10-2013

US IOOS coastal and ocean modeling testbed: Inter-model evaluation of tides, waves, and hurricane surge in the Gulf of Mexico

PC Kerr

AS Donahue

JJ Westerliink

RA Luettich

LY Zheng

See next page for additional authors

Authors

PC Kerr, AS Donahue, JJ Westerliink, RA Luettich, LY Zheng, RH Weisberg, Y Huang, Harry V. Wang, Y Teng, David R. Forrest, and et al

U.S. IOOS coastal and ocean modeling testbed: Inter-model evaluation of tides, waves, and hurricane surge in the Gulf of Mexico

P. C. Kerr,[1] A. S. Donahue,[1] J. J. Westerink,[1] R. A. Luettich Jr.,[2] L. Y. Zheng,[3] R. H. Weisberg,[3] Y. Huang,[3] H. V. Wang,[4] Y. Teng,[4] D. R. Forrest,[4] A. Roland,[5] A. T. Haase,[6] A. W. Kramer,[6] A. A. Taylor,[6] J. R. Rhome,[7] J. C. Feyen,[8] R. P. Signell,[9] J. L. Hanson,[10] M. E. Hope,[1] R. M. Estes,[1] R. A. Dominguez,[1] R. P. Dunbar,[1] L. N. Semeraro,[1] H. J. Westerink,[1] A. B. Kennedy,[1] J. M. Smith,[11] M. D. Powell,[12] V. J. Cardone,[13] and A. T. Cox[13]

[1] A Gulf of Mexico performance evaluation and comparison of coastal circulation and wave models was executed through harmonic analyses of tidal simulations, hindcasts of Hurricane Ike (2008) and Rita (2005), and a benchmarking study. Three unstructured coastal circulation models (ADCIRC, FVCOM, and SELFE) validated with similar skill on a new common Gulf scale mesh (ULLR) with identical frictional parameterization and forcing for the tidal validation and hurricane hindcasts. Coupled circulation and wave models, SWAN+ADCIRC and WWMII+SELFE, along with FVCOM loosely coupled with SWAN, also validated with similar skill. NOAA's official operational forecast storm surge model (SLOSH) was implemented on local and Gulf scale meshes with the same wind stress and pressure forcing used by the unstructured models for hindcasts of Ike and Rita. SLOSH's local meshes failed to capture regional processes such as Ike's forerunner and the results from the Gulf scale mesh further suggest shortcomings may be due to a combination of poor mesh resolution, missing internal physics such as tides and nonlinear advection, and SLOSH's internal frictional parameterization. In addition, these models were benchmarked to assess and compare execution speed and scalability for a prototypical operational simulation. It was apparent that a higher number of computational cores are needed for the unstructured models to meet similar operational implementation requirements to SLOSH, and that some of them could benefit from improved parallelization and faster execution speed.

Citation: Kerr, P. C., et al. (2013), U.S. IOOS coastal and ocean modeling testbed: Inter-model evaluation of tides, waves, and hurricane surge in the Gulf of Mexico, *J. Geophys. Res. Oceans*, *118*, 5129–5172, doi:10.1002/jgrc.20376.

---

1. Introduction

[2] Storm surge is the phenomenon of rising coastal water levels due to the cumulative effect of processes such as wind-driven setup and currents, geostrophic effects, wave setup, wave runup and breaking, atmospheric pressure changes, precipitation (freshwater flooding), and astronomical tides. The National Hurricane Center officially describes storm surge as the difference between water levels during the storm (referred to as storm tide) and predicted astronomical tides. This is not to say that astronomical tides do not influence storm surge, but rather that storm surge is considered an abnormal rise of water generated by a storm over and above astronomically predicted levels. The resulting high water, powerful currents, erosive velocities, and saline intrusion along the nearshore and shallow inland areas across the shore and via the conduit of bays, wetlands, rivers, and channels pose a significant threat to human life, infrastructure, and freshwater ecosystems. The trend of population and infrastructure growth in low-lying coastal areas, combined with the projection for increased flood susceptibility [*Mousavi et al.*, 2010; *Ning et al.*, 2012] due to changes in storm climatology and sea level rise, suggests a mounting necessity to advance our understanding of coastal storm surge and improve our risk assessment and forecasting abilities.

[3] Significant progress has been made over the past half century to understand storm evolution and coastal flooding processes. Outside the broad objective of understanding these physical processes, the two primary practical uses for coastal storm surge models are operational forecasting and design/risk analysis. Operational forecasting requires an ensemble approach of modeling variant predictions in storm track and characteristics, that must be executed on a consistently stable basis within a short window of simulation time and provide realistic results that can be processed and interpreted quickly. Operational forecasts are used by emergency managers and by the public to ensure that risk and loss of life are minimized. Because the overwhelming controlling factors in operational forecasting are "time" and "robustness/stability," metrics such as skill, resolution, and geographic range are often sacrificed so that a high number of probabilistic storm variants can be simulated. Design and risk analysis, on the other hand, has a primary objective of designing coastal protection structures and assessing risk that drives local development as well as government subsidized flood insurance. Accuracy and resolution are essential, and simulation time is of secondary concern. Design and risk analysis also takes an ensemble approach to storm surge modeling, but instead of modeling the variants of a single predicted storm, a suite of hypothetical storms are modeled, so that all contingencies are addressed. One objective of design/risk analysis is to predict flood potential statistically. A recurrence interval measure, known as the Base Flood Elevation (BFE), is a measure by which flood risk reduction systems, bridges, levees, and other structures are designed. Both operational forecasting and design/risk analysis use composites of maximum storm surge water levels, referred to as the Maximum Envelope of high Water (MEOW) for a single storm and the Maximum of the Maximum (MOM) for a suite of probabilistic storms.

[4] The National Weather Service (NWS) uses SLOSH (Sea, Lake, and Overland Surges from Hurricanes) as its official operational forecast model [*Jelesnianski et al.*, 1992]. Developed in the late 1960s and formalized in the 1980s, SLOSH uses a best track file to inform its own internal wind model and is applied via an ensemble approach on small structured curvilinear meshes overlapping the coastline [*Forbes and Rhome*, 2012; *Glahn et al.*, 2009; *Taylor and Glahn*, 2008]. While there have been improvements to the SLOSH model in terms of features such as barriers and 1-D channels, the core of the computational model has seen very little change since its inception. Recent studies, including this one and studies by *Blain et al.* [1994] and *Morey et al.* [2006], have noted the deficiency of using small domains that cannot capture large-scale processes and suggest using larger meshes for improved accuracy. In addition, the use of a structured mesh, regardless if it is curvilinear, limits the ability for localized resolution; and therefore potentially hampers accuracy in SLOSH.

[5] In the last decade, there has been rapid development and application of unstructured mesh circulation and wave models. Unlike structured models, unstructured models have the advantage of being able to provide coarser resolution at the Gulf scale and finer resolution at the coastal floodplain and channel scale without unnecessary and costly resolution throughout the Gulf; thus unstructured mesh models can simultaneously address the domain size and mesh resolution issues that have hampered the accuracy of the SLOSH model. A number of unstructured mesh coastal circulation and wave models have recently come into widespread use. The United States Army Corps of Engineers (USACE) used the ADvanced CIRCulation (ADCIRC) model [*Luettich et al.*, 1992; *Westerink et al.*, 2008] to evaluate and design improvements for the New Orleans Flood Risk Reduction System and the Federal Emergency Management Agency (FEMA) applies this model to evaluate coastal flood risk which is used to develop flood insurance rate maps [*USACE*, 2009; *FEMA*, 2009]. There has been a recent progression toward operational applications of prevalent unstructured mesh circulation and/or wave models including: the tightly coupled Simulating WAves Nearshore (SWAN) [*Booij et al.*, 1999; *Zijlema*, 2010] and ADCIRC model (SWAN+ADCIRC) [*Dietrich et al.*, 2011a]; the Finite Volume Coastal Ocean Model (FVCOM) [*Chen et al.*, 2003]; the Semi-implicit Eulerian Lagrangian Finite Element model (SELFE) [*Zhang and Baptista*, 2008]; and the tightly coupled Wind Wave Model II (WWMII) and SELFE models (WWMII+SELFE) [*Roland et al.*, 2009, 2012].

[6] The ADCIRC-based Extratropical Surge and Tide Operational Forecast System model (ESTOFS) is used to forecast tides and extratropical surges by the National Weather Service [http://www.opc.ncep.noaa.gov/estofs/estofs_surge_info.shtml]. The SWAN+ADCIRC model is used by the Advanced Surge Guidance System (ASGS), which focuses on providing operational advisory services for impending hurricane events [*Fleming et al.*, 2008]. The scope of ASGS currently includes the Gulf of Mexico and Atlantic seaboard, and it has been used to provide forecasting for Hurricanes Irene (2011), Isaac (2012), and Sandy (2012), as well as the Gulf oil spill [*Dietrich et al.*, 2012a]. FVCOM is a component of the Northeast Coastal Ocean

Forecast System (NECOFS) and is used extensively to forecast tides and extratropical storm events along the northeast coast of the United States, in particular within the Gulf of Maine region [http://www.neracoos.org/datatools/forecast/oceanforecasts]. The SELFE model is currently in use for forecasting in the Columbia River Estuary as the core ocean prediction model for the Columbia River Estuary Operational Forecast System (CREOFS) [http://tidesandcurrents.noaa.gov/ofs/creofs/creofs.html].

[7] One aim of the Southeastern Universities Research Association (SURA)-led U.S. Integrated Ocean Observing System (IOOS)-funded Coastal and Ocean Modeling Testbed (COMT) is to evaluate and improve models already in operational use as well as facilitating the transition of additional models to operational use. This inter-model comparison was initiated as part of that objective, so that the skill and behavior of leading coastal and ocean models at simulating tropical cyclone waves and surge could be assessed and the implementation requirements necessary for these models to migrate to operational use be identified. With that purpose in mind, three widely used unstructured coastal and ocean circulation models, ADCIRC, FVCOM, and SELFE, were selected for use on a new common Gulf of Mexico mesh created specifically for the Testbed. Waves are an integral aspect of hurricanes that can contribute to water levels, coastal erosion, and forces on structures, so simulations were also performed with tightly coupled wave and circulation models such as SWAN+ADCIRC and WWMII+SELFE and FVCOM loosely coupled [*Huang et al.*, 2010] with SWAN. In addition to the unstructured models, NOAA's official operational forecast storm surge model, SLOSH, was also included in this study.

[8] To ensure uniformity in the inter-model comparison, the unstructured models were implemented with identical frictional parameterization and forcing for each of the study's simulations. For consistency, SLOSH was also simulated with the same forcing as the unstructured models, but due to model constraints continued to use its own internal bottom friction formulation. Whereas the unstructured models were run on the same Gulf scale mesh, SLOSH used its local and Gulf scale meshes so that the effects of domain size for SLOSH could be formally addressed. Behavioral response to domain size and mesh selection is important to understand in order to improve SLOSH's operational implementation, especially considering the recommendations presented by *Morey et al.* [2006].

[9] Keeping with the objective of operational forecasting potential, models included in this study were benchmarked on the Texas Advanced Computing Cluster's (TACC) Ranger to assess and compare execution speed and scalability for a prototypical operational simulation. In addition, the number of computational cores needed for the unstructured mesh models to match SLOSH's operational execution speed (i.e., wall clock time) was estimated.

[10] Tides contribute to overall water levels and currents during a hurricane, and therefore this study included an inter-model tidal comparison to assess the ability for each of these models to simulate the relatively low-energy processes of tidal dynamics in the Gulf. Tides propagate into the Gulf of Mexico from the Atlantic Ocean through the Straits of Florida and from the Caribbean Sea through the Yucatan Channel and are also generated internally in the

Gulf through gravitational tidal potential forcing [*Westerink et al.*, 2008]. The dominant tidal signals in the basin are diurnal; however, resonance along the continental shelves adjacent to Florida, western Louisiana, and Texas can locally amplify the semidiurnal tides [*Reid*, 1990; *DiMarco and Reid*, 1998; *Gouillon et al.*, 2010]. Tidal validation studies have been conducted individually for ADCIRC [*Mukai et al.*, 2002; *Bunya et al.*, 2010], for FVCOM [*Chen et al.*, 2011], and for SELFE [*Bertin et al.*, 2012; *Cho et al.*, 2012]; however, no study has yet comprehensively compared the skill of these three models for the same mesh and tidal forcing. Tides-only simulations were conducted for ADCIRC, FVCOM, and SELFE, and the results of the tidal harmonic decomposition were then compared with tidal harmonic constituent data from 59 NOAA stations located throughout the Gulf of Mexico. It was not possible to include SLOSH in the tidal validation since SLOSH is unable to simulate tides itself; instead tides are generated within SLOSH by a synthesis of tidal constituent data gathered from an ADCIRC-based National Ocean Service *EC2001* database [*Mukai et al.*, 2002].

