A tempered particle filter to enhance the assimilation of SAR derived flood extent maps into flood forecasting models.

Concetta Di Mauro[1], Nancy K Nichols[2], Renaud Hostache[3], Patrick Matgen[4], Ramona-Maria Pelich[3], Marco Chini[3], Peter Jan van Leeuwen[5], and Guenter Bloeschl[6]

[1]LIST
[2]University of Reading
[3]Luxembourg Institute of Science and Technology
[4]Public Research Center - Gabriel Lippmann
[5]Colorado State University
[6]Vienna University of Technology

November 24, 2022

Abstract

Data Assimilation (DA) is a powerful tool to optimally combine uncertain model simulations and observations. Among DA techniques, the Particle Filter (PF) has gained attention for its capacity to deal with non-linear systems and for its relaxation of the Gaussian assumption. However, the PF may suffer from degeneracy and sample impoverishment. In this study, we propose an innovative approach, based on a Tempered Particle Filter (TPF), aiming at mitigating PFs issues, thus extending over time the assimilation benefits. Flood probabilistic maps derived from Synthetic Aperture Radar data are assimilated into a flood forecasting model through an iterative process including a particle mutation in order to keep diversity within the ensemble. Results show an improvement of the model forecasts accuracy, with respect to the Open Loop (OL): on average the RMSE of water levels decrease by 80% at the assimilation time and by 60% two days after the assimilation. A comparison with the Sequential Importance Sampling (SIS), is carried out showing that although SIS performances are generally comparable to the TPF ones at the assimilation time, they tend to decrease more quickly. For instance, on average TPF-based RMSE are by 20% lower compared to the SIS-based ones two days after the assimilation. The application of the TPF determines higher CSI values compared to the SIS. On average the increase in performances lasts for almost 3 days after the assimilation. Our study provides evidence that the application of the variant of the TPF enables more persistent benefits compared to the SIS.

A tempered particle filter to enhance the assimilation of SAR derived flood extent maps into flood forecasting models.

Concetta Di Mauro[1], Renaud Hostache[1], Patrick Matgen[1], Ramona Pelich[1], Marco Chini[1], Peter Jan van Leeuwen[2,4], Nancy Nichols[2], Günter Blöschl[3]

[1]Luxembourg Insitute of Science and Technology
[2]University of Reading,UK
[3]Vienna University of Technology
[4]Department of Atmospheric Science, Colorado State University, USA

Key Points:

- We assimilate flood extent maps into a flood forecasting system using a tempered particle filter
- The tempered particle filter mitigates degeneracy and enables long lasting forecast improvements
- The tempered particle filter outperforms a standard particle filter in terms of accuracy of model outputs

Corresponding author: Concetta Di Mauro, `concetta.dimauro@list.lu`

Corresponding author: Renaud Hostache, `renaud.hostache@ird.fr`

Abstract

Data Assimilation (DA) is a powerful tool to optimally combine uncertain model simulations and observations. Among DA techniques, the Particle Filter (PF) has gained attention for its capacity to deal with non-linear systems and for its relaxation of the Gaussian assumption. However, the PF may suffer from degeneracy and sample impoverishment. In this study, we propose an innovative approach, based on a Tempered Particle Filter (TPF), aiming at mitigating PFs issues, thus extending over time the assimilation benefits. Flood probabilistic maps derived from Synthetic Aperture Radar data are assimilated into a flood forecasting model through an iterative process including a particle mutation in order to keep diversity within the ensemble. Results show an improvement of the model forecasts accuracy, with respect to the Open Loop (OL): on average the RMSE of water levels decrease by 80% at the assimilation time and by 60% two days after the assimilation. A comparison with the Sequential Importance Sampling (SIS), is carried out showing that although SIS performances are generally comparable to the TPF ones at the assimilation time, they tend to decrease more quickly. For instance, on average TPF-based RMSE are by 20% lower compared to the SIS-based ones two days after the assimilation. The application of the TPF determines higher CSI values compared to the SIS. On average the increase in performances lasts for almost 3 days after the assimilation. Our study provides evidence that the application of the variant of the TPF enables more persistent benefits compared to the SIS.

Plain Language Summary

In this study, flood extent maps derived from satellite imagery were assimilated into a flood forecasting model with the aim to improve its short- to medium-range predictions. In a previous study we used a Data Assimilation technique based on Sequential Importance Sampling (SIS). While the assimilation of satellite-derived data improved the model predictions over several time steps, it was shown that such improvements did not persist over time and issues known as degeneracy and sample impoverishment led to suboptimal results. To mitigate the issues related to the application of the SIS, here we introduce a novel approach based on the so-called Tempered Particle Filter. This approach is based on iterative assimilations and updates of the initial model conditions. Our results show that the new method outperforms the previous one: water level errors over the model domain are substantially reduced up to 3 days following the assimilation and the accuracy of the flood extent maps is improved for up to 3 days. Moreover, the punctual water level and discharge accuracy are also improved. Therefore, the application of the proposed data assimilation approach not only mitigates the SIS-related issues, but it also enables longer lasting model improvements.

Introduction

Every year, floods cause important social and economic losses and the trend is increasing. Tellman et al. (2021) show that worldwide the population exposed to floods has increased by 20%–24% from 2000 to 2015, thereby highlighting the need for accurate and timely forecasts of water depth, discharge, flood wave propagation and flood extent to help reducing or preventing the adverse effects of floods. Flood forecasting models are commonly used to generate short- to mid-term predictions. However the accuracy of such predictions can be affected by multiple factors contributing to the overall model uncertainty. This challenge represents one of the major unsolved scientific problems (Blöschl et al., 2019). The assimilation of independent observations, such as field gauging data or satellite observations, can help reducing these uncertainties (Liu & Gupta, 2007). The last decade has seen a substantial increase in the number of Earth Observation (EO) satellites providing a synoptic overview of the flooding situation at increasingly high frequency. Despite possible errors in the interpretation of the SAR data (Chen et al., 2018; Grimaldi et al., 2020; Zhao et al., 2021) that should be masked out before any use of these data, frequent observations of flood extent and water depth represent substantial added value, especially over poorly gauged or ungauged catchments. For example, SAR data are relevant for observing inundation extent because of their day-night and quasi all-weather capability. As a consequence, several methods enabling an effective assimilation of such observations [e.g., Revilla-Romero et al. 2016; Hostache et al. 2018; Andreadis & Schumann 2014; Garcia-Pintado et al. 2015] for improving the predictive capability of flood models have been introduced and investigated in recent years. The most widely used methods are based on the Kalman Filter and its variants [e.g. Revilla-Romero et al. (2016); Annis et al. (2021); Wongchuig-Correa et al. (2020)] and they assume that the distributions of observation and model errors are Gaussian, which is not often the case when dealing with real word data (van Leeuwen et al., 2019).

Particle Filters (PFs) have gained attention within the research community because of their ability to handle non-linear and non-Gaussian systems (van Leeuwen et al., 2019). PFs approximate the prior and the posterior probability distribution functions (PDFs) with an ensemble of model states also called particles. An equal weight is assigned to each particle a priori. Next, as a result of the assimilation, weights are updated to represent the posterior probability given the observations. The principal limitation of PFs is the difficulty to deal with high-dimensional systems. The weights may vary significantly across particles and in the ultimate case only one particle will have a weight close to unity while the other particles will have negligible weight. As a result the ensemble may collapse. This well-known issue in PFs is often referred to as degeneracy. Degeneracy could lead to an erroneous approximation of the posterior distribution (García-Pintado et al., 2013) and a sub-optimal use of the assimilation filter. Resampling methods [e.g Gordon et al. (1993)] have been used to prevent the collapse of the ensemble: particles with significant weights are replicated and non-significant particles are discarded. Even though resampling is powerful in reducing degeneracy, it often comes with a sample impoverishment and a poor representation of the actual uncertainty of the system (Moradkhani et al., 2012). After few iterations, replicated particles will hardly diversify and particles will again collapse into a single or few particles. According to Snyder et al. (2008), the number of particles should grow exponentially with the dimension of the system, otherwise the PF may suffer from degeneracy. Of course, a higher number of particles implies an increased computational cost which may hamper the use of DA in near real-time application. As a consequence, it is important to minimize the weight variance so that each particle keeps a significant weight.

Hostache et al. (2018) and Di Mauro et al. (2021) recently developed, following a similar previous work by Giustarini et al. (2011), a data assimilation framework based on Sequential Importance Sampling (SIS), a variant

of PFs that enables an efficient assimilation of SAR data into a hydrodynamic model. In their experiment, the rainfall forcing and the SAR data are assumed to represent the only sources of uncertainty. While Di Mauro et al. (2021) showed that the SIS method provides good results when the assumptions are indeed satisfied, they also highlight the need for a method to mitigate degeneracy and sample impoverishment. The assimilation via a SIS tends to degenerate with only a few particles getting significant weights as a result of the assimilation. A preliminary attempt to mitigate the degeneracy consisted in using a tempering coefficient for the inflation of the posterior probability. The likelihood was raised to the power of a coefficient whose value enables a substantial increase of the likelihood variance. However, using this coefficient to inflate the likelihood only partially solved the degeneracy issue, and sometimes at the cost of a decrease in prediction accuracy.
In the literature, in order to mitigate the mentioned PF-related issues, the following approaches have been adopted:

1. Using a correct proposal density to steer particles in such a way that they obtain similar weights (Doucet et al., 2001);
2. Localizing PFs (Van Leeuwen, 2009; Reich, 2013);
3. Combining the PF with the Ensemble Kalman filter (Van Leeuwen, 2009; Potthast et al., 2019; Frei & Kunsch, 2013);
4. Moving the particles from the prior to the posterior by applying a smooth transition process (Beskos et al., 2014).

