Sequential Data Assimilation for Real-Time Probabilistic Flood Inundation Mapping

Keighobad Jafarzadegan*, Peyman Abbaszadeh, Hamid Moradkhani

Center for Complex Hydrosystems Research, Department of Civil, Construction, and Environmental Engineering, University of Alabama, Tuscaloosa, AL

*Corresponding author: kjafarzadegan@ua.edu

Abstract

Real-time probabilistic flood inundation mapping is crucial for flood risk warning and ~~decision~~ decision-making during the emergency of an upcoming flood event. Considering the high uncertainties involved in the modeling of a nonlinear and complex flood event, providing a deterministic flood inundation map can be erroneous and misleading for reliable and timely ~~decision~~ decision-making. The conventional flood hazard maps provided for different return periods cannot also represent the actual dynamics of flooding rivers. Therefore, a real-time modeling framework that forecasts the inundation areas before the onset of an upcoming flood is of paramount importance. Sequential Data Assimilation (DA) techniques are well-known for real-time operation of physical models while accounting for existing uncertainties. In this study, we present a ~~Data Assimilation (~~DA~~)~~-hydrodynamic modeling framework where multiple gauge observations are integrated into the LISFLOOD-FP model to improve its performance. This study utilizes the Ensemble Kalman Filter (EnKF) in a multivariate fashion for dual estimation of model state variables and parameters where the correlations among point source observations are taken into account. First, a synthetic experiment is designed to assess the performance of the proposed approach~~,~~; then the method is used to simulate the Hurricane Harvey flood in 2017. Our results indicate that the multivariate assimilation of point-source observations into hydrodynamic models can improve the accuracy and reliability of probabilistic flood inundation mapping by 5-7%, while it also provides the basis for sequential updating and real-time flood inundation mapping.

Keywords: Data Assimilation; Probabilistic Flood Inundation Mapping; Hydrodynamic Model; Ensemble Kalman Filter

1. Introduction

The on-time, accurate, and reliable characterization of an upcoming flood event is imperative for proper decision making and risk analysis. A well-calibrated hydrologic model coupled with reliable weather forecast models can be used to generate the streamflow forecast (Clark and Hay, 2004; Cuo et al., 2011; Habets et al., 2004). While streamflow forecasting during flood events is indispensable, the critical step for flood risk analysis is to estimate the flood inundation areas corresponding to the forecasted streamflow of a potential upcoming event. Hydrodynamic models are common tools used to simulate the physics of a river system and predict the spatiotemporal distribution of water surface elevation (WSE). The predicted ~~water surface elevation~~WSE can be simply converted to water depth and inundation area by overlaying with a high-resolution Digital Elevation Model (DEM) (Merwade et al., 2008; Teng et al., 2017). Since floods happen in a short period and at a certain location, it is ~~most~~ often not possible to find an appropriate remote sensing image that covers those inundated areas during the flood period. This is the main reason that research on flood inundation mapping is mostly limited to post-event analysis where specific study areas with available remote sensing data are used as testbeds.

~~According to the literature, most studies have analyzed the flood events for which the flood extent maps were available from surveying or satellite remote sensing. These studies include but are not limited to, calibration and assimilation of hydrodynamic models (Baldassarre et al., 2009; García-Pintado et al., 2013; Gobeyn et al., 2017; Hostache et al., 2009; Lai et al., 2014; Pappenberger et al., 2007; Rahman and Thakur, 2018; Tarpanelli et al., 2013). Depending on the research objectives, such studies are crucial as they address important theoretical questions and advance the~~

~~flood modeling task. For example, several studies have used satellite remote sensing data, such as Synthetic Aperture Radar (SAR) images, to find the sensitivity of hydrodynamic models to their parameters, compare calibration strategies and test the application of assimilating remote sensing data into these models (Di Baldassarre et al., 2009; Hunter et al., 2005; Mason et al., 2009; Matgen et al., 2010)~~. ~~Since floods happen in a short period and at a certain location, it is most often not possible to find an appropriate remote sensing image that covers those inundated areas during the flood period. This is the main reason that research on flood inundation mapping is mostly limited to post-event analysis where specific study areas with available remote sensing data are used as testbeds.~~

Federal Emergency Management Agency (FEMA) is the leading agency in the United States that provides flood hazard and risk maps over the Contiguous United States (CONUS). While ~~These~~ these maps display ~~the~~ flood-prone areas corresponding to specific return periods (e.g. 100 and 500-year events)~~. While the FEMA flood hazard and risk maps provide general information about risk areas~~, they are not always reliable for an upcoming flood event ~~with different return periods~~. For example, FEMA 100-year and 500-year flood hazard maps covered only one-third and half of the inundated areas induced by Hurricane Harvey in Harris County, Texas, respectively (Pinter et al., 2017). The National Water Center Innovators Program proposed the idea of real-time flood inundation mapping across the United States in 2015 (Maidment, 2017). It highlighted the importance of event-based flood inundation mapping where a model uses the forecasted river discharge to estimate the inundation areas corresponding to a specific flood just before the onset of the event. Compared to the traditional flood hazard mapping, real-time flood inundation mapping is more informative and beneficial for emergency response-related decision-making.

In real-time flood inundation mapping, the model takes advantage of forecasted forcing data and generates inundation areas corresponding to an upcoming flood event. Providing these maps ahead of time is extremely valuable for building a robust flood warning system. Data assimilation (DA) is an effective approach commonly used to improve the performance of real-time hydrologic forecasting by updating the model state variables and parameters when new observation becomes available (Moradkhani et al., 2019). The integration of DA with physical models is highly advantageous as it enables accounting for different sources of uncertainties involved in model predictions. These include (1) forcing data uncertainty due to the limitation of measurements and spatiotemporal representativeness of the data (Alemohammad et al., 2015; Kumar et al., 2017), (2) parameter uncertainty due to equifinality and non-uniqueness of parameters (Abbaszadeh et al., 2018; Leach et al., 2018), (3) model structural uncertainty due to the imperfect representation and conceptualization of a real system (Abbaszadeh et al., 2019; Pathiraja et al., 2018; Zhang et al., 2019) and (4) initial and boundary condition uncertainty (DeChant and Moradkhani, 2014; Lee et al., 2011).

Probabilistic forecasting and uncertainty quantification using DA have been the core of modeling in the atmospheric and oceanic sciences (e.g. Anderson and Anderson, 1999; Courtier et al., 1993). Later, the hydrologic community started to utilize this approach to account for the uncertainties involved in different layers of model predictions and provide a more accurate and reliable ~~model estimates such as~~estimation of soil moisture (Gavahi et al., 2020; Pauwels et al., 2001; Reichle et al., 2002; Xu et al., 2020), streamflow (Moradkhani et al., 2005a~~; Vrugt et al., 2006~~), snow (Sheffield et al., 2003; Slater and Clark, 2006) and ~~so many~~ other hydrologic variables. Despite these advances in hydrologic studies, the application of data assimilation in conjunction with hydrodynamic models has received little attention in the literature. The characterization of uncertainty in hydrodynamic models for probabilistic flood inundation mapping has been mostly limited to ~~conventional techniques, such as random~~ Monte Carlo sampling (Ahmadisharaf et al., 2018; Aronica et al., 2012; Domeneghetti et al., 2013; Neal et al., 2013; Papaioannou et al., 2017; Pedrozo-Acuña et al., 2015; Purvis et al., 2008; Savage et al., 2016) and Generalized Likelihood Uncertainty Estimation (GLUE) (Aronica et al., 2002a; Romanowicz and Beven, 2003).

The effectiveness and application of assimilating remotely sensed data (e.g. Soil Moisture Active Passive (SMAP)) into hydrologic models have been vastly investigated in the literature (Abbaszadeh et al., 2020; Azimi et al., 2020; Lievens et al., 2017). However, given the small scale of the hydrodynamic modeling process, the spatiotemporal resolution of current satellite products is not adequate for assimilating into these models. ~~To properly estimate the flood inundation extent, a spatial resolution less than river width (e.g. 100 m) is recommended. In addition, d~~Due to the short duration of floods, satellite data with ~~daily revisit time~~a sub-daily time scale and spatial resolution less than the river width (e.g. 100 m) is ~~needed~~recommended. Since remote sensing products do not provide such high spatiotemporal resolution data for hydrodynamic models, the research on hydrodynamic data assimilation is limited in the literature. ~~Due to the coarse spatial resolution of satellites that provide water surface elevation data, s~~Some studies have limited their analyses to large rivers with a width of above 1 km (e.g. study of Nile and Amazon) (Brêda et al., 2019). However, since the width of the majority of rivers is less than 100 meters, these studies cannot be practically used in many regions.

~~-~~Several studies used higher resolution synthetic Surface Water and Ocean Topography (SWOT) data to evaluate the performance of assimilation techniques (Durand et al., 2008; Munier et al., 2015; Pedinotti et al., 2014; Yoon et al., 2012). While these works provided important information about the assimilation of satellite data into hydrodynamic models, their applications are only limited to synthetic experiments, making them impractical for real case studies. Some studies have implemented indirect methods to estimate WSE from flood extents generated by high-resolution SAR satellite data (Giustarini et al., 2011; Hostache et al., 2010; Matgen et al., 2010b; Neal et al., 2009). This approach can provide high-resolution data that is suitable for the majority of rivers. However, the reliability of this data is concerning because the methods used to convert the flood extent to WSE pose additional errors ~~which~~ that downgrade~~s~~ the quality of the final observed data for assimilation practices. Besides these issues, the major drawback of remote sensing data assimilation pertains to their coarse temporal resolutions. To efficiently monitor the flood dynamics, the assimilation process should be performed at a daily/hourly time scale, however, the revisit frequency of satellites used for capturing the ~~water surface elevation~~WSE ranges from a week to a month. Therefore, there is a significantly low chance to capture multiple real-time remote sensing images for the majority of inundated catchments during flood events. In the most optimistic scenario, assimilation of satellite data is only limited to one/two updates during the simulation period which may not be sufficient for reliable probabilistic flood inundation mapping.

Application of DA in hydrodynamic modeling can be either river monitoring or flood inundation mapping. The goal of hydrodynamic data assimilation for river monitoring is to track variations in the channel roughness and bathymetry in the long run. Therefore, the weekly/monthly satellite data can be well assimilated into the models as the channel characteristics do not change on a daily basis. On the other hand, flood inundation mapping needs an hourly/daily track of WSE because floods happen rapidly and affect the river dynamics on a short time scale. The literature indicates those studies that assimilated data into hydrodynamic models have been mostly designed for river monitoring (Brêda et al., 2019; Durand et al., 2008; Yoon et al., 2012b). To capture the daily dynamics of the rivers for real-time flood inundation mapping, the discharge and water stage values measured at the gauge stations can be assimilated into the hydrodynamic models. Xu et al., (2017) performed a Particle Filtering (PF) approach to assimilate the water stage data from six gauges into a hydrodynamic model. In order to calculate the particle weights in the filtering process, they assumed that gauge observations are independent. In this study, however, we consider interconnections among the gauge stations and apply multivariate Ensemble Kalman Filter (EnKF) to a ~~2D~~ two-dimensional (2D) hydrodynamic model for better characterization and quantification of uncertainty and further improving the accuracy of model simulations.