[11] For the inter-model evaluation of skill and behavior in simulating Atlantic-based tropical cyclones, two Gulf of Mexico hurricanes were selected for analysis: Hurricane Rita (2005) and Hurricane Ike (2008). The tracks of these two hurricanes along with many of the points of interest described in this study can be found in Figures 1 and 2 and Table 1. These storms were selected for their large-scale impacts on the Louisiana-Texas coastline, unique individual characteristics, and wealth of recorded wave and water level data. Ike, while of similar strength, track, and forward speed as Rita, exhibited significantly different surge processes due to its large size. Ike's steady track across the Gulf combined with its unusually large wind field created a strong shore-parallel shelf current driven by shore-parallel winds in the 36 h before landfall. This current caused the largest ever recorded geostrophic setup or "forerunner" surge along the Texas-Louisiana coast filling coastal lakes and bays by up to 2 m before Ike made landfall [*Kennedy et al.*, 2011]. In addition to the "forerunner," Ike's relevant surge processes included: the generation of surge by steady, moderate winds in southeastern Louisiana and the influence (or lack thereof) of marshes in the region in the impedance of storm surge over extended time scales; the associated capture of surge by the protruding Mississippi River Delta and the surface gradient and associated currents created along the edge of Delta; the propagation of a coherent free wave down the Texas continental shelf; peak storm surge driven by strong shore normal winds; and a trapped resonant wave on the continental shelf caused by the recession of surge onto the shelf [*Hope et al.*, 2013]. Circulation models, ADCIRC, FVCOM, SELFE, and SLOSH, along with the tightly coupled circulation and wave models, SWAN+ADCIRC and WWMII+SELFE, and the loosely coupled FVCOM with SWAN, were applied to hindcast these storms.

2. Methods

2.1. Computational Wave and Current Models

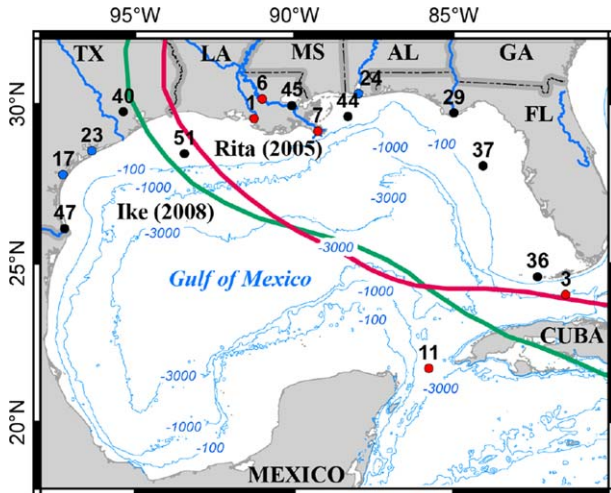[12] This study compares the behavior and performance of four circulation models (ADCIRC, FVCOM, SELFE,

Figure 1. Detail of Gulf of Mexico and tracks of Ike (green) and Rita (red). Geographic location by type and number identified in Table 1.

and SLOSH) and two wave models (SWAN and WWMII), whose characteristics are summarized in Table 2 and briefly described in this section.

2.1.1. ADCIRC

[13] The ADvanced CIRCulation Model (ADCIRC) developed by *Luettich et al.* [1992] and *Luettich and Westerink* [2004] and subsequently improved [*Atkinson et al.*, 2004a, 2004b; *Dawson et al.*, 2006; *Westerink et al.*, 2008; *Dietrich et al.*, 2011a; *Martyr et al.*, 2013] is a continuous-Galerkin finite element code that solves the Generalized Wave Continuity Equation (GWCE), which is a modified form of the Shallow Water Equations (SWE), along with the SWE momentum equations on an unstructured triangular mesh in a Cartesian or spherical coordinate system [*Kolar et al.*, 1994; *Luettich and Westerink*, 2004; *Dawson et al.*, 2006]. Advancement of the solution in time can be computed using either a semi-implicit or explicit time stepping algorithm [*Tanaka et al.*, 2011; *Dietrich et al.*, 2012b]. Code details can be found at [http://www.adcirc.org/].

2.1.2. FVCOM

[14] The Finite Volume Coastal Ocean Model (FVCOM), developed by *Chen et al.* [2003] and subsequently improved and upgraded [*Chen et al.*, 2006a, 2006b, 2007, 2008; *Lai et al.*, 2010a, 2010b; *Huang et al.*, 2008; *Chen et al.*, 2011], solves the 3-D shallow water equations on an unstructured triangular mesh using a finite volume method. An integral form of the momentum equations are solved using a second-order approximate finite volume discretization using either a Cartesian [*Chen et al.*, 2003] or spherical [*Chen et al.*, 2006b, 2011] coordinate system. Advancement in time uses the option of a mode-split solver or a semi-implicit solver [*Chen et al.*, 2009; *Lai et al.*, 2010a, 2010b; *Gao et al.*, 2011]. FVCOM has been broadly applied to simulate the 3-D baroclinic circulations in estuaries [*Weisberg and Zheng*, 2006a; *Yang and Khangaonkar*, 2008; *Zheng and Weisberg*, 2010], continental shelves [*Rego and Li*, 2010a; *Zheng and Weisberg*, 2012], and has also been applied to study storm surges [*Weisberg and Zheng*, 2006b, 2008; *Rego and Li*, 2010a]. More information can be found at http://fvcom.smast.umassd.edu/FVCOM/index.html.

2.1.3. SELFE

[15] The Semi-implicit Eulerian-Lagrangian Finite Element (SELFE) model developed by *Zhang and Baptista* [2008] and subsequently modified and improved [*Burla et al.*, 2010; *Bertin et al.*, 2009; *Brovchenko et al.*, 2011; *Pinto et al.*, 2012] solves the 3-D shallow water equations on an unstructured triangular mesh on a Cartesian or spherical coordinate system. The model uses a finite element/volume approach in conjunction with an Euler-Lagrangian method for advection terms. Advancement in time is calculated using a semi-implicit time stepping algorithm. More information can be found at http://www.stccmop.org/CORIE/modeling/selfe/ and http://ccrm.vims.edu/w/index.php/Main_Page.

2.1.4. SLOSH

[16] The Sea, Lake and Overland Surges from Hurricanes (SLOSH) model solves a simplified version of the shallow water equations, termed the transport equations, which neglects the nonlinear advection terms [*Jelesnianski*, 1966] and solves its own bottom friction formulation [*Jelesnianski*, 1967]. It uses an internal wind model to solve for meteorological forcing using a track file that
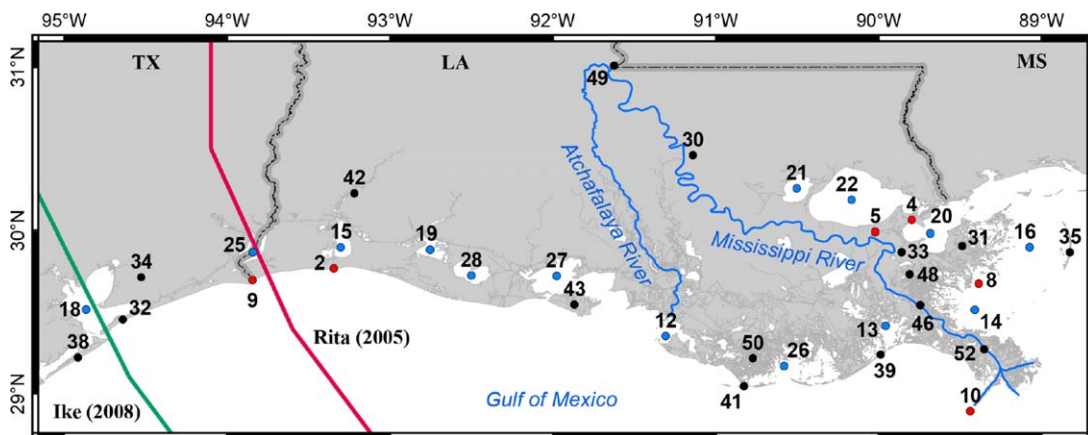

Figure 2. Detail of northern Gulf of Mexico coast. Geographic location by type and number identified in Table 1.

Table 1. Geographic Location by Type and Number Shown in Figures 1 and 2

River and Channels

1	Atchafalaya River
2	Calcasieu Shipping Channel
3	Florida Straits
4	Gulf Intracoastal Waterway (GIWW)
5	Inner Harbor Navigation Canal (IHNC)
6	Mississippi River
7	Mississippi River Bird's foot
8	Mississippi River Gulf Outlet (MRGO)
9	Sabine Pass
10	Southwest Pass
11	Yucatan Channel

Bays, Lakes, and Sounds

12	Atchafalaya Bay
13	Barataria Bay
14	Breton Sound
15	Calcasieu Lake
16	Chandeleur Sound
17	Corpus Christi Bay
18	Galveston Bay
19	Grand Lake
20	Lake Borgne
21	Lake Maurepas
22	Lake Pontchartrain
23	Matagorda Bay
24	Mobile Bay
25	Sabine Lake
26	Terrebonne Bay
27	Vermillion Bay
28	White Lake

Places

29	Apachicola, FL
30	Baton Rouge, LA
31	Biloxi Marsh
32	Bolivar Peninsula
33	Caernarvon Marsh
34	Chambers County, TX
35	Chandeleur Islands
36	Florida Keys
37	Florida shelf
38	Galveston Island
39	Grand Isle
40	Houston, TX
41	Isles Dernieres, LA
42	Lake Charles, TX
43	Marsh Island
44	Mississippi-Alabama shelf
45	New Orleans, LA
46	Pointe à La Hache
47	Port Isabel, TX
48	St. Bernard, LA
49	Tarbert Landing
50	Terrebonne Marsh
51	Texas-Louisiana shelf
52	Venice, LA

includes the location of the storm, differential atmospheric pressure, forward speed and radius to max winds. For this study, a modified version of SLOSH was developed to accept external meteorological forcing. The governing equations are solved explicitly in time with a finite difference method in space on an orthogonal curvilinear mesh [*Jelesnianski et al.*, 1992]. The model uses a one-

dimensional derivation of the equations to capture the dynamics of channel flow within the domain and is implemented with internal barriers to account for levees.

2.1.5. SWAN

[17] The Simulating WAves Nearshore (SWAN) model is a third generation, phase-averaged spectral wave model that simulates the evolution of wind generated waves in the open ocean, coastal regions, and within inland waters [*Booij et al.*, 1999; *Ris et al.*, 1999]. The wave action balance equation is solved using a finite difference solution on either a regular, curvilinear, or unstructured triangular mesh [*Zijlema*, 2010]. Solution advancement in time is resolved using an implicit backward difference scheme, with the subsequent system of equations solved using a sweeping Gauss-Seidel method.

2.1.6. WWMII

[18] The Wind Wave Model II (WWMII) is a third-generation phase-averaged spectral wave model that solves the wave action equation on an unstructured mesh. A fractional step method [*Yanenko*, 1972] is employed to split the resulting multidimensional equation into well-defined differential equations. The resulting numerical algorithm involves a combination of residual distribution schemes in geographical space and higher order difference schemes in spectral space, as used in WAVEWATCH III® [*Tolman*, 1999]. For geographical advection, implicit and explicit residual distribution schemes are used with various time discretization methods. In this study, first-order explicit schemes in geographical space have been used, details of which can be found in *Roland et al.* [2009].

2.1.7. Wave Coupling

[19] ADCIRC, FVCOM, and SELFE can receive wave radiation stress gradients as an input from a wave model, while the version of SLOSH employed here cannot; therefore only ADCIRC, FVCOM, and SELFE were coupled with wave models. Wave radiation stresses are determined within the wave models according to the theory of *Longuet-Higgins and Stewart* [1964] and passed to the hydrodynamic circulation models to update the current velocities and water levels. In turn, water levels and currents are passed back to the wave models and used to solve the wave action balance equation. SWAN+ADCIRC [*Dietrich et al.*, 2011a] is the tightly coupled version of SWAN and ADCIRC and WWMII+SELFE [*Roland et al.*, 2009, 2012] is the tightly coupled version of WWMII and SELFE, which means that these wave and circulation models run and exchange data synchronously on the same mesh and have both models embedded directly within the code [*Roland et al.*, 2009, 2012; *Dietrich et al.*, 2011a]. SWAN-FVCOM is the loose coupling of SWAN and FVCOM [*Huang et al.*, 2010, 2013]. The + symbol is used to denote tight coupling of the models, as opposed to the − symbol which is used to denote loose coupling. For this study, the loose coupling of SWAN-FVCOM required: (1) running FVCOM, (2) running SWAN with water level and current output from FVCOM, and (3) rerunning FVCOM with all forcings including the wave radiation stress gradient output from SWAN.

2.2. Meshes

[20] ADCIRC, FVCOM, SELFE, SWAN, and WWMII were run on unstructured triangular meshes. The ULtralite-Levee-Removed (*ULLR*) mesh, a 417,642 nodes (826,866

Table 2. Computational Model Characteristics

Characteristic	ADCIRC	FVCOM	SELFE	SLOSH	SWAN	WWMII
Solves	Shallow water equations (modified in some cases)				Spectral action balance	
Numerical algorithm	Finite element continuous galerkin	Finite volume	Finite volume Eulerian-Lagrangian	Finite difference	Finite difference	Finite element
Coordinate system	Cartesian or spherical			Continuously varying polar	Cartesian or spherical	
Horizontal mesh	Triangular unstructured			Curvilinear orthogonal	Regular, curvilinear or unstructured triangular	Unstructured triangular
Time stepping	Semi-implicit or explicit	Explicit	Semi-implicit	Explicit	Implicit	Fractional step
Forcing	Tidal potential Wind stress Atmospheric pressure Wave radiation stress	Tidal potential Wind stress Atmospheric pressure Wave radiation stress	Tidal potential Wind stress Atmospheric pressure Wave radiation stress	Wind stress Atmospheric pressure	Wind stress	Wind stress

elements), was created as part of the U.S. IOOS Coastal and Ocean Modeling Testbed. This mesh encompasses all of the Gulf of Mexico and has open ocean boundaries at the Straits of Florida and Yucatan Channel and extends across the entire low-lying floodplains of Texas and Louisiana (Figure 3). The mesh also has river flux boundary conditions at its domain intersections with the Atchafalaya and Mississippi Rivers, although no river fluxes were applied. The *ULLR* mesh is the final result of a series of mesh alternatives considered by the Testbed. The *ULLR* mesh was developed as a coarser version of the *SL18TX33* mesh, a high-resolution detailed mesh of the Louisiana-Texas coast, the Gulf of Mexico and Western North Atlantic [*Hope et al.*, 2013]. The *SL18TX33* mesh contains interior levee boundary features that are only incorporated by the ADCIRC model. Interior levee boundaries are a method by which ADCIRC has successfully addressed levees, a submesh-scale feature [*Westerink et al.*, 2008], which can

Figure 3. Location map of the unstructured mesh (ULLR), the three SLOSH meshes (*egm3*, *egl3*, and *ebp3*), and the tracks for Ike and Rita.

have significant implications to inundation and river surge generation [*Kerr et al.*, 2013a]. The *ULLR* mesh was created by removing these interior levee boundaries so that all unstructured models could be run on an identical mesh. The *ULLR* mesh is unstructured, which allows for mesh resolution to be strategically distributed where needed. *ULLR* resolution is coarser in the Gulf and finer inland, with higher resolution used to resolve key features such as the Mississippi and Atchafalaya Rivers (Figure 4).