In this study, the research focuses on the 4^{th} type of approach. We adopt and evaluate a novel enhanced PF following the results of the previous studies by Di Mauro et al. (2021) and Hostache et al. (2018). The novel DA approach, called tempered particle filter (TPF), applies tempering coefficients to inflate the likelihood within an iterative process so that the Bayes' formula is respected (Beskos et al., 2014). Based on the method first proposed by R. M. Neal (1996), we use the implementation of Herbst & Schorfheide (2017) in this paper. The iterative assimilation approach is based on successive Sequential Importance Resamplings (SIRs) and particle mutations. The mutations enable the ensemble to regain diversity after each resampling step and are based on a Metropolis Hasting (MH) algorithm. To the knowledge of the authors, this methodology has never been applied in hydrological sciences and, more specifically, for improving flood simulations. We hypothesize that the proposed innovative DA methodology enables the mitigation of some PF limitations, namely sample degeneracy and sample impoverishment, while preserving the assimilation performances in terms of flood extent, discharge and water level simulations.

In this study, we also further investigate the additional benefits that come from this new approach. According to Dasgupta et al. (2021a), degeneracy plays a crucial role in the persistence of the assimilation benefits over several time steps. Therefore the TPF approach could also help improving the persistence of the assimilation benefits. Moreover, DA algorithms often assume that the observations as well as the model predictions are unbiased. Many authors pointed out the importance of bias removal before the DA, but it is not a straightforward procedure, especially in model forecasts (De Lannoy et al., 2007). Bias can depend on the model structure or parameters, on the initial conditions, or on forcing errors (especially when the forcings are derived from a forecast model, as in this study). In this context, we hypothesize that the new approach based on a TPF enables the reduction of bias in the model predictions and we test this hypothesis. To enable a meaningful evaluation and to verify whether the new approach outperforms the previous one, the TPF performance is compared to that of the SIS.

In this study, we carry out twin experiments based on a synthetically generated data-set with controlled uncertainty. The SAR observations are synthetically generated from the simulated flood extent maps and assimilated into a coupled hydrologic-hydraulic model. Two different background ensembles, i.e., Open Loops, are drawn and used: in the first case the ensemble encompasses the synthetic truth most of the time, in the second case the ensemble is most of the time outside the ensemble range. The objectives of this study are therefore i) to evaluate whether the proposed method can mitigate degeneracy, ii) to evaluate whether the proposed framework improves the prediction accuracy and increases the persistence of the assimilation benefits, iii) to evaluate the efficiency of the method in reducing forecast bias. The paper is structured as follows: section 1 describes the materials and methods, section 2 showcases and discusses the results and 3 draws the conclusions of the study.

1 Materials and Methods

The first part of this section presents the structure of the flood forecasting system. The second part describes the proposed assimilation framework based on a TPF. The experimental design, case study, and the performance metrics used within this experiment are introduced in the last part.

1.1 The flood forecasting model

We use the ERA5 data-set (Hersbach et al., 2019) to derive the forcing of the flood forecasting system. Rainfall and 2 m air temperature at a spatial resolution of approximately 25 km and a temporal resolution of 1 hour are used as inputs to the flood forecasting system. A conceptual hydrological modelling (SUPERFLEX) coupled with a hydraulic model (LISFLOOD-FP) approach has been adopted: the run-off estimated with the hydrological model is used as input to the shallow water hydraulic model. In this study, the rainfall-runoff model SUPERFLEX (Fenicia et al., 2011) is a lumped conceptual model (Figure 1) and is composed of three reservoirs: an unsaturated soil reservoir with a storage S_{UR} representing the root zone, a fast reservoir with storage S_{FR} representing the fast responding components (e.g., the riparian zone and preferential flow paths), and a slow reservoir with storage S_{SR} representing slow responding components (e.g., deep groundwater). A lag function is used at the outlet of the unsaturated soil reservoir to enable a delayed hydrological response of the basin under intense rainfall conditions. The hydraulic model is based on LISFLOOD-FP (Bates & Roo, 2000; J. Neal et al., 2012) and simulates flood extent, water level and discharge within the hydraulic model domain.

ERA5 rainfall time series are used to generate the synthetic truth and are also perturbed to generate an Open Loop (OL) simulations consisting in 32 particles. These 32 particles are then used as input to the flood forecasting model to obtain the ensemble of flood extent maps. We adopt the method proposed and detailed in Di Mauro et al. (2021) to generate synthetic observations from model results. The flood extent map of the synthetic truth together with a real SAR observation are used to compute Probabilistic Flood Maps (PFMs) where each pixel represents the probability to be flooded given the recorded backscatter values. During the analysis (i.e., assimilation) step, the generated PFMs are assimilated into the ensemble of wet-dry maps via the TPF to obtain the updated particles. The following section describes the data assimilation framework.

1.2 Data assimilation framework

PFs are based on Bayes' theorem:

$$p(x^k \mid y^k) = \frac{p(y^k \mid x^k)}{p(y^k)} p(x^k) \tag{1}$$

The observation y at time k, which is the probability to be flooded given the SAR backscatter value, is combined with the forecasts of the numerical model x at time k. The posterior probability $p(x^k \mid y^k)$ is computed by multiplying the prior probability density function $p(x^k)$, which is the probability of the model before any observation is taken into account, with the likelihood $p(y^k \mid x^k)$ that is the probability density that the model state x^n produces the observation. In PFs the prior PDF is drawn from an ensemble of model states of size N

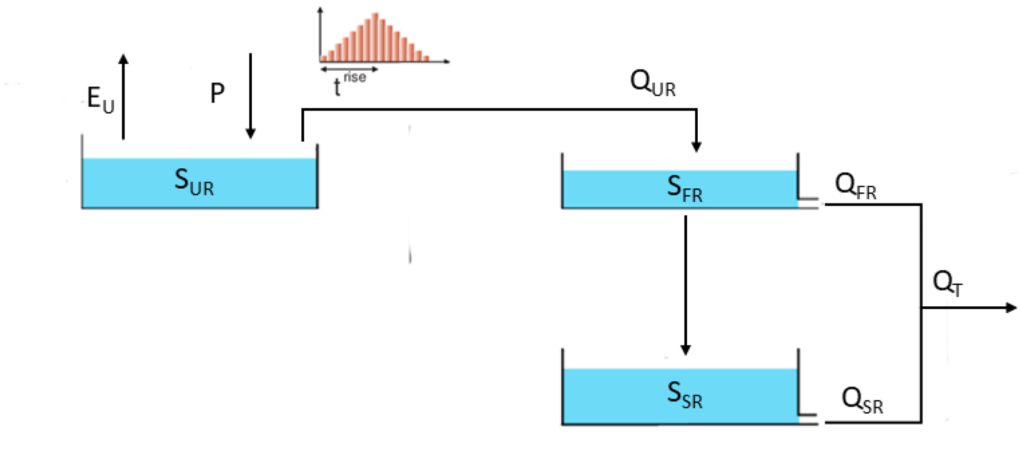

Figure 1. Scheme of the SUPERFLEX model used in this study. The hydrological model is based on three reservoirs: and unsaturated soil reservoir (S$_{UR}$), a fast run-off reservoir (S$_{FR}$) and a slow run-off reservoir (S$_{SR}$). The discharge deriving from the 3 reservoirs are: Q$_{UR}$, Q$_{FR}$, Q$_{SR}$. A triangular lag function with a base length equal to $2 \cdot t^{rise}$ is applied at the outflow of the unsaturated soil reservoir. E$_U$ and P represents the potential evaporation and rainfall respectively.

called particles. Eq. 2 represents the computation of the prior probability:

$$p(x^n) \approx \sum_{n=1}^{N} \frac{1}{N} \delta(x^k - x_n^k) \tag{2}$$

where δ is the Dirac delta function. Inserting Eq. 2 into Eq. 1 leads to the posterior probability formula:

$$p(x^k \mid y^k) \approx \sum_{n=1}^{N} W_n \delta(x^k - x_n^k) \quad \text{where} \quad W_n = \frac{p(y^k|x^k)}{p(y^k)} \tag{3}$$

The weights W_n, hereafter called global weights, were computed by the multiplication of the pixel-based local weights w_i^n, according to the formula by Hostache et al. (2018), assuming that observation errors are independent across space. Di Mauro et al. (2021) showed that the set of particles tends to degenerate: after the assimilation, the number of particles with significant weight is reduced to a few and the posterior distribution is poorly approximated. Di Mauro et al. (2021) made a first attempt to reduce degeneracy using a tempering coefficient γ according to the formula:

$$p(x^k \mid y^k) = \left(\frac{p(y^k \mid x^k)}{p(y^k)} \right)^{\gamma} p(x^k) \quad with \;\; \gamma \in [0, 1] \tag{4}$$

This technical solution enables inflating the posterior variance so that several particles keep significant weight. In the current study we aim to further improve the application of the likelihood tempering. The proposed method relies on the factorisation of the likelihood through an iterative approach according to the following formula:

$$\frac{p(y^k \mid x^k)}{p(y^k)} = \prod_{k=1}^{K} \left(\frac{p(y^k \mid x^k)}{p(y^k)} \right)^{\gamma_s} \tag{5}$$

where $0 < \gamma_s < 1$ for each iteration s, and $\sum_{s=1}^{S} \gamma_s = 1$.

This factorization enables application of the Bayes' theorem iteratively so that the transition from the prior to the posterior probability is smoothly processed. For instance, after one iteration the factorization leads to the following equation:

$$p(y^k \mid x^k) = \prod_{s=2}^{S} \left(\frac{p(y^k \mid x^k)}{p(y^k)} \right)^{\gamma_s} \left(\frac{p(y^k \mid x^k)}{p(y^k)} \right)^{\gamma_1} p(x^k) = \prod_{s=2}^{S} \left(\frac{p(y^k \mid x^k)}{p(y^k)} \right)^{\gamma_s} p_1(x^k \mid y^k) p(x^k) \tag{6}$$

and:

$$p_1(x^k \mid y^k) \approx \sum_{n=1}^{N} W_n^{(1)} \delta(x^k - x_n^k) \quad with \;\; W_n^{(1)} = \left(\frac{p(y^k \mid x^k)}{p(y^k)} \right)^{\gamma_1} \tag{7}$$

At each iteration s, the tempering coefficient γ_s enables inflation of the likelihood variance and reduction of the weight variance, therefore reducing degeneracy. The exponent γ_s is computed so that it allows to keep a substantial number of particles with significant weights at each SIS step. This is carried out through the computation of γ_s providing a target value of the ensemble inefficiency ratio ($InEff$), defined as follows:

$$InEff(\gamma_s) = \frac{1}{N} \sum_{n=1}^{N} (W_n^s(\gamma_s))^2 \tag{8}$$

If $InEff(1) \geq r^*$ (where r^* is a predefined target) then γ_s is the solution to $InEff(\gamma_s) = r^*$ otherwise $\gamma_S = 1 - \sum_{s=1}^{S-1} \gamma_s$, and the iterations are finished.

