

The Practical Predictability of Storm Tide from Tropical Cyclones in the Gulf of Mexico

KATHRYN R. FOSSELL, DAVID AHJEVYCH, REBECCA E. MORSS, CHRIS SNYDER, AND CHRIS DAVIS

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 2 March 2017, in final form 17 October 2017)

ABSTRACT

The potential for storm surge to cause extensive property damage and loss of life has increased urgency to more accurately predict coastal flooding associated with landfalling tropical cyclones. This work investigates the sensitivity of coastal inundation from storm tide (surge + tide) to four hurricane parameters—track, intensity, size, and translation speed—and the sensitivity of inundation forecasts to errors in forecasts of those parameters. An ensemble of storm tide simulations is generated for three storms in the Gulf of Mexico, by driving a storm surge model with best track data and systematically generated perturbations of storm parameters from the best track. The spread of the storm perturbations is compared to average errors in recent operational hurricane forecasts, allowing sensitivity results to be interpreted in terms of practical predictability of coastal inundation at different lead times. Two types of inundation metrics are evaluated: point-based statistics and spatially integrated volumes. The practical predictability of surge inundation is found to be limited foremost by current errors in hurricane track forecasts, followed by intensity errors, then speed errors. Errors in storm size can also play an important role in limiting surge predictability at short lead times, due to observational uncertainty. Results show that given current mean errors in hurricane forecasts, location-specific surge inundation is predictable for as little as 12–24 h prior to landfall, less for small-sized storms. The results also indicate potential for increased surge predictability beyond 24 h for large storms by considering a storm-following, volume-integrated metric of inundation.

1. Introduction

Coastal flooding caused by storm surge¹ is one of the most dangerous hurricane hazards, resulting in billions of dollars of damage and loss of life. As such, the U.S. National Weather Service and research community have expended significant effort in recent years to improve the prediction and communication of storm surge risks associated with landfalling tropical cyclones and other coastal storms (Morrow et al. 2015). The National Hurricane Center (NHC) has recently developed new storm surge

watch, warning, and probabilistic inundation products, but these storm-specific surge forecast products are not currently issued more than 48 h in advance of anticipated landfall because of the limited predictability of the hurricane's track, and thus of storm surge at different coastal locations (NHC 2017). An indication of the uncertainty in these predictions is the recent development of ensemble prediction systems for surge both in research (Di Liberto et al. 2011; Colle et al. 2015; Georgas et al. 2016) and in operations [the creation of the NHC's new probabilistic surge (P-Surge) products; Taylor and Glahn 2008].

Without adequate advanced warning of this type of hazard, it is difficult for public officials and members of the public to plan and implement evacuations and to take the other preparedness actions necessary to protect life and property when a hurricane approaches the coast. From an emergency management perspective, not only are the timing and location of potential storm surge important, but also the peak magnitude of inundation, likelihood of occurrence, length of coastline affected, and inland extent of inundation. These concerns combined with the difficulty of predicting the storm itself and resulting surge at particular locations, motivate this study,

¹ In this work, we focus primarily on the predictability of storm-induced inundation over normally dry land including the combined impact of storm surge and tide, which is scientifically referred to as *storm tide*. Following NOAA (2013), we refer to this as *coastal inundation* or simply *inundation*. When referring to surge-related predictability more generally, we conform to the NHC's use of the term *storm surge* in their storm surge watches, warnings, and potential inundation maps, where the impact of tides is typically included along with storm-induced inundation.

Corresponding author: Kathryn Fossell, fossell@ucar.edu

which aims to build understanding about the practical predictability of storm surge from landfalling tropical cyclones and the contributions of different factors to surge predictability.

While storm surge is sensitive to the bathymetry, topography, and other characteristics of the landfall region, it is also highly sensitive to attributes of the storm such as size, intensity, speed, and track. Thus, the accuracy of storm surge predictions is closely related to accuracy in predictions of the storm itself (e.g., [Zhong et al. 2010](#)). For this reason, in this study we focus primarily (but not exclusively) on the contributions of forecast errors in different storm parameters to limited surge predictability.

Several previous studies have looked at the sensitivity of storm surge (both with and without tides) to storm parameters such as track, speed, size, and intensity ([Weisberg and Zheng 2006a](#); [Irish et al. 2008](#); [Rego and Li 2009](#); [Zhong et al. 2010](#); [Sebastian et al. 2014](#); [Favre et al. 2015](#)). These studies, however, have generally focused on one storm or on specific coastal areas with similar bathymetric qualities. They also measure sensitivity using peak inundation at select observation points. Similar to these examples, NOAA's current efforts and much of the body of research on storm surge prediction focuses on simulating or predicting the water level at specific locations. [Rego and Li \(2009\)](#) evaluated the sensitivity to storm parameters using two different measures, maximum at a point and volume, and found different results depending on the metric used. This last study highlights the importance of investigating different storm surge metrics.

Building on this previous work, we investigate the practical predictability of storm surge from tropical cyclones across a range of lead times, from hours to days before landfall. Consistent with this goal, rather than focusing on accurately simulating the details of inundation, we use an idealized experimental framework designed to isolate the individual contributions of changes in four storm parameters (landfall location, intensity, size, and translation speed) to simulated surge-induced coastal inundation. We start with a parameterized version of a storm, perturb each storm parameter separately, and then generate an ensemble of storm tide simulations for the perturbed storms. Although we use ensembles, we do so to represent the uncertainty in the various storm parameters and evaluate the relative sensitivities, not to produce probabilistic forecasts of surge from the storms themselves. By recasting the variation of storm parameters into characteristic errors at different forecast lead times, we are able to quantify the contribution of forecast errors in each parameter to practical predictability.

We examine the surge modeling results in terms of inundation (including tide), which is the coastal flooding

that areas at risk may experience. Sensitivity of inundation is evaluated using two types of metrics: point-based and integrated volume. We evaluate predictability using different metrics in order to explore whether alternative methods for measuring the skill of inundation predictions can extend surge predictability to longer lead times, when location-specific inundation from surge cannot be accurately predicted.

We perform experiments with idealized versions of two historical storms, Hurricane Ike (2008) and Hurricane Charley (2004), and a third hypothetical storm ("Charlike") that has the size and intensity of Charley but the track of Ike.² These storms were chosen to represent a range of reasonably plausible storm surge scenarios in which to conduct our experiments, in two different areas of the Gulf of Mexico. Our aim is not to simulate or understand the surge from either historical storm in detail, but rather to use a systematically perturbed set of experiments with these three storms to develop more general conclusions about surge predictability given present skill in tropical cyclone forecasts.

Hurricane Ike was a medium-sized storm [radius of maximum winds (RMW) ~ 73 km (40 n mi)] that tracked west-northwest through the Gulf of Mexico. Ike made landfall in the United States near Galveston, Texas ([Fig. 1](#)), as a category 2 hurricane on the Saffir–Simpson hurricane wind scale and produced widespread storm surge along the Texas and Louisiana coasts ([Berg 2014](#)). Hurricane Charley was a much smaller storm [RMW < 11 km (6 n mi)] that crossed Cuba and turned northeast, making U.S. landfall as a category 4 storm on the southwestern coast of Florida ([Fig. 1](#)). Compared to Ike, Charley's storm surge covered a much smaller area ([Pasch et al. 2011](#)). The hypothetical storm Charlike was created to help investigate the reasons underlying the differences between the Ike and Charley surge predictability results, especially the importance of bathymetric and coastline variations compared to storm properties for limiting storm surge predictability.

By investigating how uncertainty in atmospheric predictions propagates into uncertainty in coastal inundation predictions, using different metrics of inundation, we aim to improve knowledge about the predictability of the hurricane surge system. This knowledge can help hurricane and surge modelers, forecasters, and forecast users understand what aspects of coastal inundation can

² We also considered simulations of a second hypothetical storm, "Ikechar," with the properties of Ike following the track of Charley. However, because of concerns about properly imposing the large size of Ike over the various track perturbations of Charley, including traversing over Cuba, we decided to not include these simulations in our experiments.

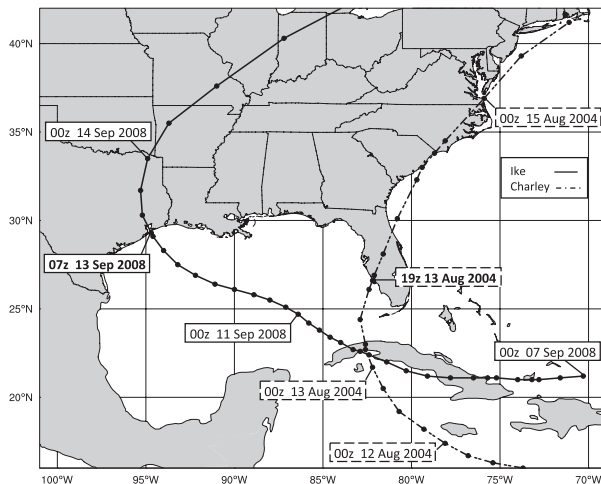


FIG. 1. Storm tracks for Hurricane Ike (solid) and Hurricane Charley (dashed).

meaningfully be predicted at different lead times, given current hurricane forecast errors and potential future hurricane forecast improvements. It also provides information about the current limitations of deterministic and probabilistic surge predictions for different types of tropical storms at different lead times. Our study does not represent many important complicating factors in real-world surge prediction, including surge model error, three-dimensional ocean effects, wave coupling, and river flows. Including these factors would likely further decrease the skill of the perturbations, and so we anticipate that the estimates of practical predictability obtained here serve as an upper bound for real-world practical predictability of surge-induced coastal inundation from hurricanes. More generally, this research contributes to knowledge about coupled system predictability, which is important given the growing emphasis on using coupled modeling and other tools to extend weather predictions into predictions of weather-related hazards and impacts (NRC 2010; NOAA 2011; Jones and Golding 2014; Morss et al. 2017).

The experiment design, model, and data used, and perturbation methodologies are described in section 2. A brief evaluation of the control simulations is presented in section 3. Results from the perturbed simulations, different metrics of surge, and relative importance of different storm parameters are discussed in section 4, followed by a summary and discussion in section 5.

2. Methodology

The overall goal of the methodology is to conduct a series of storm tide predictability experiments that evaluate the changes in coastal inundation induced by

systematically generated storm perturbations. The experimental design uses an idealized “perfect model” framework, in which the sensitivity of coastal inundation to storm forecast errors is evaluated with respect to a model control simulation (section 2a). Perfect-model experiments are an informative starting point for this type of research, since they can advance fundamental knowledge in ways that support interpretation of more complicated results (e.g., Lorenz 1963, 1965, 1982; Houtekamer and Derome 1994; Ehrendorfer 1997; Morss et al. 2001; Walser et al. 2004; Collins et al. 2006). This approach is also advantageous because it allows us to evaluate the ensemble perturbations at every model node, rather than only at the few locations where inundation observations are available from the real storm (section 3).

We use a surge modeling setup designed to represent how changes in hurricane parameters influence macro-scale aspects of storm surge on the forecast time scales of interest, while also remaining computationally feasible (section 2b). In addition, to facilitate systematic perturbation of different storm parameters, we use a parameterized version of real tropical storm wind fields (sections 2c and 2d). These choices mean that our simulations may not accurately represent the details of coastal inundation in specific locations, but they are appropriate given our study’s goals and the idealized, perfect-model approach.

The results are examined using location-based and integrated metrics of inundation (section 2e). The integrated metrics were tested to explore the potential for extending skillful surge predictions to longer lead times, when location-specific inundation forecasts have limited skill. This area-integrated approach is similar to operational weather prediction products that forecast the likelihood of a tornado or heavy rainfall in a region, which can provide useful information at longer lead times when the probability of a hazard occurring at a specific location is extremely low.

a. Experiment design

First, we use best track data from the Automated Tropical Cyclone Forecast system (ATCF; Sampson and Schrader 2000) for each hurricane (Ike, Charley, Charike³) to generate a parameterized version of the storm that is used to force the Advanced Circulation (ADCIRC; Luetlich et al. 1992; Luetlich and Westerink 2004) model, as described in sections 2b and 2c. For the two real storms, the maximum water level (surge + tide) produced by these simulations is evaluated qualitatively

³The hypothetical best track file for Charike is achieved by modifying Ike’s best track to reflect the intensity and size of Charley, and keeping the track and timing identical to Ike.

