AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2020WR027859

A Mutual Information-Based Likelihood Function for Particle Filter Flood Extent Assimilation

Antara Dasgupta[1,2,3], Renaud Hostache[4], RAAJ Ramsankaran[2], Guy J.-P Schumann[5], Stefania Grimaldi[3], Valentijn R. N. Pauwels[3], and Jeffrey P. Walker[3]

[1]IITB-Monash Research Academy, Mumbai, India, [2]Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai, India, [3]Department of Civil Engineering, Monash University, Clayton, VIC, Australia, [4]Department of Environmental Research and Innovation, Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg, [5]School of Geographical Sciences, University of Bristol, Bristol, UK

Key Points:

- A novel mutual information-based metric is proposed as the likelihood function for particle filter-based flood extent assimilation
- Distributed impacts of the assimilation on simulated flood depth and flow velocities are illustrated for different lead times
- Improvements in simulated water levels of ~80% over the open loop are shown, persistent for up to one week after the assimilation

Correspondence to:
A. Dasgupta,
antara.dasgupta@monash.edu

Citation:
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*, e2020WR027859. https://doi.org/10.1029/2020WR027859

Received 25 MAY 2020
Accepted 30 DEC 2020

Abstract Accurate flood inundation forecasts have the potential to minimize socioeconomic losses, but uncertainties in inflows propagated from the precipitation forecasts result in large prediction errors. Recent studies suggest that by assimilating independent flood observations, inherent uncertainty in hydraulic flood inundation modeling can be mitigated. Satellite observations from Synthetic Aperture Radar (SAR) sensors, with demonstrated flood monitoring capability, can thus be used to reduce flood forecast uncertainties through assimilation. However, researchers have struggled to develop an appropriate cost function to determine the innovation to be applied at each assimilation time step. Thus, a novel likelihood function based on mutual information (MI) is proposed here, for use with a particle filter-based (PF) flood extent assimilation framework. Using identical twin experiments, synthetic SAR-based probabilistic flood extents were assimilated into the hydraulic model LISFLOOD-FP using the proposed PF-MI algorithm. The 2011 flood event in the Clarence Catchment, Australia was used for this study. The impact of assimilating flood extents was evaluated in terms of subsequent flood extent evolution, floodplain water depths, flow velocities and channel water levels (WLs). Water depth and flow velocity simulations improved by ~60% over the open loop on an average and persisted for up to 7 days, following the sequential assimilation of two post-peak flood extent observations. Flood extents and channel WLs also showed mean improvements of ~10% and ~80% in accuracy, respectively, indicating that the proposed MI likelihood function can improve flood extent assimilation.

Plain Language Summary Accurate forecasts of flood inundation can minimize socioeconomic losses from the frequent and often disastrous flood events occurring world-wide. However, numerical models used to generate flood predictions are strongly dependent on the quality of inputs such as inflows and topography, which typically fail to meet the required accuracy standards. Recent studies suggest that integrating independent flood observations into these numerical models can mitigate some of the errors and increase the reliability of the resulting flood forecasts. Remotely sensed radar data, which has all-weather/all-day imaging capabilities, can thus accurately observe flooded areas. These flood extent observations can then be used to improve flood forecasts through model-data integration, but studies have struggled to develop an effective approach to combine these with numerical models, as the area under water only varies slightly with time. Consequently, this study proposed a novel model-data integration method, sensitive to slight variations in the flooded area, and verified its performance through synthetic experiments. At each time step shared information between the model predicted and the observed flooded area is quantified and used to combine their information content. This led to persistent improvements in predictions of flood extent, depth, and velocity, demonstrating the potential of the proposed integration method.

1. Introduction

Accurate predictions of flood inundation are vital to the disaster risk management cycle, so as to trigger appropriate responses from humanitarian organizations, first responders, and end users alike (Alfieri et al., 2018; Trigg et al., 2016). Flooding alone accounted for almost half of all the hydrometeorological disasters recorded over the last 2 decades, affecting approximately 2.3 billion people or one-third of the global

population (CRED and UNISDR, 2015). Future flood risk is expected to increase as a function of population growth and unbridled urbanization of flood plains and coastal zones, leading to increased flood vulnerability (CRED and UNISDR 2016; Dottori et al., 2020; Smith et al., 2019; Uhe et al., 2019). Therefore, operational flood forecasts need to provide precise information on runoff generation, runoff concentration, streamflow propagation, and floodplain inundation processes (Emerton et al., 2016; Grimaldi et al., 2016; Kauffeldt et al., 2016; Lavers et al., 2019). Consequently, typical flood forecasting systems comprise of hydrological models capable of describing the run-off generation from a given rainfall event, coupled with hydraulic models for predicting flooded zones, water depths and flow velocities within the channel network and flood plains (Grimaldi et al., 2018, 2019). However, uncertainty in inputs, parameters, process representation (e.g., model equations), numerical precision, and spatiotemporal discretization, strongly impacts the accuracy and reliability of flood forecasts, severely hampering effective decision-making during crises (Haile & Rientjes, 2007; Hostache et al., 2018; Ward et al., 2015).

In order to mitigate forecast uncertainties, the model cascade has historically been calibrated using in situ hydrometric measurements. However, an ongoing global decline in the number of gauging stations (Revilla-Romero et al., 2016), has urged developments in complementary directions such as remote sensing (Schumann & Domeneghetti, 2016) and crowdsourcing (Buytaert et al., 2016). Moreover, point measurements (0-Dimension or 0D) of streamflow and water level (WL) are not commensurate with the now widely used 2D-models, and therefore cannot effectively constrain model predictions in the floodplain (Hostache et al., 2009; Pappenberger et al., 2007). Earth Observation (EO) data has thus become a popular and attractive alternative for model implementation, calibration, and validation, and to improve forecast skill through assimilation (Grimaldi et al., 2016). Synthetic Aperture Radar (SAR) sensors are particularly preferred for flood detection due to their all-weather all-day imaging capabilities (Dasgupta et al., 2018a). For example, García-Pintado et al. (2015) assimilated WLs derived from the intersection of SAR-based flood extent maps with high accuracy LiDAR elevation data; Wood et al. (2016) used SAR-derived flood extent maps for hydraulic model calibration; Hostache et al. (2018) assimilated flood extent maps; and Cooper et al. (2019) directly used backscatter values to reduce flood forecast uncertainty.

As evident from the comprehensive review of hydraulic data assimilation (DA) studies in Table 7 of Grimaldi et al. (2016), most studies have focused on assimilating synthetic (Garambois et al., 2019; Giustarini et al., 2011; Matgen et al., 2010; Tuozzolo et al., 2019), in situ (Van Wesemael et al., 2019; Ziliani et al., 2019), or remote sensing-derived water levels (RSD-WLs) (Giustarini et al., 2012; Lai & Monnier, 2009). Unlike water depth, which is a state variable of hydraulic models, flood extents are derived prognostic variables (Lai et al., 2014). This makes the assimilation of RSD-WLs significantly more straightforward than the direct assimilation of flood extents (Hostache et al., 2018). However, the accuracy of the retrieved WLs from remotely sensed flood extents is highly dependent on the digital elevation model (DEM) accuracy and resolution. Since the DEM is used as an input to the hydraulic model as well as for the retrieval of WLs, the assimilation of satellite-derived WLs may also introduce bias, because the observation and the model are no longer independent from each other (Schumann et al., 2008). Even though the accuracy of Global DEMs is improving (Archer et al., 2018; Hawker et al., 2019; Yamazaki et al., 2017), the vertical accuracy is often insufficient for application at local scales (Schumann & Bates, 2018). Moreover, Hostache et al. (2018) argue that spatial flood information is lost during the interpretation of RSD-WLs, as the integration of flood extents and DEMs can only reliably deliver water heights at a few shoreline points (Mason et al., 2016, 2012; Shastry & Durand, 2019). Since the derivation of RSD-WLs is neither straightforward nor automatic, incorporating such assimilation frameworks into operational forecasting systems remains a challenge (Hostache et al., 2018).

Given the limitations of WL assimilation, recent studies have focused on developing techniques to directly assimilate flood extents (Lai et al., 2014; Hostache et al., 2018; Revilla-Romero et al., 2015, 2016; Shastry & Durand, 2019). While both Lai et al. (2014) and Revilla-Romero et al. (2016) interpreted inundation extents as a function of the internal model states, to develop the cost function for assimilation (i.e., water depth and discharge), Hostache et al. (2018) directly utilized the flood extents. The study by Lai et al. (2014) was a proof of concept, using the 4D-Variational technique and a single flood extent map derived from Moderate resolution Imaging Spectroradiometer (MODIS) images at 250 m resolution to optimize a lumped friction parameter. In contrast, Revilla-Romero et al. (2016) used the Ensemble Kalman Filter (EnKF) to assimilate surface water extent observations at $0.1 \times 0.1°$ from the Global Flood Detection System

(http://www.gdacs.org/flooddetection), to improve near real-time (NRT) global flood forecasts. In spite of its popularity in hydrological data assimilation literature (e.g., Patil & Ramsankaran, 2017, 2018), one of the key limitations of the EnKF is the assumption of Gaussianity for model and observation errors (Moradkhani et al., 2005b, 2019). As this assumption does not hold for assimilation of SAR-derived flood observations (Matgen et al., 2010), some studies have recommended the use of a particle filter (PF) framework (Giustarini et al., 2011; Hostache et al., 2018; Matgen et al., 2011). For example, Hostache et al. (2018) used PFs to assimilate flood extents derived from ENVISAT ASAR images (spatial resolution of 150 m – resampled to 75 m) into the hydraulic model LISFLOOD-FP and achieved improvements of up to 50% in simulated WLs at the assimilation time step.

Variations in modeled or observed flood extents are typically only limited to the boundary of the flooded area. Thus, developing a flood extent-based likelihood to isolate the best performing ensemble members which can efficiently drive the assimilation is a significant scientific challenge (Lai et al., 2014). For instance, Cooper et al. (2019) proposed the conversion of modeled binary flood extents into synthetic SAR observations, and inter compared the wet and dry backscatter observations at the flood boundary with actual SAR images. While directly utilizing backscatter values reduces the processing time, it discards any information contributed by the addition of texture (Dasgupta et al., 2018b) or coherence (Chini et al., 2019), which may reduce uncertainties. Therefore, Hostache et al. (2018) proposed the assimilation of probabilistic flood maps from SAR images in their pioneering study, whereby forecasts were updated through direct flood extent comparisons. Accordingly, local pixel-wise likelihoods were computed through a binomial comparison of the model and observation, while the joint probability density used as the global weight was computed as a product, by assuming negligible spatial correlation among observation errors (Hastie et al., 2009). This assumption is slightly unrealistic as speckle dominates SAR image errors, which is caused by backscatter interference among targets in the same or neighboring elementary resolution cells potentially resulting in spatial autocorrelation (Argenti & Alparone, 2002). During flood extent assimilation, the assumption of spatial independence of observation errors may underestimate the standard deviation and correlation length scale (Waller et al., 2016).

This study proposes a novel likelihood function based on mutual information (MI) for PF-based flood extent assimilation which accounts for spatial uncertainty correlations (Wellmann, 2013). Information theoretic cost metrics have been previously used as likelihood functions for PFs in sensor network control, target tracking, and autonomous robotics (Charrow et al., 2014; Hoffmann & Tomlin, 2010; Zhang et al., 2018), however they have not been applied in hydrology. Some studies have used MI as a diagnostic metric for performance evaluation and observation impact (e.g., Fowler & Van Leeuwen, 2012, 2013; A. M. Fowler et al., 2018; Nearing et al., 2018), but it has never been used within a sequential data assimilation framework as a cost or likelihood function to rank model ensembles. Conversely, MI has been used rather ubiquitously in remote sensing; specifically in intensity based image matching and registration studies (Chen et al., 2003; Hirschmüller, 2008; Horkaew & Puttinaovarat, 2017; Liu et al., 2018; Suri & Reinartz, 2010). An MI-based likelihood function is therefore introduced to flood data assimilation in this study. Importantly, MI provides a measure of the reduction in uncertainty about one random variable (observation) given the complete knowledge of another (model), yielding a single global weight value for each particle and eliminating the need for assuming spatial independence of observation errors (Horkaew & Puttinaovarat, 2017; Suri & Reinartz, 2010; Wellmann, 2013). As a SAR-image becomes available, it is converted to a flood probability map and the proposed MI-based likelihood function used to sequentially update the flood forecast (Giustarini et al., 2011; Hostache et al., 2018; Matgen et al., 2010, 2011; Wood, 2016). The proposed approach was evaluated in this paper using a controlled identical twin experiment, where the open loop only differed from the synthetic truth in terms of an erroneous upstream discharge. Performance of the assimilated forecast was evaluated against the "truth" model in terms of inundation extent, floodplain water depth, floodplain flow velocities, and channel flow and depth.

2. Methods

This section outlines the methods used in this study, with Figure 1 illustrating the overall workflow followed. The next sections describe the forward truth model, the synthetic satellite observation simulation, and the ensemble generation. This is followed by a detailed explanation of the proposed data assimilation framework, and finally the performance metrics used to evaluate the assimilation performance.

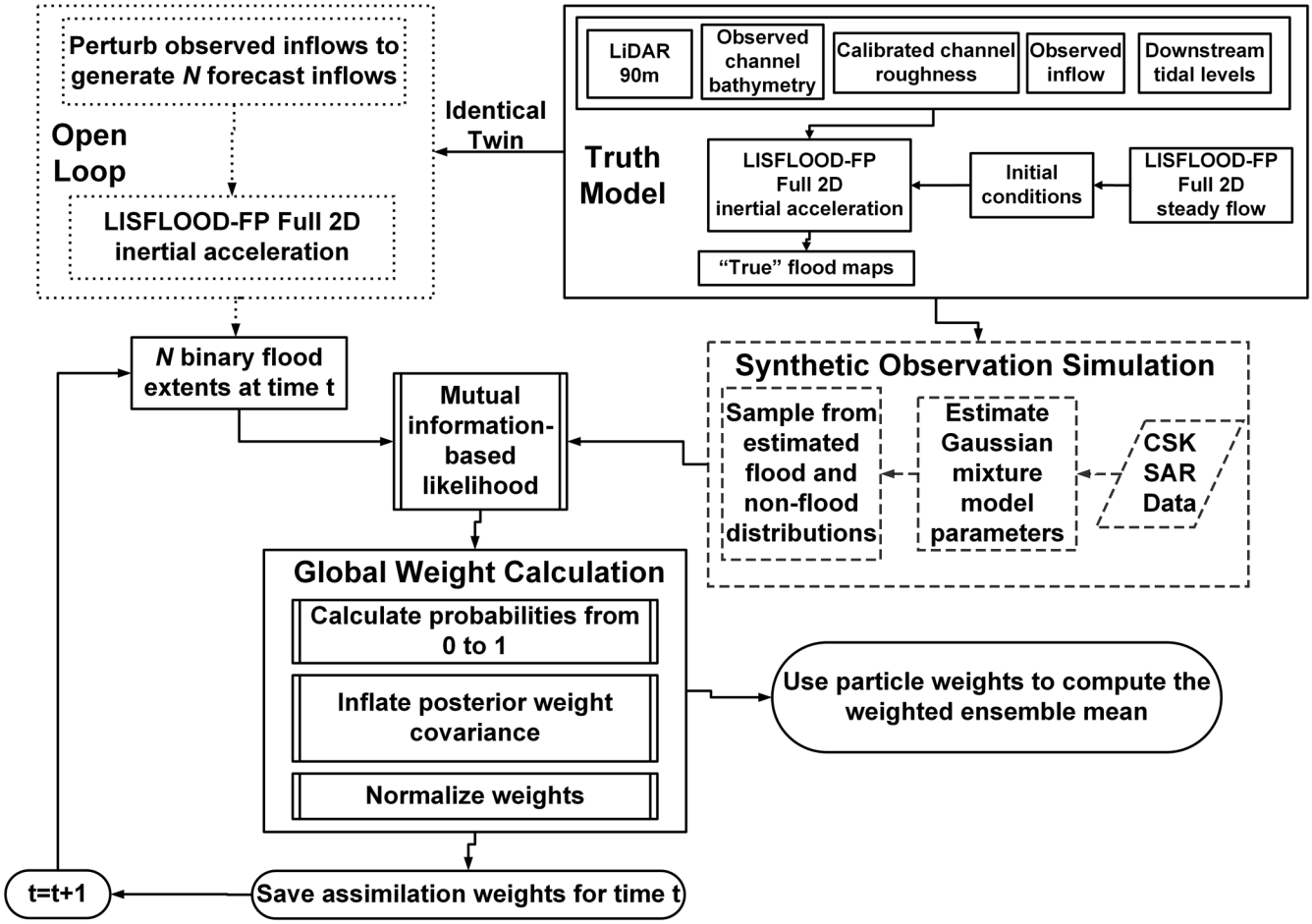

Figure 1. Schematic of the synthetic assimilation experiment using an identical twin setup, where synthetic data were generated from flood extents produced by a truth run and subsequently assimilated within the same model.