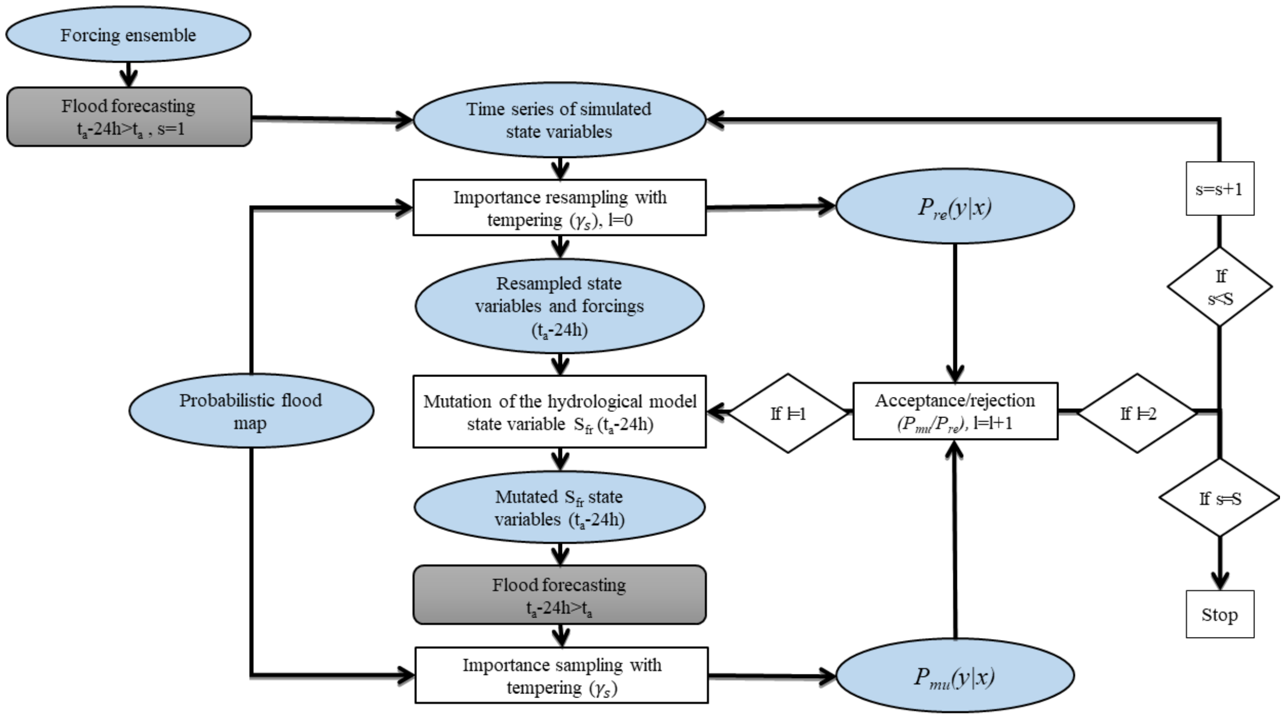

Figure 2. Flow chart of the DA framework where synthetic probabilistic flood maps are generated from flood extents, derived from a truth run, and assimilated within the same flood forecasting model. The flood forecasting model is represented with a grey rectangle, mathematical operations with a white rectangle, state variables, input and observations with a blue ellipse.

After each iteration s, the particles with high weights are resampled using the SIR algorithm proposed by Gordon et al. (1993). Particles are replicated proportionally with their weights: those with an associated low importance weight are replaced with replicas of those having higher weight. After resampling, particles are equally weighted.

Next, a mutation is applied to the fast run-off reservoir level (S_{FR}), a variable of the hydrological model, 24 hours prior to the assimilation to regain diversity within the particle ensemble and the mutated value is used as initial condition for a subsequent model simulation over the 24 hours preceding the assimilation time. Mutating the hydrological state variable 24h prior to the assimilation time and carrying out the related model simulation is done in order to update the hydrological and hydraulic models in a more consistent way. Indeed, it is important to remind here that the water depths simulated by the hydraulic model at a certain time are the result not only of the current upstream streamflow condition but also of the past time series of upstream streamflow conditions.

This mutation is carried out using a MH algorithm, based on a random perturbation via the steps of Markov chain Monte Carlo (MCMC) methods. The MH is based on two steps: first, draw a new particle from a proposal density $q(x^*)$ as $x^* \sim q(x \mid x_j^{k-1})$, and then calculate the MH acceptance ratio:

$$\alpha = \left(\frac{p(y^k \mid x^*)\, p(x^*)}{p(y^k \mid x_j^k)\, p(x_j^k)} \right) \left(\frac{q(x^n \mid x_i^{k-1})}{q(x^* \mid x_j^{k-1})} \right) \tag{9}$$

Many possibilities are available for choosing $q(x^*)$. Here we make use of the one equal to the prior to simplify the evaluation of α. The acceptance ratio becomes:

$$\alpha = \left(\frac{p(y^k \mid x^*)}{p(y^k \mid x_j^k)} \right) \tag{10}$$

where x_j^k represents the particles with high weight that have been resampled. A random variable $u \sim U[0,1]$ is drawn and the mutated particle is accepted if $\alpha > u$, otherwise we keep the particle as before its mutation. Applying the MCMC requires to draw samples from the posterior which can be burdensome because the prior is unknown. The prior can be rewritten using the prior at time $k-1$ as:

$$p(x^k) = \int p(x^k \mid x^{k-1}) p(x^{k-1}) dx^{k-1} \approx \int p(x^k \mid x^{k-1}) \frac{1}{N} \sum_{n=1}^{N} \delta(x^{k-1} - x_i^{k-1}) dx^{k-1} \approx \frac{1}{N} \sum_{n=1}^{N} p(x^n \mid x_n^{k-1}) \tag{11}$$

With this formulation we can write the posterior at first iteration as follows:

$$p_1(x^k \mid y^k) = \left[\frac{p(y^k \mid x^k)}{p(y^k)} \right]^{\gamma_1} \frac{1}{N} \sum_{n=1}^{N} p(x^k \mid x_n^{k-1}) \tag{12}$$

As proposed by Herbst & Schorfheide (2017), the mutation is carried out based on a proposed innovation $p(x^* \mid x^{k-1}) = N(0, c_k^2 \cdot \sigma)$, with c_k being a scaling factor given by the following equation:

$$c_n = c_{n0} \left(0.95 + 0.10 \cdot \frac{e^{20 \cdot (\alpha - 0.4)}}{1 + e^{20 \cdot (\alpha - 0.4)}} \right) \tag{13}$$

The mutation step is repeated for $l = 1, .., N^{MH}$. In our study $N^{MH} = 2$.

In detail, the method is structured according to the following time steps (Figure 2):

- Ensemble forcing are used as input to the flood forecasting model;
- The hydrodynamic simulations are carried out over the 24 hours prior to the assimilation.
- Calculate $p(y|x_i)$ for each particle i and find γ_1 such that $InEff(1) \geq r*$.
- Particles are resampled using the tempered weights. The particles after resampling that are duplicates of particles with high weights are perturbed at time t_a-24 hours.
- New hydrodynamic simulations with the mutated levels of the S_{FR} are carried out during the 24 hours prior to the assimilation.
- The likelihood of the mutated particles $p_{mu}(y \mid x)$ is compared to the likelihood of the resampled particles $p_{re}(y \mid x)$.
- The resampled particles are replaced by the mutated particles if the ratio of the two is larger than a value randomly taken from the interval [0, 1].
- The mutation step is repeated twice.
- The iteration with a new tempering coefficient is realized.

