Wonhyun Lee[1], Alexander Y Sun[1], and Bridget R Scanlon[1]

[1]Bureau of Economic Geology, Jackson School of Geoscience, The University of Texas at Austin

February 28, 2024

Abstract

Accurately predicting the extent of compound flooding events, including storm surge, pluvial, and fluvial flooding, is vital for protecting coastal communities. However, high computational demands associated with detailed probabilistic models highlight the need for simplified models to enable rapid forecasting. The objective of this study was to assess the accuracy and efficiency of a reduced-complexity, hydrodynamic solver – the Super-Fast INundation of CoastS (SFINCS) model – in a probabilistic ensemble simulation setting, using Hurricane Ike (2008) in the Texas Gulf Coast as a case study. Results show that the SFINCS-based framework can provide probabilistic outputs under reasonable simulation times (e.g., less than 4 hours for a 100-member ensemble on a single CPU). The model agrees well with observed data from NOAA tidal stations and USGS gage height stations. The ensemble approach significantly reduced errors (average 16%) across all stations compared to a deterministic case. The ensemble improved overall performance and revealed wider flood extents and lower depths. Sensitivity studies performed on ensemble sizes (1,000, 189, 81) and lead times (1 to 3 days before landfall) further demonstrate the reliability of flood extent predictions over varying lead times. In particular, Counties adjacent to the Trinity River Basin had [?] 80% probability in exceeding the critical 3-m flood threshold during Hurricane Ike. Our study highlights the effectiveness of the SFINCS-based framework in providing probabilistic flood extent/depth forecasts over long lead times in a timely manner. Thus, the framework constitutes a valuable tool for effective flood preparedness and response planning during compound flooding.

Probabilistic Storm Surge and Flood-Inundation Modeling for the Texas Gulf Coast Using Super-Fast INundation of CoastS (SFINCS)

Wonhyun Lee[1*], Alexander Y. Sun[1] and Bridget R. Scanlon[1]

[1]Bureau of Economic Geology, Jackson School of Geoscience, The University of Texas at Austin, USA.

*Corresponding author: Wonhyun Lee (wonhyun.lee@beg.utexas.edu)

Key Points:

- An ensemble framework was developed for rapid probabilistic flood extent/depth forecasting and benchmarked over Hurricane Ike (2008).
- A reduced-complexity hydrodynamic solver is at the core of the framework to provide fast flood simulations.
- The ensemble framework can provide reliable probabilistic results at long lead times (3-days before landfall).

Abstract

Accurately predicting the extent of compound flooding events, including storm surge, pluvial, and fluvial flooding, is vital for protecting coastal communities. However, high computational demands associated with detailed probabilistic models highlight the need for simplified models to enable rapid forecasting. The objective of this study was to assess the accuracy and efficiency of a reduced-complexity, hydrodynamic solver – the Super-Fast INundation of CoastS (SFINCS) model – in a probabilistic ensemble simulation setting, using Hurricane Ike (2008) in the Texas Gulf Coast as a case study. Results show that the SFINCS-based framework can provide probabilistic outputs under reasonable simulation times (e.g., less than 4 hours for a 100-member ensemble on a single CPU). The model agrees well with observed data from NOAA tidal stations and USGS gage height stations. The ensemble approach significantly reduced errors (average 16%) across all stations compared to a deterministic case. The ensemble improved overall performance and revealed wider flood extents and lower depths. Sensitivity studies performed on ensemble sizes (1,000, 189, 81) and lead times (1 to 3 days before landfall) further demonstrate the reliability of flood extent predictions over varying lead times. In particular, Counties adjacent to the Trinity River Basin had ≥ 80% probability in exceeding the critical 3-m flood threshold during Hurricane Ike. Our study highlights the effectiveness of the SFINCS-based framework in providing probabilistic flood extent/depth forecasts over long lead times in a timely manner. Thus, the framework constitutes a valuable tool for effective flood preparedness and response planning during compound flooding.

Plain Language Summary

Understanding and predicting compound floods caused by multiple drivers, including storm surge, extreme rainfall, and river discharge, is important for protecting coastal areas. This study tested a reduced complexity solver called SFINCS to determine if it could quickly and accurately forecast floods using Hurricane Ike (2008) in Texas as a case study. The SFINCS-based ensemble framework accurately predicted flooding patterns and depths, running on average 15-30 times faster than traditional hydrodynamic models. By simulating many ensembles, it showed which areas are at a high risk of flood inundation (> 3-m) which should help communities better prepare for future floods. Our research demonstrates that employing this SFINCS-based ensemble approach can enhance the accuracy of compound flood predictions, helping coastal communities in mitigating flood risk.

1. Introduction

Tropical cyclones (TCs), including hurricanes and tropical storms, pose substantial threats to coastal areas. In particular, compound flooding caused by the simultaneous occurrence of storm surge, pluvial flooding from heavy rainfall, and fluvial flooding from river discharge, can inflict catastrophic impact on coastal communities. Despite significant recent progress in TC forecasting, accurately predicting compound flooding remains a formidable challenge due to the complex interplay of factors such as weather conditions, ocean temperatures, geography and human intervention, especially for vulnerable regions like the Texas Gulf Coast. For instance, hurricanes Ike (2008) and Harvey (2017) that struck the Texas Gulf Coast represent some of the costliest storms in the U.S. history, with estimated ~$160 billion damage (combined both events, National Hurricane Center, 2018). These TCs highlighted the urgent need for improved nowcasting and forecasting techniques for compound flooding events (Lee et al., 2023; Wahl et al., 2015). Hurricane Ike has been extensively studied, with researches focusing on various aspects of its dynamics, coastal impacts, and infrastructure effects (Stearns & Padgett, 2012). Several studies have investigated the simulation of hydrodynamics and waves (Chen & Curcic, 2016; Hope et al., 2013; Maymandi et al., 2022; Veeramony et al., 2012; Xu et al., 2023), as well as methodologies for addressing storm surge and wave heights (Kennedy et al., 2011; Lee et al., 2017). Mitigating these risks requires a comprehensive understanding of TC-induced hazards and effective decision-making tools. However, traditional evacuation decisions during hurricanes often prioritize storm intensity, overlooking other critical factors such as compound flooding risks. One of the significant challenges in TC forecasting is the limited predictability, particularly regarding flood-related hazards. Knutson et al., (2010) and Wang et al., (2015) highlight the difficulties in aggregating forecast data across regions because of geographical variations and limited sample size of forecast TCs affecting each specific region. Understanding the interconnectivity among various forecast components is crucial for improving TC flood predictability. Lamers et al., (2023) and Nederhoff et al., (2023) highlight the importance of investigating the relationships between TC track, intensity, precipitation, and river discharge forecasts to enhance flood warning and preparedness activities.

To assess uncertainty in flood forecasts, there is a recognized need to shift towards probabilistic forecasting of downstream hazards. Titley et al., (2024) and Wright et al., (2015) stress the importance of probabilistic approaches in optimizing forecast guidance, especially for rare events like TC-induced flooding. A probabilistic surge and flood-inundation modeling system would provide coastal communities with the probability of occurrence for different surface water depth thresholds, supporting assessment of surge and flood risks, design of resilient infrastructure, as well as decision-making for coastal planning and management. Large ensemble modeling analysis is required to provide probabilities of TC-induced flooding extents and depths, particularly from compound flooding, aiding risk assessment and decision-making. While existing techniques provide essential information, challenges remain in addressing uncertainties and capturing complex interactions. Methods like the Monte Carlo Wind Speed Probability (WSP) model and landfall distribution product (LDP) effectively communicate intensity uncertainties (Trabing et al., 2023). However, challenges persist in accurately predicting TC tracks and intensity changes, necessitating further development of probabilistic forecasting methods (Torn & DeMaria, 2021). The Monte Carlo approach is employed to generate ensemble members, introducing random variations in initial conditions, and utilizing error matrices from the previous step. Specifically, these ensemble members are generated around the forecasted official track, and the method incorporates error matrices based on an autoregressive technique for along-track, cross-track, and intensity errors (DeMaria et al., 2009). Within the context of probabilistic ensemble simulation, such as Monte Carlo methods applied to hurricane forecasting, the ensemble size refers to the number of simulated scenarios or members used to capture the uncertainty associated with various input parameters (Cashwell & Everett, 1957; DeMaria et al., 2009; Nederhoff et al., 2023). In ensemble forecasting, the lead time of a hurricane's landfall plays a crucial role in determining the accuracy of mesoscale meteorological simulations (Nederhoff et al., 2023; Titley et al., 2020; Toth & Buizza, 2019) due to its direct correlation with the uncertainty associated with the impending landfall (Trabing et al., 2023). When a model accurately predicts the hurricane's path and intensity with an extended lead time, it enhances confidence in the model's reliability. Various probabilistic modeling systems and forecasting techniques have been developed to address TC-induced hazards. Global Flood Awareness System (GloFAS), despite its coarse resolution (~10 km), offers valuable insights globally (Alfieri et al., 2013; Harrigan et al., 2023) while regional systems such as the Stevens Flood Advisory System (SFAS) provide higher resolution for specific areas (Ayyad et al., 2022; Tounsi et al., 2023). However, these systems often lack explicit consideration of TC-related processes and interactions, requiring further research.

Further research is necessary for refining probabilistic modeling systems, enhancing ensemble forecasting techniques, and integrating statistical and physics-driven approaches for comprehensive TC forecasting, impact assessment, and risk management. Hybrid approaches, combining probabilistic and deterministic methods, show promise in accurately representing a wider range of scenarios (Bakker et al., 2022; Pakhale et al., 2024). Artificial Intelligence (AI) and Machine learning (ML) techniques, though increasingly popular, may overlook critical nonlinear interactions (Chen et al., 2023; Kumar et al., 2023; Lecacheux et al., 2021). Combining statistical and physics-driven modeling, as proposed by Titley et al., (2024), offers a comprehensive approach to address these challenges.

One limitation of ensemble-based approaches is the computational resources required to generate and analyze ensemble forecasts. Ensemble forecasting involves running multiple simulations to capture the uncertainty inherent in weather and climate predictions. Each ensemble member requires significant computational power and time to execute, even with the use of a high-performance computing (HPC) systems. Improved efficiency of ensemble simulations requires fast solvers, such as the open-source hydrodynamic solver, Super-Fast INundation of CoastS (SFINCS Leijnse et al., 2021) or similar models (e.g., LISFLOOD-FP: Bates et al., 2010, SLOSH: NOAA, 2006), to forecast and predict compound flood events. SFINCS integrates various flood drivers, including pluvial and fluvial drivers along with tidal, wind- and wave-driven processes into a single domain, allowing comprehensive analysis of compound flooding events (Eilander et al., 2023; Grimley et al., 2022; Leijnse et al., 2023; Nederhoff et al., 2023). Additionally, the model offers a significantly faster computational speed (~15–30 times speedup, Röbke et al., 2021) than the more complex models, such as ADCIRC (Luettich et al., 1992), Delft3D-FM (Kernkamp et al., 2011), and SCHISM (Zhang et al., 2016).

The reduced-complexity approach taken within SFINCS serves as an alternative to more complex models that are computationally intensive. SFINCS uses a first-order explicit numerical scheme (Bates et al., 2010) to solve a set of simplified depth-averaged (linear) shallow water equations. Certain physical processes (e.g., viscosity, atmospheric pressure, Coriolis, advection) that are less critical for specific predictions can be turned off to reduce computational demand, but the reduced solver with essential physics ensures both efficiency and accuracy in the model.

This innovative approach provides a distinct advantage over conventional models by significantly reducing computational requirements while maintaining a high level of accuracy. It presents a new outlook on flood prediction methodologies, placing emphasis on efficiency without compromising the reliability of forecasts. This approach significantly contributes to the timely and dependable prediction of complex storm surge and flood events, ensuring greater accuracy while optimizing computational resources.

The objective of this study was to develop a probabilistic ensemble approach to simulate compound flood events using a reduced complexity solver applied to a typical TC in the Texas Gulf Coast. Novel aspects of the work include comprehensive modeling of compound flooding, including storm surge, pluvial and fluvial components, application of a probabilistic framework to provide detailed uncertainty quantification, and assessment of different lead times to evaluate forecasting skill. We selected the SFINCS model because of its low computational intensity and ability to model compound events. SFINCS incorporates spatially and temporally detailed data, including bathymetry, storm winds, land use patterns, oceanographic conditions, and meteorological factors within the target domain. SFINCS is relatively new; therefore, detailed validation studies are warranted to test its accuracy and applicability.

Hurricane Ike's devastating impact on Texas in 2008 was primarily marked by powerful wind-driven storm surge. Although pluvial and fluvial flooding played roles as well, it is the storm surge that stands out as the defining factor in Ike's destructive force Fox News, 2015; Morss & Hayden, 2010; Rego & Li, 2010. However, in light of the global significance of compound floods Eilander et al., 2023; Gu et al., 2022, this study considered the drivers of pluvial and fluvial flooding. Ike was selected in part because its track is similar to the catastrophic 1900 Galveston hurricane (Trumbla, 2019). The evolution of Hurricane Ike from a tropical wave west of Cape Verde to a Category 4 hurricane over the central Atlantic, its fluctuations in strength, and its subsequent landfalls in Galveston, Texas, offer a rich dataset for comprehensive analysis. Moreover, Hurricane Ike's extensive aftermath, marked by its toll on human lives (mortality 195, National Hurricane Center, 2014) and infrastructure, makes it an important case study. In this analysis, all relevant processes of compound flooding events, including contributions from rainfall (pluvial), river runoff (fluvial) and surface processes (e.g. water levels, tides, storm surge) are incorporated to provide a comprehensive probabilistic modeling framework as shown in the flowchart (Fig. 1). In Section 2, we provide an in-depth exploration of the SFINCS model, its configurations, and ensemble approach. Moving to Section 3, we discuss the analysis of the model results, presenting the surge and flood maps (height and extent). In Section 4, we provide a discussion of the sensitivity analysis (ensemble sizes and lead times), alongside an assessment of its computational efficiency and accuracy.

2. Materials and Methods

2.1 Deterministic Model Configuration and Data Sources

The model extent, boundaries (e.g., county and watersheds), and best track of Hurricane Ike (2008) are shown in Fig. 2. The cartesian rectilinear grids in UTM 15N were first created with resolutions of 200 × 200 m and a subgrid mode with resolutions of 10 × 10 m was applied in this study.

Bathymetric-topographic datasets for this specific region were sourced from the National Oceanic and Atmospheric Administration - National Centers for Environmental Information (NOAA NCEI) Continuously Updated Digital Elevation Model (CUDEM, available in both 10 m and 3 m resolutions) database (CIRES, 2014a, 2014b). Manning's roughness coefficient was obtained from the National Land Cover Database (NLCD-CONUS, Homer et al., 2020). To calculate the volume of runoff during the infiltration process, a Global curve number dataset (Jaafar et al. 2019) with a spatial resolution of 250 m was utilized to derive a runoff coefficient.