2.1. The Forward Truth Hydraulic Model

The 2D grid-based hydraulic model LISFLOOD-FP (Bates et al., 2010) was used in this study. This model was chosen for the easy integration of remotely sensed data with its grid-based spatial domain discretization (P. D. Bates & De Roo 2000). Moreover, its inertial formulation has demonstrated better computational efficiency than other diffusive models, while maintaining similar levels of forecast accuracy (Fewtrell et al., 2011; Horritt & Bates, 2001). Each floodplain spatial unit is modeled as a storage cell, where flows are computed using the inertial approximation of the Saint Venant equations; only convective acceleration is ignored (Andreadis & Schumann, 2014). A finite difference numerical scheme, explicit in time and first order in space with a semi implicit treatment of the friction term, is used for increased stability (Bates et al., 2010). Flow between the cells is computed as a function of local water acceleration, skin friction, and water surface slopes (de Almeida et al., 2012). For a full description of the model LISFLOOD-FP and version history, readers are referred to Bates et al. (2013).

The model was set up at a resolution of 90 m to maintain reasonable simulation times, with one model run taking from ~8 to 12 h on a 2.20 GHz Intel(R) Xeon(R) CPU E5-2650 v4 with 64 GB of RAM. The experimental configuration included 543,222 computational cells with hourly flow depth, velocity, and hazard outputs. This resolution is also compatible with vegetation and hydrologically corrected global DEMs like MERIT (Yamazaki et al., 2017), and is often considered as the preferred grid size for downscaling coarser resolution (1 km) global flood forecasts (G. J. P. Schumann et al., 2014, 2016). Maximum flooded area averaged at around 63,000 pixels, covering about 12% of the domain.

Hydraulic models require the specification of an inflow discharge hydrograph at the upstream boundary, initial conditions, detailed topographic and bathymetric data, surface roughness assessments, and outflow data at the downstream boundary. The upstream boundary condition was fixed at Lilydale where an inflow discharge hydrograph was specified, while tidal levels were used as the downstream boundary condition at Yamba (see Figure 3 for locations). This study used the inertial acceleration solver implemented in full-2D, with surveyed bathymetry, LiDAR topography, observed tidal levels, and calibrated channel roughness. The model was calibrated at 90 m grid resolution using crowdsourced high-water marks for the 2013 flood event in the Clarence Catchment. This implementation was chosen to develop and test the proposed flood extent assimilation framework through synthetic experiments, to evaluate its efficacy independent of additional uncertainties.

Initial conditions were computed by running a steady state simulation with the median flow value at the inflow point gauge used to represent the inflow hydrograph. The resulting simulated water depths were used as a start file for the model. Based on manual trial and error exercises, a warm-up period of five weeks was found to be sufficient to ensure numerical stability. Distributed floodplain roughness values were assigned according to land-use classes identified from aerial photographs, based on the recommendations by Arcement and Schneider (1989). As the floodplain roughness is not a sensitive parameter for LISFLOOD-FP during a flood, this roughness map was kept constant throughout all the simulations (Wood et al., 2016). This choice allowed for a focused assessment of the impacts of channel roughness and inflows on the ensemble forecasts.

Identical twin experiments common to data assimilation studies were setup for the 2011 flood event in the Clarence Catchment, with simulations including warm-up starting from December 1, 2010 and running up to January 23, 2011. The flood event which occurred between 8th to 16th January 2011, was mainly used to assess the performance of the flood extent assimilation procedures outlined here. A high intensity short duration (HISD) flood was simulated to evaluate assimilation performance during "flashy" events. These two synthetic flood events were used to evaluate the performance of the proposed flood extent algorithm in comparison to Hostache et al. (2018) in Section 4.5. Observed inflows were used to generate the "truth" model simulation, along with the calibrated channel friction parameter. Observed tidal levels were applied at the downstream boundary, and floodplain elevations were specified using the LiDAR DEM and the observed channel bathymetry. Considering the observed inflows as the truth, forecast inflow uncertainties were synthetically generated (see Section 2.3), representing outputs generated by hydrological models running in forecast mode (García-Pintado et al., 2013).

2.2. Synthetic Satellite Observation

Flood extents simulated by the "truth" model were used to generate the synthetic observations, at the time steps corresponding to actual SAR acquisitions, using the approach proposed by Cooper et al. (2019). Backscatter distributions of flood and nonflood classes were assumed to follow the form of Gaussian Mixture Models. Parameters for each Gaussian curve in the observed SAR image histograms were estimated using nonlinear curve fitting. For each pixel in the modeled flood and nonflood classes, a backscatter value was sampled from the corresponding normal distribution, defined based on the observed mean and standard deviation to generate the synthetic SAR images. To emulate a realistic scenario, synthetic images at each assimilation time step were generated for the 2011 event by estimating the corresponding backscatter statistics from real SAR images. In a second synthetic event considered herein, the statistics from available images were averaged and synthetic images generated at various times, corresponding to the early rising limb, pre-peak, peak, post-peak, and late falling limb. Model grid cells with water depths >1 cm were considered flooded and the rest nonflooded following Hostache et al. (2018).

SAR-derived flood extents are inherently uncertain (see Dasgupta et al., 2018a) for an in depth discussion), and probabilistic maps are one way to characterize this uncertainty. Moreover, correct representation of observation uncertainty is vital to the assimilation process, and probabilistic flood maps allow taking advantage of complementary information layers in addition to backscatter. Therefore, the synthetic satellite-based flood observations were converted to probabilistic flood maps, using the algorithm developed by Giustarini et al. (2016). The algorithm converts backscatter values into flood probabilities, based on Bayesian conditional probabilities $p(F \mid \sigma^0)$ calculated according to

$$p\left(F \mid \sigma^0\right) = \frac{p\left(\sigma^0 \mid F\right)p\left(F\right)}{p\left(\sigma^0\right)}, \quad (1)$$

where $p\left(\sigma^0 \mid F\right)$ is the conditional probability for flood class F of observing a specific backscatter value σ^0 given that the pixel is flooded, and $p\left(\sigma^0\right)$ is the marginal probability distribution of backscatter values at a given pixel (the image histogram). The term $p(F)$ denotes the prior probability of observing a flooded pixel; assumed to be 0.5 in the absence of any ancillary information (Schlaffer et al., 2017). Since, this is a synthetic study with controlled errors, advanced algorithms such as the Neuro-Fuzzy algorithm proposed by Dasgupta et al. (2018b), which allow incorporating information from additional layers were not considered. While for real-world applications the observation accuracy would influence assimilation performance, for synthetic experiments such as the one described here, the assimilation framework is independent of the method chosen for the assignment of flood probabilities.

2.3. Ensemble Generation

The success of any assimilation experiment is driven by the perturbation technique chosen and the skill of the ensemble spread (De Lannoy et al., 2006). In this study, the upstream boundary condition, supplied in the form of hourly observed discharge values, was assumed to be the only source of uncertainty. Uncertainties generated through incorrectly specified inflows were independently considered, to better understand the impact of the proposed likelihood function on the flood extent assimilation. In order to emulate discharge forecasts produced by hydrological models, temporally correlated heteroscedastic errors were synthetically generated with a positive mean bias (Gobeyn et al., 2017). As the uncertainty of the rainfall forecasts from weather prediction models are propagated down the modeling chain, the nature of hydrological streamflow forecast errors is often rather complex with significant temporal correlation and heteroscedasticity. The forecast discharge errors were therefore simulated by considering a temporally correlated time-varying error variance. Temporal correlation of errors was simulated following the approach of Evensen (2003), while the heteroscedasticity was incorporated using the methods outlined by García-Pintado et al. (2013). The 2011 event simulation additionally included a positive multiplicative mean bias of 20% following earlier studies (e.g., García-Pintado et al., 2013; Matgen et al., 2010). This allowed emulating the uncertainty of a typical hydrological streamflow forecast and lends insights into the mean error evolution on the forecast. Since the assimilation problem was approached with no knowledge of the true mean error, the potential of data assimilation to partially solve for it could be tested.

In order to evaluate the generalizability of the results presented herein, the assimilation approach was also implemented to an HISD event, where forecast improvements are harder to gain and retain over time. For this, the 2011 hydrograph shape was scaled to 200 h, ensuring a ~30 h time-to-peak and the same peak and base flows. This was achieved by down-sampling the 2011 hydrograph to 200 h, with the median flow value observed at the upstream boundary gauge added as baseflow before the flood hydrograph, to emulate a short warm-up period of approximately ~112 h. The down-sampling was controlled such that the average ~30 h time to peak (Farr & Huxley, 2013), typically observed for the Clarence Catchment was retained. This implied that the HISD event was statistically identical (same shape and max/min flows) to the 2011 event but flashier since the longer duration flood flow dynamics were compressed within a short time window through the scaling. The 2011 event has an Annual Recurrence Interval (ARI) of ~27 years, and thus, the 25 year tidal levels were used as the downstream boundary. While the inflow ensemble generation method used in this case is identical to the 2011 event, no positive multiplicative bias was imposed on the flows, to evaluate assimilation impacts independent of bias. This implied that for the HISD event the open loop mean was already very close to the truth.

The time-dependent change in variance for both events was imposed through the coefficient of variation (cv) of discharge, estimated from the historical rating curves at Grafton, Prince Street gauge (Farr & Huxley, 2013). This *cv* value was transferred to Lilydale because published historical rating curves only existed for this particular gauge on the Clarence River, despite the fact that the geometric properties of the

cross-section could be starkly different from Grafton. As the average value of cv was observed to be close to ~0.2, a slightly higher value ($cv = 0.25$) was used to impose the heteroscedasticity as recommended by García-Pintado et al. (2013). Since operational global flood forecasting systems currently provide discharge forecasts at a daily time step (Dottori et al., 2016), a decay time of 3 days was assumed for the temporal correlation, following the work of García-Pintado et al. (2013). Figures 2a and 2b show the open loop inflow ensembles for both events used in this study, along with the observed (with or without scaling) inflows used to generate them and the open loop mean. For a detailed description of the inflow ensemble generation techniques described here and for all the equations, readers are referred to Evensen (2003) and García-Pintado et al. (2013). The inflow error generation equations have been omitted in this study for brevity as the approach is identical to García-Pintado et al. (2013).

An ensemble of 128 simulations was used in this study. While this ensemble size is small with respect to the domain size and the size of the state vector, given the results shown in Hostache et al. (2018) this was considered an appropriate trade-off for computational speed. Even though studies show that increasing the ensemble size may result in improved assimilation performance (e.g., Ziliani et al., 2019), using large ensemble sizes remains challenging even with current generation computationally efficient hydraulic models and computing power. In theory, the ideal ensemble size for particle filtering should exceed the length of the state vector by several orders of magnitude, which for this study (~63,500 wet cells) would mean an ensemble size $>10^6$ (Banister & Nichols, 2010). Not only is this computationally infeasible for research applications, even doubling or quadrupling the ensemble size (e.g., from 128 to 256 or even 528), results in only marginal performance improvements for PFs (Plaza Guingla et al., 2013). Since, ensemble skill is a function of the ensemble size and the perturbation technique chosen, the ensemble generation parameters were tuned to ensure sensible spread as suggested by Moradkhani et al. (2005).

Forecast skill was evaluated based on the upstream discharge, using the Normalized RMSE Ratio (NRR) proposed by Moradkhani et al. (2005), where an NRR value of ~1 is considered ideal. Values of $NRR >> 1$ indicate inadequate ensemble spread while $NRR << 1$ represents too much spread (Matgen et al., 2010). NRR was estimated according to

$$NRR = \frac{Ra}{E\left[Ra\right]}, \tag{2}$$

where Ra is the ratio of the time-averaged RMSE of the ensemble mean, and the mean RMSE of the ensemble members. If the observation is statistically indistinguishable from the ensemble, the expectation of Ra is given by

$$E\left[Ra\right] = \sqrt{\frac{(n+1)}{2n}}, \tag{3}$$

where n is the ensemble size. For an in-depth description of the forecast evaluation methods used in this study, readers are referred to Moradkhani et al. (2005) and Matgen et al. (2010). The value of NRR for the forecast inflow ensemble was ~0.99, being almost equal to the ideal value of unity, indicating sufficient ensemble spread and skill (Moradkhani et al., 2005). The resulting inflow ensembles were propagated through LISFLOOD-FP to generate water depth maps, subsequently converted to binary flood extents using the 1 cm threshold to avoid noncontiguous negligible water depths. This ensemble of simulated binary flood extents was then compared against the synthetically generated SAR-derived flood extents, using the proposed likelihood function at the assimilation time steps.

2.4. Data Assimilation Framework

PFs are based on sequential Monte Carlo simulations as discussed in the previous section, enabling a nonparametric representation of the continuous posterior probability distribution function (pdf). By randomly selecting a large number of discrete and independent samples from a distribution, the posterior pdf was iteratively estimated by sequentially assimilating observations (Arulampalam et al., 2002). Each

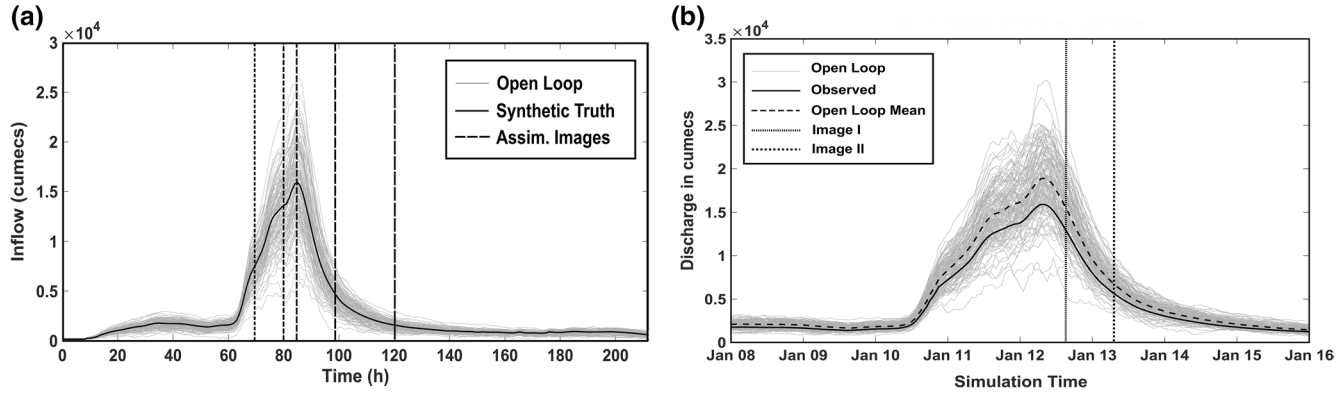

Figure 2. The inflow ensemble with 128 members generated for (a) the high intensity short duration event and (b) the 2011 real-world flood event in the Clarence Catchment.

ensemble member, or in other terms a particular model realization, represented a distinct particle with unique associated inflows. At each assimilation time step, the Bayesian conditional probability of the particular model trajectory given the observation was individually evaluated and used as the particle weight (Kantas et al., 2015). In the sequential importance sampling (SIS) algorithm, the weighted ensemble mean (or the expectation) was subsequently calculated, being representative of the total information content of the model and the observations (Plaza et al., 2012). The SIS algorithm generated particles based on a predetermined importance density, as sampling from the unknown nonGaussian posterior distribution was not possible (Moradkhani et al., 2005a). Typically, a uniform distribution is assumed when no other knowledge is available, and the initial importance weight of each particle is assigned as the reciprocal of the number of particles for example 1/128 in this case (Li et al., 2013; Plaza Guingla et al., 2013).

The SIS algorithm provides the unique advantage of avoiding hydrostatic initialization shock, which often occurs in state updating data assimilation filters applied to hydrodynamic models (Hostache et al., 2018). As data assimilation is essentially a statistical recombination procedure, the momentum across the domain is not conserved when states are updated and the model is restarted from the updated states (García-Pintado et al., 2015). This is because flow fluxes drop to zero when the model is stopped mid-simulation where conserving the momentum of large volumes of water becomes impossible. Consequently, the SIS algorithm perfectly suits the problem of hydraulic data assimilation as it only alters the ensemble statistics without disturbing the delicate model dynamics (Hostache et al., 2018).

After identifying the particular PF implementation best suited to the problem, an appropriate likelihood function needs to be defined to compute the conditional probability of the observations given the model state. When dealing with a large number of measurements, which might have correlated errors, the proposed likelihood distribution should ideally be the joint pdf of all these measurements (van Leeuwen, 2017). This implies that in SAR-based flood extent assimilation, where the pixel-wise flood inundation values constitute separate measurements and the images are often characterized by a multimodal histogram, a nonparametric approach to estimating the likelihood function must be chosen. MI was therefore chosen as the likelihood function for the SIS-based PF implementation used here (Shannon & Weaver, 1964).

MI is a measure of the amount of information one random variable contains about another (Cover & Thomas, 2005). MI has previously been used in hydrology to establish the causality of changes in flood occurrences with respect to mean precipitation patterns (Perdigão & Blöschl, 2014), and to estimate epistemic and aleatory uncertainties in hydrologic modeling (Gong et al., 2013). The following adapts the MI based likelihood function for remote sensing based hydraulic data assimilation, following the work of Hoffmann and Tomlin (2010). Note that the subscripts for time have not been included to maintain simplicity of the notations in Equations 4–14 which derive the likelihood function for PFs.

