

# Water Resources Research

## RESEARCH ARTICLE

10.1029/2020WR028238

### Key Points:

- The potential of targeted observations for optimal flood inundation forecast improvements from flood extent assimilation is demonstrated
- Impact of assimilating images in morphologically uniform reaches at different times, on channel, and floodplain water depth quantified
- Assimilating a single ideal flood extent could lead to persistent improvements comparable to assimilating multiple sub-optimal images

### Correspondence to:

A. Dasgupta,  
[antara.dasgupta@monash.edu](mailto: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). On the impacts of observation location, timing and frequency on flood extent assimilation performance. *Water Resources Research*, 57, e2020WR028238. <https://doi.org/10.1029/2020WR028238>

Received 25 JUN 2020  
Accepted 31 DEC 2020

## On the Impacts of Observation Location, Timing, and Frequency on Flood Extent Assimilation Performance

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

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

**Abstract** Flood inundation forecasts from hydrodynamic models can help with flood preparedness, but uncertainty in the inputs and parameters can lead to erroneous flood inundation estimates. However, Synthetic aperture radar (SAR)-based flood extent information can be used to constrain such model forecasts through data assimilation thus making them more accurate. Since high-resolution SAR satellites can only provide partial coverage for medium to large catchments, it is expedient to evaluate the combination of observation footprint, timing, and frequency which can lead to maximum forecast improvements. Consequently, multiple spatiotemporal SAR-based flood extent assimilation scenarios have been simulated here to identify the optimum observation design for improved flood inundation forecasts. A mutual information-based particle filter was implemented in a synthetic setup for the 2011 flood event in the Clarence Catchment, Australia, to combine SAR-based flood extents with the hydraulic model LISFLOOD-FP. The open loop ensemble was forced using uncertain inflows and the impact of assimilating flood extents in morphologically homogenous river reaches was evaluated for different first visit and revisit scenarios. Results revealed that the optimum temporal acquisition strategy strongly depends on reach morphology and flood wave arrival timing. Further, it was found that a single image at the right time could improve the 8-days forecast by ~95% when assimilated at reaches with large flat floodplains but limited tidal influence, while in reaches with narrow valleys over 10 images were needed to achieve the same outcome. Experiments such as the one presented here can therefore inform targeted observation strategies to ensure cost effective flood monitoring and maximize the forecast accuracy resulting from flood extent assimilation.

**Plain Language Summary** Satellite observations of flood inundation have the potential to increase the reliability and accuracy of flood forecasts, thereby contributing to improved flood resilience of vulnerable populations. However, new generation high-resolution satellites can only observe small portions of large river systems river during a flood. Since, the model-data integration methods used to combine flood forecasting models with satellite data are sensitive to the observation coverage, timing, and frequency best case scenarios can be constructed to obtain maximum improvements in accuracy. This study investigated the possibility of designing targeted observation strategies that can lead to more accurate flood forecasts after the model-data integration. Synthetic experiments were used to simulate multiple different image acquisition scenarios and assess the impacts on flood forecast accuracy. Results indicate that the location and timing of the images is more important than the revisit interval. Findings from this study can therefore be used to inform future satellite acquisitions, to ensure more cost effective flood monitoring from space leading to more reliable flood inundation forecasts.

### 1. Introduction

Hydrodynamic or hydraulic modeling is necessary to obtain accurate estimates of flood inundation and depth (Grimaldi et al., 2019). For hydraulic flood models running in forecast mode, the largest source of uncertainty is often the inflow error propagated from uncertain precipitation forecasts (Cloke & Pappenberger, 2009; G. J. P. Schumann et al., 2016), in addition to topographic errors from Global Digital Elevation Models (DEMs) (Kumar et al., 2019; Pilotti, 2015; Pramanik et al., 2010), and parametric errors due to

lumped or incorrect specifications (Annis et al., 2020). In order to reduce forecast uncertainties, flood models must be constrained using independent observations (Grimaldi et al., 2016). Synthetic aperture radar (SAR) sensors are ideal for this purpose, as their all-weather all-day imaging capabilities enable the spatiotemporal evolution of flood inundation to be observed (Dasgupta et al., 2018). SAR images provide information on flood extent, and when intersected with a DEM yields water level observations at select shoreline locations (Mason et al., 2012). These SAR-derived water levels can then be used to correct model predicted water levels in real-time using data assimilation, leading to more reliable flood forecasts as demonstrated by a number of studies (e.g., García-Pintado et al., 2015; Giustarini et al., 2011; Matgen et al., 2010; Neal et al., 2009). However, SAR-based water level estimation is neither straightforward nor automatic (Hostache et al., 2018) and was identified by G. Schumann et al. (2009) as an additional source of uncertainty. Consequently, a few studies have proposed the direct assimilation of SAR-derived flood extents into hydraulic models to reduce forecast uncertainty (e.g., Lai et al., 2014; Revilla-Romero et al., 2016).

Data assimilation allows for iterative system state estimation using a combination of instantaneous observations of the system and a continuous model of system dynamics to bridge the information gap during the measurement interval (Nichols, 2010). A number of different data assimilation methods have been used for the assimilation of remote sensing observations into hydraulic models. Of these the Ensemble Kalman Filter (EnKF) is the most popular as it allows the representation of model uncertainty with ensemble sizes significantly smaller than the state dimension (García-Pintado et al., 2013). However, since the EnKF requires an assumption that observation errors are Gaussian (Walker & Houser, 2005), the assimilation of SAR-based observations with known non-Gaussian errors (Xie et al., 2002) has focused attention on the non-parametric particle filter (PF). The adaptation of the PF for high dimensional applications has been questioned due to the possibility of “filter collapse,” where all ensemble members attain negligible weights while one attains a value of unity (García-Pintado et al., 2013). Indeed, the use of large ensemble sizes are theoretically mandated for PFs, with the ideal ensemble size being several orders of magnitude larger than the size of the state vector (Banister & Nichols, 2012; Nichols, 2010). In practice, however, the required ensemble size depends mainly on the model complexity, with hydrological model ensembles stabilizing with even 64 members (Reichle et al., 2002) and increases in ensemble size only leading to marginal improvements (Plaza Guingla et al., 2013). Furthermore, Dumedah and Coulibaly (2013) showed that PFs enable the retention of forecast accuracy improvements for longer lead times than the EnKF in hydrological data assimilation, which is important since the persistence of assimilation benefits is a known issue (K. M. Andreadis et al. 2007; Matgen et al., 2010).

The increasing number of spaceborne SAR sensors promise improved temporal coverage of flood events in the future. For example, SAR satellite constellations such as Cosmo-SkyMed (CSK) or the Radarsat Constellation mission offer 12- and 24-h revisit capabilities respectively, providing image sequences. For flood forecasting applications, the sensitivity of the relevant information content to first visit time and revisit time has been clearly highlighted by García-Pintado et al. (2013). However, the information content of SAR observations for flood applications is also a function of spatial coverage (Grimaldi et al., 2016). As satellites constantly evolve to provide increased spatial and temporal resolutions along with on-demand imaging capabilities, the design of targeted observation strategies could maximize forecast improvements (K. M. Andreadis 2018). According to Andreadis and Schumann (2014), the impact of assimilating water surface elevation, channel top width, and inundated area observations on forecast accuracy was keenly sensitive to observation spatiotemporal coverage. In fact, the study also highlighted that the assimilation could even degrade the forecast if observations were locally fitted to models in a highly erroneous sub-domain.

This study has therefore explored sampling design to maximize forecast improvements resulting from flood extent assimilation, by evaluating several realistic hypothetical observation footprint and temporal visit scenarios. Operational constraints related to data latency have not explicitly been considered here, since this problem is expected to further reduce in future satellites (McCabe et al., 2017). A real flood is simulated for this analysis with observed inflows serving as the baseline, with a controlled identical twin experiment used to assess the efficacy of each spatiotemporal scenario. First, the catchment was divided into three regions based on valley shape and catchment morphology, envisioned as the most important factors when assimilating flood extents (see the discussion in Wood et al., 2016 and Gobeyn et al., 2017 about the information content of flood extent for hydraulic model calibration). Single or multiple images (with revisit intervals of

12-, 24-, and 48-h) were then assimilated into any one of the three model sub-domains, and results evaluated in terms of absolute and relative errors. Impact on forecast accuracy for different lead times was evaluated by changing the time window used for calculation of the error metrics. Three brier skill score curves corresponding to each assumed observation footprint were generated for each temporal scenario and lead time combination. Ultimately, the location, timing, and frequency of observations for maximum possible improvement through flood extent assimilation is identified.

## 2. Methods

### 2.1. Experimental Design

Identical twin experiments were used with a hydraulic model simulation based on a real flood event. A uniform value of channel friction was assumed and fixed based on prior model calibration, with inflows considered as the primary source of errors to be corrected by the data assimilation in real-time. Measured inflows at the upstream boundary and calibrated friction parameters were used for the “truth” simulation, from which synthetic SAR-based probabilistic flood extents were generated. For the same event and simulation period, the observed inflow values were perturbed to generate an ensemble of erroneous inflows. Assuming the inflow measurements to represent the truth, a stationary mean error was first imposed as a positive multiplicative bias, followed by the addition of temporally correlated heteroscedastic errors. This inflow error generation method proposed by García-Pintado et al. (2013) is described in further detail later. The erroneous inflow ensemble was used to drive the open loop simulations (without data assimilation), and for all the assimilation simulation testing of different SAR visit scenarios.

The red boxes in Figure 2 (III) illustrate the sub-domains used for the assimilation, defined based on the river morphology, hereafter referred to as Sub 1, Sub 2 and Sub 3. Following the work of Domeneghetti et al. (2014) and Schumann and Andreadis (2016), three morphologically homogeneous sub-reaches were identified in the upper, central, and lower parts of the catchment. As the size of the floodplain and shape of the valley can strongly influence the inundation extent, it was reasonable to use these to define the different observation footprints considered for flood extent assimilation in this study. This visual interpretation was verified through the approach of Schumann and Andreadis (2016), where reach flow behavior was approximated using an assessment of the river longitudinal slope. For reaches with topographic controls where the flow behavior is typically kinematic, the longitudinal slope can be defined through a linear approximation. However, for reaches with more diffusive flows, the bathymetric longitudinal elevation distribution, shows a variation of  $>2\sigma$  around the fitted regression line. Based on this assessment, the sub-reaches relevant for flood extent assimilation were identified as follows.

Sub 1. The first sub-reach up to ~40 km from the upstream inflow boundary, characterized by a narrow valley (floodplain widths ranging from ~3 to 6 km) and high river longitudinal slope ( $>1\%$ ).

Sub 2. The second sub-reach from ~40 to ~90 km chainage, characterized by large flat floodplains (floodplain widths ranging from ~7 to 24 km) with gentle channel slopes ( $<1\%$ ), and limited tidal effects (see the small waves visible in the baseflow preceding the flood hydrograph at Rogan’s Bridge in Figure 2 (II)).

Sub 3. The third sub-reach from ~90 to ~150 km chainage, also had large flat floodplains (floodplain widths ranging from ~12 to 20 km) with gentle channel slopes ( $<1\%$ ), but with strong tidal backwater effects that control the inundation (see the tidal undulations visible in the receding limb of the flood hydrograph at Lawrence/Maclean, in Figure 2 (II)).

Note that this sub-reach classification is not sensitive to small shifts in chainage. This implies that altering the length of the reaches slightly is unlikely to influence the assimilation outcome and thereby the conclusions, as long as the dominant flow controls and floodplain characteristics are accurately characterized and are largely uniform across the reach. For each hypothetical observation footprint outlined here, several different temporal visit scenarios were experimentally implemented to answer the relevant research questions outlined in Table 1. Note that all references to flood extent assimilation in this study refer to observations covering only one of these sub-domains at any given time. Through this subdivision of the model into three sub-reaches and the evaluation of temporal visit scenarios, this paper identified the optimum footprint,

**Table 1**  
Summary of Different Spatial Scenarios Considered in This study

Observation location/ footprint	Sub 1 (upper)	Sub 2 (central)	Sub 3 (lower)
River Reach Morphology	Narrow valley and high river longitudinal slope	Large floodplain, gentle river longitudinal slope and very limited tidal influence	Flat and large, limited variation in flood extent due to strong tidal effects

Note. Every temporal scenario outlined in Table 2 was implemented for each spatial case listed here separately. Note that all references to flood extent assimilation in this study refer to observations covering only one of these sub-domains at any given time.

timing, and frequency of flood extent observations for optimum improvements from the flood extent assimilation, with the results generalized for river reaches of similar slopes and catchment morphologies.