[21] The *ULLR* mesh was used to simulate tides and storms with the ADCIRC, FVCOM, and SWAN models. The *ULLR-EC2001*, a modification of the *ULLR* that adds the Caribbean Sea and the eastern North Atlantic Ocean from the *EC2001* mesh [*Mukai et al.*, 2002], was used for SELFE and WWMII because of the necessity to supply SELFE with both velocity and elevation boundary conditions. The incoming boundary velocity in SELFE needs to be specified due to its large time step. Because the tracks of Hurricanes Ike and Rita featured the eye of the storm passing over the *ULLR* open ocean boundaries, it was not possible to use a predetermined boundary condition. One alternative would have been to provide SELFE with boundary conditions from another model. Instead, the *ULLR-EC2001* was created so that SELFE could use the predominately deep ocean 60°W meridian as its open ocean boundary, where it does not require a velocity boundary condition and thus not require another model's assistance.

[22] SLOSH uses structured orthogonal curvilinear meshes that can be elliptical, parabolic, or hyperbolic. In operational forecasting, the National Weather Service simulates SLOSH on local-scale meshes that overlap along the coastline. The primary meshes used in the forecasting of Ike and Rita were the *egl3* (Galveston Basin) and *ebp3* (Sabine Pass Basin), respectively (Figure 3). The *egl3* mesh lies directly in the path of Ike and the *ebp3* mesh lies directly in the path of Rita. *Blain et al.* [1994] and *Morey et al.* [2006] suggested that the shelf-scale processes of hurricanes are better simulated by meshes that exceed the local domain affected, especially within the Gulf of Mexico.
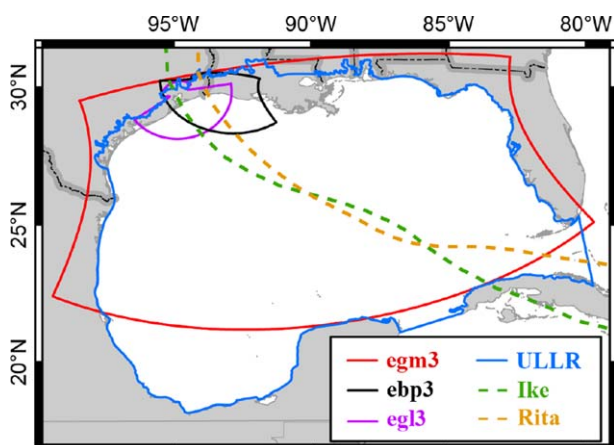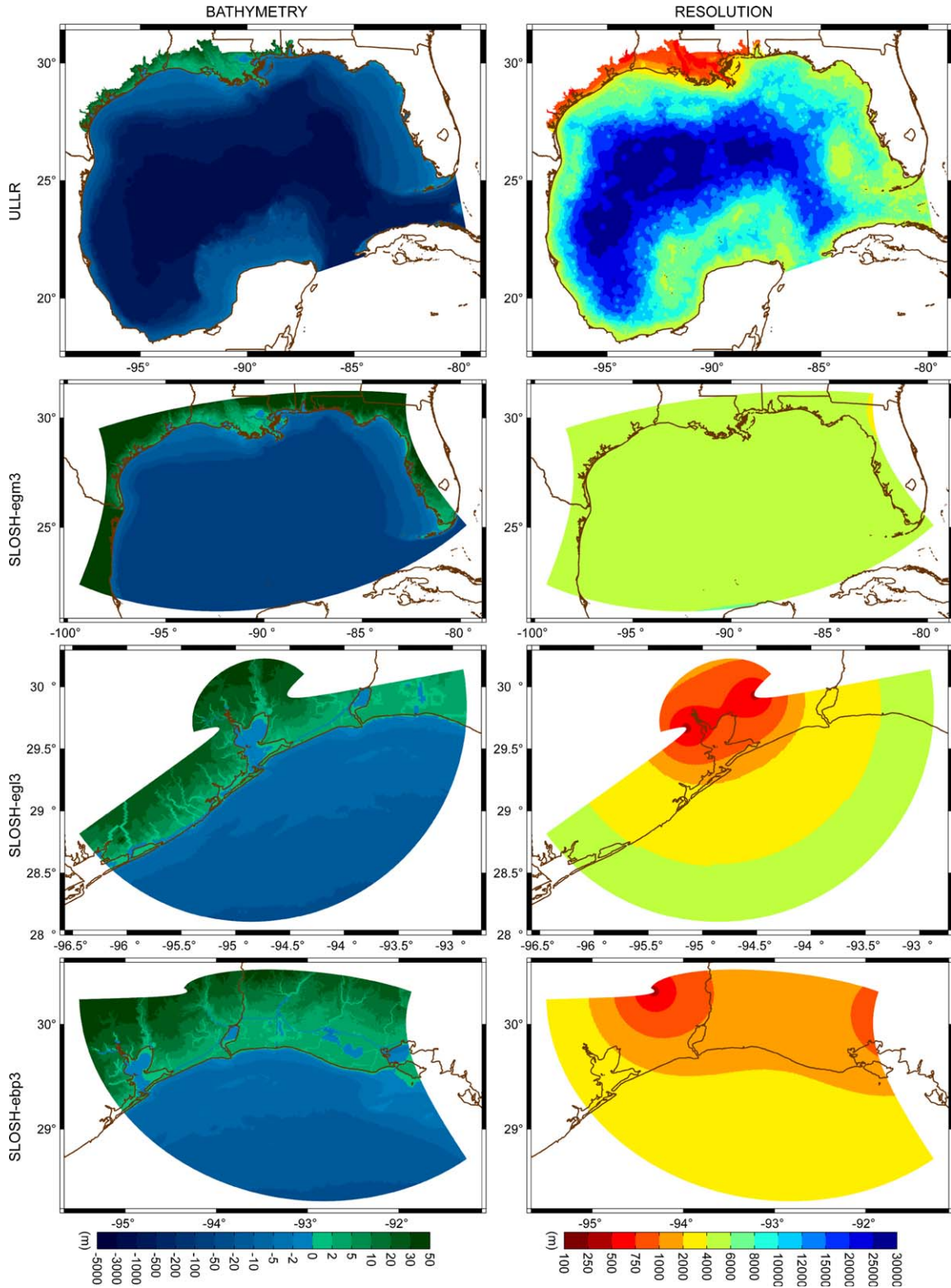
Figure 4. (left column) Mesh bathymetry in meters NAVD88–2004.65 and (right column) resolution in meters.

Therefore, the *egm3* (185,409 nodes), a Gulf wide SLOSH mesh, has been included in this study along with the local-scale SLOSH meshes to demonstrate the differences that domain can have on SLOSH behavior. Resolution of a SLOSH mesh is partially dependent on its shape, as seen in Figure 4. The *egl3* (46,222 nodes) and *ebp3* (77,827 nodes) meshes are shaped to focus higher resolution on the coastal floodplain, with resolution decreasing outward into the Gulf. The

Table 3. Principal Tidal Constituents[a]

j	n	Constituent	Name	T_{jn} (h)	B_{jn} (m)	α_{jn}	Tides		Rita (2005)		Ike (2008)	
							f_{jn}	v_{jn}	f_{jn}	v_{jn}	f_{jn}	v_{jn}
1	1	K_1	Luni-Sol. Decl.	23.934470	0.141565	0.736	1.09257	152.10	1.10985	229.51	1.09009	44.81
	2	O_1	Princ. Lun. Decl.	25.819342	0.100514	0.695	1.14990	231.92	1.17827	300.73	1.14585	347.12
	3	P_1	Princ. Sol. Decl.	24.065888	0.046843	0.706	1.00000	212.96	1.00000	128.46	1.00000	320.52
	4	Q_1	Lg. Lun. Ellipt.	26.868357	0.019256	0.695	1.14990	71.54	1.17827	15.99	1.14585	306.46
2	1	M_2	Principal Lun.	14.420601	0.242334	0.693	0.97167	26.08	0.96445	169.44	0.97268	34.11
	2	S_2	Principal Sol.	12.000000	0.112841	0.693	1.00000	0.00	1.00000	0.00	1.00000	0.00
	3	N_2	Lg. Lun. Ellipt.	12.658348	0.046398	0.693	0.97167	225.70	0.96445	244.70	0.97268	353.45
	4	K_2	Luni-Sol. Decl.	11.967235	0.030704	0.693	1.24617	124.73	1.30650	278.79	1.23784	270.17

[a]With periods (T_{jn}—hours), tidal potential constants (B_{jn}—meters), associated effective earth elasticity factors (α), time-dependent nodal factors (f_{jn}), and time-dependent equilibrium arguments (v_{jv}—degrees). Time-dependent values reflect periods of simulations found in Table 4.

egm3 mesh encompasses most of the Gulf of Mexico and has relatively constant resolution across its domain.

[23] Unstructured meshes provide the ability to apply resolution where needed; whereas structured meshes, even those that are curvilinear, are more limited, especially in rate of transition between resolution. The *ULLR* mesh's resolution ranges from 8 to 30 km in the Gulf, 2–8 km on the shelf, 500–2000 m on the floodplain, and 100–500 m in the rivers. The resolution of the *egm3* mesh is fairly constant between 4 and 6 km over the Gulf, shelf, floodplain, and channel scales, due to its structured nature; this is in contrast to the unstructured *ULLR* mesh which is coarser in the Gulf and finer inland. The *egl3* and *ebp3* meshes have more variation in their resolution than the *egm3* mesh and their resolution ranges from 500 to 4000 m, although most of the resolution at the shoreline is between 1000 and 2000 m. Comparatively, the resolution along the coastline for the *ULLR*, *egl3*, and *ebp3* meshes are similar, whereas the resolution at the coastline for the *egm3* mesh is much coarser. The overall range of resolution for each of the meshes are: *ULLR* (100–31,000 m), *egl3* (80–2700 m), *ebp3* (80–1600 m), and *egm3* (1900–4000 m).

2.3. Bathymetry, Topography

[24] The bathymetry and topography used for the development of the *ULLR* mesh matches that used by *Hope et al.* [2013] in their synoptic study of Ike and *Dietrich et al.* [2011b] in their synoptic study of Gustav. The bathymetry and topography of the *ULLR*, *ebp3*, *egl3*, and *egm3* meshes are shown in Figure 4. Depth in the *egm3* is limited to 182 m due to limits built into SLOSH. This is in stark contrast to the depth of the *ULLR*, which reaches the actual depths of over 4000 m in the center of the Gulf.

[25] Topographic and bathymetric data is referenced to the North American Vertical Datum of 1988 (NAVD88). Initial water levels are raised by 0.134 m to update elevations to the NAVD88 2004.65 epoch, and increased by an additional amount, the steric, prior to simulation. The steric reflects the seasonal variability of the of the sea level in the Gulf of Mexico, due to thermal expansion and contraction, salinity, winds, ocean currents, seasonal river runoff, and other factors. The steric for Rita was 0.146 m which results in a total adjustment of 0.28 m [*Bunya et al.*, 2010]. For Ike the steric is 0.142 m and the total adjustment is 0.276 m [*Hope et al.*, 2013].

2.4. Friction

[26] A spatially varying Manning's *n* bottom friction was applied for each of the unstructured circulation models. Manning's *n* values were derived from land-use databases as described by *Dietrich et al.* [2011b]. Capturing friction correctly on the continental shelf is essential to developing the shore-parallel current that preceded Ike and caused the geostrophic setup [*Kennedy et al.*, 2011; *Kerr et al.*, 2013b; *Hope et al.*, 2013]. This required some of the models to remove lower drag coefficient limits from their bottom friction formulations, so that Ike's forerunner could develop. Unlike the unstructured mesh model's implementation of the Manning's formula, SLOSH uses its own internal bottom friction formulation which is mildly depth dependent and is not spatially varying.

[27] *Jelesnianski* [1966, 1967] studied the effects of including or not including a bottom stress formulation in SLOSH, and concluded that a dissipation mechanism was essential for modeling slow-moving storms, shore-parallel storms, or storms with extended duration on the continental shelf, because it was necessary to control large amplitude resurgences and/or initialization phenomena that were caused by the case with no bottom friction. Thus friction is partially used in the SLOSH model to control numerical stability and initialization phenomena with its internal parameters empirically set to match peak surge levels.

2.5. Tidal Forcing

[28] The principal tidal constituents forced in this study are K_1, O_1, P_1, Q_1, M_2, K_2, N_2, and S_2. The associated periods, tidal potential constants, earth elasticity factors, and time-dependent equilibrium arguments, as shown in Table 3, were obtained or derived from *Wahr* [1981], *Foreman et al.* [1993], *Reid* [1990], and *Schureman* [1958]. The time-dependent nodal factors and equilibrium arguments shown in Table 3 are for a 105 day simulation beginning on 0000 universal time coordinated (UTC) 19 May 2008, which is the time frame used for the tidal harmonic analyses. The time-dependent coefficients used for the Ike and Rita simulations are also listed in Table 3.