Advancing the probabilistic hydrodynamic modeling with DA techniques is a necessary step to fill the gap between hydrology and hydrodynamics. To address this problem, this study aims to explore the capability of a standard sequential DA technique, namely the EnKF, for real-time probabilistic flood inundation mapping. ~~The p~~Past studies that used the DA in conjunction with hydrodynamic models, have mostly focused on the quantification of uncertainty in one or two hydrodynamic variables ; ~~(~~e.g. Giustarini et al., (2011) and Hostache et al., (2018) only investigated the uncertainty in the upstream flow and rainfall, respectively; Yoon et al., (2012) focused on the uncertainty of river bathymetry while ignoring the roughness parameter uncertainty~~)~~. In addition, the main application of DA-hydrodynamic modeling framework has been in river monitoring at long-term or water stage forecasting during the flood events (Brêda et al., 2019; Matgen et al., 2010; Xu et al., 2017). However, this study takes one step further and proposes a DA-hydrodynamic modeling framework for real-time probabilistic flood inundation mapping while accounting for ~~all~~ major sources of uncertainties involved in the model simulations including ~~. These include~~ hydrodynamic model parameters (channel roughness and river bathymetry) ~~uncertainty~~, forcing data (river boundary conditions) ~~uncertainty~~, and state variable (water depth) ~~uncertainty~~. Additionally, unlike past works that assimilated either discharge or water stage into the hydrodynamic model, this study performs a multivariate DA to incorporate the observed values of both variables into the hydrodynamic model for a reliable simulation of ~~flooding and its corresponding~~ inundation area.

2 Data and Study area

In this study, we simulate the Hurricane Harvey flood, one of the worst natural disasters in the history of the United States that caused more than 120 billion USD ~~in damage~~ (https://www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf). The Harvey storm hit Texas on August 25, 2017, caused massive precipitation for six continuous days and resulted in extreme flooding condition in Houston and surrounding areas. Given the considerable uncertainties in hydrologic and hydrodynamic processes of such an extreme flood, a deterministic modeling approach with fixed inputs provides erroneous simulations that are highly different from observations. To account for the uncertainties involved in different layers of flood simulation, this study implements a DA-hydrodynamic modeling framework and provides probabilistic flood inundation maps.

Figure 1.a shows the study area that consists of four main channels (blue lines) and eight tributaries (red lines). The study area is located in the State of Texas (Figure 1.b) in the middle of the San Jacinto watershed (Figure 1.c), a highly developed basin (USGS HUC6 #120401) with the area of 10400 km^2. The main channels simulated in the study are around 106 km ~~and~~ draining into three HUC8 watersheds; the Spring (#12040102), West Fork San Jacinto (#12040101) and East Fork San Jacinto (#12040103). The drainage areas of the channels are relatively flat with an average slope of 0.62%, and the soil is mostly impermeable due to the high rate of recent developments in this region. The upstream and downstream boundary conditions (purple points) are provided from

the daily streamflow in four United States Geological Survey (USGS) gauges ((#08068090, # 08068500, #08068740, #08068780) and water stage time series at the downstream gauge (#08069500). The daily streamflow discharge in two internal gauges (green points #08068800 and #08069000) and water stage time series in the second internal gauge are the observations that ~~will~~ ~~be~~are assimilated into the LISFLOOD-FP model. Internal gauges refer to those stations located between upstream and downstream of the simulated river system. Figures 1.b and 1.c present the geographic location of the study area within the state of Texas and San Jacinto watershed, respectively. To set up the LISFLOOD-FP model, we use a DEM with 120 m spatial resolution resampled from one arc second (30 m) USGS National Elevation Dataset. Such a coarse resolution DEM alleviates the computational intensity of the proposed probabilistic hydrodynamic modeling framework. It should be noted that the subgrid solver used for simulation of flood has the advantage of accepting narrow rivers with a width of less than 120 m while the cell sizes are 120 m. In this study, the DA-hydrodynamic modeling framework is parallelized and performed on the University of Alabama High-Performance Computing ~~(UAHPC)~~ cluster.

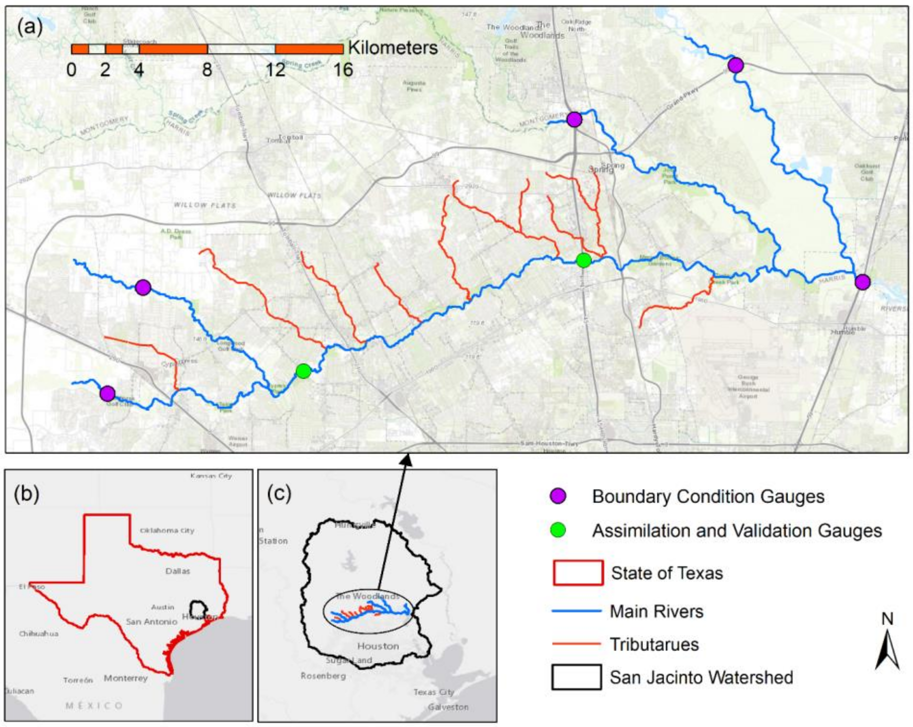

Figure 1 (a) Study area with all gauges, rivers, and tributaries. (b) Geographic location of San Jacinto Watershed within the state of Texas. (c) Geographic location of the study area within San Jacinto watershed (© NhDplus and USGS).

3. Methods

3.1 Flood inundation model

The flood inundation model used in this study is LISFLOOD-FP (Bates and De Roo, 2000), a raster-based 2D hydrodynamic model that simulates the spatiotemporal distribution of ~~water surface elevation~~WSE over the study area. The model solves the momentum and continuity equations (Saint Venont equations):

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = 0 \quad (1)$$

$$\frac{1}{A}\frac{\partial A}{\partial t} + \frac{1}{A}\frac{\partial (\frac{Q^2}{A})}{\partial x} + g\frac{\partial h}{\partial x} - g(S_0 - S_f) = 0 \quad (2)$$

where Q is the flow rate at a given cross-section with the area of A in the main channel, x denotes the location along the channel, t represents time, S_0 and S_f are channel bed and friction slopes, and g is the gravitational acceleration.

We use the sub-gird channel solver, the most recently developed numerical scheme that considers friction and water slope as well as local acceleration components in the shallow water equations (Neal et al., 2012). This solver is advantageous for large-scale and efficient modeling as it utilizes coarse resolution DEMs along with channel widths ~~values~~ that are smaller than DEM resolution. Since DA-hydrodynamic modeling requires hundreds of model simulations~~, a computationally intensive operation~~, this solver helps reduce the computational burden of each simulation and enables implementing probabilistic flood inundation mapping within a DA framework. To set up the model, we assume rectangular cross-section areas and a uniform roughness for both channel and floodplain. Given the low sensitivity of LISFLOOD-FP to the floodplain roughness (Hall et al., 2005; Horritt and Bates, 2002), this parameter is assumed a constant value. However, the uncertainty of channel roughness is ~~the only model roughness parameter whose associated uncertainty is accounted for~~taken into account within the assimilation framework. We also consider the uncertainty of bathymetry by defining an offset parameter that uniformly lowers the DEM values of the river channels. In addition to model parameters (channel roughness and bathymetry), the upstream and lateral fluxes entered the river system as the boundary conditions of the model are other main sources of uncertainty in the assimilation framework.

The upstream boundary conditions are generated from four USGS gauge stations (Figure. 1). To estimate the lateral fluxes, we calculate the deficit in the system as subtraction of the upstream from downstream flows and then, distribute the deficit among river tributaries based on their drainage areas (Please refer to Jafarzadegan et. al (2021) for detailed information about the calculation of lateral flows in this study area). In section 3.3, we will further discuss the procedure we used to initialize the model parameters and river boundary conditions.

3.2 Ensemble Kalman Filter (EnKF)

~~(~~Moradkhani et al.~~,~~ (2005b) provided a comprehensive description of the EnKF formulation for dual estimation of state and parameters in hydrologic models. Here we briefly describe the EnKF formulation for multivariate assimilation of point source water stage and discharge data into a hydrodynamic model. For a more effective assimilation proccess, both types of interconnections between observations, namely spatial correlation of a single observation (discharge or water stage) among different gauges ~~as well as~~and the correlation between both observations at a single gauge are taken into account in the EnKF equations. The ~~In this study,~~ EnKF is used to simultaneously estimate model states and parameters. For this purpose, the parameters should be treated similar to the state variables with a difference that parameter evolution is generated artificially.

Let's assume a DA-hydrodynamic modeling framework with l parameters ($p = 1,2, \ldots, l$), m states ($s = 1,2, \ldots, m$) and n observations ($j = 1,2, \ldots, n$). The following EnKF equations are described in accordance with the flowchart shown in Figure 2. In the EnKF, parameter samples can be generated by adding the noise of η_t with covariance $\sum_t^{\theta}$ to the prescribed parameters.

$$\theta_{t+1}^{i-} = \theta_t^{i+} + \tau_t^i \qquad \tau_t^i \sim N(0, \eta_{t+1}) \quad \forall \quad \eta_{t+1} = \sum\nolimits_{t+1}^{\theta} \tag{3}$$

Using θ_{t+1}^{i-} and forcing data, a model state ensemble and predictions are generated, respectively.

$$x_{t+1}^{i-} = f\left(x_t^{i+}, u_t^i, \theta_{t+1}^{i-}\right) + \omega_t^i \quad \omega_t^i \sim N(0, Q_t) \quad \forall \quad Q_t = \textstyle\sum_t^x \tag{4}$$

$$\hat{y}_{t+1}^i = h\left(x_{t+1}^{i-}, \theta_{t+1}^{i-}\right) + \nu_{t+1}^i \qquad \nu_{t+1}^i \sim N(0, R_{t+1}) \qquad \forall \quad R_{t+1} = \textstyle\sum_{t+1}^y \tag{5}$$

where x_t, u_t, θ_t and y_t are the vector of the uncertain state variables, forcing data, model parameters and observation data at time step t, respectively. ω_t represents the model errors due to the imperfect model, and ν_t is the measurement error. Most often, ω_t and ν_t are assumed to be white noises with mean zero and covariance Q_t and R_t, respectively. In addition, the two noises ω_t and ν_t are assumed to be independent.