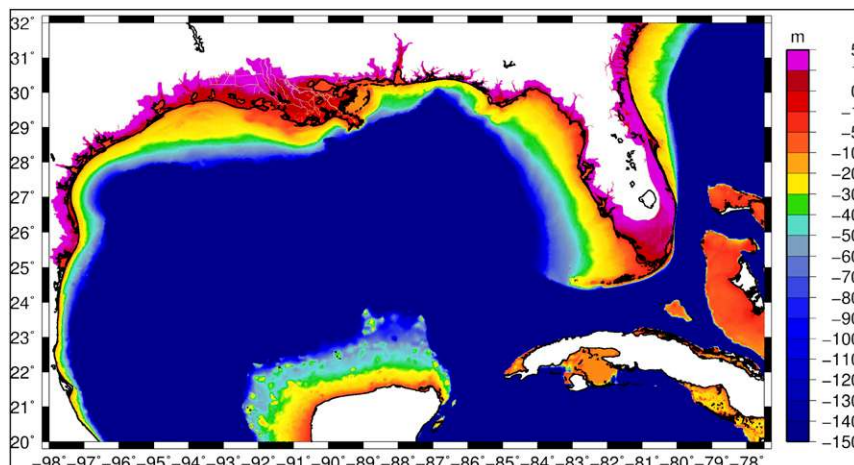


FIG. 2. Bathymetry and topography of the area of interest in the ADCIRC mesh. Negative (positive) values denote depths in meters below (above) mean sea level.

and, where possible, using available water level observations to confirm that the simulations are realistic (section 3). These best track simulations of inundation are then used as the control (CNTRL) experiments for each storm.

We then independently perturb four storm parameters as described in section 2d and rerun the ADCIRC model for each perturbed storm. This provides an ensemble of coastal flooding simulations for each storm, with the ensemble spread determined by how the differences in the meteorological inputs are translated by ADCIRC into differences in inundation. For each perturbed run, the resulting inundation is compared to the corresponding CNTRL, using the metrics presented in section 2e, to quantify the sensitivity of inundation to changes in the four perturbed storm attributes for the different storms.

b. Storm surge and inundation modeling and tidal forcing

To simulate the inundation produced by different storms, we used model version 51 of ADCIRC in two-dimensional mode. ADCIRC is a depth-integrated barotropic hydrodynamic circulation model that solves the shallow-water equations to produce domainwide water levels.

The ADCIRC model is used in conjunction with a finite-element mesh of varying resolution. The mesh used in this study was originally built and validated by Riverside and AECOM (Riverside Technology and AECOM 2015), and was provided to us by the National Ocean Service (NOS; Feyen et al. 2015), who owns and maintains the grid. It consists of almost 3 million nodes and encompasses a domain that extends along the coast

from Texas to Maine and inland to the 10-m contour with a node spacing as small as 200 m. The bathymetry and topography of the mesh for the region of interest are illustrated in Fig. 2.

We selected this mesh for this set of experiments because its higher-resolution meshing alongshore and over inland areas allows the model to resolve many aspects of inundation such as land surface heterogeneities. At the same time, the mesh covers a large geographic area, which is important for performing experiments with hurricanes in different areas of the U.S. coastline and for allowing larger track perturbations representing forecast errors at several-day lead times. A model time step of 2 s is used for the simulations.

Tidal forcing is applied at the open ocean boundary of the grid in the Atlantic Ocean with 13 tidal constituents from the TPX07.2 tidal atlas (Riverside Technology and AECOM 2015 and references within). This forcing is started approximately 15 days prior to applying meteorological forcing to allow the state to come to equilibrium. During the first 10 days of tidal forcing, a hyperbolic tangent ramping function is applied to ensure a smooth spinup of the basins. The tidal forcing continues through the simulation and thus results presented here include the tidal signal. We have chosen to include the tides in our experiments in order to represent the total inundation that coastal communities would experience.

Hurricane Ike's tidal spinup begins 23 August and ends 7 September 2008, at which point we begin applying the meteorological forcing (described in sections 2c and 2d) in addition to the tidal forcing. Hurricane Charley's spinup period is from 23 July to 9 August 2004. Charike's spinup is the same as for Ike.

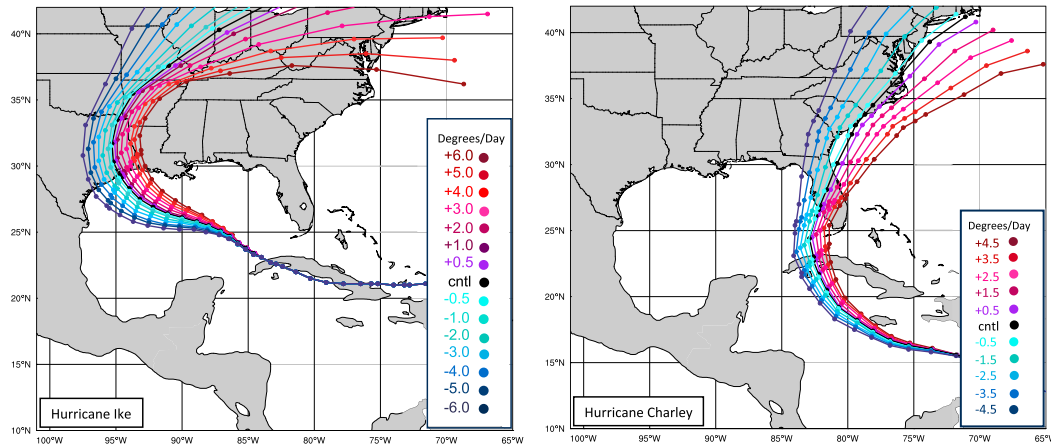


FIG. 3. TRACK perturbations for (left) Hurricane Ike and (right) Hurricane Charley.

c. Meteorological forcing

For the CNTRL simulations for each storm, the storm parameters in the best track data are used to prescribe an asymmetric gradient wind field (Mattocks et al. 2006; Mattocks and Forbes 2008) that evolves in time, which is then used to force the ADCIRC model. This asymmetric wind model modifies the Holland (1980) radial profile of gradient wind to account for the asymmetric shape of hurricane wind fields by including a different radius of maximum winds in each of the four storm quadrants. For each entry in the best track data file (typically every 6 h), ADCIRC computes the speed, direction, parameter B from Eq. (3) of Holland (1980) (Holland B parameter hereafter), and radius of the 34-, 50-, and 64-kt isotachs ($1 \text{ kt} = 0.5144 \text{ m s}^{-1}$) in each of the four quadrants. From these parameters, the model then interpolates the pressure and wind velocity to each node of the ADCIRC grid for every model time step.

The meteorological forcing is applied after the ADCIRC tidal spinup described in section 2b. For Hurricane Ike, this begins on 7 September 2008 (approximately 6 days prior to landfall) and continues until approximately 1 day after landfall, resulting in a total simulation time of approximately 22 days when considering both tidal spinup and meteorological forcing. For Hurricane Charley, the meteorological forcing begins on 9 August 2004 (approximately 5 days prior to landfall), with a total simulation time of 21 days. As with the tidal forcing, a hyperbolic tangent ramping is applied to the first 24 h of meteorological forcing to allow for a smooth adjustment of water level as the wind and pressure are interpolated to the grid.

d. Meteorological perturbations

For the perturbed simulations, we modify the storm in the CNTRL simulations by separately perturbing four

storm parameters: storm track (TRACK), maximum wind speed (VMAX), storm translation speed (SPEED), and the size of the storm (SIZE). All perturbations are applied beginning approximately 72 h prior to landfall. We focus on a 3-day lead time because initial experiments indicated that this easily contains the predictable storm surge behavior, in other words, that running for longer lead times did not provide additional information.

1) TRACK

The track perturbations are produced by applying a rate of change in “degrees per day” to the angle of storm motion, resulting in new storm tracks that veer to the left and right of the CNTRL storm. The perturbed tracks are designed so that at 72 h (landfall), their spread mimics the width of the NHC 2010–14 cone of uncertainty, which encompasses two-thirds of the historical official forecast errors in track over that 5-yr period (Cangialosi and Franklin 2016). The resulting track perturbations for each storm are shown in Fig. 3. We depict the results in terms of the distance between the landfall locations of the perturbed tracks and that of the CNTRL track.

This track perturbation approach is modeled after Fleming et al. (2008), but our method preserves storm translation speed, so that in each perturbed track the storm moves the same distance in each time period. This results in different landfall times for each perturbation (e.g., Hurricane Ike perturbed to the left will have more water to traverse before making landfall) (Fig. 3). Thus, we also compress or expand the time variation of storm intensity so that the intensity at landfall is the same in each track perturbation. However, the phase of the tide varies with the time of landfall.

2) VMAX

The VMAX perturbations are produced by applying a monotonic rate of change in maximum wind speed,

resulting in storms that have greater or weaker intensity than the CNTRL storm at landfall. The magnitude of the wind perturbations was chosen to mimic two-thirds of the mean error in NHC 2010–14 intensity forecasts (Cangialosi and Franklin 2016), as in the TRACK perturbations. In addition, we added one larger perturbation (positive and negative) to the ensemble to evaluate sensitivity to larger changes in intensity. This results in eight perturbations of ± 4 , ± 12 , ± 20 , and ± 28 kt at 72 h, which encompasses about 87% of the mean error in 2010–14 NHC intensity forecasts.

To run ADCIRC with central pressures that are consistent with the perturbed VMAX values, we used the relationships in Eqs. (7) and (8) from Holland (1980) to compute new central pressure values at each time for each perturbed maximum wind, with the Holland B parameter held constant. Since these computed central pressure values are based on Holland (1980) relationships and may not reflect the pressure in the best track data, an additional correction is made to account for this. At each time, the derived best track central pressure is computed from Eq. (1) in Holland (1980) based on the original recorded best track maximum wind. The difference between these derived and recorded pressure values for the CNTRL storm is treated as a bias correction and is added to the central pressure computed at each time on the perturbed maximum wind value.

3) SIZE

The size perturbations are generated by multiplying the wind radii (34, 50, and 64 kt) for each quadrant by a percent change, holding constant the maximum wind, to derive wind profiles that cover a larger or smaller area than the CNTRL storm. Unfortunately, because of limitations in observations, hurricane wind radii are difficult to verify (Cangialosi and Franklin 2016), and reliable estimates of errors in hurricane size forecasts are not readily available (e.g., Knaff and Sampson 2015; Cangialosi and Landsea 2016; Knaff et al. 2016). Nonetheless, the NHC does provide wind radii as part of the best track files, primarily based on forecaster analysis given available land- and satellite-based observations. Studies have shown that the forecast error in the 34-kt radii estimates is about 30–40 n mi (56–74 km) at 72-h lead time (Knaff et al. 2007; Landsea and Franklin 2013; Sampson and Knaff 2015; Cangialosi and Landsea 2016; Knaff et al. 2016).

Based on these previous studies, we perturbed storm size by choosing the extremes of the perturbations to be half and double the original size, with a sampling of perturbations in between: (-50% , -43% , -33% , -20% , 25% , 50% , 75% , and 100%). A ramping is applied during the 24 h prior to the start of the 72-h forecast period to

avoid a sudden unrealistic increase or decrease in the storm size. The size perturbation is held constant throughout the 72-h integration. Results are depicted in terms of the maximum difference in the radius of the 34-kt isotach at landfall between the perturbed storm and the CNTRL storm.