[29] The *ULLR* has open ocean boundaries along the Yucatan Channel and Straits of Florida. The principal tidal constituents, K_1, O_1, Q_1, M_2, K_2, N_2, and S_2 were generated from the *EC2001* tidal database [*Mukai et al.*, 2002]. The P_1 constituent, which is lacking from the *EC2001* database, was

generated by averaging the *FES2004* [*Lyard et al.*, 2006] and *TPXO7.2* [*Egbert et al.*, 1994; *Egbert and Erofeeva*, 2002] tidal atlases. The *ULLR-EC2001* open ocean boundary, which is located along the 60°W meridian, was forced with the K_1, O_1, P_1, Q_1, M_2, K_2, N_2, and S_2 tidal constituents generated from the *FES95.2* tidal atlas [*Le Provost et al.*, 1994, 1995].

[30] SLOSH does not simulate tides, so there was no tidal forcing to apply. For the hurricane analyses, tidal signals from ADCIRC were superimposed onto SLOSH water levels. For each cell and output step, the water level from the nearest ADCIRC node to the center of each SLOSH cell was added to that SLOSH cell's water level. This method is not ideal due to the difference in meshes, and is further compounded by the difference in resolution, with the SLOSH meshes being typically coarser than the *ULLR*. At several stations spatial misalignment between the SLOSH cell center point and the actual station was apparent. This was noticed because SLOSH-interpolated results had an erroneously high tidal signal, whereas at the same station, ADCIRC had a weaker but more correct tidal signal. The superposition of the tidal signal on the surge level does not replicate the nonlinear and mass contributions of tidal dynamics that are included in the unstructured mesh models that dynamically simulate both the tides and surge.

2.6. Meteorological Forcing

[31] OceanWeather Inc. (OWI) structured and data-assimilated wind and pressure fields were used in this study to hindcast Rita [*Bunya et al.*, 2010] and Ike [*Hope et al.*, 2013]. As described by *Bunya et al.* [2010], hindcast wind fields were defined objectively using analyzed measurements from anemometers, airborne and land-based Doppler radar, airborne stepped-frequency microwave radiometers, buoys, ships, aircraft, coastal stations, and satellite measurements. Hindcast winds are based on the blending of an inner core wind field (transformed to 30 min averaged sustained winds), such as the TC96 mesoscale models [*Thompson and Cardone*, 1996] solutions for Rita and the NOAA Hurricane Research Wind Analysis System (H*WIND) [*Powell et al.*, 1996, 1998, 2010] for Ike, with Gulf scale winds using the Interactive Objective Kinematic Analysis (IOKA) system [*Cox et al.*, 1995; *Cardone and Cox*, 2009].

[32] H*WIND analyses had a 3 h frequency and included the use of improved terrain conversions [*Vickery et al.*, 2009] and high-resolution tower data from Texas Tech University and the Florida Coastal Monitoring Program. The H*WIND analyses of Ike also included the deployment of stepped-frequency microwave radiometers aboard the Air Force Hurricane Hunter Aircraft [*Uhlhorn et al.*, 2007], which increased the availability of high radial resolution surface winds since the Katrina wind field post analysis [*Ebersole et al.*, 2007]. Peripheral winds were derived from the National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) reanalysis project [*Kalnay et al.*, 1996]. The IOKA system also includes the injection of local marine data, adjustment to a consistent 10 m elevation, 30 min sustained wind speed, marine exposure, and neutral stability. Finally, Lagrangian-based interpolation is used to produce structured wind fields every 15 min. This included for Rita a single structured grid spaced at 0.05° and for Ike a Gulf scale 0.1° grid and a local scale 0.015° grid near landfall. Sea

level atmospheric pressure fields for both Rita and Ike were derived using the widely adopted radial-based parametric model developed by *Holland* [1980] and assigned to the same structured grids as the wind fields.

[33] The methodology used by ADCIRC to employ an OWI wind field includes: (1) converting 30 min averaged sustained winds to 10 min averaged sustained winds; (2) accounting for canopy cover; (3) applying a directional reduction factor to account for surface roughness [*Bunya et al.*, 2010]; and (4) increasing winds to full marine winds as roughness elements are inundated. In this study, a sector-based directional wind drag law, which is data driven and wind speed limited, is applied in ADCIRC to compute its drag coefficient. The drag law was developed by *Powell et al.* [2003] and *Powell* [2006], and previously applied by *Dietrich et al.* [2011b] and *Hope et al.* [2013]. FVCOM, SELFE, and SLOSH models do not include these sophisticated wind reduction factors, so in order to perform a consistent comparison, wind stress (not wind velocity) was output from an ADCIRC hindcast on the *SL18TX33* mesh [*Hope et al.*, 2013] and converted to a structured format so that it could be applied to SELFE, FVCOM, and SLOSH. Each model then employed approximately the same wind stress and pressure, therefore essentially using the wind stress physics built into ADCIRC.

[34] The wave models, SWAN and WWMII, were supplied with the OWI 30 min winds. SWAN+ADCIRC used the sector-based drag law to compute its wind-wave drag coefficient, whereas the wind-wave computations in SWAN-FVCOM and WWMII+SELFE applied Wu's drag law [*Wu et al.*, 1982] with a modified cap for the SWAN-FVCOM simulations [*Huang et al.*, 2013].

2.7. Wave Parameters

[35] Identical wave parameters were used in SWAN and WWMII. Wave direction was discretized into 36 regular bins and wave frequency was logarithmically distributed into 50 bins with a range of 0.035–0.9635 Hz. The Komen formulation as modified by *Rogers et al.* [2003] was used for white capping. Wave breaking due to depth was determined spectrally according to the model of *Battjes and Janssen* [1978] with a breaking parameter of $\gamma = 0.73$ [*Battjes and Stive*, 1985]. For SWAN+ADCIRC, the spectral and directional speeds were set with a Courant-Friedrichs-Lewy (CFL) condition of 0.25 to limit spurious refractions [*Dietrich et al.*, 2012c]. Bottom friction is applied via a Joint North Sea Wave Project (JONSWAP) formulation with a $C_{fjon} = 0.019 \mathrm{m}^2/\mathrm{s}^3$, as recommended by *Kerr et al.* [2013b] for the muddy bottom of the LATEX shelf.

2.8. Summary of TESTBED Simulations

[36] A total of 21 simulations were performed as part of this study (Table 4). The inter-model comparison included tidal harmonic analyses and hindcasts of Rita and Ike. The tidal simulations were performed using ADCIRC, FVCOM, and SELFE. The hurricane hindcasts were performed using circulation models (ADCIRC, FVCOM, SELFE, and SLOSH) and coupled wave and circulation models (SWAN+ADCIRC, SWAN-FVCOM, and WWMII+SELFE). The unstructured models were all run on the *ULLR* mesh (the exception being *ULLR-EC2001* for SELFE and WWMII+SELFE); whereas SLOSH was run

Table 4. Simulations Performed

ID	Simulation	Model	Mesh	Start to End	Meteorology
1	Tidal Harmonics	ADCIRC	*ULLR*	0000 UTC 19 May 2008 to 0000 UTC 1 Sep 2008	None
2	(2008)	FVCOM	*ULLR*	0000 UTC 19 May 2008 to 0000 UTC 1 Sep 2008	None
3		SELFE	*ULLR-EC2001*	0000 UTC 19 May 2008 to 0000 UTC 1 Sep 2008	None
4	Hurricane	ADCIRC	*ULLR*	0000 UTC 13 Aug 2005 to 0000 UTC 30 Sep 2005	None
5	Rita (2005)	ADCIRC	*ULLR*	0000 UTC 13 Aug 2005 to 0000 UTC 30 Sep 2005	0000 UTC 18 Sep 2005 to 0000 UTC 25 Sep 2005
6		FVCOM	*ULLR*	0000 UTC 13 Aug 2005 to 0000 UTC 30 Sep 2005	0000 UTC 18 Sep 2005 to 0000 UTC 25 Sep 2005
7		SELFE	*ULLR-EC2001*	0000 UTC 13 Aug 2005 to 0000 UTC 30 Sep 2005	0000 UTC 18 Sep 2005 to 0000 UTC 25 Sep 2005
8		SWAN+ADCIRC	*ULLR*	0000 UTC 13 Aug 2005 to 0000 UTC 30 Sep 2005	0000 UTC 18 Sep 2005 to 0000 UTC 25 Sep 2005
9		SWAN-FVCOM	*ULLR*	0000 UTC 13 Aug 2005 to 0000 UTC 30 Sep 2005	0000 UTC 18 Sep 2005 to 0000 UTC 25 Sep 2005
10		WWMII+SELFE	*ULLR-EC2001*	0000 UTC 13 Aug 2005 to 0000 UTC 30 Sep 2005	0000 UTC 18 Sep 2005 to 0000 UTC 25 Sep 2005
11		SLOSH	*ebp3*	0900 UTC 21 Sep 2005 to 1100 UTC 25 Sep 2005	0900 UTC 21 Sep 2005 to 1100 UTC 25 Sep 2005
12		SLOSH	*egm3*	0900 UTC 21 Sep 2005 to 1100 UTC 25 Sep 2005	0900 UTC 21 Sep 2005 to 1100 UTC 25 Sep 2005
13	Hurricane	ADCIRC	*ULLR*	1200 UTC 31 Jul 2008 to 1200 UTC 17 Sep 2008	None
14	Ike (2008)	ADCIRC	*ULLR*	1200 UTC 31 Jul 2008 to 1200 UTC 17 Sep 2008	0000 UTC 10 Sep 2008 to 0000 UTC 16 Sep 2008
15		FVCOM	*ULLR*	1200 UTC 31 Jul 2008 to 1200 UTC 17 Sep 2008	0000 UTC 10 Sep 2008 to 0000 UTC 16 Sep 2008
16		SELFE	*ULLR-EC2001*	1200 UTC 31 Jul 2008 to 1200 UTC 17 Sep 2008	0000 UTC 10 Sep 2008 to 0000 UTC 16 Sep 2008
17		SWAN+ADCIRC	*ULLR*	1200 UTC 31 Jul 2008 to 1200 UTC 17 Sep 2008	0000 UTC 10 Sep 2008 to 0000 UTC 16 Sep 2008
18		SWAN-FVCOM	*ULLR*	1200 UTC 31 Jul 2008 to 1200 UTC 17 Sep 2008	0000 UTC 10 Sep 2008 to 0000 UTC 16 Sep 2008
19		WWMII+SELFE	*ULLR-EC2001*	1200 UTC 31 Jul 2008 to 1200 UTC 17 Sep 2008	0000 UTC 10 Sep 2008 to 0000 UTC 16 Sep 2008
20		SLOSH	*egm3*	0900 UTC 20 Sep 2008 to 1100 UTC 14 Sep 2008	0900 UTC 20 Sep 08 to 1100 UTC 14 Sep 2008
21		SLOSH	*egm3*	0900 UTC 20 Sep 2008 to 1100 UTC 14 Sep 2008	0900 UTC 20 Sep 08 to 1100 UTC 14 Sep 2008

on its local-scale mesh (*egl3* for Ike, and *ebp3* for Rita) and its Gulf scale mesh (*egm3*) for each hindcast. Because tides are not simulated in SLOSH, and because tides contribute to overall water levels in storms, a tides-only simulation (i.e., without meteorological forcing) was performed for Rita and Ike using the *ULLR* mesh using ADCIRC. The results of these simulations were superimposed onto SLOSH results, so that the SLOSH runs could be more fairly compared to the unstructured models.