The data assimilation problem can simply be formulated as finding the best possible estimates of the system states from uncertain observations, given an uncertain model of the system dynamics (Walker &

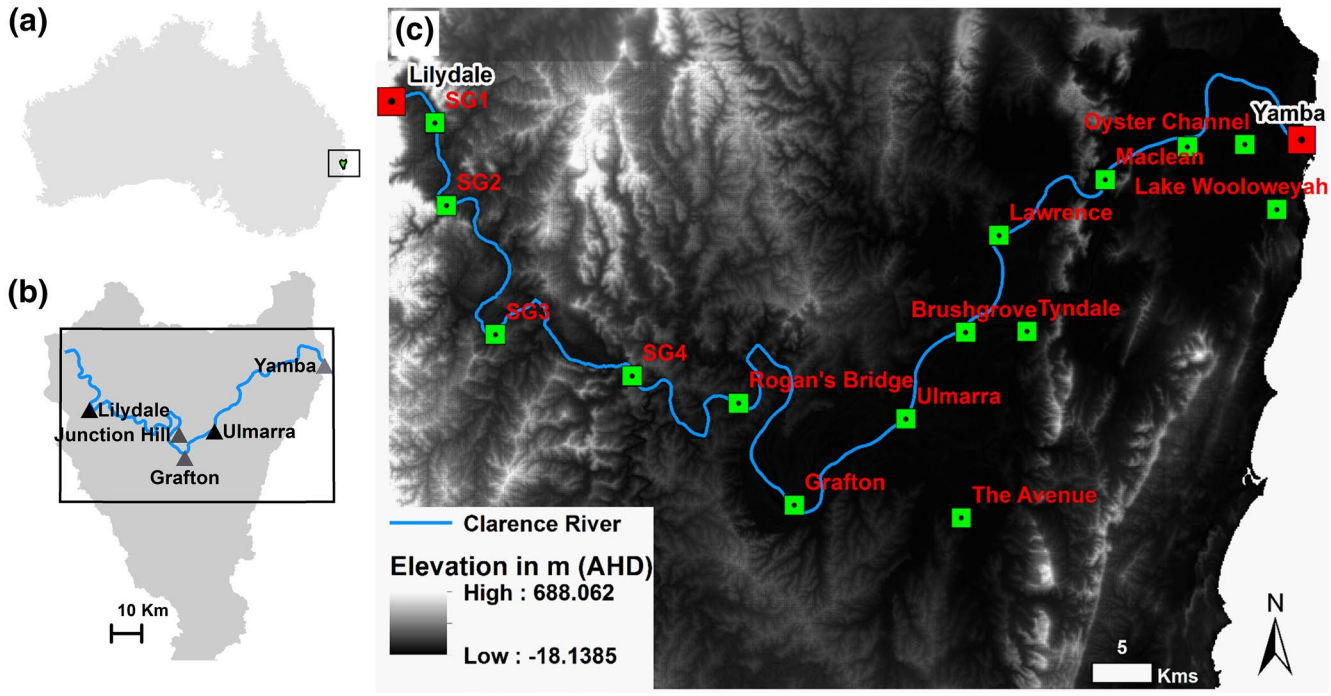

Figure 3. Geographical location of the Clarence Catchment, Australia shown in (a), with important towns marked relative to the Clarence River catchment in (b). The extent of the model domain from Lilydale to Yamba is shown as (c), with inflow and outflow boundary conditions marked in red and available gauges marked in green.

Houser, 2005). In the information theory context, this can be interpreted as reducing the uncertainty about the system states $x \in \Re^{N_x}$, using the information contained in the observations $y \in \Re^{N_y}$. The prior pdf of the state is given by $p(x)$ while observing the true state of the system yields measurements according to

$$y = h\left(x_{true}\right) + \epsilon, \tag{4}$$

where $h(.)$ maps the model state x to the observation space and ϵ is the measurement error. Representing the prior state pdf as N discrete particles by

$$p\left(x\right) = \sum_{i=1}^{N} \frac{1}{N} \delta_{x_i}, \tag{5}$$

and using them within the Bayes' Theorem,

$$p\left(x \mid y\right) = \frac{p\left(y \mid x\right) p\left(x\right)}{\int p\left(y \mid x\right) p\left(x\right) dx}, \tag{6}$$

the true state posterior can be estimated through

$$p\left(x \mid y\right) = \sum_{i=1}^{N} w_i \delta_{x_i}, \tag{7}$$

where likelihood weights are given by

$$L_i = \frac{p\left(y \mid x_i\right)}{\sum_{j=1}^{N} p\left(y \mid x_j\right)}, \tag{8}$$

with $p(y \mid x_i)$ being the pdf of the observations y given the model prediction x_i, also known as the likelihood (Arulampalam et al., 2002). Taking the log-likelihood of the posterior distribution given by Equation 6 and expanding the joint state-measurement pdf according to Bayes' rule as $p(y,x) = p(y \mid x)p(x)$ yields

$$H(x \mid y) = H(x) - I(y;x), \tag{9}$$

where

$$H(x) = -\int p(x)\log p(x)dx, \tag{10}$$

$$H(x \mid y) = -\int p(x,y)\log p(x \mid y)dxdy, \tag{11}$$

$$I(x;y) = \int p(x,y)\log\frac{p(x,y)}{p(x)p(y)}dxdy. \tag{12}$$

Here $H(x)$ is the entropy of the state distribution, $I(x;y)$ is the MI between the model state and the observation distributions, and $H(x \mid y)$ is the conditional entropy of the distribution or the expected entropy of the state given the measurements (Cover & Thomas, 2005). Entropy can be defined as a measure of uncertainty of a random variable (Hirschmüller, 2008), while MI measures the divergence (Kullback-Liebler) between the independent and joint distributions of x and y. This implies that MI quantifies the degree of interdependence between two distributions, that is MI is maximum if the modeled and observed flood extents are exactly aligned.

For a system exhibiting discrete time dynamics for the ith particle as

$$x_{t+1}^i = f_t^i\left(x_t^i, u_t^i\right), \tag{13}$$

where u_t is the set of control inputs, the expected posterior state pdf uncertainty can only be minimized when the observation information is maximized with respect to the control inputs (Hoffmann & Tomlin, 2010). This can be achieved through the MI likelihood function which measures this shared information (Zhang et al., 2018), defined as

$$L_i = I(y;x) = H(x) - H(x \mid y). \tag{14}$$

While all of the above equations are defined for continuous random variables, they can be extended to discrete variables by replacing the integrals with summations. In order to compare discrete model outputs and continuous probabilistic observations as in the present case, continuous variables can be quantized into a large but finite number of bins by assuming that the density is continuous within each bin. Here, 256 bins have been used after testing for the sensitivity of up to 1,024 bins, as increasing beyond 256 yielded no notable impact on the calculation of MI. MI is usually calculated using log with a base of 2 as introduced by the computational information theorists and is measured in bits (Souza et al., 2018; Woo et al., 2015). This was retained in the present study as the concept of bits to measure information also suits remote sensing images (Horkaew & Puttinaovarat, 2017; Liang et al., 2014; Ossadtchi et al., 2014).

The MI based likelihood function proposed here allowed for an efficient ranking of particles. However, an additional rescaling factor needed to be introduced to improve the selection of well-performing particles at each assimilation time step (Herbst & Schorfheide, 2019). When comparing flood extents using binary pattern matching measures, the coverage of the hits that is the flooded area common to the model and the observations, is typically larger than the misses or false alarms (Stephens et al., 2014). This means that it is increasingly difficult to discriminate between models using standard likelihood functions (e.g., the common Critical Success Index), since sensitivity to the differences in the predicted inundation extent is low. Moreover for PFs, the ideal proposal distribution from which prior weights are sampled, must be different from the posterior to allow effective estimation (Godsill & Clapp, 2001). However, if this difference is too large

then the importance weights will be close to zero for frequent outcomes and close to one for rare values (Johansen, 2015). Consequently, the state estimates obtained will be dominated by a small subset of the Markov chain (Woodhead, 2007). Diversifying the prior particle population in the importance-sampling step of the particle filtering is a viable option to address this issue. Studies have attempted to use evolutionary algorithms to optimize the prior distribution during the importance sampling step (Abbaszadeh et al., 2018; Kwok et al., 2005; Yin & Zhu, 2015).

Evolutionary algorithms use the principles of genetics, such as crossover and mutation, to focus the prior populations into optimal state-parameter sub-spaces (Dumedah, 2012, 2015; Dumedah & Coulibaly, 2013, 2014). However, depending on the magnitude of uncertainty in the proposal distribution, the particles should evolve over several generations to achieve the ideal sub-space. While the strength of genetic algorithms is the ability to deliver accurate results with smaller ensemble sizes, the estimation of optimal initial states is necessary to leverage this. Since the population evolves over several generations, models need to be rerun with the updated states each time, implying that for an ensemble size of 50 evolving over 5 generations 250 model runs are required (Dumedah & Coulibaly, 2013). Abbaszadeh et al. (2019) suggest using the 4D-variational method for this initial state optimization, but for the case of hydraulic models developing the adjoint model is a major scientific challenge. Moreover, selecting among several well-performing models is incredibly difficult using binary flood extents. Due to the inherent positive bias in all pattern matching measures, distinguishing between different ensemble members is especially challenging for over-predicting models in large catchments (Stephens et al., 2014).

While the use of a weight rescaling parameter to inflate the posterior weight variance is common in PFs (e.g., Fearnhead & Künsch, 2017; Herbst & Schorfheide, 2019; Woodhead, 2007), it was necessary in this study due to the nature of the flood extent assimilation problem. Particle resampling can also be used to iteratively improve the quality of the particles (e.g., the Sequential Importance Resampling algorithm). However, this necessarily involves stopping and restarting the model, driven by the resampled particle trajectories. For hydraulic data assimilation applications, this could lead to spurious inundation patterns in momentum-conserving hydrodynamic models, since it cannot be preserved during the resampling step. Even though E. S. Cooper et al. (2018) showed that applying pre-assimilation velocities to the updated water depths is a potential solution, the method still requires further testing for real case studies. Therefore, a rescaling factor to inflate the posterior variance was chosen for this study, providing a simple and elegant solution to the problem of weight sensitivity. The ratio between the likelihood of the particle and the maximum likelihood value was enhanced through an exponent ϑ and used as the rescaling factor. This was designed to lower the weights of poorly performing models while increasing the well-performing model weights. At each time step the value of the exponent ϑ in Equation 15 is optimized, to ensure that the maximum normalized particle weight in the posterior is not below 0.25 so that the best performing particles can be selected. The exponent ϑ typically varies based on flooded area coverage, which determines the sensitivity of the likelihood function and consequently, the magnitude of inflation necessary to discriminate between several well-performing models. The weights for each particle were therefore calculated using

$$w_i = \frac{\left(L_i - \min_i L\right)}{\left(\max_i L - \min_i L\right)} \times \left(\frac{L_i}{\max_i L}\right)^{\vartheta}. \tag{15}$$

The first part of the multiplier normalizes the values of L_i, while the second part is the rescaling factor introduced to heighten discriminatory power of the weights. Finally, the particle weights are normalized to calculate the conditional probabilities of a particular particle given the observation, and to ensure that the posterior probability distribution function sums to unity according to

$$W_i = \frac{w_i}{\sum_{i=1}^{N} w_i}, \tag{16}$$

where W_i denotes the global weight assigned to a particle for a given observation. Typically for the assimilation of multiple images, the posterior estimated from the assimilation of the preceding image is used as the prior. Weights are accordingly multiplied forward in time to yield cumulative weights according to

$$W_i^{\text{cum}} = \prod_{j=1}^{M} W_i^j, \tag{17}$$

where M is the total number of available observations. Note that this step may also be avoided when the observations are too far apart in time, or when no additional information (i.e., observations) about the posterior exists (Wood, 2016). This study showcases the distributed impacts of both cases, with "Image I" and "Image II" referring to the use of uniform priors while "Image I and II" refers to the use of cumulative weights that is using the assimilated forecast from "Image I" as the prior for the assimilation of "Image II." When all the observations have been processed and their information content extracted in the form of the importance weights, the expectation of any state variable gives the assimilated forecast. This can be computed as the weighted mean of the state vector ensemble given by

$$E\left[X\right] = \sum_{i=1}^{N} W_i^{\text{cum}} \times x_i. \tag{18}$$

This formula was used to compute the expectation of streamflow, flow velocities, and water depth. The expectation of WL and flood extent was derived from the expectation of water depth (Hostache et al., 2018).

2.5. Performance Evaluation Framework

A comprehensive framework for the evaluation of assimilation performance is developed in this paper, to improve the understanding of spatiotemporal error evolution following flood extent assimilation. This is particularly critical, as flood extent is a diagnostic model state, and so updating model trajectories based on an instantaneous extent might even degrade the forecast. Assimilation results were therefore evaluated against the benchmark "truth" model in a distributed fashion for the simulated floodplain states, in addition to the point gauge locations for the channel. Assimilation performance was evaluated for the assimilation time step and lead times of 1, 3, and 7 days, to assess the spatiotemporal evolution of errors in the forecast. The impact on simulated flood extent was quantified through contingency maps and contingency matrix-based statistics. Simulated WLs were evaluated spatially through water depth difference (WDD) maps, and a spatial root mean squared error (RMSE) statistic used to quantify overall performance. The spatially distributed impact of the assimilation on flow velocities was then evaluated throughflow velocity difference maps, while the percentage improvement in RMSE was used to compare the relative impact of the two assimilated images. The relative performance of the assimilation with respect to the open loop was examined through the Brier Skill Score (BSS), using hydrographs from the truth model at river gauging sites. Absolute errors in the channel performance were evaluated using RMSE and hydrograph plots. Finally, the impact on inflow uncertainties was quantified through hydrograph comparisons.

The flood extents were evaluated using the Critical Success Index (CSI), calculated based on the contingency matrix illustrated in Table 1 (which also explains the meaning of the symbols) using

$$\text{CSI} = \frac{A}{A + B + C}. \tag{19}$$

According to the above equation, CSI ignores the correctly simulated nonflooded areas to eliminate the disproportionate impacts of the typically larger nonflooded areas in the model domain. Spatial comparisons were also included through contingency maps, which illustrate the locations of the classes in Table 1. For the modeled water depth assessment, a global measure of error was required to quantify the overall model performance in absolute terms. The RMSE statistic was chosen due to its ubiquity in hydraulic data assimilation literature. Therefore, it can facilitate benchmarking vis-à-vis other studies (Cooper et al., 2018; García-Pintado et al., 2015, 2013). In the context of the present research, the RMSE was calculated between the weighted ensemble mean $E\left[WD_k\right]$, and the water depths simulated by the truth model WD_k^{truth}, at select time steps according to

Table 1
Contingency Matrix Used for the Calculation of Binary Pattern Matching Based Flood Extent Performance Measures

Contingency matrix		Modeled: Flooded	Modeled: Nonflooded
Observed	Flooded	**(A) Correct Flooded (Hits)**	***(C) Under-prediction (Misses)***
	Non-flooded	***(B) Over-prediction (False Alarms)***	**(D) Correct Nonflooded (Correct Rejects)**

The bold specifies where the model and observation are in agreement, while the italic and bold italic refer to under- and over-prediction respectively.

$$WD_{\mathrm{RMSE}} = \sqrt{\frac{\sum_{i=1}^{P}\left(E\left[WD_k\right] - WD_k^{\mathrm{truth}}\right)^2}{P}}, \tag{20}$$

where P is the total number of model grid cells. In order to better understand the spatial evolution of forecast errors, the WDD between the forecast ensemble mean and the truth at each grid was calculated. As Cooper et al. (2018) have shown using synthetic topography, WDD maps can provide valuable insights into spatiotemporal error propagation through the model domain, which lumped statistics like the RMSE are unable to detect. For each pixel, the WDD map contains the deviation of the ensemble mean from the true water depth at the corresponding location, given by

$$\mathrm{WDD}_k = \mathrm{E}\left[\mathrm{WD}_k\right] - \mathrm{WD}_k^{\mathrm{truth}}. \tag{21}$$

Flow velocities were evaluated in the exact same way as water depths. Both absolute errors and their spatial distribution were considered, through RMSE and FVD maps. Metrics were computed for the cell flow velocities, by changing the variable from water depth to flow velocity in Equations 20 and 21. These assessments are only possible for synthetic experiments such as the one detailed in this study, as spatially distributed observations of water depth and flow velocity are unavailable in practice.