4) SPEED

The perturbations in storm translation speed are produced by modifying the times in the best track files so that the storm makes landfall earlier or later than the CNTRL. Specifically, for each perturbation, each forecast hour in the best track file is multiplied by a specified (constant) percent. For example, to increase storm speed by 10% (the $+10\%$ perturbation), the position for 120 h becomes valid at 108 h. Along-track errors in hurricane forecasts have been noted to be slightly larger than the cross-track errors (Cangialosi and Franklin 2016). Thus, we chose the SPEED perturbations ($+15\%$, $+10\%$, $+5\%$, -5% , -10% , -15% , and -20%) to produce along-track spread that is slightly larger than the NHC 2010–14 mean error in position forecasts at 72 h. Note that this method changes the timing of landfall, and so it changes the impact of tides among the perturbed simulations. Similar to the TRACK perturbations, results are depicted in terms of the along-track distance of the landfall location of the perturbed storm from that of the CNTRL.

e. Metrics

Rather than evaluate storm surge predictability using the maximum water level above the geoid (e.g., mean sea level) at any offshore or land locations (e.g., Irish et al. 2008, references within), here we focus primarily on inundation. We define *inundation depth* δ as the water depth (storm surge + tide) over normally dry land in the ADCIRC grid, where normally dry land is defined as nodes above the mean higher high water (MHHW) mark, or nodes specifically initialized as dry on the mesh. Because we reference the inundation metrics to MHHW, which is the mean water level for a location at high tide, the results can be interpreted practically as the combined flooding impact from storm surge and tide over areas normally unaffected by tide. We use metrics that quantify inundation instead of the maximum water level, much of which is offshore, in order to evaluate the predictability of storm surge in terms of flooding over land—which is the surge-related information needed by emergency managers and coastal populations to make risk management decisions as a tropical cyclone approaches.

We evaluate inundation depth using two types of metrics. The first type is point-based metrics [e.g., spatial correlation coefficient r^2 and root-mean-square error (RMSE)]. By measuring differences between the

control and perturbed simulations using r^2 and RMSE, we assess sensitivity of inundation across many specific locations. For the r^2 and RMSE metrics, we use the maximum inundation at each point, at whatever time during the simulation it occurs, and do not area weight the control–perturbation pairs. No distinction is made in the location of inundation depth points based on local geographical features (e.g., estuaries, rivers), since our emphasis is not on the details of inundation produced by these complex local effects.

While point-based metrics such as r^2 and RMSE are commonly used to evaluate storm surge predictions, they can also be noisy owing to the finescale heterogeneity of the land characteristics and inundated locations. As discussed above, we also wanted to explore whether metrics that measure inundation integrated across points have the potential to enhance the predictability of storm surge, in other words, to extend skillful predictions to longer lead times. Thus, we also introduce a second type of metric, an integrated measure referred to as *inundation volume*.

To compute inundation volume, we multiply the inundation depth by the cell area and sum over nodes that are normally dry (above MHHW) where the inundation depth exceeds 1 m.⁴ To filter out inundation in outlying areas far away from landfall that are not related to the primary geographic region of surge-induced inundation, we also trim the left- and rightmost 2.5% of the inundation volume. Henceforth, we refer to this trimmed inundation volume as simply *inundation volume*. Note that while r^2 and RMSE are applied to the maximum inundation depth at each point throughout each run, which can occur at different times, the inundation volume metric occurs at a single time—the time of maximum inundation volume—which can vary from run to run.

We examine two measures of inundation volume: storm following and fixed location. The storm-following measure computes the inundation volume (above 1 m) where the inundation occurs for each (perturbed or CNTRL) storm. The fixed-location measure computes inundation in a fixed geographic region, regardless of where the storm goes. We designed the fixed-location measure to evaluate the predictability of integrated (rather than point based) coastal inundation in a given region that may be of interest to emergency management personnel and coastal populations for making evacuation decisions. The fixed-location inundation region can be defined anywhere of interest; here, we choose it to be the trimmed area inundated by the

CNTRL storm (i.e., all points in the CNTRL run at the time of greatest inundation volume where $\delta \geq 1$ m and between the 95% bounds).

3. Evaluation of control simulations of surge and inundation

Before examining surge predictability, we evaluate our CNTRL experiments, to confirm that our ADCIRC modeling is doing a sufficiently realistic job of translating storm characteristics into storm surge and inundation for the purpose of our experiments. For Hurricane Ike, the maximum water level in the control simulation surpasses 6 m, with a spatial coverage that spreads from the Galveston Bay area to the east (Fig. 4). Qualitatively, this is consistent with observations of Ike’s surge (Berg 2014).

Because the storm surge from Hurricane Ike was widespread, observations of water level were available from multiple U.S. Geological Survey (USGS) pressure sensors. To evaluate Ike’s CNTRL qualitatively, we used data from the 38 USGS sensors that recorded high water marks (East et al. 2008), were within 500 m of a mesh point in the ADCIRC grid, and were categorized as “surge only” (i.e., did not include factors not represented in our modeling, such as freshwater riverine flows, beach waves, or other unresolved processes). Each observation is compared to the modeled maximum water level at the closest ADCIRC mesh point.

Figure 5 shows a scatterplot of observed versus simulated high water marks. Timing of the simulated surge was also realistic compared to time series of observed water height from several NOAA buoys, such as at Port Arthur, Texas (8770475; NOAA/NOS/CO-OPS 2016), and the closest mesh point (not shown). Other researchers have focused on simulating the details of Ike’s surge more accurately (Kennedy et al. 2011; Hope et al. 2013; Kerr et al. 2013; Sebastian et al. 2014), for example, by using a more geographically focused, higher-resolution mesh and including a wave model. However, our qualitative evaluation and the correlation of 0.63 with observations indicate that the macroscale behavior of the Ike CNTRL simulation is sufficiently realistic to suit our goal of investigating the sensitivity of coastal inundation to large-scale errors in hurricane forecasts.

Because Hurricane Charley’s surge affected only a small area, only one NOAA tide gauge was affected (Wang et al. 2005). So, we evaluate the CNTRL simulation qualitatively. In Charley’s CNTRL, the spatial coverage of the surge is much smaller than in Ike, with maximum water levels confined to the Fort Myers region nearing 4 m (Fig. 6). This is consistent with visual reports (Pasch et al. 2011) and other simulations of Hurricane Charley storm surge (Weisberg and Zheng 2006b).

⁴The measurement of 1 m roughly matches the 3-ft NWS watch and warning threshold for coastal flooding.

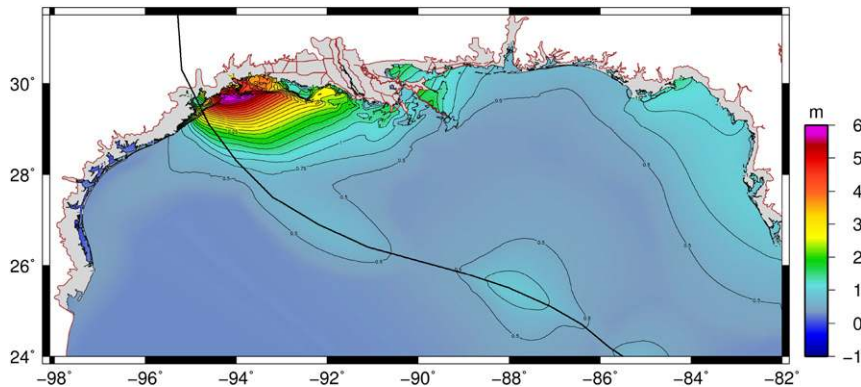


FIG. 4. Maximum water level in meters above mean sea level for CNTRL simulation of Hurricane Ike. Black line denotes storm track.

4. Results from sensitivity experiments

We now investigate the sensitivity of surge forecasts to changes in different storm parameters by comparing the perturbed simulations to the control simulations for each storm. We start with a brief qualitative comparison, and then the remainder of this section focuses on quantitative comparison using the different metrics of coastal inundation discussed in section 2e.

Examining maps of the maximum water level produced by different perturbations for Ike (Fig. 7) and Charley (Fig. 8), we see that the changes in surge produced by the storm perturbations behave as one would expect. A more intense or larger storm results in larger surge (Figs. 7d,f and 8d,f), and a storm track veered to the left moves the storm surge with it, to some extent (Figs. 7a and 8a). The amount of overlap in areas that are inundated varies with the storm and the parameter being perturbed. The primary differences between Ike and Charley that contribute to the differences in how surge responds to the storm perturbations are the size of the storm, the bathymetry within the track envelope, and, to a lesser extent, the intensity; the role of these differences is discussed further below when we compare with Charike.

a. Point-based metrics of sensitivity and predictability

1) TRACK

Figure 9 depicts the sensitivity of coastal inundation to the hurricane track perturbations using point-by-point comparisons of inundation depth (r^2 and RMSE; see section 2e), for each of the three storms examined. This location-specific inundation depth for all three storms exhibits a strong sensitivity to the storm track, as indicated by the rapid decrease in correlation coefficient (increase in RMSE) in each panel as the track deviates

farther to the left or right of the control. This broad pattern is expected since the level of inundation at any given point depends on the hurricane's track, as demonstrated by previous research (e.g., Zhong et al. 2010). For a small storm like Charley, in particular, perturbing the storm's track produces inundation in completely different locations (Figs. 8a,b).

Comparing the results across the three panels in Fig. 9, we also see significant differences among the three storms. These differences indicate that the sensitivity of coastal inundation to hurricane track perturbations is influenced by the characteristics of both the storm and the landfall location. First, we compare the ± 1.0 perturbations ($\pm \sim 50$ km from CNTRL landfall) for Ike and

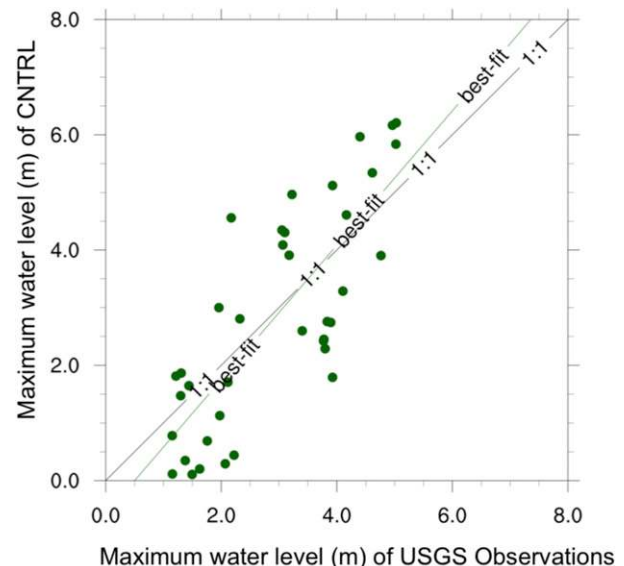


FIG. 5. Comparison of USGS pressure sensor observations for Hurricane Ike to CNTRL (see section 3). The r^2 is 0.63, the scatter index is 0.41, and the RMSE is 1.18 m.

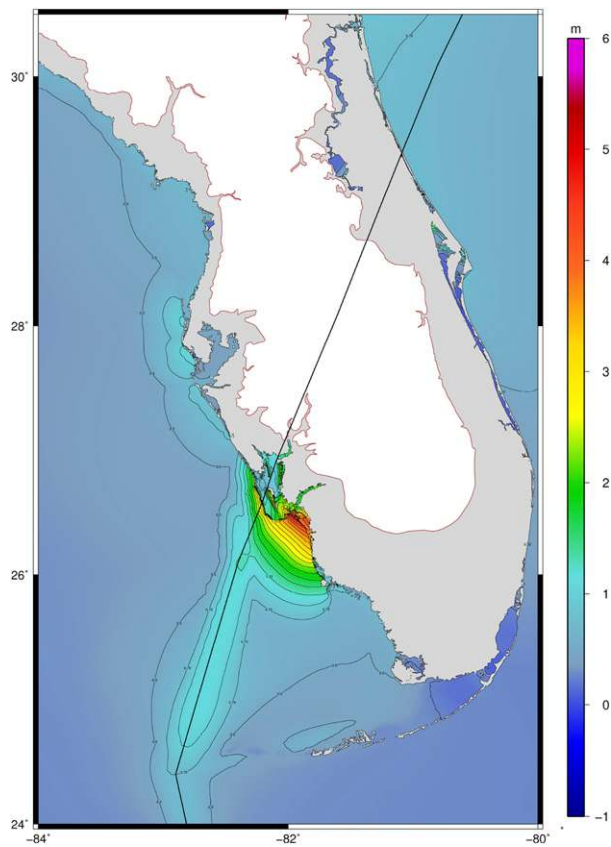


FIG. 6. As in Fig. 4, but for Hurricane Charley.