Then we update the parameter ensemble members using the standard Kalman filter equation:

$$\theta_{t+1}^{i+} = \theta_{t+1}^{i-} + K_{t+1}^{\theta}\left(y_{t+1}^i - \hat{y}_{t+1}^i\right) \tag{6}$$

where $K_{t+1}^{\theta} \in \mathbb{R}^{l \times n}$ is the Kalman gain matrix for correcting the parameter trajectories and is obtained by:

$$K_{t+1}^{\theta} = \textstyle\sum_{t+1}^{\theta y}\left[\sum_{t+1}^{yy} + R_{t+1}^{'}\right]^{-1} \tag{7}$$

where $\sum_{t+1}^{\theta y} \in \mathbb{R}^{l \times n}$ is the cross-covariance matrix of parameter ensemble and prediction ensemble (Eq. 6). Unlike other studies, and for more realistic characterization of observation and model errors here the correlation between the errors associated with *n* observation data are accounted for during the assimilation process. Therefore, the covariance matrix $R'_t \in \mathbb{R}^{n \times n}$ is a nonzero matrix, such that the values in the diagonal represent the error associated with each observation data and all elements lower/upper the main diagonal denote the cross covariance between different observations (Eq. 7). $\sum_t^{yy} \in \mathbb{R}^{n \times n}$ is also a similar covariance matrix with the inclusion of error correlation between the model simulations (Eq. 8).

$$\sum_{t+1}^{\theta y}(p,j) = \frac{1}{N}\sum_{i=1}^{N}\left[\left(\theta_{t+1}^{i-}(p) - E[\theta_{t+1}^{-}(p)]\right)\left(\hat{y}_{t+1}^{i}(j) - E[\hat{y}_{t+1}(j)]\right)\right] \tag{8}$$

$$R_{t+1}^{'}(j,j') = \begin{cases} R_{t+1} & j = j' \\ \frac{1}{N}\sum_{i=1}^{N}\left[\left(y_{t+1}^{i}(j) - E[y_{t+1}(j)]\right)\left(y_{t+1}^{i}(j') - E[y_{t+1}(j')]\right)\right] & j \neq j' \end{cases} \tag{9}$$

$$\sum_{t+1}^{yy}(j,j') = \frac{1}{N}\sum_{i=1}^{N}\left[\left(\hat{y}_{t+1}^{i}(j) - E[\hat{y}_{t+1}(j)]\right)\left(\hat{y}_{t+1}^{i}(j') - E[\hat{y}_{t+1}(j')]\right)\right] \tag{10}$$

$$E[\theta_{t+1}^{-}] = \frac{1}{N}\sum_{i=1}^{N}\theta_{t+1}^{i-} \tag{11}$$

$$E[\hat{y}_{t+1}] = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_{t+1}^{i} \tag{12}$$

Now using the updated parameter, the new model state trajectories (state forecasts) and prediction trajectories are generated:

$$x_{t+1}^{i-} = f\left(x_t^{i+}, u_t^i, \theta_{t+1}^{i+}\right) + \omega_t^i \quad \omega_t^i \sim N(0, \textstyle\sum_t^x) \quad \forall \quad Q_t = \sum_{t+1}^{x} \tag{13}$$

$$\hat{y}_{t+1}^{i} = h\left(x_{t+1}^{i-}, \theta_{t+1}^{i+}\right) + \nu_{t+1}^{i} \qquad \nu_{t+1}^{i} \sim N\left(0, \textstyle\sum_{t+1}^{y}\right) \quad \forall \quad R_{t+1} = \sum_{t+1}^{y} \tag{14}$$

Model states ensemble is similarly updated as follows:

$$x_{t+1}^{i+} = x_{t+1}^{i-} + K_{t+1}^{x}\left(y_{t+1}^{i} - \hat{y}_{t+1}^{i}\right) \tag{15}$$

$$y_{t+1}^{i} = y_{t+1}^{i} + \nu_{t+1}^{i} \qquad \nu_{t+1}^{i} \sim N(0, R_{t+1}) \quad \forall \quad R_{t+1} = \sum_{t+1}^{y} \tag{16}$$

where $K_{t+1}^{x} \in \mathbb{R}^{m \times n}$ is the Kalman gain for correcting the state trajectories and is obtained by:

$$K_{t+1}^{x} = \sum_{t+1}^{xy}\left[\sum_{t+1}^{yy} + R_{t+1}^{'}\right]^{-1} \tag{17}$$

where $\sum_{t+1}^{xy} \in \mathbb{R}^{m \times n}$ is the cross-covariance matrix of states ensemble and prediction ensemble (Eq. 16).

$$\sum_{t+1}^{xy}(s,j) = \frac{1}{N}\sum_{i=1}^{N}\left[\left(x_{t+1}^{i-}(s) - E[x_{t+1}^{-}(s)]\right)\left(\hat{y}_{t+1}^{i}(j) - E[\hat{y}_{t+1}(j)]\right)\right] \tag{18}$$

$$E[x_{t+1}^{-}] = \frac{1}{N}\sum_{i=1}^{N} x_{t+1}^{i-} \tag{19}$$

In this study the water depth along the channel is the only state variable (*m*=1). The channel roughness and bathymetry are two model parameters (*l=2)* and three point source observations including water discharge at gauge 1 and 2 as well as water stage at gauge 2 (*n*=3) are assimilated into the LISFLOOD-FP model (Table 1). Therefore, the Kalman gains used to update the model parameters and states (Eqs 5 and 15) are 2×3 and 1×3 matrices that take advantage of a multivariate point source assimilation while considering the downstream correlation between discharge observations and the correlation between water stage and discharge at gauge 2.

Formatted: Font: Italic, Complex Script Font: Italic

3.3~~. Experimental design~~DA-hydrodynamic modeling framework

~~The ultimate goal of this study is to simulate the Hurricane Harvey flood and generate probabilistic flood inundation maps through the DA-hydrodynamic modeling framework.~~ Figure. 1 illustrates the flowchart of the proposed DA-hydrodynamic modeling framework used for real-time probabilistic flood inundation mapping~~ approach~~. In this study, the EnKF is performed based on an ensemble size of 100. The boundary conditions including four upstream flows, seven lateral fluxes, and downstream flows are perturbed with adding white noises sampled from a normal distribution with a mean zero and relative error of 20%. The errors are assumed heteroscedastic meaning that their values are proportional to the flow magnitude. ~~(~~Pelletier, (1988) conducted a literature review on the uncertainty of recorded flow at rivers and demonstrated that the error varies in the range 8%-20%. Later, Di Baldassarre and Montanari, (2009) found that the uncertainty of extreme flows can exceed to 25% due to ~~the~~ extrapolating the rating curves. ~~-~~To characterize uncertainty in the initial condition, namely water depth, we add a white noise with a mean zero and standard deviation of 1 meter. In this study, using the proposed EnKF-based multivariate assimilation approach, three point-scale observations, i.e., discharge at USGS gauges 1 and 2, as well as water stage at gauge 2, are incorporated into the LISFLOOD-FP model to rectify its state variables and parameters, and hence provide more accurate and reliable flood inundation maps. ~~All these three observations are perturbed by adding a normally distributed white noise with a mean zero and a relative error of 20%.~~ First, the LISFLOOD-FP model is forced with the upstream, downstream and lateral flow ensembles. To initialize the state variables in the system, the simulated water depth values at the ending day of the warm-up period (the initial condition for the first day of the model simulation) are perturbed with adding a white noise with a mean zero and standard deviation of 1 meter. It is worth mentioning that the error terms used for the observed flows and the initial water depth are determined through a manual tuning to achieve the most reliable predictions during the simulation. The model parameters (i.e., channel roughness and bathymetry) are initialized using the Latin Hypercube Sampling method and evolved during the assimilation process. The ensemble of water depth values predicted by the model for the next time step together with observations, namely water stage and discharge at gauges are used in the multivariate Kalman equation to update the model parameters. The LISFLOOD-FP model is run for the second time with the updated parameters and the second multivariate Kalman equation uses the predicted water depth with observations to update the ensemble of water depth in the system. The ensemble of updated water depth (state), bathymetry, and channel roughness (parameters) ~~will be~~are used within the LISFLOOD-FP to predict an ensemble of water depth for the next time step. The predicted water depth is simply converted to a probabilistic flood inundation map. Using this data assimilation framework, we can generate 1-day forecast of probabilistic flood inundation maps which would be highly beneficial for real-time flood warning and decision making. ~~It is worth mentioning that the forecasted probabilistic maps account for different sources of~~

~~uncertainty including the forcing data (boundary condition flows), model parameters (channel roughness and bathymetry), and initial conditions (water depth).~~

The simulation period of the LISFLOOD-FP model is set up for 45 days from July-30-2017 to Sep-12-2017 and the entire month of July is used as a warm-up period. The model time step and the Courant number are set to 1 second and 0.7, respectively, and the model is simulated at daily scale. The water depth generated for the end of July ~~will be~~is used as the initial condition of the model. To account for the uncertainty of channel roughness and bathymetry, we sample ~~them~~ these variables from uniform distributions ranging from [0, 0.1] and [39, 42] m, respectively. The bathymetry parameter is the elevation of the channel bed at the upper location of the channel. The offset parameter is calculated by subtracting this value from DEM at the upper location. Then, the bathymetry vector that includes the channel bed elevation for all channel cells is generated by subtracting the offset from DEM values along the channel. It should be noted that the range of uniform distribution for channel roughness is chosen based on previous studies (Aronica et al., 2002b; Bales and Wagner, 2009; Di Baldassarre et al., 2009; Horritt, 2006; Pappenberger et al., 2008) while the error range assumed for the bathymetry is mostly determined based on ~~,~~ expert judgment, and trial-and-error. Since the real magnitude and distribution of these errors have not been fully understood in the literature, their estimated values may not be ~~necessarily the~~ physically correct terms and their estimation is ill-posed according to Renard et al., (2010).

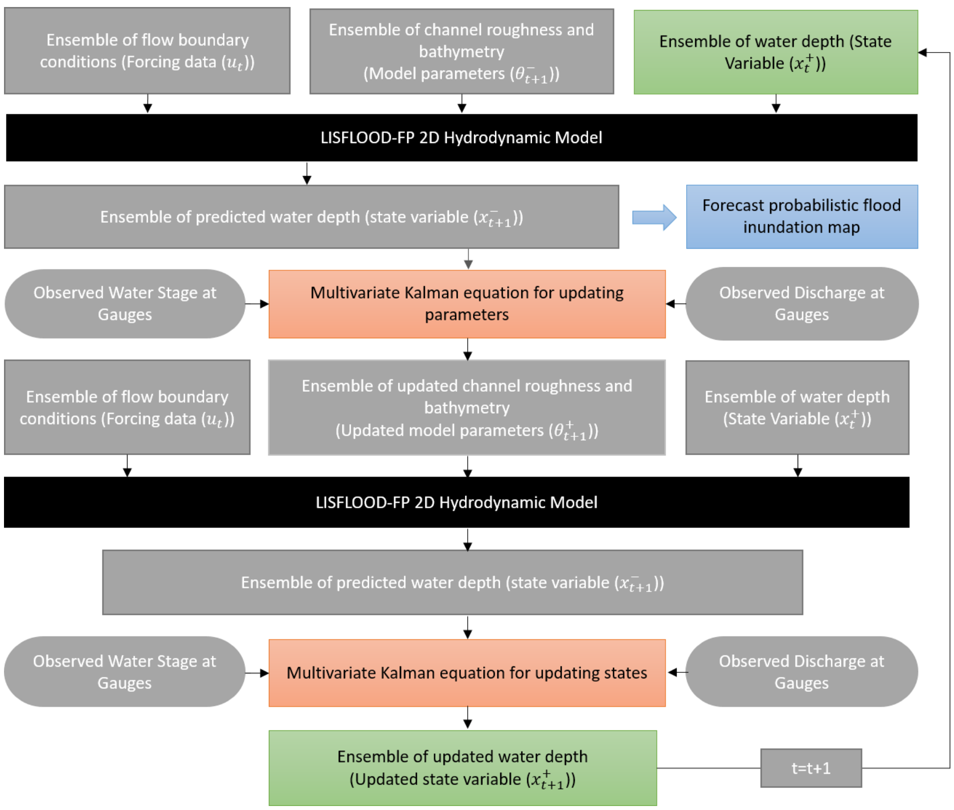

Figure 2. Schematic of the DA-hydrodynamic modeling framework for real-time probabilistic flood inundation mapping. The green boxes represent the state variables where their updated values are fed into the LISFLOOD-FP model and provide a probabilistic flood inundation map at the forecast mode (blue box). The black boxes highlight the physical model and the orange boxes represent the Kalman equations used for updating the parameter and state variables by the EnKF.

