

1 **Sequential Data Assimilation for Real-Time Probabilistic**
2 **Flood Inundation Mapping**

3 Keighobad Jafarzadegan* , Peyman Abbaszadeh, Hamid Moradkhani

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

6 *Corresponding author: kjafarzadegan@ua.edu

7
8
9
10
11
12
13
14
15
16
17
18

19 **Abstract**

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

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

40

41

42

43

44

45

46 **1. Introduction**

47 The on-time, accurate, and reliable characterization of an upcoming flood event is imperative for
48 proper decision making and risk analysis. A well-calibrated hydrologic model coupled with
49 reliable weather forecast models can be used to generate the streamflow forecast (Clark and Hay,
50 2004; Cuo et al., 2011; Habets et al., 2004). While streamflow forecasting during flood events is
51 indispensable, the critical step for flood risk analysis is to estimate the flood inundation areas
52 corresponding to the forecasted streamflow of a potential upcoming event. Hydrodynamic models
53 are common tools used to simulate the physics of a river system and predict the spatiotemporal
54 distribution of water surface elevation (WSE). The predicted WSE can be simply converted to
55 water depth and inundation area by overlaying with a high-resolution Digital Elevation Model
56 (DEM) (Merwade et al., 2008; Teng et al., 2017). Since floods happen in a short period and at a
57 certain location, it is often not possible to find an appropriate remote sensing image that covers
58 those inundated areas during the flood period. This is the main reason that research on flood
59 inundation mapping is mostly limited to post-event analysis where specific study areas with
60 available remote sensing data are used as testbeds.

61 Federal Emergency Management Agency (FEMA) is the leading agency in the United States that
62 provides flood hazard and risk maps over the Contiguous United States (CONUS). While these
63 maps display flood-prone areas corresponding to specific return periods (e.g. 100 and 500-year
64 events), they are not always reliable for an upcoming flood event. For example, FEMA 100-year
65 and 500-year flood hazard maps covered only one-third and half of the inundated areas induced
66 by Hurricane Harvey in Harris County, Texas, respectively (Pinter et al., 2017). The National
67 Water Center Innovators Program proposed the idea of real-time flood inundation mapping across
68 the United States in 2015 (Maidment, 2017). It highlighted the importance of event-based flood

69 inundation mapping where a model uses the forecasted river discharge to estimate the inundation
70 areas corresponding to a specific flood just before the onset of the event. Compared to the
71 traditional flood hazard mapping, real-time flood inundation mapping is more informative and
72 beneficial for emergency response-related decision-making.

73 In real-time flood inundation mapping, the model takes advantage of forecasted forcing data and
74 generates inundation areas corresponding to an upcoming flood event. Providing these maps ahead
75 of time is extremely valuable for building a robust flood warning system. Data assimilation (DA)
76 is an effective approach commonly used to improve the performance of real-time hydrologic
77 forecasting by updating the model state variables and parameters when new observation becomes
78 available (Moradkhani et al., 2019). The integration of DA with physical models is highly
79 advantageous as it enables accounting for different sources of uncertainties involved in model
80 predictions. These include (1) forcing data uncertainty due to the limitation of measurements and
81 spatiotemporal representativeness of the data (Alemohammad et al., 2015; Kumar et al., 2017), (2)
82 parameter uncertainty due to equifinality and non-uniqueness of parameters (Abbaszadeh et al.,
83 2018; Leach et al., 2018), (3) model structural uncertainty due to the imperfect representation and
84 conceptualization of a real system (Abbaszadeh et al., 2019; Pathiraja et al., 2018; Zhang et al.,
85 2019) and (4) initial and boundary condition uncertainty (DeChant and Moradkhani, 2014; Lee et
86 al., 2011).

87 Probabilistic forecasting and uncertainty quantification using DA have been the core of modeling
88 in the atmospheric and oceanic sciences (e.g. Anderson and Anderson, 1999; Courtier et al., 1993).
89 Later, the hydrologic community started to utilize this approach to account for the uncertainties
90 involved in different layers of model predictions and provide a more accurate and reliable
91 estimation of soil moisture (Gavahi et al., 2020; Pauwels et al., 2001; Reichle et al., 2002; Xu et

92 al., 2020), streamflow (Moradkhani et al., 2005a), snow (Sheffield et al., 2003; Slater and Clark,
93 2006) and other hydrologic variables. Despite these advances in hydrologic studies, the application
94 of data assimilation in conjunction with hydrodynamic models has received little attention in the
95 literature. The characterization of uncertainty in hydrodynamic models for probabilistic flood
96 inundation mapping has been mostly limited to Monte Carlo sampling (Ahmadisharaf et al., 2018;
97 Aronica et al., 2012; Domeneghetti et al., 2013; Neal et al., 2013; Papaioannou et al., 2017;
98 Pedrozo-Acuña et al., 2015; Purvis et al., 2008; Savage et al., 2016) and Generalized Likelihood
99 Uncertainty Estimation (GLUE) (Aronica et al., 2002a; Romanowicz and Beven, 2003).

100 The effectiveness and application of assimilating remotely sensed data (e.g. Soil Moisture Active
101 Passive (SMAP)) into hydrologic models have been vastly investigated in the literature
102 (Abbaszadeh et al., 2020; Azimi et al., 2020; Lievens et al., 2017). However, given the small scale
103 of the hydrodynamic modeling process, the spatiotemporal resolution of current satellite products
104 is not adequate for assimilating into these models. Due to the short duration of floods, satellite data
105 with a sub-daily time scale and spatial resolution less than the river width (e.g. 100 m) is
106 recommended. Since remote sensing products do not provide such high spatiotemporal resolution
107 data for hydrodynamic models, the research on hydrodynamic data assimilation is limited in the
108 literature. Some studies have limited their analyses to large rivers with a width of above 1 km (e.g.
109 study of Nile and Amazon) (Brêda et al., 2019). However, since the width of the majority of rivers
110 is less than 100 meters, these studies cannot be practically used in many regions.

111 Several studies used higher resolution synthetic Surface Water and Ocean Topography (SWOT)
112 data to evaluate the performance of assimilation techniques (Durand et al., 2008; Munier et al.,
113 2015; Pedinotti et al., 2014; Yoon et al., 2012). While these works provided important information
114 about the assimilation of satellite data into hydrodynamic models, their applications are only

115 limited to synthetic experiments, making them impractical for real case studies. Some studies have
116 implemented indirect methods to estimate WSE from flood extents generated by high-resolution
117 SAR satellite data (Giustarini et al., 2011; Hostache et al., 2010; Matgen et al., 2010b; Neal et al.,
118 2009). This approach can provide high-resolution data that is suitable for the majority of rivers.
119 However, the reliability of this data is concerning because the methods used to convert the flood
120 extent to WSE pose additional errors that downgrade the quality of the final observed data for
121 assimilation practices. Besides these issues, the major drawback of remote sensing data
122 assimilation pertains to their coarse temporal resolutions. To efficiently monitor the flood
123 dynamics, the assimilation process should be performed at a daily/hourly time scale, however, the
124 revisit frequency of satellites used for capturing the WSE ranges from a week to a month.
125 Therefore, there is a significantly low chance to capture multiple real-time remote sensing images
126 for the majority of inundated catchments during flood events. In the most optimistic scenario,
127 assimilation of satellite data is only limited to one/two updates during the simulation period which
128 may not be sufficient for reliable probabilistic flood inundation mapping.

129 Application of DA in hydrodynamic modeling can be either river monitoring or flood inundation
130 mapping. The goal of hydrodynamic data assimilation for river monitoring is to track variations in
131 the channel roughness and bathymetry in the long run. Therefore, the weekly/monthly satellite
132 data can be well assimilated into the models as the channel characteristics do not change on a daily
133 basis. On the other hand, flood inundation mapping needs an hourly/daily track of WSE because
134 floods happen rapidly and affect the river dynamics on a short time scale. The literature indicates
135 those studies that assimilated data into hydrodynamic models have been mostly designed for river
136 monitoring (Brêda et al., 2019; Durand et al., 2008; Yoon et al., 2012b). To capture the daily
137 dynamics of the rivers for real-time flood inundation mapping, the discharge and water stage

138 values measured at the gauge stations can be assimilated into the hydrodynamic models. Xu et al.,
139 (2017) performed a Particle Filtering (PF) approach to assimilate the water stage data from six
140 gauges into a hydrodynamic model. In order to calculate the particle weights in the filtering
141 process, they assumed that gauge observations are independent. In this study, however, we
142 consider interconnections among the gauge stations and apply multivariate Ensemble Kalman
143 Filter (EnKF) to a two-dimensional (2D) hydrodynamic model for better characterization and
144 quantification of uncertainty and further improving the accuracy of model simulations.

145 Advancing the probabilistic hydrodynamic modeling with DA techniques is a necessary step to fill
146 the gap between hydrology and hydrodynamics. To address this problem, this study aims to
147 explore the capability of a standard sequential DA technique, namely the EnKF, for real-time
148 probabilistic flood inundation mapping. Past studies that used the DA in conjunction with
149 hydrodynamic models, have mostly focused on the quantification of uncertainty in one or two
150 hydrodynamic variables; e.g. Giustarini et al., (2011) and Hostache et al., (2018) only investigated
151 the uncertainty in the upstream flow and rainfall, respectively; Yoon et al., (2012) focused on the
152 uncertainty of river bathymetry while ignoring the roughness parameter uncertainty. In addition,
153 the main application of DA-hydrodynamic modeling framework has been in river monitoring at
154 long-term or water stage forecasting during the flood events (Brêda et al., 2019; Matgen et al.,
155 2010; Xu et al., 2017). However, this study takes one step further and proposes a DA-
156 hydrodynamic modeling framework for real-time probabilistic flood inundation mapping while
157 accounting for major sources of uncertainties involved in the model simulations including
158 hydrodynamic model parameters (channel roughness and river bathymetry), forcing data (river
159 boundary conditions), and state variable (water depth). Additionally, unlike past works that
160 assimilated either discharge or water stage into the hydrodynamic model, this study performs a