Charike (Figs. 9a,b), which represent storms of different size and intensity making landfall in the same general location. This shows that Ike's r^2 drops from 1 in the control to about 0.8 (Fig. 9a), while Charike's r^2 drops to about 0.4 (Fig. 9b). In other words, even with similar bathymetric properties influencing inundation, Charike's much smaller size leads its point-based inundation depth to be much more sensitive to track perturbations than the larger Ike.

Second, we compare Charike with Charley (Figs. 9b,c), which represents the same storm making landfall in a different location. In this comparison, the correlation coefficient for the ± 1.5 ($\pm \sim 50$ km from CNTRL landfall) perturbations drops from 0.4 for Charike to less than 0.2 for Charley. Since the storm size and intensity remain the same in this comparison, this indicates that (at least for a storm of the size and strength of Hurricane Charley) the differences in bathymetry over which the storm travels (Fig. 2) and the landfall location also influence the sensitivity to track perturbations. In other words, performing experiments with the hypothetical Charike helps us understand how different hurricane size and intensity influences the predictability

of inundation, compared to different bathymetry and landfall location.

What do the results in Fig. 9 mean in terms of the practical predictability of location-specific surge-induced inundation, based on typical errors in hurricane track forecasts? To interpret the results from the sensitivity experiments from this perspective, we overlay on the x axis of Fig. 9 a mapping of the distance of the track perturbation from the control at landfall into the lead times at which current hurricane forecasts exhibit that mean level of error, based on the NHC 2010–14 track forecast error statistics. Viewing the sensitivity results in terms of characteristic forecast errors also allows us, in later sections, to directly compare the surge predictability implications of forecast errors in different storm parameters.

As Fig. 9 shows, when accuracy is measured using point-based metrics, the predictability of inundation is severely limited by the average uncertainty in current forecasts of hurricane track. Using a correlation coefficient of 0.6 as a general guide, at best, a large storm like Hurricane Ike retains some predictive skill for location-specific inundation for about 24 h prior to landfall. The much smaller storms, Charley and Charike, reveal a loss of predictive skill for location-specific inundation by 12 h. As discussed above, this is illustrated in Figs. 8a and 8b compared to Fig. 6, where a relatively small change in Charley's track produces inundation in different locations.

An alternative way of interpreting these results is in terms of the hurricane forecast accuracy (on average, or in a specific case) that would be required for providing useful forecasts of location-based inundation depth at a specified lead time. In order for forecasts to provide skillful predictions of coastal inundation depth at specific locations at least two days in advance of landfall, track forecasts for a small storm such as Charley would need be much more accurate, exhibiting only about 25% of the mean error in current Atlantic tropical cyclone forecasts.

2) VMAX

Results for the VMAX perturbations are shown in Fig. 10, for Ike and Charley. The point-by-point inundation depth exhibits similar sensitivity to VMAX for both storms, as measured by correlation. Overall, Figs. 7c,d and 8c,d show that increasing storm intensity produces increased surge, as in previous related work (Weisberg and Zheng 2006a; Sebastian et al. 2014; Favre et al. 2015). However, because varying intensity produces little change in the location of inundation along the coast (holding track fixed; Figs. 7c,d and 8c,d), the point-based correlation coefficient shows little

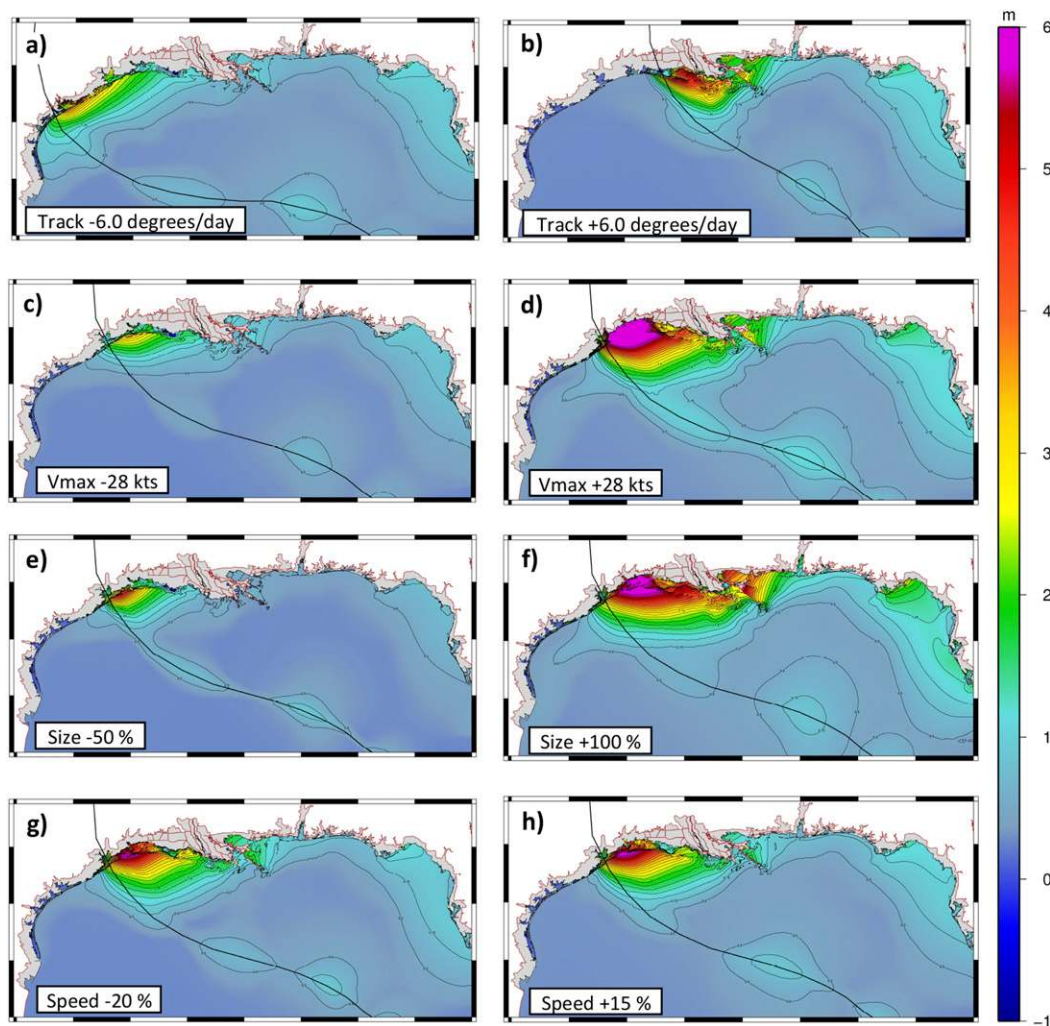


FIG. 7. Maximum water level in meters above mean sea level for Hurricane Ike perturbations: (a) TRACK $-6.0^{\circ} \text{ day}^{-1}$, (b) TRACK $+6.0^{\circ} \text{ day}^{-1}$, (c) VMAX -28 kt , (d) VMAX $+28 \text{ kt}$, (e) SIZE -50% , (f) SIZE $+100\%$, (g) SPEED -20% , and (h) SPEED $+15\%$. Black line denotes storm track.

sensitivity to VMAX perturbations for either storm. Similarly, for the smaller storm Charley, RMSE shows limited sensitivity to VMAX.

For the larger storm Ike, RMSE shows more sensitivity to VMAX, likely corresponding with the change in maximum water levels shown for the VMAX perturbations in Figs. 7c and 7d. To evaluate the practical implications of this RMSE sensitivity to changes in VMAX for Hurricane Ike (Fig. 10a), we can compare these results with those for changes in Hurricane Ike's TRACK (Fig. 9a), from the perspective of mean VMAX and TRACK forecast errors at different lead times. This comparison shows that an RMSE value of about 0.6 is reached at about 72-h lead time in VMAX, but at only 24-h lead time in TRACK. Even for the larger VMAX perturbation, TRACK perturbations produce a larger

decrease in predictive skill. This suggests that, on average, the error in current hurricane track forecasts is more limiting to storm surge predictability than the error in current intensity forecasts, as measured by the point-based RMSE or correlation coefficient metrics.

3) SIZE

Results from the SIZE perturbations for Ike and Charley are shown in Fig. 11. For Ike, which was already a large storm to begin with, increasing or decreasing the storm's size produces an increase in RMSE for inundation. For increasing or decreasing size, this likely reflects the increased or decreased spatial area inundated as well as a corresponding increase or decrease in water level at specific locations, as illustrated in Figs. 7e and 7f. Charley also exhibits sensitivity to size

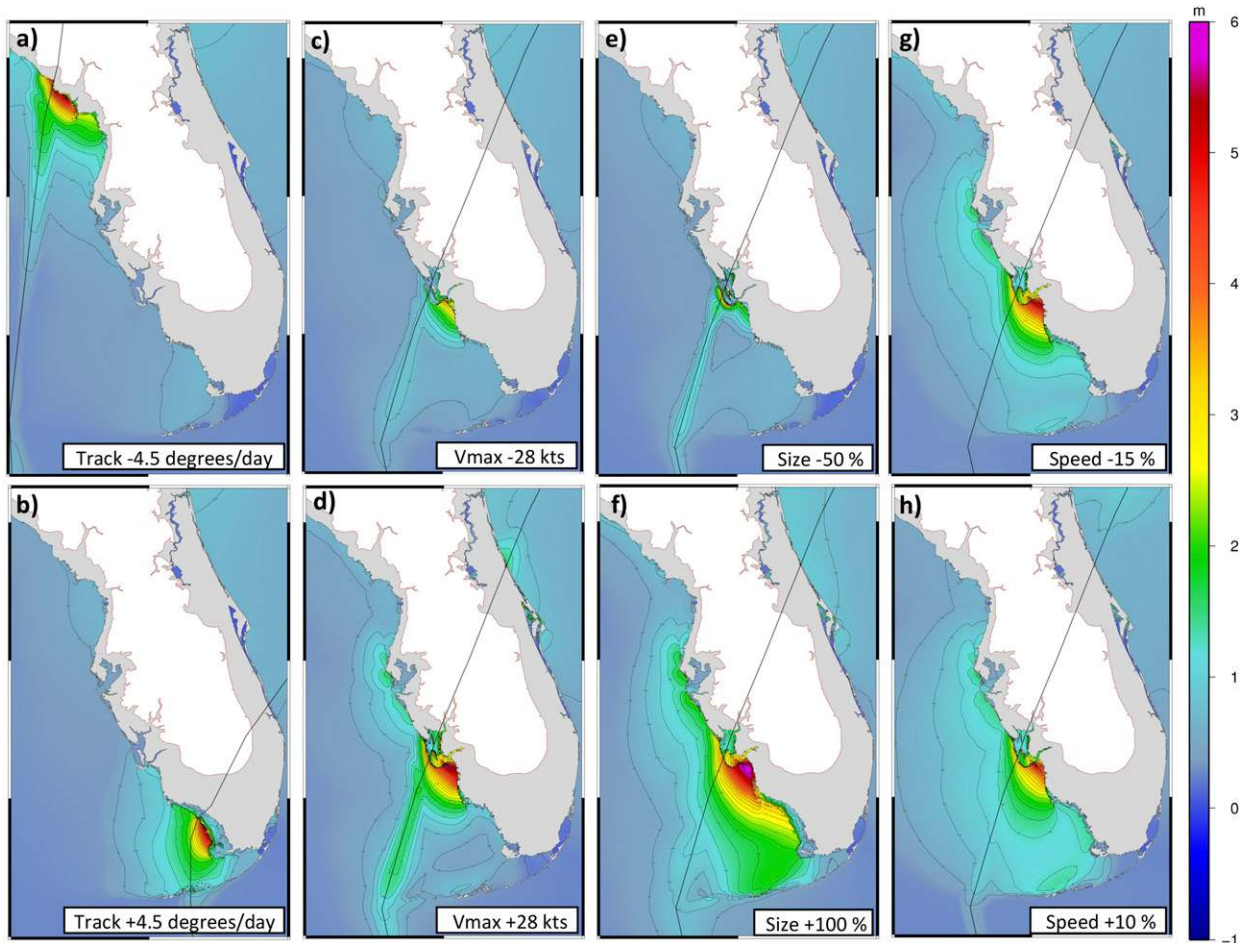


FIG. 8. Maximum water level in meters for Hurricane Charley perturbations: (a) TRACK $-4.5^{\circ} \text{ day}^{-1}$, (b) TRACK $+4.5^{\circ} \text{ day}^{-1}$, (c) VMAX -28 kt , (d) VMAX $+28 \text{ kt}$, (e) SIZE -50% , (f) SIZE $+100\%$, (g) SPEED -15% , and (h) SPEED $+10\%$. Black line denotes storm track.

perturbations (Fig. 11b), with overall spatial patterns similar to Ike (Figs. 8e,f). These results are consistent with previous studies concluding that storm surge is sensitive to storm size and generally increases or decreases with increasing or decreasing size, respectively (Weisberg and Zheng 2006a; Irish et al. 2008; Faivre et al. 2015).