3.4 Experimental Design

To assess the effectiveness and robustness of the proposed assimilation framework for probabilistic flood inundation mapping, we design ~~three~~ two ~~different~~ experiments. ~~First, an open-loop (OL) simulation is established where the model is run without assimilation.~~ In the ~~second~~ first experiment, we perform DA-hydrodynamic modeling on a synthetic case study where we assume

Formatted: Normal, Line spacing: single

Formatted: Font: (Default) +Headings CS (Times New Roman), 14 pt, Bold, Complex Script Font: +Headings CS (Times New Roman), 14 pt, Bold

Formatted: Font: 14 pt, Bold, Complex Script Font: 14 pt, Bold, Check spelling and grammar the model is perfect and has no error. In this approach, we set the model parameters (channel roughness and bathymetry), initial state (water depth) and boundary condition flows to fixed values and run the model to generate discharge and ~~water surface elevation~~WSE across the gauges within the study area. These predicted values are assumed as benchmark observations. This synthetic analysis ensures that the assimilation process performs well and the model parameters end up converging to predefined values. In the ~~next step~~second experiment, we implement the proposed assimilation framework on a real case study where the observed discharge and ~~water surface elevation~~WSE data that are recorded from the USGS gauges during Hurricane Harvey, are assimilated into the model. In both experiments, we implement an open-loop (OL) simulation where the model is run without an assimilation. The WSE and flood extent maps generated by OL are compared with the results provided by the EnKF in the synthetic and real case studies. Considering the severe flood condition during the Hurricane, we aim to investigate the extent to which the multivariate DA-Hydrodynamic modeling framework improves the model simulation and flood inundation mapping skill.

3.~~4~~ 5 Validation strategy

As mentioned before, the convergence of uncertain model parameters toward truth in the synthetic experiment demonstrates the performance of DA-hydrodynamic modeling framework. To provide a robust analysis of each assimilation run, it is necessary to assess the model performance through multiple deterministic (*KGE* and *RMSE*) and probabilistic (*NRR* and *Reliability*) measures. ~~The summary of~~The four performance measures used in this study, namely Kling Gupta Efficiency (*KGE*), Root Mean Square Error (*RMSE*), Normalized Root Mean Square Error Ratio (*NRR*) and *Reliability* ~~y is tabulated in Table 1~~are calculated using Eqs. 20-23, respectively.

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

~~Table 1: Summary of performance measures used in this study~~

~~Performance Measure~~	~~Mathematical Representation~~
~~Kling-Gupta Efficiency (KGE)~~	$1-\sqrt{\left(\left(\frac{\mathrm{Cov}_{y_t y_t'}}{\sigma\sigma'}\right)-1\right)^2+\left(\left(\frac{\sigma'}{\sigma}\right)-1\right)^2+\left(\left(\frac{\mu'}{\mu}\right)-1\right)^2}$
~~Root Mean Square Error (RMSE)~~	$\sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t'-y_t)^2}$
~~Normalized Root-Mean-Square Error Ratio (NRR)~~	$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(y_t-\overline{y'_{\blacksquare,t}}\right)^2}\times\left(\frac{1}{T}\left\{\sum_{t=1}^{T}\sqrt{\frac{1}{T}\left[\sum_{t=1}^{T}\left(y_t-\overline{y'_{\blacksquare,t}}\right)^2\right]}\right\}\sqrt{\frac{N+1}{2N}}\right)^{-1}$
~~Reliability~~	$1-\frac{2}{T}\sum_{t=1}^{T}\left\lvert\frac{Z_t}{T}-U_t\right\rvert$

$$1-\sqrt{\left(\left(\frac{\mathrm{Cov}_{y_t y_t'}}{\sigma\sigma'}\right)-1\right)^2+\left(\left(\frac{\sigma'}{\sigma}\right)-1\right)^2+\left(\left(\frac{\mu'}{\mu}\right)-1\right)^2} \tag{20}$$

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t'-y_t)^2} \tag{21}$$

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(y_t-\overline{y'_{\blacksquare,t}}\right)^2}\times\left(\frac{1}{T}\left\{\sum_{t=1}^{T}\sqrt{\frac{1}{T}\left[\sum_{t=1}^{T}\left(y_t-\overline{y'_{\blacksquare,t}}\right)^2\right]}\right\}\sqrt{\frac{N+1}{2N}}\right)^{-1} \tag{22}$$

$$1-\frac{2}{T}\sum_{t=1}^{T}\left\lvert\frac{Z_t}{T}-U_t\right\rvert \tag{23}$$

Formatted: Indent: First line: 0"

Formatted: Font: (Default) +Headings CS (Times New Roman), 12 pt, Complex Script Font: +Headings CS (Times New Roman), 12 pt

Formatted: Font: (Default) +Headings CS (Times New Roman), 12 pt, Complex Script Font: +Headings CS (Times New Roman), 12 pt where y_t and y'_t are the observed and simulated values, respectively. The Kling–Gupta Efficiency (*KGE*) varies from $-\infty$ to 1, such that a value of 1 indicates a perfect fit between observed and simulated values. The pairs of (μ, σ) and (μ', σ') represent the first two statistical moments (means and standard deviations) of y_t and y'_t, respectively. ~~Root mean squared error (~~*RMSE*~~)~~ is the square root of the mean of the square of all of the errors between the predicted and observed values.

NRR (DeChant and Moradkhani, 2012) is calculated to measure the ensemble spread and assess how confidently the ensemble mean is statistically distinguishable from the ensemble spread. *Reliability* (Renard et al., 2010b) is a measure of the fit of the Q-Q quantile plot to a uniform. A value of 1 is exactly uniform and a value of 0 is the farthest possibility from uniform. For the description of the z_t and U_t calculation, we refer the readers to Renard et al. (2010b).

The above four performance measures assess the dynamic behavior of DA-hydrodynamic modeling framework at two specific points. Moreover, to spatially evaluate the behavior of the proposed framework, we compare the maximum probabilistic flood inundation maps (union of probabilistic maps over the simulation period) with the observed floodplain map delineated aftermath of Harvey. The Receiver Operating Characteristic (ROC) graph is a common tool for validating probabilistic classifiers (Fawcett, 2006). Consider a deterministic flood map as a binary map where one and zero represent flooded and non-flooded cells, respectively. First, a threshold in the range of [0,1] is used to convert the probabilistic map to a binary deterministic map. This means all cells with the probability of inundation less than a given threshold are converted to zero and other cells are set to one. The binary map is compared with the reference map and the rate of true positive (*rtp*) and false positive (*rfp*) are calculated using Equations ~~7~~ 24 and ~~8~~ 25 (Jafarzadegan and Merwade, 2017):

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

$$rtp = \frac{True\ positive\ instances}{total\ positives}$$ (~~7~~24)

$$rfp = \frac{False\ positive\ instances}{total\ negative}$$ (~~8~~25)

where true and false positive instances represent the total number of flooded cells in the reference map that are predicted as flood and non-flooded cells, respectively. Total positives and negatives are total flooded and non-flooded cells in the reference map. This process is repeated and a set of points (*rfp*,*rtp*) are generated corresponding to different thresholds. The ROC graph connects the points in the rfp-rtp space and the area under the curve (*AUC*) represents the performance of the probabilistic classifier (Fawcett, 2006). In this study, we use *AUC* to compare the performance of OL simulation with the EnKF for probabilistic flood inundation mapping. The Fit (*F*) index is another performance measure widely used to compare two deterministic flood extent maps in the literature (Alfieri et al., 2014; Bates and De Roo, 2000; Sangwan and Merwade, 2015; Tayefi et al., 2007).

$$F = \frac{True\ positive\ instances}{Total\ positives + False\ positives} \times 100$$ (26)

In addition, we calculate the Underprediction and Overprediction Flood Indices (*UFI* and *OFI*) introduced by Jafarzadegan et al., (2018) for comparing probabilistic flood maps against deterministic reference maps:

$$UFI = \frac{\sum_{i=1}^{N}(1-P_i)}{N} \times 100 \qquad i \in Fl$$ (~~9~~27)

$$OFI = \frac{\sum_{j=1}^{M}(P_j)}{M} \times 100 \qquad j \in NFl$$ (~~10~~28)

where *Fl* and *NFl* denote the flooded and non-flooded regions in the reference map, and *i* and *j* are indicators of cells located within these regions. *N* and *M* are the total number of cells in the *Fl* and

Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic
Formatted: Font: Italic, Complex Script Font: Italic

NFl regions and P_i, P_j denote the probability of inundation for cells i and j derived from the probabilistic flood maps.

4. Results

4.1 Experiment 1: Synthetic Case Study

We conduct the synthetic experiment to ensure the usefulness and effectiveness of the proposed DA-hydrodynamic modeling framework. Figure 3.a presents uncertainty bound evolution of the parameters in the LISFLOOD-FP model (i.e., channel roughness and bathymetry) for 45 days assimilation of synthetic observations (i.e., discharge at gauges 1 and 2 and water stage at gauge 2). ~~The shaded areas correspond to 95, 75, 68, and 10 percentile predictive intervals, and the black stars at the end of each parameter subplot represent the true parameter values.~~ It is worth mentioning that the uncertainty of bathymetry shown in this Figure corresponds to the channel bed elevation at the upper location of the channel. As seen both parameters converge smoothly to the certain region in parameter space where the uncertainty bounds stabilize. While the uncertainty bound associated with the bathymetry becomes stabilized at the early stage of the assimilation process, for the channel roughness, the uncertainty bound ~~gets~~ is stabilized toward the end of the assimilation period. It is also evident from Figure 3.a that the bathymetry is a more identifiable parameter compared to the channel roughness as it shows the fastest convergence with a minimum degree of uncertainty. However, the channel roughness is less identifiable with the slowest convergence. ~~The scatter plots illustrate the evolution of parameter space at six different time segments.~~ In Figure 3.b, ~~The~~ the first day (t=1) includes all 100 ensemble members of parameters and day 30 corresponds to the highest discharge and water stage of flooding when the model parameters reach the highest improvement and get closer to the true value. Figure 3.b shows that

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

both model parameters are converging toward the true values as the assimilation proceeds. This indicates the efficacy and usefulness of the proposed DA-hydrodynamic modeling framework developed in this study.