Boundary conditions play an important role in accurately simulating tidal effects at open ocean boundaries. Within our grid-domain, an offshore water level boundary was established. The variability in water levels along these boundaries has been derived from HYCOM see surface height (SSH) data, encompassing five tidal constituents (M2, S2, O1, K1, N2) at three-hour intervals, with a spatial resolution of ~ 4 km. Additionally, we incorporated the primary river discharge (m^3/s) for each watershed, as shown in Fig. 2. The nine upstream boundary conditions (Clear Creek, Sims Bayou, Brays Bayou, White Oak Bayou, Little White Oak, Green Bayou, West Fork of San Jacinto River, Cedar Bayou, Trinity River) (streamflow, m^3/s) were defined using river discharge time series obtained from accessible USGS stream gages (15-min, USGS, 2016) during Ike 2008.

We used wind field data sourced from the National Hurricane Center-Joint Typhoon Warning Center (JTWC) best track, employing the Holland formula (Holland et al., 2010) and the R_{max} relationship (Nederhoff et al., 2019). We leverage the Wind Enhance Scheme (WES, Deltares, 2018) developed by Deltares to generate the wind and pressure fields around a specified tropical cyclone center location based on various cyclone parameters. It computes 2D surface winds and pressure fields on a moving circular "spider" web grid. These datasets were compared against the observed data from the NOAA stations. Three wind drag coefficients were specified with the SFINCS wind speed framework (0.001 at 0 m/s, 0.0025 at 28 m/s, 0.0015 at 50 m/s). In addition to the meteorological forcing condition, the ERA5 (Hersbach et al., 2020) provides downscaled hourly, 31-km precipitation rates. The effect of wind-driven waves was not factored into the storm surge and flood outcomes in this study.

2.2 Ensemble Approach

Forecasting TCs is a complex task encompassing predictions of numerous interconnected factors, including storm tracks, winds (speed/direction/pressure), and rainfall. Even though the National Hurricane Center provides forecasts every 6 hours up to 72 hours before landfall, these predictions are not perfect. Persistent errors stem from incomplete understanding of the complex formation and progression of TCs, compounded by limitations inherent in forecasting methodologies. Therefore, recognizing the inherent variability in storm track and intensity is important when employing modeling tools for cyclone simulations. Our approach, inspired by DeMaria et al., (2009), utilizes a Monte Carlo method to generate an ensemble of predictions based on error matrices (e.g., addressing along-track (AT), cross-track (CT), and intensity errors). The error vector is decomposed into AT and CT components relative to the direction of the cyclone motion vector in the forecast track, with AT and CT errors determined using a simple autoregressive technique. For predicting the error of maximum wind intensity (VE), DeMaria et al., (2009) considered the distance to landfall. However, it is simplified by a linear function of the error from the previous time step in the Delft Dashboard (Van Ormondt et al., 2020). The ensemble tracks with different lead times (days before landfall) are shown in Fig. 3. The Interagency Performance Evaluation Task Force (IPET, 2006) method employs a comprehensive rainfall analysis technique to calculate the mean rainfall intensity over a specified region. By integrating critical parameters such as the distance r (in km) from the hurricane center to the point of interest and the azimuth β (in degrees) relative to the direction of motion, the IPET method aims to provide accurate estimates of ensemble rainfalls. During the evaluation process of ensemble rainfalls, the IPET method identified an underestimation in rainfall estimates. To rectify this discrepancy, the rainfall estimates (asymmetric component factor) were doubled based on the magnitude of the underestimation. The improved rainfall intensity in the area of interest was improved using the ensemble tracks (Fig. 4). This adjustment was important in enhancing the accuracy of the IPET rainfall estimates, aligning it closely with other reanalysis data during extreme weather events.

$$m_I(r, \beta) = \begin{cases} 1.14 + 0.12\Delta P, & r \le R_{max} \\ (1.14 + 0.12\Delta P)e^{-0.3\left(\frac{r-R_{max}}{R_{max}}\right)}, & r > R_{max} \end{cases} \qquad \text{Eq. 1}$$

where m_I is the azimuthally averaged component in mm/hr, and ΔP is the central pressure deficit in millibars (Eq. 1). An example of ensemble rainfalls using the IPET approach is shown in Fig. 3.

We investigated the implementation of a simple ensemble approach for river streamflow analysis in our case study. There are five ensembles, a raw data consisting of streamflow measurements recorded at 15-min intervals, 6-hourly maximum, minimum, and mean, and daily mean streamflow values. To comprehensively represent various ensemble configurations, all combinations were parameterized as detailed in Table 1 for the purpose of this study. The table summarizes the diverse ensemble combinations explored during the analysis, facilitating a thorough examination of their impacts and efficacy in streamflow estimation and prediction.

2.3 Calibration and Validation Approaches

We conducted multiple calibration steps in the modeling analysis, incorporating various adjustments, such as adopting different tidal boundary conditions at the open ocean boundary (Fig. 2), turning the infiltration process on and off, optimizing the intensity factor of the IPET rainfall approach, and conducting sensitivity analyses on ensemble generation. These steps were undertaken to minimize systematic uncertainties and enhance the overall performance of the modeling system. Initially, boundary conditions derived from the TPXO 7.2 and 8.0 tidal models, designed to represent tidal water levels, were assessed, but they underestimated water levels at the NOAA tidal gage stations. Subsequently, in an effort to enhance the accuracy of simulated surface water elevations, water levels sourced from the HYCOM (Cummings & Smedstad, 2013) SSH data, encompassing tidal constituents such as M2, S2, O1, K1, and N2, were incorporated into the model. In compound flood modeling, the infiltration dynamics are important because the ground surface partitions rainfall into evapotranspiration, runoff, and infiltration depending on soil texture and land use. SFINCS provides several rainfall-runoff processes, and we adopted the curve number method as an empirical rainfall-runoff approach. The infiltration process by the curve number method shows a relatively average flood reduction of ~8% across the entire model domain, with on and off infiltration during the calibration process. This reduction may be attributed to Hurricane Ike not bringing intense rainfall, allowing the infiltration to decrease flood levels. However, if infiltration capacity is surpassed during high-intensity rainfall, excess water can lead to rapid surface runoff and severe flooding. Therefore, this study aims to apply the infiltration process using the curve number method more accurately and practically to achieve more realistic results. Due to the challenges in collecting precipitation data for Hurricane Ike in 2008, we additionally conducted a comparison between precipitation estimates derived from IPET and other reanalysis datasets. The IPET rainfall data significantly underestimated rainfall values from ERA5 and NCEP-CFS (Fig. 4). To rectify this discrepancy, a calibration process was undertaken by doubling the rainfall intensity factor (a similar magnitude of rainfall in Fig. 4). To perform a comprehensive sensitivity analysis considering the influence of a number of ensemble members (i.e., ranging from 81, 189, and 1,000 ensembles) and lead times prior to landfall (1 to 3 days) on storm surge and flood predictions, we configured various ensemble cases (Table 1, i.e., the number of 81 derives from all possible combinations of error variations). The objective of the sensitivity analysis was to determine the optimal combination of lead times and number of ensemble members for refining our probabilistic ensemble modeling system.

Skill score metrics employed to assess model accuracy included Root-Mean-Square Error (RMSE), Pearson's Correlation Coefficient (CC), Mean-Absolute-Error (MAE), and Refined Index of Agreement (RIA). These metrics were calculated for all available observed stations (NOAA and USGS) using data and model predictions. Simulated flood event hydrographs were compared to observations at six representative USGS gages (gage water height, G1~G6 in Fig. 2). In addition, simulated water levels were compared to observations at six NOAA tides gages. The reported data were in feet, referencing the North American Vertical Datum of 1988 (NAVD 88). Both deterministic and ensemble cases were executed over a period of 13 days, spanning September 2nd through 15th, 2008, to simulate Hurricane Ike (2008) and its associated ensembles. Model outputs include the spatial extent and depth of surge and flood levels within the inundated extents over time.

3. Results

3.1 Model Validations – Water Level and Winds at NOAA Stations

The model output was compared to water levels at NOAA tide gauges and USGS gages (Fig. 5). The average error statistics across all six stations for water levels are: RMSE:0.32, CC: 94%, MAE: 0.18, RIA: 0.78 (Table 2). Examination of individual stations reveals variations in model performance. Notably, the upstream station at Manchester (A3) has relatively higher errors (RMSE: 1.11, MAE: 0.48). This discrepancy is attributed to a combination of geological features, grid resolutions, and coarse precipitation input. Implementing an ensemble approach emerges as an important strategy for rectifying errors. The aggregate statistics derived from 1,432 simulations highlight a substantial improvement in overall performance (RMSE: 0.2, CC: 95.7%, MAE: 0.14, RIA: 0.81) relative to the deterministic scenario. Notably, the discernible improvement at Station A3 stands out (i.e., average RMSE: 0.66, MAE 0.22) compared to deterministic case, underscoring the efficacy of the ensemble methodology in refining model predictions. Our study shows slightly lower errors in water levels than those in a recent Ike study (e.g., Al-Attabi et al., 2023).

In assessing the model representation of wind speed and direction (we do not estimate the errors in winds in this study), a detailed comparison was made with NOAA stations using the spider web wind speed and direction generated by the WES scheme. Despite the absence of data at A3 and A4 during the Ike 2008 event, both deterministic and ensemble mean analyses show good agreement across the remaining stations. This robust validation reinforces the model's ability to accurately capture and simulate wind dynamics, even in challenging scenarios. While some differences in wind direction phases are observed in the early stages, these variances are deemed negligible. Ensemble wind perturbation from a specific point in time (September 12, 12:00h, 2008) contributes to a comprehensive understanding of these variations.

3.2 Model Validations – Hydrographs at USGS Gage Heights

Overall, the model shows good agreement with the USGS measurements, although variations are observed among specific stations (Fig. 5). Simulated hydrographs at Stations, G1, G2, and G3 show relatively strong correlations with observed hydrograph data. The remaining three stations (G4, G5 and G6) also show high correlations, but have relatively larger errors (RMSE, MAE and RIA), indicating greater discrepancy between model output and observations. To enhance model performance, an ensemble approach was used, resulting in a modest reduction in MAE and slightly improved error statistics for the ensemble mean. Stations G3 and G5 specifically resulted in ~ 30% and 16% reductions in MAE, and ~39% and 28% in RMSE reductions, respectively, based on the results from 1,432 ensemble members. However, other stations show negligible improvement. Station G4 has a negative RIA, indicating a larger discrepancy. Despite G4’s high spatial resolution of ~200 m and 10 m in sub-grid, the model overestimates the water levels, possibly due to complex dynamics, geological features (such as narrow channels), roughness variations, grid resolution, and interactions among other factors. Furthermore, the model has limitations in defining specific small river channels and features like weirs and dams. This limitation hinders accurate representation of upstream river and floodplain hydrodynamics, potentially impacting the model’s ability to predict hydrographs effectively. Addressing these factors is important for refining the model's accuracy and reliability in simulating hydrological processes.

3.3 Hurricane Ike 2008 – Flood inundation map

To estimate flood extent, we overlaid simulated water levels relative to a map of permanent water, utilizing the Global Surface Water Occurrence (GSWO) dataset (Pekel et al., 2016). This process involved applying a threshold depth of 0.3 m. Specifically, areas with water depths exceeding this threshold are identified and categorized as flooded. Our models (both deterministic and ensemble in Fig. 6) were further assessed through a comprehensive comparison with flood-inundation maps from multiple sources, including the Harris County Flood Control District (Fig. 6-a), NOAA estimates (Fig. 6-b), and a recent case study by Al-Attabi et al., (2023). Despite differences in map units, our models consistently show similar depths and extents of flood inundation to those from these reference sources. Flood-inundation maps from the Harris County Flood Control District serve as a benchmark for local accuracy, while NOAA estimates provide a broader perspective at a regional scale. Additionally, the recent study by Al-Attabi et al., (2023), specifically utilizing the Delft3D model, provides a valuable benchmark for comparison. The outputs from our models are similar to those from these other sources for Hurricane Ike, reinforcing the robustness of SFINCS model predictions. This alignment with an independent case study increases confidence in the accuracy and reliability of the SFINCS probabilistic modeling framework. Our model highlighted significant flooding exceeding 3 m, in Chambers County and the open Bay of Jefferson County. Coastal areas in Brazoria and Galveston counties experienced flooding in the range of 1 to 2 m. The correspondence between our model outputs and these specific observations (as Fig. 5) underscores the reliability of our models in capturing the spatial distribution and magnitude of flood inundation in diverse geographical settings. The model assesses flooding patterns in various watersheds across multiple counties, including Harris, Liberty, Galveston, and Brazoria counties. It demonstrates a satisfactory skill in identifying low levels of flooding in specific areas within these counties. Moreover, the model identifies flooding downstream of Dickinson Bayou, indicating a large-flooded area, likely attributed to storm surge effects. One key observation highlighted by the model pertains to the Upper Gulf Coast, specifically in West Bay and Galveston Barrier Island. The analysis revealed a surge flood event during Hurricane Ike in 2008. However, an interesting aspect is the discrepancy revealed by ensembles in flood extent and depth. Despite perceived non-flooding in the deterministic case, the ensemble model shows a broader flood extent with lower flood depth. This suggests that the model, through ensemble approaches, captures a more extensive spatial coverage of flooding, even in areas considered to have a lower risk. In specific watersheds within Harris, Galveston, and Brazoria counties, the ensemble approach consistently shows more widespread flooding. This underscores the effectiveness of the model in representing the potential for extensive flooding in these regions. In the Trinity River Basin, the ensemble-mean results show significant differences (> 2 m, Fig. 6-e) compared to the deterministic scenario, attributed to geographical characteristics and contributions from river discharge and storm surge-induced flooding. This implies both the substantial geographic impact and the uncertainty associated with the ensemble approach. Additionally, variations of ~ 0 to 1.3 m in flood depth are evident in Chambers and Jefferson counties, indicating a reliable extent and depth of flooding. The ensemble scenario closely aligns with the deterministic scenario in Galveston Bay and Barrier Island. Overall, our analysis demonstrates the ability of the SFINCS model to define flooding patterns, considering both spatial extent and depth variations, contributing valuable information for flood risk assessment and mitigation strategies. The analysis focuses on evaluation of maximum flood depths (Fig. 7) within the context of various ensemble numbers in comparison with deterministic and ensemble mean scenarios. In particular, ensembles #167, #169, and #184 exhibit distinct characteristics, displaying a few flooded areas in Liberty, Jefferson, and Chambers counties. This contrasts with both deterministic and ensemble mean outcomes, suggesting potential influence from disparate storm tracks and intensities. These specific cases may lead storm tracks designed to deviate downward from the best track, potentially explaining the observed wider and higher flooding depths. Furthermore, ensembles #50, #59, and #88 appear to have storm tracks slightly above the best track, resulting in less rainfall-induced flooding in Liberty County. This divergence is likely attributed to lower flood depths observed in the West Bay, Galveston Barrier Island, and Trinity River Basin. Thus, the interplay of storm tracks and flood depths underscores the significance of ensemble modeling in capturing the variability of flood scenarios.