2.9. Skill Metrics

[37] Model performance was quantified using the following skill metrics: Coefficient of Determination (R^2, which describes how well a regression line fits a set of data, with an ideal value of one), Root Mean Square Error (E_{RMS}, which is a measure of the magnitude of error, with an ideal value of equaling zero), Mean Error ($\overline{E}$), Slope of the best fit line (m, with an ideal value of one), Mean Normalized Bias (B_{MN}, which is a measure of the model's magnitude of overprediction or underprediction normalized to the observed value, with an ideal value of equaling zero),

Standard Deviation of the error in the model calculation (σ), Scatter Index (*SI*, which is the standard deviation normalized by the mean observed value, with an ideal value of equaling zero), Mean Absolute Error (*MAE*), and Mean Normalized Error (E_{NORM}, which is the mean error normalized by the mean observed value, with an ideal value of zero). Where O is the observed value, E is the error in terms of modeled minus observed, and N is the number of data points, the equation for Mean Normalized Bias is

$$B_{MN} = \frac{\frac{1}{N}\sum_{i=1}^{N} E_i}{\frac{1}{N}\sum_{i=1}^{N} |O_i|} \tag{1}$$

and the equation for Mean Normalized Error is

$$E_{NORM} = \sqrt{\frac{\frac{1}{N}\sum_{i=1}^{N} (E_i)^2}{\frac{1}{N}\sum_{i=1}^{N} (O_i)^2}} \tag{2}$$

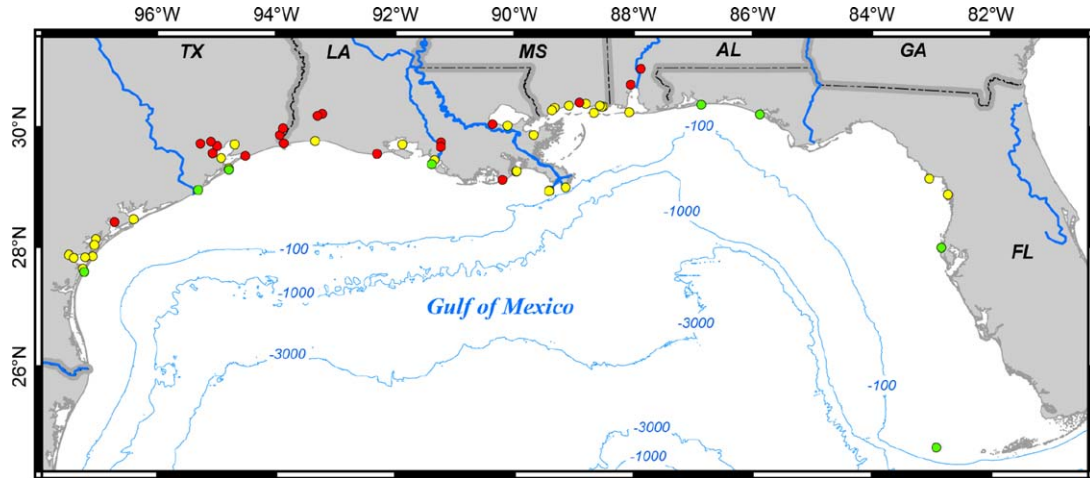and the equation for Scatter Index is

Figure 5. Location map of NOAA Tidal Observation Stations, classified as Open Coastal (Red), Protected (Green), and Inland (Blue).

$$SI = \sqrt{\frac{\frac{1}{N}\sum_{i=1}^{N}\left(E_i - \overline{E}\right)^2}{\frac{1}{N}\sum_{i=1}^{N}|O_i|}} \qquad (3)$$

[38] In the statistical analyses of water levels and high water marks, it was found that the number of stations and points in the time series which wetted varied for all the models. In some previous hurricane validation studies, dry stations and dry time series points were omitted. While this method may be appropriate for a single model validation, it became apparent that there are problems to that method when applying it for the comparison of multiple models. One example of a problem is that if one model wetted one station but the results were poor, and another model did not wet that station at all, then the model that did wet the station would be penalized with poor statistics. Another problem is that if stations are omitted that were not wetted by all models then the richness of the statistical set would be sorely depleted.

[39] To counter the problem of a disproportionate number of wet and dry points between the models, one solution is to replace dry points with a common equalizer, the bathymetry. Because the bathymetry represents the minimum water level obtainable at any point in the mesh, dry values were replaced with the ground surface elevation, and common statistical sets were obtained. This method is referred to in this study as Topo-Substitution (TS). To compare both approaches, the statistical analyses are presented with both the TS method and the Wet-Only method. The Wet-Only method refers to the method of using only those stations or points that wetted in the statistical analysis and omitting the rest.

3. Tidal Simulations

[40] A 105 day tidal simulation was performed with ADCIRC and FVCOM on the *ULLR* mesh and with SELFE on the *ULLR-EC2001* mesh. Water level time series were collected every half hour for 59 NOAA observation stations located in Texas, Louisiana, Mississippi, Alabama, and Florida, (Figure 5). The analyzed record of the water level time series began 35 days past the simulation start date of 0000 UTC 19 May 2008, so that there was sufficient time for all stations to achieve a dynamic steady state, for a total of 75 days of data. Model data was recorded at the nearest wet node to each observed station location. In a few cases, due to coarse mesh resolution, the nearest wet node could be kilometers away from the station location. Figure S1 of the supporting information gives details of each nodal location and the distance from the observed station location.

[41] NOAA uses a 37 harmonic constituent set to decompose its observed water levels (http://tidesandcurrents.-noaa.gov/), but this study used a different set of 38 harmonic constituents (Figure S2 of the supporting information) for decomposition. This constituent decomposition set was optimized based on signal duration by the MAT-LAB program T_TIDE [*Pawlowicz et al.*, 2002]. Classic harmonic analyses were performed with T_TIDE using a least squares fit to estimate the amplitude and phase of each constituent.

[42] Statistical analyses were performed to compare the skill of each the unstructured models with respect to the eight major tidal constituents; O_1, K_1, P_1, Q_1, M_2, S_2, N_2, and K_2. Comparatively, all three models performed with similar accuracy overall (Figure 6 and Table 5), regardless of region and geographic categorization [*Kerr et al.*, 2013b], and for both diurnal (e.g., K_1) and semidiurnal (e.g., M_2) constituents, but there were notable trends with regard to amplitude. ADCIRC had a tendency to predict larger amplitudes than the other two models, while SELFE tended to predict the smallest amplitude (Figure 6). Accuracy with regard to phases was very similar for all three models. Overall the three models capture the full set of amplitudes of the eight major constituents well, with an overall R^2 value of 0.80 ± 0.01 for all the combined constituents. Semidiurnal tidal signals in particular show higher R^2 values, but this is related to the much larger range of semidiurnal amplitudes associated with strong amplification along the Florida and LATEX shelves, and not necessarily a reflection of better accuracy for semidiurnal constituents. As studied by *Kerr et al.* [2013b], tidal simulation accuracy can be significantly affected by mesh
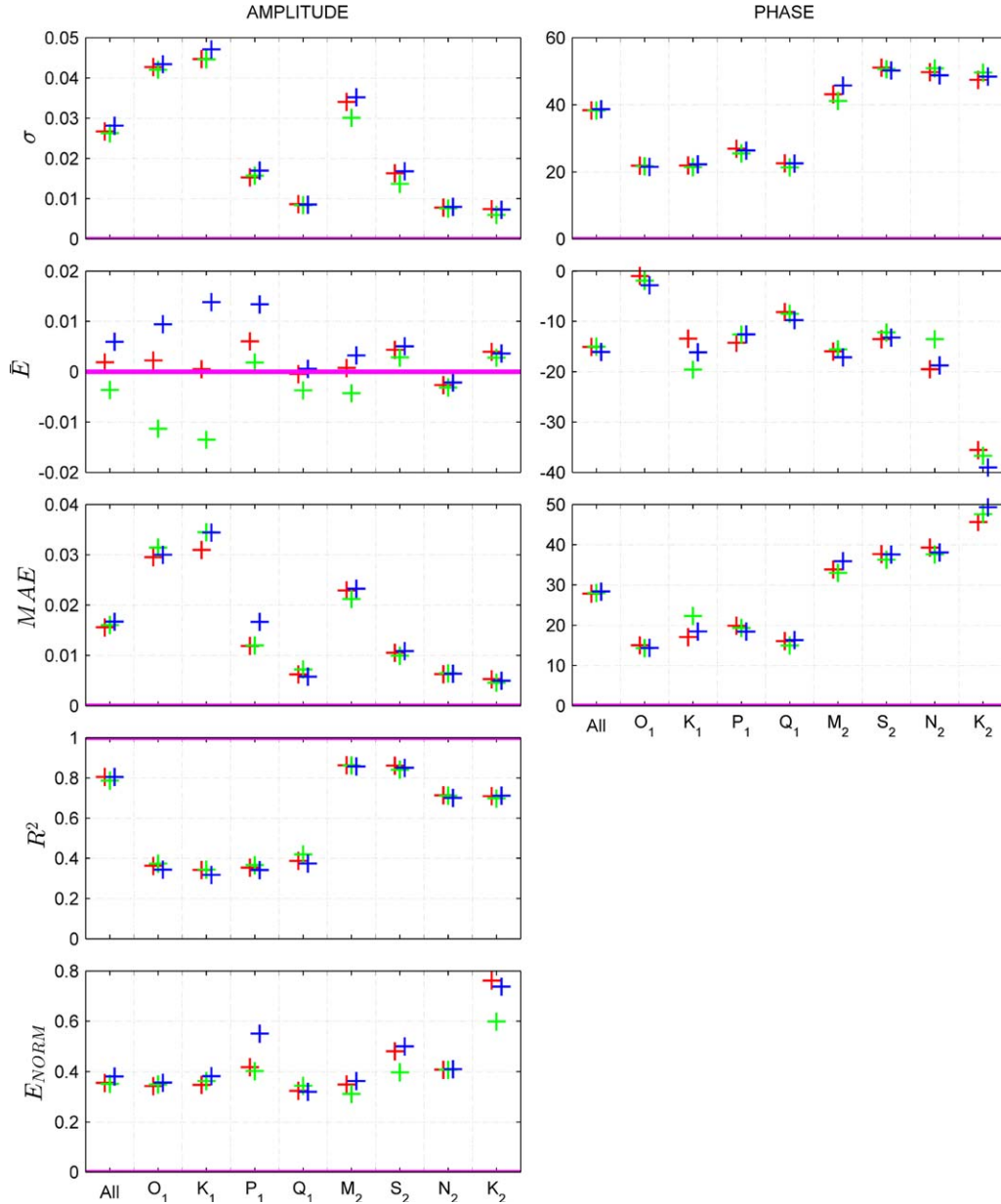
Figure 6. Amplitude and phase error statistics categorized by harmonic constituent for ADCIRC (blue), FVCOM (red), and SELFE (green) tidal validation studies.

resolution, particularly for Protected and Inland stations, or those stations where resolving topographical details is important for capturing accurate attenuation and conveyance of tidal signals.

4. Hurricane Rita (2005) Simulations and Analyses

[43] Rita entered the Gulf of Mexico via the Straits of Florida on 20 September 2005 [*Knabb et al.*, 2006]. Twenty hours prior to landfall, winds in southwestern Louisiana were northeasterly. Twelve hours prior to landfall (Figure 7a) winds in southeastern Louisiana were directed toward the northwest, allowing surge to build against the Mississippi River levee system. Southwestern Louisiana experienced strong tropical storm force northeasterly

Table 5. Tidal Harmonic Analyses Error Statistics[a]

	Amplitude			Phase		
	ADCIRC	FVCOM	SELFE	ADCIRC	FVCOM	SELFE
R^2	0.8054	0.8051	0.7872	NA	NA	NA
m	1.0684	1.0035	0.9019	NA	NA	NA
σ	0.0281	0.0267	0.0263	38.6752	38.3575	38.3117
$\bar{E}$	0.0059	0.0018	−0.0036	−16.0756	−15.0681	−15.0066
MAE	0.0167	0.0156	0.0160	28.3616	27.869	27.9955
E_{NORM}	0.3811	0.3551	0.3516	NA	NA	NA

[a]Values shown represent the "All" data set which includes the O_1, K_1, P_1, Q_1, M_2, S_2, N_2, and K_2 constituents combined. Due to the cyclical basis of phase, metrics such as correlation coefficient, slope, and mean normalized error are not applicable for phase and are not reported.
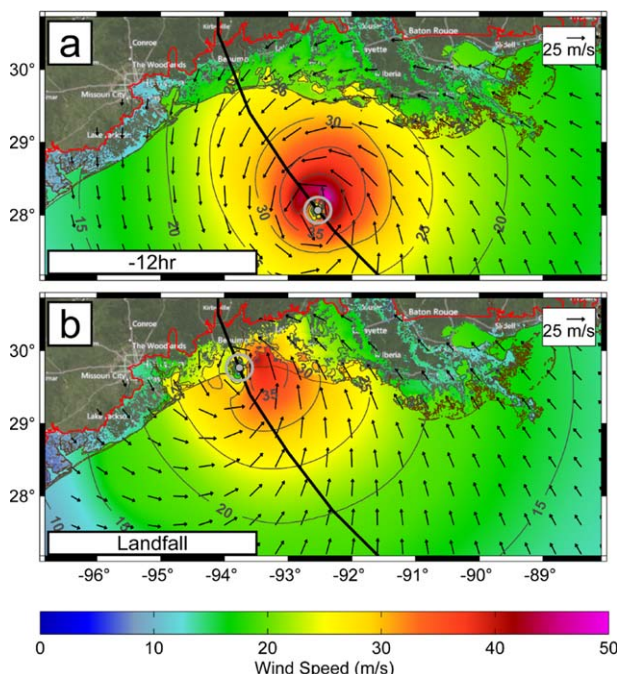
Figure 7. Contour plots of Rita 10 min averaged wind speed magnitudes and vectors in meters per second on (a) 2000 UTC 23 September 2005, roughly 12 before landfall and (b) 0800 UTC 24 September 2005, roughly at landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and ULLR domain boundary (red).

winds, forcing water out of hydraulically connected lakes and bays prior to landfall. To the west of its eventual landfall location, coastal areas in Texas experienced northerly winds ranging up to 25 m/s (10 min averaged OWI winds). Rita made landfall at the Texas-Louisiana border at 0740 UTC 24 September 2005. At landfall (Figure 7b) winds over the entire state of Louisiana ranged from southerly at the landfall location to east-southeasterly over southeastern Louisiana forcing surge shore perpendicular to the east of the storm track.

4.1. Regional Effects (12 h Before Landfall)

[44] Twelve hours prior to landfall, the eye of Rita was 170 km from the shoreline and 220 km from landfall, which as seen in Figure 8, is located just outside of the *ebp3* domain around the 100 m contour. The easterly sustained winds (Figure 7) east of the Mississippi River directed storm surge across Chandeleur and Breton Sounds, and into Biloxi Marsh, Lake Pontchartrain, and Lake Borgne, which led to overall water levels of 1.5–2 m in these areas as seen by all the unstructured models. The SLOSH *egm3* model did see some rise in water levels in Lake Borgne and Biloxi Marsh (1–1.5 m), but not in Lake Pontchartrain. The coupled wave and circulation models saw additional rise of water levels in the marshes for a total of 1.5–2 m. Currents were only mildly developed in this area (Figure 9). SLOSH does not have the ability to report currents and this region was outside of the *ebp3* domain.