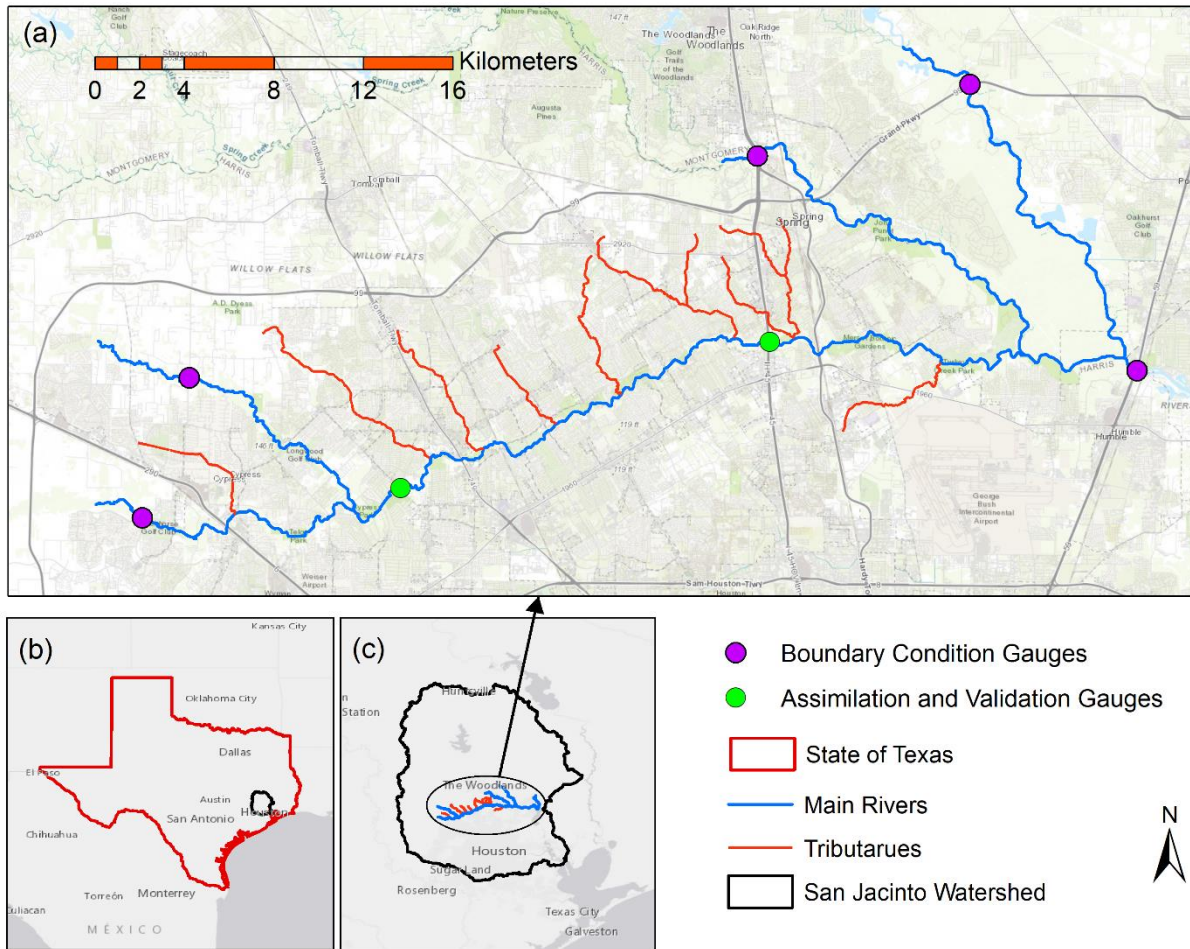
161 multivariate DA to incorporate the observed values of both variables into the hydrodynamic model
162 for a reliable simulation of inundation area.

163 **2 Data and Study area**

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

174 Figure 1.a shows the study area that consists of four main channels (blue lines) and eight tributaries
175 (red lines). The study area is located in the State of Texas (Figure 1.b) in the middle of the San
176 Jacinto watershed (Figure 1.c), a highly developed basin (USGS HUC6 #120401) with the area of
177 10400 km². The main channels simulated in the study are around 106 km draining into three HUC8
178 watersheds; the Spring (#12040102), West Fork San Jacinto (#12040101) and East Fork San
179 Jacinto (#12040103). The drainage areas of the channels are relatively flat with an average slope
180 of 0.62%, and the soil is mostly impermeable due to the high rate of recent developments in this
181 region. The upstream and downstream boundary conditions (purple points) are provided from the
182 daily streamflow in four United States Geological Survey (USGS) gauges ((#08068090, #

183 08068500, #08068740, #08068780) and water stage time series at the downstream gauge
184 (#08069500). The daily streamflow discharge in two internal gauges (green points #08068800 and
185 #08069000) and water stage time series in the second internal gauge are the observations that are
186 assimilated into the LISFLOOD-FP model. Internal gauges refer to those stations located between
187 upstream and downstream of the simulated river system. Figures 1.b and 1.c present the geographic
188 location of the study area within the state of Texas and San Jacinto watershed, respectively. To set
189 up the LISFLOOD-FP model, we use a DEM with 120 m spatial resolution resampled from one
190 arc second (30 m) USGS National Elevation Dataset. Such a coarse resolution DEM alleviates the
191 computational intensity of the proposed probabilistic hydrodynamic modeling framework. It
192 should be noted that the subgrid solver used for simulation of flood has the advantage of accepting
193 narrow rivers with a width of less than 120 m while the cell sizes are 120 m. In this study, the DA-
194 hydrodynamic modeling framework is parallelized and performed on the University of Alabama
195 High-Performance Computing cluster.



196

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

200 **3. Methods**

201 **3.1 Flood inundation model**

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

206
$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = 0 \tag{1}$$

207
$$\frac{1}{A} \frac{\partial A}{\partial t} + \frac{1}{A} \frac{\partial (\frac{Q^2}{A})}{\partial x} + g \frac{\partial h}{\partial x} - g(S_0 - S_f) = 0 \tag{2}$$

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

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

226 The upstream boundary conditions are generated from four USGS gauge stations (Figure. 1). To
 227 estimate the lateral fluxes, we calculate the deficit in the system as subtraction of the upstream

228 from downstream flows and then, distribute the deficit among river tributaries based on their
 229 drainage areas (Please refer to Jafarzadegan et. al (2021) for detailed information about the
 230 calculation of lateral flows in this study area). In section 3.3, we will further discuss the procedure
 231 we used to initialize the model parameters and river boundary conditions.

232 **3.2 Ensemble Kalman Filter (EnKF)**

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

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

$$246 \quad \theta_{t+1}^{i-} = \theta_t^{i+} + \tau_t^i \quad \tau_t^i \sim N(0, \eta_{t+1}) \quad \forall \quad \eta_{t+1} = \Sigma_{t+1}^\theta \quad (3)$$

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

$$248 \quad x_{t+1}^{i-} = f(x_t^{i+}, u_t^i, \theta_{t+1}^{i-}) + \omega_t^i \quad \omega_t^i \sim N(0, Q_t) \quad \forall \quad Q_t = \Sigma_t^x \quad (4)$$

249 $\hat{y}_{t+1}^i = h(x_{t+1}^{i-}, \theta_{t+1}^{i-}) + v_{t+1}^i \quad v_{t+1}^i \sim N(0, R_{t+1}) \quad \forall \quad R_{t+1} = \Sigma_{t+1}^y$ (5)

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

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

256 $\theta_{t+1}^{i+} = \theta_{t+1}^{i-} + K_{t+1}^\theta (y_{t+1}^i - \hat{y}_{t+1}^i)$ (6)

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

259 $K_{t+1}^\theta = \Sigma_{t+1}^{\theta y} [\Sigma_{t+1}^{yy} + R'_{t+1}]^{-1}$ (7)

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

268 $\Sigma_{t+1}^{\theta y}(p, j) = \frac{1}{N} \sum_{i=1}^N [(\theta_{t+1}^{i-}(p) - E[\theta_{t+1}^-(p)])(\hat{y}_{t+1}^i(j) - E[\hat{y}_{t+1}(j)])]$ (8)

$$269 \quad R'_{t+1}(j, j') = \begin{cases} R_{t+1} & j = j' \\ \frac{1}{N} \sum_{i=1}^N [(y_{t+1}^i(j) - E[y_{t+1}(j)])(y_{t+1}^i(j') - E[y_{t+1}(j')])] & j \neq j' \end{cases} \quad (9)$$

$$270 \quad \Sigma_{t+1}^{yy}(j, j') = \frac{1}{N} \sum_{i=1}^N [(\hat{y}_{t+1}^i(j) - E[\hat{y}_{t+1}(j)])(\hat{y}_{t+1}^i(j') - E[\hat{y}_{t+1}(j')])] \quad (10)$$

$$271 \quad E[\theta_{t+1}^-] = \frac{1}{N} \sum_{i=1}^N \theta_{t+1}^{i-} \quad (11)$$

$$272 \quad E[\hat{y}_{t+1}] = \frac{1}{N} \sum_{i=1}^N \hat{y}_{t+1}^i \quad (12)$$

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

$$275 \quad x_{t+1}^{i-} = f(x_t^{i+}, u_t^i, \theta_{t+1}^{i+}) + \omega_t^i \quad \omega_t^i \sim N(0, \Sigma_t^x) \quad \forall \quad Q_t = \Sigma_{t+1}^x \quad (13)$$

$$276 \quad \hat{y}_{t+1}^i = h(x_{t+1}^{i-}, \theta_{t+1}^{i+}) + v_{t+1}^i \quad v_{t+1}^i \sim N(0, \Sigma_{t+1}^y) \quad \forall \quad R_{t+1} = \Sigma_{t+1}^y \quad (14)$$

277 Model states ensemble is similarly updated as follows:

$$278 \quad x_{t+1}^{i+} = x_{t+1}^{i-} + K_{t+1}^x (y_{t+1}^i - \hat{y}_{t+1}^i) \quad (15)$$

$$279 \quad y_{t+1}^i = y_{t+1}^i + v_{t+1}^i \quad v_{t+1}^i \sim N(0, R_{t+1}) \quad \forall \quad R_{t+1} = \Sigma_{t+1}^y \quad (16)$$

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

$$281 \quad K_{t+1}^x = \Sigma_{t+1}^{xy} [\Sigma_{t+1}^{yy} + R'_{t+1}]^{-1} \quad (17)$$