An ensemble of 128 simulations was used in this study, representing a reasonable trade-off between computational expense and accuracy for particle filters (Hostache et al., 2018, 2015). Although it may be argued that this ensemble size is small given the number of model state variables, the practicalities of ensemble-based assimilation in operational flood forecasting need to be prioritized. Ideally the ensemble size should be several orders of magnitude larger than the state vector (Nichols, 2010) (~63,500 wet cells so at least >100,000 ensemble members), but this would be impractical for operational forecasting given the computational demands of hydraulic models. Moreover, Plaza Guingla et al. (2013) experimentally demonstrated through assessments of five different particle sets (32, 64, 128, 256, and 528) that increasing the ensemble size beyond 128 members, yields only marginal improvements which do not justify the increased computational load (see Table 4 of Plaza Guingla et al. (2013)). Therefore, 128 particles were considered sufficient for this study along with stringent ensemble spread evaluations used to ensure that the observations were captured.

The ensemble spread and mean forecast error are crucial to ensure observations are encompassed suitably (Moradkhani, Sorooshian, et al., 2005). If the observations are statistically similar to the  $N$  ensemble members over the entire forecast of length  $T$ , then the ratio ( $Ra = Rt/Rn$ ) between the RMSE of the ensemble mean  $Rt$ , calculated as

$$Rt = \frac{1}{T} \sum_{t=1}^T \sqrt{\left[ \left( \frac{1}{N} \sum_{n=1}^N \hat{y}_t^n \right) - y_t \right]^2}, \quad (1)$$

where  $y_t$  is the observed value of the state for a given ensemble member at a particular time step and  $\hat{y}_t^n$  is the model predicted value. The mean RMSE of all the ensemble members  $Rn$  is then computed according to

$$Rn = \frac{1}{N} \sum_{n=1}^N \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{y}_t^n - y_t^n)^2}, \quad (2)$$

and should be equal to the expected value of  $Ra$ , given by

$$E[Ra] = \sqrt{\frac{N+1}{2N}}. \quad (3)$$

The forecast skill can then be evaluated using the Normalized RMSE Ratio (NRR) given by

$$NRR = \frac{Ra}{E[Ra]}. \quad (4)$$

Values of  $NRR > 1$  indicate too little ensemble spread, while  $NRR < 1$  implies too much ensemble spread, with ideal ensembles producing NRR values close to unity (Moradkhani, Sorooshian, et al., 2005). An NRR value of ~0.96 was obtained for the forecast inflow ensemble in this paper which is sufficiently close to the ideal value of 1 indicating adequate ensemble spread.

## 2.2. Synthetic Satellite Observations

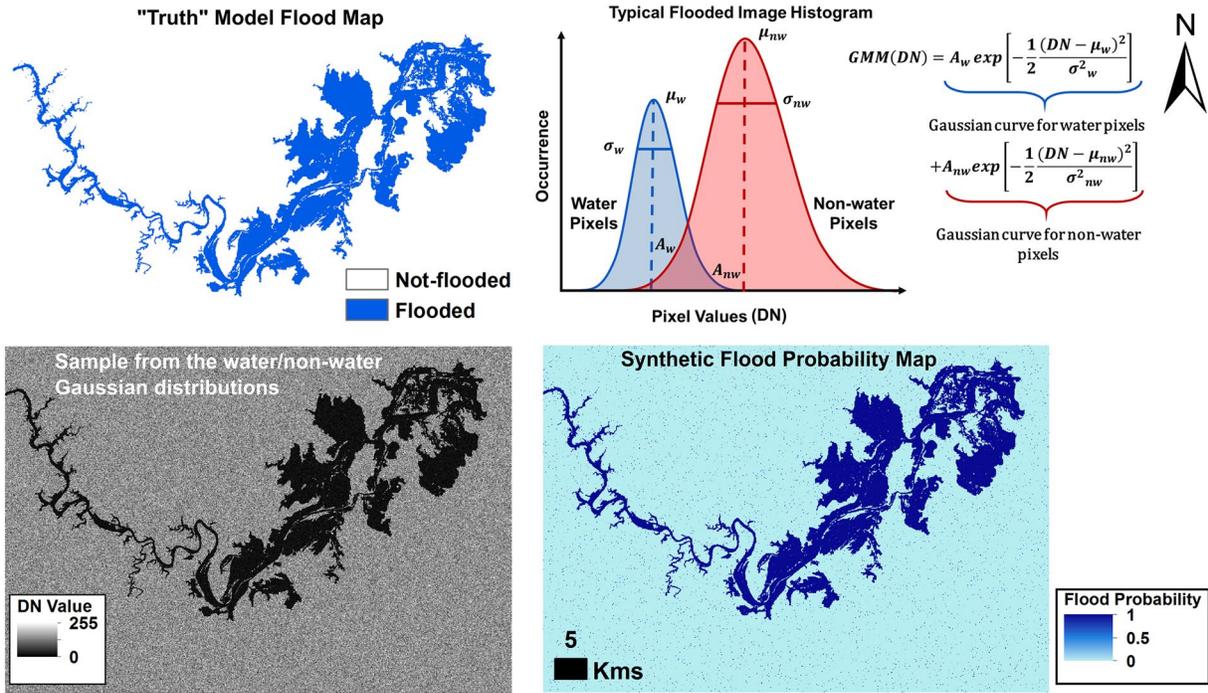
As an accurate assessment of observation error is crucial for data assimilation applications, objectively representing uncertainties in satellite observations is crucial. Therefore, the synthetic SAR observations generated in this paper were converted into probabilistic flood extents prior to the assimilation. The gridded water depth values simulated by the truth model were first extracted at the relevant assimilation time steps. These were subsequently converted into binary flood extents using a minimum water depth threshold of 1 cm, chosen to maximize extent variability between time steps. The truth model simulated binary flood extents were then processed into synthetic SAR images using the approach proposed by Cooper et al. (2019); an example of this process is illustrated in Figure 1. For each pixel in the binary “truth” model simulated flood maps, a backscatter value was sampled from the distributions of flood and non-flood classes obtained from typical flooded image histograms. These “typical” histogram statistics were computed based on two CSK images covering the real-world flood event presented in this study. Parameterizing a Gaussian Mixture Model (GMM) using the CSK image histograms, the means, standard deviations, and mixing proportions were obtained. Based on the simulated wet/dry cell status, a value was sampled from the corresponding water and non-water Gaussian distributions to generate the synthetic SAR images. Following this, the synthetic SAR images were translated into probabilistic flood maps using the backscatter-based Bayesian probabilistic mapping approach proposed by Giustarini et al. (2016). The probabilistic flood extent observations were then subset according to the spatial subsets illustrated in Figure 2 and used to simulate the various assimilation scenarios described in Section 2.1. Readers are referred to Dasgupta et al. (2021) for more details on the GMM parameterization and to Giustarini et al. (2016) for the probabilistic mapping approach.

## 2.3. Data Assimilation Framework

Particle filters use Sequential Monte Carlo sampling methods to enable the non-parametric approximate representation of continuous prior and posterior probability distribution functions (pdf) and are thus able to relax the common assumption of Gaussianity in the observation and model errors (Arulampalam et al., 2002). PFs are therefore well suited to the problem of SAR-based flood extent assimilation (Hostache et al., 2018), where flooded SAR images are known to display non-Gaussian errors (Xie et al., 2002). In particular, the sequential importance sampling (SIS) used here allows observations to be incorporated into hydraulic model forecasting chains, without disturbing the delicate system dynamics. When hydraulic models are stopped mid simulation, momentum is typically not conserved and the fluxes in the domain drop to zero, resulting in spurious inundation outcomes (E. S. Cooper et al. 2018). This issue can be partially solved by applying pre-assimilation velocities, although this approach has only been demonstrated for an idealized test case so far. In the SIS implementation, however, the state variables are not updated, but rather the ensemble mean is updated to assimilate all the observations. In other words, each model ensemble member was assigned a probabilistic weight in comparison to the observation and these weights were then used to compute the weighted mean which is representative of the assimilated state vector. While this approach has its own limitations, such as the frequently highlighted problem of particle degeneracy, these can be dealt with using pragmatic mathematical solutions (Hostache et al., 2018). However, restarting the model after fluxes artificially drop mid-simulation may result in hydraulic shock waves across the domain in addition to numerical instabilities, rendering the forecast unrealistic and useless.

In the PF-SIS, each individual model ensemble member is known as a particle and associated with a unique set of associated inputs and parameters. In this study, all the particles shared the same model parameters and geometry, with the only difference between them being the unique erroneous input inflow. The particle filter seeks to find the best estimate of the system states  $x \in \mathcal{R}_x^N$ , given uncertain knowledge of observations  $y \in \mathcal{R}_y^N$  and the system dynamics (Arulampalam et al., 2002). Starting from a uniform prior pdf when no information is available, the particle weights are updated in real-time as observations are sequentially assimilated (Moradkhani, Hsu, et al., 2005). The prior pdf of the state is given by  $p(x)$  while observing the true state of the system gives the measurements according to

$$y = h(x_{\text{true}}) + \epsilon, \quad (5)$$



**Figure 1.** An illustration of the conversion of a binary flood extent simulated by the “truth” model, into first a synthetic SAR image, and then finally, into a synthetic flood probability map using the approach of Giustarini et al. (2016). Note that the mean ( $\mu$ ), standard deviation ( $\sigma$ ), and mixing proportions ( $A$ ) of the water and non-water pixel populations (denoted by sub-scripts  $w$  and  $nw$ ), in the “Typical Flooded Image Histogram” were obtained from Cosmo-Skymed SAR images.

where  $h(\cdot)$  maps the model state  $x$  to the observation space as  $h(x) = 1$  where  $x > \text{threshold}$  and 0 otherwise, with  $\epsilon$  representing the measurement error. Discretizing the prior state pdf into  $N$  particles according to

$$p(x) = \sum_{i=1}^N \frac{1}{N} \delta_{x_i}. \quad (6)$$

Allows updating the state every time a new observation is available. At each assimilation time step, the Bayesian conditional probability of each particle given the observation is evaluated by

$$p(x | y) = \frac{p(y | x)p(x)}{\int p(y | x)p(x)dx}, \quad (7)$$

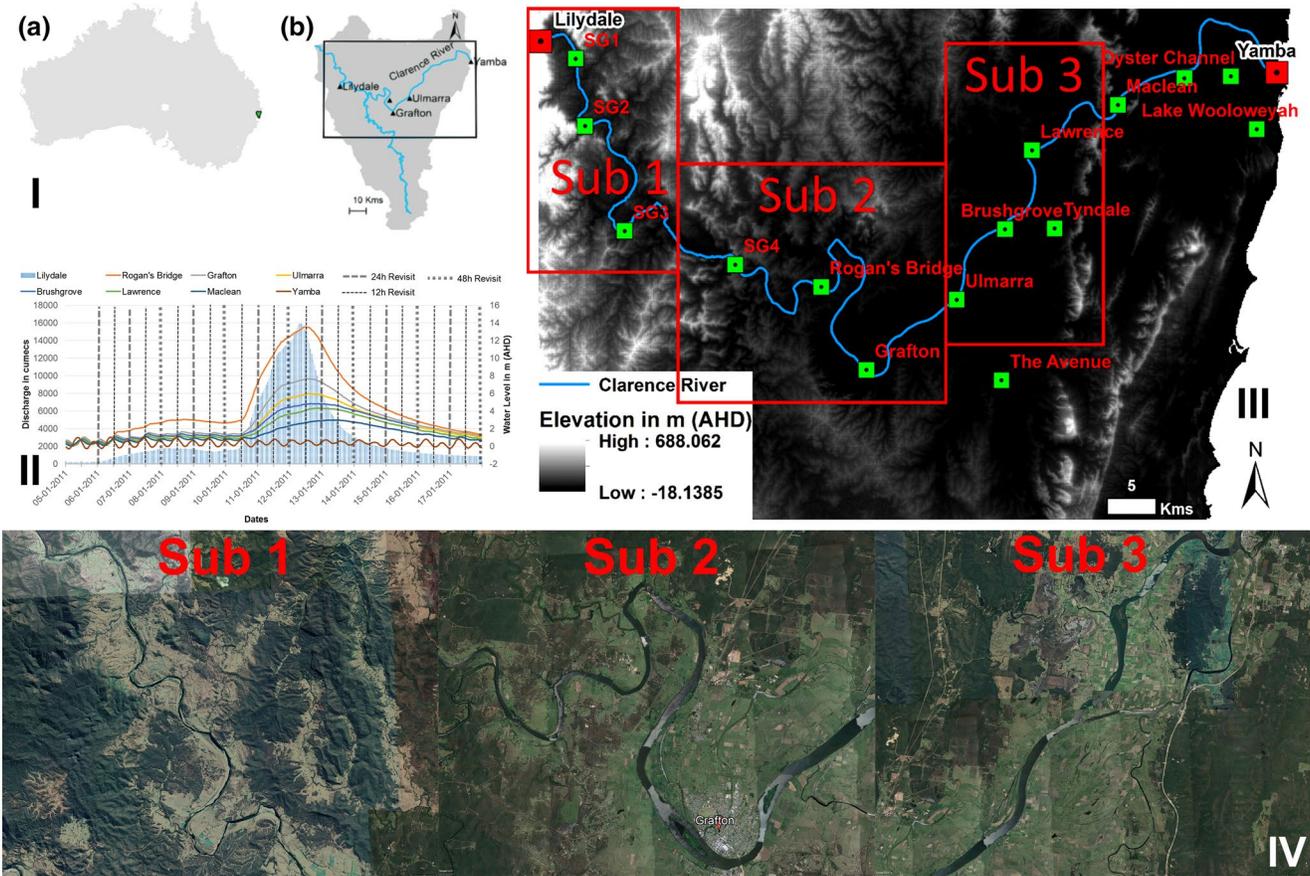
and used as the particle weight (Kantas et al., 2015) which is used to estimate the unknown posterior pdf according to

