Real-Time Simulated Storm Surge Predictions during Hurricane Michael (2018)

M. V. BILSKIE,[a] T. G. ASHER,[b] P. W. MILLER,[c] J. G. FLEMING,[d] S. C. HAGEN,[e,f,g] AND R. A. LUETTICH JR.[b]

[a] *School of Environmental, Civil, Agricultural, and Mechanical Engineering, The University of Georgia, Athens, Georgia*
[b] *Department of Marine Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina*
[c] *College of the Coast and Environment, Louisiana State University, Baton Rouge, Louisiana*
[d] *Scimaritan, Baton Rouge, Louisiana*
[e] *Center for Coastal Resiliency, Louisiana State University, Baton Rouge, Louisiana*
[f] *Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, Louisiana*
[g] *Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana*

(Manuscript received 31 August 2021, in final form 22 March 2022)

ABSTRACT: Storm surge caused by tropical cyclones can cause overland flooding and lead to loss of life while damaging homes, businesses, and critical infrastructure. In 2018, Hurricane Michael made landfall near Mexico Beach, Florida, on 10 October with peak wind speeds near 71.9 m s^{-1} (161 mph) and storm surge over 4.5 m NAVD88. During Hurricane Michael, water levels and waves were predicted near–real time using a deterministic, depth-averaged, high-resolution ADCIRC+SWAN model of the northern Gulf of Mexico. The model was forced with an asymmetrical parametric vortex model [generalized asymmetric Holland model (GAHM)] based on Michael's National Hurricane Center (NHC) forecast track and strength. The authors report errors between simulated and observed water level time series, peak water level, and timing of peak for NHC advisories. Forecasts of water levels were within 0.5 m of observations, and the timing of peak water levels was within 1 h as early as 48 h before Michael's eventual landfall. We also examined the effect of adding far-field meteorology in our TC vortex model for use in real-time forecasts. In general, we found that including far-field meteorology by blending the TC vortex with a basin-scale NWP product improved water level forecasts. However, we note that divergence between the NHC forecast track and the forecast track of the meteorological model supplying the far-field winds represents a potential limitation to operationalizing a blended wind field surge product. The approaches and data reported herein provide a transparent assessment of water level forecasts during Hurricane Michael and highlight potential future improvements for more accurate predictions.

KEYWORDS: Forecast pdfs/skill; Hindcasts; Operational forecasting; Model errors; Model evaluation/performance; Ocean models

1. Introduction

Storm surge is the abnormal rise of water level from high winds and reduced atmospheric pressure during meteorological events such as tropical cyclones and typhoons. Storm surge and high waves can lead to loss of life while damaging homes, businesses, and critical infrastructure (Anarde et al. 2018; Rappaport 2000; Reed et al. 2010; Resio and Westerink 2008). Since the 2005 Atlantic hurricane season, substantial progress has been made in developing more accurate numerical models to predict storm surge, wind waves, and overland coastal flooding. Such progress has included advancing mesh generation techniques, enhancing the accuracy of topography/bathymetry in storm surge models, and wind-wave coupling, among other advancements (Bilskie et al. 2020; Bilskie et al. 2016a; Bunya et al. 2010; Dietrich et al. 2011a; Morey et al. 2006; Mori et al. 2014; Weisberg and Zheng 2006). These advancements have made their way into improved storm surge forecasts and permitted the dissemination of forecast results to coastal emergency managers (and other stakeholder groups) during an impending tropical cyclone (TC). The availability of accurate forecasts aids planning and preparation for flooding to mitigate property destruction and loss of life. Therefore, providing accurate and timely storm surge forecasts along with proper communication of model limitations and potential errors allows stakeholders to make informed decisions (DeLorme et al. 2020; Hamill et al. 2012; Morrow et al. 2015; Munroe et al. 2018; Ramos et al. 2010; Wolshon et al. 2005).

Computing accurate water levels in real-time is a challenge. First, simulations must be performed with results generated and processed quicker than actual time (or else the forecast is no longer valid). Second, the water level response is reliant on accurate TC forecasts (i.e., the storm's track and wind field structure) and regional geometric configurations (i.e., shoreline, bathymetry, and topography) (Bilskie et al. 2016a; Mori et al. 2014; Rappaport et al. 2009; Resio and Westerink 2008). Wind speed magnitude and the location over which the wind is blowing is arguably the most important component in the generation of storm surge. Water levels can become magnified in shallow water since the wind stress term in the governing equations is divided by water depth and increases with decreasing depth (Pugh 1996). In most cases, geometric configurations (e.g., shoreline, bathymetry, topography) are well known, and the challenge resides in accurately forecasting the TC's track, landfall location, and intensity (Torn and Snyder 2012).

Hagen: Deceased.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Matthew V. Bilskie, mbilskie@uga.edu

DOI: 10.1175/WAF-D-21-0132.1

Probabilistic storm surge prediction methods have been developed to alleviate uncertainties associated with TC forecasts and uncertainty propagation into water level predictions (Cloke and Pappenberger 2009; Dale et al. 2014; Davis et al. 2010; Flowerdew et al. 2010; Plumlee et al. 2021), such as the National Hurricane Center's (NHC) P-Surge (NHC 2019). Probabilistic methods are practical but require a large number of simulations (tens to thousands), which necessarily leads to compromises in model fidelity (e.g., choices of resolution and physics). Additionally, probabilistic guidance products require careful messaging and end-user training because they are intrinsically nonphysical (Cloke and Pappenberger 2009; Dale et al. 2014; Dietrich et al. 2018; Kerr et al. 2013; Resio et al. 2017).

An alternative approach to probabilistic guidance is deterministic, scenario-based forecasting wherein a small number of high-fidelity simulations are conducted in real-time for a given storm event (e.g., Rey and Mulligan 2021). Because these simulations are high-fidelity, they provide emergency managers with more detailed (e.g., fine-scale) surge guidance. Significant challenges in employing high-fidelity deterministic simulations are increased run times due to additional physics invoked in the computations and higher-resolution model grids (or unstructured meshes), increasing the degrees of freedom (Kerr et al. 2013). Broader availability of high-performance computers (HPC) and improvements in hardware and software have made it realistic to employ high-fidelity storm surge forecast models in real time (Morales-Hernández et al. 2020; Tanaka et al. 2011). Furthermore, recent efforts have been placed on developing improved mesh discretization practices for surge forecasting models (Bilskie et al. 2020), subgrid correction schemes (Kennedy et al. 2019), dynamic load balancing (Roberts et al. 2021), and flood prediction downscaling (Rucker et al. 2021).

It can be argued that the use of deterministic storm surge guidance becomes more relevant for emergency response and planning as TC forecasts improve. Therefore, it is important to examine how errors in TC forecasts propagate into storm surge forecasts, specifically the magnitudes and arrival of peak water levels. In this study, we focus on the evolution of Hurricane Michael (2018) forecast advisories issued by the NHC and the resulting predicted water levels using a real-time scenario-based modeling framework. Comparisons were made between simulated and observed water levels for NHC forecast advisories as well as for a hindcast simulation using the post-storm NHC operational best track. We also examine the effect of including far-field meteorology in our tropical cyclone–only wind model for use in real-time forecasts. We discuss the forecast water level and their errors with respect to time before landfall.

2. Hurricane Michael meteorological forecasts

Hurricane Michael was the third-most intense Atlantic-basin hurricane (with a central pressure of 919 hPa) to make landfall along the contiguous U.S. coast on record, and the most intense in terms of maximum sustained (1-min) wind speed since Hurricane Andrew (1992). Michael made landfall near Mexico Beach, Florida, at 1700 UTC 10 October 2018 with peak winds near 259 km h^{-1} (161 mph). Michael generated storm surges of 4.7 m and 2.6 m near Mexico Beach and Port St. Joe, respectively. The NHC began issuing forecasts for eventual Hurricane Michael at 2100 UTC 6 October 2018, 68 h before eventual landfall (Fig. 1 and Table 1). Early forecast advisories showed Michael's landfall east of Mexico Beach and closer to Apalachicola with landfall location errors ranging from 42 to 81 km (advisories 5–8). The landfall track error for these early forecasts (68–50 h prior to landfall) was well within the 5-yr (2012–16) mean official forecast error of 363–47 km at forecast periods of 120–12 h, respectively (Fig. 2a) (Cangialosi 2018). Landfall error fell to within 20 km by advisory 10 (38 h prior to landfall). The 5-yr mean official forecast error is 99 km at for 36-h forecast period (Cangialosi 2018).

Advisories consistently underestimated Michael's strength (Figs. 3a and 2b) with the NHC even increasing Michael's strength to category 5 based on a detailed post-storm analysis (Beven et al. 2019). Initial forecasts placed the maximum sustained wind speed 43.7 m s^{-1} at landfall, the last advisories before landfall (15 and 16) predicted 64.3 m s^{-1}, and the operational best track estimated 69.5 m s^{-1}. Michael underwent multiple cycles of rapid intensification in the four days prior to landfall, with wind speeds increasing 20.6 m s^{-1} and central pressures deepening 42 hPa in the 24-h period preceding landfall. Rapid intensification, defined by the NHC as an increase in maximum sustained wind speeds of 15.4 m s^{-1} in 24 h, is difficult to predict (Gall et al. 2013; Kaplan et al. 2010). Complicating the forecast, the rapid intensification associated with Hurricane Michael occurred in a moderate- to high-shear environment exceeding 10.3 m s^{-1} (Hazelton et al. 2020) as well as a sector of the Gulf of Mexico where no rapid intensification had been observed historically (Rappaport et al. 2010; Rhome et al. 2006).

Intensity errors nearly tripled the 5-yr average error (2012–16) (Fig. 2b). A sharp contrast to the narrow track prediction errors, this juxtaposition mirrors TC forecast improvement over the last several decades, whereby intensity forecast accuracy has lagged behind track accuracy (DeMaria et al. 2014). The operational best track shows the storm intensifying through landfall, an unusual trait for a major TC, but one that is mirrored by nearly all advisories and has been theorized to occur more frequently in a future climate regime (Emanuel 2017).

3. Data and methods

a. Water level data

The 14 stations are a collection of NOAA tide gauges and storm tide sensors deployed by the USGS before the storm (shown in Fig. 5) (Byrne 2019). The NOAA tide gauges are located in always-wet areas near the open coast; however, the rapidly deployed USGS storm tide sensors were deployed on normally dry land. Since they are on normally dry land, water level data are only collected when the adjacent area becomes inundated, which may occur for a short time (i.e., less than 24 h). The rapidly deployed USGS sensors collected data via a Hobo U20 pressure transducer with sampling interval of 30 s. The raw