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

$$284 \quad \Sigma_{t+1}^{xy}(s, j) = \frac{1}{N} \sum_{i=1}^N [(x_{t+1}^{i-}(s) - E[x_{t+1}^-(s)])(\hat{y}_{t+1}^i(j) - E[\hat{y}_{t+1}(j)])] \quad (18)$$

$$285 \quad E[x_{t+1}^-] = \frac{1}{N} \sum_{i=1}^N x_{t+1}^{i-} \quad (19)$$

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

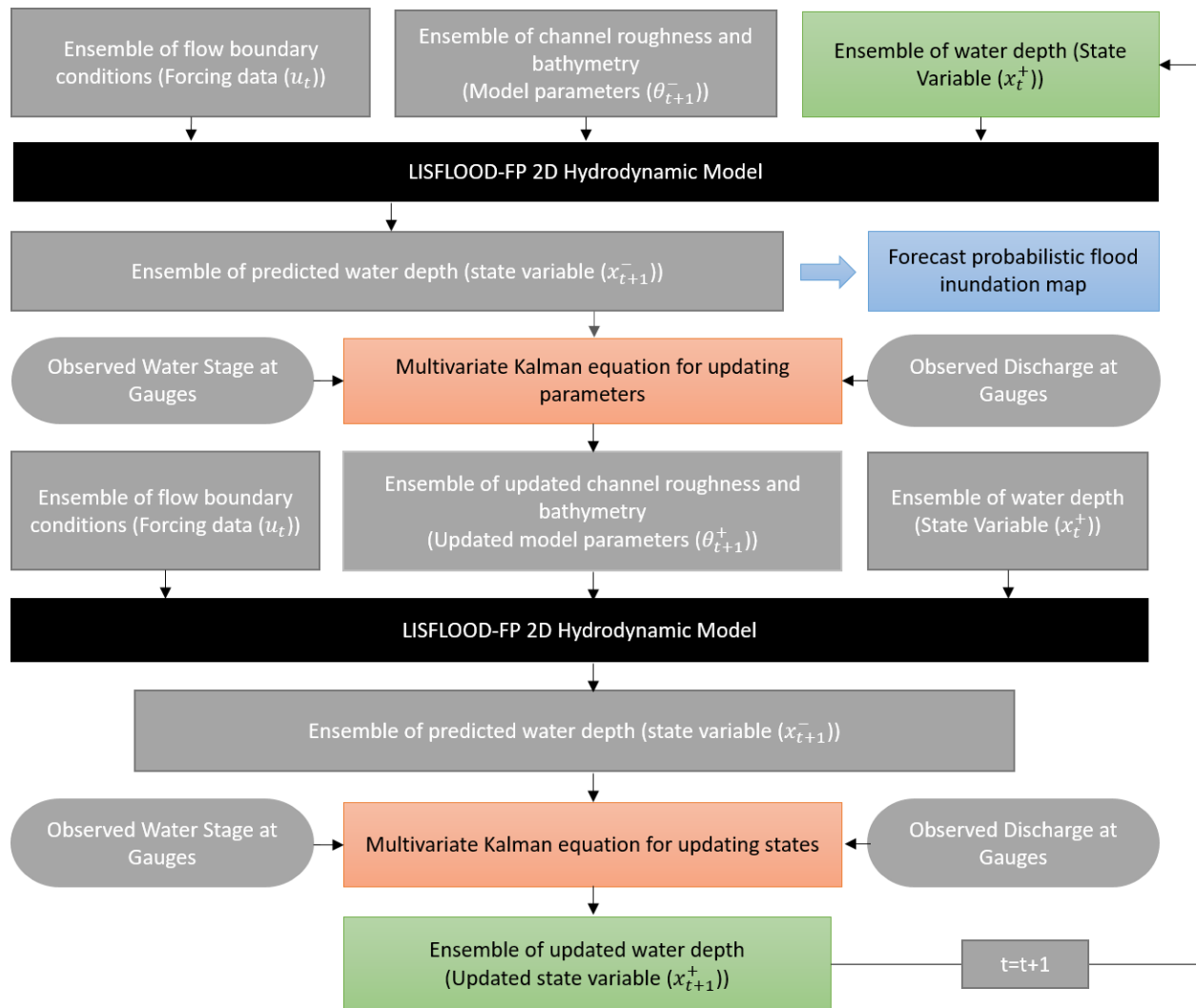
293 **3.3 DA-hydrodynamic modeling framework**

294 Figure. 1 illustrates the flowchart of the proposed DA-hydrodynamic modeling framework used
295 for real-time probabilistic flood inundation mapping. In this study, the EnKF is performed based
296 on an ensemble size of 100. The boundary conditions including four upstream flows, seven lateral
297 fluxes, and downstream flows are perturbed with adding white noises sampled from a normal
298 distribution with a mean zero and relative error of 20%. The errors are assumed heteroscedastic
299 meaning that their values are proportional to the flow magnitude. Pelletier, (1988) conducted a
300 literature review on the uncertainty of recorded flow at rivers and demonstrated that the error varies
301 in the range 8%-20%. Later, Di Baldassarre and Montanari, (2009) found that the uncertainty of
302 extreme flows can exceed to 25% due to extrapolating the rating curves. To characterize
303 uncertainty in the initial condition, namely water depth, we add a white noise with a mean zero
304 and standard deviation of 1 meter. In this study, using the proposed EnKF-based multivariate
305 assimilation approach, three point-scale observations, i.e., discharge at USGS gauges 1 and 2, as
306 well as water stage at gauge 2, are incorporated into the LISFLOOD-FP model to rectify its state
307 variables and parameters, and hence provide more accurate and reliable flood inundation maps.
308 First, the LISFLOOD-FP model is forced with the upstream, downstream and lateral flow

309 ensembles. To initialize the state variables in the system, the simulated water depth values at the
310 ending day of the warm-up period (the initial condition for the first day of the model simulation)
311 are perturbed with adding a white noise with a mean zero and standard deviation of 1 meter. It is
312 worth mentioning that the error terms used for the observed flows and the initial water depth are
313 determined through a manual tuning to achieve the most reliable predictions during the simulation.
314 The model parameters (i.e., channel roughness and bathymetry) are initialized using the Latin
315 Hypercube Sampling method and evolved during the assimilation process. The ensemble of water
316 depth values predicted by the model for the next time step together with observations, namely
317 water stage and discharge at gauges are used in the multivariate Kalman equation to update the
318 model parameters. The LISFLOOD-FP model is run for the second time with the updated
319 parameters and the second multivariate Kalman equation uses the predicted water depth with
320 observations to update the ensemble of water depth in the system. The ensemble of updated water
321 depth (state), bathymetry, and channel roughness (parameters) are used within the LISFLOOD-FP
322 to predict an ensemble of water depth for the next time step. The predicted water depth is simply
323 converted to a probabilistic flood inundation map. Using this data assimilation framework, we can
324 generate 1-day forecast of probabilistic flood inundation maps which would be highly beneficial
325 for real-time flood warning and decision making.

326 The simulation period of the LISFLOOD-FP model is set up for 45 days from July-30-2017 to
327 Sep-12-2017 and the entire month of July is used as a warm-up period. The model time step and
328 the Courant number are set to 1 second and 0.7, respectively, and the model is simulated at daily
329 scale. The water depth generated for the end of July is used as the initial condition of the model.
330 To account for the uncertainty of channel roughness and bathymetry, we sample these variables
331 from uniform distributions ranging from [0, 0.1] and [39, 42] m, respectively. The bathymetry

332 parameter is the elevation of the channel bed at the upper location of the channel. The offset
333 parameter is calculated by subtracting this value from DEM at the upper location. Then, the
334 bathymetry vector that includes the channel bed elevation for all channel cells is generated by
335 subtracting the offset from DEM values along the channel. It should be noted that the range of
336 uniform distribution for channel roughness is chosen based on previous studies (Aronica et al.,
337 2002b; Bales and Wagner, 2009; Di Baldassarre et al., 2009; Horritt, 2006; Pappenberger et al.,
338 2008) while the error range assumed for the bathymetry is mostly determined based on expert
339 judgment, and trial-and-error. Since the real magnitude and distribution of these errors have not
340 been fully understood in the literature, their estimated values may not be physically correct terms
341 and their estimation is ill-posed according to Renard et al., (2010).



342

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

349 **3.4 Experimental Design**

350 To assess the effectiveness and robustness of the proposed assimilation framework for
 351 probabilistic flood inundation mapping, we design two experiments. In the first experiment, we
 352 perform DA-hydrodynamic modeling on a synthetic case study where we assume the model is
 353 perfect and has no error. In this approach, we set the model parameters (channel roughness and

354 bathymetry), initial state (water depth) and boundary condition flows to fixed values and run the
 355 model to generate discharge and WSE across the gauges within the study area. These predicted
 356 values are assumed as benchmark observations. This synthetic analysis ensures that the
 357 assimilation process performs well and the model parameters end up converging to predefined
 358 values. In the second experiment, we implement the proposed assimilation framework on a real
 359 case study where the observed discharge and WSE data that are recorded from the USGS gauges
 360 during Hurricane Harvey, are assimilated into the model. In both experiments, we implement an
 361 open-loop (OL) simulation where the model is run without an assimilation. The WSE and flood
 362 extent maps generated by OL are compared with the results provided by the EnKF in the synthetic
 363 and real case studies. Considering the severe flood condition during the Hurricane, we aim to
 364 investigate the extent to which the multivariate DA-Hydrodynamic modeling framework improves
 365 the model simulation and flood inundation mapping skill.

366 **3.5 Validation strategy**

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

$$374 \quad 1 - \sqrt{\left(\left(\frac{\text{Cov}_{y_t y'_t}}{\sigma \sigma'}\right) - 1\right)^2 + \left(\left(\frac{\sigma'}{\sigma}\right) - 1\right)^2 + \left(\left(\frac{\mu'}{\mu}\right) - 1\right)^2} \quad (20)$$

$$375 \quad \sqrt{\frac{1}{T} \sum_{t=1}^T (y'_t - y_t)^2} \quad (21)$$

$$376 \quad \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y}'_{\bullet,t})^2} \times \left(\frac{1}{T} \left\{ \sum_{t=1}^T \sqrt{\frac{1}{T} \left[\sum_{t=1}^T (y_t - \bar{y}'_{\bullet,t})^2 \right]} \right\} \sqrt{\frac{N+1}{2N}} \right)^{-1} \quad (22)$$

$$377 \quad 1 - \frac{2}{T} \sum_{t=1}^T \left| \frac{z_t}{T} - U_t \right| \quad (23)$$

378 where y_t and y'_t are the observed and simulated values, respectively. The Kling–Gupta Efficiency
 379 (*KGE*) varies from $-\infty$ to 1, such that a value of 1 indicates a perfect fit between observed and
 380 simulated values. The pairs of (μ, σ) and (μ', σ') represent the first two statistical moments
 381 (means and standard deviations) of y_t and y'_t , respectively. *RMSE* is the square root of the mean
 382 of the square of all of the errors between the predicted and observed values.

