of the Space for Climate Observatory (SCO). The authors gratefully thank the Electricité de France (EDF) for providing the Telemac2D model on the Garonne Downstream catchment, the SCHAPI, and the Garonne-Tarn-

Lot and Gironde-Adour-Dordogne SPCs for providing in-situ data. They also would like to thank S. El Garroussi and M. De Lozzo for former developments on the Telemac System with API dedicated to data-driven simulations.

References

Asch, M., Bocquet, M., & Nodet, M. (2016). *Data assimilation: methods, algorithms, and applications*. SIAM.

Bates, P. (2004). Remote sensing and flood inundation modelling. *Hydrological Processes*, *18*, 2593 – 2597. (Publisher: John Wiley and Sons)

Besnard, A., & Goutal, N. (2011). Comparaison de modèles 1D à casiers et 2D pour la modélisation hydraulique d'une plaine d'inondation–Cas de la Garonne entre Tonneins et La Réole. *La Houille Blanche*(3), 42–47.

Cian, F., Marconcini, M., Ceccato, P., & Giupponi, C. (2018). Flood depth estimation by means of high-resolution SAR images and LiDAR data. *Natural Hazards and Earth System Sciences*, *18*(11), 3063–3084.

Cooper, E. S., Dance, S. L., García-Pintado, J., Nichols, N. K., & Smith, P. J. (2019). Observation operators for assimilation of satellite observations in fluvial inundation forecasting. *Hydrology and Earth System Sciences*, *23*(6), 2541–2559.

Dasgupta, A., Hostache, R., Ramsankaran, R., Grimaldi, S., Matgen, P., Chini, M., ... Walker, J. P. (2021). Chapter 12 - earth observation and hydraulic data assimilation for improved flood inundation forecasting. In G. J.-P. Schumann (Ed.), *Earth observation for flood applications* (p. 255-294). Elsevier. Retrieved from `https://www.sciencedirect.com/science/article/pii/B9780128194126000122` doi: https://doi.org/10.1016/B978-0-12-819412-6.00012-2

Dasgupta, A., Hostache, R., Ramsankaran, R., Schumann, G. J.-P., Grimaldi, S., Pauwels, V. R. N., & Walker, J. P. (2021). A mutual information-based likelihood function for particle filter flood extent assimilation. *Water Resources Research*, *57*(2), e2020WR027859.

Di Mauro, C., Hostache, R., Matgen, P., Pelich, R., Chini, M., Jan van Leeuwen, P., ... Bloschl, G. (2021). Assimilation of probabilistic flood maps from SAR data into a coupled hydrologic-hydraulic forecasting model: a proof of concept. *Hydrol. Earth Syst. Sci.*, *25*, 4081-4097.

Gauckler, P. (1867). *Etudes Théoriques et Pratiques sur l'Ecoulement et le Mouvement des Eaux*. Gauthier-Villars.

Grimaldi, S., Li, Y., Pauwels, V. R., & Walker, J. P. (2016). Remote sensing-derived water extent and level to constrain hydraulic flood forecasting models: Opportunities and challenges. *Surveys in Geophysics*, *37*(5), 977–1034.

Hervouet, J.-M. (2007). *Hydrodynamics of free surface flows: modelling with the finite element method* (Vol. 360). Wiley Online Library.

Hostache, R., Chini, M., Giustarini, L., Neal, J., Kavetski, D., Wood, M., ... Matgen, P. (2018). Near-real-time assimilation of SAR-derived flood maps for improving flood forecasts. *Water Resources Research*, *54*(8), 5516–5535.

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*(3-4), 257–268.

Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, *32*(1), 1–6.

Huang, T., Baillarin, S., Altinok, A., Blanchet, G., Hausman, J., Kettig, P., & Shah, S. (2020, December). Distributed Machine Learning and Data Fusion for Flood Detection and Monitoring. In *AGU fall meeting abstracts* (Vol. 2020, p. IN041-09).