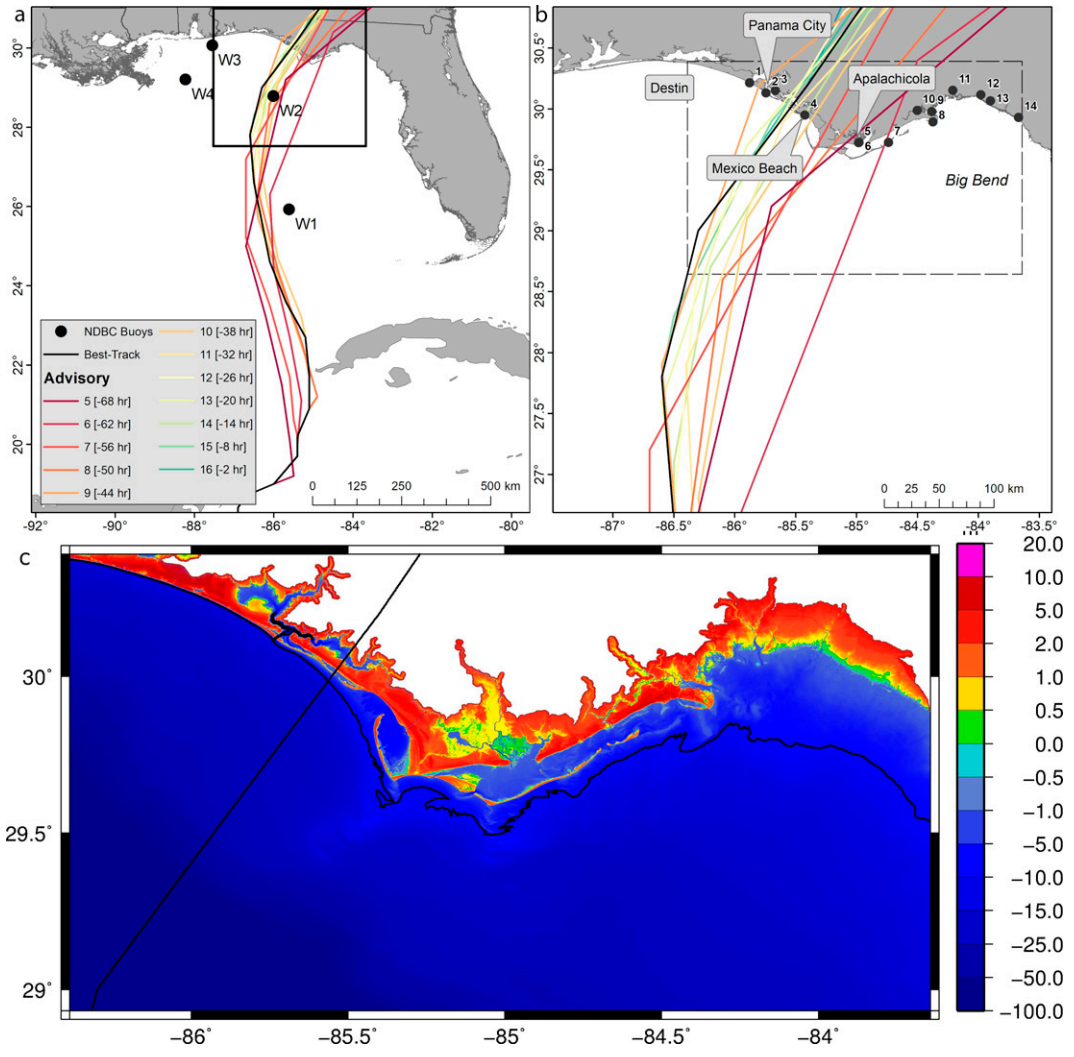

FIG. 1. (a) NHC forecast tracks for advisories 5–16 and operational best track for Hurricane Michael. Advisories are shown by NHC forecast advisory number and time from eventual landfall. NDBC wave buoys are labeled as W1 (42003), W2 (42039), W3 (42012), and W4 (42040). (b) Zoom in of forecast and operational best tracks within the study domain outlined by the black inset box of (a). Locations of observations are shown from 1 to 14. Refer to Fig. 5 and Table 1 for details on the observation stations. (c) Bathymetry and topography as represented by the real-time ACDIRC+SWAN northern Gulf of Mexico mesh (m, NAVD88) for the dashed box of (b). The operational best track of Hurricane Michael is shown in black as well as the 10-m bathymetric contour.

water level data were smoothed using a low-pass filter (see the metadata for additional details on processing done by USGS to obtain a filtered water surface elevation; https://stn.wim.usgs.gov/STNServices/Files/105420/Item).

b. *Storm surge and wave models*

Numerical simulations were carried out using a two-dimensional, depth-averaged, tightly coupled ADCIRC+SWAN model of the northern Gulf of Mexico (Fig. 1c) (Bilskie et al. 2016a; Bunya et al. 2010; Dietrich et al. 2011a; Luettich and Westerink 2004). ADCIRC solves the depth-integrated momentum and the generalized wave continuity equations for depth-averaged currents and water levels (Dawson et al. 2006; Kinnmark 1986; Kolar et al. 1994; Luettich and Westerink 2004; Westerink et al. 2008). ADCIRC utilizes an unstructured finite element mesh, which permits local mesh refinements in regions of interest or large solution gradients. The model simulations performed here use a 1-s time step, include nonlinear advection terms, wetting and drying is activated with a minimum depth of 0.1 m, and a Smagorinsky-type turbulence closure (value of 0.20). Spatially varying roughness parameters that include Manning's n bottom friction, wind reduction, and canopy coverage are derived from the 2010 Coastal Change Analysis Program (CCAP) land use land cover data (Bilskie et al. 2015; Bilskie et al. 2016b; Dietrich et al. 2011b; National Oceanic and Atmospheric Administration 2010).

TABLE 1. Details for NHC advisories 5–17, and the hindcast (H—also advisory 18) and estimated landfall wind speed (m s^{-1}), forward speed (m s^{-1}) at landfall, landfall error (km), and landfall timing error (h). The latitude and longitude coordinates are at the forecasted mainland landfall. Positive landfall error indicates a forecast of early arrival and negative error indicates a late arrival. Timing errors were based on the NHC forecast tracks and a linear interpolation of forward speed and timing along the track.

					Values are at forecasted landfall			
NHC advisory	Issued	H (h)	Landfall latitude (°)	Landfall longitude (°)	Wind speed (m s^{-1})	Forward speed (m s^{-1})	Landfall error (km)	Timing error (h)
5	2100 UTC 7 Oct 2018	−68	−85.5	30.0	43.7	5.7	54.4	4
6	0300 UTC 8 Oct 2018	−62	−85.1	29.7	43.7	5.1	81.7	4
7	0900 UTC 8 Oct 2018	−56	−84.7	29.8	48.9	5.1	42.0	4
8	1500 UTC 8 Oct 2018	−50	−85.3	29.7	54.0	5.7	47.6	1
9	2100 UTC 8 Oct 2018	−44	−85.2	29.7	51.4	6.2	−31.6	1
10	0300 UTC 9 Oct 2018	−38	−85.8	30.2	54.0	6.2	19.9	−1
11	0900 UTC 9 Oct 2018	−32	−85.4	29.9	54.0	5.7	3.3	1
12	1500 UTC 9 Oct 2018	−26	−85.5	30.0	56.6	5.7	−2.8	2
13	2100 UTC 9 Oct 2018	−20	−85.6	30.0	56.6	5.7	−7.2	3
14	0300 UTC 10 Oct 2018	−14	−85.6	30.1	59.2	5.7	13.1	2
15	0900 UTC 10 Oct 2018	−8	−85.4	29.9	64.3	6.2	−4.9	2
16	1500 UTC 10 Oct 2018	−2	−85.6	30.0	64.3	5.1	−3.4	1
Landfall	1700 UTC 10 Oct 2018	—	—	—	—	—	—	
17	2100 UTC 10 Oct 2018	4	−85.5	30.0	64.3	6.2	0.0	
H	0300 UTC 11 Oct 2018	10	−85.5	30.0	64.3	6.2	0.0	

SWAN, governed by the wave action balance equation, represents the wave field as a phase-averaged spectrum. Wave frequencies were discretized into 40 bins of relative frequencies on a logarithmic scale from 0.031 to 1.42 Hz and directions into 36 bins of 10°. Source terms within SWAN include wind-induced wave growth (Cavaleri and Rizzoli 1981; Komen et al. 1984), energy loss to white-capping (Rogers et al. 2003), depth-induced breaking (index of 0.73) (Battjes and Janssen 1978), and bottom friction (Dietrich et al. 2011a; Madsen et al. 1988). Nonlinear triad interactions are not included. SWAN is tightly coupled to ADCIRC by using the same unstructured mesh as ADCIRC, they share the same computing infrastructure, and are run sequentially in time (pass information every 600 s—also the SWAN time step) (Dietrich et al. 2011a; Zijlema 2010). Although SWAN is employed in this study, the focus is not on the real-time prediction of waves. This is difficult to assess as there were only a few wave observations during Michael and no wave observations in the nearshore region Even so, we include a wave model within the forecasting framework. The contribution of waves have been found to increase water levels by 15% or greater in coastal regions (Bertin et al. 2015; Funakoshi et al. 2008).

The northern Gulf of Mexico real time (NGOM-RT) unstructured finite element mesh was specifically developed for real-time coastal flood forecasting (Bilskie et al. 2020). The mesh spans the western North Atlantic Ocean up to the 60°W meridian and includes high resolution (20–100 m) along the shoreline and into the coastal floodplain (i.e., normally dry land) of Mississippi, Alabama, and the Florida Panhandle. ADCIRC+SWAN simulations using the NGOM-RT model setup have been extensively validated with model results comparing well to observations for Hurricanes Ivan, Dennis, Katrina, and Isaac. A comparison with peak water levels yields an R^2 of 0.96 and best-fit slope of 0.97 (Fig. S3in Bilskie et al. 2020, 2016a,b).

c. *Parametric tropical cyclone representation*

Tropical cyclone surface wind and pressure fields exhibit large gradients near the center of the storm, with winds going from near zero in the eye of the storm to their maximum within as little as 10 km. Global NWP models lack the resolution to capture these gradients, and so parametric models of storm wind and pressure fields are often employed. Since meteorological forcing far from the storm center can have a sizeable influence on water levels, far-field input can be supplied from another meteorological model (Asher et al. 2019; Morey et al. 2006). The generalized asymmetric Holland model (GAHM) developed by Gao (2018) was used in this study to transform NHC tropical cyclone best track and forecasts into a parametric wind field. GAHM takes advantage of recent improvements that include quadrant-varying storm asymmetry and removing the assumption of cyclostrophic balance when $r = R_{max}$ (Hu et al. 2012; Xie et al. 2006) while reintroducing two scaling parameters, in the original Holland formulation that satisfy $V = V_{max}$ and $dV/dr = 0$ at $r = R_{max}$. Details regarding GAHM can be found in Dietrich et al. (2018) and Gao (2018). The GAHM algorithm has been implemented within ADCIRC so the wind and pressure fields are internally computed for each computational point within the unstructured mesh (Dietrich et al. 2018). Figure 4a presents an example of the GAHM-generated wind field for Hurricane Michael at 0000 UTC 10 October based on NHC advisory 7.

d. *Basin-scale meteorology*

The GAHM wind model only accounts for the tropical cyclone and does not include the effects of far-field winds. To include winds away from the tropical cyclone, we used the North American Mesoscale Forecast System (NAM) with a grid spacing of 12 km. While the 12-km grid spacing is

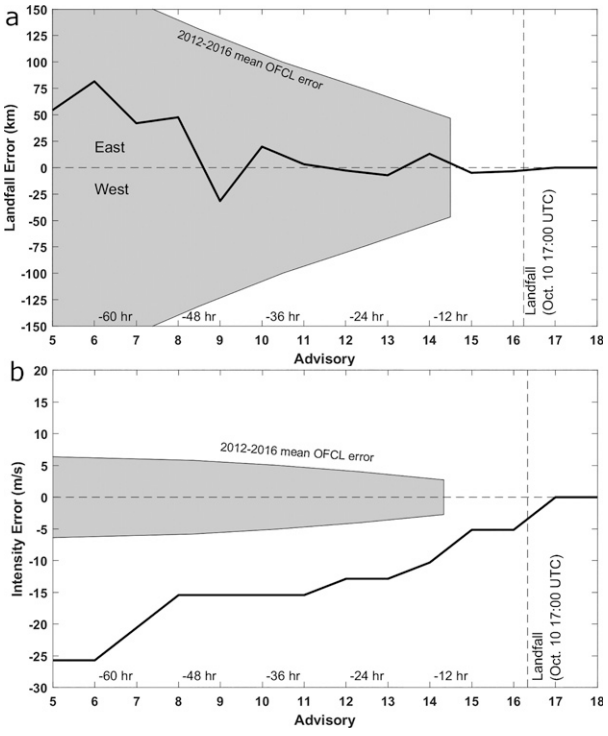

FIG. 2. (a) Landfall error (km) and (b) intensity error (m s^{-1}) for advisory 5–18. Negative intensity indicates an underprediction. Time before landfall is shown every 12 h and their related advisory. The shaded gray regions correspond to the 2012–16 mean NHC forecast errors (Cangialosi 2018). Although the NHC forecast errors are calculated as a function of forecast period (i.e., hours into future the forecast is targeting) rather than observed lead time (i.e., hours before observed landfall), we equate the two time scales for interpretability (e.g., assume that the advisory issued ~24 h prior to landfall also forecast landfall in ~24 h).

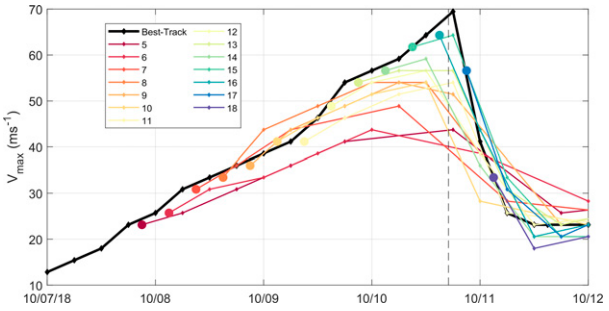

FIG. 3. NHC forecast maximum wind speed (m s^{-1}) for each advisory and operational best track hindcast. The circles indicate the start of each NHC advisory forecast, and the dashed line indicates landfall.

appropriate for far-field winds, violent winds within the TC eyewall occur on scales less than 12 km (e.g., Wu et al. 2019), and the GAHM wind model must be superimposed for regions nearest the TC. While the NAM is not typically consulted for operational TC forecasting guidance, its combination of high spatial resolution and swift completion time make it a desirable source for far-field meteorology. Historically, compared to its global counterpart, the Global Forecast System (GFS), NAM's horizontal grid spacing is finer and its 84-h forecast is available ~75 min faster. While recent upgrades to the GFS have improved its spatial resolution (now 13 km) to compete with the 12-km NAM, the NAM's faster completion time remains advantageous in a real-time forecasting setting. The fine spatial resolution is also important for accurately depicting the land–sea mask, which can strongly influence modeled wind speeds near the coast via changes in the landscape's surface roughness.