The model performance within the channel was evaluated using hydrometric gauges. River gauges recording water depth and discharge were considered at the locations where real gauges along the Clarence River already exist. Synthetic gauges (SG1-4) were also added at a few locations to enable a more comprehensive gauge-based evaluation. The relative skill of the assimilated ensemble with respect to the open loop was quantified through the BSS, calculated according to

$$\mathrm{BSS} = 1 - \frac{\overline{\left(\mathbf{Assim.} - \mathbf{Truth}\right)\left(\mathbf{Assim.} - \mathbf{Truth}\right)^{\mathbf{T}}}}{\overline{\left(\mathbf{OL} - \mathbf{Truth}\right)\left(\mathbf{OL} - \mathbf{Truth}\right)^{\mathbf{T}}}}, \tag{22}$$

where the variables in bold denote state vectors for the assimilated (**Assim.**), open loop (**OL**), and truth (**Truth**) forecasts while the overbar denotes an average. Values of $BSS \in (-\infty, 1]$, where $BSS = 0$ indicates no change in forecast skill with respect to the open loop while $BSS = 1$ is the ideal score. Note that the results were interpreted by comparing objective function values obtained by evaluating the open loop and the assimilated ensemble against the synthetic truth model.

3. Study Area and Data

3.1. Study Area

The Clarence Catchment is situated in New South Wales, Australia, as shown in Figure 3, and drains a total of 22,700 km^2. The Clarence River is 394 km long of which a stretch of ~160 km, between Lilydale and Yamba, was used for the hydraulic model set up. This catchment has most recently experienced flooding in February 2020 with the flood in 2013 being the highest on record, when the river reached WLs of 8.09 m Australian Height Datum (AHD) at the Prince Street Gauge, Grafton (Huxley & Beaman, 2014). Floods in this catchment move fast and take about 30 h to peak on average, generating a relatively flashy

catchment response (Rogencamp, 2004). Low-intensity, long duration rainfall events are the dominant cause of flooding in the area, closely followed by the back propagation of ocean storm tides. Indeed, the ocean tides control the inundation dynamics up until Ulmarra, and slight tidal oscillations are visible as far upstream as Rogan's Bridge during low-flow conditions. Tidal oscillations are evident from both the observed and simulated hydrographs, with tidal levels providing the downstream boundary condition at Yamba. The catchment is characterized by largely variable flow velocities, ranging from 2 to 5 m/s in the channel and along the levees, to almost zero in the backwaters (Sinclair Knight Merz and Roads and Traffic Authority of NSW 2011).

3.2. Datasets Used

Remotely sensed data in the form of SAR imagery and LiDAR elevations, as well as field-recorded hydrometric data were used in this study. Bathymetric data was collected during a field campaign in 2015 (Grimaldi et al., 2017), described extensively in Grimaldi et al. (2018), and supplemented with pre-existing bathymetric datasets Farr and Huxley (2013), forming a rich database. Topographic information was available in the form of a 1 m Light Detection And Ranging (LiDAR) bare earth DEM, acquired in 2010 with a vertical accuracy of ±30 cm at 95% confidence (1.96 × RMSE) and horizontal accuracy of ±80 cm at 95% confidence (1.73 × RMSE) (New South Wales Land and Property Management Authority, 2010). The vertical accuracy was assessed by comparing LiDAR point returns against survey check points as per the ICSM Guidelines for Digital Elevation Data (Intergovernmental Committee On Surveying & Mapping 2008). This data set is freely available under a Creative Commons Attribution 4.0 license, for commercial and noncommercial applications at https://elevation.fsdf.org.au/, provided by Geoscience Australia. The channel bathymetry was reconstructed by interpolating between field observed cross-sections and incorporated within the LiDAR DEM, for the part of the domain where it was available. The area upstream of Copmanhurst where LiDAR coverage was unavailable, was infilled with the SRTM-derived 30 m product enhanced by Geoscience Australia (DEM-H), described in the next paragraph. For a detailed description of this combined topographic data set, field data collection, and bathymetry generation, readers are referred to Grimaldi et al. (2018). The bathymetric data set is freely available at https://figshare.com/articles/Bathymetric_survey_of_the_Upper_Clarence/5648251 (Grimaldi et al., 2017).

The SRTM-derived DEM-H (Gallant et al., 2011), a 1 arc second (~30 m) gridded DEM corrected for hydrological applications produced for the Australian continent by Geoscience Australia, was used in this study to represent the "best available" global topography for the region (Jarihani et al., 2015). The DEM-H is considered superior to the globally available SRTM product for hydraulic modeling, as drainages were enforced using 1:250,000 scale watercourse lines and smoothed using ANUDEM software (Hutchinson, 2011; Zheng et al., 2016), to ensure seamless hydraulic connectivity (Dowling et al., 2011). As coarse resolution SRTM products are often unable to accurately capture flow lines, which lead to erroneous inundation forecasts, the DEM-H with region specific hydrological enhancements was considered an appropriate choice (Mukherjee et al., 2013). The accuracy of DEM-H is similar to the raw SRTM 1-arc second product with geolocation errors <12.6 m and 90% of tested heights within 9.8 m for Australia (Gallant et al., 2011). This data set is available for free, downloadable from http://elevation.fsdf.org.au/.

Two COSMO-SkyMed (CSK) X-band (9.6 GHz with a wavelength of 3.1 cm) HH-Polarized images, acquired by the CSK-3 satellite, were available for this study. The data were acquired in Stripmap HIMAGE mode at 3 m resolution on January 12, 2011 at 18:03 h and January 13, 2011 at 07:33 h Australian Eastern Daylight Time (AEDT). The images were acquired at the peak of the 2011 flood event at Grafton and just after it on the falling limb as depicted in Figure 2b. The CSK Level 1D Georeferenced Terrain Corrected (GTC) product delivered as an 8-bit image of digital numbers was used in this study. The calibration process for the GTC product followed by E-Geos, corrects for local incidence angle impacts using a DEM, by normalising the backscatter to a 40° reference incidence angle (Italian Space Agency 2009). The domain comprised of 74,056,858 pixels each having an area of 9 m^2, bringing the total tile coverage to approximately 666.5 km^2. These images were only used in this paper to parameterize the backscatter statistical distributions of water and nonwater pixels, which were then used to derive synthetic SAR images from "true" flood extents (i.e., the "truth" model result).

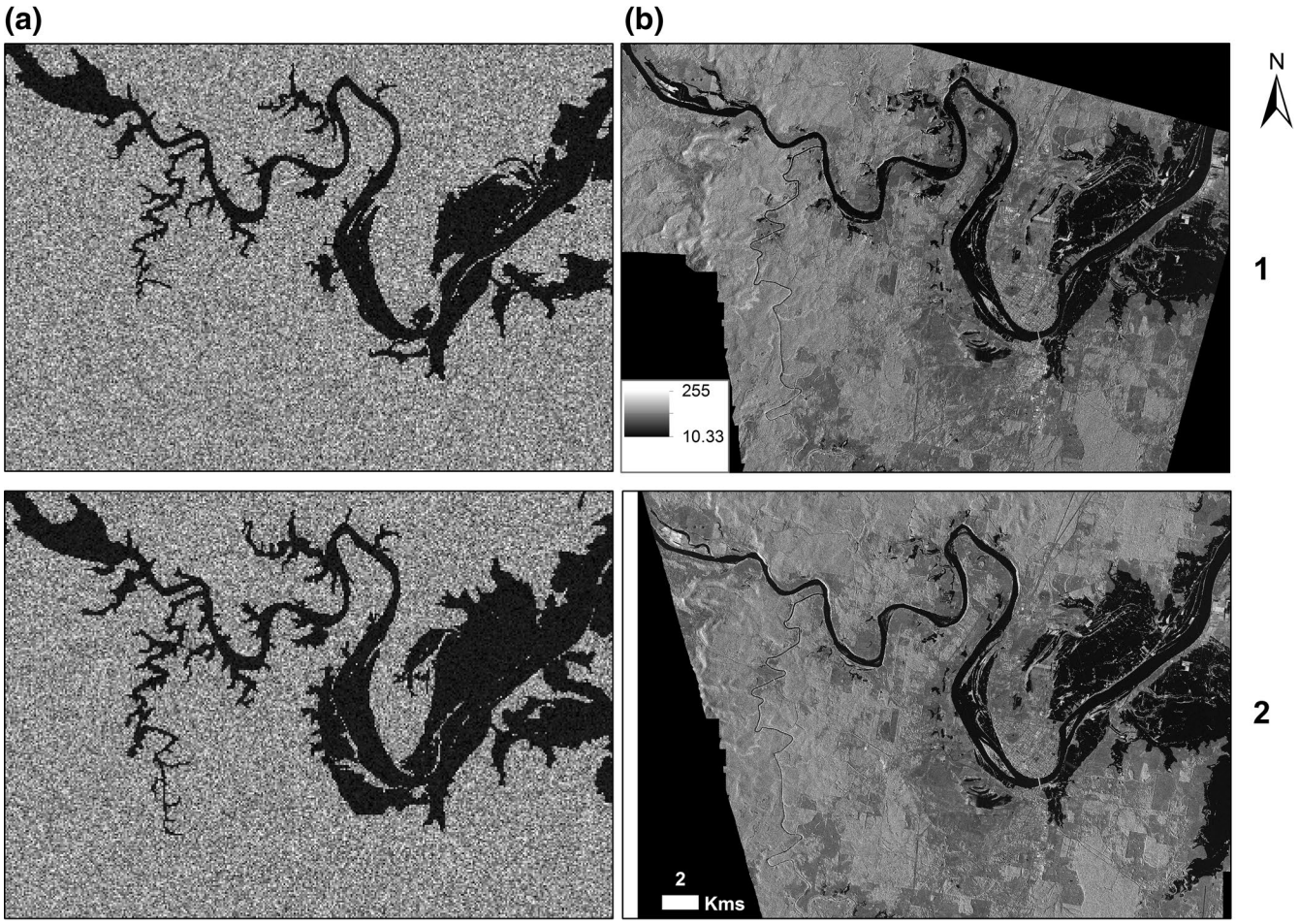

Figure 4. Synthetic and real SAR images juxtaposed in columns (a and b), respectively, for images acquired on January 12, 2011 at 18:03 h and January 13, 2011 at 07:33 h, hereafter referred to as assimilation time steps 1 and 2, respectively.

4. Results and Discussion

The impacts of assimilating synthetic SAR observations on the open loop ensemble, generated by propagating uncertain simulated forecast inflows through LISFLOOD-FP, are summarized here. The real SAR images for each assimilation time step upon which the synthetic images were derived are shown in Figures 4a and 4b respectively, with only the synthetic images used for the assimilation experiment in this study. Forecast performance was first evaluated in terms of flood extent, followed by a spatial analysis of water depth. Floodplain flow velocity errors were also spatially assessed in this study for the first time in assimilation literature. Finally, the impacts on in-channel performance were quantified via discharge and WL estimation capabilities.

4.1. Impact on Simulated Inundation Extent

Contingency maps from comparing the assimilated and open loop forecasts to the synthetic truth are displayed in Figure 5, with the assimilation of Image II also showing the effect of using different priors. From the contingency maps, it is evident that the open loop consistently over-estimated the true inundated area from the large number of "false alarm" pixels, although this was expected due to the positively biased inflow errors used in this experiment. The closed loop, otherwise known as the assimilated ensembles, exhibited lower numbers of false alarm pixels as a consequence of the assimilation. Positively biased models are notoriously difficult to distinguish using binary pattern matching measures, and thus further highlights the

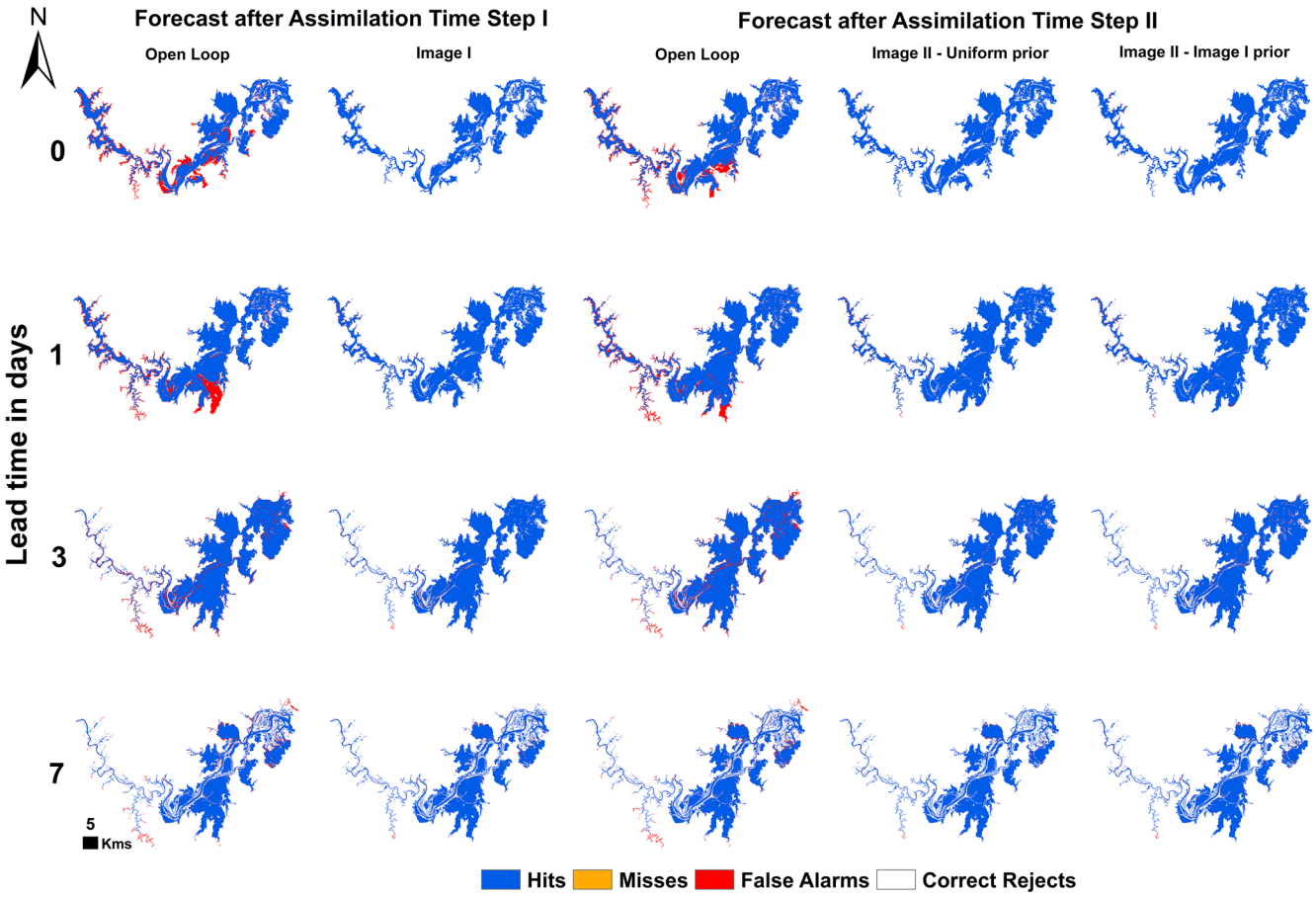

Figure 5. Contingency maps comparing the forecast and true flood extents for the open loop and assimilation runs, at (a) the first assimilation time step and (b) at the second assimilation time step. Note that for the second assimilation time step, the effect of using different prior distributions is also illustrated.

model selection potential of the proposed algorithm (see Stephens et al. (2014) for an in depth discussion). The main shift after the assimilation was from the false alarm class to correct rejects, while the number of hits remained fairly consistent. This was expected as the open loop constantly over-estimated the extent with most of the true extent being correctly simulated and the performance of the ensemble members differing primarily in terms of false alarm rates.

The assimilation of synthetic "observed" flood extents generated from the true flood extents should ideally nudge the model trajectory toward the truth, with more observations helping to constrain the forecast and resulting in fewer false alarms. While the assimilated forecast is consistently better than the open loop, based on the spatially distributed figures, no obvious differences can be observed between the assimilation of Images I and II or the use of different priors from Figure 5. A limited number of pixels changed their wet-dry status in the narrow valley between Lilydale and Rogan's Bridge (see Figure 3 for locations). This was in line with expectations as the assimilation of extent can only have limited impact in regions where flood extents are not sensitive to subtle changes in the water depth, as observed in the topographically constrained upstream part of the reach. Similarly, the assimilation was not expected to produce large impacts near the downstream boundary between Brushgrove and Yamba, where the inundation is dominated by tidal backwater effects, since the tidal level uncertainty is not considered in this experiment. The largest reductions in over-prediction (Boxes 1 and 2 in both contingency maps) were observed in the central part of the model domain around the Grafton-Ulmarra area, with a large and gentle-slope floodplain and limited backwater effects.