$$p(x | y) = \sum_{i=1}^N w_i \delta_{x_i}, \quad (8)$$

The model simulated state estimates are evaluated against observations at each assimilation time step using the function  $L_i$ ,

$$L_i = \frac{p(y | x_i)}{\sum_{j=1}^N p(y | x_j)}, \quad (9)$$

with the density  $p(y|x_i)$  being the probability of the observations given the model or the likelihood. The weighting function assigns higher weights to models which better reproduce the observations, using the mutual information based likelihood proposed by Dasgupta et al. (2021), which quantifies the reduction



**Figure 2.** Geographical location of the Clarence Catchment, in Australia shown in I(a), with important towns marked with respect to the Clarence River system in I(b). The observed hydrographs and corresponding temporal image acquisition scenarios are shown in II. The extent of the model domain from Lilydale to Yamba is shown in III, along with the subsets identified for the assimilation based on morphological features and gauge locations used for evaluation. Sub-figure IV shows Google Earth photos highlighting the hydraulic complexities of the various sub-reaches.

in uncertainty about the model given the observation. The likelihoods are then rescaled and inflated to enhance posterior variance, and weights computed using

$$w_i = \frac{(L_i - \min L_i)}{(\max L_i - \min L_i)} \times \left( \frac{L_i}{\max L_i} \right)^\vartheta, \quad (10)$$

where  $\vartheta$  is a rescaling factor which is internally optimized based on site-specific ensemble and observation properties. The weights are finally normalized according to

$$W_i = \frac{w_i}{\sum_{i=1}^N w_i}, \quad (11)$$

where  $W_i$  denotes the global weight assigned to a particle for a given observation, to ensure that the posterior pdf sums to unity. As new observations become available, particle weights evolve to incorporate the new information and cumulative weights are computed. Here the weights are multiplied forward, implying that the assimilated forecast including information from both the model and the observation, forms the prior (i.e.  $p(x)$  in Equation 7) for each subsequent observation being the previous best state estimate. More simply, this means that the information from past observations is carried forward in time. When the observations are too far apart in time, the system may become independent of the previous information and a uniform

prior can be used again in the absence of any additional information. Combining the weights through forward multiplication beyond the decorrelation length may negatively impact the forecast (Wood, 2016). This “decorrelation length” is also investigated in this study (Section 4.2.2), as the number of images which yield (i) positive improvements and (ii) the maximum improvements, when the assimilated forecast is used as the prior in Equation 7. Note that the “decorrelation length” or the time window until which the forecast is correlated to the instantaneous observational information, is not used in this paper but rather evaluated experimentally, to provide a framework for the future development of temporal weight decay algorithms.

In the SIS algorithm used in this study, the updated weights are used to compute the expectation of all state variables as the weighted ensemble mean given by,

$$E[X] = \sum_{i=1}^N W_i \times x_i. \quad (12)$$

The expectation therefore represents the best state estimates using the available information in the uncertain models and observations (Plaza et al., 2012). Moreover, the method proposed here based on the PF-SIS assimilation scheme is particularly suitable for the proposed assessment, as it allows to update the state predictions everywhere even if only a subset of the model domain is observed. The experiments presented herein can also help to evaluate the effects of global updating based on local observations offering only partial coverage of the domain. Interested readers are referred to Dasgupta et al. (2021) for a more in depth discussion of the assimilation strategy used here.

#### 2.4. Ensemble Generation

Quality of ensemble forecasts is strongly dependent on the choice of the ensemble generation method, the forecast model, and the analysis scheme (García-Pintado et al., 2013). In this study, a major structural assumption is that all model errors arise from the upstream inflow boundary data, while the model structure and parameterization is assumed to be perfect. Moreover, there could also be errors due to unspecified or erroneous lateral flows into the domain, which are not accounted for in this study, since upstream flow measurements were considered to be the sole source of forecast uncertainty. This assumption is unrealistic for real-world applications, since the DEM/river geometry, friction parameters and for some reaches, the lateral inflows heavily contribute to the forecast uncertainties. However, assuming a single uncertainty source is fairly standard in synthetic data assimilation studies (e.g., García-Pintado et al., 2015, 2013; Hostache et al., 2018). Not only does this facilitate an assessment of forecast uncertainties arising from a particular source of error, but it also helps to better understand assimilation performance independent of complex uncertainty interactions, which is especially critical for the development of new assimilation techniques. In such cases though, the perturbation model chosen is a key determinant of the assimilation outcomes, and therefore must be as realistic as possible to ensure transferability of results to real-world case studies.

Observed inflows are known to exhibit measurement errors due to the instrumentation, but for high-flows typical of floods the flow values are often derived from extrapolated rating curves which additionally contribute to the uncertainty (see Di Baldassarre & Montanari, 2009 for an extended discussion). In assimilation studies, authors have chosen to perturb the input forcings of hydrological models e.g. precipitation fields (K. M. Andreadis et al. 2007; Hostache et al., 2018) or the forcings, parameters, and initial conditions (Matgen et al., 2010), to obtain an ensemble of inflows for the hydraulic modeling. While some studies have argued that using model cascades result in more “realistic” forecast inflow ensembles, this also means that the error characteristics of the resulting inflows are neither controlled nor explicitly understood. As García-Pintado et al. (2013) pointed out, it is vital to have clarity on the error characteristics of the inflow perturbations for studies focused on hydraulic data assimilation, since the main objective is to efficiently mitigate them.

Essentially, the choice of the inflow error generation model depends on the nature of the uncertainty targeted. For instance, for hindcasting or calibration studies it makes sense to emulate the measurement uncertainties discussed earlier. However, for forecasting applications simulating uncertainties typically observed in hydrological model forecasts is more suitable. The errors are expected to display some degree of temporal autocorrelation in both cases, with hydrological model forecasts exhibiting higher autocorrelation in time than measurements. Some degree of spatial correlation is also expected, although since the domain used

in this study has only one inflow point, this discussion is presently out of scope. Since, the flood extent assimilation scenarios emulated in this study are positioned specifically to improve operational inundation forecasts, temporally autocorrelated and heteroscedastic random errors characteristic of hydrological model outputs were used to perturb the measured inflows (as in García-Pintado et al., 2013). The main advantage of adopting this approach was that the impacts of selecting a “specific” hydrological model on the error characteristics is negated, leading to more generic and controlled error representation.

The temporal error characteristics of hydrological ensemble forecasts for real-world applications are typically unknown. Given the best available knowledge of discharge errors, a deterministic error was imposed here as a multiplicative bias (20% as recommended by García-Pintado et al., 2013) along with temporally correlated random errors to obtain the input inflow ensemble. The heteroscedasticity was ensured by scaling the inflow errors according to the coefficient of variation observed from historical rating curves. Even though the mean error dynamics are emulated through a deterministic stationary bias in this case, a temporal correlation factor is introduced which decays to zero between assimilation time steps. This allows for a more realistic evaluation of the assimilation performance, as the real-world inflow error evolution is somewhat captured. Here a temporal correlation of 3 days was assumed following García-Pintado et al. (2013), who argue that while arbitrary this value is representative of typical distributed hydrological models. It is noted that longer temporal correlation windows, such as the one used herein, will inevitably lead to more persistent bias correction, especially due to the stationary mean bias. However, García-Pintado et al. (2013) postulate that this “intentional mismatch between the error forecast model and the stationary bias serves to emulate the lack of knowledge of the mean error evolution in real cases.” The inflow ensemble generation parameters used here are identical to García-Pintado et al. (2013), except for the coefficient of variation (cv) which was estimated from the available gauges (cv = 0.25).

### 2.5. Performance Metrics

The performance metrics used in this study reflect model accuracy in simulating channel and floodplain inundation dynamics. Impacts on the gauge water levels were evaluated with respect to the truth model at point locations using the Brier Skill Scores while floodplain water levels were assessed within the assimilation sub-domain for the local case and over the entire model domain in the global case using the root mean squared errors (RMSE). First, the RMSE was chosen for a lumped evaluation of absolute model errors (E. S. Cooper et al. 2018; García-Pintado et al., 2013, 2015), between the weighted ensemble mean  $E[WD_i]$  and the truth simulation  $WD_i^{\text{truth}}$  according to

$$WD_{\text{RMSE}} = \frac{1}{T} \sum_{i=1}^T \sqrt{\frac{\sum_{i=1}^P (E[WD_i] - WD_i^{\text{truth}})^2}{P}}, \quad (13)$$

where  $P$  refers to the total number of model grid cells and  $i$  refers to a particular cell number.  $WD_{\text{RMSE}}$  was computed over the specified time windows of length  $T$  (see Table 2) using hourly data and was used to quantify algorithm sensitivity to spatial coverage locally (within the assimilation sub-domain) and globally (across the entire model domain).

Synthetic gauges were considered at a few locations in addition to the real gauge locations, to ensure an equitable distribution of gauges in all sub-domains. Impacts on channel performance were evaluated through Brier Skill Scores (BSS), which quantify the relative improvement in the skill of the assimilated forecast with respect to the open loop. BSS was calculated according to

$$BSS = 1 - \frac{\overline{(\text{Assim.} - \text{Truth})(\text{Assim.} - \text{Truth})^T}}{\overline{(\text{OL} - \text{Truth})(\text{OL} - \text{Truth})^T}}, \quad (14)$$

where the variables in bold denote state vectors while the overline denotes a time average. Values of  $BSS \in (-\infty, 1]$ , where  $BSS = 0$  indicates no change in forecast skill while  $BSS = 1$  is the ideal score indicating

**Table 2**  
Summary of Temporal Scenarios Implemented for Each Spatial Scenario Outlined in Table 1

Assimilation case	Observation characteristics			Assessment area		Forecast lead time		Research question
	Footprint	Acquisition/ first visit time	Revisit interval	Channel	Channel & floodplain	Full forecast (days)	Revisit interval (hours)	
Single image (Section 4.1)	All subs Table 1	Every 12 h	No revisit	Point (BSS)	Local RMSE Global RMSE	1–16 (Section 4.1.1)	12 (Section 4.1.2)	When and where is a single observation most useful for the assimilation?
Multiple images (Section 4.2)		Every 12/24/48 h	12/24/48 h	Point (BSS)	Local RMSE Global RMSE	1–16 (Section 4.2.1)	12/24/48 (Section 4.2.2)	How do the observation footprint, timing, and frequency impact assimilation efficiency?
Maximum improvement (Section 4.3)			Single or 12/24/48 h	Point (BSS)	–	–	12/24/48 (Table 3)	How much can the forecast be improved with optimal observation characteristics?

Local RMSE = RMSE calculated within the assimilation sub-domain; Global RMSE = RMSE calculated over the entire model domain; Assimilation efficiency = Forecast improvements resulting from the assimilation

BSS, Brier Skill Scores.