The role of far-field (i.e., background) wind and pressure on a hindcast of Hurricane Michael water levels to improve forecast results was examined. Surface pressure and 10-m zonal and meridional wind speed were extracted from the NAM analysis fields (i.e., the 0-h forecast) for each 6-hourly initialization coinciding with the hydrodynamic simulations and interpolated to a 12-km structured grid for use in the ADCIRC+SWAN simulations (see Fig. 4b for the NAM forecast at 0000 UTC 10 October).

During run time, ADCIRC seams the GAHM and NAM products. GAHM is used for the tropical cyclone vortex, and NAM is blended in over several TC radii for a smooth transition between the TC vortex and far-field wind and pressure. Herein, this meteorological product is referred to as the "blended" or "GAHM+NAM" winds. The values of d_{vortex} and $d_{\text{background}}$ were set to 3 and 12, respectively (Fig. 4c), meaning that GAHM winds were blended with NAM winds in the region lying between 3 and 12 times the radius of maximum winds from the storm center. Other values were tested, but there were little to no differences in the simulated water levels. Please reference the appendix for details regarding the wind blending approach.

e. *Model forcing and simulations*

All ADCIRC+SWAN simulations were initiated (cold start) at 1800 UTC 22 September and forced by astronomic tides and 6-hourly wind and pressure from NAM for 14 days until 1800 UTC 6 October. The 14-day period is to ramp the model forcing and response up to a dynamic steady state using a hyperbolic tangent ramp function. Astronomic tides were supplied as body forcing, as well as along the open ocean boundary located along the 60°W meridian and obtained through the Oregon State TPXO7.2 tidal atlas (Egbert and Erofeeva 2002; Egbert et al. 1994). Hurricane Michael surge forecasts (advisory 5–18) were hot-started on 1800 UTC 6 October with the available NHC best track information until 2100 UTC 7 October (advisory 5) where the forecast simulations began (68 h prior to eventual landfall). A new model simulation was conducted for each forecast advisory, released by the NHC every 6 h, using the available operational best track information and 5-day NHC advisory forecast for the forecast time period (the operational best track information is used since this was the data product available during real-time). Simulations were originally conducted in real-time using the ADCIRC Surge Guidance System (ASGS) (Blanton et al. 2012; Fleming et al. 2007). The ASGS was run in real-time,

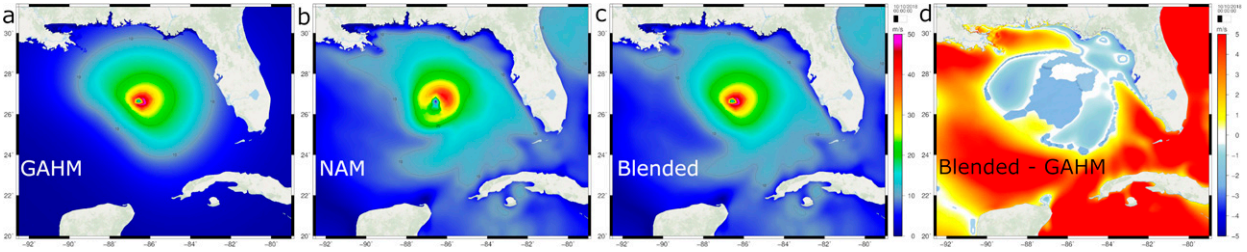

FIG. 4. The wind magnitude (m s^{-1}) at 0000 UTC 10 Oct 2018 generated by (a) GAHM, (b) NAM, (c) the blended product of GAHM+NAM, and (d) wind speed difference between GAHM+NAM and GAHM only (positive values indicate the addition of winds from NAM within the blended wind field). For (d), differences of 0 m s^{-1} are shown as transparent, and the background imagery is shown.

providing forecast guidance to coastal emergency managers during Hurricane Michael with model results available via the Coastal Emergency Risks Assessment online visualization platform (https://cera.coastalrisk.live/). For comparison purposes, the operational best track (10 h after landfall) simulation was used as our real-time hindcast (best available hindcast immediately post-landfall), which was the best track up to advisory 18.

4. Results

To summarize this section, section 4a communicates the results of 14 GAHM-only storm surge forecasts based on NHC advisories 5–18. The last of these advisories, advisory 18, also serves as the operational best track and is used to force the GAHM-only hindcast which is presented in section 4b. Last, a blended hindcast whereby the GAHM-only hindcast winds are superimposed upon background winds from the contemporaneous NAM analyses, is shown in section 4c. Though the GAHM-only hindcast and blended hindcast are plotted in graphics referenced in section 4a, they are discussed independently in sections 4b and 4c, respectively.

a. *Forecasts simulations*

1) TIME SERIES WATER LEVELS

Simulated water level errors were computed from observations at 14 gauge stations (Fig. 5) for each advisory. Errors were computed using root-mean-square error (RMSE) (m) of the time series data, peak water level difference (m), and peak water level timing error (h).

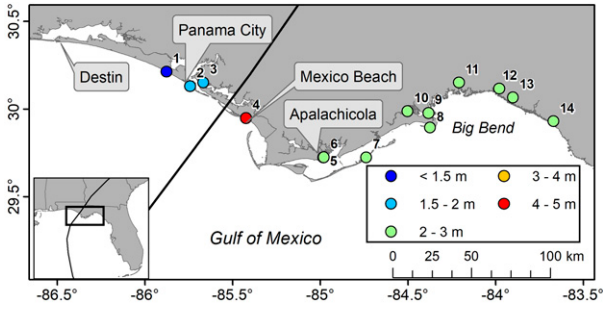

FIG. 5. Observation station locations. Station points are colored by the observed maximum water level (m, NAVD88). Refer to Table 2 for details regarding each station.

Average water level RMSE was below 1 m for the duration of the forecast advisories (solid line in Fig. 6). The RMSE was near 1 m for advisories 5–8 before reducing to 0.5 m for advisory 9 (42 h before landfall). The average RMSE was relatively constant at 0.5 m from advisory 9 through landfall (other than a slight increase for advisory 13) where the RMSE reduced to 0.37 m. NHC advisory 9 was substantially altered from the previous landfall location guidance for advisories 5–8. Advisories 5–8 had a landfall to the east of Mexico Beach with landfall errors of 54.4, 81.8, and 42.0 km, respectively (Table 1; Fig. 7). For NHC advisory 9, landfall shifted to the west with a track error of −31.6 km (negative value indicated a landfall error to the west). The forecast advisories continued to improve and landfall forecasts honed in around Mexico Beach. The changes in forecast track inherently improved the wind speed forecast between advisories 8 and 9 (Fig. 7), even though there was little change in forecasted wind speed at landfall (2.6 m s^{-1}; Table 1).

Figure 8 presents gauge-specific RMSE as it evolves with each updated forecast advisory. Figure 8a shows RSME and Fig. 8b includes RMSE normalized by the peak water level observed at each gauge station. Water level errors generally improved for all stations after advisory 8 (48 h before landfall). The observation gauge at Mexico Beach (gauge ID 4: FLBAY03283) resulted in the largest RMSE among all gauges with values ranging from 1.25 m to greater than 2.5 m. However, the water level forecast at Mexico Beach did generally well considering the observed peak stillwater elevation was 4.4 m NAVD88. This is, to some degree, expected since this gauge observed the highest surge, and peak surge is sensitive to storm inputs. When peak water level errors are normalized to the observed peak surge, errors at Mexico Beach are comparable to other locations, as shown in Fig. 8b.

The normalized RMSE for gauges 1–3 ranged from 20% to 30% for the duration of the storm. The normalized results for gauge 4 started high (>50%) for the initial set of advisories and reduced after advisory 8 with normalized errors 24 h prior to landfall declining to 30%–40%. The water level forecasts

TABLE 2. Observation station locations (see ID column) and data source URL. Refer to Fig. 5 for map of locations. NOAA water level data were obtained at https://tidesandcurrents.noaa.gov/ and USGS storm tide sensor data were obtained using the USGS Flood Event Viewer at https://stn.wim.usgs.gov/fev/#MichaelOct2018.

ID	Station site No.	Longitude (°)	Latitude (°)	Observed peak water level (m, NAVD88)
1	NOAA8729210	−85.8783	30.2133	1.47
2	FLBAY26247	−85.7437	30.1307	1.57
3	NOAA8729108	−85.6671	30.1512	1.86
4	FLBAY03283	−85.4246	29.9490	4.38
5	FLFRA03276	−84.9830	29.7232	2.51
6	NOAA8728690	−84.9800	29.7250	2.61
7	FLFRA26257	−84.7376	29.7242	2.36
8	FLFRA26263	−84.3736	29.8939	2.64
9	FLWAK03369	−84.3836	29.9774	2.44
10	CTMF1	−84.5025	29.9883	2.41
11	FLWAK03364	−84.2091	30.1518	2.76
12	FLTAY17325	−83.9795	30.1165	2.74
13	FLTAY03362	−83.9002	30.0665	2.67
14	FLTAY03359	−83.6702	29.9302	2.52

for gauges 5–14, located east of Michael's landfall, did well within 48 h of landfall. Several locations result in normalized RMSE as low as 10%–20% and relative errors as high as 30%–40% (not including advisory 14).

2) TIMING AND MAGNITUDE OF PEAK WATER LEVEL

A critical component of storm surge forecasting is an accurate prediction of peak water level and its timing. Figure 9 shows box plots of peak water level error and the error in the timing of peak water level for each forecast advisory. For each advisory, the box plot shows the peak water level (Fig. 9a) and timing of peak errors (Fig. 9b) measured for each observation gauge in Fig. 5 (n = 14). Errors in peak water level are large for advisories 5, 6, and 8 with median errors (horizontal line in each box) near 1 m. The median error reduces for advisory 9 to less than 0.5 m and remains relatively low through landfall. The spread of peak water level errors via the interquartile (IQ) range (bounds of each box) reduces from 1.28 m for advisory 11 to an IQ range of 0.28 with a median error of −0.32 m for advisory 12 (~36-h pre-landfall). The IQ range remains compact and the median error approaches zero from advisory 12 to 15, and reduced slightly for advisory 16. The error is slightly less for advisory 17 than the hindcast simulation. This is caused by differing wind speeds around Pensacola between advisory 17 and advisory 18 (hindcast) during the peak wind.

The error of peak water level timing is shown in Fig. 9b for each forecast advisory. Negative values indicate the forecast peak arrived after the observed peak. Positive values specify a forecast peak water level arriving earlier than the observed peak. The initial advisories of 5–7 yield timing errors greater less than −2 h, with median errors of −3.3, −2.1, and −3.0 h, respectively. The timing error is reduced starting at advisory 8 (48-h pre-landfall) with a median of −0.84 h and an IQ range of 0.92 h and stays within ±1–2 h until advisory 13 (24-h pre-landfall). After advisory 13, the median timing error continues to improve through landfall. The median peak timing error for advisory 16 (~6 h before landfall) was 0 h with an IQ range of −0.3 h. Overall, the timing of the peak surge was well captured as early as advisory 8, which was 48 h before landfall. In addition, advisories 5–7 depicted a reasonable forecast of peak water level timing given the uncertainty in Michael's forecast track and intensity (Fig. 2).

b. *GAHM hindcast*

An immediate post-storm operational hindcast was performed using the NHC best track information up to advisory 18 (10-h post-landfall). The storm track and strength information were input into GAHM which forced the ADCIRC+SWAN hindcast simulation. Advisory 18 had minimal track and intensity error (Figs. 2 and 3a), and at the time of Hurricane Michael (post-landfall), was the best and most immediately available product. The simulated maximum water level and significant wave height for Hurricane Michael is shown in Fig. 12.