The translation of forecast SIZE errors into lead time in Fig. 11 is based on Cangialosi and Landsea (2016). This emphasizes the large contribution of observational uncertainty in storm size forecast errors. Errors characteristic of 12-h forecasts are 37 km, which encompasses all of the SIZE perturbations for Charley, and much of the range of perturbations for Ike. Comparison of Figs. 9–11 indicates that at short forecast lead times (e.g., less than 12 h), SIZE errors, mainly due to observational uncertainty, will dominate uncertainty in point-based coastal inundation forecasts. At lead times of 12 h or greater, TRACK forecast errors become the greatest

limitation of predictive skill, followed by SIZE errors and then VMAX.

4) SPEED

The results from the SPEED perturbation runs (Fig. 12) show that location-based inundation depth exhibits little sensitivity to changes in storm translation speed for either Ike or Charley, as indicated by the high correlation values and low RMSE. When SPEED is perturbed, the spatial extent and magnitude of the maximum water levels remain relatively unchanged in both Hurricane Ike and Hurricane Charley (Figs. 7g,h and 8g,h, respectively). Thus, for these storms the point-based inundation is relatively insensitive to errors in forecasts of storm speed, compared to the other storm attributes.

This differs from conclusions from other studies (e.g., Zhong et al. 2010; Faivre et al. 2015) who conclude that storm surge is sensitive to storm speed when considering peak surge values at specific locations. One reason that

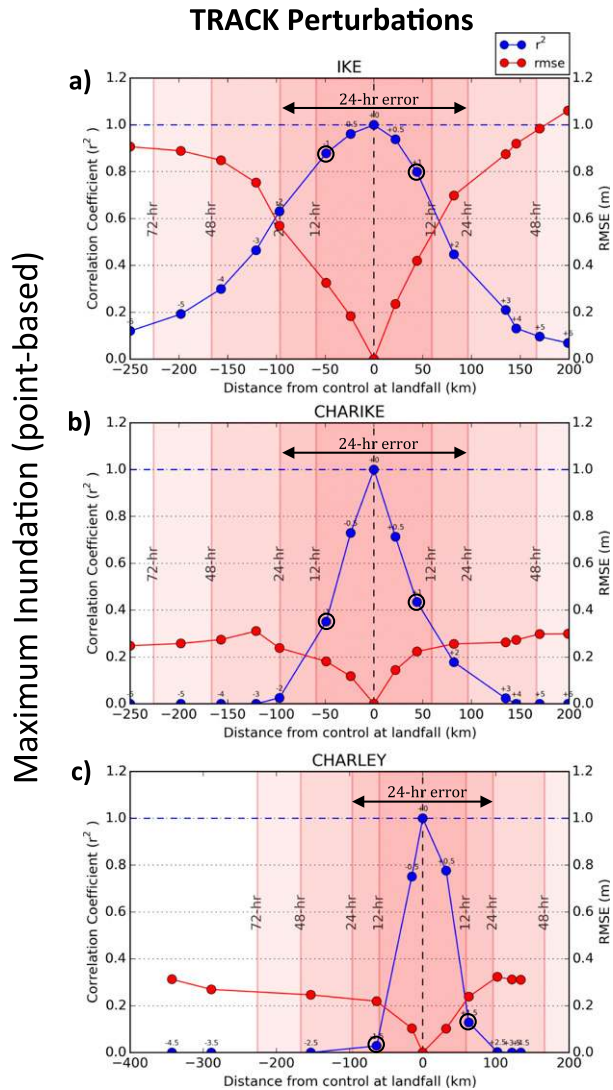


FIG. 9. Correlation coefficient (blue) and RMSE in meters (red) of the point-by-point inundation produced by TRACK perturbations compared to CNTRL for (a) Ike, (b) Charike, and (c) Charley. Shaded red areas translate forecast errors in terms of distance into lead time, based on mean track forecast errors at different lead times. Black circles surround the correlation coefficients for the perturbations representing $\pm \sim 50$ km from CNTRL landfall. Blue horizontal dashed line represents a correlation coefficient of 1.0.

our results may differ is that we consider the r^2 and RMSE of inundation depth across the inundated region, as opposed to selected observation points. Note that changing SPEED does alter the phasing of landfall relative to the tide, which is included in our experiments. Thus, the low sensitivity of inundation to SPEED suggests that the storm-induced surge makes a greater contribution to inundation than the tide for these storms. Sensitivity to SPEED is likely to be greater in coastal regions where the tidal signal is stronger.

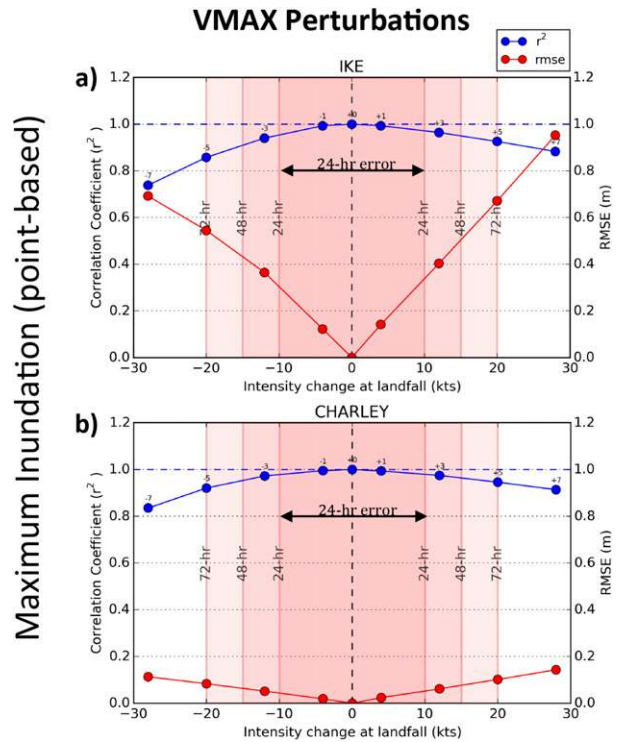


FIG. 10. As in Fig. 9, but for VMAX perturbations for (a) Hurricane Ike and (b) Hurricane Charley.

The change in storm speed does change the timing of landfall, and thus uncertainty in SPEED leads to uncertainty in timing of inundation (which can be important for evacuation decisions). These differences are not clearly shown in the r^2 and RMSE metrics, and are discussed in the next section.

b. Inundation volume: Toward a storm-following integrated metric of coastal inundation

The results shown above begin to quantify the practical predictability of location-specific coastal inundation across a range of lead times. However, Figs. 7 and 8 suggest that the point-based metrics do not fully capture the information that the experiments provide about the predictability of surge. In particular, Figs. 7 and 8 suggest that there may be a spatially integrated metric that retains predictive skill for longer lead times. To quantify the visual impression, we examine results from the same set of experiments using inundation volume metrics (section 2e), that is, a bulk characterization of the inundation. This allows us to explore whether any predictability is gained by integrating inundation across a region, instead of using point-based metrics.

We use two measures of inundation volume (storm following and fixed location; see section 2e), to explore the predictability of coastal inundation where the storm

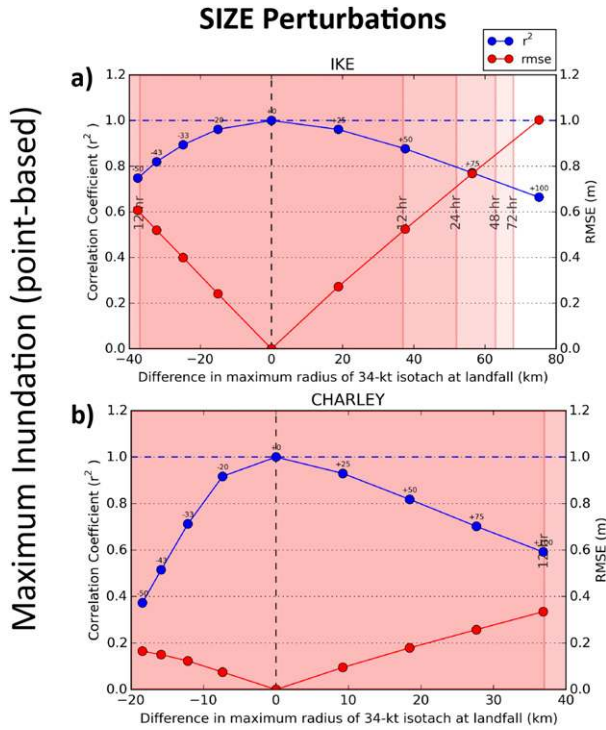


FIG. 11. As in Fig. 9, but for SIZE perturbations for (a) Hurricane Ike and (b) Hurricane Charley.

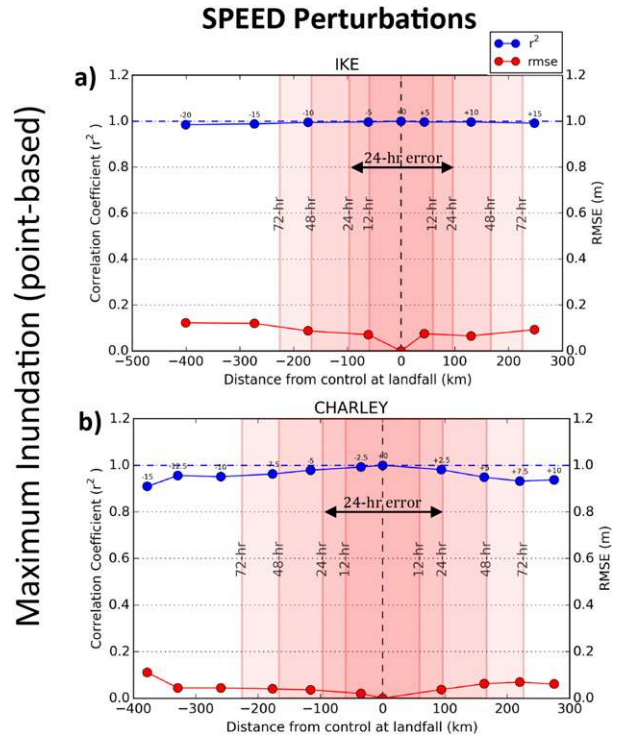


FIG. 12. As in Fig. 9, but for SPEED perturbations for (a) Hurricane Ike and (b) Hurricane Charley.

goes (storm following) with that in a specific region (fixed location). The storm-following measure allows us to evaluate predictability of the volume of coastal inundation from storm tide associated with a certain type of storm wherever that inundation occurs (even when there is significant uncertainty in the landfall location). The fixed-location measure may help to discern predictability for a particular region for example, for emergency managers and coastal populations who are making evacuation decisions for that region. Since the storm-following inundation volume captures the full inundation, wherever it occurs, it can be larger than the fixed-location inundation volume because the area of inundation expands outside the fixed location when the storm is larger or stronger than the CNTRL storm.