[45] Just west of the Mississippi River Delta, storm surge was directed perpendicular to Terrebonne and Baratara

Bays. This led to higher water levels (~1 m) in that region for the unstructured mesh models with water levels increasing to 1–1.5 m for the coupled circulation and wave models. The SLOSH *egm3* model did not produce similar high water levels in this region. Comparing the currents (Figure 10) between the unstructured models, a difference is noticed in the currents at the Mississippi River Delta. ADCIRC and FVCOM and their respective coupled wave models developed high currents (>2 *m/s*) between the Bird's Foot and the Continental shelf break that extended along the edge of the shelf in both directions. In contrast, the SELFE model developed a much weaker current (~1.5 m/s) at the Bird's Foot, while the WWMII+SELFE model only had up to a ~2 m/s current. ADCIRC and FVCOM developed a strong current that extended west of Bird's Foot, which SELFE did not. This region was also outside of the *ebp3* domain.

[46] The area between Vermillion and Terrebonne Bays experienced increased water levels (Figure 8). All unstructured models saw an increase in water levels near Trinity and Ship Shoals (~1.5 m), with FVCOM and SELFE having a slightly greater water level at Trinity Shoal than ADCIRC. Strong shore-parallel winds (Figure 7a) in this region developed a strong shore-parallel current (Figure 9), which in turn began the drive of a short-lived geostrophic setup. SLOSH *egm3* developed a rise in water levels in this area (1 m), but to a lesser degree than the unstructured models. The SLOSH *ebp3* model covers a portion of this region and sees a mild increase in water levels near the *ebp3* open ocean boundary (~0.5 m) but none near the shore. Currents developed by each of the unstructured models in this region differ slightly. Considering the area just directly north of the eye of the storm, SELFE had the weakest currents (~2 m/s), while FVCOM had the strongest (~2.5 m/s).

[47] West of Vermillion Bay and near the landfall location, the unstructured mesh models saw moderate water levels nearshore (0.5–1 m). One difference between the unstructured models is that SELFE's 0.5 m water level contour parallels the shore (~50 km) near Matagorda Bay, whereas the 0.5 m water level contour intersects Matagorda Bay for ADCIRC and FVCOM. SLOSH *ebp3* and SLOSH *egm3* saw little or no rise in water levels along the nearshore. West of Rita's track, SELFE's currents are slightly lower than the currents for ADCIRC and FVCOM, especially along the shelf adjacent to Matagorda Bay.

4.2. Regional Effects (At Landfall)

[48] Rita made landfall near Sabine Pass at the border of Louisiana and Texas. East of the Mississippi River Delta, easterly sustained winds forced storm surge into Chandeleur and Breton Sounds, leading to elevated water levels in the marshes, bays, and lakes (~2 m Lake Pontchartrain). All the unstructured models showed similar results in terms of moderate water levels (Figure 10) and mild currents in this region (Figure 11). The SLOSH *egm3* had increased water levels (1.5–2 m) in Biloxi Marsh and Lake Borgne. In addition, SLOSH had a much lower water level in Lake Pontchartrain (0.5–1.0 m) than the unstructured models. Another differentiating result between the models is the location of the 0.5 m contour, which for ADCIRC, FVCOM, and SLOSH, was located off the continental
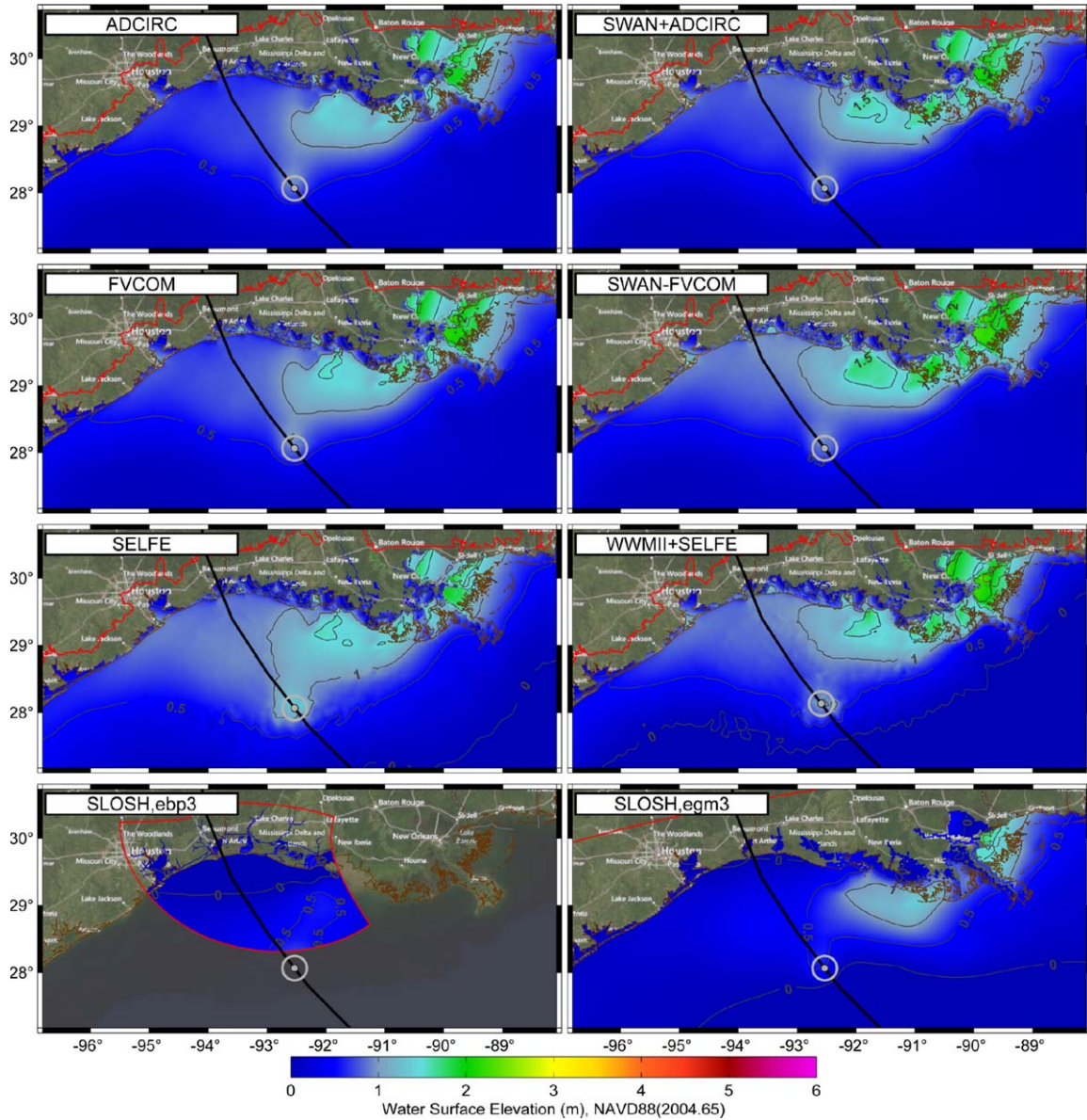
Figure 8. Contour plots of Rita water surface elevation in meters NAVD88 (2004.65) on 2000 UTC 23 September 2005, roughly 12 h before landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and domain boundary (red).

shelf, while for SELFE it was located closer to the shelf break.

[49] Between the Mississippi River Delta and the Atchafalaya Delta, winds were shore perpendicular (Figure 7b), leading to storm surge progressing toward and into the West Bank of New Orleans, Louisiana. Water levels were only mildly higher in the coupled wave and circulation models at this point in the storm (<0.5 m difference) than their counterparts. Water levels (Figure 10) for all unstructured models were fairly similar; however, currents similar to 12 h prior to landfall, differed for SELFE compared to ADCIRC and FVCOM. A strong current exists for the ADCIRC and FVCOM simulations at the Bird's Foot but not for SELFE (Figure 11). SLOSH *egm3* water levels were lower in the marshes and bays compared to the unstructured models.

[50] The area directly east of landfall received high shore-perpendicular winds ($\sim$40 m/s), which led to high water levels between Sabine Pass and the Atchafalaya Delta for all models. The water levels in this region were similar for all the unstructured models, with the coupled wave and circulation models showing slightly higher water levels. Currents for these models (Figure 11) are similar in this region, with the only notable difference being that the SELFE currents are slightly lower east of the track and near the shelf break. The general characteristics of the SLOSH *egm3* water bulge is similar but lower than the unstructured models, but SLOSH *egm3* had higher water levels extending further out on the coast than for the unstructured models, giving the appearance of a wider bulge. The high water levels for the SLOSH *ebp3* are less spread out than for the SLOSH *egm3* model, both seaward
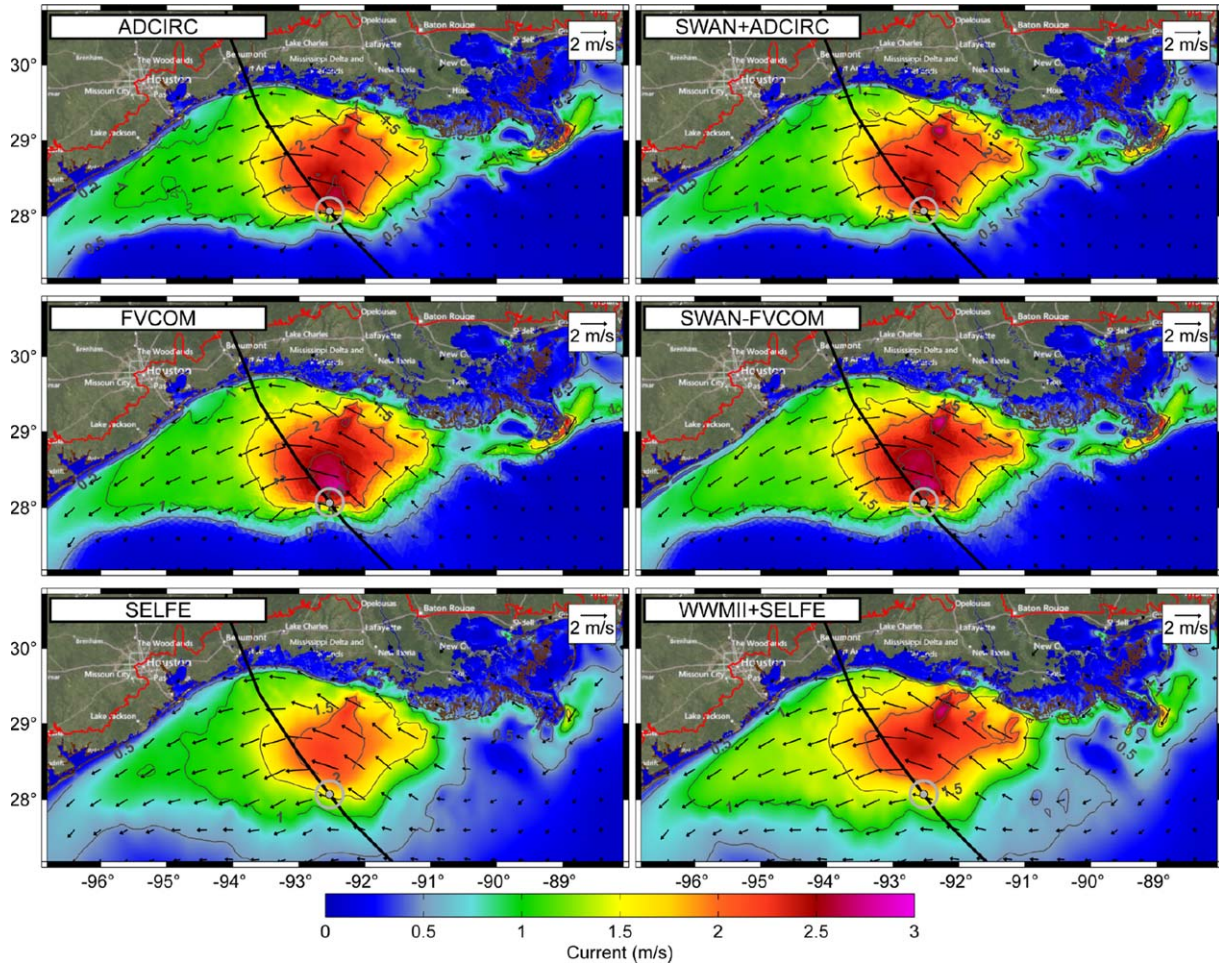
Figure 9. Contour plots of Rita depth-averaged water velocity magnitudes and vectors in meters per second on 2000 UTC 23 September 2005, roughly 12 h before landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and domain boundary (red).

and eastward along the coastline. The SLOSH *ebp3* model has a lower peak (∼4 m) than the SLOSH *egm3* model (∼4.5 m), and both were lower than the unstructured models (∼5 m).

[51] West of the track, each of the unstructured models had similar water levels and currents. One small difference in currents was the unusual increase in current magnitude west of the track for the WWMII+SELFE simulation compared to the SELFE simulation. Another small difference between the unstructured models is the slightly lower and wider current for SELFE along the shelf adjacent to Matagorda Bay compared to ADCIRC and FVCOM. With regard to the SLOSH, elevations west of the track were similar for SLOSH *ebp3* and SLOSH *egm3* but slightly lower than for the unstructured models.