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., & Rodes, I. (2017). Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sensing*, *9*(1), 95.

Jung, H. C., Jasinski, M., Kim, J.-W., Shum, C. K., Bates, P., Neal, J., ... Alsdorf, D. (2012). Calibration of two-dimensional floodplain modeling in the central atchafalaya basin floodway system using SAR interferometry. *Water Resources Research*, *48*(7). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012WR011951 doi: https://doi.org/10.1029/2012WR011951

Kettig, P., Baillarin, S., Blanchet, G., Taillan, C., Ricci, S., Nguyen, T.-H., ... Roumagnac, A. (2021). The SCO-FloodDAM Project: New Observing Strategies for Flood Detection, Alert and Rapid Mapping. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS* (p. 1464-1467). doi: 10.1109/IGARSS47720.2021.9553036

Madsen, H., & Skotner, C. (2005). Adaptive state updating in real-time river flow forecasting - a combined filtering and error forecasting procedure. *Journal of Hydrology*, *308*, 302–312. doi: 10.1016/j.jhydrol.2004.10.030

Mason, D., Schumann, G.-P., Neal, J., Garcia-Pintado, J., & Bates, P. (2012). Automatic near real-time selection of flood water levels from high resolution synthetic aperture radar images for assimilation into hydraulic models: A case study. *Remote Sensing of Environment*, *124*, 705–716.

Mirouze, I., Ricci, S., & Goutal, N. (2019). The impact of observation spatial and temporal densification in an ensemble Kalman Filter. In *XXVIth TELEMAC-MASCARET User Conference, 15th to 17th october 2019, toulouse.*

Neal, J., & Jeffrey, C. (2007). Flood inundation model updating using an ensemble kalman filter and spatially distributed measurements. *Journal of Hydrology*, *336*, 401–415. doi: 10.1016/j.jhydrol.2007.01.012

Neal, J., Jeffrey, C., Atkinson, P., & Hutton, C. (2009). Evaluating the utility of the ensemble transform kalman filter for adaptive sampling when updating a hydrodynamic model. *Journal of Hydrology*, *375*(3-4), 589–600. doi: 10.1016/j.jhydrol.2009.07.008

Nguyen, T. H., Delmotte, A., Fatras, C., Kettig, P., Piacentini, A., & Ricci, S. (2021). Validation and Improvement of Data Assimilation for Flood Hydrodynamic Modelling Using SAR Imagery Data. In *Proceedings of the telemac-mascaret user conference october 2021* (pp. 100–108).

Nguyen, T. H., Ricci, S., Fatras, C., Piacentini, A., Delmotte, A., Lavergne, E., & Kettig, P. (2022). Improvement of Flood Extent Representation With Remote Sensing Data and Data Assimilation. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-22. doi: 10.1109/TGRS.2022.3147429

Revilla-Romero, B., Wanders, N., Burek, P., Salamon, P., & de Roo, A. (2016). Integrating remotely sensed surface water extent into continental scale hydrology. *Journal of hydrology*, *543*, 659–670.

Schumann, G., Bates, P. D., Horritt, M. S., Matgen, P., & Pappenberger, F. (2009). Progress in integration of remote sensing–derived flood extent and stage data and hydraulic models. *Reviews of Geophysics*, *47*(4). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008RG000274 doi: https://doi.org/10.1029/2008RG000274

Schumann, G. J.-P., Bates, P. D., Di Baldassarre, G., & Mason, D. C. (2012). The use of radar imagery in riverine flood inundation studies. In *Fluvial remote sensing for science and management* (p. 115-140). John Wiley and Sons, Ltd. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119940791.ch6 doi: https://doi.org/10.1002/9781119940791.ch6

Stephens, E., Schumann, G., & Bates, P. (2014). Problems with binary pattern measures for flood model evaluation. *Hydrological Processes*, *28*(18), 4928–4937.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., ... others

(2012). GMES sentinel-1 mission. *Remote Sensing of Environment*, *120*, 9–24.
Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025–3033.