The inundation volume metrics for the TRACK perturbations of Ike, Charley, and Charike are plotted in Fig. 13. Note that for our initial exploration of inundation volume, we are using a total volume measure which does not specify inland or coastal extent to the left or right of the storm center; as shown in Figs. 4 and 6–8, the vast majority of the inundation is to the right of the storm track.

For Hurricane Ike, when the storm track is perturbed to the left of the control, the inundation volume decreases rapidly, in both measures. However, when the storm is perturbed to the right, both the storm-following

and fixed-location inundation volumes remain fairly constant for the first few perturbations (Fig. 13a). Compared with Fig. 9a, this indicates that Ike’s inundation exhibits less sensitivity to those TRACK perturbations when using the integrated metrics than the point-based metrics. In other words, even if the storm track veers 100 km to the right of the forecast, Ike’s inundation volume remains predictable, overall and in the specific area near Galveston inundated by Ike’s storm surge. This suggests that, at least for a large storm like Ike, spatial integration of the inundation may provide a metric with enhanced predictability compared with point-specific measures.

For Hurricane Charley and Charike, however, the storm-following and fixed-location inundation volumes are more sensitive to TRACK perturbations (Figs. 13b,c). This suggests that even with an integrated metric, the coastal inundation produced by small storms is highly sensitive to the storm’s track and the characteristics (bathymetry, topography) of the landfall region. In other words, even though the surge bulge looks similar in maps such as Figs. 8a and 8b compared to Fig. 6, the inundation over land can be quite different.

To interpret the inundation volume results in the context of practical predictability at different lead times, we again overlay on the x axis of Fig. 13 a translation of

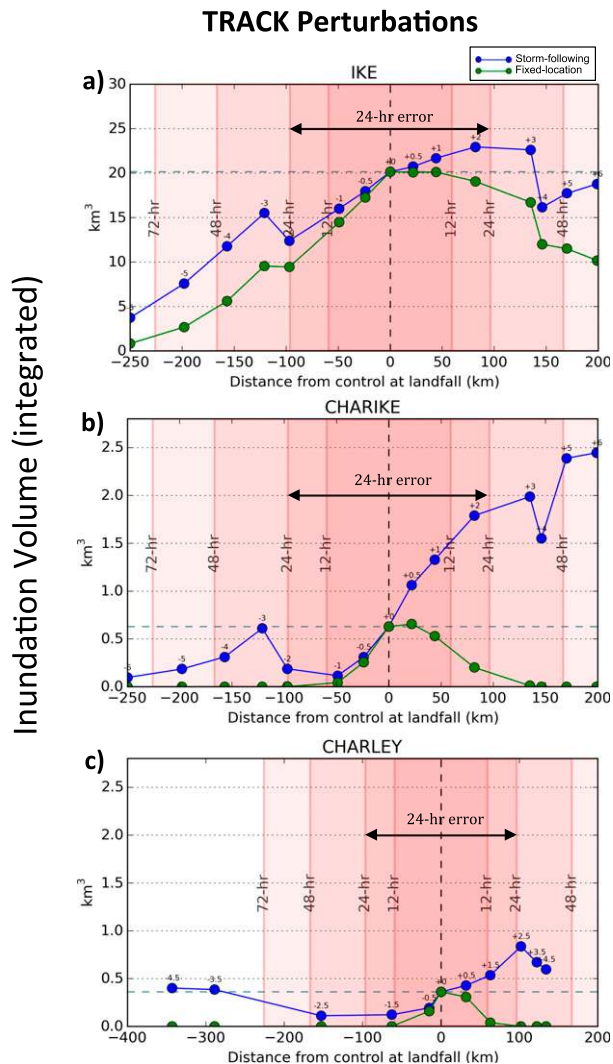


FIG. 13. Storm-following (blue) and fixed-zone (green) inundation volumes (km^3) produced by TRACK perturbations compared to CNTRL for (a) Ike, (b) Charrike, and (c) Charley. As in Fig. 9, shaded red areas translate forecast errors in terms of distance into lead time. Green horizontal dashed line represents inundation volume of CNTRL.

the mean hurricane track forecast error in terms of distance into lead time, based on the NHC 2010–14 track error statistics. The results suggest that at least for track perturbations to the right of the control, storm-following inundation volume retains predictability out to 72 h for Ike, which is much greater than with the location-specific inundation metrics discussed in section 4a(1). However this is not the case for Charley or Charrike, for which predicting coastal inundation within 24 h of landfall remains difficult even with integrated metrics.

For the intensity (VMAX) and storm size (SIZE) perturbations, the sensitivity results for the inundation volume metrics reveal a quasi-linear behavior that is similar

for both Ike and Charley (Figs. 14 and 15). The results are similar to those presented in Rego and Li (2009), where an increase or decrease in hurricane intensity or size leads to a corresponding increase or decrease in inundation volume. The change of inundation volume across the range of perturbations is approximately a factor of 10 for both Ike and Charley, indicating that both intensity and size forecast errors can be important in ways not fully captured by the point-based r^2 and RMSE metrics.

Next, we compare the results in Figs. 13–15 in terms of lead time. For Ike it appears that VMAX and SIZE perturbations have a greater effect on inundation volume than TRACK perturbations, particularly when we consider a storm-following metric. For instance, VMAX or SIZE perturbations characteristic of 72-h forecast errors produce a factor of 2 or greater change in the inundation volume relative to CNTRL. TRACK perturbations produce a comparable change only for perturbations to the left of the track in CNTRL. For tracks to the right, the storm-following inundation volume exhibits less change.

For Charley, on the other hand, changes in storm-following inundation volume produced by TRACK perturbations still dominate those produced by changes in VMAX or SIZE. These results suggest that the relative importance of perturbations (or forecast errors) in different storm parameters varies with the surge metric and the storm.

The inundation volume results from the SPEED perturbations are shown in Fig. 16. Similar to the point-based results (Fig. 12), inundation volume exhibits little sensitivity to the translation speed of the storm aside from the timing of inundation (not shown). For both Ike and Charley, slowing the storm produces an increase in inundation volume, similar, but lower in magnitude, to results shown by Rego and Li (2009). For Ike, comparing Fig. 16 with Figs. 13–15 indicates that inundation volume is less sensitive to changes in SPEED than changes in TRACK, VMAX, or SIZE. For Charley, changes in SPEED produce changes in inundation volume that are smaller than changes in TRACK or SIZE, but of similar magnitude to changes in VMAX; for example, the slowest storm shows an increase in inundation volume that is approximately equivalent to a 28-kt increase in storm intensity, to a category 5 storm (Fig. 14).

When interpreting the results from the translation speed perturbations, it is important to consider the role of tides in the inundation. This tidal signal can be seen with the increase of inundation volume in the slowest and fastest storms for Hurricane Charley (Fig. 16b); the temporal displacement of landfall time for these storms is approximately 10 h, which covers most of the tidal period of that region. For Hurricane Ike, on the other

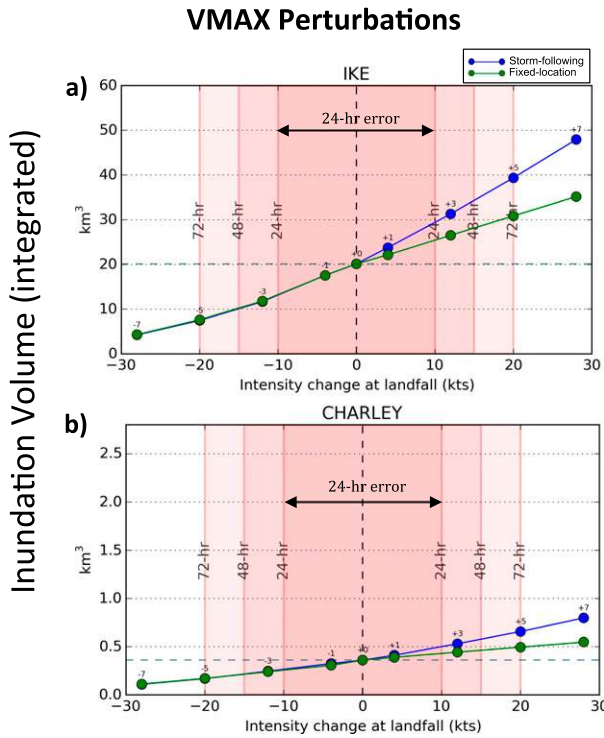


FIG. 14. As in Fig. 13, but for VMAX perturbations for (a) Hurricane Ike and (b) Hurricane Charley.

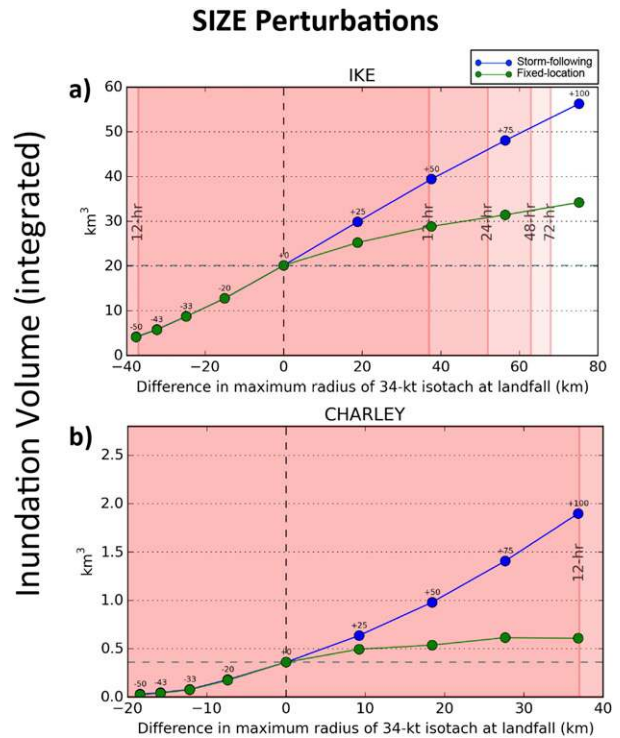


FIG. 15. As in Fig. 13, but for SIZE perturbations for (a) Hurricane Ike and (b) Hurricane Charley.

hand, the tidal signal is not as evident. These results indicate that the tidal contribution is a larger portion of the inundation for a smaller storm with a small inundation magnitude compared to a large storm for which the magnitude of the surge from the storm itself may overwhelm the tidal contribution.

The impact of tides on inundation volume undoubtedly plays a larger role for regions in which tides are larger (e.g., the U.S. East Coast). For example, Colle et al. (2008, 2015) show the importance of the tidal phase for coastal inundation in the New York City region. Given the large population at risk along the East Coast, the contribution of tides to the predictability of coastal inundation is an important area for future work, as is the influence of changes in storm translation speed and landfall phasing with the tide on inundation timing.

Overall, the inundation volume results indicate that for medium or large storms with less complicated bathymetric variations along the coast, a volume-integrated storm-following metric of inundation may offer greater predictability. This metric does not provide information about the inundation at any given location. But at longer lead times where the landfall location is still highly uncertain and thus the predictability of inundation in a given location is low, it may still be useful to forecast that the inundation is likely to be large wherever it occurs. For a

small storm like Hurricane Charley, however, a storm-following volume-integrated metric fails to offer greater predictability.

5. Summary and discussion

An ensemble of coastal inundation (storm surge + tide) simulations was run for three hurricanes in the Gulf of Mexico, systematically perturbing four storm parameters, to quantify the sensitivity of coastal inundation to each of the storm attributes using multiple measures of inundation. The three storms (Ike, Charley, and the hypothetical Charike) were selected so that the ensembles would span a range of possible storms (in terms of size and intensity) making landfall in different regions of the Gulf of Mexico. This allows us to interpret the findings in terms of sensitivity of inundation to different types of hurricane forecast errors, for different types of storms. By translating mean hurricane forecast errors into lead times, we also examine the implications of those sensitivities for the practical predictability of storm surge.