Hindcast time series water levels were compared to observations at the 14 observations station (Fig. 10). Peak water levels were minimal (<2 m NAVD88) west of the storm track (NOAA8729210, FLBAY26247, and NOAA8729108). The largest water level of 4.4 m NAVD88 was observed near Mexico Beach, the nearest gauge to Michael's landfall location (FLBAY03282). Peak water levels of 2.6 m NAVD88 were observed near Apalachicola (FLFRA03276 and NOAA8728690). They remained high along the eastern edge of the storm track into Florida's Big Bend region (see Fig. 5) with peak water levels around 2.6 m NAVD88 (FLWAK03362, FLTAY17235, FLTAY03362, and FLTAY03359). Florida's Big Bend region has a larger tide range than the Florida Panhandle (see NOAA8729210 versus FLWAK03364), and surges are amplified

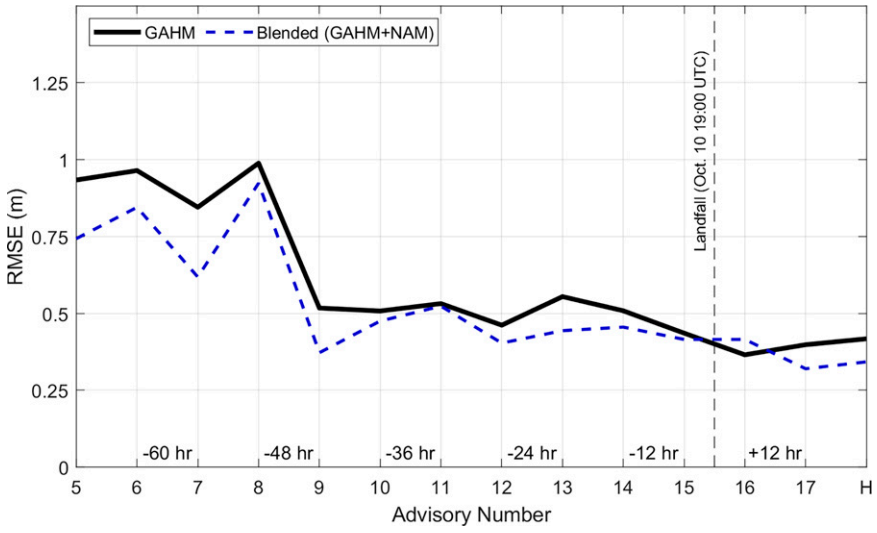

FIG. 6. Average root-mean-square error (RMSE; m) across all 14 gauge stations from advisory 5–17 and the hindcast (H, advisory 18) for the GAHM-driven simulation (blank line) and blended simulation (dashed blue line). Landfall is indicated by the vertical dashed black line.

due to Kelvin waves and funneling effects into the region (Bilskie et al. 2016a; Lin et al. 2014).

The simulated wind waves were also compared to available observations for significant wave height, mean wave direction, and peak wave period at four NDBC buoys (Fig. 11). All simulated wave values show good agreement across all stations. Unfortunately, station NDBC42039 (W2) failed during 10 October. The average RMSE across all 14 stations is reduced from 0.44 to 0.37 m by including wave radiation stresses generated by the SWAN model.

Next, we examine how blending the GAHM vortex winds (driven by the post-storm NHC best track information) with a large-scale wind field, such as NAM, may increase the accuracy of the simulated storm surge.

c. *Blended hindcast*

After performing the standard hindcast, the ADCIRC+SWAN simulation was repeated using the GAHM-generated wind field blended (using the NHC operational best track) to the NAM hindcast (see section 3d). The simulated time series water levels are shown in Fig. 10, time series of waves in Fig. 11, and maximum water levels and significant wave heights are shown in Fig. 12b. The simulated peak water levels are similar (within 0.1 m) between the GAHM-only and

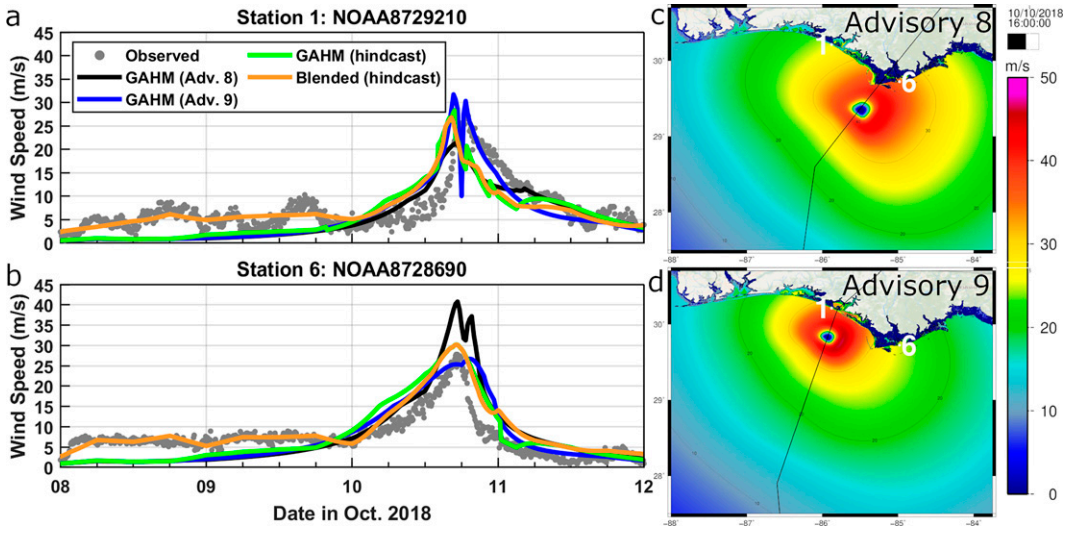

FIG. 7. Time series of wind speeds (m s^{-1}) based on observations (gray circles), GAHM advisory 8 (black line), GAHM advisory 9 (blue line), GAHM hindcast (green line), and blended GAHM+NAM hindcast (orange line) for 8–12 Oct 2018 for (a) NOAA8729210 and (b) NOAA8728690. Simulated wind speeds produced by GAHM for (c) advisory 8 and (d) advisory 9 for 1600 UTC 10 Oct. The forecasts track is shown as the solid black line.

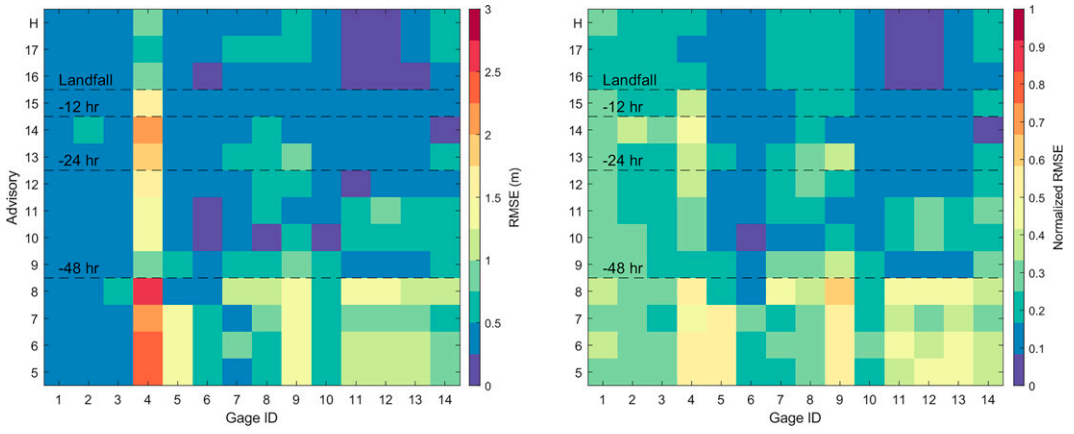

FIG. 8. (a) Root-mean-square error (RMSE) (m) and (b) RMSE normalized by the observed peak water level at each gauge across all 14 gauges (x axis) and advisories 5–17 and the hindcast (H, advisory 18).

blended hindcast simulations, particularly near landfall (Fig. 12c). Water levels are 0.20 m higher west of landfall and reach as high as 0.4 m to the east, near Florida's Big Bend (near station 9 and 10). Overall, the blended winds cause an increase (i.e., pre-storm set up) in peak water levels across the entire study region. The increase in the peak water level is caused by far-field (i.e., away from the tropical cyclone) winds rather than just winds associated with the tropical cyclone. Figures 4d–f shows the wind magnitude representation of Hurricane Michael at 0000 UTC 10 October 2018 as it approaches the Florida Panhandle. The blended product includes the far-field winds, which results in a more accurate representation of the basin-scale meteorological conditions than the tropical cyclone vortex alone. For instance, Fig. 7 includes the observed, GAHM-only, and blended hindcast wind speeds at two NOAA stations (both are west of landfall). The peak winds are nearly identical. However, pre- and post-peak winds are virtually zero for the GAHM simulation, and the blended winds agree more closely with observations at these sites.

The simulated time series using the blended hindcast winds (Fig. 10) show an overall improvement in the pre-peak water levels and peak water levels. Specifically, the pre-peak water levels improve for stations 1–4 and 6 (NOAA8729210, FLBAY26247, NOAA8729108, FLBAY03283, and NOAA87288690). The peak storm surge is improved at stations 7, 8, and 10 (FLFRA26257, FLFRA26263, and CTMF1). Similar to the pre-storm tides, the recession limb of the main storm surge hydrograph is also improved on many of the stations, namely, stations 8 and 10 (FLFRA26263 and CTMF1). The average RMSE across all 14 stations is reduced from 0.37 to 0.29 m by using the blended winds. The largest reduction in RMSE occurs at station 6 (NOAA8728690), where the RMSE decreases by 0.20 m, and is followed by a 0.19 m reduction in RMSE at stations 7 and 8 (FLFRA26257 and FLFRA26263). We do note that the RMSE and increased at station 9 (FLWAK03369) as well as the peak storm surge is overpredicted when using the blended winds.

There is little difference in the simulated significant wave height and peak wave period between the GAHM-only and blended wind fields at the four buoy locations (Fig. 11). However, there is a large difference between the two model simulations in mean wave direction, particularly at NDBC42003 (W1) and NDBC42040 (W4). Unfortunately, it is difficult to definitively state that the GAHM-only or blended wind fields produced a more favorable result do to the lack of data on 10 October for NDBC42003 (W1) and noise in the observed wave directions for NDBC42040 (W4). Furthermore, there is little difference (less than 0.75 m) in the simulated peak significant wave height between the GAHM-only and blended wind fields (Fig. 12f). In general, the GAHM-only run results in lower peak wave heights offshore. There is a slight increase in wave heights resulting from the blended wind simulation within the Big Bend region near Apalachee Bay.

5. Discussion

a. *Barriers to operationalization*

A limitation of surge guidance is the description of the far-field winds, which can cause a water level setup prior to the arrival of tropical storm strength winds, especially in regions with wide and shallow continental shelves (Asher et al. 2019). To this end, we explored the performance gains that could have been achieved during Hurricane Michael if our scenario-based framework was expanded to include far-field winds. The GAHM vortex model using the earliest (post-landfall) available best track information was blended with the NAM analysis for the same time period. The results show that including far-field winds improve the overall water level forecast and does not increase run time.

As a natural extension of this hindcasting effort, we preliminarily implemented the blended winds in mock forecasting scenarios to identify barriers to operationalization. To achieve this, wind and pressure fields were created using GAHM for NHC advisories 6, 8, 10, 12, and 14 (as described in sections 4a and 4d). However, the GAHM-only forecasts fields were then blended with the NAM wind and pressure forecast fields readily available at the release of the respective NHC advisories. The resulting time series of water levels for three select gauges are shown in Fig. 13. Most notably, water level predictions are