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

388 The above four performance measures assess the dynamic behavior of DA-hydrodynamic
 389 modeling framework at two specific points. Moreover, to spatially evaluate the behavior of the
 390 proposed framework, we compare the maximum probabilistic flood inundation maps (union of
 391 probabilistic maps over the simulation period) with the observed floodplain map delineated
 392 aftermath of Harvey. The Receiver Operating Characteristic (ROC) graph is a common tool for
 393 validating probabilistic classifiers (Fawcett, 2006). Consider a deterministic flood map as a binary
 394 map where one and zero represent flooded and non-flooded cells, respectively. First, a threshold
 395 in the range of $[0,1]$ is used to convert the probabilistic map to a binary deterministic map. This

396 means all cells with the probability of inundation less than a given threshold are converted to zero
397 and other cells are set to one. The binary map is compared with the reference map and the rate of
398 true positive (*rtp*) and false positive (*rfp*) are calculated using Equations 24 and 25 (Jafarzadegan
399 and Merwade, 2017):

$$400 \quad rtp = \frac{\text{True positive instances}}{\text{total positives}} \quad (24)$$

$$401 \quad rfp = \frac{\text{False positive instances}}{\text{total negative}} \quad (25)$$

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

$$412 \quad F = \frac{\text{True positive instances}}{\text{Total positives} + \text{False positives}} \times 100 \quad (26)$$

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

$$416 \quad UFI = \frac{\sum_{i=1}^N (1-P_i)}{N} \times 100 \quad i \in Fl \quad (27)$$

$$417 \quad OFI = \frac{\sum_{j=1}^M (P_j)}{M} \times 100 \quad j \in NFl \quad (28)$$

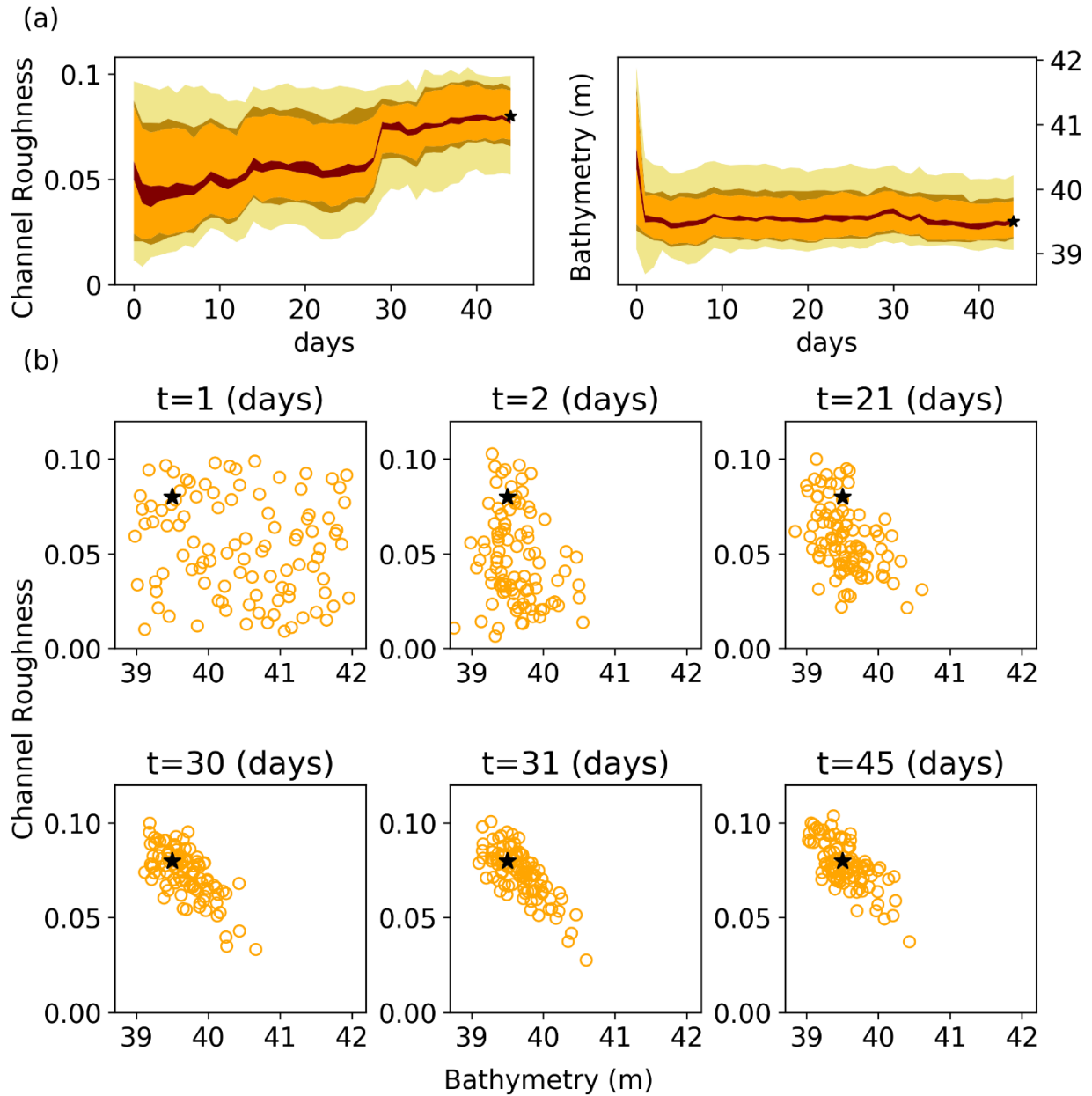
418 where Fl and NFl denote the flooded and non-flooded regions in the reference map, and i and j are
 419 indicators of cells located within these regions. N and M are the total number of cells in the Fl and
 420 NFl regions and P_i, P_j denote the probability of inundation for cells i and j derived from the
 421 probabilistic flood maps.

422 **4. Results**

423 **4.1 Experiment 1: Synthetic Case Study**

424 We conduct the synthetic experiment to ensure the usefulness and effectiveness of the proposed
 425 DA-hydrodynamic modeling framework. Figure 3.a presents uncertainty bound evolution of the
 426 parameters in the LISFLOOD-FP model (i.e., channel roughness and bathymetry) for 45 days
 427 assimilation of synthetic observations (i.e., discharge at gauges 1 and 2 and water stage at gauge
 428 2). It is worth mentioning that the uncertainty of bathymetry shown in [this Figure](#) corresponds to
 429 the channel bed elevation at the upper location of the channel. As seen both parameters converge
 430 smoothly to the certain region in parameter space where the uncertainty bounds stabilize. While
 431 the uncertainty bound associated with the bathymetry becomes stabilized at the early stage of the
 432 assimilation process, for the channel roughness, the uncertainty bound is stabilized toward the end
 433 of the assimilation period. It is also evident from Figure 3.a that the bathymetry is a more
 434 identifiable parameter compared to the channel roughness as it shows the fastest convergence with
 435 a minimum degree of uncertainty. However, the channel roughness is less identifiable with the
 436 slowest convergence. In Figure 3.b, the first day ($t=1$) includes all 100 ensemble members of

437 parameters and day 30 corresponds to the highest discharge and water stage of flooding when the
438 model parameters reach the highest improvement and get closer to the true value. Figure 3.b shows
439 that both model parameters are converging toward the true values as the assimilation proceeds.
440 This indicates the efficacy and usefulness of the proposed DA-hydrodynamic modeling framework
441 developed in this study.



442

443 *Figure 3. Temporal evolution of the LISFLOOD parameters for the synthetic experiment during*
 444 *Hurricane Harvey using the EnKF. (a) Temporal evolution of model parameter predictive*
 445 *intervals (shaded areas) corresponding to 95, 75, 68, and 10 percentile (b) Temporal evolution*
 446 *of particle positions in the model parameter space at six different days during the Hurricane.*
 447 *The black stars at the end of each parameter subplot represent the true parameter values.*