4. Discussion

4.1 Influence of Ensemble Size

Ensemble size plays an important role in probabilistic ensemble simulation, significantly impacting the reliability and accuracy of the simulation results (Buizza & Palmer, 1998; Milinski et al., 2020; Tebaldi et al., 2021). In particular, ensemble approaches show significant error reduction in Stations A1 and A3 (Fig. 8, and Table 2). The diverse ranges within the ensemble are visually depicted in pink, highlighting the variability encapsulated by different ensemble sizes. The ensemble approach seems to demonstrate a slightly better error performance when compared to the deterministic case, although the difference seems quite small (Table 2). This observation is particularly evident in the USGS hydrographs (Fig. 9), indicating a notable influence from upstream boundary conditions rather than wind and rainfall, a characteristic consistently observed across a comparison of six different USGS stations. Notably, when compared with Fig. 6, it is evident that the majority of flooding in Liberty and Harris counties is confined to levels below 1 m. Beyond flooding probabilities of 2 m, distinctive variations emerge in the Trinity River Basin, Chambers, and Jefferson counties. A reduction in ensemble size may reveal an expanding range in flooding probability, likely indicating an increase in marginal error attributed to increased uncertainty. Hurricane Ike led to substantial flooding in Chambers County and the open bay of Jefferson County, featuring a flooding probability exceeding 3 m. Diverse ensemble sizes consistently depict the elevated flood risks in these specific regions. Although larger ensemble sizes may yield marginal enhancements in accuracy and convergence rates, they are associated with increased computational demands. Alternatively, smaller ensemble size, while conceding a degree of precision, rapidly deliver practical and reasonable information regarding flood extents and probability maps. The aftermath of Hurricane Ike distinctly revealed severe flooding in southwestern Chambers and Jefferson counties, highlighting the efficacy of the ensemble approach in verifying elevated flood risks within these critical areas.

4.2 Influence of Lead Time before Landfall

Comparison of the output from 81 ensembles to the output from the deterministic case across different lead time shows improved performance of the ensemble (RMSE: 0.2, CC: 95.6%, MAE: 0.14, RIA: 0.82). Overall, the error statistics exhibit a remarkably similar pattern, but as the lead time decreases, the accuracy is slightly improved. This is particularly evident when examining stations A1 and A3. In the comparison of USGS hydrographs, the ensemble approach does not show significant improvement (Fig. 12). However, it shows a slight enhancement compared to deterministic scenarios (avg. RMSE: 1.6, CC: 88.3%, MAE: 1.13, RIA: 0.43, over six stations). The average errors for the ensemble approach are RMSE of 1.35, CC of 88.1%, MAE of 1.02, and RIA of 0.48. Similar to the analysis of ensemble size in section 4.1, there is no noteworthy variation in hydrographs based on lead times. It appears that the hydrographs are significantly influenced by upstream boundary forcing. With increasing lead times, the probability of flooding above 1 m slightly decreases in Chambers and Jefferson counties.

Significant changes are observed in the Trinity River Basin for the above 2 m flooding probability. A shorter lead time indicates increased variability in flood extent, particularly observable in the Trinity River Basin, Chambers, and Jefferson counties. Analysis of different lead times for Hurricane Ike reveals that a relatively long lead time (3-days) results in similar flood extent and flooding probability, compared with 1-day lead time. This reliability allows us to rely on the flooding probability prediction three days ahead of flooding, leveraging the extended lead time. During Hurricane Ike, Harris County, with its dense population, did not face significant risks. Drawing a comparison to Hurricane Harvey in 2017, understanding the factors that caused tremendous damages in Harris County can aid in more effective probabilistic predictions with sufficient lead times. This consistency in findings aligns with previous research (Huang et al., 2016; White et al., 2017), supporting preparedness, decision-making (Rezuanul Islam et al., 2023), and response efforts.

Ultimately, these contributions enhance the safety and well-being of communities. Additionally, time-series of 95% confidence intervals were estimated for each ensemble scenario. While the ensemble means of each scenario exhibit relatively similar values, variations in ensemble spread are noticeable. It is apparent that the 95% bandwidth relative to ensemble members decreases as the ensemble size increases as shown for NOAA Station 3 (Fig. 14). This suggests a narrower bandwidth, indicating a smaller margin of error from the mean flood depth within 5% of ensemble scenarios. Scenarios with different lead times exhibit similar patterns. However, as lead time decreases (approaching landfall), the ensemble spread tends to increase, and there is a slight rise in the predicted water levels. This trend is also reflected in a marginal decrease in MAE (Table 3). Despite differences in ensemble spread and extent of the 95% range, these trends were consistently observed when comparing water levels across various NOAA stations. In the Hurricane Ike case study, discernible differences in ensemble means were not apparent among various ensemble scenarios. The similarity in wind-driven water level changes among ensemble members can arise when the initial conditions and error matrices used in the Monte Carlo method exhibit minimal variation with different lead times. Therefore, a stable set of initial conditions, consistent error matrices, and use of autoregressive techniques collectively contribute to the production of ensemble members that display comparable wind-driven water level changes, even with varying lead times in the ensemble approach.

While the spatial patterns of flood inundation remain similar across various ensemble sizes and lead times, there is a notable increase in uncertainty for the relatively small number of ensembles (81). Moreover, the 3-days lead time (long lead time) displays a broad confidence interval, signaling a higher level of uncertainty (Fig. 14). Hurricane Ike is characterized by flooding predominantly driven by wind-driven storm surge Morss & Hayden, 2010; Rego & Li, 2010. This is exemplified through diverse ensemble scenarios, particularly evident in the flooding incidents in West Bay, Galveston Barrier Island, Chambers, and Jefferson counties.

4.3 Efficiency and Accuracy of Probabilistic Ensemble Modeling

The use of a reduced-complexity probabilistic modeling system in this study facilitates rapid predictions of surge, flooding levels, and extents. Furthermore, it demonstrates promising efficiency in statistically representing water surface elevations in coastal areas and water height in inland regions. For example, when dealing with compound conditions in a 13-day simulation, the 1,000-, 189-, and 81-ensemble modes required 1.9 days, 8.4 hours, and 3.6 hours of computational time (using a single multi-core CPU), respectively. The resultant error statistics (Tables 2 and 3) prove adequate in assessing the model’s reliability and the quality of its predictions within a condensed timeframe. This highlights the system’s effectiveness in producing dependable predictions while efficiently managing computational resources. The choice of ensemble size should be made based on the specific requirements of the application. If high accuracy is critical and computational resources are available, a larger ensemble, such as 1,000, might be preferred. However, in time-sensitive situations or when computational resources are limited, smaller ensembles, such as 189 or 81, may be more practical choices, providing a balance between accuracy and computational efficiency. Additionally, accurate representation of boundary conditions, especially water level and tidal conditions, is important for reliable predictions in coastal and inland regions.

4.4 Future Work

In light of the global impact and growing relevance of compound floods (Eilander et al., 2023; Gu et al., 2022; Lai et al., 2021; Lee et al., 2023), this study is of continued importance in understanding the probabilistic nature of compound events, with important implications for effective flood management strategies. To further advance this field, we will focus on achieving higher resolutions in bathymetry within localized areas. This will encompass a comprehensive analysis of infiltration processes and integration of novel data sources, such as bathymetry, wind patterns, high-resolution curve numbers, updated bottom roughness, river data, and projections of forthcoming extreme rainfall events. These will enable enhancement of existing models by investigating frequency and return periods of each driver contributing to compound flooding. As part of future initiatives, priorities will include development of an advanced river ensemble system to enhance river management, leveraging artificial intelligence (AI)/machine learning (ML) for flood-inundation mapping using SFINCS model outputs, and expanding the scope of hurricane-impact assessment by considering other hurricanes along the Texas Gulf Coast. Employing AI/ML techniques, as demonstrated by Sun et al., (2023), will facilitate rapid flood-inundation mapping based on model outputs from the SFINCS model. Furthermore, a robust framework rooted in coastal digital twins will be established, extending the visualization capabilities for storm surges and compound flooding, offering insights into flood levels and extents. These progressive steps aim to improve our understanding and response capabilities in the face of compound flooding scenarios, thereby contributing to more resilient and efficient disaster management strategies.

5. Conclusions

Our study demonstrates the effectiveness of SFINCS in providing accurate and timely predictions of storm surge and flood events. Adopting an ensemble-based approach was pivotal in reducing errors, leading to substantial improvements across all stations. Similarly, a comparison of wind speed and direction with NOAA stations affirmed the model's capability to accurately capture wind dynamics. Despite some early-stage variances, the robust validation underscored the model's reliability, even in challenging scenarios. The validation of hydrographs revealed a good level of comparability with USGS measurements, albeit with variations among specific stations. Implementing an ensemble approach resulted in modest error reduction and improved error statistics (MAE ~ 0.22 m in water level, ~ 0.21 m in hydrograph) relative to the deterministic case. To assess flood extent, a method of masking water depth based on permanent water maps was employed. This is in agreement with other Ike products from HCFCD, NOAA and Al-Attabi et al., (2023), further bolstering confidence in the model predictions, especially concerning flood dynamics during Hurricane Ike. Chambers and Jefferson counties, along with the Trinity River Basin, exhibit significant flood risks, particularly exceeding 3 m, highlighting the importance of ensemble modeling in identifying critical flood-prone areas. While larger ensemble sizes could potentially result in marginal improvements in accuracy, they also come with increased computational demands. Therefore, opting for smaller ensemble sizes within our framework, which incorporate probabilistic sampling errors, could be a more pragmatic approach for rapid flood risk assessment. In flooding probabilities at different lead times before Hurricane Ike's landfall, significant changes were found in the Trinity River Basin for probabilities above 2~3 m. In addition, the uncertainty quantification and long lead times play important roles in enhancing flood preparedness efforts. The study highlights the importance of estimating time-series of 95% confidence intervals for each ensemble scenario, which provide valuable insights into the variability and reliability of flood predictions. As the ensemble size increases, the 95% confidence interval narrows, indicating a smaller margin of error in predicting flood depths. However, as lead time decreases, there is a tendency for the ensemble spread to increase, accompanied by a slight rise in predicted water levels. While spatial patterns of flood inundation remain consistent across various ensemble sizes and lead times, there is a notable increase in uncertainty with smaller ensemble sizes and longer lead times. Specifically, the 3-days lead time exhibits a broad confidence interval, indicating higher uncertainty. However, this study underscores the critical role of ensemble size in probabilistic ensemble simulation, with both small and large member ensembles yielding similar results. The analysis demonstrates that ensembles with lead times ranging from 1 day to 3 days provide comparable statistics, suggesting reliable flood extent and depth predictions three days in advance. This analysis underscores the reliability of flood extent predictions three days before landfall, particularly beneficial for preparedness and response efforts. Sensitivity studies on ensemble size and lead times provided valuable insights into forecasting precision. Ensemble approaches demonstrated better performance (relative to a single deterministic approach) across different lead times, contributing to enhanced accuracy and reliability in hurricane forecasting methodologies.

The utilization of reduced-complexity model-based probabilistic modeling systems facilitated rapid predictions while efficiently managing computational resources with up to 1,000 members of ensemble simulations in 1.9 days (e.g., with ~100 member ensembles totaling ~ 4 hour). Depending on specific requirements and computational constraints, the choice of ensemble size can be varied, offering a balance between accuracy and efficiency. Overall, our study highlights the effectiveness of ensemble approaches in improving model accuracy and reliability, offering valuable insights into flood dynamics and enhancing preparedness and response efforts for water-related events.

Acknowledgments

This research has been supported by Department of Energy, Advanced Scientific Computing Research Program under grant no. DESC0022211. The authors thank all collaborators for sharing data sources, and interactive discussion. The authors would also like to thank anonymous reviewers for their valuable feedback during the submission and editing process of this manuscript.

Conflict of interest

The authors declare that we have no conflict of interest/competing interests.

Authorship contribution statement

W. Lee designed the concept and methodology of integrated modeling system and constructed the SFINCS model, its validations. And W. Lee analyzed the model results and provided the initial draft of the manuscript. A. Sun and B. R. Scanlon contributed to editing and writing processes. All authors actively contributed to analyzing model outcomes and revisions in writing process, and agreed to publish the manuscript.

Data Availability

All dataset used in this study are publicly available as follow:

a) CUDEM (1/9 and 1/3 arc-sec) bathymetry data is available at https://coast.noaa.gov/htdata/raster2/elevation/.

b) wind and surface pressure, and precipitation inputs (ERA5); best storm track from NHC/JTWC (https://www.nhc.noaa.gov/data/, https://www.metoc.navy.mil/jtwc/jtwc.html?western-pacific) and era5 reanalysis dataset https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form.

c) dataset of precipitation rate: NCEP-Climate Forecast System Version 2 (CFSv2) hourly time-series products https://rda.ucar.edu/datasets/ds094-1/dataaccess/

d) NLCD Land Cover dataset for Manning coefficient: https://www.mrlc.gov/data/nlcd-2016-land-cover-conus

e) Global curve number: https://www.ndbc.noaa.gov/

f) water level validation with NOAA stations: https://tidesandcurrents.noaa.gov/stations.html?type=Water+Levels

g) hydrographs with USGS stations: https://waterdata.usgs.gov/nwis

h) Global surface water: https://global-surface-water.appspot.com/download

i) HYCOM SSH data: https://www.hycom.org/hycom/overview

References

Al-Attabi, Z., Xu, Y., Tso, G., & Narayan, S. (2023). The impacts of tidal wetland loss and coastal development on storm surge damages to people and property: a Hurricane Ike case-study. *Scientific Reports*, *13*(1), 4620. https://doi.org/10.1038/s41598-023-31409-x

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., & Pappenberger, F. (2013). GloFAS – global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, *17*(3), 1161–1175. https://doi.org/10.5194/hess-17-1161-2013

Ayyad, M., Orton, P. M., El Safty, H., Chen, Z., & Hajj, M. R. (2022). Ensemble forecast for storm tide and resurgence from Tropical Cyclone Isaias. *Weather and Climate Extremes*, *38*, 100504. https://doi.org/10.1016/j.wace.2022.100504

Bakker, T. M., Antolínez, J. A. A., Leijnse, T. W. B., Pearson, S. G., & Giardino, A. (2022). Estimating tropical cyclone-induced wind, waves, and surge: A general methodology based on representative tracks. *Coastal Engineering*, *176*(104154), 1–17. https://doi.org/10.1016/j.coastaleng.2022.104154

Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, *387*(1), 33–45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Buizza, R., & Palmer, T. N. (1998). Impact of Ensemble Size on Ensemble Prediction. *Monthly Weather Review*, *126*(9), 2503–2518. https://doi.org/10.1175/1520-0493(1998)126<2503:IOESOE>2.0.CO;2

Cashwell, E. D., & Everett, C. J. (1957). *A Practical Manual on the Monte Carlo Method for Random Walk Problems*. Los Alamos Scientific Laboratory of the University of California.