The assimilation had visible positive impacts until a lead time of 7 days as evident from Figure 5, although the open loop errors progressively decreased with lead time as the inflow magnitudes and the errors

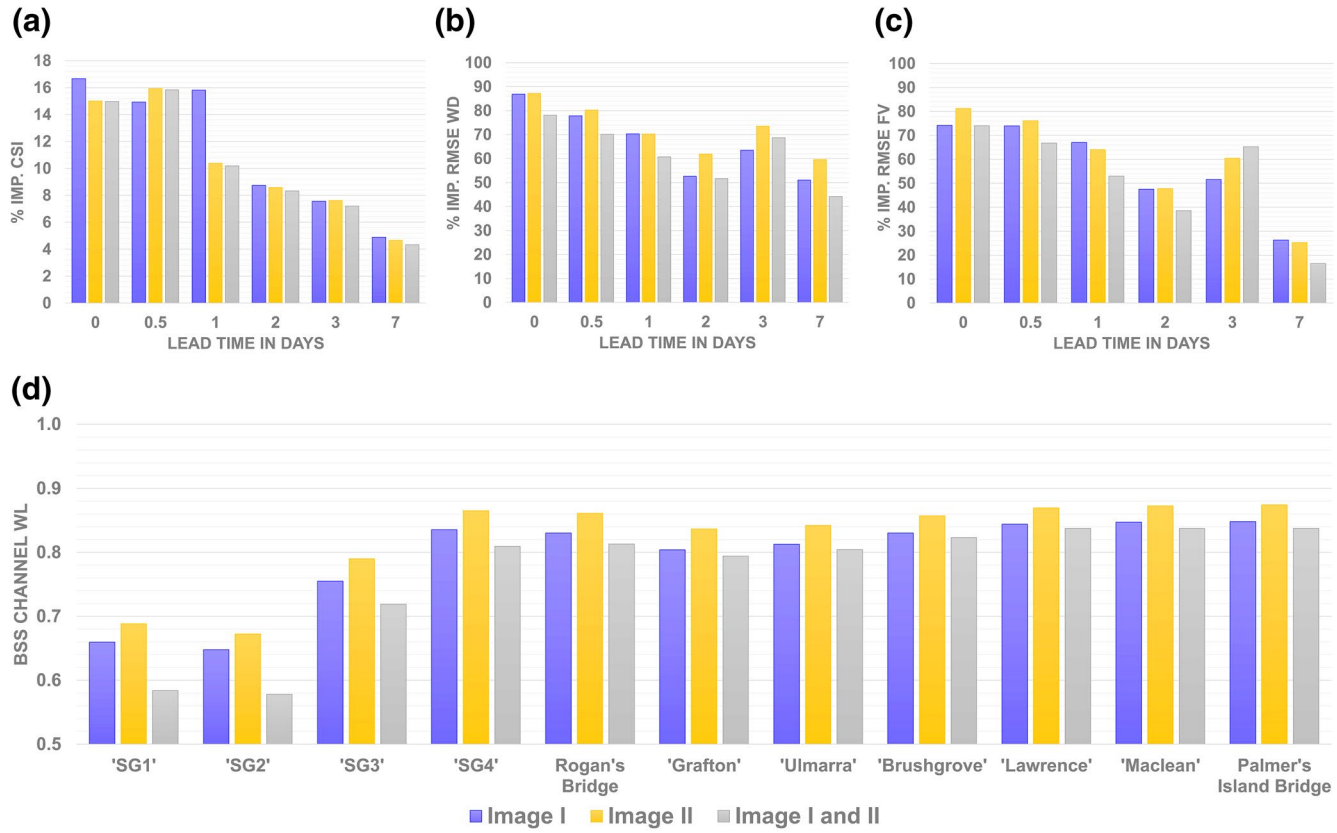

Figure 6. Percentage improvement before and after the assimilation for different lead times from the instance of assimilation, with the different plots showing the improvements in (a) the Critical Success Index for flood extent comparisons, (b) RMSE for water depths, (c) RMSE for flow velocities, and (d) the BSSs for the channel WL simulations, at the gauges along the main stem of the river (locations shown in Figure 3). Image I refers to the assimilation of only Image I, Image II refers to the assimilation of Image II using a uniform prior, while Images I and II refer to the sequential assimilation of Image II using the assimilated forecast from Image I as the prior. BSS was calculated from the assimilation time step to the end of the forecast.

decayed during the falling limb. The percentage improvement in the CSI (Figure 6a) reveals some interesting differences between the various assimilation configurations. In terms of acquisition timing with respect to the hydrograph, assimilating the first image ~6 h after the peak reduced false alarms by ~17% (purple bars), while the second image ~18 h post peak reduced overestimation by ~15% at the assimilation time steps using either the uniform priors (yellow bars) or the Image I priors (gray bars). This implied that the temporal position of Image I was more favorable for the assimilation. The percentage improvement in the CSI values decreased with increasing lead time but continued to produce improvements up to 7 days (~4%).

Until the 24 h (1 day) lead time, there were notable differences in the magnitude of improvements achieved from the three cases examined herein. However, subsequently the assimilation resulted in nearly similar improvements in the flood extent forecast for all of the different configurations. As the assimilated images were acquired close together during the falling limb, the correlation between them was expected to be high, and thus incorporating information from both images (i.e., using the Image I prior for the assimilation of Image II) was expected to yield greater forecast improvements. However, the gray bars are consistently lower (albeit marginally) than the yellow bars resulting from the use of a uniform prior, demonstrating that this assumption is not always true. The correlation time between images is thus an important factor to consider for sequential flood extent assimilation, as this experiment highlights the influence of the selected prior on the forecast quality.

4.2. Impact on Floodplain Water Depth Simulation

The forecast minus true WDD is shown in Figure 7 for the first and second assimilation time steps, with the same experimental configurations as in the previous section and for different lead times. Positive errors

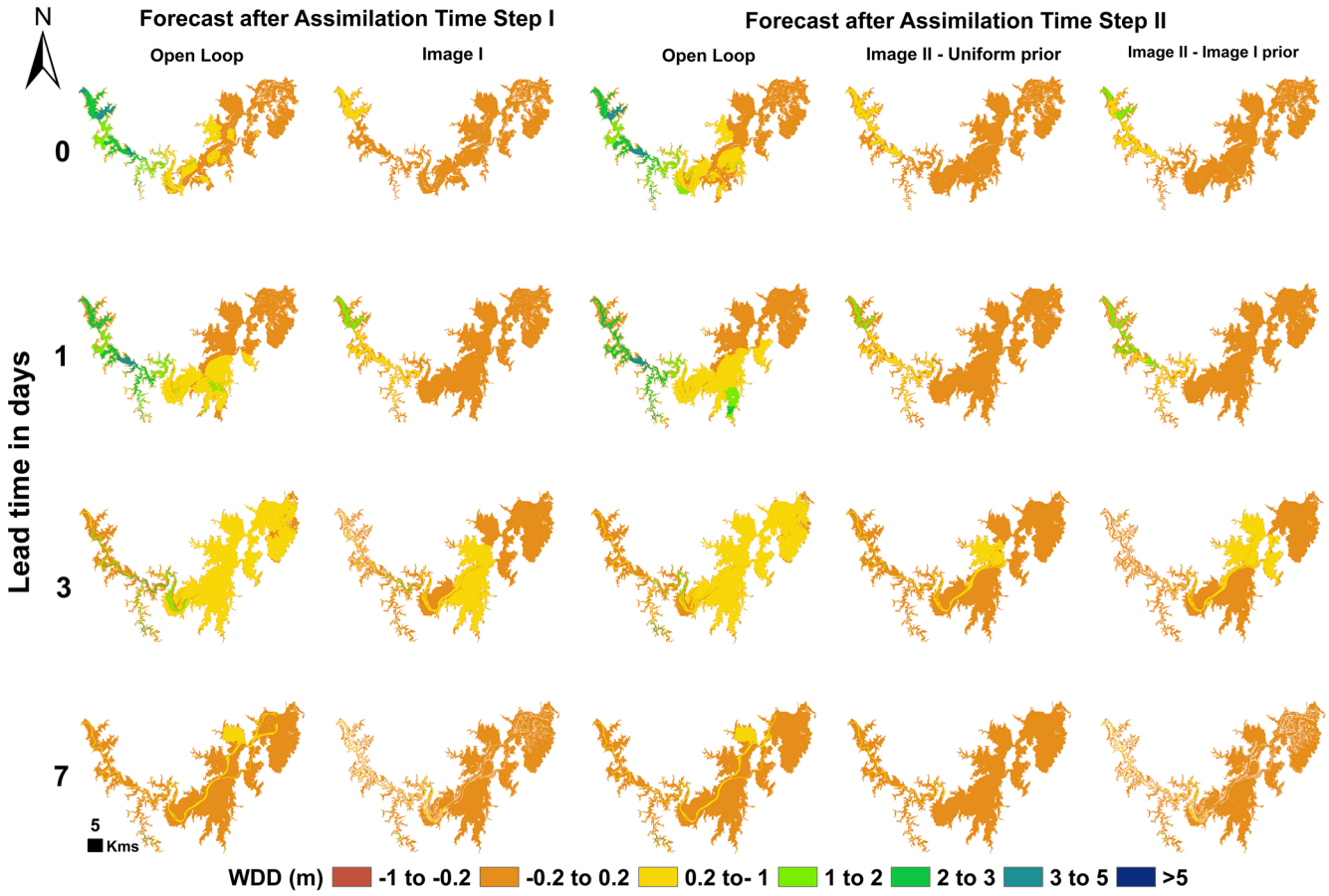

Figure 7. Forecast ensemble mean minus true water depth for different lead times, showing (a) lead times measured from the first time step after assimilating only the first image, and (b) lead times after the sequential assimilation of both images. Colors refer to the direction of the deviation from the truth; that is orange is within ±20 cm of the truth, the different shades of green through to indigo represent excess water while rust indicates a lack of water. The difference is calculated as forecast minus truth, so positive errors represent over-estimation while negative errors show under-estimation. Note that for the second assimilation time step, the effect of using different prior distributions is also illustrated.

imply that the forecast over-estimates the true water depth, while negative errors indicate under-estimation. Darker colors in each direction are representative of the magnitude of the deviation from the truth. Since, the forecast has more positive errors than negative, the color discretization which best highlights the differences in the forecasts was chosen.

The first observation that can be made from Figure 7 is that the open loop was positively biased in most of the domain and the assimilation resulted in a strong reduction in the magnitude of errors. As an example, if considering the central portion of the catchment it is easily observed that a large number of pixels moved to lighter hues between the open loop and assimilated forecasts, implying a definitive nudge toward the truth, persistent up to the maximum examined lead time of 7 days. The WDD was reduced by the assimilation all over the domain, with notable error reductions in the upstream and downstream parts of the catchment as well. Even though flood extent improvements were not evident in these regions, the simulated water depths were improved and hence the flood hazard estimates could be considered to be more reliable.

In the WDD maps, the difference between the assimilation of Images I and II is much clearer, as is the difference due to the use of the two different priors. In fact, the assimilation performance for the assimilation of Images I and II using uniform priors led to nearly similar improvements but using the Image I forecast as a prior for Image II reduced the magnitude of improvement. For example, if the forecast at the second assimilation time step is compared across the open loop, assimilation with the uniform prior, and assimilation with the Image I prior, it is evident that the uniform prior led to better results. Consider the case of a lead

time of 1 day that is Row 2 of Figure 7. While the assimilated forecast was better than the open loop in both cases, using the uniform prior led to larger error reductions. On examining the upstream part of the catchment, it is clear that the uniform prior case reduced positive WDD errors by several orders of magnitude, bringing a large number of pixels from dark green (>2 m errors) to yellow (0.2–1 m), while using the Image I prior only changed these pixels to light green (1–2 m) for many pixels.

Comparing the percentage improvement in the lumped spatial RMSE water depth values between the assimilation and the open loop forecasts in Figure 6b, revealed that persistent improvements of ∼60% could be retained after the assimilation of Image II for at least 7 days. Indeed, the assimilation at the second time step resulted in greater improvements using the uniform prior. In fact, improvements >80% in the spatial water depth simulations could be achieved at the assimilation time steps, which is significant in terms of the potential of flood extent assimilation to improve global flood inundation forecasts. Assimilating Image I and Image II led to an average improvement of ∼60% in the RMSE WD statistic, while using Image I as a prior for Image II reduced the magnitude of improvements (∼50%) confirming the findings from Figure 7. The WD forecast improvements first decayed with time up to a lead time of 2 days and then started to increase again. This is expected to be a consequence of the selection of particles which continue to perform well even toward the end of the simulation. Moreover, the assimilation of Image I was more beneficial for shorter lead times while Image II led to greater forecast improvements for longer lead times.

Note that the images available to this study were at favorable temporal positions with respect to the flood hydrograph that is post peak, they proved to be particularly informative for the inundation forecast. One of the reasons for this is that after the flood peak has passed and flows are rapidly transitioning between the channel and the floodplains, inflow errors which are proportional to the flow magnitudes also start to decay with time after the peak. Forecast errors can thus be better constrained by observations acquired after the flood peak, than by images covering the rising limb, since inflow errors dominate during this time. Moreover, since errors increased proportional to the increasing inflows during the rising limb, the particle weights assigned based on a given observation remained relevant to the forecast for very short lead times. The assimilation of early images with limited flood extents could, therefore, potentially lead to inconsistent increments. This finding is consistent with earlier studies which found that flood extent observations acquired just after the peak had higher information content for hydraulic model parameterization (Gobeyn et al., 2017; Wood, 2016; Wood et al., 2016a). Moreover, at this time, flood extents were more sensitive to water depth changes due to transitioning flows. This allows the assimilation of post-peak images to better constrain the spatial water depth forecast. Although this implies that the algorithm is sensitive to observation spatiotemporal characteristics, it is expected that the assimilation will nonetheless improve the forecast for most cases, but the magnitude of improvements may vary. Note that all references to the flood peak in this paragraph refer to the peak inflow value in the input hydrograph, while the inundation peak observed in the floodplain follows much later (several hours to several days) based on the catchment morphology. This implies that "post-peak" image assimilation can still lead to forecast improvements which enhance emergency preparedness and save lives.

4.3. Impact on Floodplain Flow Velocity Simulation

The distributed impact of flood extent assimilation on flow velocities was estimated through the (FVD) maps shown in Figure 8, computed as the pixel-wise difference between the expectation and the truth at different lead times from the assimilation. The product of flow velocities and depths constitute the flood hazard, which determines the potential for casualties during floods and are therefore important to be simulated well. The cell velocities were computed by combining the x- and y-components of cell edge velocities produced as outputs by LISFLOOD-FP. The color scheme used here is similar to the previous section, although it is important to note that the discretization of classes differs based on the range of bias values observed. The light green therefore corresponds to the minimum error class ±1 cm/s while darker shades going toward indigo indicate over-estimation of velocities and the darker shades toward rust indicate under-estimation. Readers should note that this study makes no claims about the accuracy of the flow velocities simulated by the truth model, because they cannot be validated practically in the field at present.

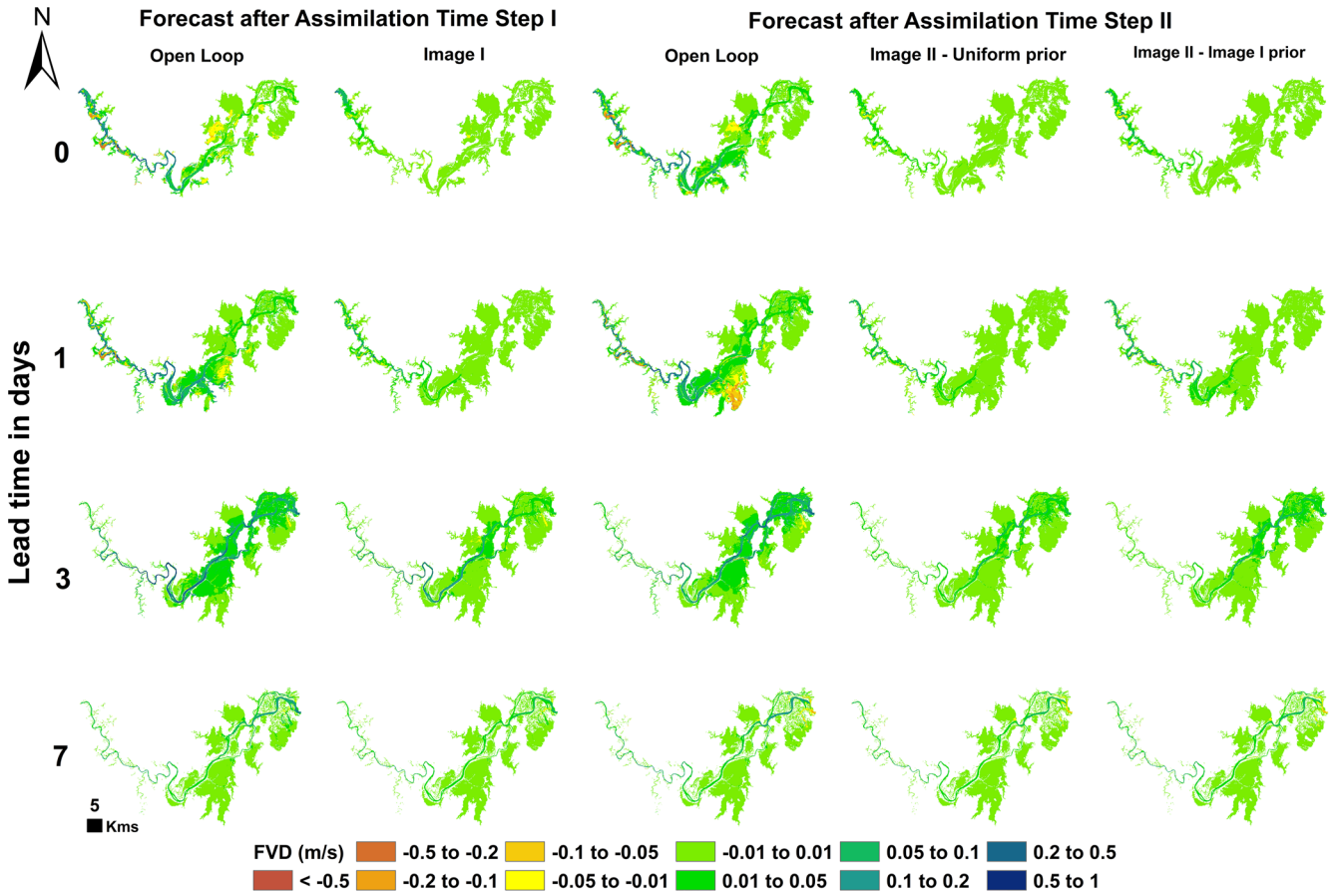

Figure 8. As for Figure 7 but for absolute flow velocity magnitudes rather than water depth.