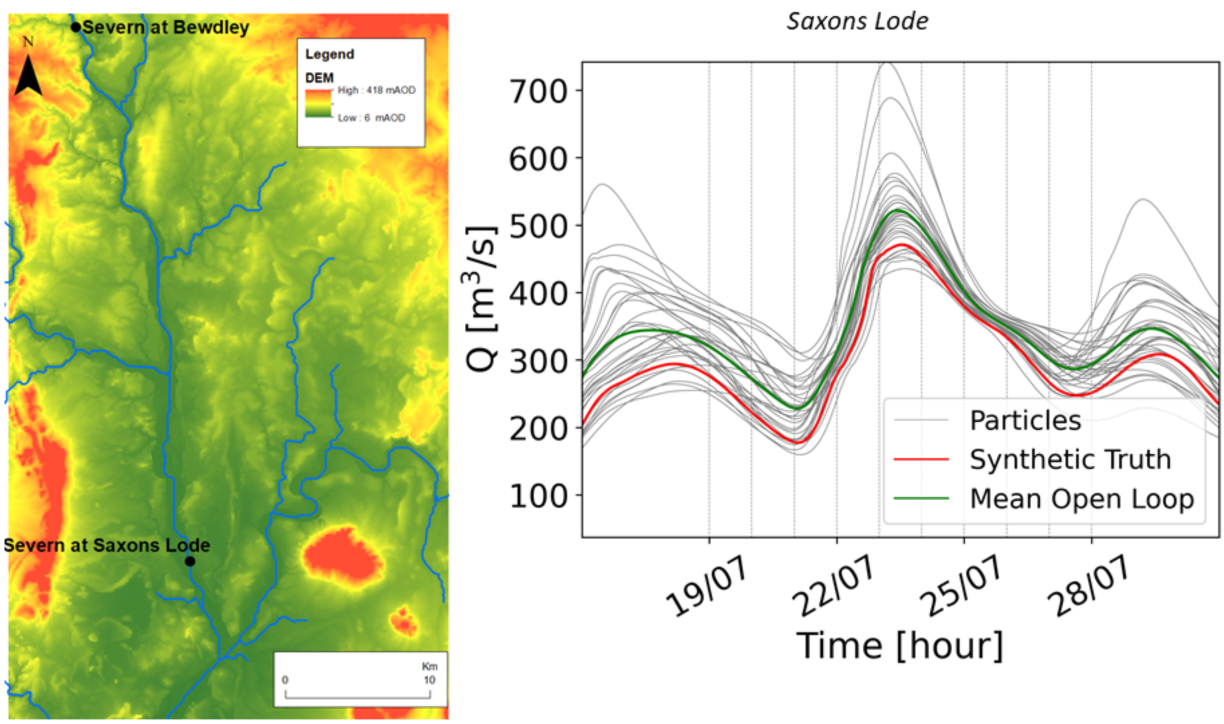

Figure 3. Study area of the synthetic experiment (left). Black dots correspond to the points where evaluation of the DA performances is carried out ("Severn at Bewdley" and "Severn at Saxons Lode"). Ensemble time series of discharge in Saxons Lode and assimilation times (right). Gray lines correspond to the Open Loop (OL), the red line corresponds to the synthetic truth, the green line corresponds to the mean of the OL. The dashed lines correspond to the different assimilation time steps performed independently every 24 hours from 19/07 00:00 to 28/07 00:00.

- The entire process is repeated until the sum of the tempering coefficients is equal to unity.

1.3 Experimental design, case study and performance metrics

The study area is the lower river Severn located in the United Kingdom (Figure 3, on the left). To analyze the filter performances at different assimilation times, SAR images have been synthetically generated [see Di Mauro et al. (2021)] every 24 hours from 07/19 00:00 to 07/28 00:00 (Figure 3, on the right) and the 10 corresponding independent assimilations are carried out and evaluated. The flood event has been simulated using the rainfall and temperature (ERA-5 dataset) time series corresponding to the July 2007 event as input data to the flood forecasting system.

Further details concerning the hydrological and hydraulic model set-up as well as the study area of the synthetic experiment, are provided in our previous study (Di Mauro et al., 2021). In this study, the ensemble contains 32 particles. The proposed TPF is characterised by a particle mutation at each iteration. The mutation step could have a key-role, especially when the ensemble is biased with respect to the observations. On the one hand, in the SIS case the weighted mean (also called expectation) is based on the initial particles of the ensemble meaning that if the truth falls outside the ensemble range the expectation cannot reach the synthetic truth. On the other hand, in the TPF case the particles can mutate and move outside the initial ensemble range. This way the expectation can potentially reach the synthetic truth. For evaluating the capability of the TPF to compensate for bias within the ensemble, two different cases are investigated. The difference between the OL and the synthetic truth (O) rainfall time series averaged over the flood event period (K) represents the mean bias error (MBE, equation 14) and it is used to estimate the bias. For a "markedly" biased case MBE is 0.92 $\frac{mm}{h}$ while for a "limited" bias case the MBE is 0.14 $\frac{mm}{h}$, meaning that the error of the markedly biased case is 6.56 times larger than for the other case.

$$MBE = \frac{1}{K}\sum_{k=1}^{K}(OL_k - O_k) \tag{14}$$

In the limited case the synthetic truth is most of the time within the ensemble range; in the other case the ensemble is conspicuously biased and the synthetic truth falls outside the ensemble range most of the time. The assimilation steps are performed at the same time for both cases and the same observations are used. Results are analyzed according to different spatial (global and local) and temporal scales (at the assimilation time and for the subsequent time steps). The filter performances are evaluated in terms of predicted flood extent and water depth maps, as well as local discharge and water levels time series. The performance metrics are assessed by comparing the results of the TPF with those of the OL. Moreover, the TPF is compared with the SIS method applied in our previous study Di Mauro et al. (2021).The local evaluation of the prediction accuracy of water levels and discharge is performed by comparing the simulated discharge and water level time series with respect to the synthetic truth.
The following performance metrics are used:

- Confusion matrices: a matrix providing the number of false negatives (under-prediction) and false positives (over-prediction), together with correct positives and negatives;
- Contingency maps: maps comparing the simulated flood map with the synthetic truth map;
- Critical success index (CSI): a metric that evaluates the accuracy of the flood map predictions and is defined as the ratio between the number of pixels correctly predicted as flooded over the sum of predicted flooded pixels (correct positives, false positives and false negatives). It ranges from 0, complete disagreement, to one, perfect match;
- Root mean square error (RMSE): it is given by the square root of the mean of the squares of the deviations of the predicted water levels against the synthetic truth over the hydraulic model domain. It evaluates the prediction errors of a state variable, in our case the water levels.

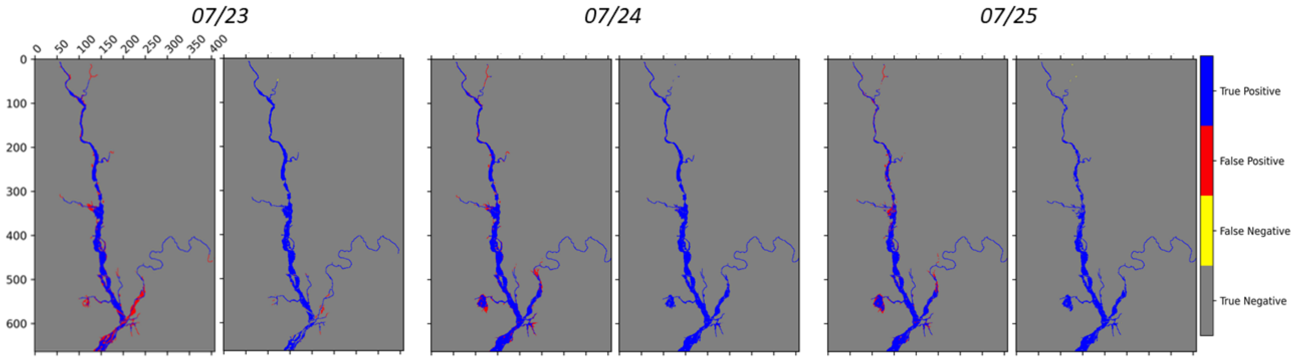

Figure 4. Contingency maps of the Open Loop (left) and after the assimilation (right) for three different assimilations at time 07/23 00:00, 07/24 00:00, 07/25 00:00. Red pixels correspond to over-prediction (false positives) errors, yellow pixels to under-prediction (false negatives) errors, pixels correctly classified as not-flooded are in grey and when the contrary occurs pixels are in blue.

2 Results and discussions

2.1 TPF-based assimilation performances

2.1.1 Flood extent map predictions

The flood extent maps are evaluated via different performance metrics: the contingency maps, the CSI and the confusion matrix. The contingency map is derived from the comparison between the simulated flood extent map (i.e. expectation) and the validation map which is derived from the synthetic truth simulation in our case. The contingency maps, corresponding to 3 different assimilation time steps (rising limb, peak, falling limb), are shown in Figure 4.

Yellow and red pixels correspond to errors of under-prediction (when the model wrongly predicts the pixels as not-flooded) and over-prediction (the opposite case), respectively. In Figure 4, the reported images for each assimilation time correspond to the OL (on the left) and the TPF analysis (on the right). Over-prediction represents the most frequent type of error and it is significantly reduced as a result of the TPF-based assimilation.

The decrease of wrongly predicted pixels is quantified in the confusion matrix reported in Table 1. In line with Figure 4, after any of the three assimilation time steps, the number of over-prediction errors is reduced by 90% or more, while the number of under-predicted pixels increases in the upstream part of the river. However, they represent only 0.3% or less of the total number of flooded pixels.

Time series of CSI are also used to evaluate the TPF performances (Figure 5). They allow to evaluate the predicted flood extent maps not only at the assimilation time step (as for the contingency maps and the confusion matrices) but also for subsequent time steps. Moreover, they provide an assessment of the persistence of the improvements over longer lead times after the assimilation. Figure 5 shows the time series of CSI before (black line) and after (blue line) the assimilation of SAR images taken during the rising limb (07/23 00:00),

Method		07/23 00:00		07/24 00:00		07/25 00:00	
		PF	**PN**	**PF**	**PN**	**PF**	**PN**
Open Loop	**TF**	7497	0	9374	0	8390	1
	TN	2441	260974	1356	260182	1219	261302
TPF	**TF**	7475	22	9374	22	8378	13
	TN	204	263211	78	261460	30	262491

Table 1. Confusion matrix of the Open Loop and Tempered Particle Filter analysis for three different time steps (07/23 00:00, 07/24 00:00, 07/25 00:00): TF= flooded pixels in the truth map, TN= not-flooded pixels in the truth map, PF= predicted flooded pixels, PN=predicted non-flooded pixels.