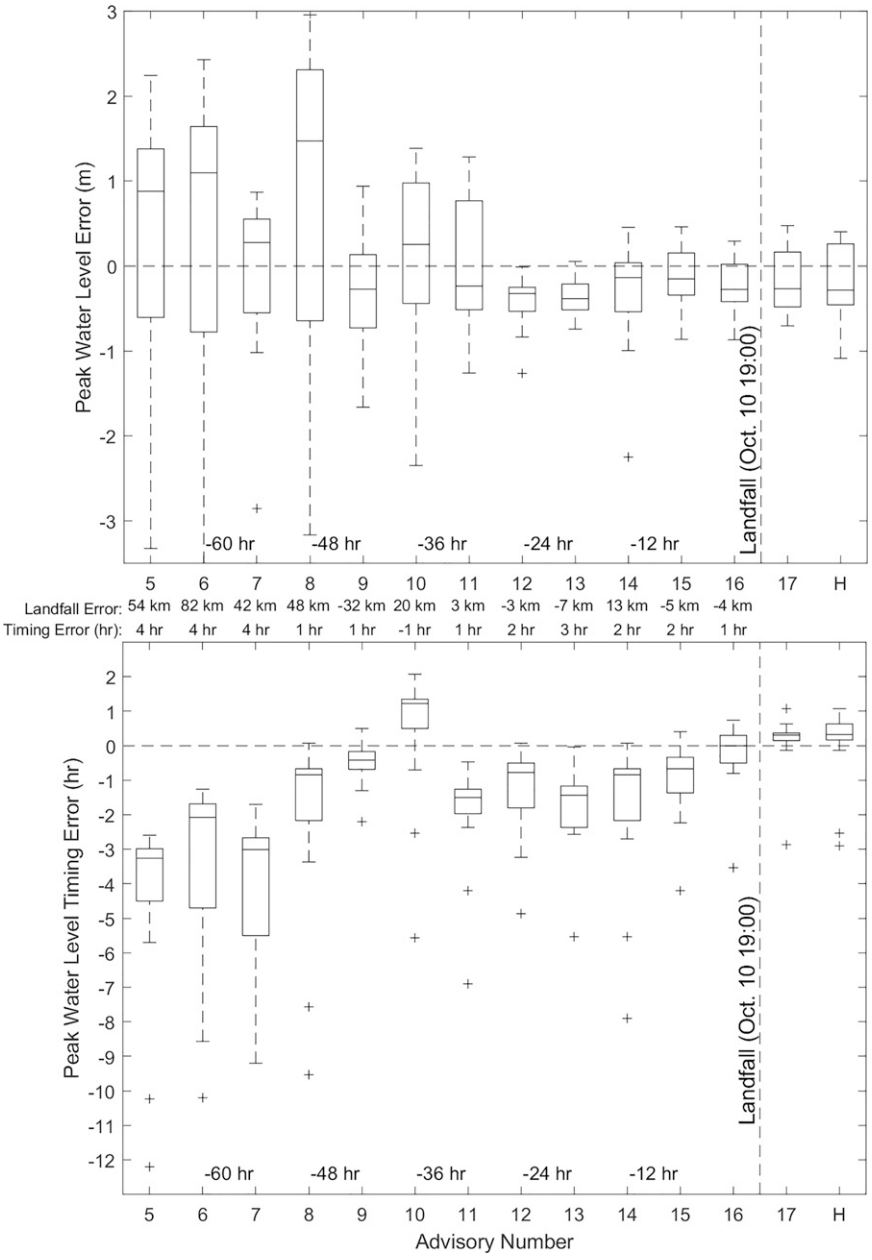

FIG. 9. (a) Peak water level error (m) and (b) peak water level timing error (h) of the forecast simulations of NHC advisory 5–17 and the hindcast (H, advisory 18) as compared to observed water levels (hourly) from 10 to 12 Oct 2018. Outliers are shown as the "+" symbol and are values outside 1.5 times the interquartile range. The landfall error (km) and timing until landfall (h) for each advisory are shown (also in Table 1).

more accurate in representing the pre- and post-peak values. Additionally, the magnitude and peak surge timing is better represented by the blended winds, specifically for stations east of landfall and away from the high winds of the tropical cyclone (NOAA8728690 and CTMF1). While these stations show a reduction in error, further investigation is warranted using a broader set of observations. For example, differences in the GAHM-only and blended-forced simulated water levels may be a result of capturing diverging forecast tracks between the NAM and NHC forecast.

Although the mock forecast and hindcast simulations indicate a need to include non-tropical far-field winds when forecasting hurricane storm surge, there are some limitations in using this approach. One major potential limitation is a divergence in the NHC forecast track and the selected meteorological model forecasts. For example, there was a relatively large

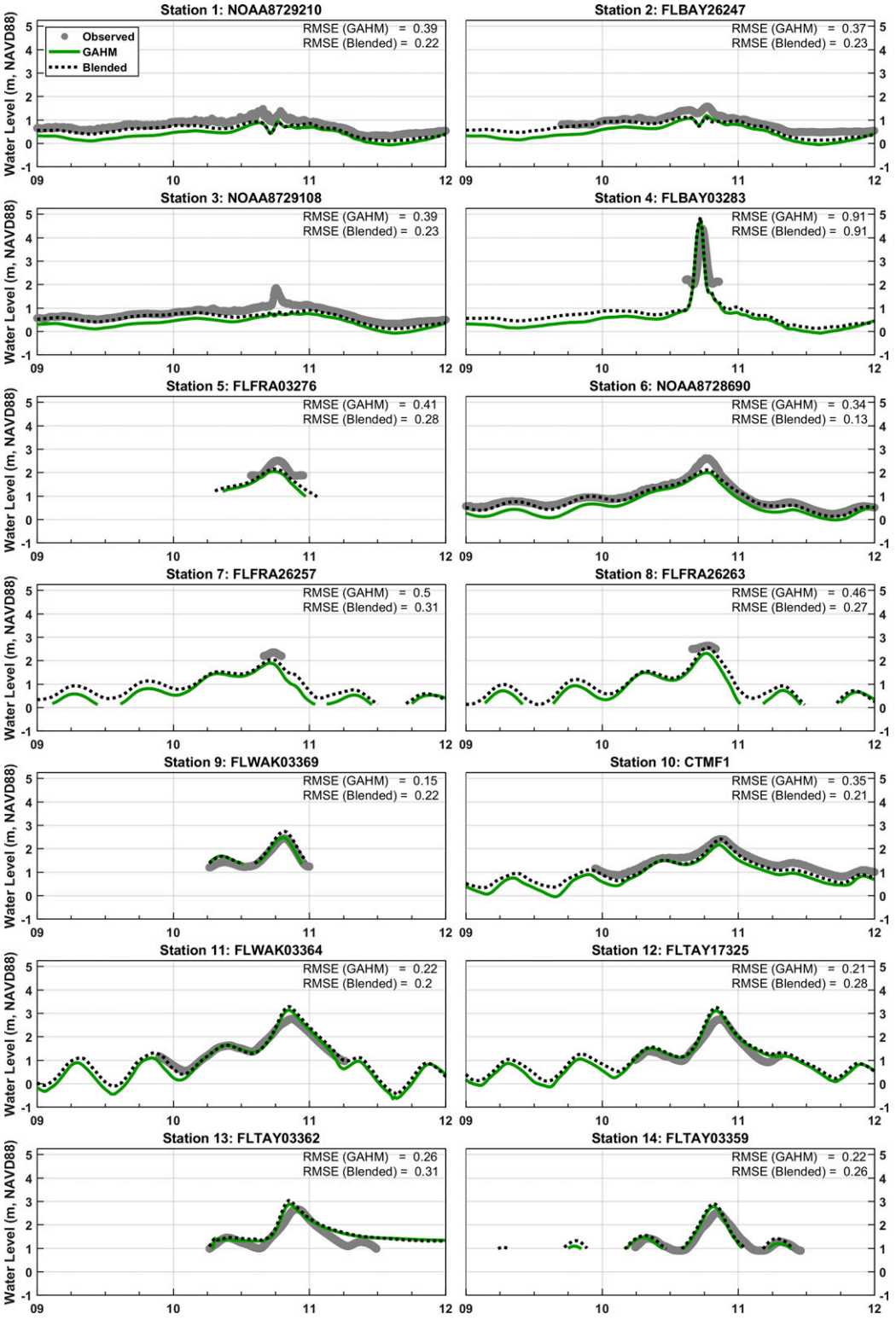

FIG. 10. Water level time series (m, NAVD88) for observed (gray circles) and ADCIRC+SWAN hindcast results from 9 to 12 Oct 2018. The green line is ADCIRC+SWAN results from GAHM and the black dashed line are from the blended GAHM/NAM simulation. Both results are from the operational best track simulation (i.e., hindcast). Station locations are shown in Fig. 5 and Table 2. Root-mean-square error (RMSE) is shown in the top right for the GAHM and GAHM+NAM (blended) simulation.

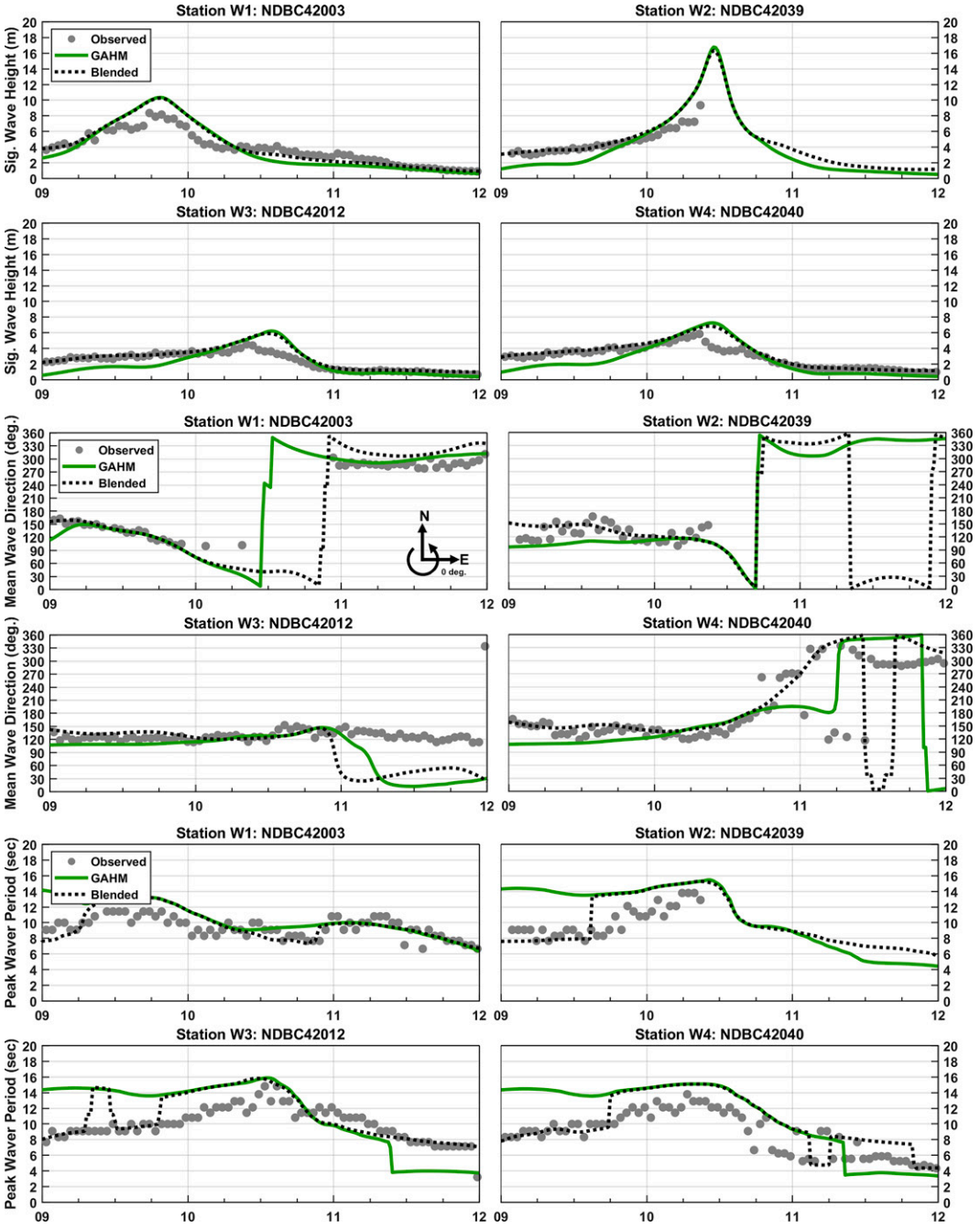

FIG. 11. Significant wave height (m), mean wave direction (°), and peak wave period (s) for observed (gray circles) and ADCIRC+SWAN hindcast results from 9 to 12 Oct 2018. The green line is ADCIRC+SWAN results from GAHM and the black dashed line are from the blended GAHM+NAM simulation. Both results are from the operational best track simulation (i.e., hindcast). Buoy locations are shown in Fig. 1a.

difference (140 km) between the NHC advisory 6 forecast track and NAM's representation of Hurricane Michael (cycle 2018100800; Fig. 14). While this might be expected given that the NAM was selected for logistical considerations rather than its reliability as a TC guidance tool, this same limitation could be circumstantially true for any NWP model. Because the NHC consults numerous global (e.g., GFS) and vortex-following (e.g., Hurricane-WRF) numerical weather models before issuing an advisory, the deterministic output from any single model is always likely to deviate from the NHC's advisory, which typically reflects the spread among the ensemble members. For instance, the NHC's forecast discussion issued contemporaneously with advisory 6 stated that " … the official forecast track leans more toward the HCCA/IVCN and