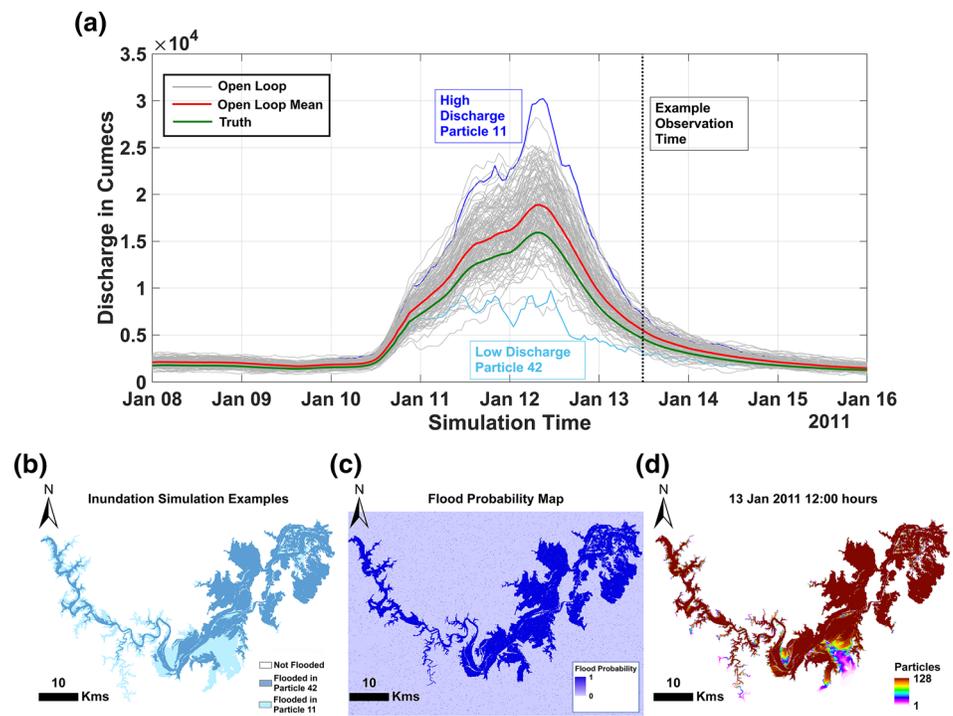
no error remains after the assimilation. BSS is therefore a relevant proxy for characterizing assimilation framework efficiency.

### 3. Hydraulic Model, Domain, and Data

#### 3.1. The Forward Truth Hydraulic Model

The two-dimensional grid-based hydraulic model LISFLOOD-FP (Bates et al., 2010) was chosen as the grid-based domain discretization allows easy integration with remotely sensed datasets (P. D. Bates & De Roo 2000), and used for all the experiments described here. The inertial formulation implemented in LISFLOOD-FP is computationally more efficient than other diffusive models with similar levels of accuracy (Fewtrell et al., 2011; Horritt & Bates, 2001). Flows through each cell face are computed using an inertial approximation of the Saint Venant equations where only convective acceleration is ignored (Bates et al., 2010). A finite difference numerical scheme is used for increased stability, explicit in time and first order in space with a semi implicit treatment of the friction term (Bates et al., 2010). Flows between cells are computed as a function of local water acceleration, and of the friction and water 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 LISFLOOD-FP model was implemented here in full two-dimensional using a 90 m pixel size as the Clarence River is wider than 100 m at most locations and this allowed an acceptable trade-off between model resolution, accuracy, and computational time. At Lilydale, the upstream boundary condition was specified based on the observed discharge hydrograph for the 2011 flood event in the Clarence Catchment, while gauged tidal levels were used as the downstream boundary at Yamba (see Figure 1 for locations). Distributed floodplain friction coefficients were assigned based on land-use identified from field and aerial photographs, while a lumped value for the channel friction specified as Manning's  $n$  was calibrated using high water marks (HWMs) derived from crowdsourced photos. A Manning's  $n$  value of  $0.026 \text{ m}^{1/3} \text{ s}^{-1}$  led to



**Figure 3.** An illustration of the flood inundation extents on January 13, 2011 12:00 h, with the open loop inflow ensemble shown in (a), example extents for high and low discharge shown in (b), the corresponding flood probability map presented in (c), and the number of particles which classify a given pixel as inundated at the given timestep are illustrated in (d).

the best mean RMSE value of 9 cm with respect to the HWMs for 2013 flood event in the Clarence Catchment and was thus used as the channel friction in this study. Terrain elevations were supplied using the LiDAR DEM available to this study, upscaled to a 90 m grid resolution with integrated observed channel bathymetry, which ensured that the channel conveyance is adequately represented. Even though this factor has limited importance for synthetic twin experiments, a poorly resolved channel might alter the system wave propagation dynamics by forcing water onto the floodplain at low discharge, which in turn would impact the channel overtopping time. Readers are referred to Grimaldi et al. (2018) for a detailed description of the bathymetric data set. Initial conditions were computed by running a steady state simulation with the same model implementation and an input inflow of the most commonly observed flow value at Lilydale. A warm-up period of five weeks was deemed necessary to ensure numerical stability and compensate for uncertain initial conditions, as described in Dasgupta et al. (2021) where the model implementation was first presented.

### 3.2. Study Area and Model Domain

The Clarence Catchment is situated in New South Wales, Australia, as shown in Figure 2, and drains a total of 22,700 sq. kms. The Clarence River is 394 km long of which a reach of 160 km, between Lilydale and Yamba, is covered by the hydraulic model domain. This catchment most recently experienced severe flooding in 2020, with the highest on record flood in 2013 which reached water levels of 8.09 m Australian Height Datum at the Prince Street Gauge, Grafton (Huxley & Beaman, 2014). Floods in this catchment move quickly taking about 30 h to peak on average, generating a flashy catchment response (Rogencamp, 2004). Low-intensity, long duration rainfall events are the dominant cause of flooding in the area. The back propagation of ocean storm tides also strongly influence inundation dynamics at the downstream gauges around Lawrence, and the impacts are visible as far upstream as Rogan's Bridge (see Figure 2 for locations). In fact, the tidal oscillations are evident from the observed hydrographs, as well as the simulated hydrographs

where tidal levels are used as the downstream boundary condition at Yamba. Indeed, the magnitude of the tidal influence is visible from the observed hydrographs in Figure 3 (II), where the gauges arranged upstream to downstream show progressively increasing discharge/water level oscillations. 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).

### 3.3. Datasets

The setup of a hydrodynamic model typically requires the specification of inflow discharge at the upstream boundary, catchment topography, channel bathymetry, floodplain surface roughness, and a downstream boundary condition. In this study, synthetic hydrographs were generated for the open loop runs by perturbing the observed inflows at Lilydale for the 2011 flood event in the Clarence River using the error model described by García-Pintado et al. (2013). The truth model was based on the observed inflows and tidal levels at the downstream boundary. Topographic information for the truth and open loop runs was identical and specified using a 1 m Light Detection and Ranging (LiDAR) Digital Elevation Model (DEM) with a vertical accuracy of  $\pm 30$  cm and horizontal accuracy of  $\pm 80$  cm resampled to 30 m (New South Wales Land and Property Management Authority, 2010; Figure 2). The channel bathymetry was reconstructed by interpolating field-surveyed cross sections at strategic locations and stitched to the LiDAR DEM, for the part of the domain where it was available. The area upstream of Copmanhurst where LiDAR coverage was unavailable, was filled with the 1 Second SRTM Derived Hydrological Digital Elevation Model (DEM-H) product available from Geoscience Australia (Gallant et al., 2011). The Lidar DEM was first resampled to the same grid as the DEM-H using nearest neighbor, following which the DEMs were merged ensuring identical reference systems and datums. For a detailed description of this topographic data set and the relative coverage for LiDAR versus the HDEM, readers are referred to (<https://elevation.fsd.org.au/>) where the datasets are also freely downloadable. Details of the bathymetry data collection and interpolation are provided in Grimaldi et al. (2018), while the bathymetric data set is available at [https://figshare.com/articles/Bathymetric\\_survey\\_of\\_the\\_Upper\\_Clarence/5648251](https://figshare.com/articles/Bathymetric_survey_of_the_Upper_Clarence/5648251) (Grimaldi et al., 2017). In a real world application, this DEM merging would have notable impacts on the assimilation given the stark difference in the uncertainty characteristics. However, it is expected to have negligible impacts in this study, due to the use of a synthetic twin experimental setup where the topography is identical for the open loop and assimilated forecasts.

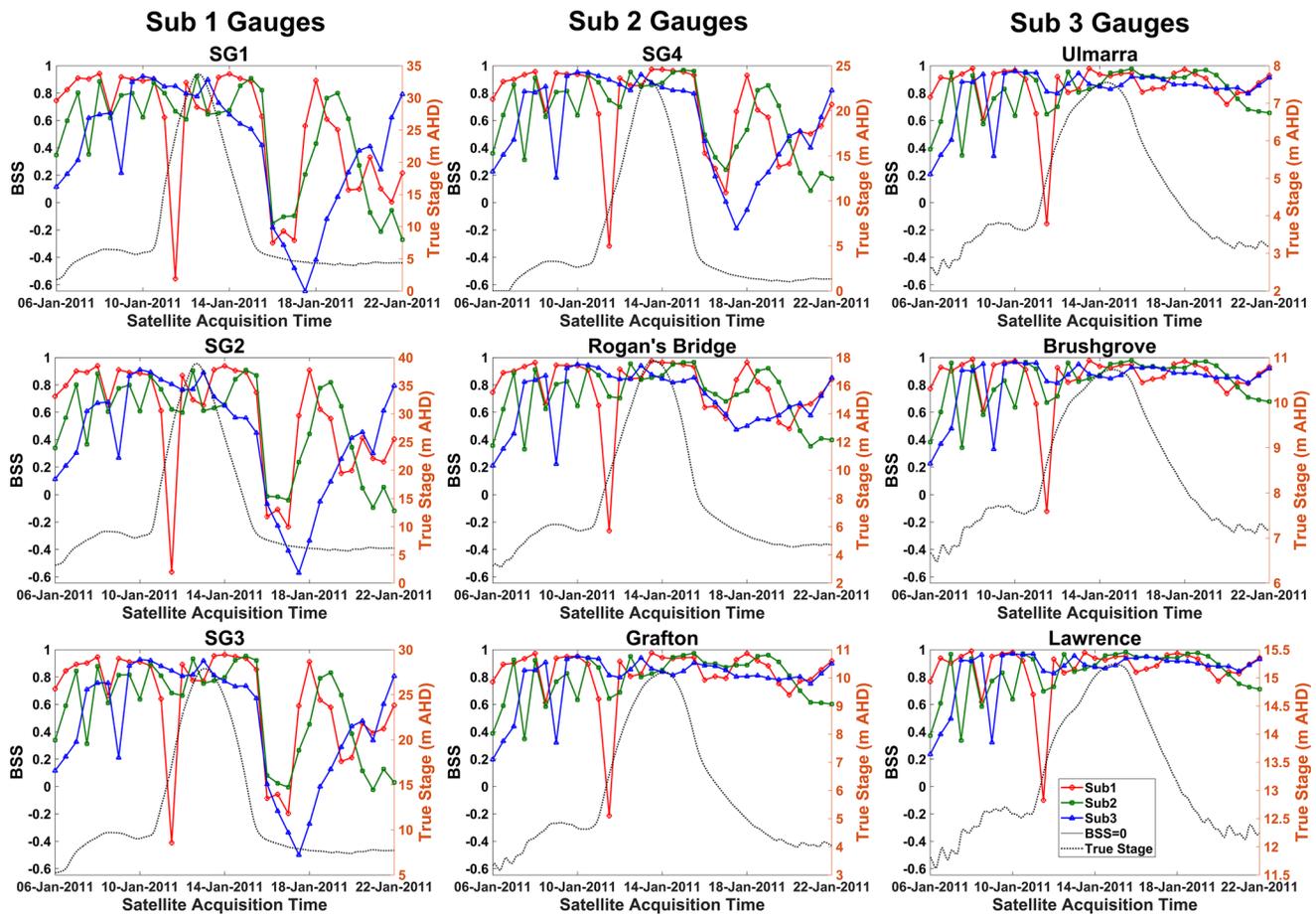
## 4. Results and Discussion

### 4.1. Single Image Assimilation

#### 4.1.1. Impact of Observation Footprint and Timing on Long Lead Times

A single image was assimilated into the hydraulic model at different locations and acquisition times (considered at every 12 h from the start of the event), and the impact on the channel water level forecasts across the event was evaluated. This allowed the persistence of assimilation benefits for long lead times to be assessed. As noted in Table 2, the assessment period started from the image acquisition time and extended until the end of the forecast. This implies that for the first image timing considered (i.e. January 6, 2011) the forecast lead time was of 16 days, while for the final acquisition evaluated (i.e. 22nd January) the forecast lead time was 1 day. The difference forecast lead times was compensated by using a relative metric for the gauge evaluation, i.e. the BSS, which implicitly considers the lead time since it is essentially the ratio of errors in the assimilated and open loop forecasts for the same assessment period. For the absolute error metric, i.e. the  $WD_{RMSE}$  used to evaluate the spatiotemporal impact on water depth simulations, this was dealt with by additionally providing the open loop response. This facilitated point-wise comparisons between the open loop and assimilated forecast  $WD_{RMSE}$  at each acquisition time for the same assessment periods. Moreover, all comments of forecast improvements were made relative to the open loop errors, even for the absolute error quantified by the  $WD_{RMSE}$ , thereby, ensuring that comparative statements could be made.