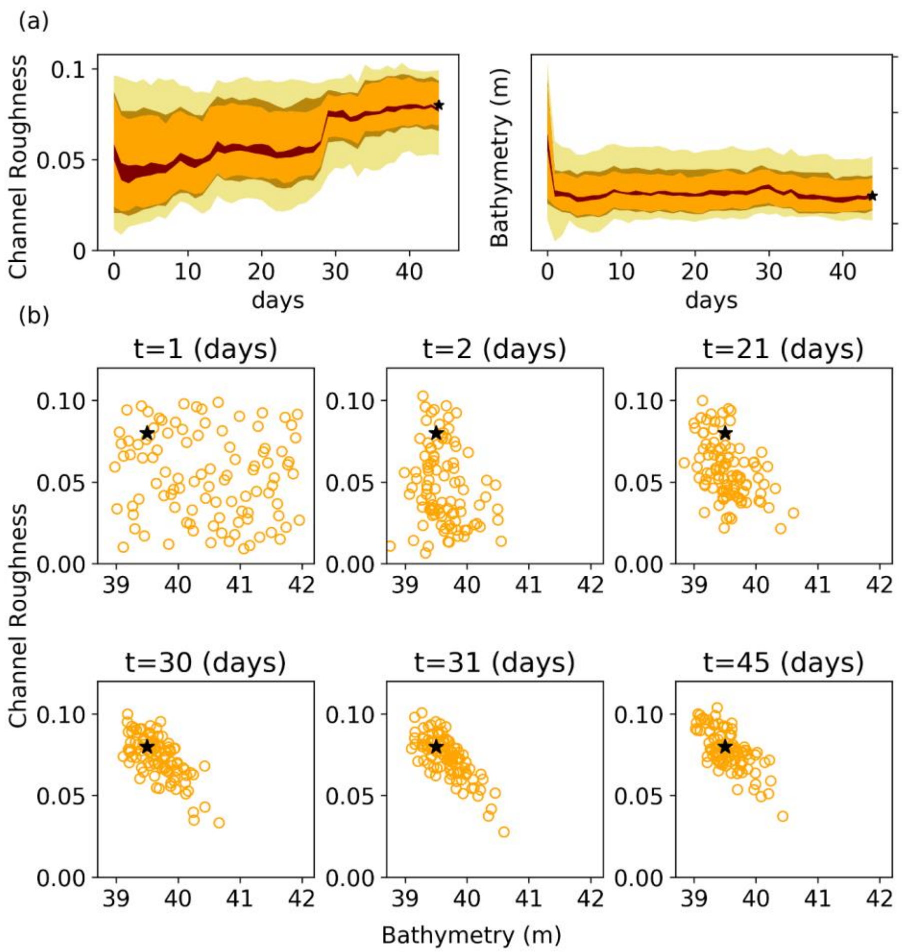

Figure 3. Temporal evolution of the LISFLOOD parameters for the synthetic experiment during Hurricane Harvey using the EnKF. (a) Temporal evolution of model parameter predictive intervals (shaded areas) corresponding to 95, 75, 68, and 10 percentile (b) Temporal evolution

of particle positions in the model parameter space at six different days during the Hurricane. ~~The shaded areas correspond to 95, 75, 68, and 10 percentile predictive intervals, and t~~The black stars at the end of each parameter subplot represent the true parameter values.

4.2 Experiment 2: Real Case Study

In the real experiment, we assimilate the discharge and water stage readings from two ~~internal~~ USGS gauges into the LISFLOOD-FP model. We also run the OL simulation and calculate the ensemble mean to predict the discharge and water stage at these two gauges. Figure 4 presents a comparison of simulated discharge (Figures 4a and~~,~~ 4b) and water stage (Figures 4c and~~,~~ 4d) with observations using both OL and our EnKF-based approach. Figures 4a and 4c are the prior estimates of discharge and water stage, while Figures 4b and 4d show their posterior distributions~~s~~ ~~which~~ that reflect the updated variables after assimilating the observations into the model. It is worth mentioning that although prior distributions~~s~~ represent the results before assimilating new observations into the model, their values are dependent on the initial conditions updated from observations in the previous time step. ~~In this study, s~~Since forecasting (1-day lead time) is the main objective of DA-hydrodynamic modeling framework, we specifically focus on behavior of priors. As can be seen, the simulated peak discharge by the OL is highly overestimated by around 200 ~~cms~~ m^3/s while assimilating the observations improve the results so that their difference with observation is less than 50 m^3/s ~~cms~~ at the peak of the flood (*KGE* =0.76 and *RMSE*=40.9 m^3/s~~cms~~)). In contrast, the simulated water stage in Figures 4c and 4d are underestimated by OL by around 2 meters at the peak. Compared to the OL, ~~Using~~ using the developed EnKF approach raises the peak of water stage at peak and reduces the errors significantly (*KGE*=0.96 and *RMSE*=0.5 m). The accurate estimates of prior discharge and water stage confirm the applicability

Formatted: Superscript

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

of the proposed assimilation framework in forecast mode when real-time flood warning and decision making is the priority. The *NRR* measure for the prior discharge and water stage are 1.17 and 0.65, showing that the uncertainty bound is underestimated and overestimated, respectively. The ~~reliability~~ *Reliability* of both variables is above 70 percent since the uncertainty bounds encompass the observations for almost the entire simulation period.

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

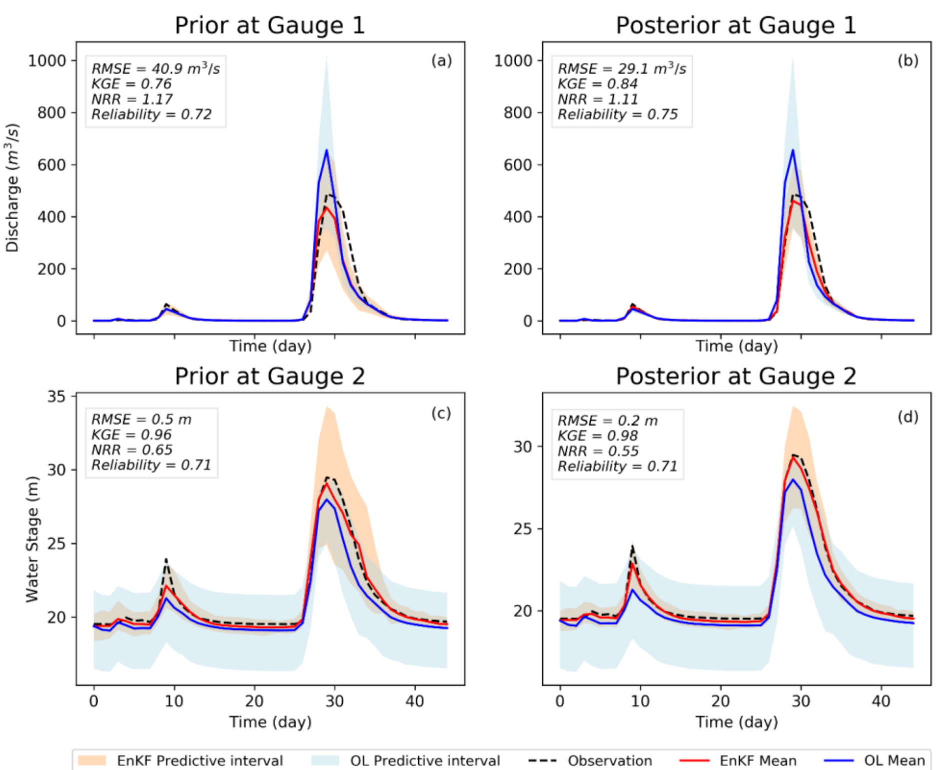

Figure 4 Simulation results of LISFLOOD-FP for the real experiment during Hurricane Harvey using the EnKF and open-loop. (a) Prior simulated discharge at gauge 1 (b) Posterior simulated discharge at gauge 1 (c) Prior simulated water stage at gauge 2 (d) Posterior simulated water

stage at gauge 2. The shaded areas represent the predictive interval of simulated discharge and water stage by EnKF.

Figure 5 illustrates the prior and posterior distributions of discharge and water stage in the beginning, peak, and ending days of Hurricane Harvey flood. In all three days, the uncertainty bounds of both discharge and water stage are narrowed down by assimilating the observations so that posterior distributions are more precise compared to the priors. In the beginning and ending days (Aug 26 and Sep 1), the mean of prior distributions is substantially shifted toward truth in the posterior distributions. Figure 5 reveals that our developed approach provides more accurate and reliable posterior discharge and water stage distributions compared to prior distributions where the simulations are either overestimated or underestimated. It is noted that, on August 28 (day of flood peak), although the prior distributions accurately represent the observation, they have a wide uncertainty bound. After correcting/updating the model state variables and parameters, as posterior distributions show, the uncertainty bound is reduced while the ensemble mean remains closer to the observation.

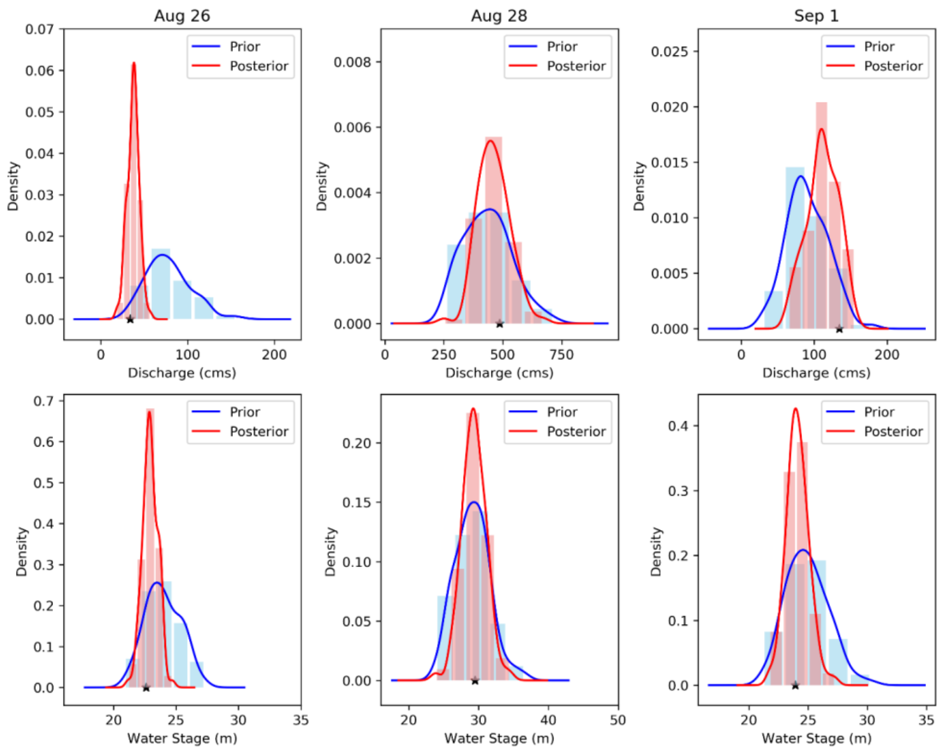

Figure 5. Prior and posterior distribution of discharge (a,b,c) and water stage (c,d,f) at the beginning (Aug 26), peak (Aug 28), and ending (Sep1) days of Hurricane Harvey using the EnKF

4.3 Probabilistic Flood Inundation Mapping

In this ~~study~~section, we ~~propose a DA hydrodynamic modeling framework to~~ account for the uncertainties involved in flood modeling and generate real-time probabilistic flood inundation maps. Since the majority of flooding conditions occurred within ~~6~~ six days from August 27-Sep 1, we display the spatial distribution of water depth in this period and provide probabilistic flood inundation maps using both OL and our developed approach (~~see~~ Figures 6 and 7). Figure 6 represents the first three days of Harvey, which corresponds to the upper limb of the flood hydrograph. On August 27, the major difference between the OL and EnKF appears in the regions around the upstream of the lower channel where the EnKF provides a more reliable prediction of the inundated area. Moving toward the peak of flood on Aug 29, the OL generates a large region of uncertain cells around the banks of the upper channel, while both the extent and density of uncertain values in the probabilistic maps generated by the EnKF is smaller during the peak of Harvey.