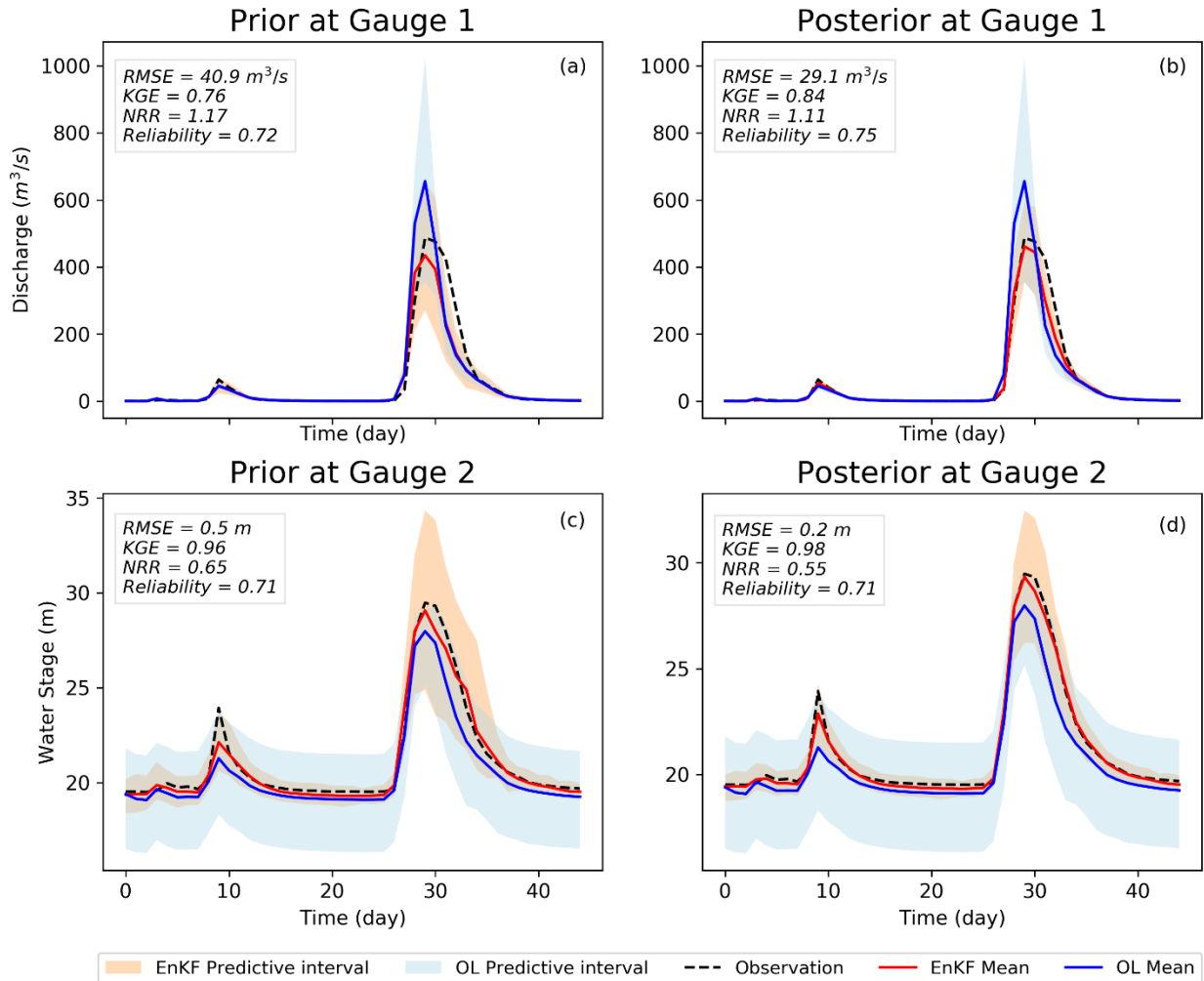
448

449

450 **4.2 Experiment 2: Real Case Study**

451 In the real experiment, we assimilate the discharge and water stage readings from two USGS
452 gauges into the LISFLOOD-FP model. We also run the OL simulation and calculate the ensemble
453 mean to predict the discharge and water stage at these two gauges. Figure 4 presents a comparison
454 of simulated discharge (Figures 4a and 4b) and water stage (Figures 4c and 4d) with observations
455 using both OL and our EnKF-based approach. Figures 4a and 4c are the prior estimates of
456 discharge and water stage, while Figures 4b and 4d show their posterior distributions that reflect
457 the updated variables after assimilating the observations into the model. It is worth mentioning
458 that although prior distributions represent the results before assimilating new observations into the
459 model, their values are dependent on the initial conditions updated from observations in the
460 previous time step. Since forecasting (1-day lead time) is the main objective of DA-hydrodynamic
461 modeling framework, we specifically focus on behavior of priors. As can be seen, the simulated
462 peak discharge by the OL is highly overestimated by around 200 m³/s while assimilating the
463 observations improve the results so that their difference with observation is less than 50 m³/s at
464 the peak of the flood ($KGE=0.76$ and $RMSE=40.9$ m³/s). In contrast, the simulated water stage
465 in Figures 4c and 4d are underestimated by OL by around 2 meters at the peak. Compared to the
466 OL, using the developed EnKF approach raises the peak of water stage at peak and reduces the
467 errors significantly ($KGE=0.96$ and $RMSE=0.5$ m). The accurate estimates of prior discharge and
468 water stage confirm the applicability of the proposed assimilation framework in forecast mode
469 when real-time flood warning and decision making is the priority. The *NRR* measure for the prior
470 discharge and water stage are 1.17 and 0.65, showing that the uncertainty bound is underestimated
471 and overestimated, respectively. The *Reliability* of both variables is above 70 percent since the
472 uncertainty bounds encompass the observations for almost the entire simulation period.

473

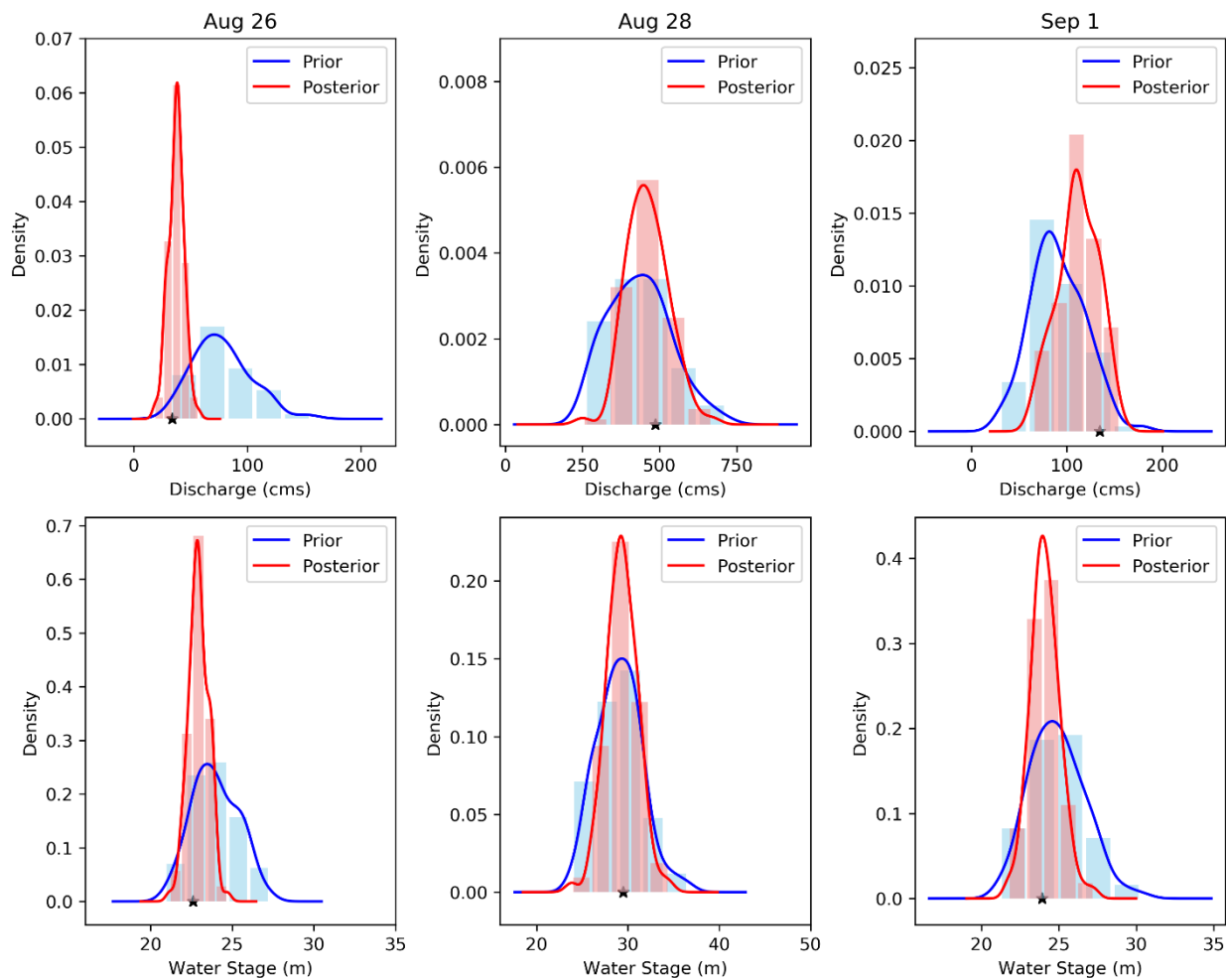


474

475 *Figure 4 Simulation results of LISFLOOD-FP for the real experiment during Hurricane Harvey*
 476 *using the EnKF and open-loop. (a) Prior simulated discharge at gauge 1 (b) Posterior simulated*
 477 *discharge at gauge 1 (c) Prior simulated water stage at gauge 2 (d) Posterior simulated water*
 478 *stage at gauge 2. The shaded areas represent the predictive interval of simulated discharge and*
 479 *water stage by EnKF.*

480 Figure 5 illustrates the prior and posterior distributions of discharge and water stage in the
 481 beginning, peak, and ending days of Hurricane Harvey flood. In all three days, the uncertainty
 482 bounds of both discharge and water stage are narrowed down by assimilating the observations so
 483 that posterior distributions are more precise compared to the priors. In the beginning and ending
 484 days (Aug 26 and Sep 1), the mean of prior distributions is substantially shifted toward truth in the
 485 posterior distributions. Figure 5 reveals that our developed approach provides more accurate and

486 reliable posterior discharge and water stage distributions compared to prior distributions where the
 487 simulations are either overestimated or underestimated. It is noted that, on August 28 (day of flood
 488 peak), although the prior distributions accurately represent the observation, they have a wide
 489 uncertainty bound. After correcting/updating the model state variables and parameters, as posterior
 490 distributions show, the uncertainty bound is reduced while the ensemble mean remains closer to
 491 the observation.