Figure 3 shows the variation in extents across the particles in different parts of the catchment and the BSS values calculated based on the water level hydrographs for all single image assimilation time steps are shown in Figure 4. BSS values were also computed for channel discharge but have been omitted for brevity due to identical trends. Figure 4 clearly shows that the flood extent assimilation was largely able to signif-



**Figure 4.** Brier Skill Scores (BSS) obtained for single image assimilation in each sub-reach, from the time of the satellite overpass to the end of the forecast. Observations were independently considered each 12 h starting from the 6<sup>th</sup> of Jan with BSS calculated at nine water level gauges along the channel (three in each sub-domain); the true stage at the location is shown in all subplots as a reference. Each point on each curve is representative of a satellite acquisition time and the corresponding BSS obtained from the time of the satellite overpass to the end of the forecast.

icantly improve forecast quality, with some images leading to even a 100% reduction in open loop errors. The ideal image acquisition timing to enable the highest improvement differs for the three sub-reaches, indicating sensitivity of the assimilation algorithm to observation footprint. Assimilating images in Sub 2 and Sub 3 indeed resulted in more consistent improvements of higher magnitude as compared to Sub 1, suggesting that reaches with larger and flatter floodplains are better candidates for flood extent assimilation. This finding can be attributed to the extent variability shown in Figure 3, which essentially implies that the reaches exhibiting more variability among model simulated extents, are more likely to allow for the effective ranking of particles.

In Sub 1 (the red line—Figure 4), the assimilation produced extremely inconsistent improvements due to a narrow and constrained valley, where extent sensitivity to changes in the simulated water depth varied notably during the event (see Figure 3d). Consequently, the assimilation in Sub 1 results in an almost uniform posterior distribution, when the observation becomes uninformative about particle likelihoods. The more concerning factor is that forcing the models to be ranked according the extents observed in Sub 1, could potentially degrade the model as shown by the negative BSS values obtained. When assimilating an image in Sub 1, the impacts of the assimilation were propagated downstream through the domain by the numerical model, leading to higher BSS values for the downstream gauges. In fact, the images (16/17 January) which degrade the forecast at the local gauges, actually result in >80% improvements at the Sub 3 gauges (notice the changing shape of the red curve upstream to downstream—Figure 4).

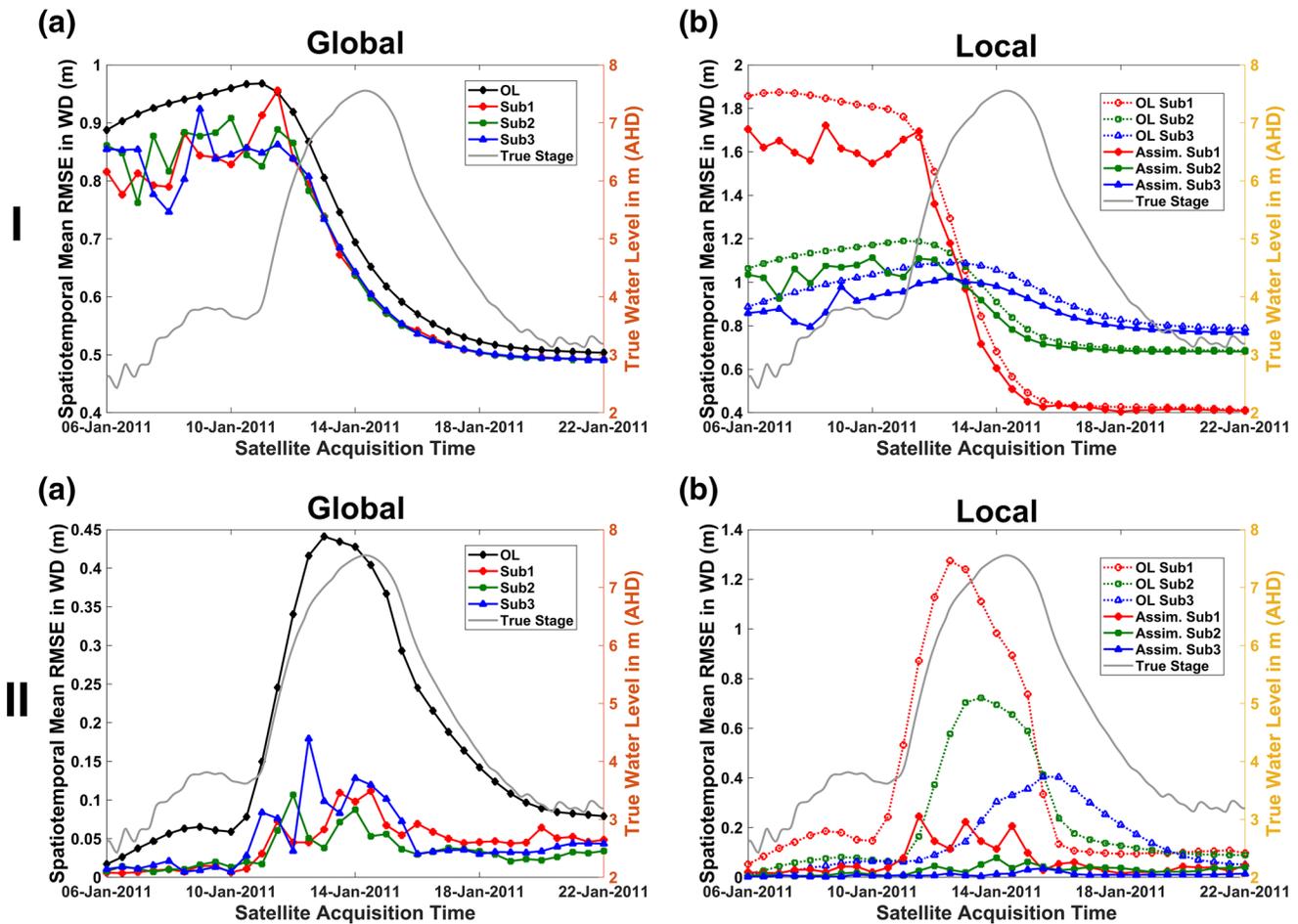
When images were assimilated in Sub 2 and Sub 3 with the flatter floodplains (green and blue lines in Figure 4, respectively), the forecast within and downstream of the observation footprint improved notably. However, the BSS substantially declined for some specific time steps at the upstream gauges located within Sub 1, most visibly for the images acquired after the flood peak in Sub 3, as observed from the noticeable dip in the blue line at the Sub 1 gauges. One of the reasons for this is that the particles which perform well in Sub 3 differ from those that perform well in Sub 1. This effect is most pronounced when the various reaches experienced different phases of the flood event (e.g., peak vs. falling limb) inducing rapidly varying flood extents. After the valley filling occurs in Sub 3 and the flood extent stabilizes, this effect is reduced and universally well-performing particles were again selected, which also performed well in Sub 1 as evident from the steep rise in the blue line. The green line of BSS in Sub 2 is mostly higher than in Sub 3, indicating that the assimilation is able to better update the forecast in reaches largely independent of tidal influence. Since, the deterministic downstream tidal levels are the dominant flow controls in Sub 3 (see the low sensitivity of the blue curve in Figure 3d), the impact on the forecast from assimilating in Sub 3 is limited as tidal levels are not updated by the assimilation. Indeed, adding the downstream level as a source of uncertainty, would allow for larger forecast improvements as a result of the assimilation.

Assimilating single images using the proposed algorithm demonstrated a large positive impact as indicated by the magnitude of BSS values. Except the images acquired in the rising limb before the flood peak, all images assimilated in Sub 2 and Sub 3 after the 14th of Jan reduced the quadratic errors of the water level forecast by more than 50% for notable lead times. The positive impact of flood extent assimilation increased in general as the flooded area increased in the domain, as evident from the increasing trend in the BSS curves with time for Sub 2 and Sub 3. Images at and after the peak consistently produced persistent and strong positive impacts, as the inflow values and the corresponding error added to the domain at each consecutive time step started to diminish. Notably, the BSS values just after the peak were highest for the images covering Sub 2 (green curve), with the maximum BSS using a single image resulting in ~95% relative improvement over the open loop. This was in line with expectations as this reach is characterized by gravitational flow and not heavily influenced by tides.

Evaluating the spatial water depth RMSE over time, facilitates an evaluation of the spatiotemporal impacts of flood extent assimilation on distributed inundation depth simulations for different observational configurations. Such analyses can aid the identification of an optimum targeted observation design for floodplains as well, which had never been previously investigated in the context of flood extent assimilation. Figure 5 therefore shows the spatiotemporal (i.e., computed across space and time)  $WD_{RMSE}$  of water level over the entire domain in (a) and within the assimilation sub-domain in (b), with Part I of the figure focusing on the mean water depth RMSE from the satellite acquisition time to the end of the forecast. The red, blue, and green lines representing assimilation in the three unique sub-reaches, must be interpreted in comparison with the corresponding black line of the open loop, since the statistic is not normalized for flow magnitude or image footprint size.

The reduction in  $WD_{RMSE}$  values both in the global and local cases, was maximum around the time when the flood wave traversed the particular reach, but no clear trends were evident in terms of where or when to assimilate (the three colored lines are always lower than the open loop in black in Figure 5 but one is not consistently better than the other). Images just before the peak resulted in greater global improvements over longer lead times as the forecast was kept on track by the assimilation. There were minor differences in the improvements from the assimilation before the flood peak, which became negligible after the peak, but early images were generally better for the forecast. It is worth noting that the BSS values were consistently high in this region of the hydrograph, especially for assimilation in Sub 2 and Sub 3. However, it is expected that local improvement within the channel might have been smoothed out in the global  $WD_{RMSE}$  due to averaging over a large number of grid cells. Interestingly, this implies that the timing and location for maximum improvements from flood extent assimilation differ for the channel and the floodplain.

Figure 5 I(b) shows the  $WD_{RMSE}$  computed locally across the specific model sub-domains used for the assimilation. The magnitude of local improvements obtained through assimilation in Sub 1 were lower than in the global case, indicating that the assimilation in Sub 1 has an overall positive impact across the domain even though the local improvements are not substantial. Similarly, the assimilation in Sub 2 continued to demonstrate improvements in the global  $WD_{RMSE}$ , even as the local improvement decayed to zero due to