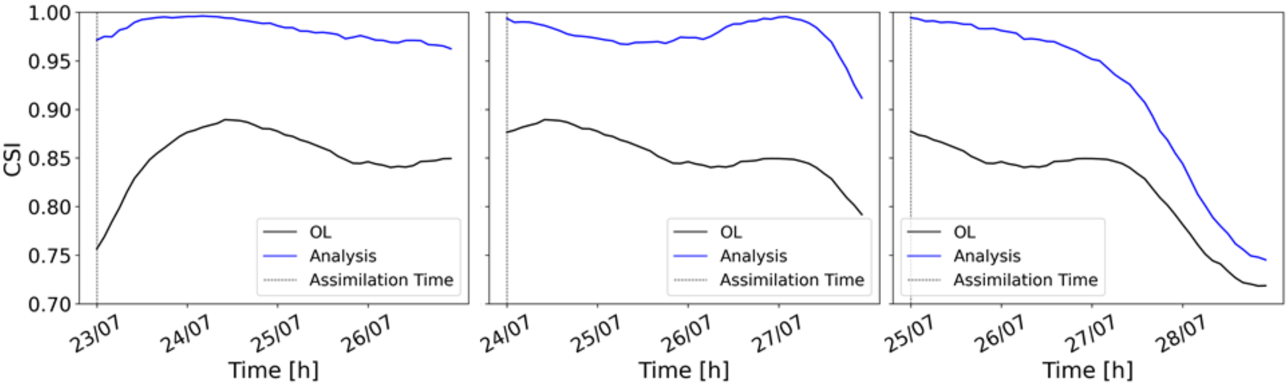

Figure 5. Hourly time series of the Critical Success Index of the Open Loop (black line) and Tempered Particle Filter analysis (blue line) due to the assimilation of 3 different images: during the rising limb (07/23 00:00), at the peak (07/24 00:00) and during the falling limb (07/25 00:00).

at the peak (07/24 00:00) and during the falling limb (07/25 00:00) of the flood event. This figure shows an improvement of the analysis compared to the OL not only at the assimilation time but also over subsequent time steps: on average, CSI improvements persist for more than 3 days after the TPF application.

2.1.2 Water level and discharge predictions

To further investigate the TPF assimilation performance we evaluate water level and discharge predictions. This evaluation is carried out first at specific points along the river Severn: in Bewdley (the gauge station located at the upstream boundary of the hydraulic model domain), and in Saxons Lode (within the hydraulic domain). In Figures 6, the discharge at Bewdley (on the left) and at Saxons Lode (on the right) are plotted. The analysis expectation of discharge (blue line) moves closer to the synthetic truth (red line) at the two stations as a result of the assimilation showing a substantial improvement of the predictions. Here we show the results from the assimilation on July 23^{th} 00:00 as an illustrative example since the other assimilations produce similar effects.

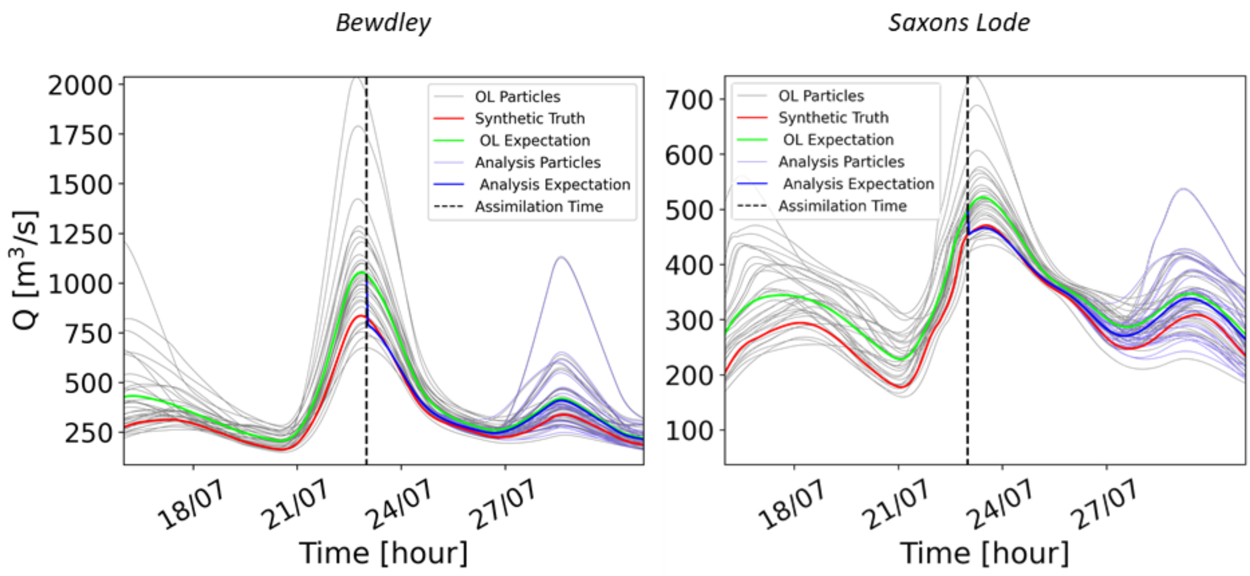

Figure 6. Time series of discharge at the peak at Bewdley and at Saxons Lode with the assimilation of an image at 07/23 00:00. The vertical dashed lines indicate the time of the assimilation. The gray lines correspond to the OL particles, the green line to the OL mean, the light blue lines to the analysis particles and the blue line to the analysis expectation. The synthetic truth is represented by a red line.

In Figure 6, it can be observed that the degeneracy is mitigated. At the assimilation time, the analysis particles are very similar and close to the synthetic truth, but rapidly regain diversity, thereby avoiding degeneracy. After more than 3 days, the particles returns to their initial trajectories (i.e. the OL) mainly because precipitation uncertainty seems to prevail in the forecasts from that moment on.

To generalize the evaluation made for the gauging stations, we evaluate the accuracy of water level predictions globally, using time series of RMSE computed over the entire hydraulic model domain. This index has been calculated at the assimilation time and for subsequent time steps, in order to assess if the assimilation benefits persist in time. In Figure 7, the RMSE of the analysis is lower than the OL and this improvement lasts for more than 3 days following the assimilation. As for the CSI plots, the improvements of RMSE start dropping more quickly for the assimilation during the falling limb (07/25 00:00) in Saxons Lode compared to the assimilation of the SAR image at the peak or during the rising limb. The standard deviation of the errors has also been computed in order to evaluate the accuracy of the second moment. In this case the standard deviation represents the dispersion of the errors (given as the difference between the expectation and the true water levels). Results show that the TPF application determines less dispersed and more clustered results around the synthetic truth.

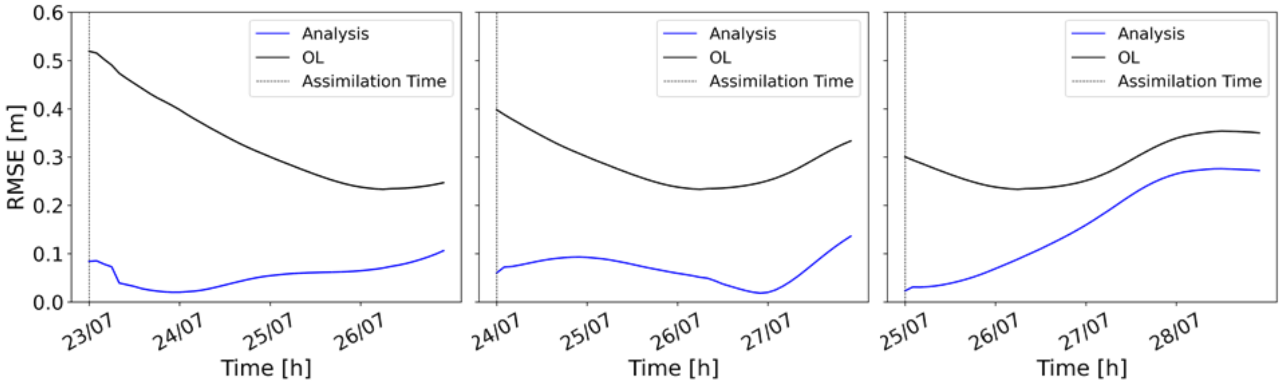

Figure 7. Hourly time series of the standard deviation of the errors due to the assimilation of 3 different images: 07/23 00:00, 07/24 00:00, and 07/25 00:00. The standard deviation of the errors as difference between the OL and the true water levels (black line) and as difference between the analysis expectation and the true water levels (blue line).

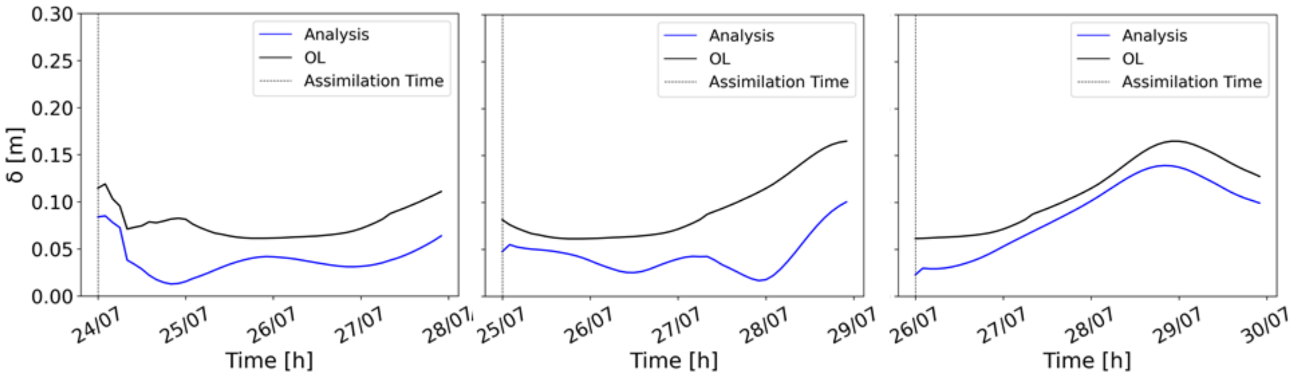

Figure 8. Hourly time series of the RMSE. Black line refers to the OL and blue line to the analysis results after the assimilations of 3 different images (07/23 00:00, 07/24 00:00, and 07/25 00:00).

2.2 Comparison between TPF- and SIS-based assimilation experiments with unbiased background

We showed in section 2.1 that the TPF improves the predictions of water levels and discharge, as well as flood extent. In this section, the new TPF-based DA framework is compared with the SIS approach previously proposed by Di Mauro et al. (2021). To do so, we apply the SIS method as proposed in Di Mauro et al. (2021) on the same 32 background particles (i.e., OL) and the same synthetically generated flood extent observations. The choice of comparing the TPF with this SIS is related to the fact that other methods reported in Di Mauro et al. (2021) were providing comparable performances, and therefore, SIS has been chosen as a benchmark. In terms of flood extent, the comparison is realized using the hourly time series of the CSI index (Figure 9).