Chen, L., Han, B., Wang, X., Zhao, J., Yang, W., & Yang, Z. (2023). Machine Learning Methods in Weather and Climate Applications: A Survey. *Applied Sciences*, *13*(21), 12019. https://doi.org/10.3390/app132112019

Chen, S. S., & Curcic, M. (2016). Ocean surface waves in Hurricane Ike (2008) and Superstorm Sandy (2012): Coupled model predictions and observations. *Ocean Modelling*, *103*, 161–176. https://doi.org/10.1016/j.ocemod.2015.08.005

CIRES. (2014a). Continuously Updated Digital Elevation Model (CUDEM) - 1/3 Arc-Second Resolution Bathymetric-Topographic Tiles. Texas. *NOAA National Centers for Environmental Information*. https://doi.org/10.25921/0mpp-h192

CIRES. (2014b). Continuously Updated Digital Elevation Model (CUDEM) - 1/9 Arc-Second Resolution Bathymetric-Topographic Tiles. Texas. *NOAA National Centers for Environmental Information*. https://doi.org/10.25921/ds9v-ky35

Cummings, J. A., & Smedstad, O. M. (2013). Variational Data Assimilation for the Global Ocean. In S. K. Park & L. Xu (Eds.), *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)* (pp. 303–343). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35088-7_13

Deltares. (2018). Wind Enhance Scheme for cyclone modelling - User Manual. 1–110.

DeMaria, M., Knaff, J. A., Knabb, R., Lauer, C., Sampson, C. R., & DeMaria, R. T. (2009). A New Method for Estimating Tropical Cyclone Wind Speed Probabilities. *Weather and Forecasting*, *24*(6), 1573–1591. https://doi.org/10.1175/2009WAF2222286.1

Eilander, D., Couasnon, A., Leijnse, T., Ikeuchi, H., et al. (2023). A globally applicable framework for compound flood hazard modeling. *Natural Hazards and Earth System Sciences*, *23*(2), 823–846. https://doi.org/10.5194/nhess-23-823-2023

Fox News. (2015, March 25). Texas Governor: Federal Response to Hurricane Ike “Underwhelming” [Text.Article]. Retrieved February 21, 2024, from https://www.foxnews.com/story/texas-governor-federal-response-to-hurricane-ike-underwhelming

Grimley, L. E., Leijnse, T., Ratcliff, J., Sebastian, A., & Luettich, R. (2022). Flooding at the Fringe: A Reduced-physics Model for Assessing Compound Flooding from Pluvial, Fluvial, and Coastal Hazards, *2022*, NH36A-03. Presented at the AGU Fall Meeting Abstracts.

Gu, L., Chen, J., Yin, J., Slater, L. J., et al. (2022). Global Increases in Compound Flood-Hot Extreme Hazards Under Climate Warming. *Geophysical Research Letters*, *49*(8), e2022GL097726. https://doi.org/10.1029/2022GL097726

Harrigan, S., Zsoter, E., Cloke, H., Salamon, P., & Prudhomme, C. (2023). Daily ensemble river discharge reforecasts and real-time forecasts from the operational Global Flood Awareness System. *Hydrology and Earth System Sciences*, *27*(1), 1–19. https://doi.org/10.5194/hess-27-1-2023

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Holland, J., Belanger, I., & Fritz, A. (2010). A revised model for radial profiles of hurricane winds. *Monthly Weather Review*, 4393–4401. https://doi.org/10.1175/2010MWR3317.1

Homer, C., Dewitz, J., Jin, S., Xian, G., et al. (2020). Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 184–199. https://doi.org/10.1016/j.isprsjprs.2020.02.019

Hope, M. E., Westerink, J. J., Kennedy, A. B., Kerr, P. C., et al. (2013). Hindcast and validation of Hurricane Ike (2008) waves, forerunner, and storm surge. *Journal of Geophysical Research: Oceans*, *118*(9), 4424–4460. https://doi.org/10.1002/jgrc.20314

Huang, S.-K., Lindell, M. K., & Prater, C. S. (2016). Who Leaves and Who Stays? A Review and Statistical Meta-Analysis of Hurricane Evacuation Studies. *Environment and Behavior*, *48*(8), 991–1029. https://doi.org/10.1177/0013916515578485

Interagency Performance Evaluation Task Force (IPET). (2006). Performance evaluation of the New Orleans and Southeast Louisiana hurricane protection system draft final report of the Interagency Performance Evaluation Task Force. Retrieved October 23, 2023, from https://lccn.loc.gov/2006618548

Jaafar, H. H., Ahmad, F. A., & El Beyrouthy, N. (2019). GCN250, new global gridded curve numbers for hydrologic modeling and design. *Scientific Data*, *6*(1), 145. https://doi.org/10.1038/s41597-019-0155-x

Kennedy, A., Rogers, S., Sallenger, A., Gravois, U., Zachry, B., Dosa, M., & Zarama, F. (2011). Building Destruction from Waves and Surge on the Bolivar Peninsula during Hurricane Ike. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, *137*(3), 132–141. https://doi.org/10.1061/(ASCE)WW.1943-5460.0000061

Kernkamp, H. W. J., Van Dam, A., Stelling, G. S., & de Goede, E. D. (2011). Efficient scheme for the shallow water equations on unstructured grids with application to the Continental Shelf. *Ocean Dynamics*, *61*(8), 1175–1188. https://doi.org/10.1007/s10236-011-0423-6

Knutson, T. R., McBride, J. L., Chan, J., Emanuel, K., et al. (2010). Tropical cyclones and climate change. *Nature Geoscience*, *3*(3), 157–163. https://doi.org/10.1038/ngeo779

Kumar, V., Azamathulla, H. M., Sharma, K. V., Mehta, D. J., & Maharaj, K. T. (2023). The State of the Art in Deep Learning Applications, Challenges, and Future Prospects: A Comprehensive Review of Flood Forecasting and Management. *Sustainability*, *15*(13), 10543. https://doi.org/10.3390/su151310543

Lai, Y., Li, J., Gu, X., Liu, C., & Chen, Y. D. (2021). Global Compound Floods from Precipitation and Storm Surge: Hazards and the Roles of Cyclones. *Journal of Climate*, *34*, 8319–8339. https://doi.org/10.1175/JCLI-D-21-0050.1

Lamers, A., Devi. S, S., Sharma, M., Berg, R., et al. (2023). Forecasting tropical cyclone rainfall and flooding hazards and impacts. *Tropical Cyclone Research and Review*, *12*(2), 100–112. https://doi.org/10.1016/j.tcrr.2023.06.005

Lecacheux, S., Rohmer, J., Paris, F., Pedreros, R., Quetelard, H., & Bonnardot, F. (2021). Toward the probabilistic forecasting of cyclone-induced marine flooding by overtopping at Reunion Island aided by a time-varying random-forest classification approach. *Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards*, *105*(1), 227–251.

Lee, W., McLaughlin, P. W., & Kaihatu, J. M. (2017). Parameterization of Maximum Significant Wave Heights in Coastal Regions due to Hurricanes. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, *143*(2), 04016016. https://doi.org/10.1061/(ASCE)WW.1943-5460.0000362

Lee, W., Sun, A. Y., Scanlon, B. R., & Dawson, C. (2023). Hindcasting compound pluvial, fluvial and coastal flooding during Hurricane Harvey (2017) using Delft3D-FM. *Natural Hazards*. https://doi.org/10.1007/s11069-023-06247-9

Leijnse, T., van Ormondt, M., Nederhoff, K., & van Dongeren, A. (2021). Modeling compound flooding in coastal systems using a computationally efficient reduced-physics solver: Including fluvial, pluvial, tidal, wind- and wave-driven processes. *Coastal Engineering*, *163*, 103796. https://doi.org/10.1016/j.coastaleng.2020.103796

Leijnse, T., Nederhoff, K., Thomas, J., Parker, K., et al. (2023). Rapid modeling of compound flooding across broad coastal regions and the necessity to include rainfall driven processes: a case study of hurricane florence (2018). In *Coastal Sediments 2023* (pp. 2576–2584). WORLD SCIENTIFIC. https://doi.org/10.1142/9789811275135_0235

Luettich, R. A. (Richard A., Westerink, J. J., & Scheffner, N. W. (1992). *ADCIRC : an advanced three-dimensional circulation model for shelves, coasts, and estuaries. Report 1, Theory and methodology of ADCIRC-2DD1 and ADCIRC-3DL* (Report). *This Digital Resource was created from scans of the Print Resource*. Coastal Engineering Research Center (U.S.). Retrieved from https://erdc-library.erdc.dren.mil/jspui/handle/11681/4618

Maymandi, N., Hummel, M. A., & Zhang, Y. (2022). Compound Coastal, Fluvial, and Pluvial Flooding During Historical Hurricane Events in the Sabine–Neches Estuary, Texas. *Water Resources Research*, *58*(12), e2022WR033144. https://doi.org/10.1029/2022WR033144

Milinski, S., Maher, N., & Olonscheck, D. (2020). How large does a large ensemble need to be? *Earth System Dynamics*, *11*(4), 885–901. https://doi.org/10.5194/esd-11-885-2020

Morss, R. E., & Hayden, M. H. (2010). Storm Surge and “Certain Death”: Interviews with Texas Coastal Residents following Hurricane Ike. *Weather, Climate, and Society*, *2*(3), 174–189. https://doi.org/10.1175/2010WCAS1041.1

National Hurricane Center. (2014, March 18). Tropical Cyclone Report Hurricane Ike (AL092008) 1 - 14 September 2008. National Hurricane Center, United States National Oceanic and Atmospheric Administration’s National Weather Service.

National Hurricane Center. (2018, January 26). Costliest U.S. tropical cyclones tables updated. National Hurricane Center.

Nederhoff, K., Giardino, A., van Ormondt, M., & Vatvani, D. (2019). Estimates of tropical cyclone geometry parameters based on best-track data. *Natural Hazards and Earth System Sciences*, *19*(11), 2359–2370. https://doi.org/10.5194/nhess-19-2359-2019

Nederhoff, K., van Ormondt, M., Veeramony, J., van Dongeren, A., Antolínez, J., Leijnse, T., & Roelvink, D. (2023). Accounting for Uncertainties in Forecasting Tropical Cyclone-Induced Compound Flooding. *EGUsphere*, 1–41. https://doi.org/10.5194/egusphere-2023-2341

Nederhoff, K., Leijnse, T., Parker, K., Thomas, J., et al. (2023). Tropical or extratropical cyclones: what drives the compound flood hazard, impact, and risk for the United States Southeast Atlantic coast? Retrieved from https://eartharxiv.org/repository/view/5123/

NOAA. (2006). SLOSH, Sea, Lake, and Overland Surges from Hurricanes, User and Technical Software Documentation. NOAA.

Pakhale, G., Khosa, R., & Gosain, A. K. (2024). In Today’s World, Is It Worth Performing Flood Frequency Analysis Using Observed Streamflow Data? *Environmental Advances*, *15*, 100485. https://doi.org/10.1016/j.envadv.2024.100485

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10.1038/nature20584

Rego, J., & Li, C. (2010). Storm surge propagation in Galveston Bay during Hurricane Ike. *Journal of Marine Systems*, *82*, 265–279. https://doi.org/10.1016/j.jmarsys.2010.06.001

Rezuanul Islam, Md., Duc, L., & Sawada, Y. (2023). Assessing Storm Surge Multiscenarios Based on Ensemble Tropical Cyclone Forecasting. *Journal of Geophysical Research: Atmospheres*, *128*(23), e2023JD038903. https://doi.org/10.1029/2023JD038903

Röbke, B. R., Leijnse, T., Winter, G., van Ormondt, M., van Nieuwkoop, J., & de Graaff, R. (2021). Rapid Assessment of Tsunami Offshore Propagation and Inundation with D-FLOW Flexible Mesh and SFINCS for the 2011 Tōhoku Tsunami in Japan. *Journal of Marine Science and Engineering*, *9*(5), 453. https://doi.org/10.3390/jmse9050453

Stearns, M., & Padgett, J. E. (2012). Impact of 2008 Hurricane Ike on Bridge Infrastructure in the Houston/Galveston Region. *Journal of Performance of Constructed Facilities*, *26*(4), 441–452. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000213

Sun, A. Y., Li, Z., Lee, W., Huang, Q., Scanlon, B. R., & Dawson, C. (2023). Rapid Flood Inundation Forecast Using Fourier Neural Operator (Version 1). https://doi.org/10.48550/ARXIV.2307.16090

Tebaldi, C., Dorheim, K., Wehner, M., & Leung, R. (2021). *Extreme Metrics and Large Ensembles* (preprint). Earth system change: climate prediction. https://doi.org/10.5194/esd-2021-53

Titley, H. A., Cloke, H. L., Stephens, E. M., Pappenberger, F., & Zsoter, E. (2024). Using Ensembles to Analyze Predictability Links in the Tropical Cyclone Flood Forecast Chain. *Journal of Hydrometeorology*, *25*(1), 191–206. https://doi.org/10.1175/JHM-D-23-0022.1

Titley, Helen A., Bowyer, R. L., & Cloke, H. L. (2020). A global evaluation of multi-model ensemble tropical cyclone track probability forecasts. *Quarterly Journal of the Royal Meteorological Society*, *146*(726), 531–545. https://doi.org/10.1002/qj.3712

Torn, R. D., & DeMaria, M. (2021). Validation of Ensemble-Based Probabilistic Tropical Cyclone Intensity Change. *Atmosphere*, *12*(3), 373. https://doi.org/10.3390/atmos12030373

Toth, Z., & Buizza, R. (2019). Chapter 2 - Weather Forecasting: What Sets the Forecast Skill Horizon? In A. W. Robertson & F. Vitart (Eds.), *Sub-Seasonal to Seasonal Prediction* (pp. 17–45). Elsevier. https://doi.org/10.1016/B978-0-12-811714-9.00002-4

Tounsi, A., Temimi, M., Abdelkader, M., & Gourley, J. J. (2023). Assessment of deterministic and probabilistic precipitation nowcasting techniques over New York metropolitan area. *Environmental Modelling & Software*, *168*, 105803. https://doi.org/10.1016/j.envsoft.2023.105803

Trabing, B. C., Musgrave, K. D., DeMaria, M., Zachry, B. C., Brennan, M. J., & Rappaport, E. N. (2023). The Development and Evaluation of a Tropical Cyclone Probabilistic Landfall Forecast Product. *Weather and Forecasting*, *38*(8), 1363–1374. https://doi.org/10.1175/WAF-D-22-0199.1

Trumbla, R. (2019). The Great Galveston Hurricane of 1900. NOAA Celebrates 200 Years of Science, Service and Stewardship. National Oceanic and Atmospheric Administration.

USGS. (2016). USGS Water Data for the Nation. Retrieved December 7, 2023, from https://waterdata.usgs.gov/nwis

Van Ormondt, M., Nederhoff, K., & van Dongeren, A. (2020). Delft Dashboard: a quick set-up tool for hydrodynamic models. *Journal of Hydroinformatics*, *22*(3), 510–527.