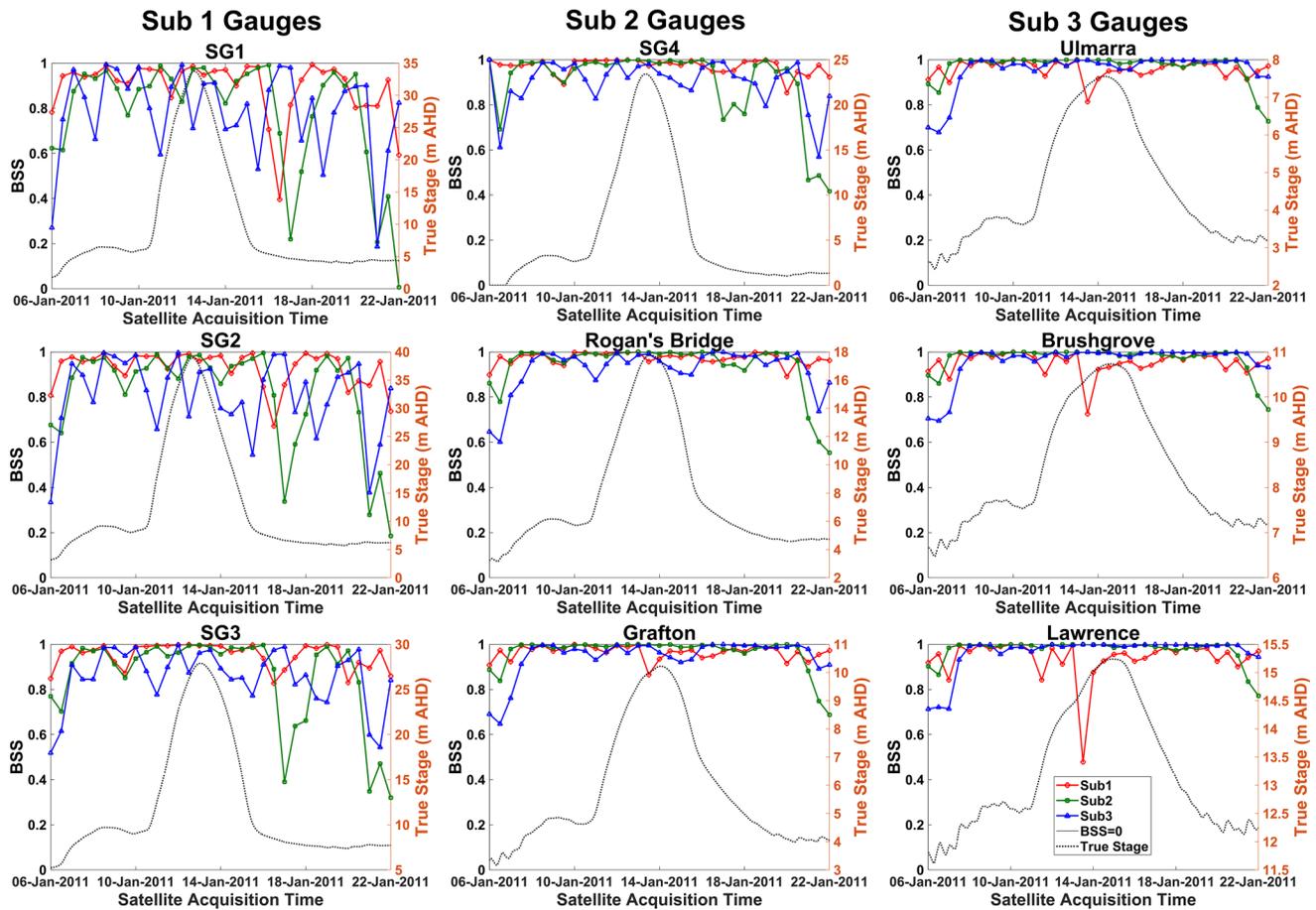


**Figure 5.** Plots of  $WD_{RMSE}$  computed for the (a) global case (over the entire domain) and (b) local case (over the assimilation sub-domain), from the time of the satellite overpass to the end of the forecast in (I) and the 12 h between one assimilation time step to the next in (II). In both cases, observations are assimilated only within a given sub-domain. The true water level at Grafton is shown as a reference in each sub-plot. Here OL refers to the Open Loop forecast, while Sub1/Sub2/Sub3 refer to the assimilated forecast obtained after assimilating images in Sub-reach 1, 2, and 3, respectively. RMSE, root mean squared error.

positive impacts downstream. This is aligned with the findings of E. S. Cooper et al. (2018), where SAR-based water level observations assimilated upstream had more persistent positive impacts consistent from upstream to downstream. For Sub 1, the local impacts became nearly negligible after the inflow peak, which was expected as the valley filling occurred, and the same effect was observable in Sub 2 and Sub 3 at later stages. The local and global  $WD_{RMSE}$  for Sub 3 were consistent as it was the most downstream assimilation location, and therefore did not have notable positive impacts upstream. A maximum reduction of  $\sim 15$  cm in the  $WD_{RMSE}$  was obtained in the best case i.e. assimilating flood extents in Sub 3 during the rising limb, which is substantial from an emergency management perspective.

#### 4.1.2. Impact of Observation Footprint and Timing on Short Lead Times

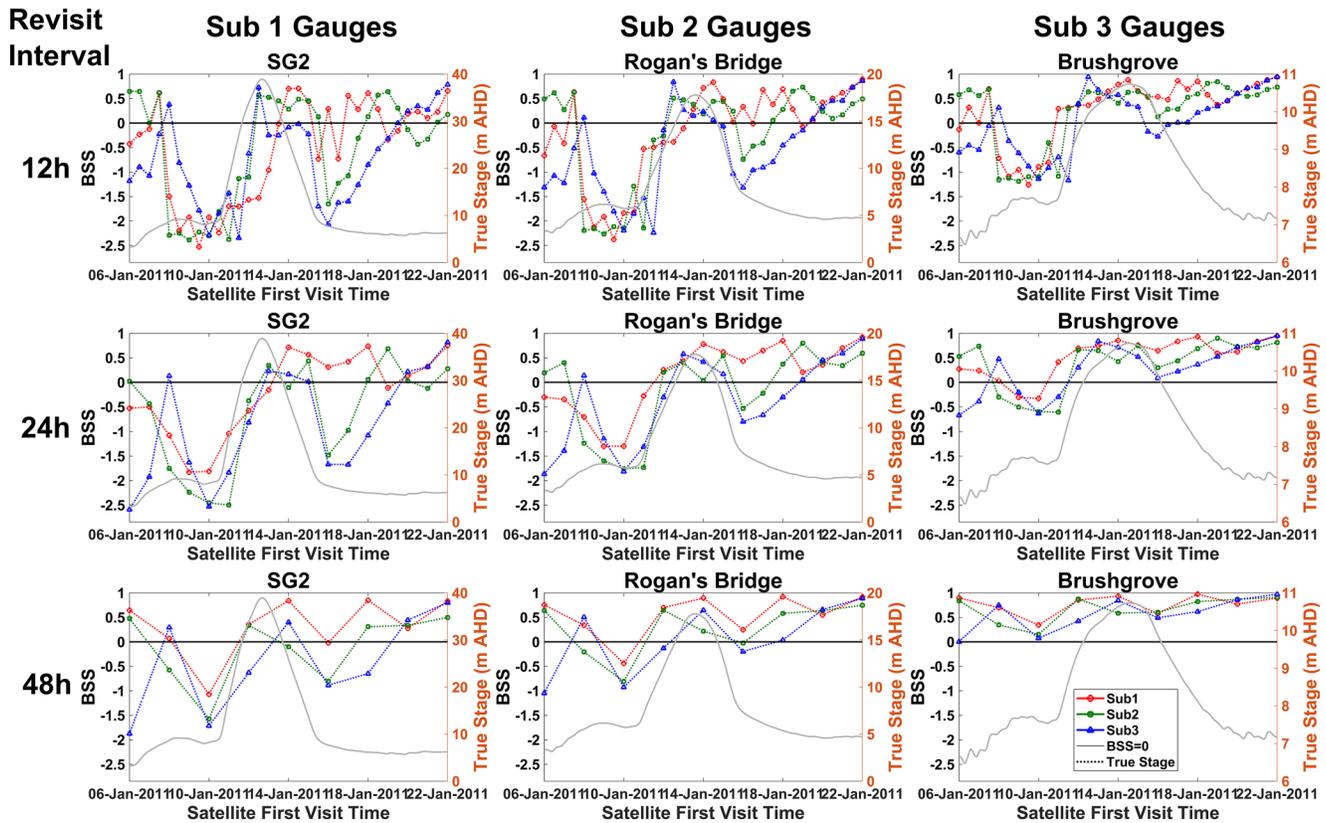
This section investigated the ideal combination of image footprint and timing which maximized forecast improvements for short lead times of up to 12 h after the assimilation. As the previous section has demonstrated possible improvements for longer lead times, this section focused on the performance of single image assimilation for forecasts with short lead times. Channel water level forecasts were evaluated through the BSS as shown in Figure 6. The magnitude of improvement was substantially larger for 12-h lead times in comparison to the longer lead times in Section 4.1.1 This was expected since the temporal distance from the observation time is lesser, thereby the forecast is still “on-track.” There was almost no difference in BSS values when assimilating images in Sub 2 and Sub 3 acquired around the peak. For images assimilated after the peak the difference again increased as the floodplain inundation rapidly increased. Assimilation in Sub



**Figure 6.** As for Figure 4 except for the time window used for the BSS calculation. Here BSS was calculated from the assimilation time step of a single image until the next image became available. As images were considered every 12 h, this time window was restricted to 12 h after each assimilation time step. BSS, Brier Skill Scores.

1 consistently resulted in lower BSS values leading to lesser forecast improvements than the other reaches, which was expected due to the low sensitivity of flood extents to depth in the narrow valley. Assimilation impacts propagate from upstream to downstream, resulting in nearly identical BSS curves at all the gauges for images covering Sub 1. In contrast, the impact of assimilating in Sub 2 was consistent for gauges in Sub 2 and Sub 3, while the forecast at the upstream gauges did not improve as much. Most images assimilated in Sub 2 and Sub 3 (except for a few very early images with limited or no inundation) resulted in  $BSS > 0.9$ , implying  $>90\%$  improvements over the quadratic errors in the open loop.

Spatial performance for the 12 h forecast was also evaluated through the  $WD_{RMSE}$ , computed globally—across the entire domain and locally—within each assimilation sub-domain, as shown in Figure 5 II(a) and (b), respectively. Figure 5 II(a) shows that the assimilation always improves the forecast since the red, green, and blue lines show lower RMSE values than the open loop in black, but the differences resulting from the image footprint in the global case are rather minimal. Images acquired at or around the peak produced considerable forecast improvements in all sub-reaches. However, when the flood peak was entering a particular sub-reach and images were assimilated elsewhere, the assimilation resulted in lower  $WD_{RMSE}$  reductions. For example, the January 13, image assimilated in Sub 3 (blue line in Figure 5IIa), does not improve the RMSE as much as assimilating in Sub 1 and 2. Since the flood wave is leaving Sub 1 at this point in time and entering Sub 2, as evident from the reference true water levels at Grafton located in Sub 2 (gray line in the Figure 5 sub-plots), an observation covering Sub 3 is not very informative for the forecast at this time. In fact, for images acquired just before the flood peak, assimilating in Sub 2 reduced the global  $WD_{RMSE}$  by  $\sim 40$  cm, which is substantial considering that it is an average value computed across  $\sim 550,000$



**Figure 7.** As for Figure 4 but for the multiple image assimilation case with weights carried forward by multiplication; impacts of different revisit intervals are shown in each row. Each point on each curve corresponds to the first visit time and the BSS obtained from the time of the last image assimilated on January 22, 2011 00:00 until the end of the forecast on January 23, 2011 00:00. BSS, Brier Skill Scores.

modeled cells. Even for flood extent assimilation in Sub 1, the  $WD_{RMSE}$  reduced by a maximum of  $\sim 18$  cm, highlighting the potential of this approach for improved forecasts.

Figure 5IIB illustrates that the local improvements were typically larger than the impacts produced globally, with the maximum reductions of  $\sim 1$  m in the local  $WD_{RMSE}$  in Sub 1. Assimilating images at most time steps in Sub 2 and Sub 3 reduced the local  $WD_{RMSE}$  to almost zero for the 12-h forecast, indicating that this lead time falls within the window of system memory. Images after the peak had negligible impact on local  $WD_{RMSE}$  values when assimilated in Sub 1. Assimilation of images after the peak produced the maximum improvements in the channel water level forecast as we saw in the BSS plots of Figure 6, while those assimilated before and at the peak most reduced the absolute errors in spatial flood inundation depth relative to the open loop. Images assimilated at and around the peak demonstrated strong positive impacts on the flood forecast irrespective of assimilation location, although the impact was notably greater in the reaches with larger and gentles floodplains. In fact, the quadratic errors were likely most reduced when the extent varied rapidly, and the floodplain was involved through overbank flows.

## 4.2. Multiple Image Assimilation

### 4.2.1. Impact of Observation Footprint, Timing, and Revisit Interval on the Daily Forecast

The BSS plots for the different multiple image assimilation scenarios considered here are illustrated in Figure 7, to examine the impact of observation footprint, first visit, and revisit intervals on channel water level forecasts for long lead times. Each row illustrates the BSS plots for a given revisit frequency. Each point on each curve corresponds to the first visit time starting from which multiple images were considered with a particular revisit interval, and the BSS value obtained after all the images were assimilated. The final image

considered in each case is on January 22, one day before the end of the event and thus the lead time for the forecast is 24-h. The first visit times are shifted by the duration of the revisit interval starting from the first image considered on January 6, very early in the hydrograph, implying that the number of observations assimilated for each consecutive point decreases by 1. For instance, if the red curve in the BSS plot for SG2 is considered, the first point indicates the BSS value for the assimilation of 33 images in Sub 1 starting from January 6, while the final point on January 22, represents the impact of assimilating 1 image.