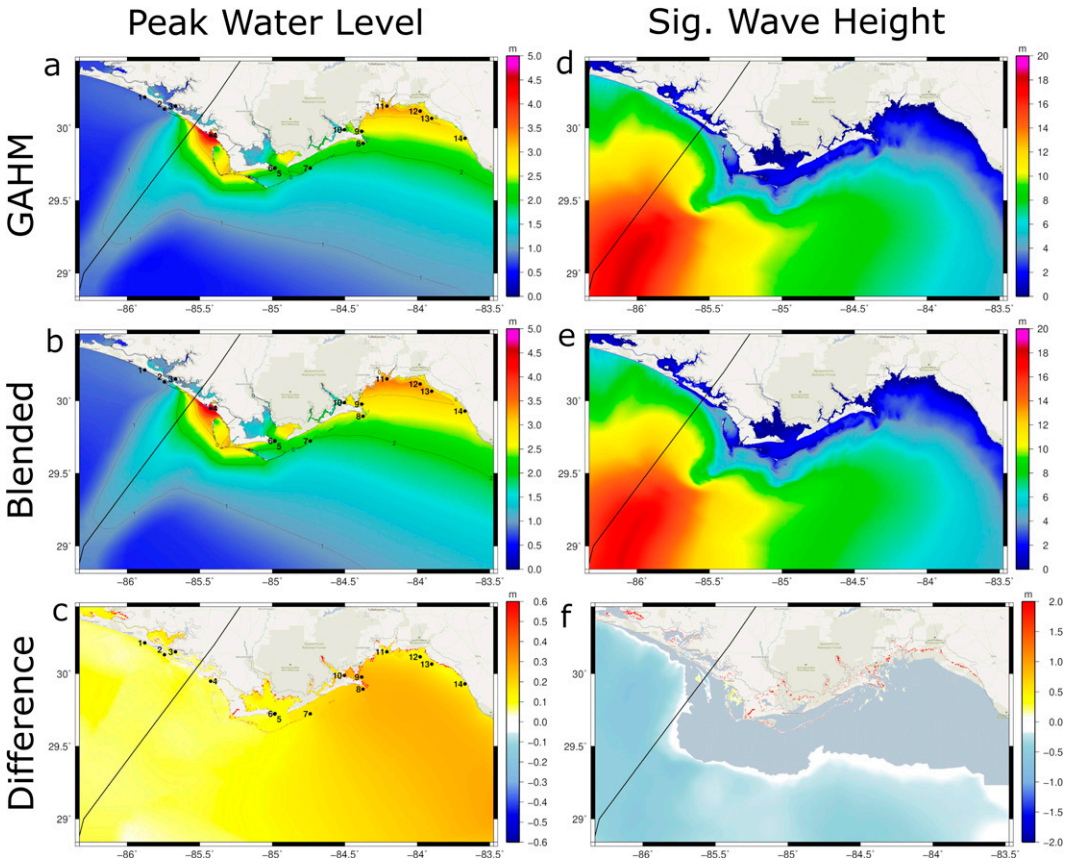

FIG. 12. (a)–(f) Maximum ADCIRC+SWAN simulated water levels (m, NAVD88) and significant wave height (m) resulting from the operational best track simulation NAM and GAHM+NAM (blended) simulation (e.g., hindcast simulation). In (c) and (f), the peak water level and peak significant wave height difference between the blended and GAHM simulations are shown. The black line is the storm track. Refer to Fig. 1 and Table 2 for gauge station information.

UKMET model solutions." In this case, only a small subset of the weather models available for far-field winds would have mirrored the official forecast track used to derive the GAHM wind field.

The primary hazard associated with this disagreement occurs when the meteorological model's predicted track diverges $> 3R_{max}$ (see section 3d) from the official NHC forecast. At this threshold, the near-field winds associated with meteorological model's depiction of the storm encroach upon the zone where far-field winds are blended into the NHC-based GAHM wind field. In cases of severe divergence, the "blended" winds effectively contain two cyclones: one for the NHC-based GAHM forecast and one for the meteorological model that was intended to supply the far-field winds only.

b. *Future improvements*

Based on the preliminary discussion in section 5a, future work should explore methods for leveraging the far-field wind data from a weather model even when its own TC path deviates from the NHC advisory. Such efforts might attempt to mask winds above a certain threshold or supplement using a climatological wind vector.

Future work should also quantify meteorological forecast errors for a wide range of historic hurricanes and how they propagate into the surge forecasts while considering various modeling approaches. Future research ought to focus on methods to improve scenario-based storm surge predictions provided known uncertainties in the meteorological forecasts. For instance, NHC forecast discussions often contain language characterizing the average error of the track forecast at certain lead times (e.g., "It should be noted that the average NHC track errors are 175 miles and 200 miles at days 4 and 5, respectively."). The development of similar metrics for storm surge forecasts (e.g., "The average water level forecast error at Day 2 is 1 m.") could benefit end-users. Such statements would require a standardized approach among the forecasting and modeling community.

Hurricane evacuation measures typically require lead times of 48–96 h while flood gate closures and marine safe harbor access require lead times of 36–72 h (Boukhanovsky and Ivanov 2012; Czajkowski 2011; Morrow et al. 2015; Munroe et al. 2018; Regnier 2008; Wolshon et al. 2005). Early watches

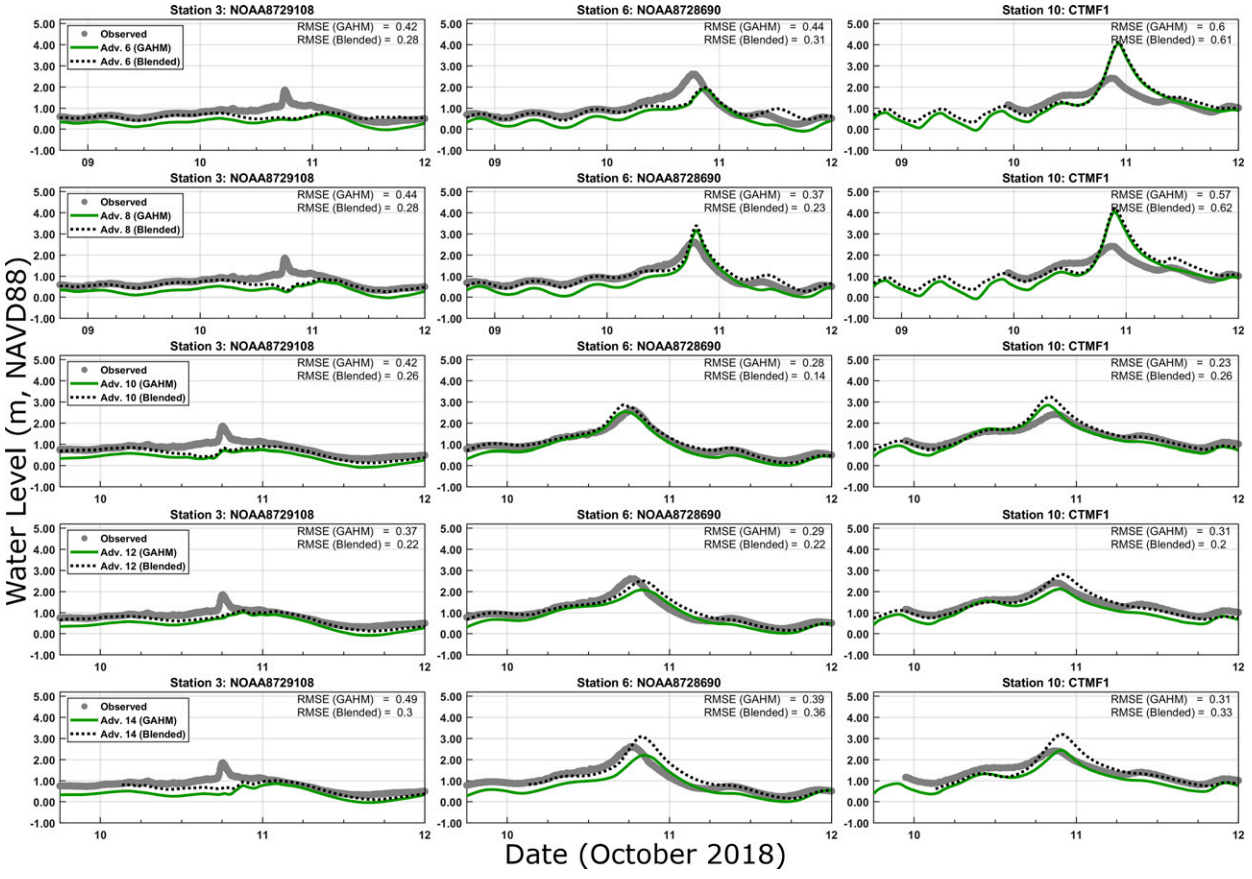

FIG. 13. Water level time series (m, NAVD88) for observed (gray circles) and ADCIRC+SWAN results from 1800 UTC 9–12 Oct 2018. The green line is ADCIRC+SWAN results from GAHM and the black dashed line are from the blended GAHM+NAM forecast simulation (using the NAM forecast most readily available at the time of the NHC issuance). Station locations are shown in Fig. 1 and Table 2. Root-mean-square error (RMSE) is shown in the top right for the GAHM and GAHM+NAM (blended) simulation. RMSE for all simulations was computed from the NHC forecast advisory until 12 Oct. Each row represents a difference NHC advisory (6, 8, 10, 12, and 14).

and warnings are issued and plans to protect critical infrastructure are implemented when hazardous conditions are expected within 24–48 h (https://www.weather.gov/safety/hurricane-ww) (Golding 2009). Having a better understanding of the evolution of the accuracy (and related uncertainties) of storm surge forecasts enables stakeholders to make more informed and timelier decisions (Ramos et al. 2010).

Scenario-based storm surge guidance provides a means to examine credible realization of the regional surge response, and when augmented with expert opinion, allows forecasters and emergency managers to consider the sensitivity of localized peak surges to potential shifts in TC forecast track and intensity. Additionally, TC track predictions have already, or may be near, their limit of predictability. Therefore, efforts should be made to enhance surge forecasts in ways beyond improved TC track forecasts while developing a better understanding how forecast track uncertainty contributes to surge forecast error. Nonetheless, there is much improvement that can be made in intensity forecasts such as increased observations, better understanding and simulation of the atmospheric–ocean–wave interactions, adaptive mesh refinement, and downscaling results to finer resolutions (Landsea and Cangialosi 2018; Rogers et al. 2006; Rucker et al. 2021; Thomas et al. 2021; Wada et al. 2010; Zambon et al. 2014).

6. Conclusions

Herein, we provided an examination and assessment of ADCIRC+SWAN hurricane storm surge forecasts that were performed in real time during Hurricane Michael (2018) and investigated one particular method for improvement (i.e., blending cyclone with far-field meteorology). Numerical simulations were forced by a parametric wind field derived solely from information provided by the NHC for each forecast advisory. The small NHC forecast track errors during the evolution of Hurricane Michael were notably less than the reported 5-yr average hurricane season track errors. The largest radial landfall error during Michael was 82 km at 62 h prior to landfall. Wind intensity forecasts generally underpredicted observed operational best track values, with the at-landfall error consistently more than double the previous 5-yr average intensity errors.

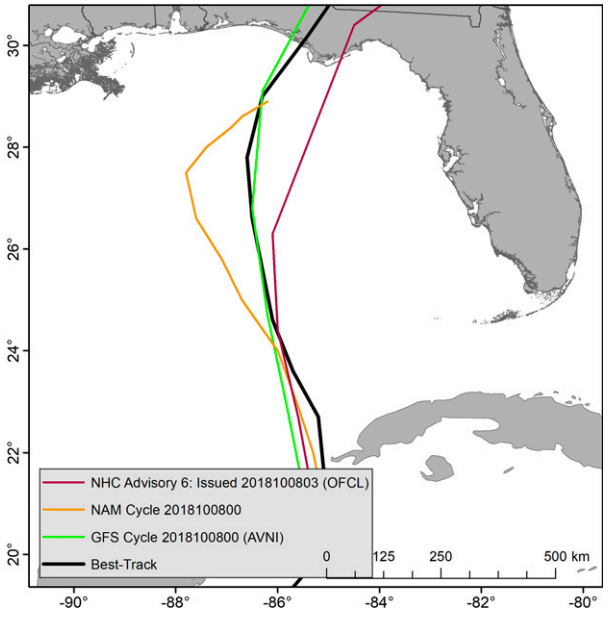

FIG. 14. NHC forecast track (OFCL) for advisory 6 (62 h before landfall, issued on 0300 UTC 8 Oct 2018), the NAM forecast track (cycle 2018100800), the GFS (AVNI) forecast track (cycle 2018100800), and the operational best track for Hurricane Michael. The NAM forecast is for 84 h (3.5 days), which is why it stops short of landfall.

The effects of the evolution of Hurricane Michael NHC forecasts on real-time storm surge forecasts were investigated. Model simulations indicate that water level predictions and coastal flooding were in good agreement with observed data 48 h prior to landfall. Additionally, comparison of forecast water levels to observations indicate that RMS errors fell below 0.5 m with 42 h until landfall. Similarly, differences in peak water level and peak water level timing were less than 0.5 m and 2 at 42 h prior to landfall, respectively. The reduction in water level errors is attributed to the improvement in the storm's forecast intensity error and stable track forecast error beginning with NHC advisories 9 and 10. These results indicate that real-time, deterministic, simulated water level forecasts closely match observations. Therefore, we provide evidence that scenario-based surge guidance may provide reasonable results for pre-storm emergency response. This is especially true when considering the effect of far-field winds and its impact on total water levels and the timing of peak surge.