Veeramony, J., Condon, A., & Hebert, D. (2012). Effect of Coupling Wave And Flow Dynamics On Hurricane Surge And Inundation. Presented at the The Twenty-second International Offshore and Polar Engineering Conference, OnePetro. Retrieved from https://dx.doi.org/

Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E. (2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, *5*(12), 1093–1097. https://doi.org/10.1038/nclimate2736

Wang, C.-C., Lin, B.-X., Chen, C.-T., & Lo, S.-H. (2015). Quantifying the Effects of Long-Term Climate Change on Tropical Cyclone Rainfall Using a Cloud-Resolving Model: Examples of Two Landfall Typhoons in Taiwan. *Journal of Climate*, *28*(1), 66–85. https://doi.org/10.1175/JCLI-D-14-00044.1

White, C. J., Carlsen, H., Robertson, A. W., Klein, R. J. T., et al. (2017). Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorological Applications*, *24*(3), 315–325. https://doi.org/10.1002/met.1654

Wright, D. B., Knutson, T. R., & Smith, J. A. (2015). Regional climate model projections of rainfall from U.S. landfalling tropical cyclones. *Climate Dynamics*, *45*(11), 3365–3379. https://doi.org/10.1007/s00382-015-2544-y

Xu, C., Nelson-Mercer, B. T., Bricker, J. D., Davlasheridze, M., Ross, A. D., & Jia, J. (2023). Damage Curves Derived from Hurricane Ike in the West of Galveston Bay Based on Insurance Claims and Hydrodynamic Simulations. *International Journal of Disaster Risk Science*, *14*(6), 932–946. https://doi.org/10.1007/s13753-023-00524-8

Zhang, Y. J., Ye, F., Stanev, E. V., & Grashorn, S. (2016). Seamless cross-scale modeling with SCHISM. *Ocean Modelling*, *102*, 64–81. https://doi.org/10.1016/j.ocemod.2016.05.002

Figures

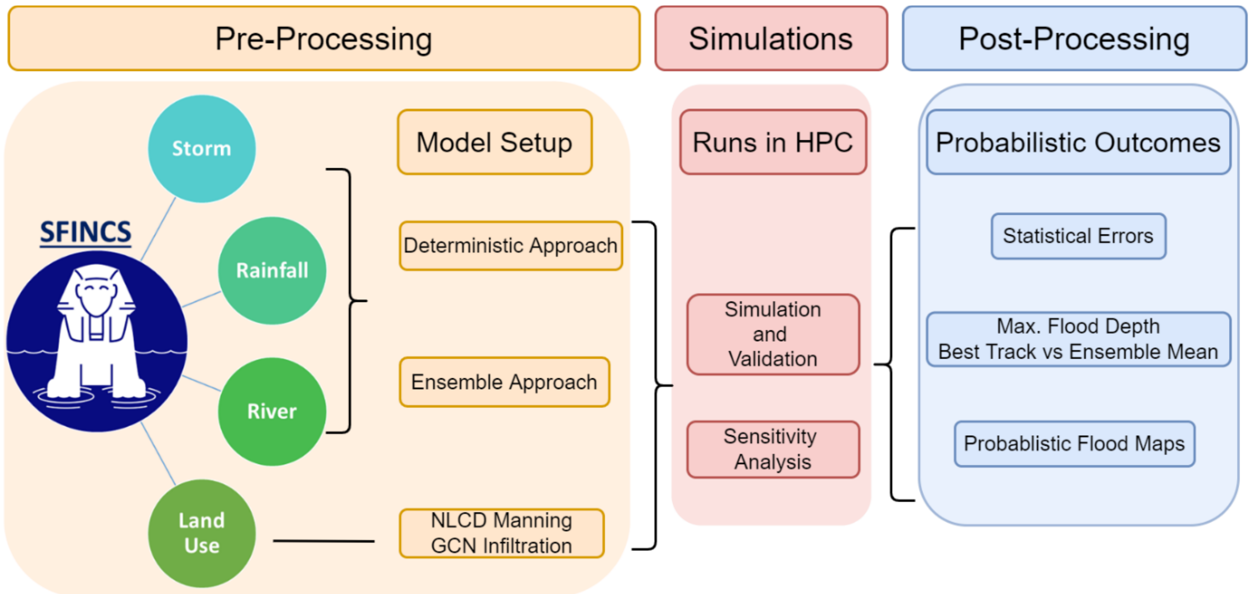

Fig. 1. Probabilistic modeling framework for storm surge and compound flooding at the regional scale. Acronym – Super-Fast INundation of CoastS (SFINCS), National Land Cover Database (NLCD), Global Curve Number (GCN).

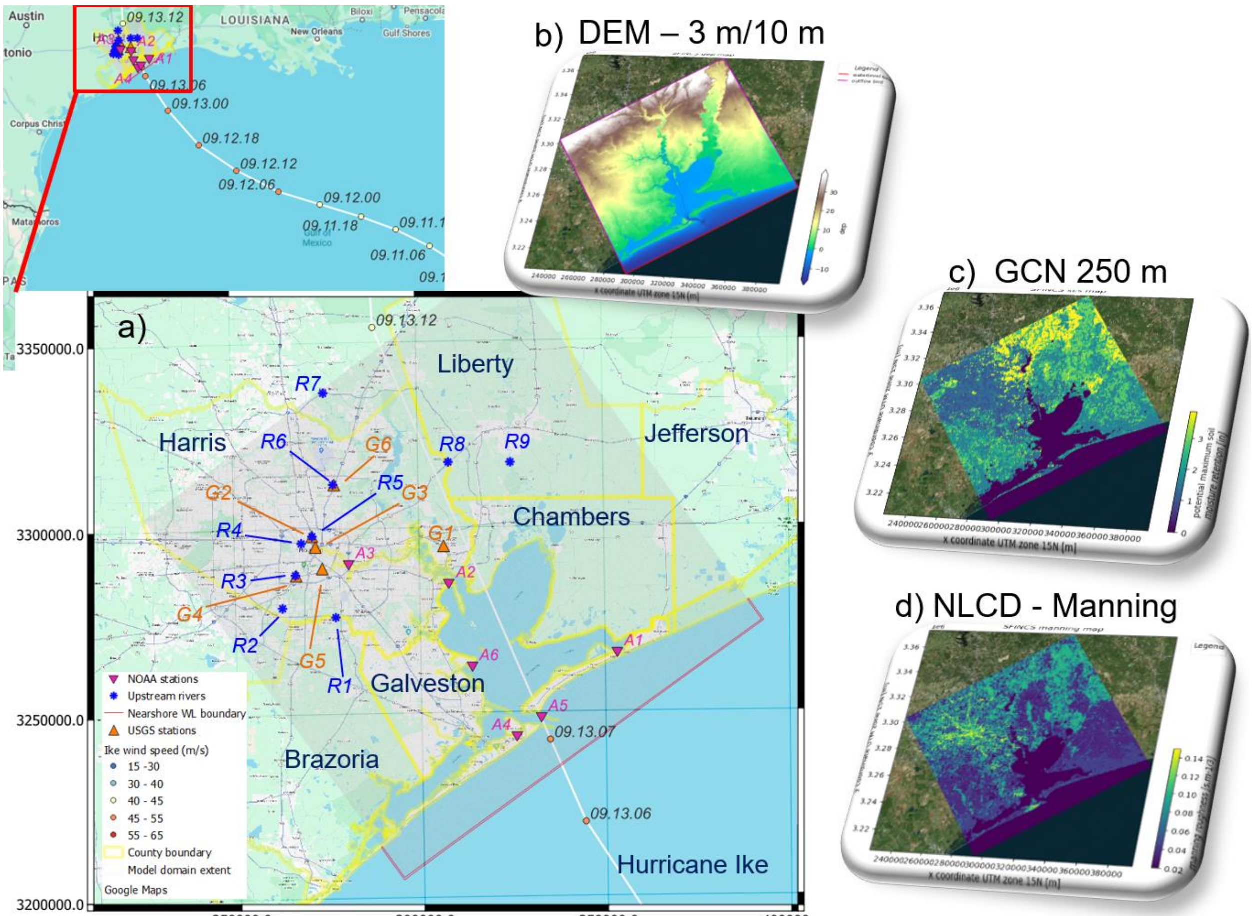

Fig. 2. a) Model extent and boundaries, including the watershed, nine upstream boundaries (R1: Clear Creek, R2: Sims Bayou, R3: Brays Bayou, R4: White Oak Bayou, R5: Little White Oak, R6: Green Bayou, R7: West Fork of San Jacinto River, R8: Cedar Bayou, R9: Trinity River). The best track of Hurricane Ike 2008 is shown with the time (mm.dd.hh). The observation stations used in this study include NOAA tide gages (denoted as A1~A6), and USGS gage-height stations (denoted as G1~G6), b) Maps for digital elevation model (DEM) bathymetry, c) global curve number (GCN), and d) National Land Cover Database (NLCD) Manning bottom frictions.

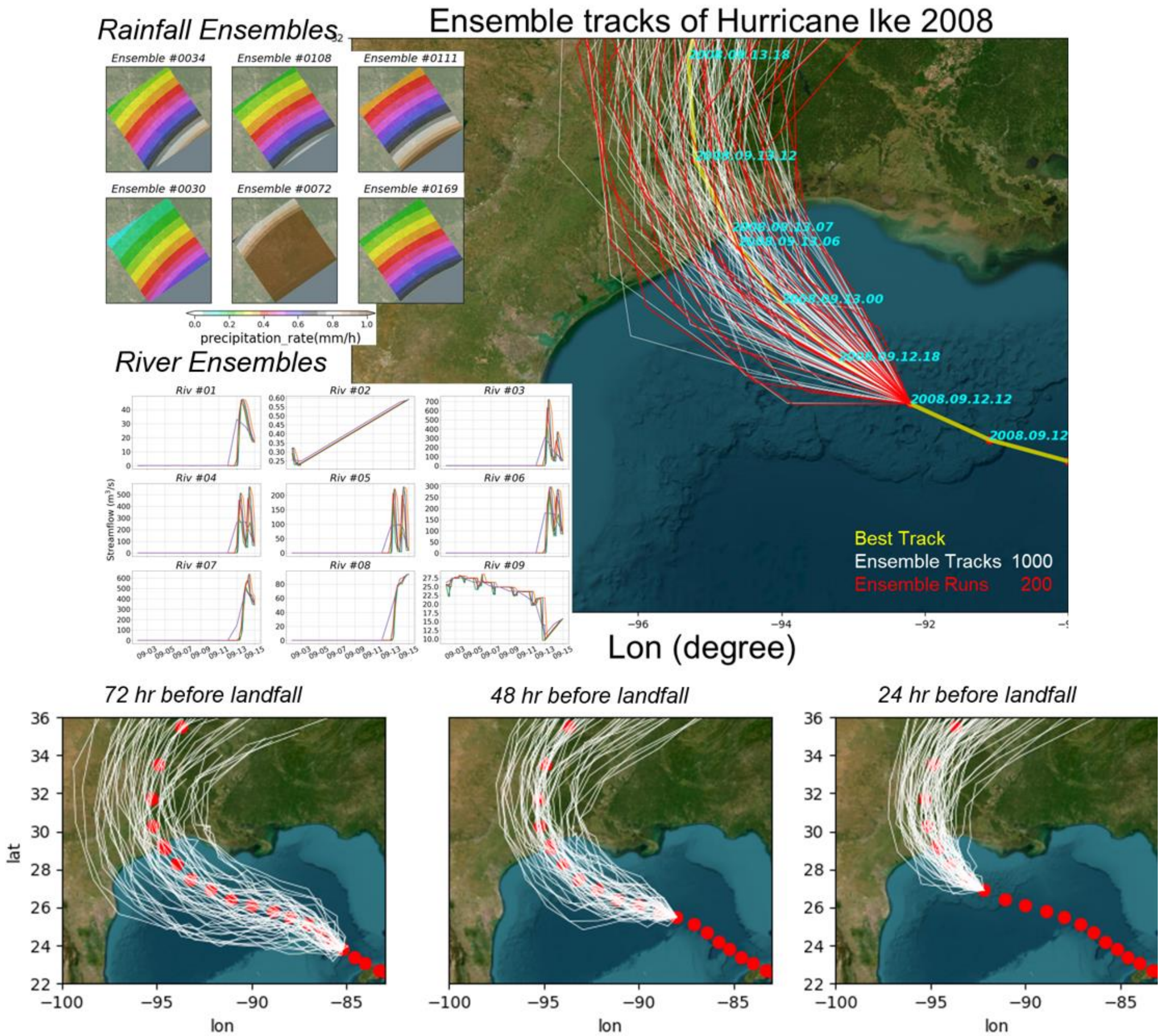

Fig. 3. Generation of ensemble wind-pressure field, rainfall and upstream river discharge. Ensemble wind-generations with different lead-times (1-day, 2-days, 3-days before) of Ike landfall.

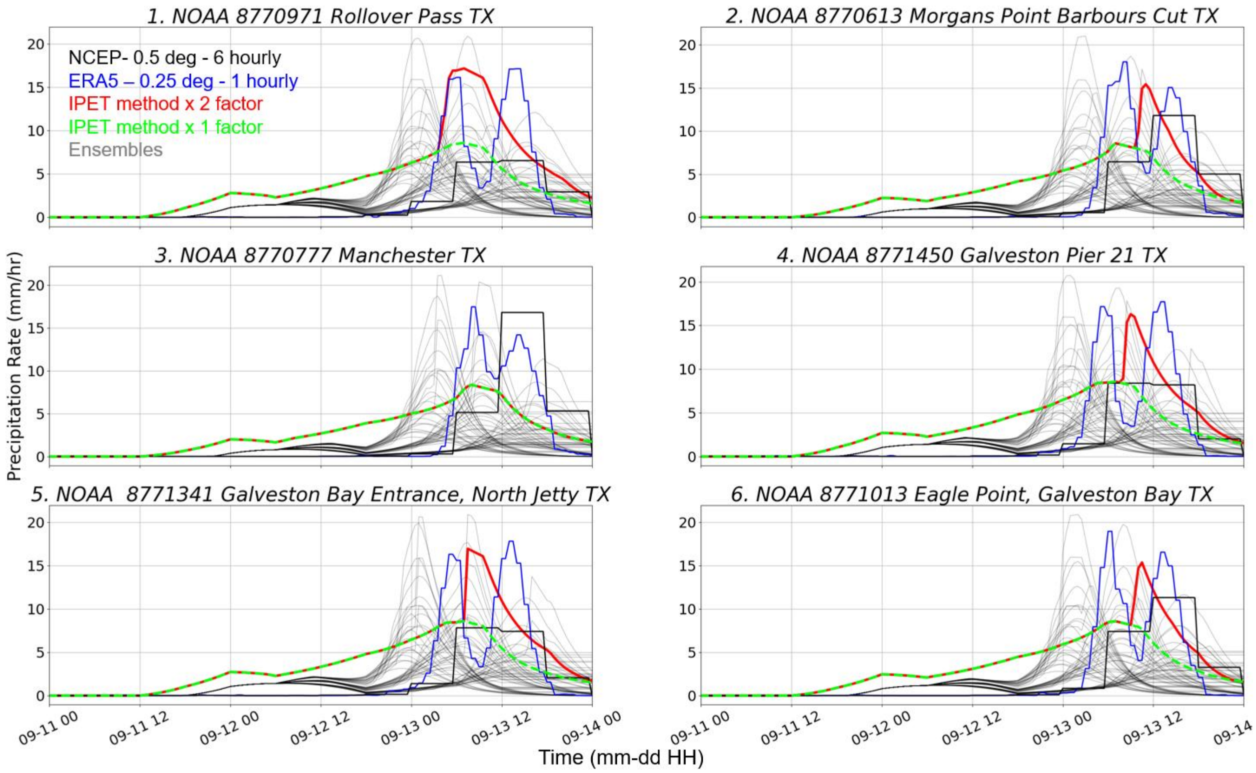

Fig. 4. Comparisons of precipitation rates (mm/hr) between Interagency Performance Evaluation Task Force (IPET) method and other reanalysis data, during Hurricane Ike (09/03/2008-09/15/2008). Please see Fig. 2 for site locations.