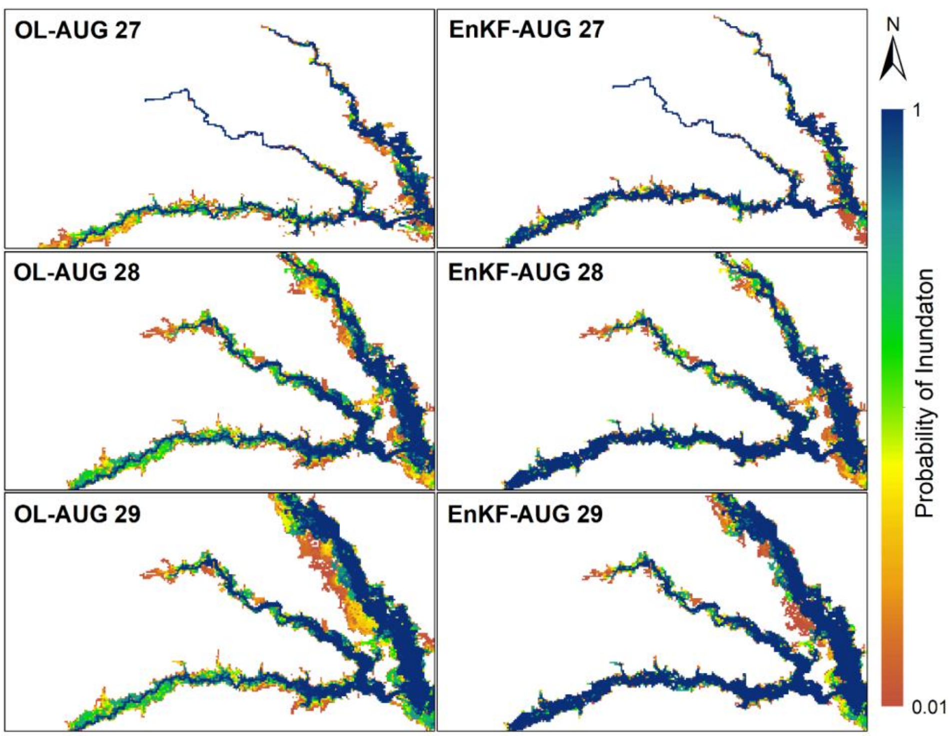

Figure 6 Probabilistic flood inundation maps generated by OL and EnKF techniques to simulate the upper limb of Harvey flood hydrograph from Aug 27 to Aug 29.

Figure 7 shows the probabilistic inundation areas in the last three days corresponding to the lower limb of the flood hydrograph. ~~In this figure, t~~The discrepancies between the OL and EnKF flood maps increase showing that performing DA is more effective in improving the inundation mapping skill from peak to ending point of the flood hydrograph. A large number of inundated cells generated by the OL are vanished after the peak of Harvey which results in a set of scattered discontinuous maps in Aug 31 and Sep 1. On the other hand, the probabilistic maps generated by the EnKF maintain their continuous shapes so that the probability of inundation is reduced without changing the extent. The merit of the EnKF in improving the flood inundation areas at the lower limb of the flood hydrograph agrees with results in Figures 4c and 4d where the EnKF widens the simulated water stage hydrographs and removes the lag difference that exists between the open-loop and observations.

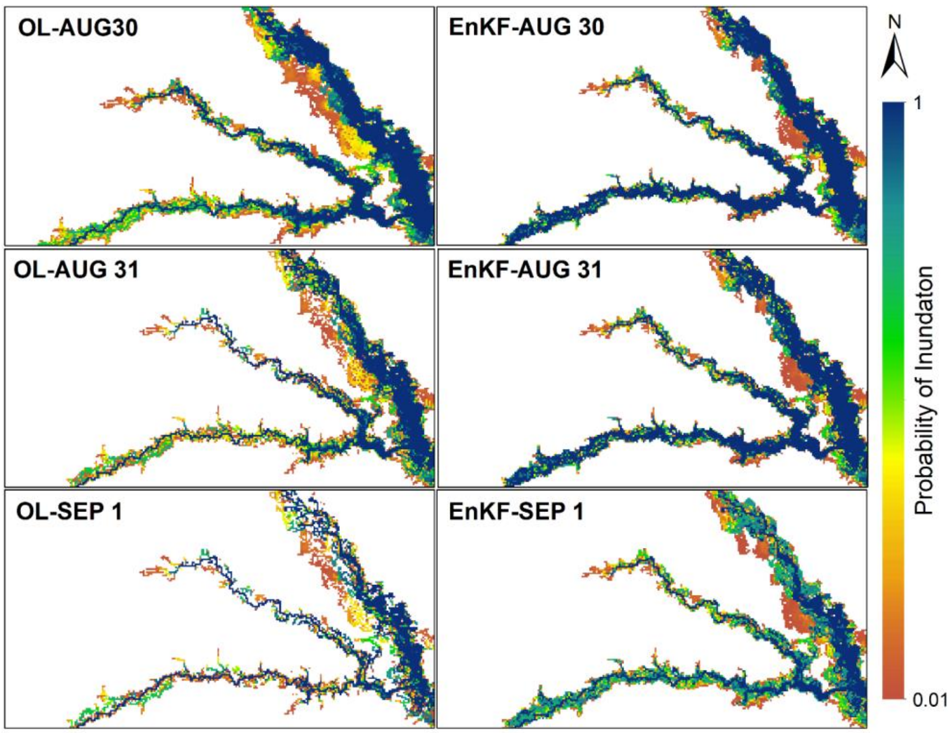

Figure 7 Probabilistic flood inundation maps generated by OL and EnKF techniques to simulate the lower limb of Harvey flood hydrograph from Aug 30 to Sep 1.

Finally, to quantify the performance of EnKF and OL for generating a spatial distribution of water depth over the domain, we illustrate the ROC graphs, the *AUC* values, and Fit indices in Figure 8. To calculate these measures, we ignore the temporal distributions and only report the maximum inundation maps that represent the union of flooded areas over the entire period of Harvey. Comparing the EnKF and OL in Figure 8.a, the EnKF line (blue) is closer to the northwest of the *rfp-rtp* space where its *AUC* is 5% higher than the OL approach. In Figure 8.b, each point represents the *F*~~it~~ indices for the OL and the EnKF approaches corresponding to a given threshold. Using ~~hundred number of~~100 ~~-~~thresholds that ~~each ranging~~range from [0.01,1], the probabilistic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

maps are converted to 100 deterministic maps and the *F*~~it~~ indices are calculated. The position of scatters above the dash line confirms the EnKF outperforms the OL. In addition to these measures, the [*UFI*, *OFI*] indices calculated for OL and EnKF approaches are [30.3, 0.26] %, and [23.4, 0.4]% respectively. The low values of *OFI* for both approaches (< 1%) show that the simulations mostly underestimate the flood inundation areas. In addition, comparing the indices of both approaches reveal that the EnKF reduces the overall underestimation by around 7%.

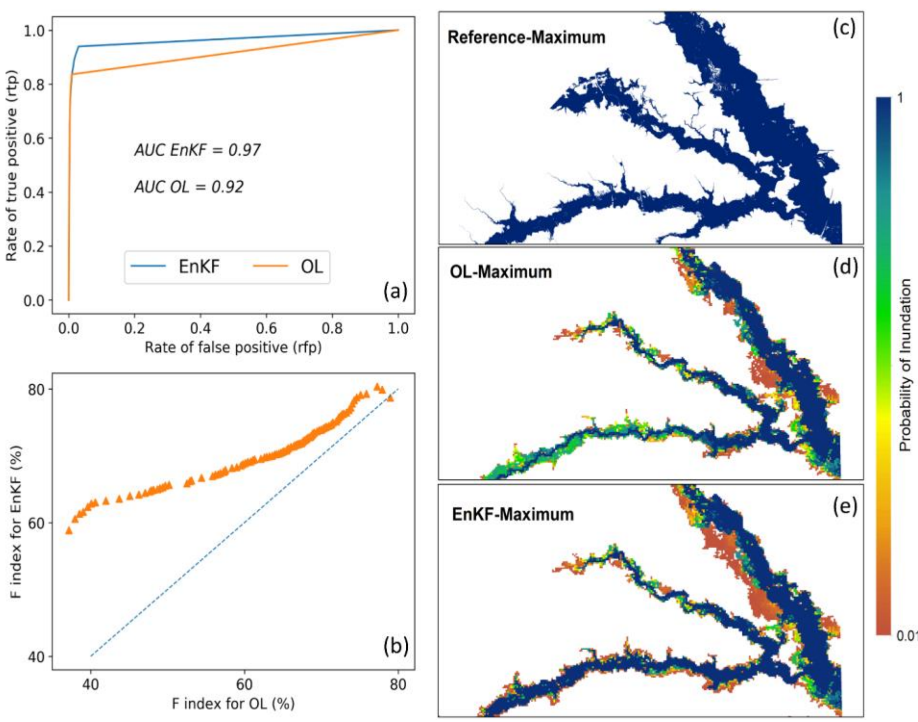

Figure 8 The Receiver Operating Curves (ROC) indicating the performance of OL and EnKF techniques for probabilistic flood inundation mapping

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Font: Italic, Complex Script Font: Italic

Formatted: Normal, Left, Line spacing: single

Formatted: Font: (Default) +Body (Calibri), 11 pt, Complex Script Font: +Body CS (Arial), 11 pt

5. Discussion and Conclusions

The main motivation in this study is to propose a DA-hydrodynamic modeling framework for real-time probabilistic flood inundation mapping. Considering the coarse spatiotemporal resolution of satellite data for capturing the ~~water surface elevation~~WSE, assimilating them into the hydrodynamic models may not be a practical solution for an upcoming flood event. On the other hand, the availability of daily discharge and ~~water surface elevation~~WSE data at gauge stations is a great opportunity to establish a multivariate DA-hydrodynamic modeling framework that updates the initial condition of modeling at daily scale and forecast the flood inundation areas at 1 day lead time. Here, we used the EnKF data assimilation method in conjunction with a hydrodynamic model to account for different sources of uncertainties involved in different layers of model simulations, including the boundary conditions, model parameters, and initial condition,~~–~~ and generate real-time probabilistic flood inundation maps~~-~~. To further enhance the performance of the developed framework, the discharge and water stage at two different gauges are simultaneously assimilated into the LISFLOOD-FP model. The multivariate EnKF approach considers the correlation between discharge at two gauges and between discharge and ~~water surface elevation~~WSE at one gauge using a modified covariance matrix and Kalman gain equation.

In the synthetic experiment, we examined the convergence of model parameters toward truth and found that the proposed DA-hydrodynamic modeling framework can be successfully used to improve the accuracy and reliability of model predictions while accounting for uncertainties associated with model parameters. The channel roughness coefficient varied more ~~-~~rapidly than the bathymetry during the temporal evolutions of these parameters showing the better iden~~m~~tifiability of this parameter. The validation results of the real experiment revealed that the assimilation with the EnKF approach improves the model predictions ~~at~~ across temporal and spatial scales (i.e., discharge and water stage time series at gauges and flood maps showing the maximum water depth over the simulation period). These improvements are more pronounced during the falling limb of the flood hydrograph where the EnKF widens the simulated hydrograph and removes the existing lag compared to the observations. Similarly, the simulated flood inundation maps confirm that the OL provides discontinuous scattered maps during the flood recession period while the EnKF provides a more accurate representation of the inundation areas. The validation results also demonstrate that the EnKF reduces the underestimation by 7% and outperformed the OL approach by around 5% for probabilistic flood inundation mapping.