4.3. Observation Data

[52] Figure 12 and Figure S4 of the supporting information describe the observed data collected during Rita and the spatial distribution of the observation locations. Much of this data was originally used for analysis in *Bunya et al.* [2010] and *Dietrich et al.* [2010]. Original raw data was assessed for accuracy and reliability. Data was

described by six categories: *Found*—the number of total stations found within the study domain; *Failed*—Stations that failed to record data for a period during the study, so if too few data were recorded during the storm the station was dismissed; *No Data*—Stations that failed to record data at any point during the study period; *Runoff*—Stations located in river channels where the effects of precipitation are clearly dominant; *Distrust*—Station that showed signs of inaccurate recording or other spurious observations during the study period; and *Used*—the number of stations actually used for the skill assessments. The categories are not additive as it was possible for a station to fail and still provide reasonable and useful data up until its time of failure. In addition to this station classification, each station's data were examined individually and individual data points were removed if identified as spurious to ensure all data analyzed were accurate. In total, 38 wave parameter time series, 121 water level time series, and 84 high water marks were used in the analysis (H. Das, personal communication, 2006). This included 21 fast-deploy gauges that were dispatched in the days prior to landfall, the first time this procedure was executed on a large scale for a major storm [*McGee*, 2006]
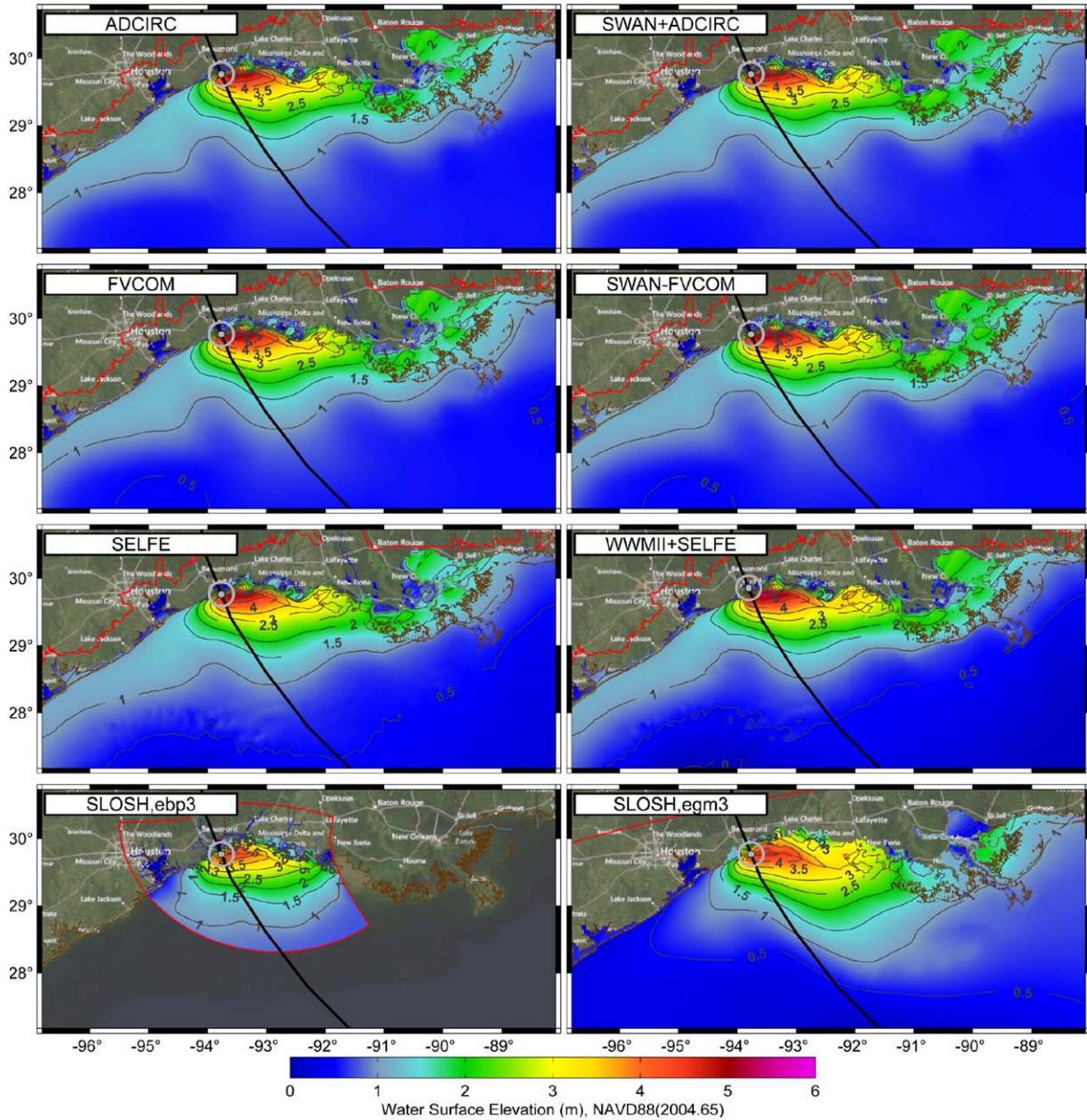
Figure 10. Contour plots of Rita water surface elevation in meters NAVD88 (2004.65) on 0800 UTC 24 September 2005, roughly at landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and domain boundary (red).

Collected observation data from Rita stretches from Key West, Florida, to the south Texas coast incorporating inland, floodplain, coastal, and deep water locations. The majority of locations are found in extreme eastern Texas and southern and southwest Louisiana characterizing surge in the highly dynamic area near landfall.

4.4. Waves

[53] Comparisons of SWAN+ADCIRC, SWAN-FVCOM, and WWMII+SELFE to measured significant wave heights are shown in Figure 13. As expected, SWAN+ADCIRC and SWAN-FVCOM perform similarly and accurately with slight differences likely caused by the variation in wind drag formulations between the wave models. A small departure is seen between WWMII+SELFE and the two other models. This departure appears to be spatially dependent with departures seen at stations further from the track. Table 6 lists the results of a statistical analysis of wave model performance. WWMII+SELFE has a large negative B_{MN} for significant wave height and also has the lowest R^2 for peak period. Aside from this, all models perform comparably across all wave parameters with the exception of the large wave direction $\overline{E}$ for SWAN-FVCOM. See Figure S5 of the supporting information for a comparison of wave characteristic performance categorized by data source.
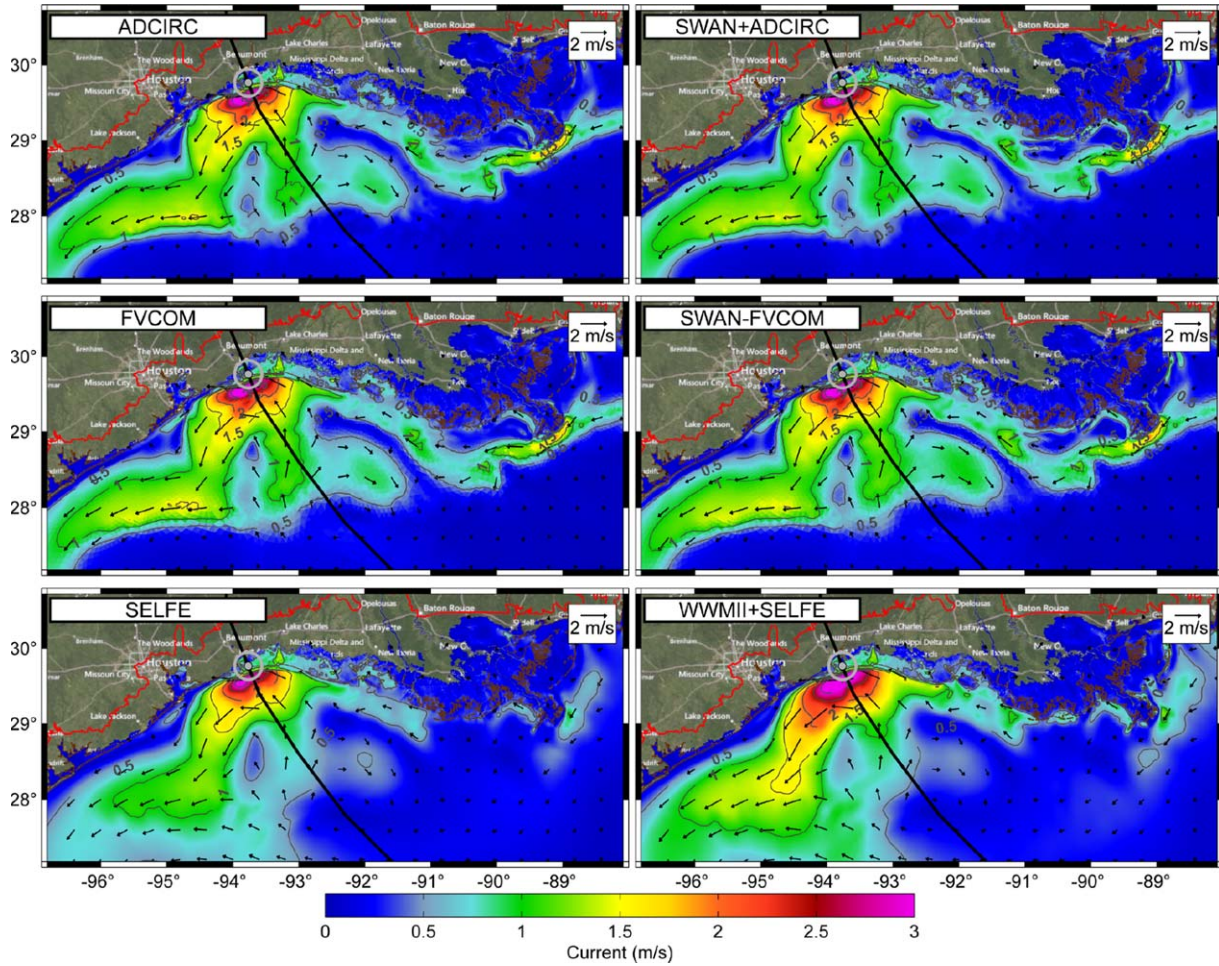
Figure 11. Contour plots of Rita depth-averaged water velocity magnitudes and vectors in meters per second on 0800 UTC 24 September 2005, roughly at landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and domain boundary (red).

Figure 12. Locations of water level and wave observation stations for Rita categorized by data source. Also shown are the *ULLR* domain boundary (black line) and the track of Rita (red line).

Figure 13. Observed and modeled significant wave height time histories (date in 2005) for Rita at select stations.

4.5. Water Level Time Series

[54] As shown in Figure 14, modeled water level time series were compared to observed water levels at southwest and southern Louisiana locations for circulation mode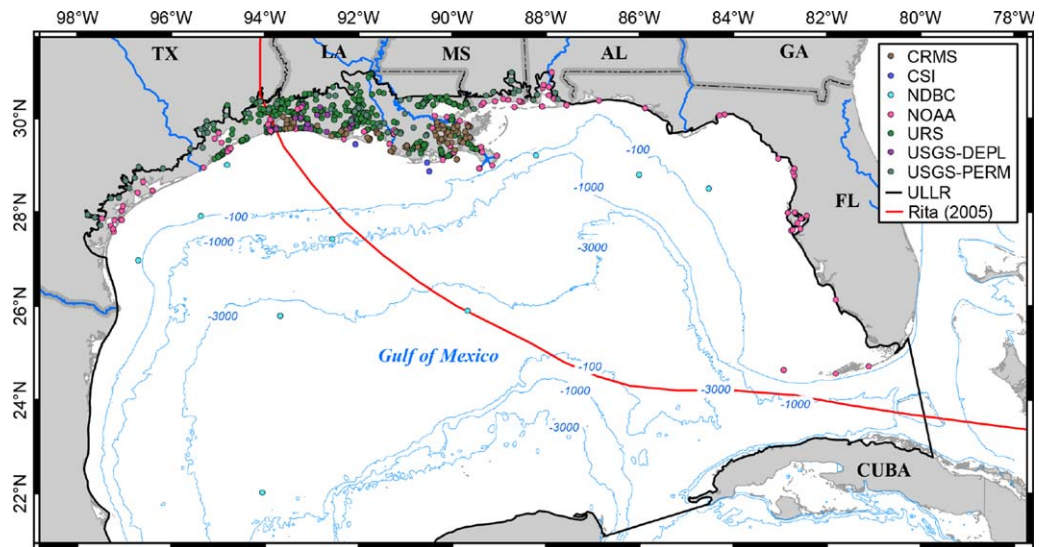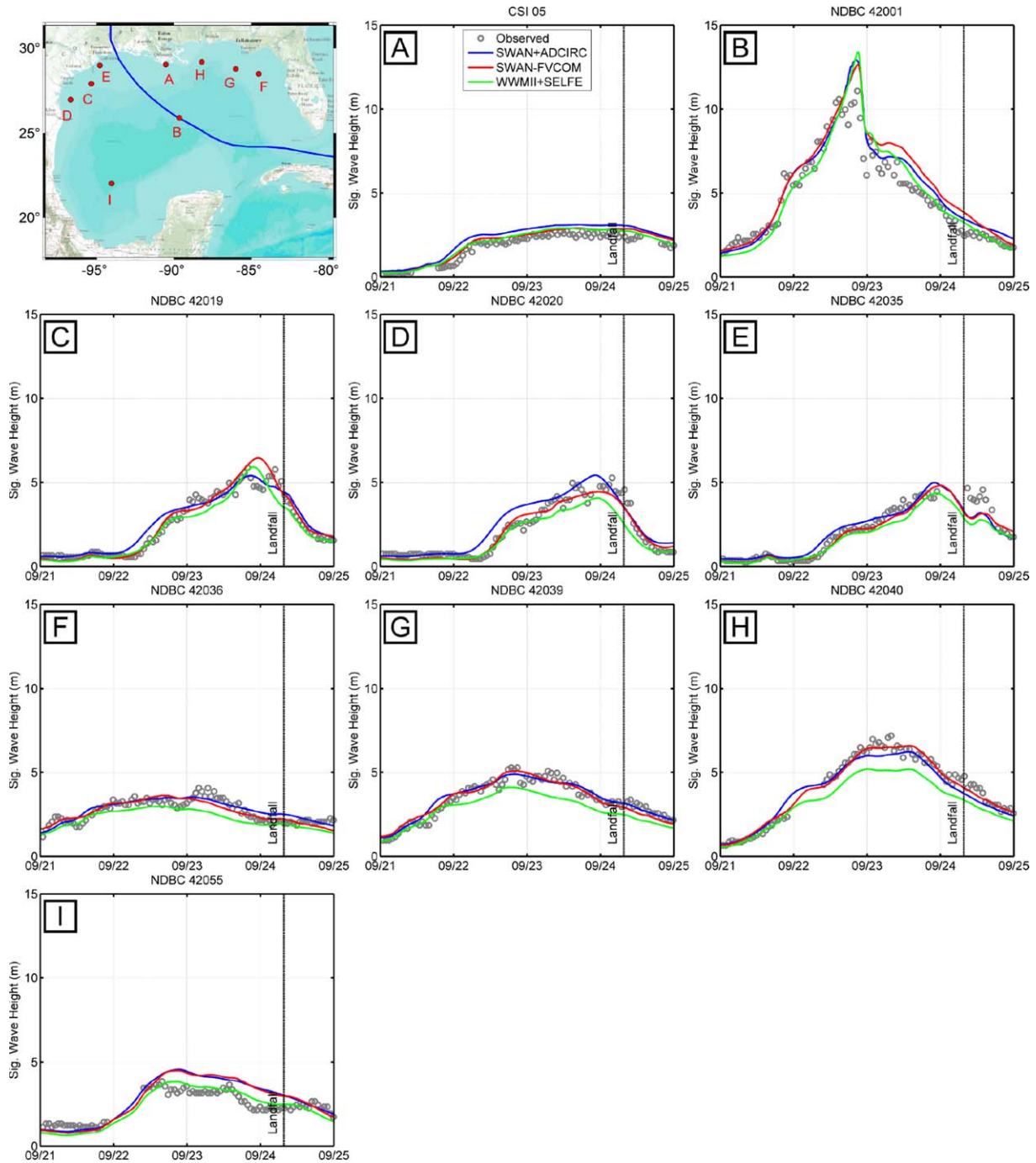ls without waves for Rita. Away from landfall in southern Louisiana (locations A, B, J, and K), overall model results are acceptable, accurately modeling the small surge driven by moderate winds, with a slight phase lag in the arrival of peak surge at location J, indicating a possible error in the hydraulic connectivity in the marsh. In the region of land-