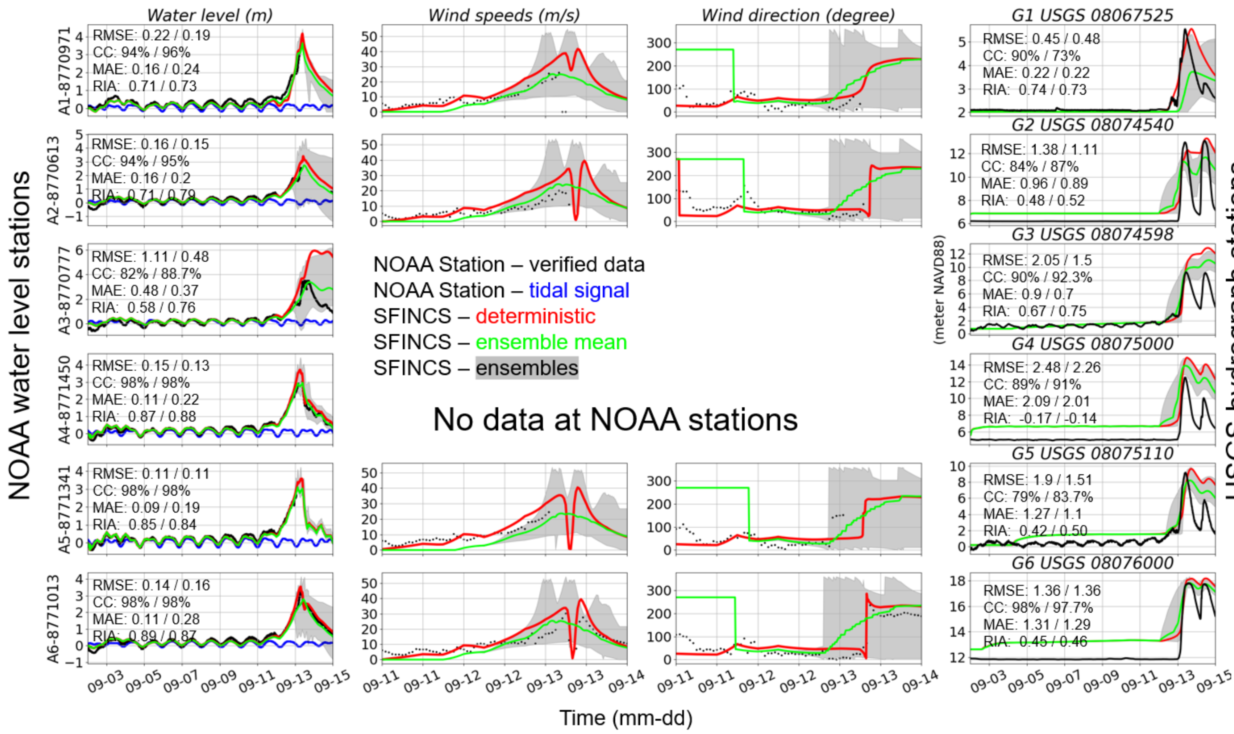

Fig. 5. Comparisons of water levels and winds speeds and directions at six NOAA stations (A1: 8770971 Rollover Pass TX, A2: 8770613 Morgans Point Barbours Cut TX, A3: 8770777 Manchester TX, A4: 8771450 Galveston Pier 21 TX, A5: Galveston Bay Entrance, North Jetty TX, A6: 8771013 Eagle Point, Galveston Bay TX) and hydrographs at six USGS gage-height stations (labeled A1-A6: NOAA, G1-G6: USGS stations in Fig. 2). Error statistics indicate: deterministic- case / ensemble (1,000, 189, 81) average.

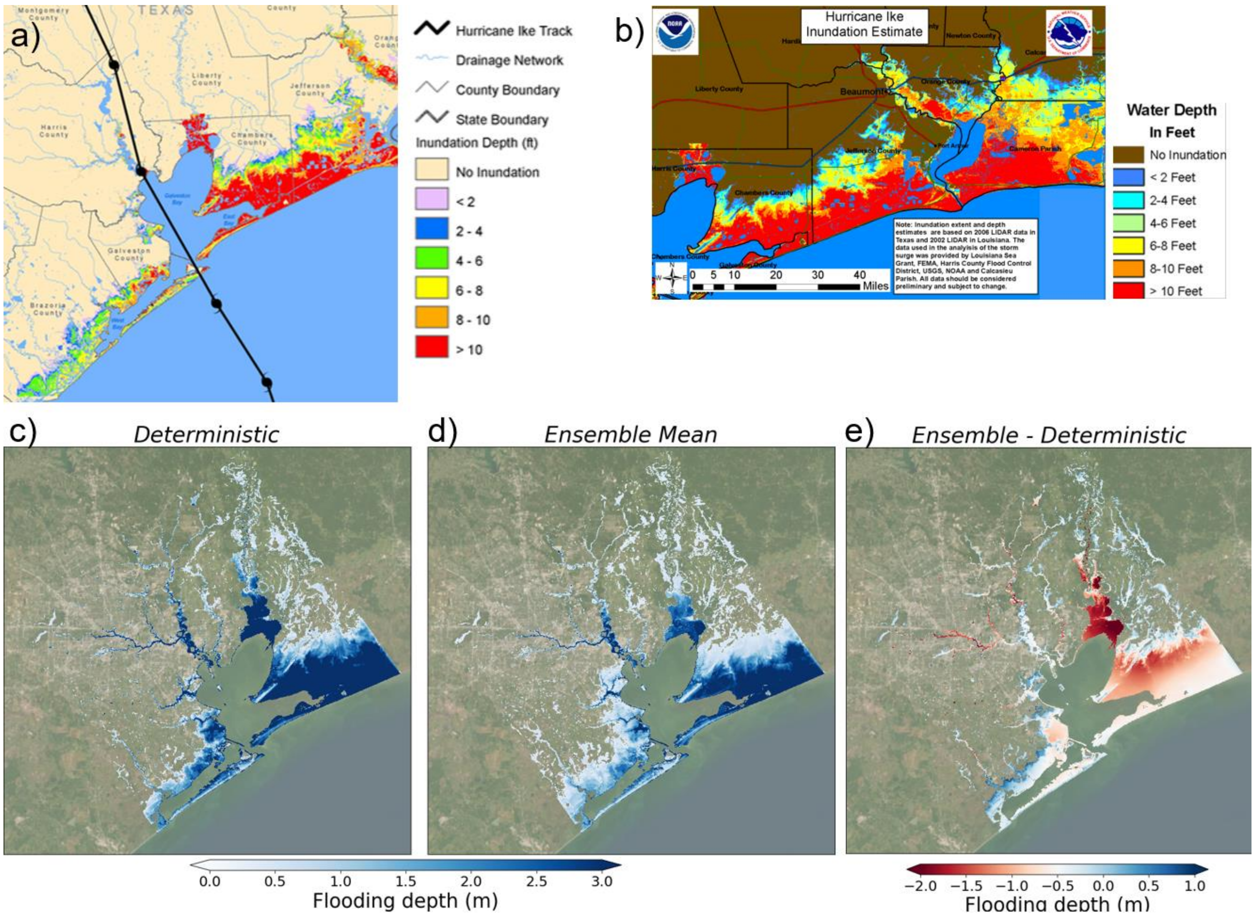

Fig. 6. Comparisons of flood-inundation depth during Hurricane Ike: a) Inundation depth (ft) from Harris County Flood Control District, b) Inundation estimate (ft) from NOAA, c) SFINCS modeled flood-inundation depth (m): deterministic scenario, d) ensemble (mean) scenario, e) Differences between ensemble mean and deterministic scenarios in SFINCS showing lower depths in the ensemble, particularly near the coast and in the Trinity River Basin.

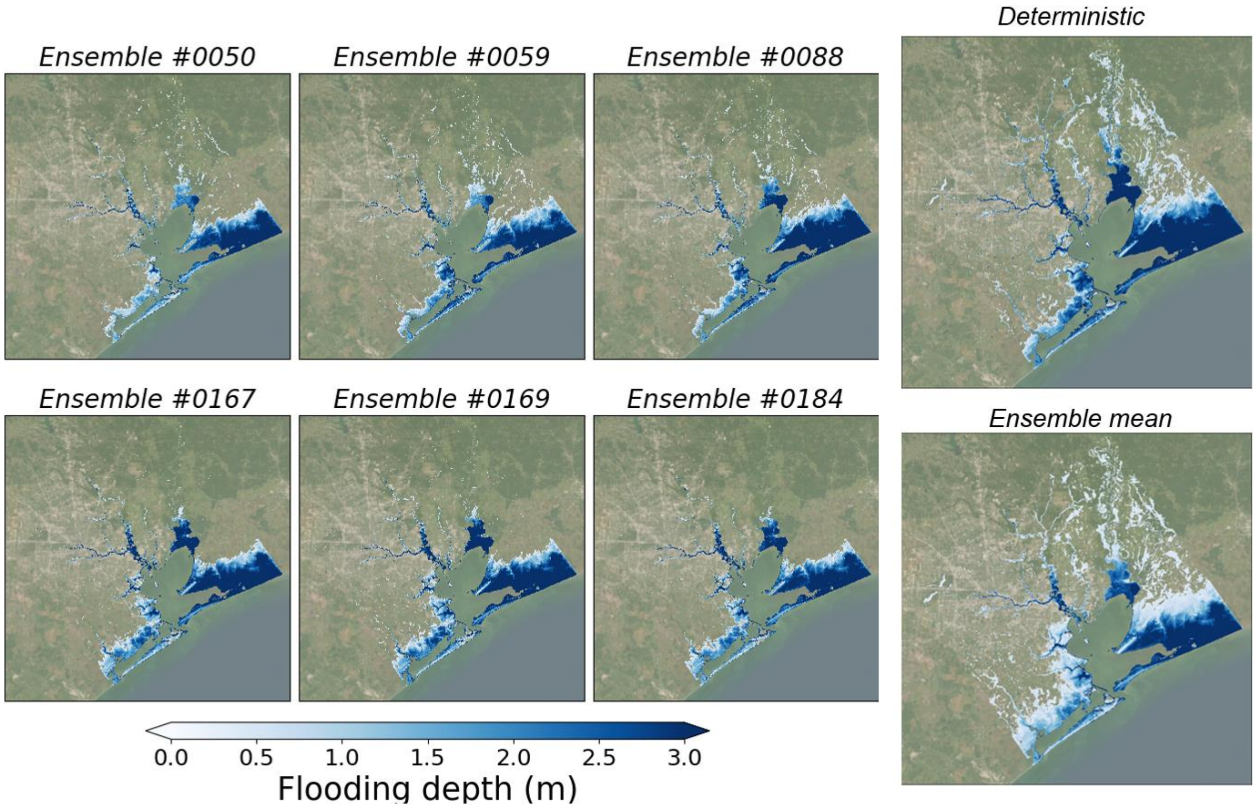

Fig. 7. Comparisons of the maximum flooding depth from ensemble members, ensemble-mean, and deterministic cases.

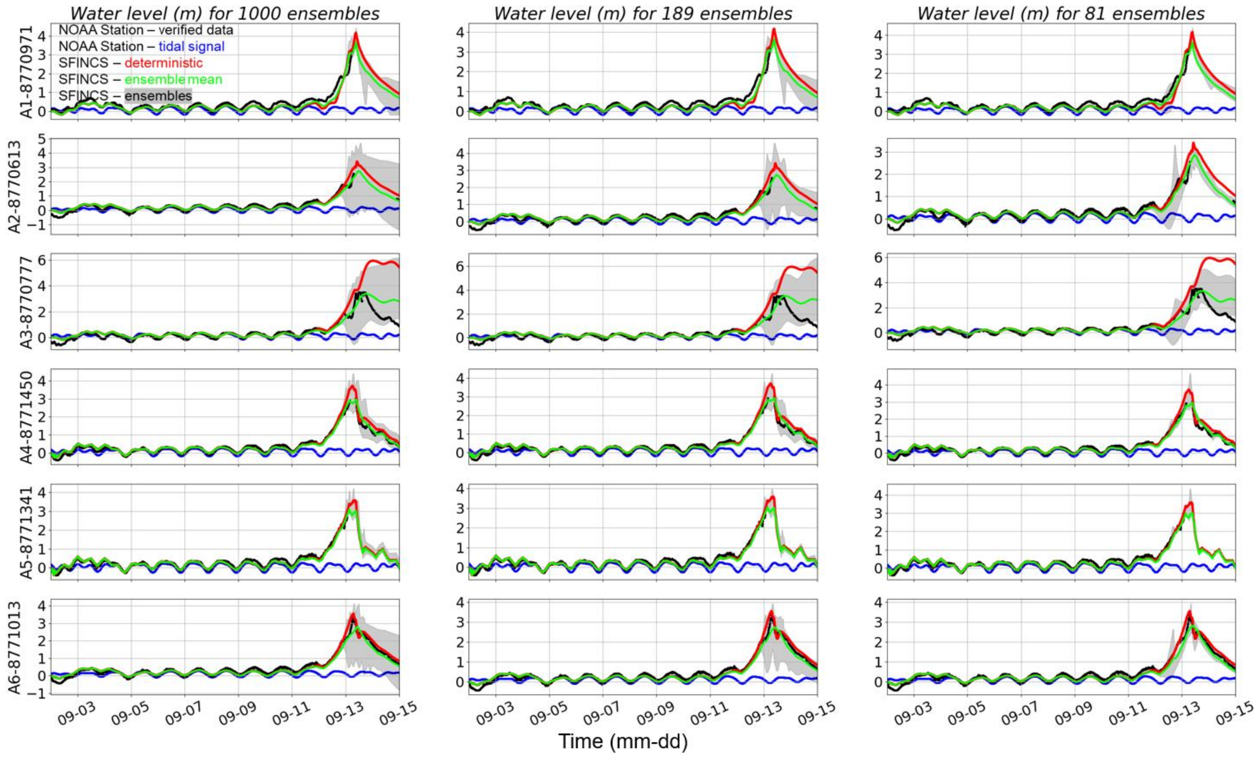

Fig. 8. Comparison of water levels at the NOAA stations (A1 through A6) from different ensemble sizes: NOAA observed data depicted in black, tidal signal in blue, deterministic case in red, ensembles visually filled in pink, and ensemble mean in green.