For real-time flood inundation mapping, timely decision making is of paramount importance. The time between the issuance of the warning and the occurrence of the flood is typically a short period less than a day. Additionally, the flood waves propagate, inundate the affected regions and cause damages rapidly. Thus, the main requirement for real-time probabilistic inundation mapping is to develop a fast and efficient modeling framework that is beneficial for decision makers and emergency managers. Considering the high computational expense of hydrodynamic models and the need for generating a multitude of simulations in the probabilistic fashion, this study uses a coarse resolution 120m DEM to maintain the efficiency of the modeling and meet the requirements for practical benefits. In this study, the DA-hydrodynamic modeling framework is executed on the University of Alabama High Performance Computing (UAHPC) cluster. Considering the ensemble size of 100, we submit a job array with 100 cores where each core is assigned to a specific member of the DA-hydrodynamic modeling simulation. The efficient hydrodynamic model setup with coarse resolution DEM helps ~~to~~ simulate the Harvey and generate probabilistic results in 4-5 hours (~ 4 hours for the hydrodynamic simulation and ~20 minutes for the DA). Applying this computationally efficient framework is highly beneficial, specially for the emergency response agencies (e.g. FEMA), insurance companies, Water Centers, and other private companies that need to forecast the inundation areas and take timely decisions a few hours before the onset of floods.

~~(Savage et al., 2016)~~

~~To simulate flood hazards during the emergency of an upcoming flood event, using an efficient flood modeling framework is of paramount importance~~

~~.~~The coarse DEM used in this study cannot perfectly represent the watershed topography and bathymetry, and can be the main reason for ~~the~~ underestimation of inundation areas (F index less than 80%). Savage et al., (2016) investigated the impacts of DEM resolution on the accuracy and efficiency of probabilistic flood inundation maps generated with the LISFLOOD-FP model. They demonstrated that models with resolution less than 50 offer little gain in performance yet are more than an order of magnitude computationally expensive which can become infeasible when undertaking probabilistic analysis. They also found that the reliability of flood maps deteriorates at resolutions coarser than 100 m. Considering the medium scale of our study (> 100 km river) compared to the reach scale (~10 km river) of the work by Savage et al., (2016), here we slightly increased their suggested threshold for the DEM and demonstrated that the accuracy of results is still acceptable.

~~However, a simplified model setup (i.e. using coarse resolution DEM,~~The simulation of an extreme flooding condition such as Hurricane Harvey with a simplified model setup (i.e. using a coarse DEM, assuming uniform roughness coefficient for channel and floodplain, and estimating bathymetry by lowering DEM with one parameter)~~)~~ ~~for efficient flood modeling~~ is prone to losing accuracy. ~~Particularly, for an extreme flooding condition such as Hurricane Harvey, the simplified modeling may pose significant errors.~~ The results obtained from the simulation of the real experiment demonstrated that, despite using a simplified efficient modeling setup, we can still simulate the discharge, water stage, and inundation areas for an extreme flood event with an acceptable accuracy while accounting for uncertainties involved in model predictions. This shows that assimilating the gauge data into a simplified model setup improves the accuracy, and provides an efficient probabilistic framework for real-time flood inundation mapping that considers potential sources of uncertainties in different layers of modeling.

The time dependency that exists between the upstream and downstream gauges along a channel can affect the performance of multivariate assimilation with those gauges. For future studies, using a more advanced DA technique that fully characterizes the model structural uncertainty (Abbaszadeh et al., 2019), and considering the time lag dependency between multiple gauges can improve the performance of modeling and provide more realistic assimilation of the hydrodynamic models. Another limitation of this study is the simple assumptions made for perturbing the initial condition (water depth), parameters (channel roughness and river bathymetry) and observations (WSE and discharge). More investigation on the physically meaningful distribution of these values can enhance the performance of the DA-hydrodynamic modeling framework in future studies. A joint assimilation of point source gauges and remotely sensed data can also improve the reliability and accuracy of the results. Finally, proposing a DA-hydrodynamic modeling framework that considers the DEM and channel width uncertainty can provide a more comprehensive uncertainty quantification for probabilistic flood inundation mapping in future studies.

An advantage of the proposed DA-hydrodynamic modeling framework is its generic format so that other studies can follow the flowchart in Figure~~.~~ 2 and use information in Section 3.2 and 3.3 to set up the hydrodynamic model and the EnKF algorithm, respectively. To properly apply this framework to other studies, first, the point source observations of WSE and discharge should be available at daily/sub-daily scales. ~~In other words, the proposed framework cannot be implemented in ungauged basins.~~ Second, the modeler should have access to high performance computing facilities for parallel simulation of ensemble members. Third, the hydrodynamic model should be sequentially executed within the DA algorithm. The modeler should check the hydrodynamic model manual and make sure that the outputs and initial conditions can be up~~gra~~dated in a sequential manner. Taking these three considerations into account, the proposed DA-hydrodynamic modeling framework can be applied to any other study areas that are prone to frequent flooding and provide a robust and generic tool for real-time probabilistic flood inundation mapping.

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt

Data availability

All the data used in this study, including the gauge streamflow and water stage data and the DEMs, are publicly available from the USGS website and National Elevation Dataset (NED). The reference flood maps provided for Hurricane Harvey is available from the USGS report at https://pubs.usgs.gov/sir/2018/5070/sir20185070.pdf.

Author contribution

KJ, PA and HM conceptualized the study and designed the synthetic and real experiments. KJ developed, set up, evaluated and implemented the DA-hydrodynamic modeling framework for both experiments. PA and HM provided guidance on the assimilation experiments. KJ wrote the first draft of the manuscript. HM and PA provided comments and edited the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

Partial financial support for this study was provided by the USACE contract #W912HZ2020055. We would like to thank the anonymous reviewers for their constructive comments on the original version of the manuscript.

References

Abbaszadeh, P., Gavahi, K., Moradkhani, H., 2020. Multivariate remotely sensed and in-situ data assimilation for enhancing community WRF-Hydro model forecasting. Adv. Water Resour. 145, 103721.

Abbaszadeh, P., Moradkhani, H., Daescu, D.N., 2019. The Quest for Model Uncertainty Quantification: A Hybrid Ensemble and Variational Data Assimilation Framework. Water Resour. Res. 55, 2407–2431. https://doi.org/10.1029/2018WR023629

Abbaszadeh, P., Moradkhani, H., Yan, H., 2018. Enhancing hydrologic data assimilation by evolutionary Particle Filter and Markov Chain Monte Carlo. Adv. Water Resour. 111, 192–204. https://doi.org/10.1016/j.advwatres.2017.11.011

Ahmadisharaf, E., Kalyanapu, A.J., Bates, P.D., 2018. A probabilistic framework for floodplain mapping using hydrological modeling and unsteady hydraulic modeling. Hydrol. Sci. J. 63, 1759–1775. https://doi.org/10.1080/02626667.2018.1525615

Alemohammad, S.H., McLaughlin, D.B., Entekhabi, D., 2015. Quantifying precipitation uncertainty for land data assimilation applications. Mon. Weather Rev. 143, 3276–3299.

Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., Feyen, L., 2014. Advances in pan-European flood hazard mapping. Hydrol. Process. 28, 4067–4077. https://doi.org/10.1002/hyp.9947

Anderson, J.L., Anderson, S.L., 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon. Weather Rev. 127, 2741–2758.

Aronica, G., Bates, P.D., Horritt, M.S., 2002a. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. Hydrol. Process. 16, 2001–2016. https://doi.org/10.1002/hyp.398

Aronica, G., Bates, P.D., Horritt, M.S., 2002b. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. Hydrol. Process. 16, 2001–2016. https://doi.org/10.1002/hyp.398

Aronica, G.T., Franza, F., Bates, P.D., Neal, J.C., 2012. Probabilistic evaluation of flood hazard in urban areas using Monte Carlo simulation. Hydrol. Process. 26, 3962–3972. https://doi.org/10.1002/hyp.8370

Formatted: Font: 12 pt, Complex Script Font: 12 pt

Azimi, S., Dariane, A.B., Modanesi, S., Bauer-Marschallinger, B., Bindlish, R., Wagner, W., Massari, C., 2020. Assimilation of Sentinel 1 and SMAP–based satellite soil moisture retrievals into SWAT hydrological model: The impact of satellite revisit time and product spatial resolution on flood simulations in small basins. J. Hydrol. 581, 124367.

Bales, J.D., Wagner, C.R., 2009. Sources of uncertainty in flood inundation maps. J. Flood Risk Manag. 2, 139–147. https://doi.org/10.1111/j.1753-318X.2009.01029.x

Bates, P.D., De Roo, A.P.J., 2000. A simple raster-based model for flood inundation simulation. J. Hydrol. 236, 54–77. https://doi.org/10.1016/S0022-1694(00)00278-X

Brêda, J.P.L.F., Paiva, R.C.D., Bravo, J.M., Passaia, O.A., Moreira, D.M., 2019. Assimilation of Satellite Altimetry Data for Effective River Bathymetry. Water Resour. Res. 55, 7441–7463. https://doi.org/10.1029/2018WR024010

Clark, M.P., Hay, L.E., 2004. Use of Medium-Range Numerical Weather Prediction Model Output to Produce Forecasts of Streamflow. J. Hydrometeorol. 5, 15–32. https://doi.org/10.1175/1525-7541(2004)005<0015:UOMNWP>2.0.CO;2

Courtier, P., Derber, J., Errico, R.O.N., Louis, J.-F., VukiĆEviĆ, T., 1993. Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. Tellus Dyn. Meteorol. Oceanogr. 45, 342–357.

Cuo, L., Pagano, T.C., Wang, Q.J., 2011. A Review of Quantitative Precipitation Forecasts and Their Use in Short- to Medium-Range Streamflow Forecasting. J. Hydrometeorol. 12, 713–728. https://doi.org/10.1175/2011JHM1347.1

DeChant, C.M., Moradkhani, H., 2014. Toward a reliable prediction of seasonal forecast uncertainty: Addressing model and initial condition uncertainty with ensemble data assimilation and sequential Bayesian combination. J. Hydrol. 519, 2967–2977.

Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrol. Earth Syst. Sci. 13, 913–921. https://doi.org/10.5194/hess-13-913-2009

Di Baldassarre, G., Schumann, G., Bates, P.D., 2009. A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent. J. Hydrol. 367, 276–282. https://doi.org/10.1016/j.jhydrol.2009.01.020

Domeneghetti, A., Vorogushyn, S., Castellarin, A., Merz, B., Brath, A., 2013. Probabilistic flood hazard mapping: effects of uncertain boundary conditions. Hydrol. Earth Syst. Sci. 17, 3127–3140.

Durand, M., Andreadis, K.M., Alsdorf, D.E., Lettenmaier, D.P., Moller, D., Wilson, M., 2008. Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. Geophys. Res. Lett. 35.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett., ROC Analysis in Pattern Recognition 27, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Gavahi, K., Abbaszadeh, P., Moradkhani, H., Zhan, X., Hain, C., 2020. Multivariate Assimilation of Remotely Sensed Soil Moisture and Evapotranspiration for Drought Monitoring. J. Hydrometeorol. 21, 2293–2308. https://doi.org/10.1175/JHM-D-20-0057.1

Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza Guingla, D.A., Pauwels, V., De Lannoy, G., De Keyser, R., Pfister, L., Hoffmann, L., 2011. Assimilating SAR-derived water level data into a hydraulic model: a case study. Hydrol. Earth Syst. Sci. 15, 2349–2365.