In Figure 9, the blue line corresponds to the CSI of the forecast obtained from the TPF-based case, the orange line to the one obtained from the SIS-based case and the black line to the one of the OL. The 3 plots correspond respectively to the assimilation on July 23 00:00, july 24 00:00 and July 25 00:00. The

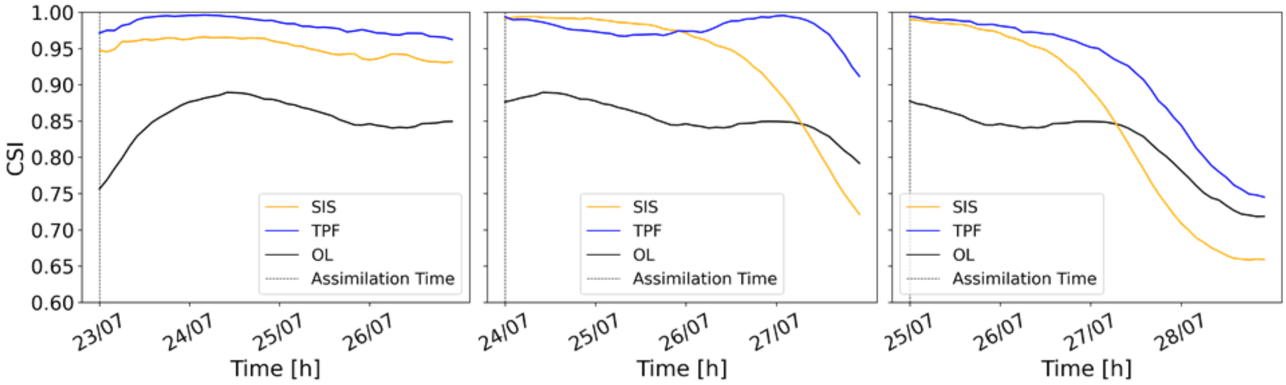

Figure 9. Comparison of the hourly time series of the Critical Success Index of the OL (black line), TPF analysis (blue line) and SIS analysis (orange line) due to the assimilation of 3 different images: 07/23 00:00, 07/24 00:00, and 07/25 00:00.

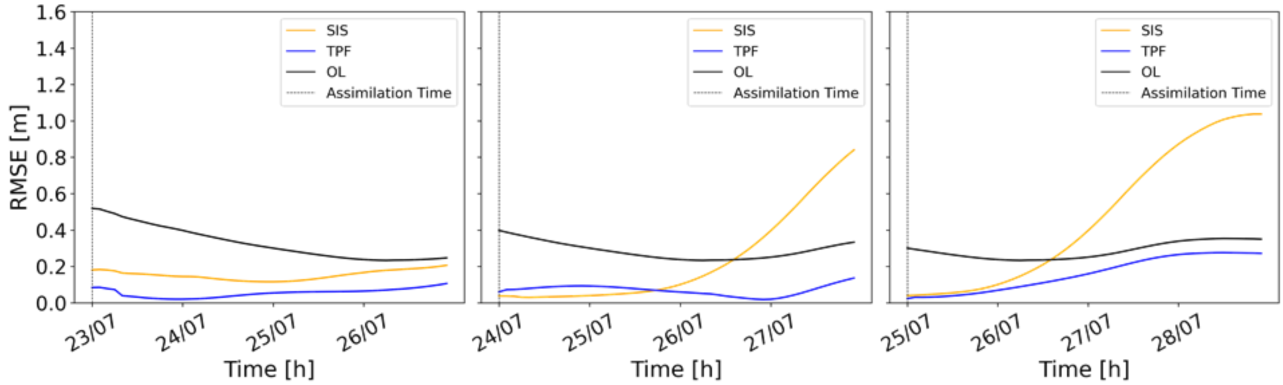

Figure 10. Hourly Root Mean Square Error (RMSE) time series. The black line represents the RMSE of the OL, the blue line the TPF-based RMSE and the orange line the SIS-based RMSE. 3 different assimilation cases are plotted: 07/23 00:00, 07/24 00:00, and 07/25 00:00.

CSI values obtained when assimilating an image during the rising limb are systematically higher for the TPF. When the image is assimilated close to the peak and during the falling limb, CSI values of the TPF and SIS-based assimilation are very similar at the assimilation time and for subsequent time steps. After 2 days, the performance of the SIS becomes substantially worse than that of the TPF.

We have also compared the performances of the SIS and the TPF using time series of RMSE (Figure 10). As expected, the RMSE time series exhibit very similar trend to the CSI: the RMSE is lower with the TPF experiment when assimilating an image during the rising limb. For the other two assimilation steps RMSE values are comparable, but performances of the SIS decrease more rapidly, especially after 2 days. Overall, Figures 9 and 10 clearly show the beneficial effects of the TPF assimilation on the long-term.

Table 2 reports the ratios between the analysis-RMSE and the OL-RMSE for each assimilated SAR image and for different lead times. These ratios were calculated at each hour and for all the different assimilation dates. In the table the values at the assimilation time and for lead times of 6 hours, 1 day, 2 days, 3 days and 4 days are reported. The ratios obtained with the TPF method are shown in the gray cells. The cyan cells contain the ratios obtained with the SIS experiment. The last row of the table shows the mean of the RMSE ratios over the different assimilation times at given prediction lead times. The lower the RMSE ratio values, the better the performance. Ratios of RMSEs lower than unity indicate that the assimilation improves forecasts. Table 2 shows that the TPF-based ratios are most of the time substantially lower than those of the SIS-based ones. For instance, the SIS-based mean ratios for 3 and 4 days of lead times are almost twice that of the TPF-based one. The benefit of the TPF-based assimilation persists for more than 4 days after the assimilation time. Moreover, the TPF-based ratios are always lower than unity, whereas the SIS-based ratios get also values higher than unity.

Table 2. Ratios between the analysis and Open Loop RMSE for each assimilation date and for various lead times. Gray cells refer to the TPF-based method, cyan cells to the SIS-based method.

Image	**Lead time**											
date	0		6 hours		1 day		2 days		3 days		4 days	
07/19	0.25	0.24	0.25	0.24	0.23	0.26	0.20	0.22	0.59	0.57	0.80	0.83
07/20	0.23	0.26	0.22	0.26	0.19	0.22	0.60	0.57	0.83	0.85	0.90	1.08
07/21	0.19	0.22	0.28	0.24	0.62	0.57	0.77	0.85	0.79	1.10	0.76	1.26
07/22	0.27	0.25	0.30	0.29	0.35	0.35	0.31	0.36	0.23	0.39	0.27	0.67
07/23	0.16	0.35	0.15	0.36	0.05	0.36	0.18	0.39	0.27	0.70	0.43	0.84
07/24	0.15	0.09	0.19	0.09	0.31	0.13	0.25	0.42	0.08	1.58	0.41	2.52
07/25	0.08	0.13	0.11	0.16	0.29	0.42	0.63	1.58	0.78	2.57	0.78	2.96
07/26	0.17	0.23	0.17	0.25	0.25	0.20	0.54	0.24	0.63	0.38	0.64	0.72
07/27	0.11	0.18	0.12	0.16	0.26	0.24	0.38	0.41	0.49	0.69	0.56	1.20
07/28	0.15	0.24	0.23	0.29	0.36	0.41	0.54	0.69	0.63	1.26	-	-
Mean	0.17	0.21	0.19	0.22	0.25	0.29	0.39	0.48	0.44	0.85	0.58	1.1

2.3 Comparison between TPF- and SIS-based assimilation experiments with biased background

In this last experiment, we use the same set-up as in the previous experiment but with the exception of a modified OL. We have introduced a perturbation error to the ERA-5 rainfall time series so that the bias in the ensemble is 6.56 times larger than in the previous case. The ensemble has significant bias and the synthetic truth is most of the time located outside of the ensemble range as can be see in Figure 11. For the evaluation of the

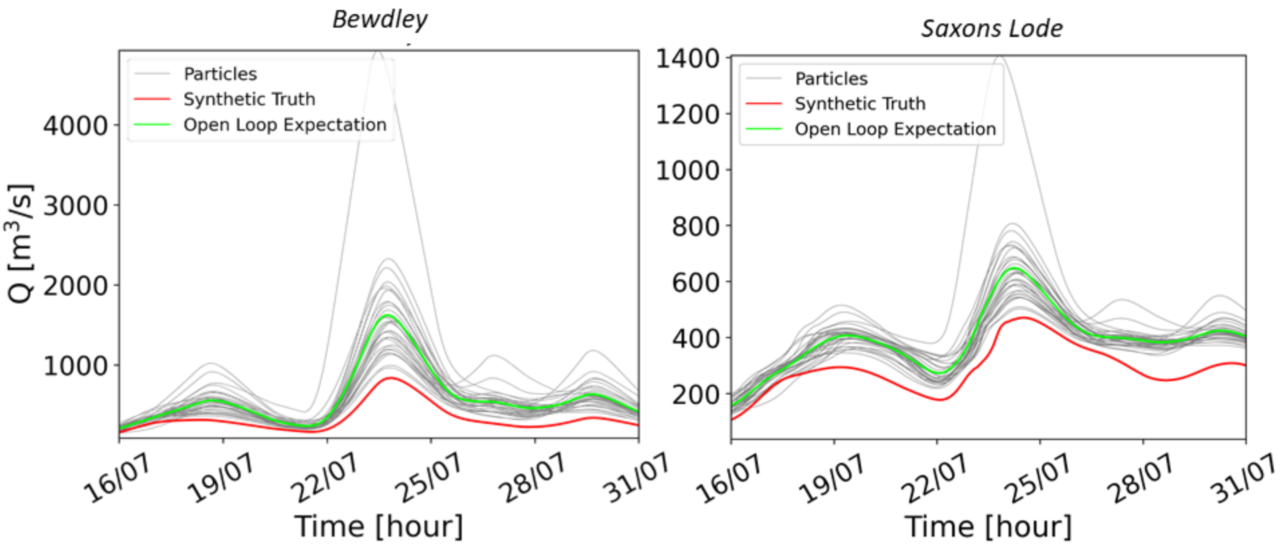

Figure 11. Discharge time series ensemble at Bewdley (on the left) and at Saxons Lode (on the right). The OL particles are represented with gray lines, the synthetic truth is represented by the red line. The OL expectation is in green. In this case, the ensemble is markedly biased; the synthetic truth falls outside the ensemble range most of the time.

results, the same performance indices and the same plots are used. The ratios between the analysis-RMSE and the OL-RMSE for each assimilated SAR image and for different lead times are reported in the Table 3. At the assimilation time and for more than one day after that, the TPF-based assimilation is capable of substantially reducing the forecast bias. The SIS is less efficient in that respect, as RMSE ratios are larger for the SIS-based assimilation. For longer lead times, the error in water levels increases due to the bias in the rainfall ensemble and the RMSE ratios of the TPF-based and the SIS-based assimilation become similar. This is clearly visible in Figure 12 that shows the RMSE time series on July 23^{th}, 24^{th}, and 25^{th} at 00:00. When the bias is limited and the synthetic truth falls inside the ensemble range most of the time, as in the previous case (Figure 7), the forecast improvement lasts for longer lead times. However, when the ensemble is markedly biased (Figure 12), the TPF improves the results at the assimilation time but the level of improvement degrades more quickly compared to the limited biased case.