Results show that when the skill of surge forecasts is assessed with point-based metrics of maximum inundation depth (including tide), the practical predictability of storm surge is limited, on average, primarily by current errors in hurricane track forecasts, followed by

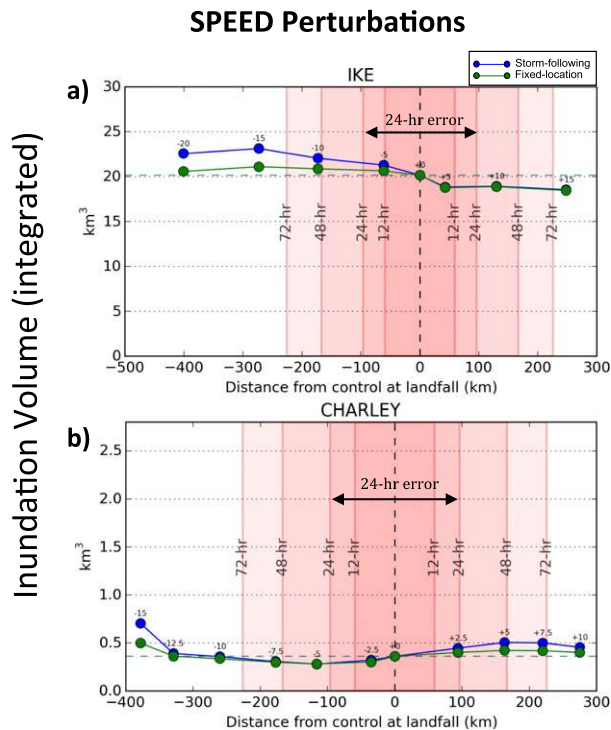


FIG. 16. As in Fig. 13, but for SPEED perturbations for (a) Hurricane Ike and (b) Hurricane Charley.

intensity forecast errors, and then translation-speed forecast errors. However, if we consider potential improvements in track forecasts, such that errors in 72-h forecasts are similar to mean errors in today's 24-h NHC forecasts, then the observational uncertainty in storm size, measured by the radius of gales, becomes the leading source of error in coastal inundation forecasts.

Using point-based metrics and a correlation coefficient of 0.6 as a guideline for useful forecast skill, inundation from storm tide exhibits practical predictability only for lead times of approximately 24 h for a large storm such as Hurricane Ike. Predictability lead times are even lower for smaller storms such as Charley or Charike. Note that these conclusions are based on mean forecast errors; in reality, hurricane track forecasts vary significantly from storm to storm. Thus, if one can accurately assess in advance that the track of a specific hurricane is more predictable than average, location-specific inundation forecasts may be possible further in advance. Alternatively, if track uncertainty is greater than the perturbations used here, storm surge predictability will be even more limited.

One caveat for interpreting these results is that we only examined storms in the Gulf of Mexico, for which the tidal phase makes a relatively small contribution to coastal inundation. Thus, aspects of our results—especially those

from experiments in which the timing of landfall changes—are not directly transferrable to other regions where tides are more significant. Future research would benefit from examination of hurricanes making landfall in locations where tides make a larger contribution to the inundation (e.g., New York City; Colle et al. 2008, 2015). In such cases, we anticipate the sensitivity of inundation to perturbations in translation speed, and in some cases track, to be greater than shown here, due to changes in how the surge is phased with the tide. The implications for practical predictability likely vary with the characteristics of the storm and landfall region, but we anticipate that when the tidal signal is larger, this greater sensitivity will cause the practical predictability limits to be even lower than those shown here for the Gulf of Mexico storms.

Our results also illustrate how the relative importance of different factors that contribute to the limits of storm surge predictability can vary with the storm and landfall location. For example, for a small storm like Hurricane Charley in a region of substantial along-coast variations (e.g., bathymetry or topography), practical predictability of coastal inundation measured at specific points is nearly zero. Additionally, our experiments with the hypothetical Charike compared to Ike indicate that for track perturbations, a smaller storm over the same landfall location exhibits a larger sensitivity. This suggests that both storm size and bathymetric characteristics are influential. For intensity perturbations, on the other hand, inundation exhibits more sensitivity for a large storm than for a smaller storm.

Given the current predictability limits for location-specific coastal inundation, we also began investigating the predictability of surge evaluated using spatially integrated metrics, such as inundation volume. We asked the following: By using a storm-following integrated metric, might one be able to predict the potential for a larger or smaller inundation based on a storm's characteristics, even at longer lead times when location-specific surge predictability is low due to uncertainty in the landfall location? Our results suggest that yes, spatially integrated metrics can extend surge predictability, but this appears possible only for large storms, in certain situations. For example, large storms such as Ike produce many times the inundation volume of small storms like Charley and Charike, even if the smaller storm is more intense.

These results suggest that there is promise for storm-following volume-integrated metrics to provide guidance about the potential for inundation from storm surge in the most widespread, destructive cases, beyond the 12–24-h lead times at which location-specific coastal inundation is predictable with current levels of hurricane forecast skill. Since evacuations in many U.S. coastal areas must be

initiated 36–72 h before the anticipated arrival of tropical storm-force winds, such information has potential to be useful for emergency management and populations at risk (e.g., Wolshon et al. 2005; Lindell et al. 2007; Demuth et al. 2012; Morrow et al. 2015). Further work would be needed, however, to explore the potential of different integrated surge metrics, as well as their limitations.

To systematically explore storm surge predictability in an idealized context, we used a perfect-model framework to conduct the experiments and evaluate the results. Although model error is a contributor to errors in real-world surge predictions, our purpose was to assess the sensitivity of surge-induced inundation to systematically applied storm perturbations. Unless our simulation is severely biased or fails to represent key dynamics of the system of interest (which section 3 indicates is not the case), sensitivity may be assessed relative to our CNTRL simulation. If one wishes to assess actual forecast accuracy, the added error in the CNTRL must also be considered. In this way, our conclusions about practical predictability are likely upper bounds; real-world surge predictability, in the presence of model error, is likely to be even more limited.

Hurricane forecast uncertainty is represented in our study by using a parametric wind model to generate an idealized representation of historical storms, and then systematically perturbing the storm's track, intensity, size, and speed. We scaled the perturbations based on mean errors in current forecasts; for example, track perturbations are based on the NHC "cone of uncertainty." This approach does not include the full range of possible forecast errors. However, we selected this approach to provide a simplified, but still realistic, setting for exploring surge predictability, with limited degrees of freedom and sources of error interactions. We anticipate that future work will investigate similar questions with a more complex representation of atmospheric uncertainty (e.g., using ensembles from a numerical weather prediction model). Results from this more simplified set of perturbations can then serve as a useful guide for interpreting results from more complicated experiments.

In addition, the bathymetric characteristics and coastal landscape of the landfall location significantly influence surge-induced inundation in a region, in ways that we do not fully explore here. Although we chose not to emphasize these factors in this study, we created a hypothetical storm, Charike, as a first step toward disambiguating the roles of storm parameters and landfall locations. This study also does not address surge associated with extratropical storms and storms undergoing extratropical transition. As with addressing the impact of tides, examining more storms and landfall locations in future studies will help further investigate these important aspects of surge predictability.

Despite these limitations, this study makes important contributions to the current understanding of storm surge predictability. Scientifically, the experiments using a perfect-model framework combined with systematic perturbations of idealized representations of historical storms makes valuable conceptual and methodological contributions. For hurricane and surge researchers and forecasters, the results emphasize the importance of probabilistic (rather than deterministic) storm surge forecasting for lead times longer than 12–24 h, shorter for smaller storms, given current errors in hurricane forecasts. They also suggest that in order to efficiently sample the probability space, designing an ensemble for probabilistic forecasting (e.g., the relative importance of cross track vs other types of perturbations) depends on storm size, landfall location, and other factors. For emergency managers and populations at risk, the results indicate that there are currently important practical limitations to our abilities to accurately predict location-specific inundation, especially at the lead times that are required to implement effective evacuations in many areas.

The findings also raise several important questions that can guide future research. How do these findings translate to other storms and landfall locations? How can we increase the predictability of storm surge given significant uncertainty in the landfall location at longer lead times? Is it practical to try and improve storm surge predictions at specific locations at longer lead times (>24 h) and if not, are there metrics, ensemble, or experimental designs other than those explored here that could provide more predictability? What are the predictability characteristics of surge when measured using additional metrics relevant for emergency management decisions, such as timing, depth, area covered, and along-coast and inland extent of surge? This study takes a step toward providing a framework for answering these and other important questions, by investigating and quantifying storm surge predictability across a range of lead times, from different perspectives.

Acknowledgments. Work presented here was sponsored by the NSF Hazards SEES Grant 1331490. We also thank Jamie Rhome from the National Hurricane Center for his insightful discussions and operational perspective, Jesse Feyen and Sergey Vinogradov from NOAA for providing data and helpful guidance, and Jason Fleming for technical assistance related to running ADCIRC. Some ADCIRC visualizations were produced with FigureGen (Dietrich et al. 2013). We would also like to acknowledge high-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, which is sponsored by the National Science Foundation.