Figure 7 shows that the assimilation of flood extents is extremely sensitive to first visit time and revisit intervals. Interestingly, for the sequential assimilation of multiple images, the BSS values obtained for Sub 1 were sometimes higher than the assimilation in the other two sub-domains. This is mainly expected to be a function of the way in which PF-SIS is implemented here and elsewhere in literature, with the posterior for a given observation considered as a prior for the next and cumulative weights obtained through multiplication (Equation 12). As the water rapidly transfers between the channel and the floodplain, the system errors quickly evolve and become independent of the previous observations, thus making the weight multiplication detrimental to the forecast. Since, the shape of the valley in that particular sub-domain is not conducive for high temporal variability in simulated extents, the assimilation impacts were consistently positive for the assimilation of multiple images in Sub 1, implying a longer correlation window of model errors in that region. In contrast, single image assimilation in Sub 2 and Sub 3 was more efficient, as the flood extent rapidly varied between time steps and different particles performed well at different times. Therefore, the forward weight multiplication frequently implemented in PF-SIS algorithms is infeasible for hydraulic data assimilation in river reaches with flat floodplains and gentle slopes, even substantially degrading the forecast when many images are cumulatively assimilated as the model error correlation quickly decays.

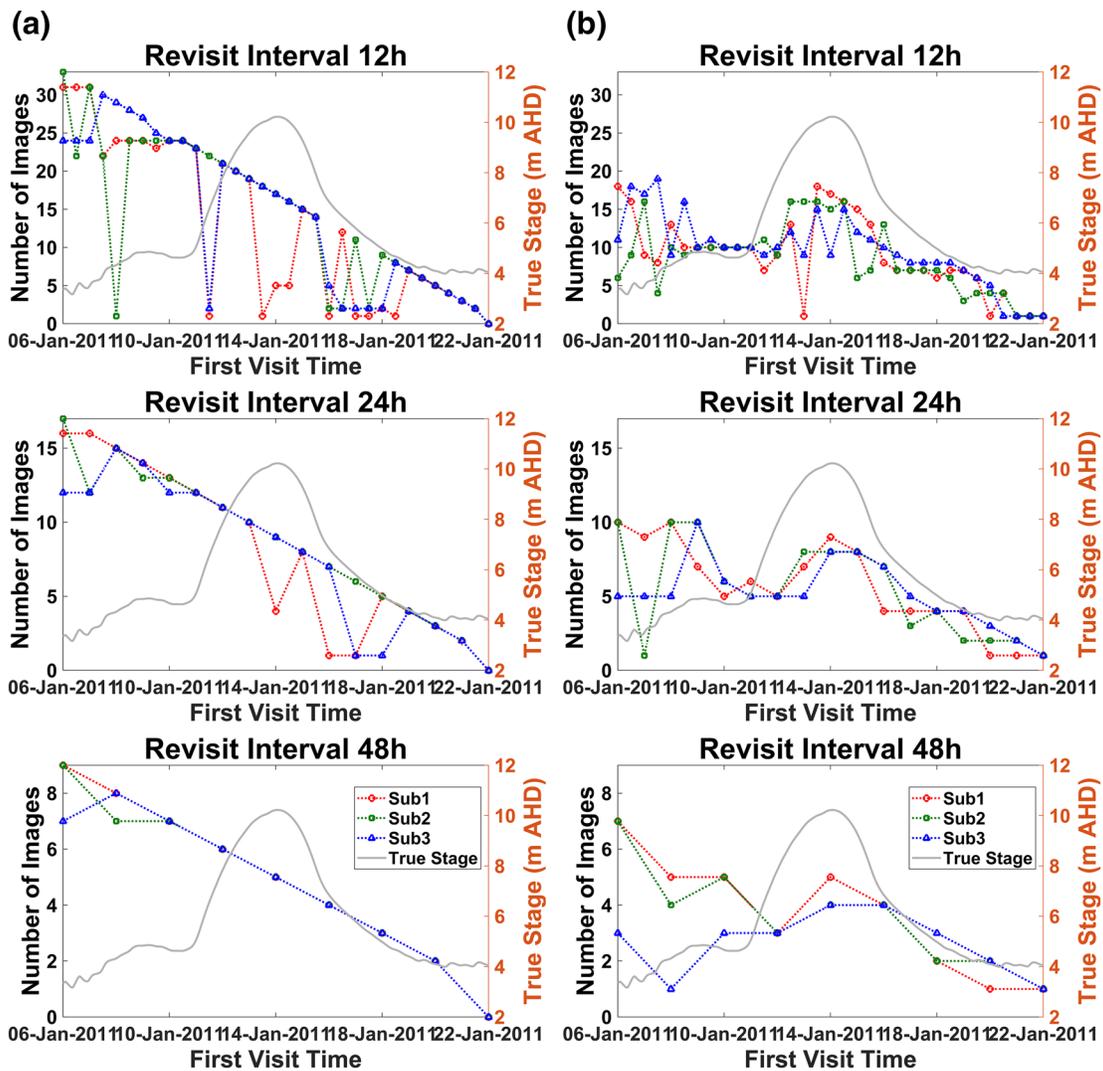
Increasing the length of the revisit interval from 12 to 48 h also did not exhibit noticeable impacts on the timing of the maximum BSS, and the shapes of the plots obtained for the different sub-reaches remained consistent. The magnitude of maximum BSS declined with a decrease in assimilation frequency. This effect was most pronounced for first visit times before the peak and decreased in the post peak images, in line with the findings of García-Pintado et al. (2013). The study found that after the flood peak, SAR-derived water level assimilation lost sensitivity toward revisit frequency. Assimilation impact for images considered within Sub 1 was consistent across all gauges in the domain as impacts propagated upstream to downstream, while images assimilated elsewhere degraded the forecast at gauges located upstream.

#### 4.2.2. Impact of Observation Frequency Based on River Reach Morphology

The previous section revealed that using the posterior at the previous assimilation time step as the prior at the next, degraded the forecast after several images were assimilated. This section thus investigated the maximum number of images which continue to improve the forecast when cumulative weights are computed through forward weight multiplication.

Additionally, this section also investigated the optimum number of images which lead to the maximum positive impact in each sub-domain. Figures 8 (a) and (b) illustrate the number of images which continue to improve the forecast and those which produces the maximum improvements with respect to different first visit times, respectively. BSS values were calculated for the duration of the revisit interval in each case, implying that lead times of 12, 24, and 48 h were considered for forecast error assessments. It is encouraging to note that the assimilation continued to produce positive forecast benefits for >2 weeks for each revisit interval examined, which is much longer than the temporal error correlation enforced on the inflows.

The number of images which continue to lead to forecast improvements differs significantly from the number which produce the most improvements in the forecast. As the revisit interval increased the number of images which continued to have positive impacts increased, except for first visit times very early in the event. For the 48-h revisit interval, for example, all the images assimilated in Sub 1 starting from any first visit time, continued to produce positive impacts since there was only a limited number of observations covering the variable phase of the inundation. However, for the 12-h revisit case, even the number of images leading to forecast improvements in Sub 1 decreased, since the increased temporal coverage meant that multiple observations covering completely different inundation extents were cumulatively considered



**Figure 8.** An illustration of the impact of image frequency on forecast improvements from the assimilation with respect to different first visit times and revisit intervals, where (a) shows the number of images which continue to produce positive impacts, while (b) shows the number of images which produce the maximum positive impact when assimilated sequentially for different revisit intervals. Positive impact was defined in terms of the BSS calculated for the following revisit interval. The true stage at Grafton is shown as a reference in all the sub-plots.

through weight multiplication. For assimilation in Sub 1, the number of images which produced positive impacts was nearly the same as the number of images which produced the largest forecast improvements. This implies that the assimilation efficiency increases in reaches with narrow valleys as more images are considered with weight multiplication.

The differences between columns (a) and (b) in Figure 8, were largest for the 12-h revisit and decreased with increasing revisit intervals. This highlights that the forecast improvements reached a local maximum value and then started to decrease when multiple images were considered in this PF-SIS framework. For example, if the first visit occurred on January 11, with a revisit interval of 24 h (also the forecast lead time), the forward multiplication continues to yield forecast improvements for 12 image dates ( $BSS \gg 0$ ), while the highest BSS was observed after 5 days/images. For first visit times early in the rising limb, maximum forecast improvements were consistently observed in the falling limb for Sub 2 and Sub 3, with a greater number of images continuing to produce forecast improvements in Sub 2 than in Sub 3.

**Table 3**

*Summary Table of Maximum Possible Improvements in BSS Through Flood Extent Assimilation for Gauged Water Level Simulations Within the Channel*

Sub-reach	Revisit	Max. BSS	Ideal acquisition/first visit time	Number of images assimilated
1	12 h	1.00	Very early rising limb	11
	24 h	0.98	Very early rising limb	6
	48 h	0.85	Very early rising limb	4
	Single	1.00	Multiple images	1
2	12 h	1.00	Very early rising limb	11
	24 h	1.00	Just before the peak	8
	48 h	0.99	Post-peak	4
	Single	1.00	Multiple images	1
3	12 h	1.00	Very early rising limb	8
	24 h	1.00	Early rising limb	7
	48 h	1.00	Very early rising limb	7
	Single	1.00	Multiple images	1

*Note.* Morphological characteristics of each sub-reach can be found in Table 1. BSS values were computed for the revisit interval following the assimilation (see Table 1) and averaged across all gauges in the sub-domain. For multiple image assimilation, weights were carried forward through multiplication. Temporal positions mentioned are all with respect to the inflow peak which differs significantly from the peak inundation timing in the floodplain. BSS, Brier Skill Scores.

### 4.3. Maximum Forecast Improvements Possible through Flood Extent Assimilation

The maximum channel water level forecast improvements which can be obtained through flood extent assimilation are quantified in this section in terms of BSS and reported in Table 3. Here, the time window for the BSS calculation was the length of the revisit interval as in Section 4.2.2, and the BSS values were averaged across all gauges. While the previous section explored the timing associated with the maximum BSS value obtained, here the magnitudes are discussed for both single and multiple image assimilation. The maximum BSS that was observed across all cases was  $\sim 1$  indicating a 100% reduction in quadratic errors relative to the open loop. In fact, BSS values  $\sim 1$  were obtained for almost all multiple image assimilation scenarios considered in Sub 2 and Sub 3. This implies that flood extent assimilation exhibits relatively higher sensitivity toward first visit times as compared to revisit intervals or frequency, also noting that for longer revisit intervals the forecast lead time was also longer in this experimental setup. For Sub 1, however, the  $\sim 100\%$  relative improvement in the assimilated forecast was only achieved a few times for each revisit case for the assimilation of 11 images, and the BSS declined with decreasing observation frequency especially for the 48-h revisit or a single visit.

In general, the best performing first visit time for multiple image assimilation in all the reaches was a few days before the inflow peak, for which the maximum BSS was observed during the falling limb of the hydrograph. The only exception to this was the 48-h revisit interval in Sub 2, where starting the assimilation after the peak was most beneficial to the forecast. For single image assimilation, multiple observations covering Sub 2, individually produced  $\sim 100\%$  improvements in the channel water level forecast. It should also be noted here that the forecast lead time for this case is the same as in Section 4.1.1, which is  $>7$  days for this particular image acquired on January 15, about  $\sim 2.5$  days after the inflow peak. This implies that in a budget limited scenario, assimilating a single image at the right time can produce improvements comparable to the assimilation of multiple images. Given these results it appears that it would be expedient to consider when and where the most improvements in forecast accuracy are desired, in order to identify the optimum targeted observation strategy for flood extent assimilation. Readers should note that the exact optimum assimilation time may vary with respect to the upstream inflow error generation model, and thus, more general assertions have been made in Table 3 based on the flood phase.

particular image acquired on January 15, about  $\sim 2.5$  days after the inflow peak. This implies that in a budget limited scenario, assimilating a single image at the right time can produce improvements comparable to the assimilation of multiple images. Given these results it appears that it would be expedient to consider when and where the most improvements in forecast accuracy are desired, in order to identify the optimum targeted observation strategy for flood extent assimilation. Readers should note that the exact optimum assimilation time may vary with respect to the upstream inflow error generation model, and thus, more general assertions have been made in Table 3 based on the flood phase.