At the assimilation time, the TPF always improves the accuracy of the results of the flood forecasts (in terms of flood extent, water levels, discharge) with respect to the OL and it is comparable to the SIS performances. An important aspect that emerges from the results is the persistence of the assimilation benefits. They remain significant even 3 days after the TPF assimilation when compared to the SIS performances; nonetheless, performances start degrading with the onset of rainfall over the headwater catchment and rainfall uncertainty prevails in the forecast uncertainty. Moreover, the accuracy of the results is higher when the

Table 3. Ratio between the analysis and Open Loop of the the RMSE for each assimilation date and for various lead times for a markedly biased case. Gray cells refer to the TPF-based method, cyan cells to the SIS-based method.

Image date	**Lead time**											
	0		6 hours		1 day		2 days		3 days		4 days	
07/19	0.19	0.42	0.13	0.42	0.10	0.44	0.26	0.52	0.82	0.53	0.94	0.58
07/20	0.29	0.44	0.25	0.46	0.21	0.52	0.72	0.53	0.88	0.59	0.91	0.67
07/21	0.47	0.52	0.49	0.53	0.54	0.53	0.70	0.59	0.82	0.68	0.71	0.82
07/22	0.47	0.53	0.49	0.52	0.53	0.59	0.70	0.68	0.82	0.83	0.88	0.95
07/23	0.32	0.31	0.31	0.29	0.30	0.26	0.47	0.38	0.71	0.57	0.81	0.64
07/24	0.17	0.26	0.20	0.27	0.39	0.38	0.61	0.57	0.71	0.64	0.78	0.70
07/25	0.15	0.38	0.21	0.43	0.41	0.57	0.55	0.64	0.65	0.71	0.76	0.80
07/26	0.16	0.57	0.18	0.59	0.28	0.64	0.44	0.71	0.61	0.81	0.68	0.87
07/27	0.24	0.52	0.16	0.55	0.34	0.70	0.68	0.96	0.83	1.05	0.78	1.04
07/28	0.34	0.70	0.36	0.77	0.51	0.96	0.65	1.05	0.58	1.04	-	-
Mean	0.26	0.46	0.24	0.48	0.34	0.56	0.55	0.66	0.72	0.74	0.81	0.79

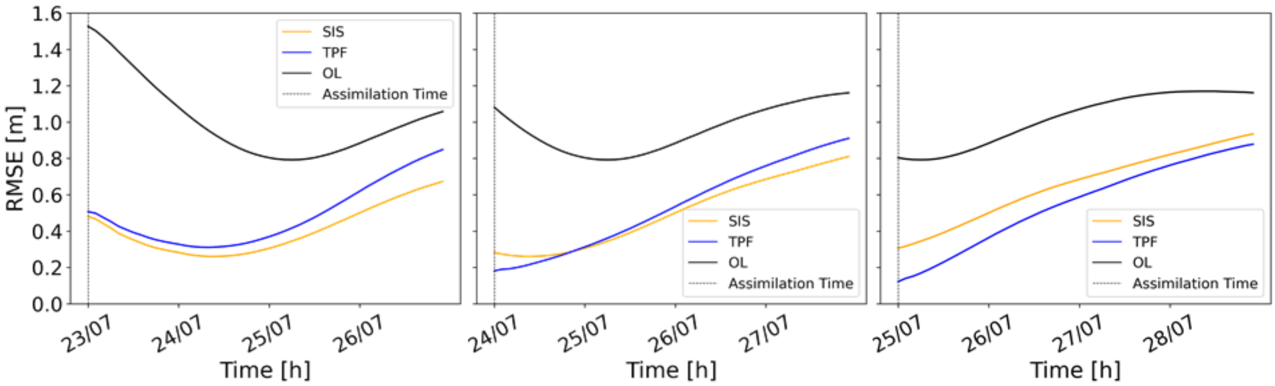

Figure 12. Hourly RMSE time series for a markedly biased ensemble case. The black line represents the RMSE of the OL, the blue line the RMSE after the TPF application and the orange line the RMSE after the SIS application. Assimilation at 07/23 00:00, 07/24 00:00 and 07/25 00:00 are plotted.

observations are assimilated after flood peak when inflow errors are dominating and flood extent is becoming more sensitive to changes in water depth due to the connectivity between the river channel and its floodplain (Dasgupta et al., 2021b). We argue that the marked improvement in the forecast skill is due to the update of the initial conditions of the hydrological model including S_{FR} 24 h prior to the assimiliation time. Better initial conditions of the model forecast are defined at each assimilation time. The runoff that is used as upstream boundaries of the hydraulic model is a function of the storage S_{FR} of the hydrological model. Updating the S_{FR}, and consequently the fast run-off, represents an effective way to increase the long-lasting effects of DA since runoff has the highest uncertainty deriving from poorly known rainfall as already pointed out by Matgen et al.

(2010). This aspect, together with the mitigation of degeneracy, as hypothesized by Dasgupta et al. (2021a), could explain the longer-term persistence of DA benefits via the TPF.

In the markedly biased ensemble case, although the particles move towards the synthetic truth after the TPF application, the amount of rainfall entering the system is too large and the update of the reservoir level is not able to compensate for the error in the rainfall forcing. As a consequence, results obtained using the TPF are sometimes similar to those obtained using the SIS, or even slightly less satisfying when the rainfall intensity is high and rainfall uncertainty dominates the system. The improvements resulting from the update of the initial conditions are then vanished after a few days and the model moves back to the OL state. To increase the time window of the assimilation benefits, the update of hydrological model state variable could be completed by a forcing update or by a parameter update, as in Cooper et al. (2019) where channel friction is updated togheter with a state variable, but with the consequent risk of multiple acceptable solutions of the system according to the equifinality concept (Beven & Freer, 2001).

3 Conclusions

In this paper, we have proposed a new approach based on a Tempered Particle Filter (TPF) to assimilate flood extent maps into a flood forecasting system. The objective of this new data assimilation framework is to mitigate degeneracy and sample impoverishment, well known issues in particle filtering. We have evaluated the performances of the filter in two different cases: with a limited forecast bias and with a more important forecast bias. In addition, the TPF has been compared with the a standard Particle Filter, namely the Sequential Importance Sampling (SIS) as used in previous studies (Hostache et al., 2018; Di Mauro et al., 2021). The following key conclusions are drawn from our experiments:

1. At the time of the assimilation, forecasts are very accurate locally: the forecast overlaps the synthetic truth for all the different assimilation cases and for both analysed locations. Results are very satisfying at a larger scale as well: RMSE and CSI improve systematically as a result of the assimilation. On average, RMSE values decrease by 80% whereas CSI values increase by 30% as a result of the assimilation;
2. Results are also satisfying across time: the CSI and RMSE are improved up to 3 days after the assimilation;
3. Performances are improved compared to the OL and the SIS filter. The benefits of the newly introduced TPF-based assimilation are longer persisting when compared to the effects obtained with assimilation techniques used in the previous studies;
4. The new assimilation framework significantly outperforms the SIS. SIS performance indices are generally comparable to the TPF ones at the assimilation time, but they tend to drop more rapidly, in general 2 days after the assimilation. For example, TPF-based RMSE are 20% lower compared to the SIS-based ones, 2 days after the assimilation;

5. When the ensemble is markedly biased results are significantly improved by the TPF at the assimilation times and for few days after. Afterwards, TPF and SIS based results are similar because the model state update cannot compensate for a too large bias in the precipitation ensemble.

The proposed data assimilation framework based on a TPF holds promise for improving prediction accuracy for longer lead times. In this study, we have shown a synthetic experiment where rainfall and SAR observations are the only sources of uncertainty. In a future study, it will be interesting to apply and evaluate this enhanced approach on a real test case in a weakly controlled environment.

Acknowledgments

The research reported herein was funded by the National Research fund of Luxembourg through the HyDRO-CSI and CASCADE projects. Funding from the Austrian Science Funds as part of the Vienna Doctoral Programme on Water Resources System (DK W1219-N22) is acknowledged. Funding was also provided by the UK Engineering and Physical Sciences Research Council (EPSRC) DARE project (EP/P002331/1). Peter Jan van Leeuwen thanks the European Research Council (ERC) for funding of the CUNDA ERC 694509 project under the European Unions Horizon 2020 research and innovation programme. Nancy K. Nichols was funded in part by the UK Natural Environmental Research Council (NERC) National Centre for Earth Observation (NCEO). The work of Renaud Hostache was supported by the National Research Fund of Luxembourg through the CASCADE Project under Grant C17/SR/11682050.
The Lisflood-FP model can be freely downloaded at `http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood`. The river cross-section data, the digital elevation model, and the gauging station water level, streamflow, and rating curve data are freely available upon request from the Environment Agency (enquiries@environmentagency.gov.uk). The ERA-5 data set is freely available at `https://confluence.ecmwf.int/display/CKB/ERA5`.