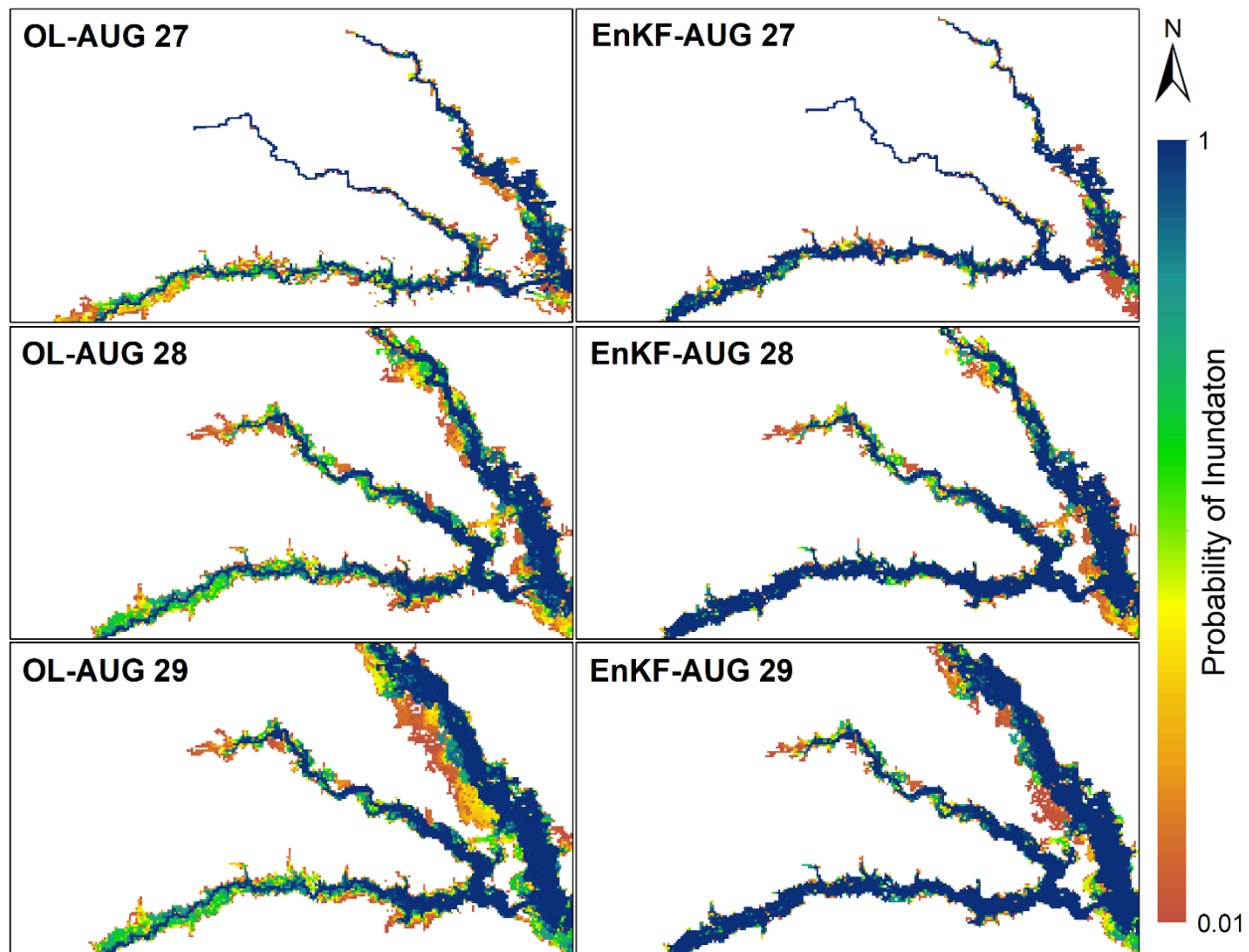
492

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

496

497 **4.3 Probabilistic Flood Inundation Mapping**

498 In this section, we account for the uncertainties involved in flood modeling and generate real-time
499 probabilistic flood inundation maps. Since the majority of flooding conditions occurred within six
500 days from August 27-Sep 1, we display the spatial distribution of water depth in this period and
501 provide probabilistic flood inundation maps using both OL and our developed approach (Figures
502 6 and 7). Figure 6 represents the first three days of Harvey, which corresponds to the upper limb
503 of the flood hydrograph. On August 27, the major difference between the OL and EnKF appears
504 in the regions around the upstream of the lower channel where the EnKF provides a more reliable
505 prediction of the inundated area. Moving toward the peak of flood on Aug 29, the OL generates a
506 large region of uncertain cells around the banks of the upper channel, while both the extent and
507 density of uncertain values in the probabilistic maps generated by the EnKF is smaller during the
508 peak of Harvey.

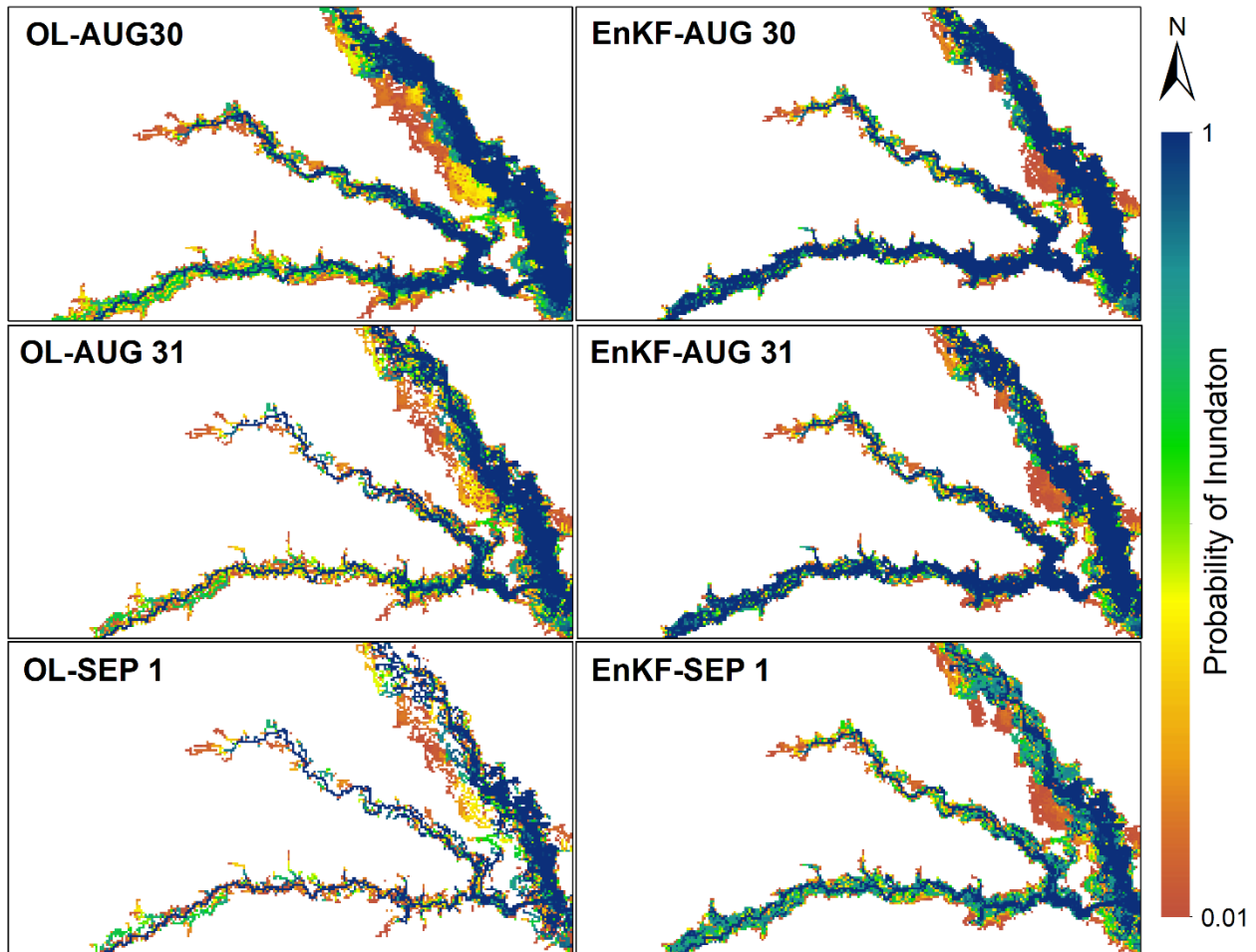


509

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

512 Figure 7 shows the probabilistic inundation areas in the last three days corresponding to the lower
 513 limb of the flood hydrograph. The discrepancies between the OL and EnKF flood maps increase
 514 showing that performing DA is more effective in improving the inundation mapping skill from
 515 peak to ending point of the flood hydrograph. A large number of inundated cells generated by the
 516 OL are vanished after the peak of Harvey which results in a set of scattered discontinuous maps in
 517 Aug 31 and Sep 1. On the other hand, the probabilistic maps generated by the EnKF maintain their
 518 continuous shapes so that the probability of inundation is reduced without changing the extent.
 519 The merit of the EnKF in improving the flood inundation areas at the lower limb of the flood

520 hydrograph agrees with results in Figures 4c and 4d where the EnKF widens the simulated water
521 stage hydrographs and removes the lag difference that exists between the open-loop and
522 observations.



523

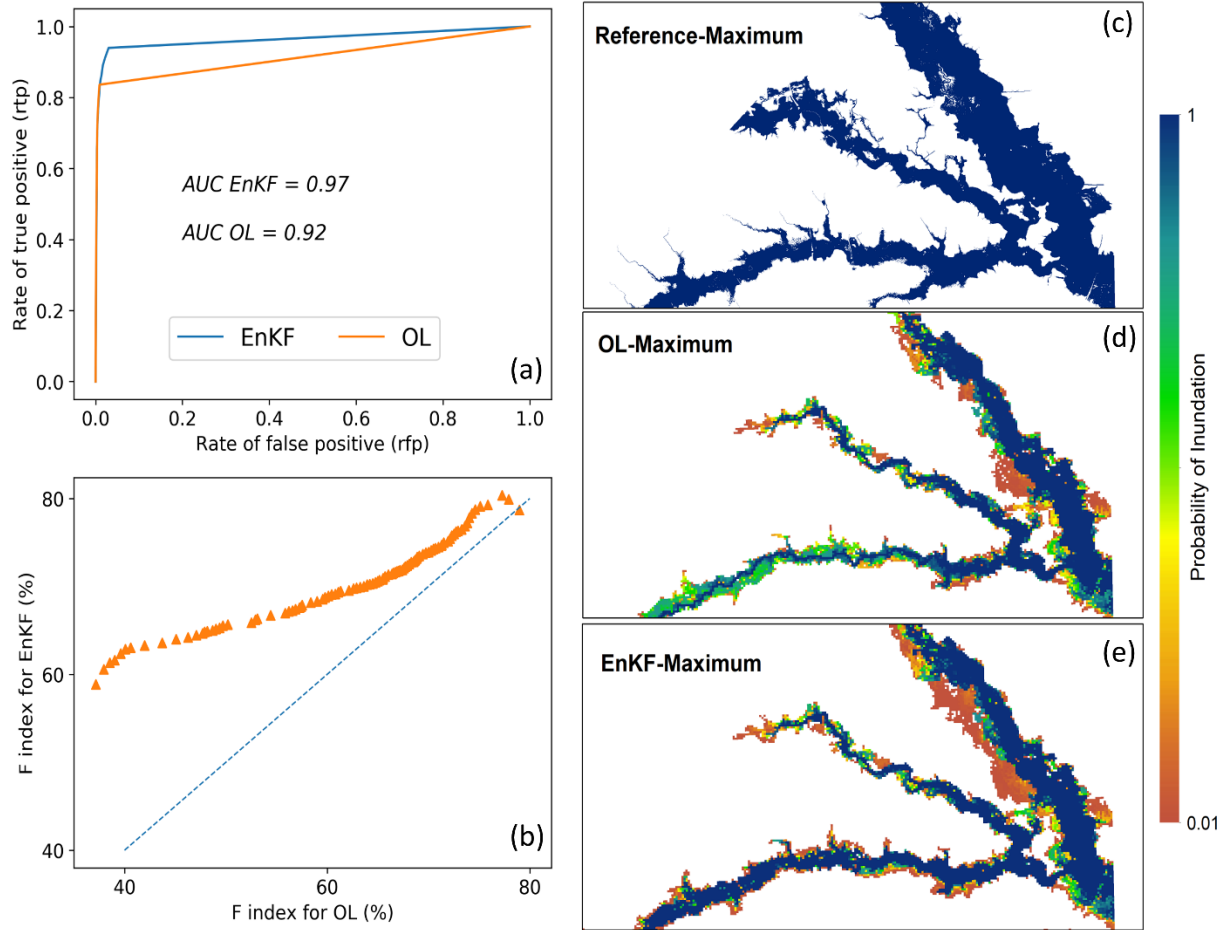
524 *Figure 7 Probabilistic flood inundation maps generated by OL and EnKF techniques to simulate*
525 *the lower limb of Harvey flood hydrograph from Aug 30 to Sep 1.*

526 Finally, to quantify the performance of EnKF and OL for generating a spatial distribution of water
527 depth over the domain, we illustrate the ROC graphs, the *AUC* values, and Fit indices in Figure 8.