Habets, F., LeMoigne, P., Noilhan, J., 2004. On the utility of operational precipitation forecasts to served as input for streamflow forecasting. J. Hydrol. 293, 270–288. https://doi.org/10.1016/j.jhydrol.2004.02.004
Hall, J.W., Tarantola, S., Bates, P.D., Horritt, M.S., 2005. Distributed Sensitivity Analysis of Flood Inundation Model Calibration. J. Hydraul. Eng. 131, 117–126. https://doi.org/10.1061/(ASCE)0733-9429(2005)131:2(117)
Horritt, M.S., 2006. A methodology for the validation of uncertain flood inundation models. J. Hydrol. 326, 153–165. https://doi.org/10.1016/j.jhydrol.2005.10.027
Horritt, M.S., Bates, P.D., 2002. Evaluation of 1D and 2D numerical models for predicting river flood inundation. J. Hydrol. 268, 87–99. https://doi.org/10.1016/S0022-1694(02)00121-X
Hostache, R., Chini, M., Giustarini, L., Neal, J., Kavetski, D., Wood, M., Corato, G., Pelich, R.-M., Matgen, P., 2018. Near-Real-Time Assimilation of SAR-Derived Flood Maps for Improving Flood Forecasts. Water Resour. Res. 54, 5516–5535. https://doi.org/10.1029/2017WR022205
Hostache, R., Lai, X., Monnier, J., Puech, C., 2010. Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of Mosel River. J. Hydrol. 390, 257–268. https://doi.org/10.1016/j.jhydrol.2010.07.003
Jafarzadegan, K., Merwade, V., 2017. A DEM-based approach for large-scale floodplain mapping in ungauged watersheds. J. Hydrol. 550, 650–662. https://doi.org/10.1016/j.jhydrol.2017.04.053
Jafarzadegan, K., Merwade, V., Saksena, S., 2018. A geomorphic approach to 100-year floodplain mapping for the Conterminous United States. J. Hydrol. 561, 43–58. https://doi.org/10.1016/j.jhydrol.2018.03.061
Kumar, S.V., Dong, J., Peters-Lidard, C.D., Mocko, D., Gómez, B., 2017. Role of forcing uncertainty and background model error characterization in snow data assimilation. Hydrol. Earth Syst. Sci. 21, 2637–2647.
Leach, J.M., Kornelsen, K.C., Coulibaly, P., 2018. Assimilation of near-real time data products into models of an urban basin. J. Hydrol. 563, 51–64.
Lee, H., Seo, D.-J., Koren, V., 2011. Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the data and initial model soil moisture states. Adv. Water Resour. 34, 1597–1615.
Lievens, H., Reichle, R.H., Liu, Q., De Lannoy, G.J., Dunbar, R.S., Kim, S.B., Das, N.N., Cosh, M., Walker, J.P., Wagner, W., 2017. Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. Geophys. Res. Lett. 44, 6145–6153.
Maidment, D.R., 2017. Conceptual Framework for the National Flood Interoperability Experiment. JAWRA J. Am. Water Resour. Assoc. 53, 245–257. https://doi.org/10.1111/1752-1688.12474
Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V.R.N., De Lannoy, G., De Keyser, R., Savenije, H.H.G., 2010a. Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept. Hydrol. Earth Syst. Sci. 14, 1773–1785.
Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V.R.N., De Lannoy, G., De Keyser, R., Savenije, H.H.G., 2010b. Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept. Hydrol. Earth Syst. Sci. 14, 1773–1785.

Merwade, V., Cook, A., Coonrod, J., 2008. GIS techniques for creating river terrain models for hydrodynamic modeling and flood inundation mapping. Environ. Model. Softw. 23, 1300–1311. https://doi.org/10.1016/j.envsoft.2008.03.005

Moradkhani, H., Hsu, K.-L., Gupta, H., Sorooshian, S., 2005a. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. Water Resour. Res. 41.

Moradkhani, H., Nearing, G.S., Abbaszadeh, P., Pathiraja, S., 2019. Fundamentals of data assimilation and theoretical advances. Handb. Hydrometeorol. Ensemble Forecast. Springer Berl. Heidelb. Berl. Heidelb. 675–699.

Moradkhani, H., Sorooshian, S., Gupta, H.V., Houser, P.R., 2005b. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. Adv. Water Resour. 28, 135–147. https://doi.org/10.1016/j.advwatres.2004.09.002

Munier, S., Polebistki, A., Brown, C., Belaud, G., Lettenmaier, D.P., 2015. SWOT data assimilation for operational reservoir management on the upper Niger River Basin. Water Resour. Res. 51, 554–575. https://doi.org/10.1002/2014WR016157

Neal, J., Keef, C., Bates, P., Beven, K., Leedal, D., 2013. Probabilistic flood risk mapping including spatial dependence. Hydrol. Process. 27, 1349–1363. https://doi.org/10.1002/hyp.9572

Neal, J., Schumann, G., Bates, P., 2012. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. Water Resour. Res. 48. https://doi.org/10.1029/2012WR012514

Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., Pappenberger, F., 2009. A data assimilation approach to discharge estimation from space. Hydrol. Process. 23, 3641–3649. https://doi.org/10.1002/hyp.7518

Papaioannou, G., Vasiliades, L., Loukas, A., Aronica, G.T., 2017. Probabilistic flood inundation mapping at ungauged streams due to roughness coefficient uncertainty in hydraulic modelling. Adv. Geosci. 44, 23–34.

Pappenberger, F., Beven, K.J., Ratto, M., Matgen, P., 2008. Multi-method global sensitivity analysis of flood inundation models. Adv. Water Resour. 31, 1–14. https://doi.org/10.1016/j.advwatres.2007.04.009

Pathiraja, S., Moradkhani, H., Marshall, L., Sharma, A., Geenens, G., 2018. Data-driven model uncertainty estimation in hydrologic data assimilation. Water Resour. Res. 54, 1252–1280.

Pauwels, V.R., Hoeben, R., Verhoest, N.E., De Troch, F.P., 2001. The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation. J. Hydrol. 251, 88–102.

Pedinotti, V., Boone, A., Ricci, S., Biancamaria, S., Mognard, N., 2014. Assimilation of satellite data to optimize large-scale hydrological model parameters: a case study for the SWOT mission. Hydrol. Earth Syst. Sci. 18, 4485–4507. https://doi.org/10.5194/hess-18-4485-2014

Pedrozo-Acuña, A., Rodríguez-Rincón, J.P., Arganis-Juárez, M., Domínguez-Mora, R., Villareal, F.J.G., 2015. Estimation of probabilistic flood inundation maps for an extreme event: Pánuco River, México. J. Flood Risk Manag. 8, 177–192. https://doi.org/10.1111/jfr3.12067

Pelletier, P.M., 1988. Uncertainties in the single determination of river discharge: a literature review. Can. J. Civ. Eng. 15, 834–850. https://doi.org/10.1139/l88-109

Pinter, N., Santos, N., Hui, R., 2017. Preliminary analysis of Hurricane Harvey flooding in Harris County, Texas. Retrieved UC Davis Cent. Watershed Sci.-Calif. WaterBlog Httpscaliforniawaterblog Com20170901preliminary-Anal.--Hurricaneharvey-Flooding--Harris-Cty.-Tex.

Purvis, M.J., Bates, P.D., Hayes, C.M., 2008. A probabilistic methodology to estimate future coastal flood risk due to sea level rise. Coast. Eng. 55, 1062–1073. https://doi.org/10.1016/j.coastaleng.2008.04.008

Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002. Hydrologic data assimilation with the ensemble Kalman filter. Mon. Weather Rev. 130, 103–114.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks, S.W., 2010a. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. Water Resour. Res. 46. https://doi.org/10.1029/2009WR008328

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks, S.W., 2010b. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. Water Resour. Res. 46. https://doi.org/10.1029/2009WR008328

Romanowicz, R., Beven, K., 2003. Estimation of flood inundation probabilities as conditioned on event inundation maps. Water Resour. Res. 39. https://doi.org/10.1029/2001WR001056

Sangwan, N., Merwade, V., 2015. A Faster and Economical Approach to Floodplain Mapping Using Soil Information. JAWRA J. Am. Water Resour. Assoc. 51, 1286–1304. https://doi.org/10.1111/1752-1688.12306

Savage, J.T.S., Bates, P., Freer, J., Neal, J., Aronica, G., 2016. When does spatial resolution become spurious in probabilistic flood inundation predictions? Hydrol. Process. 30, 2014–2032. https://doi.org/10.1002/hyp.10749

Sheffield, J., Pan, M., Wood, E.F., Mitchell, K.E., Houser, P.R., Schaake, J.C., Robock, A., Lohmann, D., Cosgrove, B., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Ramsay, B.H., 2003. Snow process modeling in the North American Land Data Assimilation System (NLDAS): 1. Evaluation of model-simulated snow cover extent. J. Geophys. Res. Atmospheres 108. https://doi.org/10.1029/2002JD003274

Slater, A.G., Clark, M.P., 2006. Snow data assimilation via an ensemble Kalman filter. J. Hydrometeorol. 7, 478–493.

Tayefi, V., Lane, S.N., Hardy, R.J., Yu, D., 2007. A comparison of one- and two-dimensional approaches to modelling flood inundation over complex upland floodplains. Hydrol. Process. 21, 3190–3202. https://doi.org/10.1002/hyp.6523

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., Kim, S., 2017. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. Environ. Model. Softw. 90, 201–216. https://doi.org/10.1016/j.envsoft.2017.01.006

~~Vrugt, J.A., Gupta, H.V., Nualláin, B., Bouten, W., 2006. Real-time data assimilation for operational ensemble streamflow forecasting. J. Hydrometeorol. 7, 548–565.~~

Xu, L., Abbaszadeh, P., Moradkhani, H., Chen, N., Zhang, X., 2020. Continental drought monitoring using satellite soil moisture, data assimilation and an integrated drought index. Remote Sens. Environ. 250, 112028. https://doi.org/10.1016/j.rse.2020.112028

Xu, X., Zhang, X., Fang, H., Lai, R., Zhang, Y., Huang, L., Liu, X., 2017. A real-time probabilistic channel flood-forecasting model based on the Bayesian particle filter approach. Environ. Model. Softw. 88, 151–167. https://doi.org/10.1016/j.envsoft.2016.11.010

Formatted: Normal, Indent: Before: 0", Hanging: 0.5"

Yoon, Y., Durand, M., Merry, C.J., Clark, E.A., Andreadis, K.M., Alsdorf, D.E., 2012a. Estimating river bathymetry from data assimilation of synthetic SWOT measurements. J. Hydrol. 464, 363–375.

Yoon, Y., Durand, M., Merry, C.J., Clark, E.A., Andreadis, K.M., Alsdorf, D.E., 2012b. Estimating river bathymetry from data assimilation of synthetic SWOT measurements. J. Hydrol. 464, 363–375.

Zhang, Q., Shi, L., Holzman, M., Ye, M., Wang, Y., Carmona, F., Zha, Y., 2019. A dynamic data-driven method for dealing with model structural error in soil moisture data assimilation. Adv. Water Resour. 132, 103407.