References

Andreadis, K. M., & Schumann, G. J.-P. (2014). Estimating the impact of satellite observations on the predictability of large-scale hydraulic models. *Advances in Water Resources*, *73*(C), 44-54. doi: 10.1016/j.advwatres.2014.06.006

Annis, A., Nardi, F., & Castelli, F. (2021). Simultaneous assimilation of water levels from river gauges and satellite flood maps for near-real time flood mapping. *Hydrology and Earth System Sciences Discussions*, *2021*, 1–37. Retrieved from `https://hess.copernicus.org/preprints/hess-2021-125/` doi: 10.5194/hess-2021-125

Bates, P., & Roo, A. D. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, *236*(1), 54 - 77. Retrieved from `http://www.sciencedirect.com/science/article/pii/S002216940000278X` doi: https://doi.org/10.1016/S0022-1694(00)00278-X

Beskos, A., Crisan, D., & Jasra, A. (2014). On the stability of sequential monte carlo methods in high dimensions. *The Annals of Applied Probability*, *24*(4), 1396–1445. Retrieved from http://www.jstor.org/stable/42920481

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology. *Journal of Hydrology*, *249*(1), 11-29. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022169401004218 doi: https://doi.org/10.1016/S0022-1694(01)00421-8

Blöschl, G., Bierkens, M. F., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., ... Zhang, Y. (2019). Twenty-three unsolved problems in hydrology (uph) – a community perspective. *Hydrological Sciences Journal*, *64*(10), 1141-1158. Retrieved from https://doi.org/10.1080/02626667.2019.1620507 doi: 10.1080/02626667.2019.1620507

Chen, X., Sun, Q., & Hu, J. (2018). Generation of complete sar geometric distortion maps based on dem and neighbor gradient algorithm. *Applied Sciences*, *8*(11). Retrieved from https://www.mdpi.com/2076-3417/8/11/2206 doi: 10.3390/app8112206

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. Retrieved from https://hess.copernicus.org/articles/23/2541/2019/ doi: 10.5194/hess-23-2541-2019

Dasgupta, A., Hostache, R., Ramsankaran, R., Schumann, G. J.-P., Grimaldi, S., Pauwels, V. R. N., & Walker, J. P. (2021a). A mutual information-based likelihood function for particle filter flood extent assimilation. *Water Resources Research*, *57*(2), e2020WR027859. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020WR027859 (e2020WR027859 2020WR027859) doi: https://doi.org/10.1029/2020WR027859

Dasgupta, A., Hostache, R., Ramsankaran, R., Schumann, G. J.-P., Grimaldi, S., Pauwels, V. R. N., & Walker, J. P. (2021b). A mutual information-based likelihood function for particle filter flood extent assimilation. *Water Resources Research*, *57*(2), e2020WR027859. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020WR027859 (e2020WR027859 2020WR027859) doi: https://doi.org/10.1029/2020WR027859

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Pauwels, V. R. N., & Verhoest, N. E. C. (2007). Correcting for forecast bias in soil moisture assimilation with the ensemble kalman filter. *Water Resources Research*, *43*(9). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006WR005449 doi: https://doi.org/10.1029/2006WR005449

Di Mauro, C., Hostache, R., Matgen, P., Pelich, R., Chini, M., van Leeuwen, P. J., ... Blöschl, G. (2021). Assimilation of probabilistic flood maps from sar data into a coupled hydrologic–hydraulic forecasting model: a proof of concept. *Hydrology and Earth System Sciences*, *25*(7), 4081–4097. Retrieved from https://hess.copernicus.org/articles/25/4081/2021/ doi: 10.5194/hess-25-4081-2021

Doucet, A., De Freitas, N., & Gordon, N. (2001). Sequential monte carlo methods in practice. , xxvii, 581 p. :.

doi: https://doi.org/10.1007/978-1-4757-3437-9

Fenicia, F., Kavetski, D., & Savenije, H. H. G. (2011). Elements of a flexible approach for conceptual hydrological modeling: 1. motivation and theoretical development. *Water Resources Research*, *47*(11). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR010174 doi: 10.1029/2010WR010174

Frei, M., & Kunsch, H. R. (2013, Jul). Bridging the ensemble kalman and particle filters. *Biometrika*, *100*(4), 781–800. Retrieved from http://dx.doi.org/10.1093/biomet/ast020 doi: 10.1093/biomet/ast020

Garcia-Pintado, J., Mason, D., Dance, S. L., Cloke, H., Neal, J. C., Freer, J., & Bates, P. D. (2015). Satellite-supported flood forecasting in river networks: a real case study. *Journal of Hydrology*, *523*, 706–724. Retrieved from http://centaur.reading.ac.uk/39388/ doi: 10.1016/j.jhydrol.2015.01.084

García-Pintado, J., Neal, J. C., Mason, D. C., Dance, S. L., & Bates, P. D. (2013). Scheduling satellite-based sar acquisition for sequential assimilation of water level observations into flood modelling. *Journal of Hydrology*, *495*, 252-266. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022169413002783 doi: https://doi.org/10.1016/j.jhydrol.2013.03.050

Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza, D., Pauwels, V. R. N., … Savenije, H. H. G. (2011). Assimilating sar-derived water level data into a hydraulic model: a case study. *Hydrology and Earth System Sciences*, *15*(7), 2349–2365. Retrieved from https://hess.copernicus.org/articles/15/2349/2011/ doi: 10.5194/hess-15-2349-2011

Gordon, N., Salmond, D., & Smith, A. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proc. F Radar Signal Process. UK*, *140*(2), 107. Retrieved from http://dx.doi.org/10.1049/ip-f-2.1993.0015 doi: 10.1049/ip-f-2.1993.0015

Grimaldi, S., Xu, J., Li, Y., Pauwels, V., & Walker, J. (2020). Flood mapping under vegetation using single sar acquisitions. *Remote Sensing of Environment*, *237*, 111582. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425719306029 doi: https://doi.org/10.1016/j.rse.2019.111582

Herbst, E., & Schorfheide, F. (2017, May). Tempered particle filtering [Working Paper]. (23448). Retrieved from http://www.nber.org/papers/w23448 doi: 10.3386/w23448

Hersbach, H., Bell, W., Berrisford, P., Horányi, A., J., M.-S., Nicolas, J., … Dee, D. (2019, 04). Global reanalysis: goodbye era-interim, hello era5. , 17-24. Retrieved from https://www.ecmwf.int/node/19027 doi: 10.21957/vf291hehd7

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. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017WR022205 doi: 10.1029/2017WR022205

Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, *43*(7). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006WR005756 doi: https://doi.org/10.1029/2006WR005756

Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., ... Savenije, H. H. G. (2010). Towards the sequential assimilation of sar-derived water stages into hydraulic models using the particle filter: proof of concept. *Hydrology and Earth System Sciences*, *14*(9), 1773–1785. Retrieved from https://hess.copernicus.org/articles/14/1773/2010/ doi: 10.5194/hess-14-1773-2010

Moradkhani, H., Dechant, C., & Sorooshian, S. (2012). Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-markov chain monte carlo method. Retrieved from https://escholarship.org/uc/item/76j5z2t7

Neal, J., Schumann, G., & Bates, P. (2012). A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research*, *48*(11). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012WR012514 doi: 10.1029/2012WR012514

Neal, R. M. (1996). Sampling from multimodal distributions using tempered transitions. *Statistics and Computing*, *6*, 353-366. doi: 10.1175/2009MWR2835.1

Potthast, R., Walter, A., & Rhodin, A. (2019). A localized adaptive particle filter within an operational nwp framework. *Monthly Weather Review*, *147*(1), 345 - 362. Retrieved from http://proxy.bnl.lu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=iih&AN=134406412&site=ehost-live&scope=site

Reich, S. (2013). A nonparametric ensemble transform method for bayesian inference. *SIAM Journal on Scientific Computing*, *35*(4), A2013-A2024. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84886860357&doi=10.1137%2f130907367&partnerID=40&md5=a6ab826cdcb46994c3aef93d949a912c (cited By 64) doi: 10.1137/130907367

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. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022169416306862 doi: https://doi.org/10.1016/j.jhydrol.2016.10.041

Snyder, C., Bengtsson, T., Bickel, P., & Anderson, J. (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, *136*(12), 4629 - 4640. Retrieved from http://proxy.bnl.lu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=iih&AN=36092236&site=ehost-live&scope=site

Tellman, B., Sullivan, J., & Kuhn, K. A. J. D. C. S. B. G. R. E. T. A. S. D. A., C. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature*, *596*, 80–86. Retrieved from https://doi-org.proxy.bnl.lu/10.1038/s41586-021-03695-w doi: 10.1093/biomet/ast020

Van Leeuwen, P. J. (2009). Particle filtering in geophysical systems. *Mon. Wea. Rev.*, *137*, 4089-4114. doi: 10.1175/2009MWR2835.1

van Leeuwen, P. J., Künsch, H. R., Nerger, L., Potthast, R., & Reich, S. (2019). Particle filters for high-dimensional geoscience applications: A review. *Quarterly Journal of the Royal Meteorological Society*, *145*(723), 2335-2365. Retrieved from https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/

qj.3551 doi: 10.1002/qj.3551

Wongchuig-Correa, S., Cauduro Dias de Paiva, R., Biancamaria, S., & Collischonn, W. (2020). Assimilation of future swot-based river elevations, surface extent observations and discharge estimations into uncertain global hydrological models. *Journal of Hydrology*, *590*, 125473. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022169420309331 doi: https://doi.org/10.1016/j.jhydrol.2020.125473

Zhao, J., Pelich, R., Hostache, R., Matgen, P., Cao, S., Wagner, W., & Chini, M. (2021). Deriving exclusion maps from c-band sar time-series in support of floodwater mapping. *Remote Sensing of Environment*, *265*, 112668. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425721003886 doi: https://doi.org/10.1016/j.rse.2021.112668