However, the future looks promising with ongoing research to estimate flow velocities from drone-based (Tauro & Salvatori, 2017; Tauro et al., 2016a, 2016b) and crowdsourced (Le Boursicaud et al., 2016; Le Coz et al., 2016; Muste et al., 2011) video observations, which could soon make this type of validation possible. Since this study presents a synthetic experiment, the benchmark flow velocities can be used for evaluation. This analysis may provide useful insights on (a) the distributed spatial impact of the assimilation on this state variable which has never previously been examined, and (b) investigate the temporal evolution of these errors with increasing lead time.

The first and most obvious observation from the FVD maps illustrated in Figure 8 is that the assimilation improved the forecast for all examined lead times, consistent with the findings of previous sections. Interestingly, the assimilation was able to locally correct for both under- and over-estimation, introduced in some regions by the evolution of the inflow errors in the open loop. Since the absolute values of velocity in the floodplain were very low, the largest errors in the velocity simulations were observed in and around the channel. This was expected since accurately representing the channel conveyance plays a major role and presents a significant scientific challenge due to the large quantity of field information on bathymetry and roughness necessary to achieve this perfectly. The general trend of over-estimating velocities in the open loop was reduced after the assimilation, with no visible differences' observable between the assimilation of Image I and II or the use of different priors. Encouragingly, the impact of the assimilation remained positive and rather constant across all examined lead times. More importantly, the forecast was never degraded by the assimilation, despite studies having shown that degradation is a distinct possibility for hydraulic data assimilation (see Andreadis & Schumann, 2014, for details). This spatial assessment of flow velocities illustrated that the assimilation improved all model states and did not negatively impact the delicate equilibrium between them.

The percentage improvement in the spatial mean RMSE in the flow velocity is illustrated in Figure 6c. Average improvements of ~50% could be observed in the spatial velocity simulations after the assimilation, with Image II generally producing greater improvements than Image I over the different lead times. One possible reason for this could be the timing of Image II, which was almost a day after the observed peak. At this time of the flood, the flows are out of bank and rapidly transitioning between the channel and the floodplains, leading to inundation patterns primarily driven by model propagation as opposed to the inflow errors. As current inflow errors have comparatively less importance, the assimilation of flood extent observations was able to effectively constrain the model trajectory longer. The use of different priors also influenced the assimilation performance in this case. Using the uniform prior is generally better but for the 3 days lead time the Image I prior led to greater improvements in the velocity forecast. This trend is similar to what was observed for the water depths in the previous section where the 3 days lead time showed a greater forecast improvement than some of the shorter (1 and 2 days) and longer (7 days) lead times examined.

4.4. Impact on Channel Water Level Simulation

The assimilation results for the channel WLs, at 11 gauging stations (four synthetic and seven real gauges; locations shown in Figure 3) along the main stem of the river are illustrated in Figure 9. Note that the uniform prior was used for both Image I and Image II in the assimilated channel WL forecast trajectory shown in the plots. Discharge plots have been excluded for brevity as they exhibited similar characteristics. A cursory examination reveals that the assimilation always improved the estimation of flow and depth in the channel, with the assimilation of Image II exhibiting stronger impacts than the assimilation of Image I. As expected, due to the positively biased inflows, the open loop WL hydrograph was consistently higher than the truth, but the assimilation was able to recover the truth from the ensemble. Moreover, improvements persisted for up to ~7 days after the assimilation. This is crucial from an emergency management perspective, especially for the populated area downstream of Grafton where the flood peak arrived after both images had been assimilated. Of course, this may be catchment specific and dependent on the inflow error generation model chosen here, requiring further testing in catchments with different hydraulic characteristics to demonstrate scalability.

The WL time series was also statistically evaluated using the BSS statistic as shown in Figure 6d, which compares the forecast skill of the assimilated ensemble with the open loop. With values ranging from $-\infty$ to an ideal value of 1, BSS quantifies the percentage improvement in the quadratic errors of the forecast, in this case for the simulated channel WLs. The errors as well as the corresponding percentage improvement as a consequence of the assimilation were consistent from upstream to downstream and were ~70% on average (mean of all bars ~0.7 in Figure 6d). Indeed, as the errors were introduced as inflows at the upstream boundary, the improvements from assimilation were sustained for longer durations downstream (E. S. Cooper et al., 2018). The persistence in the improvements could also be a function of the 3 days temporal correlation imposed in this study following the work of García-Pintado et al. (2013). Indeed, the stationary bias in the true mean error imposed here undoubtedly resulted in more consistent improvements, since the correlation in time means that the different particles differed in terms of magnitude but not in terms of timing, which is unrealistic for real forecasting scenarios. While this serves to emulate the forecasts of hydrological models using precipitation forecasts as inputs, where the knowledge on the mean error evolution is typically lacking, future investigations should examine more complex methods to approximate the real world error dynamics. There was a general increasing trend in assimilation benefits from upstream to downstream in the discharge statistics, although the magnitude of the trend was marginal. This is expected to be a function of the forecast velocity improvements achieved due to the assimilation, as the WLs do not exhibit this trend.

4.5. Likelihood Function Efficiency as Compared to Hostache et al. (2018)

Figure 10 shows the results from a comparison with the flood extent assimilation method proposed by Hostache et al. (2018), hereafter referred to as PF-ST (PF-Standard), and the approach of this study, hereafter called PF-MI. In this section, both of the flood extent assimilation approaches were applied to the HISD (Part I) and the 2011 (Part II) events described in Section 2.3. Images were assimilated at different temporal

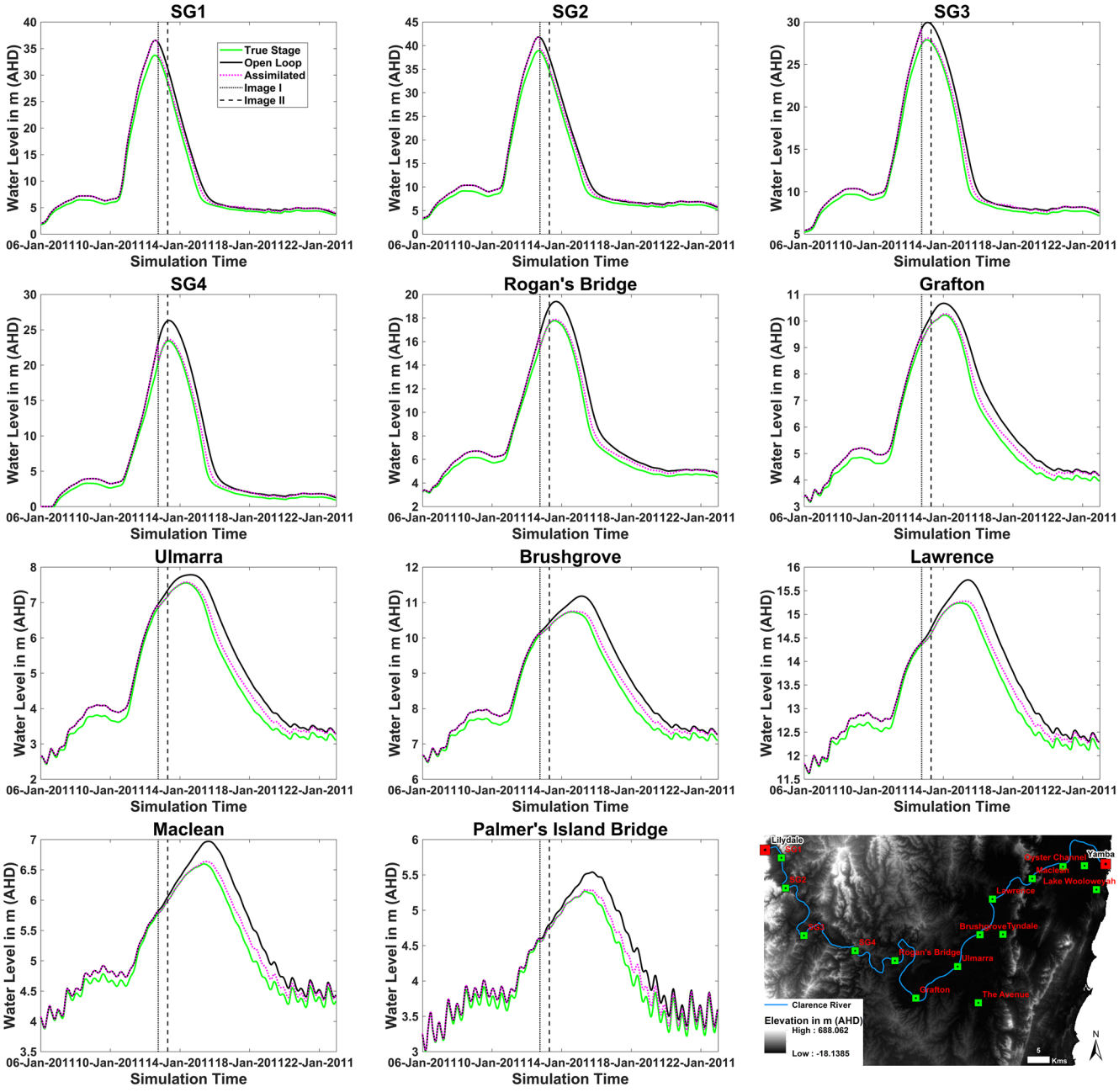

Figure 9. Water level time series at the synthetic and real gauge locations along the main stem of the Clarence River. The dashed black vertical lines show the timings of the two assimilated images. The plots for each gauge have been arranged upstream to downstream in the different rows of the figures, with the final subplot illustrating the gauge locations for reference.

instances across the hydrograph, including during the early rising limb, pre-peak, peak, post-peak, and late falling limb using PF-LIST and PF-MI for the HISD event. For the 2011 event, the observations corresponding to the real-world SAR acquisition times were assimilated. Note that all images were assimilated using a uniform prior, to assess the individual impact of each standalone observation. To evaluate the two likelihood functions, the distribution of simulated WLs at Lawrence (an example downstream gauge—selected due to clear differences in the performances of both algorithms) was plotted, along with the particle weights assigned by both algorithms, the true WL and the open loop mean. This allowed visualizing the particles selected by each algorithm after the assimilation and lends insights to understand the performance in terms of the RMSE at the assimilation time step and the simulated WL hydrographs.

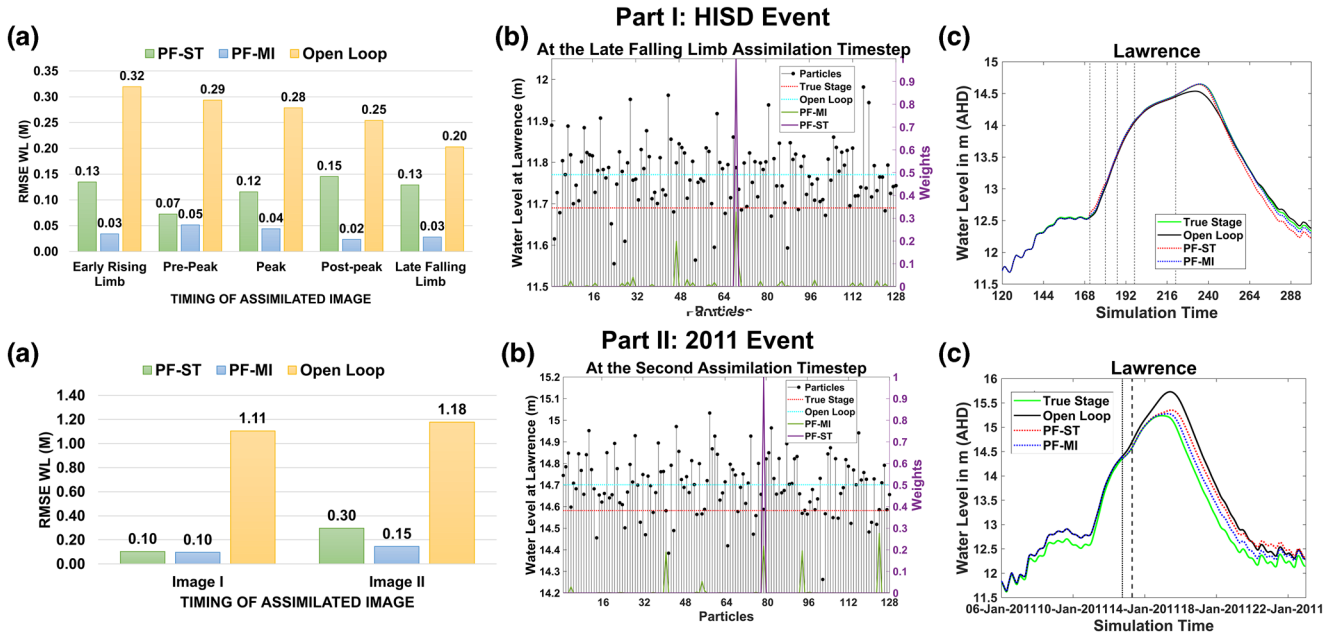

Figure 10. An illustration of the performance of the PF-MI and PF-ST, applied to two different flood events, evaluated in terms of the channel water level forecasts. Parts I and II showcase applications to the HISD and the 2011 flood events, respectively. Panels (a and c), show the RMSE at the assimilation timestep and the channel water level hydrographs at Lawrence, respectively. Panel (b) shows the water levels at Lawrence for all the particles, along with the true water level, the open loop mean, and the particle weights computed by both algorithms at a given assimilation time step.

It is clear from Figure 10, that the flood extent assimilation led to forecast improvements for both events and that the PF-MI approach consistently produced performances equivalent to or better than the PF-ST approach. For the unbiased inflows used in the HISD event (Part I), the open loop already performs quite well, and it was thus harder to make improvements. However, the PF-MI produced consistently greater RMSE improvements in comparison to the PF-ST for all the images considered in (c), despite forecasting of "flashy" events being notoriously difficult. Panel (b) reveals that the PF-MI selected a range of particles (see the number of nonnegligible weights in green), which allowed the forecast to stay on track for longer lead times. In contrast, the PF-ST (purple line) selected only one particle with a full probability weight of 1, which may have performed excellently at the assimilation time step but may not remain relevant for long lead times. The consequences of this are evident from (d), which shows that the PF-MI forecast closely followed the truth for more than 3 days after the assimilation of the final observation. While it could be argued that this 3-days improvement was a function of the inflow error temporal correlation, this seems highly unlikely since the improvement was only visible for PF-MI even though the same ensemble was used for the PF-ST tests.

Similar observations can be made from Part II of Figure 10, where the assimilation methods were applied to the 2011 event. In this case, (b) shows that the PF-ST selected 1 well-performing particle (ensemble member #79), which reproduced the true WL at Lawrence almost exactly at the assimilation time step, while PF-MI selected multiple well-performing particles. However, after performing well for a lead time of 2 days (d), the PF-ST forecast started deviating from the truth and after a lead time of 7 days the errors exceeded the open loop errors with a large positive bias. These findings further highlight the need to avoid particle degeneracy, and the crucial role it plays in the persistence of assimilation benefits. Moreover, the PF-MI produced greater improvements both at the assimilation time step and for longer lead times, implying that it has the potential to improve forecasts with rapidly varying hydrological conditions. Future work should investigate the impact of using different model likelihood functions on events of varying intensity and duration across diverse catchments. Furthermore, effective temporal decay functions should be designed which can objectively relax the observation influence on the forecast as the trajectories become independent of the instantaneous initial conditions observed at the image acquisition time.

5. Conclusions

A method for assimilating SAR image-derived flood extent into a hydraulic model was proposed, using MI as a novel model likelihood function implemented within a SIS PF framework. MI quantifies the reduction of uncertainty in predicting the observation knowing the model simulation at each assimilation time step. Each particle, representing a particular model realization with a unique erroneous inflow input, was assigned a weight using the MI-based metric. The efficiency of the proposed assimilation algorithm was demonstrated through synthetic experiments, which evaluated assimilation impacts on forecast flood extent, floodplain water depths, floodplain flow velocities, channel depth, and flow. The sensitivity of flood extent assimilation algorithms, toward the prior distribution chosen for the sequential assimilation of SAR images, was also demonstrated in this study. Interestingly, using the uniform prior led to better forecast quality as compared to the use of the assimilated forecast from the previous observation as the prior. The influence of the selected prior distribution on the estimation of the true state posterior needs to be better understood through further investigations.