## 5. Conclusions

This study strived to identify a targeted observation strategy for optimal forecast improvements from SAR-based flood extent assimilation. Specifically, the impacts of observation footprint, timing, and frequency on flood forecasts was evaluated for different lead times. For this study, a mutual information based PF-SIS framework was implemented to assimilate synthetic observations based on a real event in the Clarence Catchment, simulated by the two-dimensional flood inundation model Lisflood-FP. Several conclusions can be inferred from the results. First, the performance of the assimilation algorithm was found to be sensitive to the location and timing of the image, particularly with respect to reach morphology and flood wave arrival time. For example, assimilating images in Sub 2 and Sub 3 with larger and flatter floodplains resulted in strongly different outcomes than assimilating in Sub 1 with a narrow constrained valley. Second, the images which produced maximum forecast improvements within the channel were different from those which most improved spatial water depth simulations. In line with previous literature, this study showed that improvements from the assimilation increased from upstream to downstream. Multiple image assimilation

was shown to be heavily influenced by first visit time and revisit frequency, with the latter being a function of the reach morphology. It was also demonstrated that for budget limited scenarios, a single image with the ideal coverage and timing could produce improvements comparable to assimilating multiple images, implying that observation coverage and timing were more important than frequency. Results presented herein can be extended to catchments around the world with reaches exhibiting similar morphological and flow characteristics. For instance, even for large and complex river systems such as the Ganges or the Brahmaputra, if inflow errors dominate the forecast uncertainty, reaches with gentle gradients, large floodplains, and limited tidal influence are expected to benefit the most from systematic inundation monitoring.

Currently, users have limited control over image availability and acquisition parameters, but this is expected to change in the near future with multiple radar Earth Observation satellites being launched every year. Satellites can also be tasked with specific acquisitions under the International Charter on “Space and Major Disasters”. This highlights the importance of designing targeted observation strategies such as the one presented here, which are absolutely vital to maximize the forecast accuracy resulting from the flood extent assimilation. Future research should additionally examine the influence of parametric and topographic uncertainties on the optimum satellite acquisition scenario for flood extent assimilation. Moreover, techniques developed here should be tested for different catchment morphologies under a range of flow conditions to identify the opportunities and pitfalls for operational applications.

### Data Availability Statement

All the data and codes for used for this paper are available in a public Figshare repository at <https://doi.org/10.6084/m9.figshare.13301342>

### Acknowledgments

This study was 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. Antara Dasgupta was funded by a PhD scholarship from the IITB-Monash Research Academy. Renaud Hostache's contribution was supported by the National Research Fund of Luxembourg (FNR) through the CASCADE project, grant number C17/SR/11682050.

### References

- Andreadis, K. M. (2018). Data Assimilation and River Hydrodynamic Modeling Over Large Scales. In G. J. -P. Schumann, P. D. Bates, H. Apel, & G. T. Aronica (Eds.), *Global Flood Hazard*. <https://doi.org/10.1002/9781119217886.ch13>
- Andreadis, K. M., Clark, E. A., Lettenmaier, D. P., & Alsdorf, D. E. (2007). Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, *34*, 1–5. <https://doi.org/10.1029/2007GL029721>
- 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>
- Annis, A., Nardi, F., Volpi, E., & Fiori, A. (2020). Quantifying the relative impact of hydrological and hydraulic modelling parameterizations on uncertainty of inundation maps. *Hydrological Sciences Journal*, *65*, 507–523. <https://doi.org/10.1080/02626667.2019.1709640>
- Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *Signal Process IEEE Trans*, *50*, 174–188. <https://doi.org/10.1109/78.978374>
- Banister, A. R., & Nichols, N. (2012). *Ensemble methods in data assimilation* (Vols. 1–38). Data Assimilation Research Center, University of Reading.
- 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*. University of Bristol.
- Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, *375*, 613–626.
- Cooper, E. S., Dance, S. L., Garcia-Pintado, J., Nichols, N. K., & Smith, P. J. (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., Garcia-Pintado, J., Nichols, N. K., & Smith, P. J. (2019). Observation operators for assimilation of satellite observations in fluvial inundation forecasting. *Hydrology and Earth System Sciences*, *23*, 2541–2559. <https://doi.org/10.5194/hess-2018-589>
- Dasgupta, A., Grimaldi, S., Ramsankaran, R., Pauwels, V. R., Walker, J. P., Chini, M., et al. (2018). 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., 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*, *56*, e2020WR027859. <https://doi.org/10.1029/2020WR027859>
- 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 Resources Research*, *48*, W05528. <https://doi.org/10.1029/2011WR011570>
- Di Baldassarre, G., & Montanari, A. (2009). Uncertainty in river discharge observations: A quantitative analysis. *Hydrology and Earth System Sciences*, *13*, 913–921. <https://doi.org/10.5194/hess-13-913-2009>
- Domenghetti, A., Tarpanelli, A., Brocca, L., Barbetta, S., Moramarco, T., Castellarin, A., & Brath, A. (2014). The use of remote sensing-derived water surface data for hydraulic model calibration. *Remote Sensing of Environment*, *149*, 130–141. <https://doi.org/10.1016/j.rse.2014.04.007>

- 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>
- Fewtrell, T. J., Duncan, A., Sampson, C. C., Neal, J. C., & Bates, P. D. (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>
- Gallant, J. C., Dowling, T. I., Read, A. M., Wilson, N., Tickle, P., & Inskip, C. (2011). *1 second SRTM Derived Digital Elevation Models User Guide*. Geoscience Australia. Retrieved from [www.ga.gov.au/topographic-mapping/digital-elevation-data.html](http://www.ga.gov.au/topographic-mapping/digital-elevation-data.html)
- García-Pintado, J., Mason, D. C., Dance, S. L., Cloke, H. L., Neal, J. C., Freer, J., & Bates, P. D. (2015). Satellite-supported flood forecasting in river networks: A real case study. *Journal of Hydrology*, *523*, 706–724. <https://doi.org/10.1016/j.jhydrol.2015.01.084>
- García-Pintado, J., Neal, J. C., Mason, D. C., Dance, S. L., & Bates, P. D. (2013). Scheduling satellite-based SAR acquisition for sequential assimilation of water level observations into flood modelling. *Journal of Hydrology*, *495*, 252–266. <https://doi.org/10.1016/j.jhydrol.2013.03.050>
- Giustarini, L., Hostache, R., Kavetski, D., Chini, M., Corato, G., Schlaffer, S., & Matgen, P. (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., Montanari, M., Plaza Guingla, D. A., Pauwels, V., 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>
- Gobeyn, S., Van Wesemael, A., Neal, J., Lievens, H., Eerdenbrugh, K., De Vleeschouwer, N., 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>
- 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. (2018). Effective representation of river geometry in hydraulic flood forecast models. *Water Resources Research*, *54*, 1031–1057. <https://doi.org/10.1002/2017WR021765>
- Grimaldi, S., Li, Y., Walker, J. P., Pauwels, V. R. N., & Monash University (2017). *Bathymetric survey of the Upper Clarence*. <https://doi.org/10.4225/03/5a20708405eccd>
- Grimaldi, S., Schumann, G. J.-P., Shokri, A., Walker, J. P., & Pauwels, V. R. N. (2019). Challenges, opportunities, and pitfalls for global coupled hydrologic-hydraulic modeling of floods. *Water Resources Research*, *55*, 5277–5300. <https://doi.org/10.1029/2018WR024289>
- 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., Neal, J., Kavetski, D., Wood, M. (2018). Near-Real-time assimilation of SAR-derived flood maps for improving flood forecasts. *Water Resources Research*, *54*, 5516–5535. <https://doi.org/10.1029/2017WR022205>
- Hostache, R., Matgen, P., Giustarini, L., Teferle, F. N., Tailliez, C., Iffly, J. F., & Corato, G. (2015). A drifting GPS buoy for retrieving effective riverbed bathymetry. *Journal of Hydrology*, *520*, 397–406. <https://doi.org/10.1016/j.jhydrol.2014.11.018>
- Huxley, C., & Beaman, F. (2014). *Additional crossing of the Clarence River at Grafton: Flood impact, levee upgrade, and structural considerations*. Hydraulic structures and society—Engineering challenges and extremes. Brisbane, Australia (pp. 1–8).
- Kantas, N., Doucet, A., Singh, S. S., Maciejowski, J., & Chopin, N. (2015). On particle methods for parameter estimation in state-space models. *Statistical Science*, *30*, 328–351. <https://doi.org/10.1214/14-STS511>
- Kumar, A., Dasgupta, A., Lokhande, S., & Ramsankaran, R. A. A. J. (2019). Benchmarking the Indian National CartoDEM against SRTM for 1D hydraulic modelling. *International Journal of River Basin Management*, *0*, 1–10. <https://doi.org/10.1080/15715124.2019.1606816>
- 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>
- Mason, D. C., Schumann, G. J.-P., Neal, J. C., García-Pintado, J., & Bates, P. D. (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>
- Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., 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>
- McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W. (2017). The future of Earth observation in hydrology. *Hydrology and Earth System Sciences*, *21*, 3879–3914. <https://doi.org/10.5194/hess-21-3879-2017>
- Moradkhani, H., Hsu, K.-L., Gupta, H., & Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research*, *41*, 1–17. <https://doi.org/10.1029/2004WR003604>
- Moradkhani, H., Sorooshian, S., Gupta, H. V., & Houser, P. R. (2005). 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>
- Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., & Pappenberger, F. (2009). A data assimilation approach to discharge estimation from space. *Hydrological Processes*, *23*, 3641–3649. <https://doi.org/10.1002/hyp.7518>
- Nichols, N. K. (2010). Mathematical concepts of data assimilation. In W. Lahoz (Ed.), *Data assimilation* (pp. 13–39). Berlin: Springer Berlin Heidelberg.
- Pilotti, M. (2015). Extraction of cross sections from digital elevation model for one-dimensional dam-break wave propagation in mountain valleys. *Water Resources Research*, *52*, 52–68. <https://doi.org/10.1002/2015WR017273>. Received
- Plaza Guingla, D. A., De Keyser, R., De Lannoy, G. J. M., Giustarini, L., Matgen, P., & Pauwels, V. R. (2013). Improving particle filters in rainfall-runoff models: Application of the resample-move step and the ensemble Gaussian particle filter. *Water Resources Research*, *49*, 4005–4021. <https://doi.org/10.1002/wrcr.20291>
- Plaza, D. A., De Keyser, R., De Lannoy, G. J. M., Giustarini, L., Matgen, P., & Pauwels, V. R. (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>
- Pramanik, N., Panda, R. K., & Sen, D. (2010). One dimensional hydrodynamic modeling of river flow using DEM extracted river cross-sections. *Water Resources Management*, *24*, 835–852. <https://doi.org/10.1007/s11269-009-9474-6>
- Reichle, R. H., Walker, J. P., Koster, R. D., & Houser, P. R. (2002). Extended versus ensemble Kalman filtering for land data assimilation. *Journal of Hydrometeorology*, *3*, 728–740. [https://doi.org/10.1175/1525-7541\(2002\)003<0728:EVEKFF>2.0.CO;2](https://doi.org/10.1175/1525-7541(2002)003<0728:EVEKFF>2.0.CO;2)
- Revilla-Romero, B., Wanders, N., Burek, P., Salamon, P., & de Roo, A. (2016). Integrating remotely sensed surface water extent into continental scale hydrology. *Journal of Hydrology*, *543*, 659–670. <https://doi.org/10.1016/j.jhydrol.2016.10.041>

- Rogencamp, G. (2004). Lower Clarence River flood study review—Final report: March 2004: Volume 1 of 2 main text.
- Schumann, G. J. P., & Andreadis, K. M. (2016). A method to assess localized impact of better floodplain topography on flood risk prediction. *Advances in Meteorology*, 2016, 1–8. <https://doi.org/10.1155/2016/6408319>
- Schumann, G., Bates, P. D., Horritt, M. S., Matgen, P., & Pappenberger, F. (2009). Progress in integration of remote sensing-derived flood extent and stage data and hydraulic models. *Reviews of Geophysics*, 47, 1–20. <https://doi.org/10.1029/2008RG000274>
- Schumann, G. J. P., Stampoulis, D., Smith, A. M., Sampson, C. C., Andreadis, K. M., Neal, J. C., & Bates, P. D. (2016). Rethinking flood hazard at the global scale. *Geophysical Research Letters*, 43, 10249–10256. <https://doi.org/10.1002/2016GL070260>
- Sinclair Knight Merz F, and Roads and Traffic Authority of NSW TPS. (2011) Wells Crossing to Iluka Road: Upgrading the Pacific Highway: Tyndale to Maclean alternative alignment: Decision report.
- Walker, J. P., & Houser, P. R. (2005). *Hydrologic data assimilation*. In *Advances in water science methodologies* (Vol. 233). CRC Press. <https://doi.org/10.5772/1112>
- Wood, M. (2016). *Improving hydraulic model parameterization using SAR data*. University of Bristol.
- Wood, M., Hostache, R., Neal, J., Wagener, T., Giustarini, L., Chini, M., 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>
- Xie, H., Pierce, L. E., & Ulaby, F. T. (2002). Statistical properties of logarithmically transformed speckle. *IEEE Transactions on Geoscience and Remote Sensing*, 40, 721–727. <https://doi.org/10.1109/TGRS.2002.1000333>