528 To calculate these measures, we ignore the temporal distributions and only report the maximum
529 inundation maps that represent the union of flooded areas over the entire period of Harvey.

530 Comparing the EnKF and OL in Figure 8.a, the EnKF line (blue) is closer to the northwest of the

531 *rfp-rtp* space where its *AUC* is 5% higher than the OL approach. In Figure 8.b, each point
532 represents the *F* indices for the OL and the EnKF approaches corresponding to a given threshold.
533 Using 100 thresholds that range from [0.01,1], the probabilistic maps are converted to 100
534 deterministic maps and the *F* indices are calculated. The position of scatters above the dash line
535 confirms the EnKF outperforms the OL. In addition to these measures, the [*UFI*, *OFI*] indices
536 calculated for OL and EnKF approaches are [30.3, 0.26] %, and [23.4, 0.4]% respectively. The
537 low values of *OFI* for both approaches (< 1%) show that the simulations mostly underestimate the
538 flood inundation areas. In addition, comparing the indices of both approaches reveal that the EnKF
539 reduces the overall underestimation by around 7%.



540

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

543

544

545 **5. Discussion and Conclusions**

546 The main motivation in this study is to propose a DA-hydrodynamic modeling framework for real-
 547 time probabilistic flood inundation mapping. Considering the coarse spatiotemporal resolution of
 548 satellite data for capturing the WSE, assimilating them into the hydrodynamic models may not be
 549 a practical solution for an upcoming flood event. On the other hand, the availability of daily
 550 discharge and WSE data at gauge stations is a great opportunity to establish a multivariate DA-

551 hydrodynamic modeling framework that updates the initial condition of modeling at daily scale
552 and forecast the flood inundation areas at 1 day lead time. Here, we used the EnKF data
553 assimilation method in conjunction with a hydrodynamic model to account for different sources
554 of uncertainties involved in different layers of model simulations, including the boundary
555 conditions, model parameters, and initial condition, and generate real-time probabilistic flood
556 inundation maps. To further enhance the performance of the developed framework, the discharge
557 and water stage at two different gauges are simultaneously assimilated into the LISFLOOD-FP
558 model. The multivariate EnKF approach considers the correlation between discharge at two gauges
559 and between discharge and WSE at one gauge using a modified covariance matrix and Kalman
560 gain equation.

561 In the synthetic experiment, we examined the convergence of model parameters toward truth and
562 found that the proposed DA-hydrodynamic modeling framework can be successfully used to
563 improve the accuracy and reliability of model predictions while accounting for uncertainties
564 associated with model parameters. The channel roughness coefficient varied more rapidly than the
565 bathymetry during the temporal evolutions of these parameters showing the better identifiability
566 of this parameter. The validation results of the real experiment revealed that the assimilation with
567 the EnKF approach improves the model predictions across temporal and spatial scales (i.e.,
568 discharge and water stage time series at gauges and flood maps showing the maximum water depth
569 over the simulation period). These improvements are more pronounced during the falling limb of
570 the flood hydrograph where the EnKF widens the simulated hydrograph and removes the existing
571 lag compared to the observations. Similarly, the simulated flood inundation maps confirm that the
572 OL provides discontinuous scattered maps during the flood recession period while the EnKF
573 provides a more accurate representation of the inundation areas. The validation results also

574 demonstrate that the EnKF reduces the underestimation by 7% and outperformed the OL approach
575 by around 5% for probabilistic flood inundation mapping.

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

593 The coarse DEM used in this study cannot perfectly represent the watershed topography and
594 bathymetry, and can be the main reason for underestimation of inundation areas (F index less than
595 80%). Savage et al., (2016) investigated the impacts of DEM resolution on the accuracy and
596 efficiency of probabilistic flood inundation maps generated with the LISFLOOD-FP model. They

597 demonstrated that models with resolution less than 50 offer little gain in performance yet are more
598 than an order of magnitude computationally expensive which can become infeasible when
599 undertaking probabilistic analysis. They also found that the reliability of flood maps deteriorates
600 at resolutions coarser than 100 m. Considering the medium scale of our study (> 100 km river)
601 compared to the reach scale (~ 10 km river) of the work by Savage et al., (2016), here we slightly
602 increased their suggested threshold for the DEM and demonstrated that the accuracy of results is
603 still acceptable.

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

614 The time dependency that exists between the upstream and downstream gauges along a channel
615 can affect the performance of multivariate assimilation with those gauges. For future studies, using
616 a more advanced DA technique that fully characterizes the model structural uncertainty
617 (Abbaszadeh et al., 2019), and considering the time lag dependency between multiple gauges can
618 improve the performance of modeling and provide more realistic assimilation of the hydrodynamic
619 models. Another limitation of this study is the simple assumptions made for perturbing the initial

620 condition (water depth), parameters (channel roughness and river bathymetry) and observations
621 (WSE and discharge). More investigation on the physically meaningful distribution of these values
622 can enhance the performance of the DA-hydrodynamic modeling framework in future studies. A
623 joint assimilation of point source gauges and remotely sensed data can also improve the reliability
624 and accuracy of the results. Finally, proposing a DA-hydrodynamic modeling framework that
625 considers the DEM and channel width uncertainty can provide a more comprehensive uncertainty
626 quantification for probabilistic flood inundation mapping in future studies.

627 An advantage of the proposed DA-hydrodynamic modeling framework is its generic format so that
628 other studies can follow the flowchart in Figure 2 and use information in Section 3.2 and 3.3 to set
629 up the hydrodynamic model and the EnKF algorithm, respectively. To properly apply this
630 framework to other studies, first, the point source observations of WSE and discharge should be
631 available at daily/sub-daily scales. Second, the modeler should have access to high performance
632 computing facilities for parallel simulation of ensemble members. Third, the hydrodynamic model
633 should be sequentially executed within the DA algorithm. The modeler should check the
634 hydrodynamic model manual and make sure that the outputs and initial conditions can be updated
635 in a sequential manner. Taking these three considerations into account, the proposed DA-
636 hydrodynamic modeling framework can be applied to any other study areas that are prone to
637 frequent flooding and provide a robust and generic tool for real-time probabilistic flood inundation
638 mapping.

639 **Data availability**

640 All the data used in this study, including the gauge streamflow and water stage data and the DEMs,
641 are publicly available from the USGS website and National Elevation Dataset (NED). The

642 reference flood maps provided for Hurricane Harvey is available from the USGS report at
643 <https://pubs.usgs.gov/sir/2018/5070/sir20185070.pdf>.

644 **Author contribution**

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

649 **Competing interests**

650 The authors declare that they have no conflict of interest.

651 **Acknowledgments**

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

655

656

657

658

659

660

661 **References**

- 662 Abbaszadeh, P., Gavahi, K., Moradkhani, H., 2020. Multivariate remotely sensed and in-situ data
663 assimilation for enhancing community WRF-Hydro model forecasting. *Adv. Water*
664 *Resour.* 145, 103721.
- 665 Abbaszadeh, P., Moradkhani, H., Daescu, D.N., 2019. The Quest for Model Uncertainty
666 Quantification: A Hybrid Ensemble and Variational Data Assimilation Framework.
667 *Water Resour. Res.* 55, 2407–2431. <https://doi.org/10.1029/2018WR023629>
- 668 Abbaszadeh, P., Moradkhani, H., Yan, H., 2018. Enhancing hydrologic data assimilation by
669 evolutionary Particle Filter and Markov Chain Monte Carlo. *Adv. Water Resour.* 111,
670 192–204. <https://doi.org/10.1016/j.advwatres.2017.11.011>
- 671 Ahmadisharaf, E., Kalyanapu, A.J., Bates, P.D., 2018. A probabilistic framework for floodplain
672 mapping using hydrological modeling and unsteady hydraulic modeling. *Hydrol. Sci. J.*
673 63, 1759–1775. <https://doi.org/10.1080/02626667.2018.1525615>
- 674 Alemohammad, S.H., McLaughlin, D.B., Entekhabi, D., 2015. Quantifying precipitation
675 uncertainty for land data assimilation applications. *Mon. Weather Rev.* 143, 3276–3299.
- 676 Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., Feyen, L., 2014. Advances in pan-
677 European flood hazard mapping. *Hydrol. Process.* 28, 4067–4077.
678 <https://doi.org/10.1002/hyp.9947>
- 679 Anderson, J.L., Anderson, S.L., 1999. A Monte Carlo implementation of the nonlinear filtering
680 problem to produce ensemble assimilations and forecasts. *Mon. Weather Rev.* 127, 2741–
681 2758.
- 682 Aronica, G., Bates, P.D., Horritt, M.S., 2002a. Assessing the uncertainty in distributed model
683 predictions using observed binary pattern information within GLUE. *Hydrol. Process.* 16,
684 2001–2016. <https://doi.org/10.1002/hyp.398>
- 685 Aronica, G., Bates, P.D., Horritt, M.S., 2002b. Assessing the uncertainty in distributed model
686 predictions using observed binary pattern information within GLUE. *Hydrol. Process.* 16,
687 2001–2016. <https://doi.org/10.1002/hyp.398>
- 688 Aronica, G.T., Franza, F., Bates, P.D., Neal, J.C., 2012. Probabilistic evaluation of flood hazard
689 in urban areas using Monte Carlo simulation. *Hydrol. Process.* 26, 3962–3972.
690 <https://doi.org/10.1002/hyp.8370>
- 691 Azimi, S., Dariane, A.B., Modanesi, S., Bauer-Marschallinger, B., Bindlish, R., Wagner, W.,
692 Massari, C., 2020. Assimilation of Sentinel 1 and SMAP-based satellite soil moisture
693 retrievals into SWAT hydrological model: The impact of satellite revisit time and product
694 spatial resolution on flood simulations in small basins. *J. Hydrol.* 581, 124367.
- 695 Bales, J.D., Wagner, C.R., 2009. Sources of uncertainty in flood inundation maps. *J. Flood Risk*
696 *Manag.* 2, 139–147. <https://doi.org/10.1111/j.1753-318X.2009.01029.x>
- 697 Bates, P.D., De Roo, A.P.J., 2000. A simple raster-based model for flood inundation simulation.
698 *J. Hydrol.* 236, 54–77. [https://doi.org/10.1016/S0022-1694\(00\)00278-X](https://doi.org/10.1016/S0022-1694(00)00278-X)
- 699 Brêda, J.P.L.F., Paiva, R.C.D., Bravo, J.M., Passaia, O.A., Moreira, D.M., 2019. Assimilation of
700 Satellite Altimetry Data for Effective River Bathymetry. *Water Resour. Res.* 55, 7441–
701 7463. <https://doi.org/10.1029/2018WR024010>
- 702 Clark, M.P., Hay, L.E., 2004. Use of Medium-Range Numerical Weather Prediction Model
703 Output to Produce Forecasts of Streamflow. *J. Hydrometeorol.* 5, 15–32.
704 [https://doi.org/10.1175/1525-7541\(2004\)005<0015:UOMNWP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0015:UOMNWP>2.0.CO;2)