REFERENCES

- Berg, R., 2014: Tropical cyclone report: Hurricane Ike (1–14 September 2008). NHC Tech. Rep. AL092008, 55 pp., http://www.nhc.noaa.gov/data/tcr/AL092008_Ike.pdf.
- Cangialosi, J. P., and J. L. Franklin, 2016: National Hurricane Center forecast verification report: 2015 hurricane season. NOAA/NHC Tech. Rep., 69 pp., http://www.nhc.noaa.gov/verification/pdfs/Verification_2015.pdf.
- , and C. W. Landsea, 2016: An examination of model and official National Hurricane Center tropical cyclone size forecasts. *Wea. Forecasting*, **31**, 1293–1300, <https://doi.org/10.1175/WAF-D-15-0158.1>.
- Colle, B. A., F. Buonaiuto, M. J. Bowman, R. E. Wilson, R. Flood, R. Hunter, A. Mintz, and D. Hill, 2008: New York City's vulnerability to coastal flooding: Storm surge modeling of past cyclones. *Bull. Amer. Meteor. Soc.*, **89**, 829–841, <https://doi.org/10.1175/2007BAMS2401.1>.
- , and Coauthors, 2015: Exploring water level sensitivity for metropolitan New York during Sandy (2012) using ensemble storm surge simulations. *J. Mar. Sci. Eng.*, **3**, 428–443, <https://doi.org/10.3390/jmse3020428>.
- Collins, M., and Coauthors, 2006: Interannual to decadal climate predictability in the North Atlantic: A multimodel-ensemble study. *J. Climate*, **19**, 1195–1203, <https://doi.org/10.1175/JCLI3654.1>.
- Demuth, J. L., R. E. Morss, B. H. Morrow, and J. K. Lazo, 2012: Creation and communication of hurricane risk information. *Bull. Amer. Meteor. Soc.*, **93**, 1133–1145, <https://doi.org/10.1175/BAMS-D-11-00150.1>.
- Dietrich, J. C., and Coauthors, 2013: Real-time forecasting and visualization of hurricane waves and storm surge using SWAN+ADCIRC and FigureGen. *Computational Challenges in the Geosciences*, C. Dawson and M. Gerritsen, Eds., The IMA Volumes in Mathematics and its Applications Series, Vol. 156, Springer, 49–70, doi:10.1007/978-1-4614-7434-0_3.
- Di Liberto, T., B. A. Colle, N. Georgas, A. F. Blumberg, and A. A. Taylor, 2011: Verification of a multimodel storm surge ensemble around New York City and Long Island for the cool season. *Wea. Forecasting*, **26**, 922–939, <https://doi.org/10.1175/WAF-D-10-05055.1>.
- East, J. W., M. J. Turco, and R. R. Mason Jr., 2008: Monitoring inland storm surge and flooding from Hurricane Ike in Texas and Louisiana, September 2008. USGS Open-File Rep. 2008-1365, 38 pp., <http://pubs.usgs.gov/of/2008/1365/>.
- Ehrendorfer, M., 1997: Predicting the uncertainty of numerical weather forecasts: A review. *Meteor. Z.*, **6**, 147–183.
- Faivre, G., J. Burston, D. Ware, and R. Tomlinson, 2015: Queensland storm surge forecasting model design using sensitivity analysis. *21st Int. Congress on Modeling and Simulation*, Gold Coast, Australia, Modeling and Simulation Society of Australia and New Zealand, 1240–1246, <https://www.mssanz.org.au/modsim2015/F2/faivre.pdf>.
- Feyen, J., S. Vinogradov, T. G. Asher, J. Halgren, and Y. Funakoshi, 2015: Development of operational storm surge guidance to support total water predictions. *14th Int. Workshop on Wave Hindcasting and Forecasting/Fifth Coastal Hazard Symp.*, Key West, FL, JCOMM, X3, http://www.waveworkshop.org/14thWaves/Presentations/X3_Feyen_Waves_OpsSurgeGuidance_Nov2015-FINAL.pdf.
- Fleming, J. G., C. W. Fulcher, R. A. Luettich Jr., B. D. Estrade, G. D. Allen, and H. S. Winer, 2008: A real time storm surge forecasting system using ADCIRC. *Proc. 10th Int. Conf. on Estuarine and Coastal Modeling*, Newport, RI, ASCE, 373–392, [https://doi.org/10.1061/40990\(324\)48](https://doi.org/10.1061/40990(324)48).
- Georgas, N., and Coauthors, 2016: The Stevens Flood Advisory System: Operational H3E flood forecasts for the Greater New York/New Jersey Metropolitan region. *Int. J. Safety Sec. Eng.*, **6**, 648–662, <https://doi.org/10.2495/SAFE-V6-N3-648-662>.
- Holland, G. J., 1980: An analytical model of the wind and pressure profiles in hurricanes. *Mon. Wea. Rev.*, **108**, 1212–1218, [https://doi.org/10.1175/1520-0493\(1980\)108<1212:AAMOTW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1980)108<1212:AAMOTW>2.0.CO;2).
- Hope, M. E., and Coauthors, 2013: Hindcast and validation of Hurricane Ike (2008) waves, forerunner, and storm surge. *J. Geophys. Res. Oceans*, **118**, 4424–4460, <https://doi.org/10.1002/jgrc.20314>.
- Houtekamer, P. L., and J. Derome, 1994: Prediction experiments with two-member ensembles. *Mon. Wea. Rev.*, **122**, 2179–2191, [https://doi.org/10.1175/1520-0493\(1994\)122<2179:PEWTME>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<2179:PEWTME>2.0.CO;2).
- Irish, J. L., D. T. Resio, and J. J. Ratcliff, 2008: The influence of storm size on hurricane surge. *J. Phys. Oceanogr.*, **38**, 2003–2013, <https://doi.org/10.1175/2008JPO3727.1>.
- Jones, S., and B. Golding, 2014: HIWeather: A research activity on High Impact Weather within the World Weather Research Programme. WMO Rep., 87 pp., http://www.wmo.int/pages/prog/arep/wwrp/new/documents/HiW_IP_v1_4.pdf.
- Kennedy, A. B., and Coauthors, 2011: Origin of the Hurricane Ike forerunner surge. *Geophys. Res. Lett.*, **38**, L08608, <https://doi.org/10.1029/2011GL047090>.
- Kerr, P. C., and Coauthors, 2013: U.S. IOOS coastal and ocean modeling testbed: Evaluation of tide, wave, and hurricane surge response sensitivities to mesh resolution and friction in the Gulf of Mexico. *J. Geophys. Res. Oceans*, **118**, 4633–4661, <https://doi.org/10.1002/jgrc.20305>.
- Knaff, J. A., and C. R. Sampson, 2015: After a decade are Atlantic tropical cyclone gale force wind radii forecasts now skillful? *Wea. Forecasting*, **30**, 702–709, <https://doi.org/10.1175/WAF-D-14-00149.1>.
- , —, M. DeMaria, T. P. Marchok, J. M. Gross, and C. J. McAdie, 2007: Statistical tropical cyclone wind radii prediction using climatology and persistence. *Wea. Forecasting*, **22**, 781–791, <https://doi.org/10.1175/WAF1026.1>.
- , C. J. Slocum, K. D. Musgrave, C. R. Sampson, and B. R. Strahl, 2016: Using routinely available information to estimate tropical cyclone wind structure. *Mon. Wea. Rev.*, **144**, 1233–1247, <https://doi.org/10.1175/MWR-D-15-0267.1>.
- Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, <https://doi.org/10.1175/MWR-D-12-00254.1>.
- Lindell, M., C. Prater, and W. Peacock, 2007: Organizational communication and decision making for hurricane emergencies. *Nat. Hazards Rev.*, **8**, 50–60, [https://doi.org/10.1061/\(ASCE\)1527-6988\(2007\)8:3\(50\)](https://doi.org/10.1061/(ASCE)1527-6988(2007)8:3(50)).
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2).
- , 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–333, <https://doi.org/10.3402/tellusa.v17i3.9076>.
- , 1982: Atmospheric predictability experiments with a large numerical model. *Tellus*, **34**, 505–513, <https://doi.org/10.3402/tellusa.v34i6.10836>.
- Luettich, R. A., and J. J. Westerink, 2004: Formulation and numerical implementations of the 2D/3D ADCIRC finite

- element model version 44.XX. University of North Carolina at Chapel Hill Tech. Rep., 74 pp., http://www.unc.edu/ims/adcirc/adcirc_theory_2004_12_08.pdf.
- , —, and N. W. Scheffner, 1992: ADCIRC: An advanced three-dimensional circulation model for shelves, coasts and estuaries. Report 1: Theory and methodology of ADCIRC-2DDI and ADCIRC-3DL. U.S. Army Corps of Engineers Tech Rep. DRP-92-6, 137 pp.
- Mattocks, C., and C. Forbes, 2008: A real-time, event-triggered storm surge forecasting system for the state of North Carolina. *Ocean Modell.*, **25**, 95–119, <https://doi.org/10.1016/j.ocemod.2008.06.008>.
- , —, and L. Ran, 2006: Design and implementation of a real-time storm surge and flood forecasting capability for the State of North Carolina. UNC-CEP Tech. Rep., 103 pp., http://www.unc.edu/ims/adcirc/publications/2006/2006_Mattocks.pdf.
- Morrow, B. H., J. K. Lazo, J. Rhome, and J. Feyen, 2015: Improving storm surge risk communication: Stakeholder perspectives. *Bull. Amer. Meteor. Soc.*, **96**, 35–48, <https://doi.org/10.1175/BAMS-D-13-00197.1>.
- Morss, R. E., K. A. Emanuel, and C. Snyder, 2001: Idealized adaptive observation strategies for improving numerical weather prediction. *J. Atmos. Sci.*, **58**, 210–232, [https://doi.org/10.1175/1520-0469\(2001\)058<0210:IAOSFI>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<0210:IAOSFI>2.0.CO;2).
- , and Coauthors, 2017: Hazardous weather prediction and communication in the modern information environment. *Bull. Amer. Meteor. Soc.*, <https://doi.org/10.1175/BAMS-D-16-0058.1>, in press.
- NHC, 2017: Storm surge product decision support wedge. NOAA/NHC, accessed 23 September 2017, <http://www.nhc.noaa.gov/surge/products.php>.
- NOAA, 2011: NWS Central Region service assessment: Joplin, Missouri, tornado—May 22, 2011. NOAA/NWS Tech Rep., 41 pp., https://www.weather.gov/media/publications/assessments/Joplin_tornado.pdf.
- , 2013: Defining storm surge, storm tide and inundation. NOAA/NWS Tech Rep., 2 pp., http://www.nhc.noaa.gov/news/20130806_pa_defineSurge.pdf.
- NOAA/NOS/CO-OPS, 2016: Tides and currents. NOAA, accessed 20 October 2016, <http://tidesandcurrents.noaa.gov>.
- NRC, 2010: *When Weather Matters: Science and Service to Meet Critical Societal Needs*. National Academies Press, 198 pp.
- Pasch, R. J., D. P. Brown, and E. S. Blake, 2011: Tropical cyclone report: Hurricane Charley (9–14 August 2004). NHC Tech. Rep. AL032004, 23 pp., http://www.nhc.noaa.gov/data/tcr/AL032004_Charley.pdf.
- Rego, J. L., and C. Li, 2009: On the importance of the forward speed of hurricanes in storm surge forecasting: A numerical study. *Geophys. Res. Lett.*, **36**, L07609, <https://doi.org/10.1029/2008GL036953>.
- Riverside Technology and AECOM, 2015: Mesh development, tidal validation, and hindcast skill assessment of an ADCIRC model for the hurricane storm surge operational forecast system on the US Gulf-Atlantic Coast. NOAA/NOS Tech. Rep., 180 pp., https://www.earthsystemcog.org/site_media/projects/umac_model_advisory/HSSOFS_Development_Evaluation_20150410.pdf.
- Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (version 3.2). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240, [https://doi.org/10.1175/1520-0477\(2000\)081<1231:TATCFS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<1231:TATCFS>2.3.CO;2).
- , and J. A. Knaff, 2015: A consensus forecast for tropical cyclone gale wind radii. *Wea. Forecasting*, **30**, 1397–1403, <https://doi.org/10.1175/WAF-D-15-0009.1>.
- Sebastian, A. G., J. M. Proft, J. C. Dietrich, W. Du, P. B. Bedient, and C. N. Dawson, 2014: Characterizing storm surge behavior in Galveston Bay using the SWAN + ADCIRC model. *Coastal Eng.*, **88**, 171–181, <https://doi.org/10.1016/j.coastaleng.2014.03.002>.
- Taylor, A., and B. Glahn, 2008: Probabilistic guidance for hurricane storm surge. *19th Conf. on Probability and Statistics*, New Orleans, LA, Amer. Meteor. Soc., 7.4, https://ams.confex.com/ams/88Annual/techprogram/paper_132793.htm.
- Walser, A., D. Lüthi, and C. Schär, 2004: Predictability of precipitation in a cloud-resolving model. *Mon. Wea. Rev.*, **132**, 560–577, [https://doi.org/10.1175/1520-0493\(2004\)132<0560:POPIAC>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0560:POPIAC>2.0.CO;2).
- Wang, R., M. Manausa, and J. Cheng, 2005: Hurricane Charley characteristics and storm tide evaluation. Beaches and Shores Resource Center, Institute of Science and Public Affairs, Florida State University Rep., 15 pp., <https://softlive.dep.state.fl.us/sites/default/files/HurricaneCharleyCharacteristics-StormTideEvaluation.pdf>.
- Weisberg, R. H., and L. Zheng, 2006a: Hurricane storm surge simulations for Tampa Bay. *Estuaries Coasts*, **29**, 899–913, <https://doi.org/10.1007/BF02798649>.
- , and —, 2006b: A simulation of the Hurricane Charley storm surge and its breach of North Captiva Island. *Fla. Sci.*, **69**, 152–165, <http://ocgweb.marine.usf.edu/Products/StormSurge/CharleyFLS1-10-06.pdf>.
- Wolshon, B., E. Urbina, C. Wilmot, and M. Levitan, 2005: Review of policies and practices for hurricane evacuation. I: Transportation planning, preparedness, and response. *Nat. Hazards Rev.*, **6**, 129–142, [https://doi.org/10.1061/\(ASCE\)1527-6988\(2005\)6:3\(129\)](https://doi.org/10.1061/(ASCE)1527-6988(2005)6:3(129)).
- Zhong, L., M. Li, and D.-L. Zhang, 2010: How do uncertainties in hurricane model forecasts affect storm surge predictions in a semi-enclosed bay? *Estuarine Coastal Shelf Sci.*, **90**, 61–72, <https://doi.org/10.1016/j.ecss.2010.07.001>.