The assimilation resulted in mean flood extent forecast improvements of ~16% at the assimilation time steps, which decayed to ~4% for a 7 days lead time. Simulated water depths evaluated both in the channel and in the floodplain, improved by more than 70% in some cases due to the flood extent assimilation, and were consistent from upstream to downstream. The distributed assessment of forecast flood depths and flow velocities in the floodplain showcased the robustness of the proposed algorithm toward positive and negative biases. This implies that the observation was able to facilitate the selection of particles that performed well across the entire domain for post-peak images.

Forecast improvements were shown to persist for up to 7 days after the assimilation, even though the magnitude of the improvements decayed with lead time. Some of the persistence in assimilation benefits can be attributed to the (rather strong) temporal error correlation imposed on the inflows, however the correlation length is only of 3 days and thus cannot completely explain the weeklong benefits. Using the PF-MI thus seems viable for operational applications and can lead to improved forecasts due to larger information uptake along with spatial uncertainty considerations. Comparisons with the PF-ST algorithm proposed by Hostache et al. (2018), revealed that the PF-MI not only produced greater or equivalent improvements at the assimilation time step, but also that these improvements were also persistent for longer lead times.

The experiments presented in this study demonstrate that the assimilation of SAR-derived flood extents improve hydraulic flood inundation forecast skill despite being a prognostic state variable. Forecast skill was consistently and persistently improved through assimilation in the channel and in the floodplain, implying that the assimilation of flood extents into hydraulic inundation forecasts could be significant from an emergency management perspective.

Acknowledgments
This study has been conducted within the framework of the project "Improving flood forecast skill using remote sensing data," funded by the Bushfire and Natural Hazards CRC of Australia. The authors also acknowledge E-Geos for providing SAR imagery and timely support during pre-processing. Additionally, gratitude is extended to the Australian Bureau of Meteorology (http://www.bom.gov.au/waterdata/) and New South Wales Manly Hydraulics Laboratory (http://new.mhl.nsw.gov.au/) for the gauge data, in addition to Geoscience Australia and the Clarence Valley Council for sharing field and ancillary data. Specifically, Chris Huxley from BMT WBM and Kieran McAndrew from the Clarence Valley Council are thanked, for providing the design flood inflows and outflows for the Lower Clarence Valley model implementation presented in Farr and Huxley (2013). Antara Dasgupta's time was funded by an IITB-Monash Research Academy PhD scholarship for the project entitled "Optimizing SAR-based Flood Extent Assimilation for Improved Hydraulic Flood Inundation Forecasts," project number IMURA0351. Renaud Hostache's contribution was supported by the National Research Fund of Luxembourg (FNR) through the CASCADE project, grant number C17/SR/11682050.

Data Availability Statement

All the data and codes used for this study are available at https://doi.org/10.6084/m9.figshare.13152908.v1 distributed under the CC BY 4.0 License.

References

Abbaszadeh, P., Moradkhani, H., & Daescu, D. N. (2019). The quest for model uncertainty quantification: A hybrid ensemble and variational data assimilation framework. *Water Resource Research, 55*, 2407–2431. https://doi.org/10.1029/2018WR023629

Abbaszadeh, P., Moradkhani, H., & Yan, H. (2018). Enhancing hydrologic data assimilation by evolutionary Particle Filter and Markov Chain Monte Carlo. *Advances in Water Resources*, *111*, 192–204. https://doi.org/10.1016/j.advwatres.2017.11.011

Alfieri, L., Cohen, S., Galantowicz, J, et al. (2018). A global network for operational flood risk reduction. *Environmental Science & Policy*, *84*, 149–158. https://doi.org/10.1016/j.envsci.2018.03.014

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*, 44–54. https://doi.org/10.1016/j.advwatres.2014.06.006

Arcement, G. J., & Schneider, V. R. (1989). *Guide for selecting Manning's roughness coefficients for natural channels and flood plains.* Water-Supply Papers (Vol. 2339, pp. 1–38). Books and Open-File Reports Section: US Geological Survey. Retrieved from https://dpw.lacounty.gov/lacfcd/wdr/files/WG/041615/Guide%20for%20Selecting%20n-Value.pdf

Archer, L., Neal, J. C., Bates, P. D., & House, J. I. (2018). Comparing TanDEM-X data with frequently-used DEMs for flood inundation modelling. *Water Resource Research*, *54*, 10205–10222. https://doi.org/10.1029/2018WR023688

Argenti, F., & Alparone, L. (2002). Speckle removal from SAR images in the undecimated wavelet domain. *IEEE Transactions on Geoscience and Remote Sensing*, *40*, 2363–2374. https://doi.org/10.1109/TGRS.2002.805083

Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-GaussianBayesian tracking. *Signal Process IEEE Transaction*, *50*, 174–188. https://doi.org/10.1109/78.978374

Bates, P. D., & De Roo, A. P. J. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, *236*, 54–77.

Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, *387*, 33–45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Bates, P., Trigg, M., Neal, J., & Dabrowa, A. (2013). *LISFLOOD-FP User manual*. Bristol School of Geographical Sciences; University of Bristol (2013).

Buytaert, W., Dewulf, A., De Bièvre, B, et al. (2016). Citizen Science for Water Resources Management: Toward Polycentric Monitoring and Governance? *Journal of Water Resources Planning and Management*, *142*, 01816002. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000641

Charrow, B., Kumar, V., & Michael, N. (2014). Approximate representations for multi-robot control policies that maximize mutual information. *Autonomous Robots*, *37*, 383–400. https://doi.org/10.1007/s10514-014-9411-2

Chen, H. M., Varshney, P. K., & Arora, M. K. (2003). Performance of mutual information similarity measure for registration of multitemporal remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 2445–2454. https://doi.org/10.1109/TGRS.2003.817664

Chini, M., Pelich, R., Pulvirenti, L, et al. (2019). Sentinel-1 InSAR Coherence to Detect Floodwater in Urban Areas: Houston and Hurricane Harvey as A Test Case. *Remote Sens*, *11*, 107. https://doi.org/10.3390/RS11020107

Cooper, E. S., Dance, S. L., Garcia-Pintado, J, et al. (2018). Observation impact, domain length and parameter estimation in data assimilation for flood forecasting. *Environmental Modelling & Software*, *104*, 199–214. https://doi.org/10.1016/j.envsoft.2018.03.013

Cooper, E. S., Dance, S. L., García-Pintado, J, et al. (2019). Observation operators for assimilation of satellite observations in fluvial inundation forecasting. *Hydrology and Earth System Sciences*, *23*, 2541–2559. https://doi.org/10.5194/hess-2018-589

Cover, T. M., & Thomas, J. A. (2005). *Elements of information theory*. Wiley. https://doi.org/10.1002/047174882X

CRED, & UNISDR (2015). In M. Wahlstrom & D. Guha-Sapir (Eds.), *The human costs of weather related disasters*. Intergovernmental Panel on Climate Change. Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.004

Dasgupta, A., Grimaldi, S., Ramsankaran, R, et al. (2018a). Flood mapping using synthetic aperture radar sensors from local to global scales. In G. J. Schumann, P. D. Bates, H. Apel, & G. T. Aronica (Eds.), *Global flood hazard: Applications in modeling, mapping, and forecasting, volume 233 of geophysical monograph series*. 1st ed. (pp. 55–77), John Wiley & Sons.

Dasgupta, A., Grimaldi, S., Ramsankaran, R. A. A. J, et al. (2018b). Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches. *Remote Sensing of Environment*, *215*, 313–329. https://doi.org/10.1016/j.rse.2018.06.019

de Almeida, G. a. M., Bates, P., Freer, J. E., & Souvignet, M. (2012). Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling. *Water Resource Research*, *48*, W05528. https://doi.org/10.1029/2011WR011570

De Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N., & Verhoest, N. E. C. (2006). Assessment of model uncertainty for soil moisture through ensemble verification. *Journal of Geophysical Research*, *111*, D10101. https://doi.org/10.1029/2005JD006367

Dottori, F., Mentaschi, L., Bianchi, A., Alfieri, L., & Feyen, L. (2020). In *Adapting to rising river flood risk in the EU under climate change*. Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/14505

Dottori, F., Salamon, P., Bianchi, A, et al. (2016). Development and evaluation of a framework for global flood hazard mapping. *Advances in Water Resources*, *94*, 87–102. https://doi.org/10.1016/j.advwatres.2016.05.002

Dowling, T. I., Brooks, M., & Read, A. M. (2011). Continental hydrologic assessment using the 1 second (30 m) resolution Shuttle Radar Topographic Mission DEM of Australia. In *MODSIM 2011 - 19th International Congress Model Simulation - Sustain Our Future Understanding Living With Uncertain 2395–2401*.

Dumedah, G. (2012). Formulation of the evolutionary-based data assimilation, and its implementation in hydrological forecasting. *Water Resources Management*, *26*, 3853–3870. https://doi.org/10.1007/s11269-012-0107-0

Dumedah, G. (2015). Toward essential union between evolutionary strategy and data assimilation for model diagnostics: An application for reducing the search space of optimization problems using hydrologic genome map. *Environmental Modelling & Software*, *69*, 342–352. https://doi.org/10.1016/j.envsoft.2014.09.025

Dumedah, G., & Coulibaly, P. (2013). Evaluating forecasting performance for data assimilation methods: The ensemble Kalman filter, the particle filter, and the evolutionary-based assimilation. *Advances in Water Resources*, *60*, 47–63. https://doi.org/10.1016/j.advwatres.2013.07.007

Dumedah, G., & Coulibaly, P. (2014). Integration of an evolutionary algorithm into the ensemble Kalman filter and the particle filter for hydrologic data assimilation. *Journal of Hydroinformatics*, *16*, 74–94. https://doi.org/10.2166/hydro.2013.088

Emerton, R. E., Stephens, E. M., Pappenberger, F, et al. (2016). Continental and global scale flood forecasting systems. *Wiley Interdiscip Review Water*, *3*, 391–418. https://doi.org/10.1002/wat2.1137

Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*, 343–367. https://doi.org/10.1007/s10236-003-0036-9

Farr, A., & Huxley, C. (2013). In *Lower clarence flood model update 2013*, (Technical report), BMT WBM. Available at: https://www.clarence.nsw.gov.au/page.asp?f=RES-FRE-61-54-36

Fearnhead, P., & Künsch, H. R. (2018). Particle Filters and Data Assimilation. *Annual Review of Statistics and Its Application*, *5*(1), 421–449. https://doi.org/10.1146/annurev-statistics-031017-100232

Fewtrell, T. J., Duncan, A., Sampson, C. C., et al. (2011). Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data. *Physics and Chemistry of the Earth*, *36*, 281–291. https://doi.org/10.1016/j.pce.2010.12.011

Fowler, A. M., Dance, S. L., & Waller, J. A. (2018). On the interaction of observation and prior error correlations in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *144*, 48–62. https://doi.org/10.1002/qj.3183

Fowler, A., & Van Leeuwen, P. J. (2012). Measures of observation impact in non-gaussian data assimilation. *Tellus Series A Dynamic Meteorology and Oceanography*, *64*, 17192. https://doi.org/10.3402/tellusa.v64i0.17192

Fowler, A., & Van Leeuwen, P. J. (2013). Observation impact in data assimilation: The effect of non-gaussian observation error. *Tellus Series A Dynamic Meteorology and Oceanography*, *65*, 20035. https://doi.org/10.3402/tellusa.v65i0.20035

Gallant, J. C., Dowling, T. I., Read, A. M, Wilson, N., Tickle, P., & Inskeep, C. (2011). 1 second SRTM Derived Digital Elevation Models User Guide. Geoscience Australia report, available at www.ga.gov.au/topographic-mapping/digital-elevation-data.html

Garambois, P. A., Larnier, K., Monnier, J., Finaud-Guyot, P., Verley, J., Montazem, A. S., & Calmant, S. (2019). *Effective channel and ungauged braided river discharge estimation by assimilation of multi-satellite water heights of different spatial sparsity. (hal-02308560)*

García-Pintado, J., Mason, D. C., Dance, S. L, et al. (2015). Satellite-supported flood forecasting in river networks: A real case study. *Journal of Hydrology*, *523*, 706–724. https://doi.org/10.1016/j.jhydrol.2015.01.084

García-Pintado, J., Neal, J. C., Mason, D. C, et al. (2013). Scheduling satellite-based SAR acquisition for sequential assimilation of water level observations into flood modelling. *Journal of Hydrology*, *495*, 252–266. https://doi.org/10.1016/j.jhydrol.2013.03.050

Giustarini, L., Hostache, R., Kavetski, D, et al. (2016). Probabilistic Flood Mapping Using Synthetic Aperture Radar Data. *IEEE Transactions on Geoscience and Remote Sensing*, *54*, 6958–6969.

Giustarini, L., Matgen, P., Hostache, R., & Dostert, J. (2012). From SAR-derived flood mapping to water level data assimilation into hydraulic models. *Remote Sens Agric Ecosyst Hydrol XIV*, *8531*, 85310U. https://doi.org/10.1117/12.974655

Giustarini, L., Matgen, P., Hostache, R, et al. (2011). Assimilating SAR-derived water level data into a hydraulic model: A case study. *Hydrology and Earth System Sciences*, *15*, 2349–2365. https://doi.org/10.5194/hess-15-2349-2011

Glasser, R., & Guha-Sapir, D. (2016). *Poverty and death: Disaster mortality 1996-2015*. The United Nations Office for Disaster Risk Reduction: Brussels, Belgium.

Gobeyn, S., Van Wesemael, A., Neal, J, et al. (2017). Impact of the timing of a SAR image acquisition on the calibration of a flood inundation model. *Advances in Water Resources*, *100*, 126–138. https://doi.org/10.1016/j.advwatres.2016.12.005

Godsill, S., & Clapp, T. (2001). Improvement strategies for Monte Carlo particle filters. In A. Doucet, N. de Freitas, & N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science (pp. 139–158). Springer, New York, NY, https://doi.org/10.1007/978-1-4757-3437-9_7.

Gong, W., Gupta, H. V., Yang, D, et al. (2013). Estimating epistemic and aleatory uncertainties during hydrologic modeling: An information theoretic approach. *Water Resource Research*, *49*, 2253–2273. https://doi.org/10.1002/wrcr.20161

Grimaldi, S., Li, Y., Pauwels, V. R. N., & Walker, J. P. (2016). Remote sensing-derived water extent and level to constrain hydraulic flood forecasting models: opportunities and challenges. *Surveys in Geophysics*, *37*, 977–1034. https://doi.org/10.1007/s10712-016-9378-y

Grimaldi, S., Li, Y., Walker, J. P., & Pauwels, V. R. N. (2017). Bathymetric survey of the Upper Clarence. *Dataset*, https://doi.org/10.4225/03/5a20708405ecd

Grimaldi, S., Li, Y., Walker, J. P., & Pauwels, V. R. N. (2018). Effective representation of river geometry in hydraulic flood forecast models. *Water Resource Research, 54*, 1031–1057. https://doi.org/10.1002/2017WR021765

Grimaldi, S., Schumann, G. J. -P., Shokri, A, et al. (2019). Challenges, opportunities, and pitfalls for global coupled hydrologic-hydraulic modeling of floods. *Water Resource Research*, *55*, 5277–5300. https://doi.org/10.1029/2018WR024289

Haile, A. T., & Rientjes, T. H. M. (2007). Uncertainty issues in hydrodynamic flood modeling. In *Proceedings of the 5th international symposium on spatial data quality SDQ 2007, modelling qualities in space and time, ITC*. Enscheda.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning. *Elements*, *1*, 337–387. https://doi.org/10.1007/b94608

Hawker, L., Neal, J. C., & Bates, P. (2019). Accuracy assessment of the TanDEM-X 90 Digital Elevation Model for selected floodplain sites. *Remote Sensing of Environment*, *232*, 111319. https://doi.org/10.1016/j.rse.2019.111319

Herbst, E., & Schorfheide, F. (2019). Tempered particle filtering. *Journal of Econometrics*, *210*, 26–44. https://doi.org/10.1016/j.jeconom.2018.11.003

Hirschmüller, H. (2008). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*, 328–341. https://doi.org/10.1109/TPAMI.2007.1166

Hoffmann, G. M., & Tomlin, C. J. (2010). Mobile sensor network control using mutual information methods and particle filters. *IEEE Transactions on Automatic Control*, *55*, 32–47. https://doi.org/10.1109/TAC.2009.2034206

Horkaew, P., & Puttinaovarat, S. (2017). Entropy-based fusion of water indices and DSM derivatives for automatic water surfaces extraction and flood monitoring. *ISPRS International Journal of Geo-Information*, *6*, 301. https://doi.org/10.3390/ijgi6100301

Horritt, M. S., & Bates, P. D. (2001). Predicting floodplain inundation: Raster-based modelling versus the finite-element approach. *Hydrological Processes*, *15*, 825–842. https://doi.org/10.1002/hyp.188

Hostache, R., Chini, M., Giustarini, L, et al. (2018). Near-real-time assimilation of SAR-derived flood maps for improving flood forecasts. *Water Resource Research*, *54*, 5516–5535. https://doi.org/10.1029/2017WR022205

Hostache, R., Matgen, P., Schumann, G, et al. (2009). Water level estimation and reduction of hydraulic model calibration uncertainties using satellite SAR images of floods. *IEEE Transactions on Geoscience and Remote Sensing*, *47*, 431–441. https://doi.org/10.1109/TGRS.2008.2008718

Hutchinson, M. F. (2011). *ANUDEM version 5.3 user guide*. Canberra: Fenner School of Environment and Society at the Australian National University.