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

708 Cuo, L., Pagano, T.C., Wang, Q.J., 2011. A Review of Quantitative Precipitation Forecasts and
709 Their Use in Short- to Medium-Range Streamflow Forecasting. *J. Hydrometeorol.* 12,
710 713–728. <https://doi.org/10.1175/2011JHM1347.1>

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

714 Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: a
715 quantitative analysis. *Hydrol. Earth Syst. Sci.* 13, 913–921. [https://doi.org/10.5194/hess-](https://doi.org/10.5194/hess-13-913-2009)
716 [13-913-2009](https://doi.org/10.5194/hess-13-913-2009)

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

720 Domeneghetti, A., Vorogushyn, S., Castellarin, A., Merz, B., Brath, A., 2013. Probabilistic flood
721 hazard mapping: effects of uncertain boundary conditions. *Hydrol. Earth Syst. Sci.* 17,
722 3127–3140.

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

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

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

731 Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza Guingla, D.A., Pauwels, V., De
732 Lannoy, G., De Keyser, R., Pfister, L., Hoffmann, L., 2011. Assimilating SAR-derived
733 water level data into a hydraulic model: a case study. *Hydrol. Earth Syst. Sci.* 15, 2349–
734 2365.

735 Habets, F., LeMoigne, P., Noilhan, J., 2004. On the utility of operational precipitation forecasts
736 to served as input for streamflow forecasting. *J. Hydrol.* 293, 270–288.
737 <https://doi.org/10.1016/j.jhydrol.2004.02.004>

738 Hall, J.W., Tarantola, S., Bates, P.D., Horritt, M.S., 2005. Distributed Sensitivity Analysis of
739 Flood Inundation Model Calibration. *J. Hydraul. Eng.* 131, 117–126.
740 [https://doi.org/10.1061/\(ASCE\)0733-9429\(2005\)131:2\(117\)](https://doi.org/10.1061/(ASCE)0733-9429(2005)131:2(117))

741 Horritt, M.S., 2006. A methodology for the validation of uncertain flood inundation models. *J.*
742 *Hydrol.* 326, 153–165. <https://doi.org/10.1016/j.jhydrol.2005.10.027>

743 Horritt, M.S., Bates, P.D., 2002. Evaluation of 1D and 2D numerical models for predicting river
744 flood inundation. *J. Hydrol.* 268, 87–99. [https://doi.org/10.1016/S0022-1694\(02\)00121-X](https://doi.org/10.1016/S0022-1694(02)00121-X)

745 Hostache, R., Chini, M., Giustarini, L., Neal, J., Kavetski, D., Wood, M., Corato, G., Pelich, R.-
746 M., Matgen, P., 2018. Near-Real-Time Assimilation of SAR-Derived Flood Maps for
747 Improving Flood Forecasts. *Water Resour. Res.* 54, 5516–5535.
748 <https://doi.org/10.1029/2017WR022205>

749 Hostache, R., Lai, X., Monnier, J., Puech, C., 2010. Assimilation of spatially distributed water
750 levels into a shallow-water flood model. Part II: Use of a remote sensing image of Mosel
751 River. *J. Hydrol.* 390, 257–268. <https://doi.org/10.1016/j.jhydrol.2010.07.003>

752 Jafarzadegan, K., Merwade, V., 2017. A DEM-based approach for large-scale floodplain
753 mapping in ungauged watersheds. *J. Hydrol.* 550, 650–662.
754 <https://doi.org/10.1016/j.jhydrol.2017.04.053>

755 Jafarzadegan, K., Merwade, V., Saksena, S., 2018. A geomorphic approach to 100-year
756 floodplain mapping for the Conterminous United States. *J. Hydrol.* 561, 43–58.
757 <https://doi.org/10.1016/j.jhydrol.2018.03.061>

758 Kumar, S.V., Dong, J., Peters-Lidard, C.D., Mocko, D., Gómez, B., 2017. Role of forcing
759 uncertainty and background model error characterization in snow data assimilation.
760 *Hydrol. Earth Syst. Sci.* 21, 2637–2647.

761 Leach, J.M., Kornelsen, K.C., Coulibaly, P., 2018. Assimilation of near-real time data products
762 into models of an urban basin. *J. Hydrol.* 563, 51–64.

763 Lee, H., Seo, D.-J., Koren, V., 2011. Assimilation of streamflow and in situ soil moisture data
764 into operational distributed hydrologic models: Effects of uncertainties in the data and
765 initial model soil moisture states. *Adv. Water Resour.* 34, 1597–1615.

766 Lievens, H., Reichle, R.H., Liu, Q., De Lannoy, G.J., Dunbar, R.S., Kim, S.B., Das, N.N., Cosh,
767 M., Walker, J.P., Wagner, W., 2017. Joint Sentinel-1 and SMAP data assimilation to
768 improve soil moisture estimates. *Geophys. Res. Lett.* 44, 6145–6153.

769 Maidment, D.R., 2017. Conceptual Framework for the National Flood Interoperability
770 Experiment. *JAWRA J. Am. Water Resour. Assoc.* 53, 245–257.
771 <https://doi.org/10.1111/1752-1688.12474>

772 Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V.R.N.,
773 De Lannoy, G., De Keyser, R., Savenije, H.H.G., 2010a. Towards the sequential
774 assimilation of SAR-derived water stages into hydraulic models using the Particle Filter:
775 proof of concept. *Hydrol. Earth Syst. Sci.* 14, 1773–1785.

776 Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V.R.N.,
777 De Lannoy, G., De Keyser, R., Savenije, H.H.G., 2010b. Towards the sequential
778 assimilation of SAR-derived water stages into hydraulic models using the Particle Filter:
779 proof of concept. *Hydrol. Earth Syst. Sci.* 14, 1773–1785.

780 Merwade, V., Cook, A., Coonrod, J., 2008. GIS techniques for creating river terrain models for
781 hydrodynamic modeling and flood inundation mapping. *Environ. Model. Softw.* 23,
782 1300–1311. <https://doi.org/10.1016/j.envsoft.2008.03.005>

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

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

789 Moradkhani, H., Sorooshian, S., Gupta, H.V., Houser, P.R., 2005b. Dual state–parameter
790 estimation of hydrological models using ensemble Kalman filter. *Adv. Water Resour.* 28,
791 135–147. <https://doi.org/10.1016/j.advwatres.2004.09.002>

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

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

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

801 Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., Pappenberger, F., 2009. A data
802 assimilation approach to discharge estimation from space. *Hydrol. Process.* 23, 3641–
803 3649. <https://doi.org/10.1002/hyp.7518>

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

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

810 Pathiraja, S., Moradkhani, H., Marshall, L., Sharma, A., Geenens, G., 2018. Data-driven model
811 uncertainty estimation in hydrologic data assimilation. *Water Resour. Res.* 54, 1252–
812 1280.

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

816 Pedinotti, V., Boone, A., Ricci, S., Biancamaria, S., Mognard, N., 2014. Assimilation of satellite
817 data to optimize large-scale hydrological model parameters: a case study for the SWOT
818 mission. *Hydrol. Earth Syst. Sci.* 18, 4485–4507. [https://doi.org/10.5194/hess-18-4485-](https://doi.org/10.5194/hess-18-4485-2014)
819 2014

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

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

826 Pinter, N., Santos, N., Hui, R., 2017. Preliminary analysis of Hurricane Harvey flooding in
827 Harris County, Texas. Retrieved UC Davis Cent. Watershed Sci.-Calif. WaterBlog
828 [https://california.waterblog](https://california.waterblog.com/2017/09/01/preliminary-analysis-hurricane-harvey-flooding-harris-county-texas/)
829 [Com20170901preliminary-Anal.--Hurricaneharvey-Flooding--](https://california.waterblog.com/2017/09/01/preliminary-analysis-hurricane-harvey-flooding-harris-county-texas/)
[Harris-Cty.-Tex.](https://california.waterblog.com/2017/09/01/preliminary-analysis-hurricane-harvey-flooding-harris-county-texas/)

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

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

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

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

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

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

847 Savage, J.T.S., Bates, P., Freer, J., Neal, J., Aronica, G., 2016. When does spatial resolution
848 become spurious in probabilistic flood inundation predictions? *Hydrol. Process.* 30,
849 2014–2032. <https://doi.org/10.1002/hyp.10749>

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

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

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

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

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

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

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

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

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

879
880