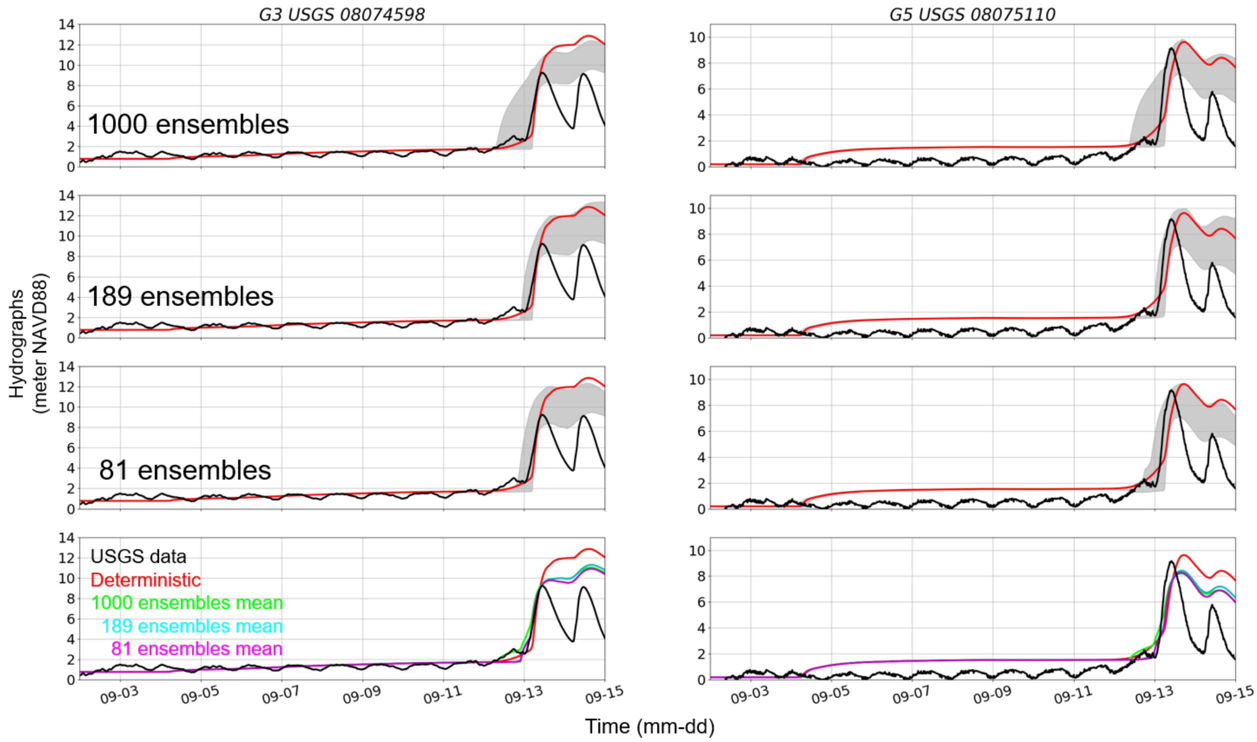

Fig. 9. Comparison of hydrographs at the USGS stations (G3 and G5): USGS data depicted in black, deterministic case in red, ensembles visually filled in pink, and ensemble mean: 1,000, 189 and 81 presented in green, cyan and magenta, respectively.

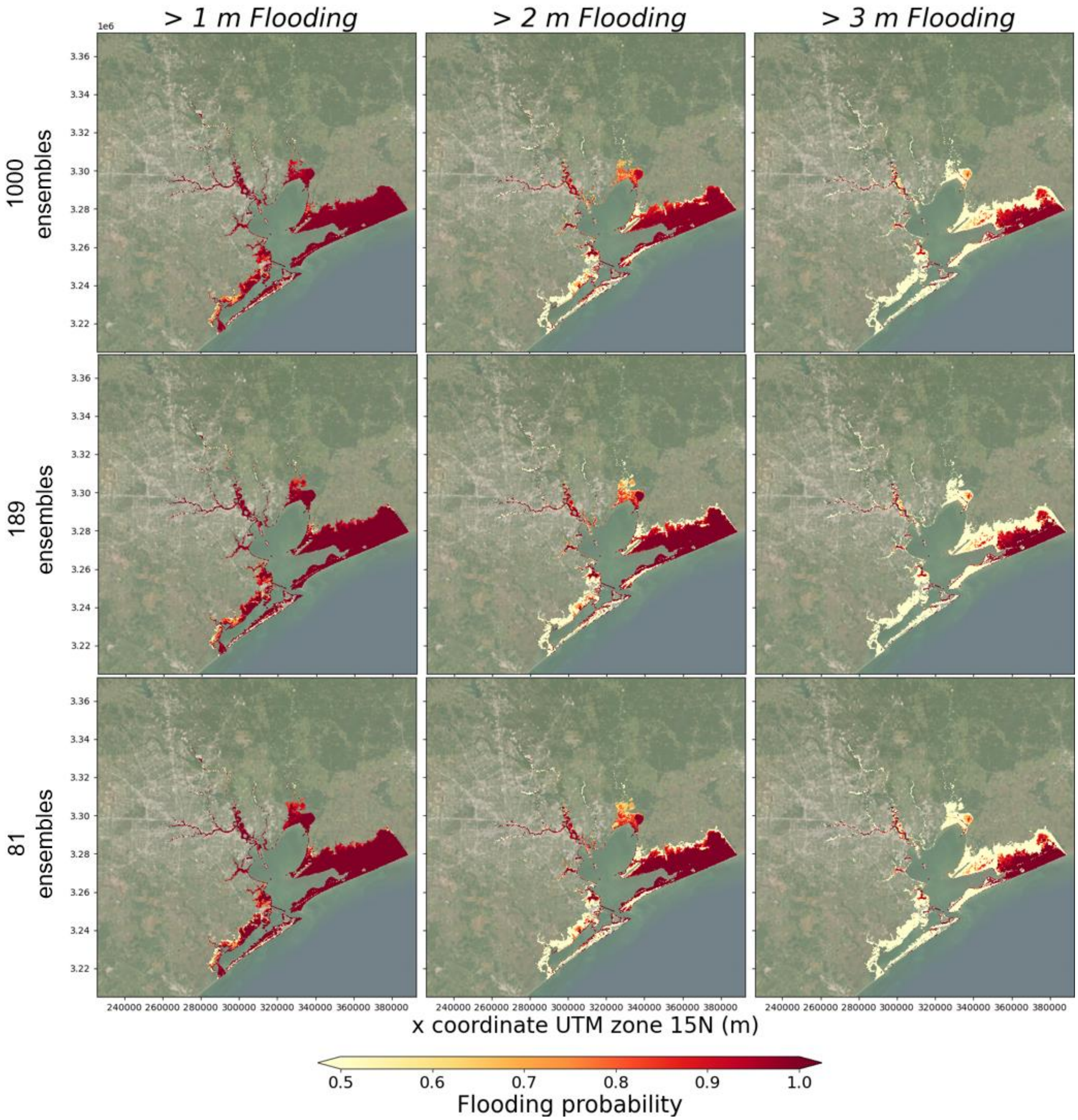

Figure 10. A comprehensive comparison of flooding probability with different ensemble sizes: > 1 m, >2 m, and > 3 m flooding during Hurricane Ike.

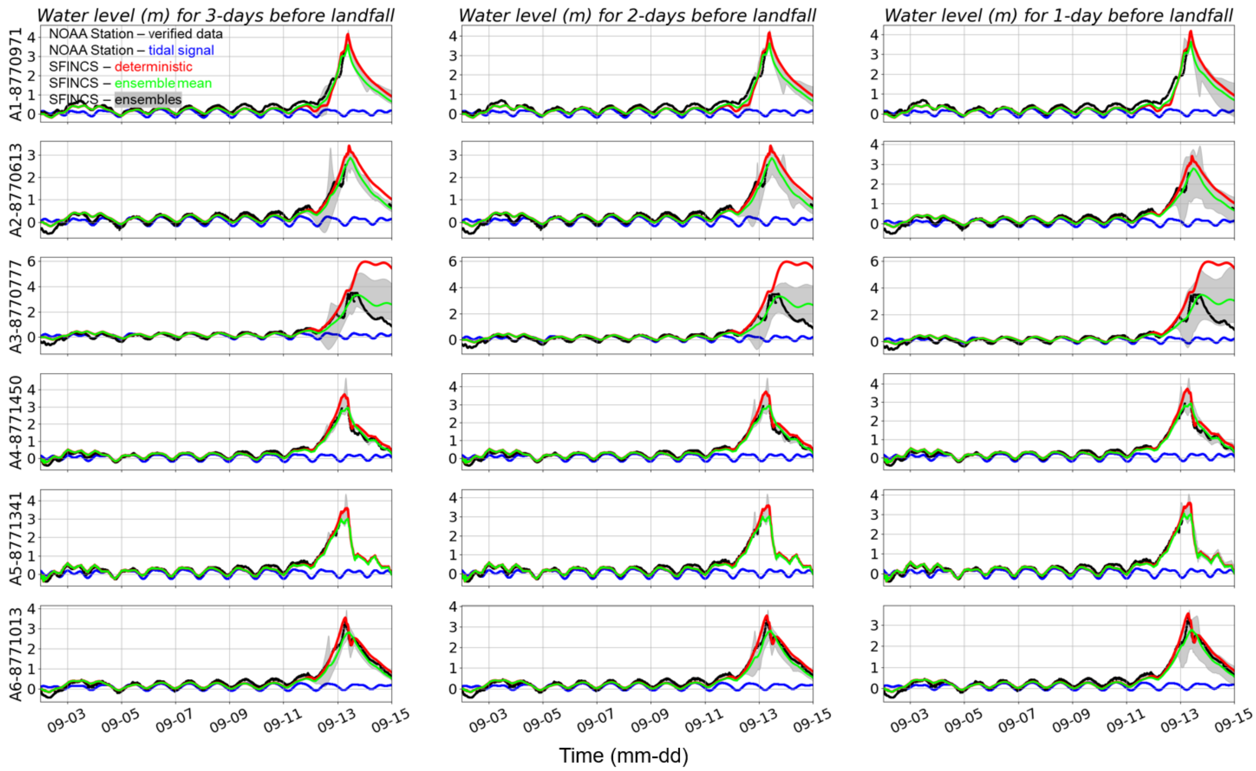

Fig. 11. Comparison of water levels at the NOAA stations (A1 through A6) with different lead time of Ike landfall: NOAA data depicted in black, tidal signal in blue, deterministic case in red, ensembles visually filled in pink, and ensemble mean presented in green.

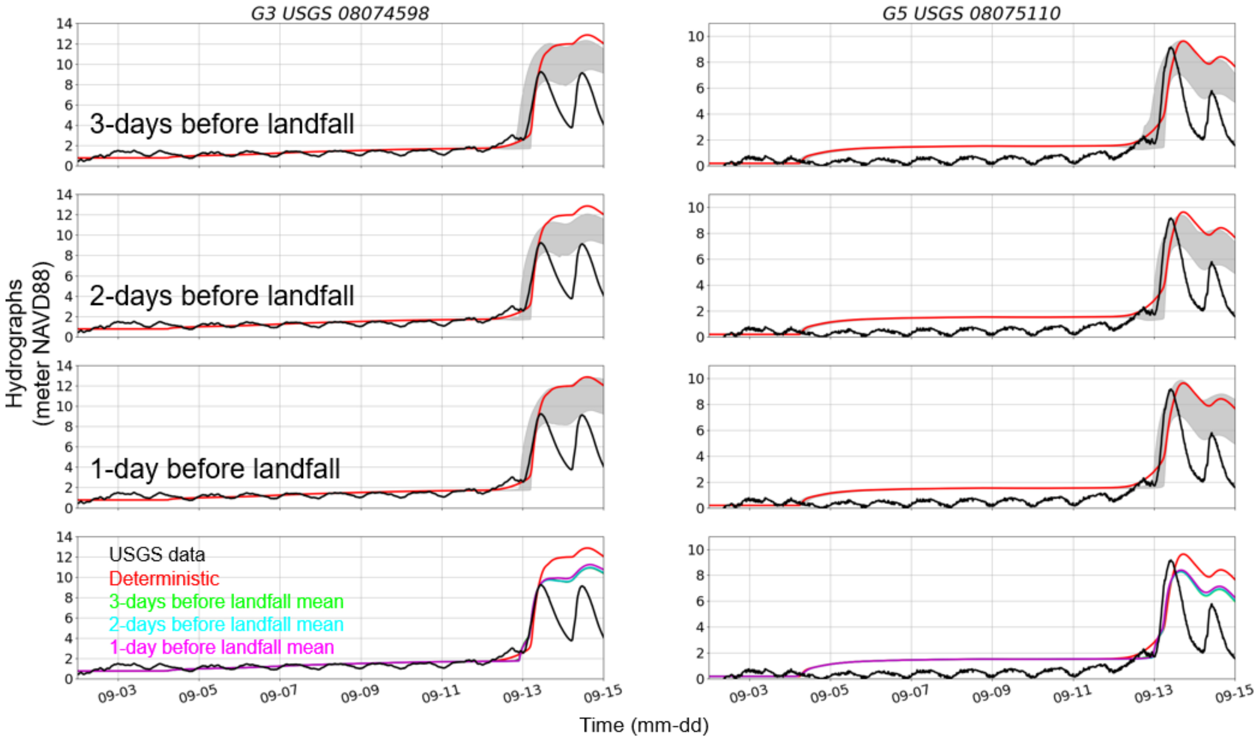

Fig. 12. Comparison of hydrographs at the USGS stations (G3 and G5): USGS data depicted in black, deterministic case in red, ensembles visually filled in pink, and ensemble mean: 3-days, 2-days and 1-day before Ike landfall presented in green, cyan and magenta, respectively.