Deterministic-based forecasts are limited by the forecast track and intensity. We found that small changes in Michael's forecast track can alter the region of maximum storm surge (and its timing), which can vary from forecast to forecast. This is especially true with forecasts beyond 48 h prior to landfall. As Michael's forecast track error was unprecedently small, such errors will be exacerbated for storms that have larger error in the forecast track, which may have serious consequences on the flood forecast (e.g., misplaced and/or delayed storm surge warnings). With more uncertain forecast tracks, probabilistic guidance may be more beneficial than deterministic. This is an area of future work. In particular, to examine the advantages and disadvantages for probabilistic and deterministic surge forecasts for various storm tracks across different regions.

Although specific to Hurricane Michael, this study contributes to the ongoing research of providing accurate and reliable guidance of hurricane storm surge flooding with lead times of hours to days. Such efforts are critical to flood hazard and risk management, especially in low-lying coastal regions that are frequently impacted by TCs.

Acknowledgments. This material is based upon work supported by the U.S. Department of Homeland Security under Grant Award 2015-ST-061-ND0001-01 and the National Oceanic Partnership Program Award 5120731. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S Department of Homeland Security. This research was also funded in part under the National Oceanic and Atmospheric Administration (NOAA) Ecological Effects of Sea Level Rise (EESLR) Program (Awards NA10NOS4780146 and NA16NOS4780208), and the Louisiana Sea Grant Laborde Chair. This work used High Performance Computing at the University of Georgia's (UGA) Georgia Advanced Computing Research Center, Louisiana State University (LSU), and the Louisiana Optical Network Initiative (LONI). The statements and conclusions are those of the authors and do not necessarily reflect the views of DHS, NOAA, Louisiana Sea Grant, UGA, LSU, or LONI. The authors also wish to thank the three anonymous reviewers for their invaluable input into this manuscript.

APPENDIX

Generalized Asymmetrical Holland Model

During run time, ADCIRC seams the GAHM and NAM products. GAHM is used for the tropical cyclone vortex, and NAM is blended in over several TC radii for a smooth transition between the TC vortex and far-field wind and pressure. First, the spherical distance from the storm center to each mesh node, r_{node}, is computed. A blending factor f_{blend}, ranging from zero to one, is assigned to each mesh node based on its distance to the storm center. A blending factor of one indicates that the node is entirely within the radial extent of the vortex, whereas a blending factor of 0 means the mesh node is outside the vortex where its effects are insignificant. Two user-defined input parameters are necessary for this process, d_{vortex} and $d_{\text{background}}$, which are coefficients used to define the transition between the vortex and background meteorology. When multiplied by the radius of maximum winds (R_{max}) for the current time step, D_{vortex} [Eq. (A1)] and $D_{\text{background}}$ [Eq. (A2)] represent the radial distance to the pure edge of the vortex and the radial extent to where the vortex is not significant, respectively:

$$D_{\text{vortex}} = R_{\text{max}} d_{\text{vortex}}, \tag{A1}$$

$$D_{\text{background}} = R_{\text{max}} d_{\text{background}}. \tag{A2}$$

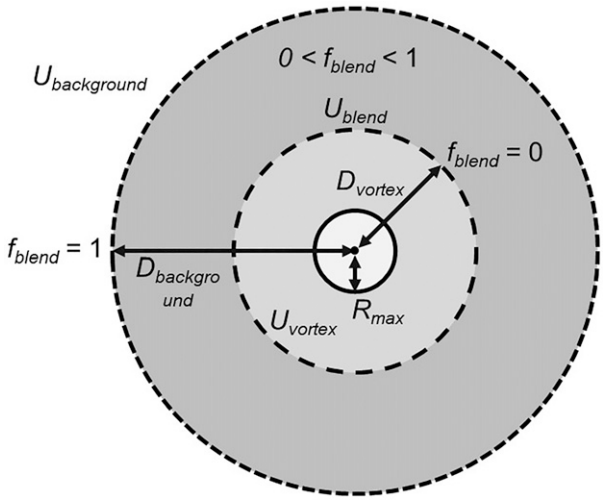

FIG. A1. Wind-blending schematic.

For mesh nodes within these radial distances (i.e., blended region), values between zero and one are determined based on the node's radial position within the blended region to obtain f_{blend} (Fig. A1). The blending factor [Eq. (A3)] is then used to blend the vortex and background winds (U_{blend}) using Eq. (A4):

$$f_{\text{blend}} = \frac{D_{\text{background}} - r_{\text{node}}}{D_{\text{background}} - D_{\text{vortex}}}, \tag{A3}$$

$$U_{\text{blend}} = f_{\text{blend}} U_{\text{vortex}} + (1 - f_{\text{blend}}) U_{\text{background}}, \tag{A4}$$

where U_{vortex} is the wind speed vector from the vortex wind model, and $U_{\text{background}}$ is the wind speed vector from the background (i.e., NAM) winds. This calculation is performed in the x and y directions.

REFERENCES

Anarde, K. A., S. Kameshwar, J. N. Irza, J. A. Nittrouer, J. Lorenzo-Trueba, J. E. Padgett, A. Sebastian, and P. B. Bedient, 2018: Impacts of hurricane storm surge on infrastructure vulnerability for an evolving coastal landscape. *Nat. Hazards Rev.*, **19**, 04017020, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000265.

Asher, T. G., R. A. Luettich Jr., J. G. Fleming, and B. O. Blanton, 2019: Low frequency water level correction in storm surge models using data assimilation. *Ocean Modell.*, **144**, 101483, https://doi.org/10.1016/j.ocemod.2019.101483.

Battjes, J. A., and J. P. F. M. Janssen, 1978: Energy loss and set-up due to breaking of random waves. *Proc. 16th Conf. on Coastal Engineering*, Hamburg, Germany, American Society of Civil Engineers, 569–587.

Bertin, X., K. Li, A. Roland, and J.-R. Bidlot, 2015: The contribution of short-waves in storm surges: Two case studies in the Bay of Biscay. *Cont. Shelf Res.*, **96**, 1–15, https://doi.org/10.1016/j.csr.2015.01.005.

Beven, J. L., II, R. Berg, and A. Hagen, 2019: Tropical cyclone report: Hurricane Michael (7–11 October 2018). NHC Tech. Rep. AL142018, 86 pp., https://www.nhc.noaa.gov/data/tcr/AL142018_Michael.pdf.

Bilskie, M. V., D. Coggin, S. C. Hagen, and S. C. Medeiros, 2015: Terrain-driven unstructured mesh development through semi-automatic vertical feature extraction. *Adv. Water Resour.*, **86**, 102–118, https://doi.org/10.1016/j.advwatres.2015.09.020.

——, S. C. Hagen, S. C. Medeiros, A. T. Cox, M. Salisbury, and D. Coggin, 2016a: Data and numerical analysis of astronomic tides, wind-waves, and hurricane storm surge along the northern Gulf of Mexico. *J. Geophys. Res. Oceans*, **121**, 3625–3658, https://doi.org/10.1002/2015JC011400.

——, ——, K. Alizad, S. C. Medeiros, D. L. Passeri, H. F. Needham, and A. Cox, 2016b: Dynamic simulation and numerical analysis of hurricane storm surge under sea level rise with geomorphologic changes along the northern Gulf of Mexico. *Earth's Future*, **4**, 177–193, https://doi.org/10.1002/2015EF000347.

——, ——, and S. C. Medeiros, 2020: Unstructured finite element mesh decimation for real-time Hurricane storm surge forecasting. *Coastal Eng.*, **156**, 103622, https://doi.org/10.1016/j.coastaleng.2019.103622.

Blanton, B., and Coauthors, 2012: Urgent computing of storm surge for North Carolina's coast. *Procedia Comput. Sci.*, **9**, 1677–1686, https://doi.org/10.1016/j.procs.2012.04.185.

Boukhanovsky, A. V., and S. V. Ivanov, 2012: Urgent computing for operational storm surge forecasting in Saint-Petersburg. *Procedia Comput. Sci.*, **9**, 1704–1712, https://doi.org/10.1016/j.procs.2012.04.188.

Bunya, S., and Coauthors, 2010: A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southeastern Louisiana and Mississippi. Part I: Model development and validation. *Mon. Wea. Rev.*, **138**, 345–377, https://doi.org/10.1175/2009MWR2906.1.

Byrne, M. J., Sr., 2019: Monitoring storm tide from Hurricane Michael along the northwest coast of Florida, October 2018. U.S. Geological Survey Open-File Rep. 2019-1059, USGS, https://doi.org/10.3133/ofr20191059.

Cangialosi, J. P., 2018: National Hurricane Center Forecast pdfs report: 2017 hurricane season. NOAA/NWS, 73 pp., https://www.nhc.noaa.gov/pdfs/pdfs/pdfs_2017.pdf.

Cavaleri, L., and P. M. Rizzoli, 1981: Wind wave prediction in shallow water: Theory and applications. *J. Geophys. Res.*, **86**, 10 961–10 973, https://doi.org/10.1029/JC086iC11p10961.

Cloke, H. L., and F. Pappenberger, 2009: Ensemble flood forecasting: A review. *J. Hydrol.*, **375**, 613–626, https://doi.org/10.1016/j.jhydrol.2009.06.005.

Czajkowski, J., 2011: Is it time to go yet? Understanding household hurricane evacuation decisions from a dynamic perspective. *Nat. Hazards Rev.*, **12**, 72–84, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000037.

Dale, M., J. Wicks, K. Mylne, F. Pappenberger, S. Laeger, and S. J. N. H. Taylor, 2014: Probabilistic flood forecasting and decision-making: An innovative risk-based approach. *Nat. Hazards*, **70**, 159–172, https://doi.org/10.1007/s11069-012-0483-z.

Davis, J. R., V. A. Paramygin, D. Forrest, and Y. P. Sheng, 2010: Toward the probabilistic simulation of storm surge and inundation in a limited-resource environment. *Mon. Wea. Rev.*, **138**, 2953–2974, https://doi.org/10.1175/2010MWR3136.1.

Dawson, C., J. J. Westerink, J. C. Feyen, and D. Pothina, 2006: Continuous, discontinuous and coupled discontinuous–continuous Galerkin finite element methods for the shallow water

equations. *Int. J. Numer. Methods Fluids*, **52**, 63–88, https://doi.org/10.1002/fld.1156.

DeLorme, D. E., S. H. Stephens, M. V. Bilskie, and S. C. Hagen, 2020: Coastal decision-makers' perspectives on updating storm surge guidance tools. *J. Contingencies Crisis Manage.*, **28**, 158–168, https://doi.org/10.1111/1468-5973.12291.

DeMaria, M., C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398, https://doi.org/10.1175/BAMS-D-12-00240.1.

Dietrich, J. C., and Coauthors, 2011a: Modeling hurricane waves and storm surge using integrally-coupled, scalable computations. *Coastal Eng.*, **58**, 45–65, https://doi.org/10.1016/j.coastaleng.2010.08.001.

——, and Coauthors, 2011b: Hurricane Gustav (2008) waves and storm surge: Hindcast, synoptic analysis, and validation in southern Louisiana. *Mon. Wea. Rev.*, **139**, 2488–2522, https://doi.org/10.1175/2011MWR3611.1.

——, A. Muhammad, M. Curcic, A. Fathi, C. N. Dawson, S. S. Chen, and R. A. Luettich, 2018: Sensitivity of storm surge predictions to atmospheric forcing during Hurricane Isaac. *J. Waterw., Port, Coastal, Ocean Eng.*, **144**, 04017035, https://doi.org/10.1061/(ASCE)WW.1943-5460.0000419.

Egbert, G. D., and S. Y. Erofeeva, 2002: Efficient inverse modeling of barotropic ocean tides. *J. Atmos. Oceanic Technol.*, **19**, 183–204, https://doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2.

——, A. F. Bennett, and M. G. G. Foreman, 1994: TOPEX/POSEIDON tides estimated using a global inverse model. *J. Geophys. Res.*, **99**, 24 821–24 852, https://doi.org/10.1029/94JC01894.

Emanuel, K., 2017: Will global warming make hurricane forecasting more difficult? *Bull. Amer. Meteor. Soc.*, **98**, 495–501, https://doi.org/10.1175/BAMS-D-16-0134.1.

Fleming, J. G., C. W. Fulcher, R. A. Luettich, B. D. Estrade, G. D. Allen, and H. S. Winer, 2007: A real time storm surge forecasting system using ADCIRC. *10th Int. Conf. on Estuarine and Coastal Modeling,* Newport, RI, American Society of Civil Engineers, 893–912.