Huxley, C., & Beaman, F. (2014). Additional crossing of the Clarence River at Grafton: flood impact, levee upgrade, and structural considerations. In *Hydraulic structures and society - engineering challenges and extremes* (pp. 1–8). Brisbane, Australia.

Intergovernmental Committee On Surveying & Mapping (2008). *ICSM Guidelines for Digital Elevation Data*, Elevation, (pp. 1–49). Australia: Intergovernmental Committee On Surveying & Mapping. Retrieved from https://www.icsm.gov.au/sites/default/files/2017-03/ICSM-GuidelinesDigitalElevationDataV1.pdf

Italian Space Agency (2009). In: COSMO-SkyMed SAR products handbook, Italian Space Agency.

Jarihani, A. A., Callow, J. N., Mcvicar, T. R, et al. (2015). Satellite-derived Digital Elevation Model (DEM) selection, preparation and correction for hydrodynamic modelling in large, low-gradient and data-sparse catchments. *Journal of Hydrology*, *524*, 489–506. https://doi.org/10.1016/j.jhydrol.2015.02.049

Johansen, A. M. (2015). On Blocks, Tempering and Particle MCMC for Systems Identification. *IFAC-PapersOnLine*, *48*, 969–974. https://doi.org/10.1016/j.ifacol.2015.12.256

Kantas, N., Doucet, A., Singh, S. S, et al. (2015). On particle methods for parameter estimation in state-space models. *Statistical Science*, *30*, 328–351. https://doi.org/10.1214/14-STS511

Kauffeldt, A., Wetterhall, F., Pappenberger, F, et al. (2016). Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software*, *75*, 68–76. https://doi.org/10.1016/j.envsoft.2015.09.009

Kwok, N. M., Gu, F., & Zhou, W. (2005). Evolutionary particle filter: re-sampling from the genetic algorithm perspective. In *2005 IEEE/RSJ international conference on intelligent robots and systems* (pp. 2935–2940). IEEE.

Lai, X., Liang, Q., Yesou, H., & Daillet, S. (2014). Variational assimilation of remotely sensed flood extents using a 2-D flood model. *Hydrology and Earth System Sciences*, *18*, 4325–4339. https://doi.org/10.5194/hess-18-4325-2014

Lai, X., & Monnier, J. (2009). Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: Mathematical method and test case. *Journal of Hydrology*, *377*, 1–11. https://doi.org/10.1016/j.jhydrol.2009.07.058

Lavers, D. A., Harrigan, S., Andersson, E, et al. (2019). A vision for improving global flood forecasting. *Environmental Research Letters*, *14*, 121002. https://doi.org/10.1088/1748-9326/ab52b2
Le Boursicaud, R., Pénard, L., Hauet, A, et al. (2016). Gauging extreme floods on YouTube: Application of LSPIV to home movies for the post-event determination of stream discharges. *Hydrological Processes*, *30*, 90–105. https://doi.org/10.1002/hyp.10532
Le Coz, J., Patalano, A., Collins, D, et al. (2016). Crowdsourced data for flood hydrology: Feedback from recent citizen science projects in Argentina, France and New Zealand. *Journal of Hydrology*, *541*, 766–777. https://doi.org/10.1016/j.jhydrol.2016.07.036
Liang, J., Liu, X., Huang, K, et al. (2014). Automatic registration of multisensor images using an integrated spatial and mutual information (SMI) metric. *IEEE Transactions on Geoscience and Remote Sensing*, *52*, 603–615. https://doi.org/10.1109/TGRS.2013.2242895
Li, M., Pang, B., He, Y., & Nian, F. (2013). Particle filter improved by genetic algorithm and particle swarm optimization algorithm. *Journal of Software*, *8*, 666–672. https://doi.org/10.4304/jsw.8.3.666-672
Liu, X., Chen, S., Zhuo, L, et al. (2018). Multi-sensor image registration by combining local self-similarity matching and mutual information. *Frontiers of Earth Science*, *12*, 779–790. https://doi.org/10.1007/s11707-018-0717-9
Mason, D. C., Schumann, G. J.-P., Neal, J. C, et al. (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. https://doi.org/10.1016/j.rse.2012.06.017
Mason, D. C., Trigg, M., Garcia-Pintado, J, et al. (2016). Improving the TanDEM-X Digital Elevation Model for flood modelling using flood extents from Synthetic Aperture Radar images. *Remote Sensing of Environment*, *173*, 15–28. https://doi.org/10.1016/j.rse.2015.11.018
Matgen, P., Hostache, R., Schumann, G, et al. (2011). Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies. *Physics and Chemistry of the Earth*, *36*, 241–252. https://doi.org/10.1016/j.pce.2010.12.009
Matgen, P., Montanari, M., Hostache, R, et al. (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*, 1773–1785. https://doi.org/10.5194/hess-14-1773-2010
Moradkhani, H., Hsu, K.-L., Gupta, H., & Sorooshian, S. (2005a). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resource Research*, *41*, 1–17. https://doi.org/10.1029/2004WR003604
Moradkhani, H., Nearing, G. S., Abbaszadeh, P., & Pathiraja, S. (2019). Fundamentals of data assimilation and theoretical advances. In Q. Duan, F. Pappenberger, A. Wood, H. Cloke, & J. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39925-1_30
Moradkhani, H., Sorooshian, S., Gupta, H. V., & Houser, P. R. (2005b). Dual state-parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, *28*, 135–147. https://doi.org/10.1016/j.advwatres.2004.09.002
Mukherjee, S. S., Joshi, P. K., Mukherjee, S. S, et al. (2013). Evaluation of vertical accuracy of open source Digital Elevation Model (DEM). *International Journal of Applied Earth Observation and Geoinformation*, *21*, 205–217. https://doi.org/10.1016/j.jag.2012.09.004
Muste, M., Ho, H. C., & Kim, D. (2011). Considerations on direct stream flow measurements using video imagery: Outlook and research needs. *Journal of Hydro-Environment Research*, *5*, 289–300. https://doi.org/10.1016/j.jher.2010.11.002
Nearing, G., Yatheendradas, S., Crow, W, et al. (2018). The efficiency of data assimilation. *Water Resource Research*, *54*, 6374–6392. https://doi.org/10.1029/2017WR020991
Nichols, N. K. (2010). Mathematical Concepts of Data Assimilation. In W. Lahoz (Ed.), *Data Assimilation* (pp. 13–39). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74703-1_2
Ossadtchi, A., Pronko, P., Baillet, S, et al. (2014). Mutual information spectrum for selection of event-related spatial components. Application to eloquent motor cortex mapping. *Frontiers in Neuroinformatics*, *7*, 1–11. https://doi.org/10.3389/fninf.2013.00053
Pappenberger, F., Frodsham, K., Beven, K. J, et al. (2007). Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. *Hydrology and Earth System Sciences*, *11*, 739–752. https://doi.org/10.5194/hess-11-739-2007
Patil, A., & Ramsankaran, R. (2017). Improving streamflow simulations and forecasting performance of SWAT model by assimilating remotely sensed soil moisture observations. *Journal of Hydrology*, *555*, 683–696. https://doi.org/10.1016/j.jhydrol.2017.10.058
Patil, A., & Ramsankaran, R. (2018). Improved streamflow simulations by coupling Soil Moisture Analytical Relationship in EnKF based hydrological data assimilation framework. *Advances in Water Resources*, *121*, 173–188. https://doi.org/10.1016/J.ADVWATRES.2018.08.010
Perdigão, R. A. P., & Blöschl, G. (2014). Spatiotemporal flood sensitivity to annual precipitation: Evidence for landscape-climate coevolution. *Water Resource Research*, *50*, 5492–5509. https://doi.org/10.1002/2014WR015365
Plaza Guingla, D. A., De Keyser, R., De Lannoy, G. J. M, et al. (2013). Improving particle filters in rainfall-runoff models: Application of the resample-move step and the ensemble Gaussian particle filter. *Water Resource Research*, *49*, 4005–4021. https://doi.org/10.1002/wrcr.20291
Plaza, D. A., De Keyser, R., De Lannoy, G. J. M, et al. (2012). The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter. *Hydrology and Earth System Sciences*, *16*, 375–390. https://doi.org/10.5194/hess-16-375-2012
Revilla-Romero, B., Hirpa, F. A., del Pozo, J. T, et al. (2015). On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions. *Remote Sens*, *7*, 15702–15728. https://doi.org/10.3390/rs71115702
Revilla-Romero, B., Wanders, N., Burek, P, et al. (2016). Integrating remotely sensed surface water extent into continental scale hydrology. *Journal of Hydrology*, *543*, 659–670. https://doi.org/10.1016/j.jhydrol.2016.10.041
Rogencamp, G. (2004). *Lower clarence river flood study review* (Technical report). WBM.
Schlaffer, S., Chini, M., Giustarini, L., & Matgen, P. (2017). Probabilistic mapping of flood-induced backscatter changes in SAR time series. *International Journal of Applied Earth Observation and Geoinformation*, *56*, 77–87. https://doi.org/10.1016/j.jag.2016.12.003
Schumann, G. J. P., Andreadis, K. M., & Bates, P. D. (2014). Downscaling coarse grid hydrodynamic model simulations over large domains. *Journal of Hydrology*, *508*, 289–298. https://doi.org/10.1016/j.jhydrol.2013.08.051
Schumann, G. J.-P., & Bates, P. D. (2018). The Need for a High-Accuracy, Open-Access Global DEM. *Frontiers of Earth Science*, *6*, 225. https://doi.org/10.3389/FEART.2018.00225
Schumann, G. J. P., & Domeneghetti, A. (2016). Exploiting the proliferation of current and future satellite observations of rivers. *Hydrological Processes*, *30*, 2891–2896. https://doi.org/10.1002/hyp.10825
Schumann, G., Matgen, P., Cutler, M. E. J. E. J, et al. (2008). Comparison of remotely sensed water stages from LiDAR, topographic contours and SRTM. *ISPRS Journal of Photogrammetry and Remote Sensing*, *63*, 283–296. https://doi.org/10.1016/j.isprsjprs.2007.09.004
Schumann, G. J. P., Stampoulis, D., Smith, A. M, et al. (2016). Rethinking flood hazard at the global scale. *Geophysical Research Letters*, *43*, 10249–10256. https://doi.org/10.1002/2016GL070260
Shannon, C. E., & Weaver, W. (1964). *The mathematical theory of communication* (2nd ed.). Urbana: University of Illinois Press.
Shastry, A., & Durand, M. (2019). Utilizing flood inundation observations to obtain floodplain topography in data-scarce regions. *Frontiers of Earth Science*, *6*, 1–10. https://doi.org/10.3389/feart.2018.00243

Sinclair Knight Merz F, & Roads and Traffic Authority of NSW TPS (2011). *Wells crossing to iluka road: Upgrading the pacific highway: Tyndale to MacLean alternative alignment: Decision report*. Sinclair Knight Merz F and Roads and Traffic Authority of NSW TPS.
Smith, A., Bates, P. D., Wing, O, et al. (2019). New estimates of flood exposure in developing countries using high-resolution population data. *Nature Communications*, *10*, 1–7. https://doi.org/10.1038/s41467-019-09282-y
Souza, B. C., Pavão, R., Belchior, H., & Tort, A. B. L. (2018). On Information Metrics for Spatial Coding. *Neuroscience*, *375*, 62–73. https://doi.org/10.1016/j.neuroscience.2018.01.066
Stephens, E., Schumann, G., & Bates, P. (2014). Problems with binary pattern measures for flood model evaluation. *Hydrological Processes*, *28*, 4928–4937. https://doi.org/10.1002/hyp.9979
Suri, S., & Reinartz, P. (2010). Mutual-information-based registration of TerraSAR-X and ikonos imagery in Urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, *48*, 939–949. https://doi.org/10.1109/TGRS.2009.2034842
Tauro, F., Petroselli, A., & Arcangeletti, E. (2016a). Assessment of drone-based surface flow observations. *Hydrological Processes*, *30*, 1114–1130. https://doi.org/10.1002/hyp.10698
Tauro, F., Porfiri, M., & Grimaldi, S. (2016b). Surface flow measurements from drones. *Journal of Hydrology*, *540*, 240–245. https://doi.org/10.1016/j.jhydrol.2016.06.012
Tauro, F., & Salvatori, S. (2017). Surface flows from images: Ten days of observations from the Tiber River gauge-cam station. *Hydrology Research*, *48*, 646–655. https://doi.org/10.2166/nh.2016.302
Trigg, M. A., Birch, C. E., Neal, J. C, et al. (2016). The credibility challenge for global fluvial flood risk analysis. *Environmental Research Letters*, *11*, 094014. https://doi.org/10.1088/1748-9326/11/9/094014
Tuozzolo, S., Lind, G., Overstreet, B, et al. (2019). Estimating river discharge with swath altimetry: A proof of concept using AirSWOT observations. *Geophysical Research Letters*, *46*, 1459–1466. https://doi.org/10.1029/2018GL080771
Uhe, P., Mitchell, D., Bates, P, et al. (2019). Enhanced flood risk with 1.5°C global warming in the Ganges-Brahmaputra-Meghna basin. *Environmental Research Letters*, *14*, 074031. https://doi.org/10.1088/1748-9326/ab10ee
van Leeuwen, P. (2017). Particle Filters for nonlinear data assimilation in high-dimensional systems. *Annales de la Faculté des Sciences de Toulouse*, *26*, 1051–1085. https://doi.org/10.5802/afst.1560
Van Wesemael, A., Landuyt, L., Lievens, H., & Verhoest, N. E. C. (2019). Improving flood inundation forecasts through the assimilation of in situ floodplain water level measurements based on alternative observation network configurations. *Advances in Water Resources*, *130*, 229–243. https://doi.org/10.1016/j.advwatres.2019.05.025
Walker, J. P., & Houser, P. R. (2005). Hydrologic data assimilation. In *Advances in water science methodologies*. (pp. 45–68). CRC Press
Waller, J. A., Dance, S. L., & Nichols, N. K. (2016). Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. *Quarterly Journal of the Royal Meteorological Society*, *142*, 418–431. https://doi.org/10.1002/qj.2661
Ward, P. J., Jongman, B., Salamon, P, et al. (2015). Usefulness and limitations of global flood risk models. *Nature Climate Change*, *5*, 712–715. https://doi.org/10.1038/nclimate2742
Wellmann, J. (2013). Information theory for correlation analysis and estimation of uncertainty reduction in maps and models. *Entropy*, *15*, 1464–1485. https://doi.org/10.3390/e15041464
Wood, M. (2016). *Improving hydraulic model parameterization using SAR data*. University of Bristol.
Woodhead, S. P. B. (2007). *Bayesian calibration of flood inundation simulators using an observation of flood extent*. University of Bristol.
Wood, M., Hostache, R., Neal, J, et al. (2016). Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium-resolution SAR data and identifiability techniques. *Hydrology and Earth System Sciences*, *20*, 4983–4997. https://doi.org/10.5194/hess-20-4983-2016
Woo, J., Stone, M., & Prince, J. L. (2015). Multimodal registration via mutual information incorporating geometric and spatial context. *IEEE Transactions on Image Processing*, *24*, 757–759. https://doi.org/10.1109/TIP.2014.2387019
Yamazaki, D., Ikeshima, D., Tawatari, R, et al. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, *44*, 5844–5853. https://doi.org/10.1002/2017GL072874
Yin, S., & Zhu, X. (2015). Intelligent Particle Filter and Its Application on Fault Detection of Nonlinear System. *IEEE Transactions on Industrial Electronics*, *62*, 1. https://doi.org/10.1109/TIE.2015.2399396
Zhang, D., Liu, M. qin, Zhang, S., et al. (2018). Mutual-information based weighted fusion for target tracking in underwater wireless sensor networks. *Front Information Technology Electron Engineering*, *19*, 544–556. https://doi.org/10.1631/FITEE.1601695
Zheng, X., Xiong, H., Yue, L., & Gong, J. (2016). An improved ANUDEM method combining topographic correction and DEM interpolation. *Geocarto International*, *31*, 492–505. https://doi.org/10.1080/10106049.2015.1059899
Ziliani, M. G., Ghostine, R., Ait-El-Fquih, B, et al. (2019). Enhanced flood forecasting through ensemble data assimilation and joint state-parameter estimation. *Journal of Hydrology*, *577*, 123924. https://doi.org/10.1016/j.jhydrol.2019.123924