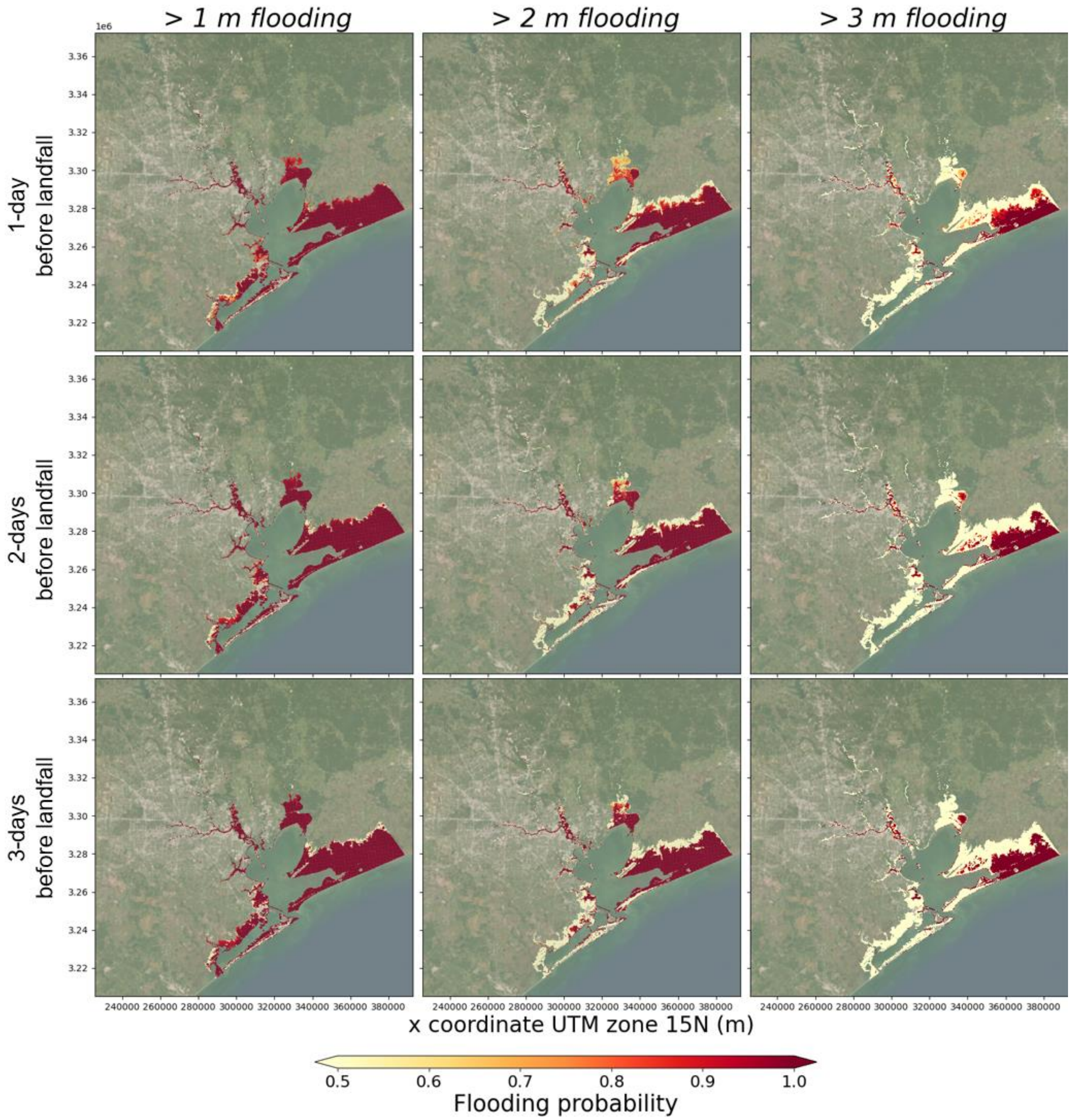

Fig. 13. Comparison of flooding probability with different lead times: > 1 m, 2 m, and 3 m flooding during Hurricane Ike.

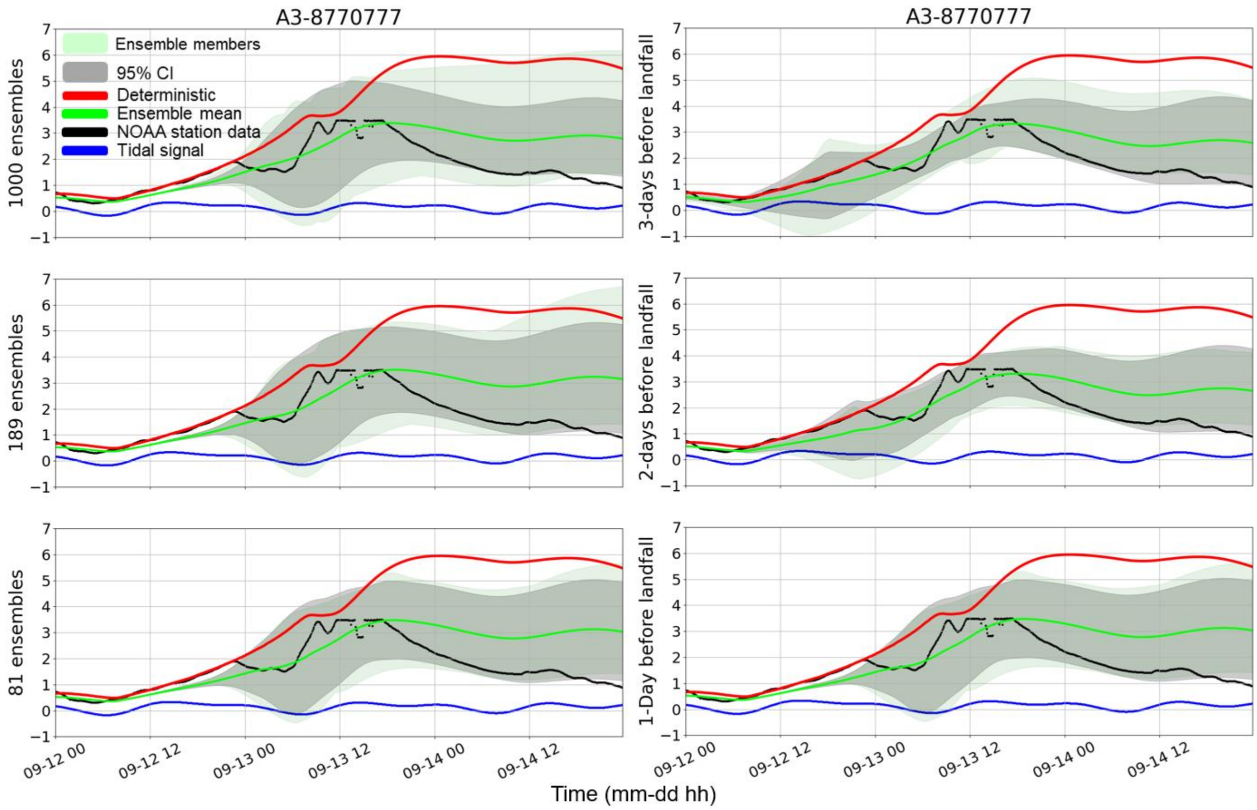

Fig. 14. SFINCS water level results at station A3: 1,000, 189, and 81 ensembles; 3-days, 2-days, and 1-day before landfall. Deterministic case in red; ensemble mean in green; tidal signal in blue; NOAA station data in black; the mint-green band is the ensemble members; the pink band is the 95% CI.

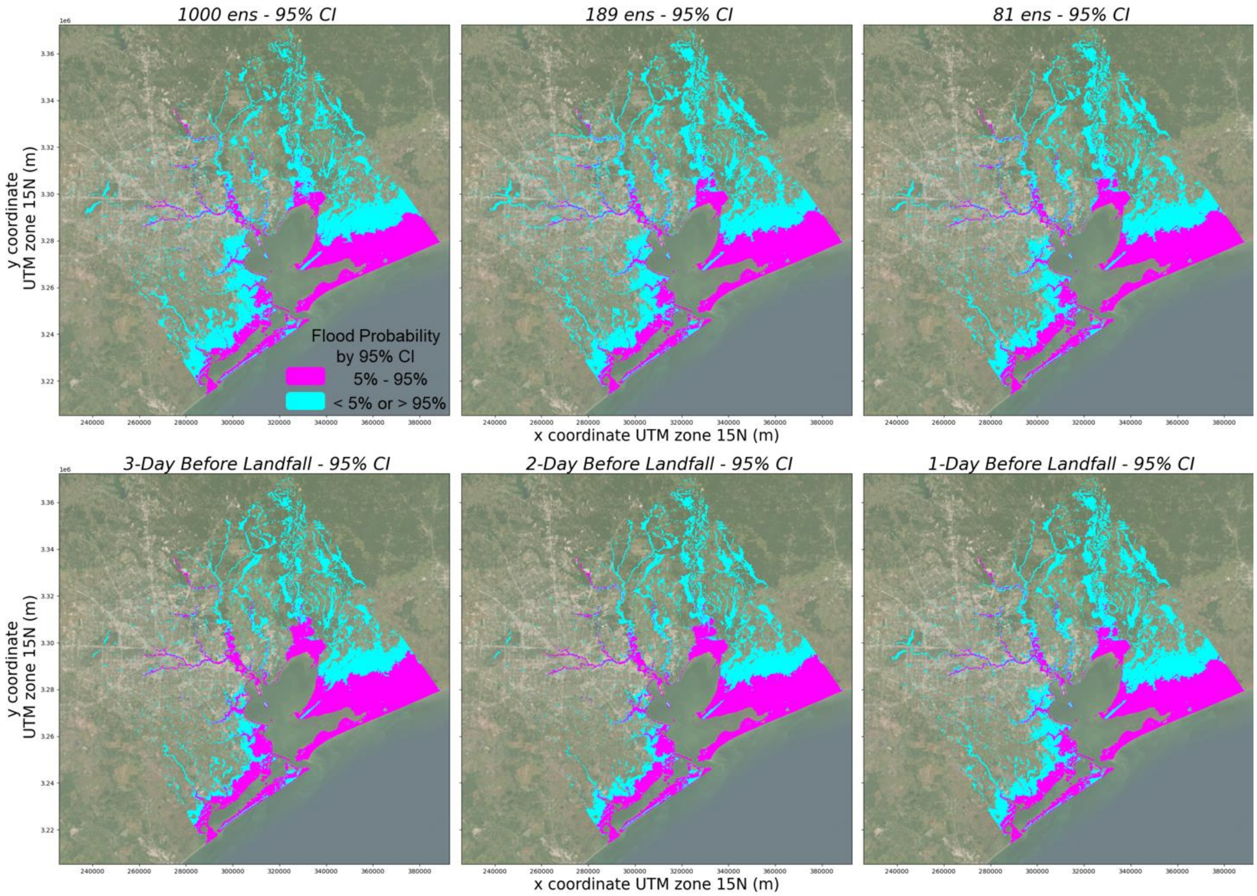

Fig. 15. Maps of uncertainty bandwidth within 95% confidence interval for Hurricane Ike: across different ensemble sizes (1,000, 189, 81) and different lead times (3-days, 2-days, 1-day) before landfall.

Tables

Table 1. Summary of the combination of ensembles in winds, pressure, precipitation and river discharges. Vmax is the maximum wind speed (intensity).

Cases	Total Ensemble	Number of Cross Track Error (CTE)	Number of Along Track Error (ATE)	Number of Vmax	River Streamflow
Case 1	1,000	8	5	5	5
Case 2	189	7	3	3	3
Case 3	81	3	3	3	3
Total	1,270	1-day + (2-days, 3-days before landfall) using Case 3 = 1,270 + 2× 81 = 1,432			

Table 2. Summary of skill scores in different ensemble members (RMSE, correlation coefficient, mean absolute error, refined index of agreement) for water levels from NOAA stations, hydrographs from USGS stations, including station labels.

		NOAA Stations - Water Level (m)						USGS Stations – Hydrograph (m)					
		A1	A2	A3	A4	A5	A6	G1	G2	G3	G4	G5	G6
Deter-ministic	RMSE	0.22	0.16	1.11	0.15	0.11	0.14	0.45	1.38	2.05	2.48	1.90	1.36
	CC	0.94	0.94	0.82	0.98	0.98	0.98	0.90	0.84	0.90	0.89	0.79	0.98
	MAE	0.16	0.12	0.48	0.11	0.09	0.11	0.22	0.96	0.90	2.09	1.27	1.31
	RIA	0.71	0.79	0.58	0.87	0.85	0.89	0.74	0.48	0.67	-0.17	0.42	0.45
Mean Ens-81	RMSE	0.17	0.15	0.4	0.14	0.11	0.17	0.6	1.05	1.41	2.2	1.46	1.34
	CC	0.96	0.94	0.91	0.97	0.98	0.98	0.78	0.89	0.93	0.92	0.83	0.98
	MAE	0.14	0.12	0.24	0.1	0.09	0.13	0.25	0.87	0.67	1.97	1.08	1.27
	RIA	0.75	0.79	0.78	0.87	0.84	0.87	0.7	0.54	0.76	-0.12	0.51	0.47
Mean Ens-189	RMSE	0.19	0.15	0.51	0.13	0.11	0.16	0.47	1.10	1.52	2.25	1.53	1.34
	CC	0.96	0.95	0.88	0.98	0.98	0.98	0.74	0.87	0.92	0.92	0.83	0.98
	MAE	0.15	0.12	0.28	0.09	0.09	0.13	0.22	0.88	0.71	2.00	1.11	1.28
	RIA	0.73	0.79	0.75	0.88	0.84	0.87	0.73	0.53	0.75	-0.13	0.50	0.47
Mean Ens-1000	RMSE	0.19	0.15	0.45	0.13	0.11	0.16	0.47	1.15	1.49	2.28	1.47	1.40
	CC	0.96	0.95	0.90	0.98	0.98	0.98	0.74	0.86	0.93	0.90	0.85	0.97
	MAE	0.14	0.12	0.26	0.09	0.09	0.13	0.22	0.92	0.69	2.03	1.09	1.32
	RIA	0.74	0.79	0.77	0.88	0.84	0.87	0.74	0.51	0.75	-0.15	0.51	0.45

Table 3. Summary of skill scores in different lead time of landfall (root mean square error (RMSE), correlation coefficient (CC), mean absolute error (MAE), refined index of agreement (RIA)) for water level from NOAA stations, hydrographs from USGS stations, including station labels.

Ensemble 81		NOAA Stations - Water Level (m)						USGS Stations - Hydrograph					
		A1	A2	A3	A4	A5	A6	G1	G2	G3	G4	G5	G6
3-days before landfall	RMSE	0.19	0.15	0.49	0.13	0.11	0.16	0.49	1.09	1.50	2.24	1.52	1.34
	CC	0.96	0.95	0.88	0.98	0.98	0.98	0.71	0.88	0.92	0.91	0.83	0.98
	MAE	0.15	0.12	0.27	0.09	0.09	0.12	0.23	0.88	0.70	1.99	1.10	1.27
	RIA	0.73	0.79	0.76	0.88	0.84	0.87	0.73	0.53	0.75	-0.13	0.50	0.47
2-days before landfall	RMSE	0.19	0.15	0.51	0.13	0.11	0.16	0.56	1.05	1.41	2.20	1.47	1.33
	CC	0.96	0.94	0.90	0.97	0.98	0.98	0.75	0.89	0.93	0.92	0.83	0.98
	MAE	0.14	0.12	0.25	0.10	0.09	0.13	0.23	0.87	0.67	1.97	1.08	1.27
	RIA	0.74	0.79	0.78	0.88	0.84	0.87	0.73	0.54	0.76	-0.12	0.51	0.47
1-day before landfall	RMSE	0.17	0.15	0.4	0.14	0.11	0.17	0.6	1.05	1.41	2.2	1.46	1.34
	CC	0.96	0.94	0.91	0.97	0.98	0.98	0.78	0.89	0.93	0.92	0.83	0.98
	MAE	0.14	0.12	0.24	0.1	0.09	0.13	0.25	0.87	0.67	1.97	1.08	1.27
	RIA	0.75	0.79	0.78	0.87	0.84	0.87	0.7	0.54	0.76	-0.12	0.51	0.47