Flowerdew, J., K. Horsburgh, C. Wilson, and K. Mylne, 2010: Development and evaluation of an ensemble forecasting system for coastal storm surges. *Quart. J. Roy. Meteor. Soc.*, **136**, 1444–1456, https://doi.org/10.1002/qj.648.

Funakoshi, Y., S. C. Hagen, and P. Bacopoulos, 2008: Coupling of hydrodynamic and wave models: Case study for Hurricane Floyd (1999) hindcast. *J. Waterw., Port, Coastal, Ocean Eng.*, **134**, 321–335, https://doi.org/10.1061/(ASCE)0733-950X(2008)134:6(321).

Gall, R., J. Franklin, F. Marks, E. N. Rappaport, and F. Toepfer, 2013: The Hurricane Forecast Improvement Project. *Bull. Amer. Meteor. Soc.*, **94**, 329–343, https://doi.org/10.1175/BAMS-D-12-00071.1.

Gao, J., 2018: *On the Surface Wind Stress for Storm Surge Modelling*. University of North Carolina–Chapel Hill, 109 pp.

Golding, B. W., 2009: Long lead time flood warnings: Reality or fantasy? *Meteor. Appl.*, **16**, 3–12, https://doi.org/10.1002/met.123.

Hamill, T. M., M. J. Brennan, B. Brown, M. DeMaria, E. N. Rappaport, and Z. Toth, 2012: NOAA's future ensemble-based hurricane forecast products. *Bull. Amer. Meteor. Soc.*, **93**, 209–220, https://doi.org/10.1175/2011BAMS3106.1.

Hazelton, A. T., X. Zhang, S. Gopalakrishnan, W. Ramstrom, F. Marks, and J. A. Zhang, 2020: High-resolution ensemble HFV3 forecasts of Hurricane Michael (2018): Rapid intensification in shear. *Mon. Wea. Rev.*, **148**, 2009–2032, https://doi.org/10.1175/MWR-D-19-0275.1.

Hu, K., Q. Chen, and S. K. Kimball, 2012: Consistency in hurricane surface wind forecasting: An improved parametric model. *Nat. Hazards*, **61**, 1029–1050, https://doi.org/10.1007/s11069-011-9960-z.

Kaplan, J., M. DeMaria, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **25**, 220–241, https://doi.org/10.1175/2009WAF2222280.1.

Kennedy, A. B., D. Wirasaet, A. Begmohammadi, T. Sherman, D. Bolster, and J. C. Dietrich, 2019: Subgrid theory for storm surge modeling. *Ocean Modell.*, **144**, 101491, https://doi.org/10.1016/j.ocemod.2019.101491.

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.

Kinnmark, I., 1986: *The Shallow Water Wave Equations: Formulation, Analysis, and Application.* Springer-Verlag, 188 pp., https://doi.org/10.1007/978-3-642-82646-7.

Kolar, R. L., W. G. Gray, J. J. Westerink, M. E. Cantekin, and C. A. Blain, 1994: Aspects of nonlinear simulations using shallow-water models based on the wave continuity equation. *Comput. Fluids*, **23**, 523–538, https://doi.org/10.1016/0045-7930(94)90017-5.

Komen, G. J., K. Hasselmann, and K. Hasselmann, 1984: On the existence of a fully developed wind-sea spectrum. *J. Phys. Oceanogr.*, **14**, 1271–1285, https://doi.org/10.1175/1520-0485(1984)014<1271:OTEOAF>2.0.CO;2.

Landsea, C. W., and J. P. Cangialosi, 2018: Have we reached the limits of predictability for tropical cyclone track forecasting. *Bull. Amer. Meteor. Soc.*, **99**, 2237–2243, https://doi.org/10.1175/BAMS-D-17-0136.1.

Lin, N., P. Lane, K. A. Emanuel, R. M. Sullivan, and J. P. Donnelly, 2014: Heightened hurricane surge risk in northwest Florida revealed from climatological-hydrodynamic modeling and paleorecord reconstruction. *J. Geophys. Res. Atmos.*, **119**, 8606–8623, https://doi.org/10.1002/2014JD021584.

Luettich, R. A., and J. J. Westerink, 2004: Formulation and numerical implementations of the 2D/3D ADCIRC finite element model version 44.XX. 74 pp., https://adcirc.org/wp-content/uploads/sites/2255/2018/11/adcirc_theory_2004_12_08.pdf.

Madsen, O. S., Y. K. Poon, and H. C. Graber, 1988: Spectral wave attenuation by bottom friction: Theory. *Proc. 21st Int. Conf. on Coastal Engineering Proceedings*, Malaga, Spain, American Society of Civil Engineers, 492–504.

Morales-Hernández, M., and Coauthors, 2020: High-performance computing in water resources hydrodynamics. *J. Hydroinf.*, **22**, 1217–1235, https://doi.org/10.2166/hydro.2020.163.

Morey, S. L., S. Baig, M. A. Bourassa, D. S. Dukhovskoy, and J. J. O'Brien, 2006: Remote forcing contribution to storm-induced sea level rise during Hurricane Dennis. *Geophys. Res. Lett.*, **33**, L19603, https://doi.org/10.1029/2006GL027021.

Mori, N., M. Kato, S. Kim, H. Mase, Y. Shibutani, T. Takemi, K. Tsuboki, and T. Yasuda, 2014: Local amplification of storm surge by Super Typhoon Haiyan in Leyte Gulf. *Geophys. Res. Lett.*, **41**, 5106–5113, https://doi.org/10.1002/2014GL060689.

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.

Munroe, R., B. Montz, and S. Curtis, 2018: Getting more out of storm surge forecasts: Emergency support personnel needs in North Carolina. *Wea. Climate Soc.*, **10**, 813–820, https://doi.org/10.1175/WCAS-D-17-0074.1.

National Oceanic and Atmospheric Administration, 2010: Coastal Change Analysis Program (C-CAP) regional land cover classification scheme. NOAA, 4 pp., https://coast.noaa.gov/data/digitalcoast/pdf/ccap-class-scheme-regional.pdf.

NHC, 2019: Tropical cyclone storm surge probabilities. NOAA, accessed 15 February 2022, https://www.nhc.noaa.gov/templates/graphics_psurge_em.shtml.

Plumlee, M., T. G. Asher, W. Chang, and M. V. Bilskie, 2021: High-fidelity hurricane surge forecasting using emulation and sequential experiments. *Ann. Appl. Stat.*, **15**, 460–480, https://doi.org/10.1214/20-AOAS1398.

Pugh, D. T., 1996: *Tides, Surges and Mean Sea-Level*. John Wiley & Sons Ltd., 472 pp.

Ramos, M.-H., T. Mathevet, J. Thielen, and F. Pappenberger, 2010: Communicating uncertainty in hydro-meteorological forecasts: Mission impossible? *Meteor. Appl.*, **17**, 223–235, https://doi.org/10.1002/met.202.

Rappaport, E. N., 2000: Loss of life in the United States associated with recent Atlantic tropical cyclones. *Bull. Amer. Meteor. Soc.*, **81**, 2065–2074, https://doi.org/10.1175/1520-0477(2000)081<2065:LOLITU>2.3.CO;2.

——, and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419, https://doi.org/10.1175/2008WAF2222128.1.

——, J. L. Franklin, A. B. Schumacher, M. DeMaria, L. K. Shay, and E. J. Gibney, 2010: Tropical cyclone intensity change before U.S. Gulf Coast landfall. *Wea. Forecasting*, **25**, 1380–1396, https://doi.org/10.1175/2010WAF2222369.1.

Reed, D. A., M. D. Powell, and J. M. Westerman, 2010: Energy infrastructure damage analysis for Hurricane Rita. *Nat. Hazards Rev.*, **11**, 102–109, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000012.

Regnier, E., 2008: Public evacuation decisions and hurricane track uncertainty. *Manage. Sci.*, **54**, 16–28, https://doi.org/10.1287/mnsc.1070.0764.

Resio, D. T., and J. J. Westerink, 2008: Modeling the physics of storm surges. *Phys. Today*, **61**, 33–38, https://doi.org/10.1063/1.2982120.

——, N. J. Powell, M. A. Cialone, H. S. Das, and J. J. Westerink, 2017: Quantifying impacts of forecast uncertainties on predicted storm surges. *Nat. Hazards*, **88**, 1423–1449, https://doi.org/10.1007/s11069-017-2924-1.

Rey, A. J. M., and R. P. Mulligan, 2021: Influence of hurricane wind field variability on real-time forecast simulations of the coastal environment. *J. Geophys. Res. Oceans*, **126**, e2020JC016489, https://doi.org/10.1029/2020JC016489.

Rhome, J. R., C. A. Sisko, and R. D. Knabb, 2006: On the calculation of vertical shear: An operational perspective. *27th Conf. on Hurricanes and Tropical Meteorology*, Monterey, CA, Amer. Meteor. Soc., 14A.4, https://ams.confex.com/ams/27Hurricanes/techprogram/paper_108724.htm.

Roberts, K., J. Dietrich, D. Wirasaet, W. Pringle, and J. Westerink, 2021: Dynamic load balancing for predictions of storm surge and coastal flooding. *Environ. Modell. Software*, **140**, 105045, https://doi.org/10.1016/j.envsoft.2021.105045.

Rogers, R., and Coauthors, 2006: The Intensity Forecasting Experiment: A NOAA multiyear field program for improving tropical cyclone intensity forecasts. *Bull. Amer. Meteor. Soc.*, **87**, 1523–1538, https://doi.org/10.1175/BAMS-87-11-1523.

Rogers, W. E., P. A. Hwang, and D. W. Wang, 2003: Investigation of wave growth and decay in the SWAN model: Three regional-scale applications. *J. Phys. Oceanogr.*, **33**, 366–389, https://doi.org/10.1175/1520-0485(2003)033<0366:IOWGAD>2.0.CO;2.

Rucker, C. A., N. Tull, J. C. Dietrich, T. E. Langan, H. Mitasova, B. O. Blanton, J. G. Fleming, and R. A. Luettich, 2021: Downscaling of real-time coastal flooding predictions for decision support. *Nat. Hazards*, **107**, 1341–1369, https://doi.org/10.1007/s11069-021-04634-8.

Tanaka, S., S. Bunya, J. J. Westerink, C. Dawson, and R. A. Luettich Jr., 2011: Scalability of an unstructured grid continuous Galerkin based hurricane storm surge model. *J. Sci. Comput.*, **46**, 329–358, https://doi.org/10.1007/s10915-010-9402-1.

Thomas, A., J. C. Dietrich, M. Loveland, A. Samii, and C. N. Dawson, 2021: Improving coastal flooding predictions by switching meshes during a simulation. *Ocean Modell.*, **164**, 101820, https://doi.org/10.1016/j.ocemod.2021.101820.

Torn, R. D., and C. Snyder, 2012: Uncertainty of tropical cyclone best-track information. *Wea. Forecasting*, **27**, 715–729, https://doi.org/10.1175/WAF-D-11-00085.1.

Wada, A., N. Kohno, and Y. Kawai, 2010: Impact of wave-ocean interaction on typhoon Hai-Tang in 2005. *SOLA*, **6A**, 13–16, https://doi.org/10.2151/sola.6A-004.

Weisberg, R. H., and L. Zheng, 2006: Hurricane storm surge simulations for Tampa Bay. *Estuaries Coasts*, **29**, 899–913, https://doi.org/10.1007/BF02798649.

Westerink, J. J., and Coauthors, 2008: A basin- to channel-scale unstructured grid hurricane storm surge model applied to Southern Lousiana. *Mon. Wea. Rev.*, **136**, 833–864, https://doi.org/10.1175/2007MWR1946.1.

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).

Wu, L., Q. Liu, and Y. Li, 2019: Tornado-scale vortices in the tropical cyclone boundary layer: Numerical simulation with the WRF–LES framework. *Atmos. Chem. Phys.*, **19**, 2477–2487, https://doi.org/10.5194/acp-19-2477-2019.

Xie, L., S. Bao, L. J. Pietrafesa, K. Foley, and M. Fuentes, 2006: A real-time hurricane surface wind forecasting model: Formulation and pdfs. *Mon. Wea. Rev.*, **134**, 1355–1370, https://doi.org/10.1175/MWR3126.1.

Zambon, J. B., R. He, and J. C. Warner, 2014: Investigation of Hurricane Ivan using the Coupled Ocean–Atmosphere–Wave–Sediment Transport (COAWST) model. *Ocean Dyn.*, **64**, 1535–1554, https://doi.org/10.1007/s10236-014-0777-7.

Zijlema, M., 2010: Computation of wind-wave spectra in coastal waters with SWAN on unstructured grids. *Coastal Eng.*, **57**, 267–277, https://doi.org/10.1016/j.coastaleng.2009.10.011.