fall, a definitive model shortcoming of SLOSH can be identified: the underdissipation of surge over land. At coastal locations (D, H, and I, with the exception of F at which an improperly resolved coastal road allowed surge to penetrate inland), peak water levels are captured by all models, with slight overprediction by all models except SLOSH-*epb3* which is accurate at H and I, but underpredicts at D. At inland stations (locations C, E, and G), while some models have dissipated some of the surge, none have dissipated the surge so that the proper inland value is modeled. At location E, north of Calcasieu Lake, SLOSH *ebp3*, has

Table 6. Rita Simulation Wave Characteristic Time Series and Water Level Hydrograph Error Statistics[a]

	Model	R^2	E_{RMS}	$\overline{E}$	B_{MN}	σ	SI	MAE	E_{NORM}	Pts
Wave direction	SWAN+ADCIRC		43.214	−1.068		39.931	0.228	28.766		9
	SWAN-FVCOM		54.441	21.082		45.171	0.257	37.372		9
	WWMII+SELFE		61.242	−2.031		57.281	0.345	41.629		9
Significant wave height	SWAN+ADCIRC	0.904	0.638	0.314	0.122	0.5	0.178	0.507	0.209	10
	SWAN-FVCOM	0.917	0.598	0.24	0.08	0.485	0.172	0.468	0.191	10
	WWMII+SELFE	0.909	0.664	−0.178	−0.058	0.502	0.174	0.52	0.206	10
Mean period	SWAN+ADCIRC	0.83	1.068	0.304	0.036	0.971	0.114	0.843	0.121	9
	SWAN-FVCOM	0.788	2.115	−1.878	−0.222	0.933	0.11	1.908	0.24	9
	WWMII+SELFE	0.809	2.13	−1.907	−0.226	0.915	0.107	1.942	0.242	9
Peak period	SWAN+ADCIRC	0.708	2.029	0.704	0.068	1.869	0.18	1.278	0.187	10
	SWAN-FVCOM	0.686	2.457	−1.324	−0.13	1.868	0.18	1.892	0.229	10
	WWMII+SELFE	0.439	3.528	0.647	0.063	3.338	0.323	2.133	0.324	10
Water level (wet-only)	ADCIRC	0.759	0.361	−0.016	0.044	0.162	0.163	0.315	0.328	83
	FVCOM	0.753	0.368	0.059	0.113	0.164	0.177	0.319	0.354	82
	SELFE	0.754	0.364	−0.055	−0.066	0.179	0.183	0.315	0.314	82
	SLOSH,*ebp3*	0.5	0.995	−0.405	−0.295	0.389	0.288	0.898	0.632	32
	SLOSH,*egm3*	0.522	0.7	−0.08	−0.055	0.258	0.27	0.633	0.589	77
	SWAN+ADCIRC	0.743	0.365	0.104	0.149	0.175	0.173	0.315	0.332	83
	SWAN-FVCOM	0.743	0.418	0.181	0.23	0.179	0.189	0.368	0.39	83
	WWMII+SELFE	0.676	0.404	0.096	0.04	0.204	0.258	0.346	0.372	83
Water level (TS)	ADCIRC	0.528	1.651	1.292	1.099	0.212	0.184	1.605	1.345	119
	FVCOM	0.518	1.662	1.33	1.139	0.216	0.196	1.615	1.363	119
	SELFE	0.527	1.656	1.26	1.025	0.219	0.192	1.609	1.331	119
	SLOSH,*ebp3*	0.394	1.498	0.246	0.107	0.438	0.325	1.396	0.904	41
	SLOSH,*egm3*	0.353	2.298	1.702	1.434	0.29	0.282	2.236	1.892	120
	SWAN+ADCIRC	0.531	1.65	1.373	1.171	0.219	0.19	1.603	1.345	119
	SWAN-FVCOM	0.523	1.686	1.426	1.227	0.225	0.204	1.638	1.384	119
	WWMII+SELFE	0.475	1.673	1.372	1.1	0.239	0.247	1.619	1.37	119

[a]The rows listed as Topo-Substitution (TS) use the station bathymetry at stations where it is not wetted by the model, while the rows listed "Wet-Only" omit any station that is not wetted by the model from the analysis. Some statistics were not applicable for wave direction.

dissipated surge to its proper level; however, the large lag in time of arrival of peak surge indicates improper model physics in the region.

[55] Modeled water level time series were also compared to observed water levels at southwest and southern Louisiana locations for coupled wave and circulation models for Rita (Figure 15). Almost identical behavior is seen between the coupled waves and circulation models presented in Figure 15 as the circulation only models show in Figure 14, with the exception of Stations A and B, which show improvement with the addition of waves in southeastern Louisiana. Peak water levels at the coast are accurately captured, however at inland stations little to no dissipation is seen as the addition of wave radiation stress gradients increase water levels in wave breaking and inland areas.

[56] Analyzing the statistics of model performance to observed data (Table 6 and Figure S6 of the supporting information) confirms much of what was identified in the previous figures. Based on B_{MN} and $\overline{E}$, SLOSH tends to overpredict water levels. In fact, almost without exception all models tend to be overpredictive based on B_{MN} and $\overline{E}$. This explains the deterioration of results when wave models are coupled to the circulation models. Because the circulation models are already statistically overpredictive, when wave radiation stress gradients are added, water levels can be expected to increase in inland areas, thus making already overpredictive models more overpredictive. While an increase in $\overline{E}$ and B_{MN} values is seen when adding the wave models, the change is small and the coupled wave

and circulation models would still be classified as accurate. In general, ADCIRC is the top statistical performer; however, the differences between the unstructured models are small and all unstructured mesh simulations can be classified as accurate. A large separation is seen between the structured and unstructured models. SLOSH *egm3* performs better than SLOSH *ebp3*; however, the low number of stations analyzed in the *ebp3* domain limits its statistical set in comparison to the other models. The lack of inland attenuation can seen by the Topo-Substitution method (TS) water level B_{MN} statistics for USGS-PERM stations (See Figure S6 of the supporting information). USGS-PERM stations are located primarily in inland areas such as channels and lakes and the high values for B_{MN} produced by all models quantifies this overestimation. Compare this to model performance at coastal stations (CRMS, CSI, NOAA, and USGS-DEPL), where B_{MN} values are much closer to zero, indicating better model performance. It is to be noted that the number of locations at which model comparisons were made differs from model to model, particularly between the structured and unstructured models.

4.6. High Water Marks

[57] As shown in Figure 16, the maximum modeled water levels during Rita were compared for each model and those water levels compared to measured high water marks (HWMs) spatially in Figure 17 and correlatively in Figure 18 with the statistics of these HWM comparisons presented in Table 7. Qualitatively and quantitatively,
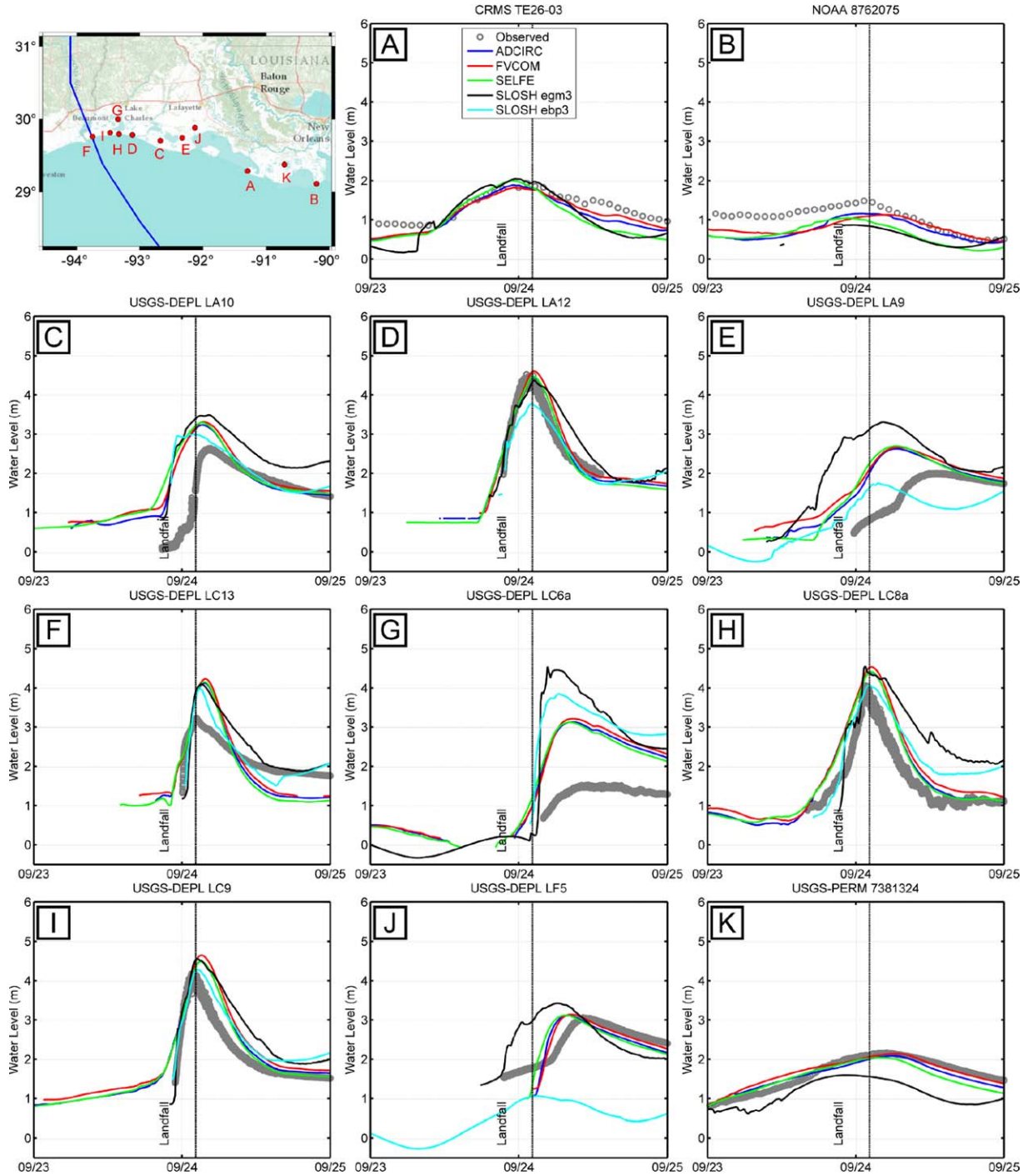
Figure 14. Observed and circulation (without waves) modeled water level time histories (date in 2005) for Rita at select stations.

ADCIRC, FVCOM, SELFE, and their associated coupled wave and circulation models perform similarly; however a large degradation is seen in correlation between the models analyzed here and those used in previous high-resolution studies [*Dietrich et al.*, 2011b, 2012b; *Hope et al.*, 2013]. The line of best fit for all unstructured models shows the models to be slightly overpredictive, although this fits into the idea that the models are overpredictive inland as this is where the majority of HWMs are located. SLOSH results, despite having a good zero-intercept slope of best fit and a relatively low mean error, had significantly higher *SI*, *MAE*, and E_{NORM} than the unstructured models. This is further illustrated by the map, where SLOSH underestimated HWMs near landfall, and highly overestimated HWMs further inland and adjacent to the track.

4.7. Inundation Extents

[58] Geographic snapshots of Rita's maximum extent of inundation for each of the unstructured models and SLOSH simulations are shown in Figures S7 and S8 of the
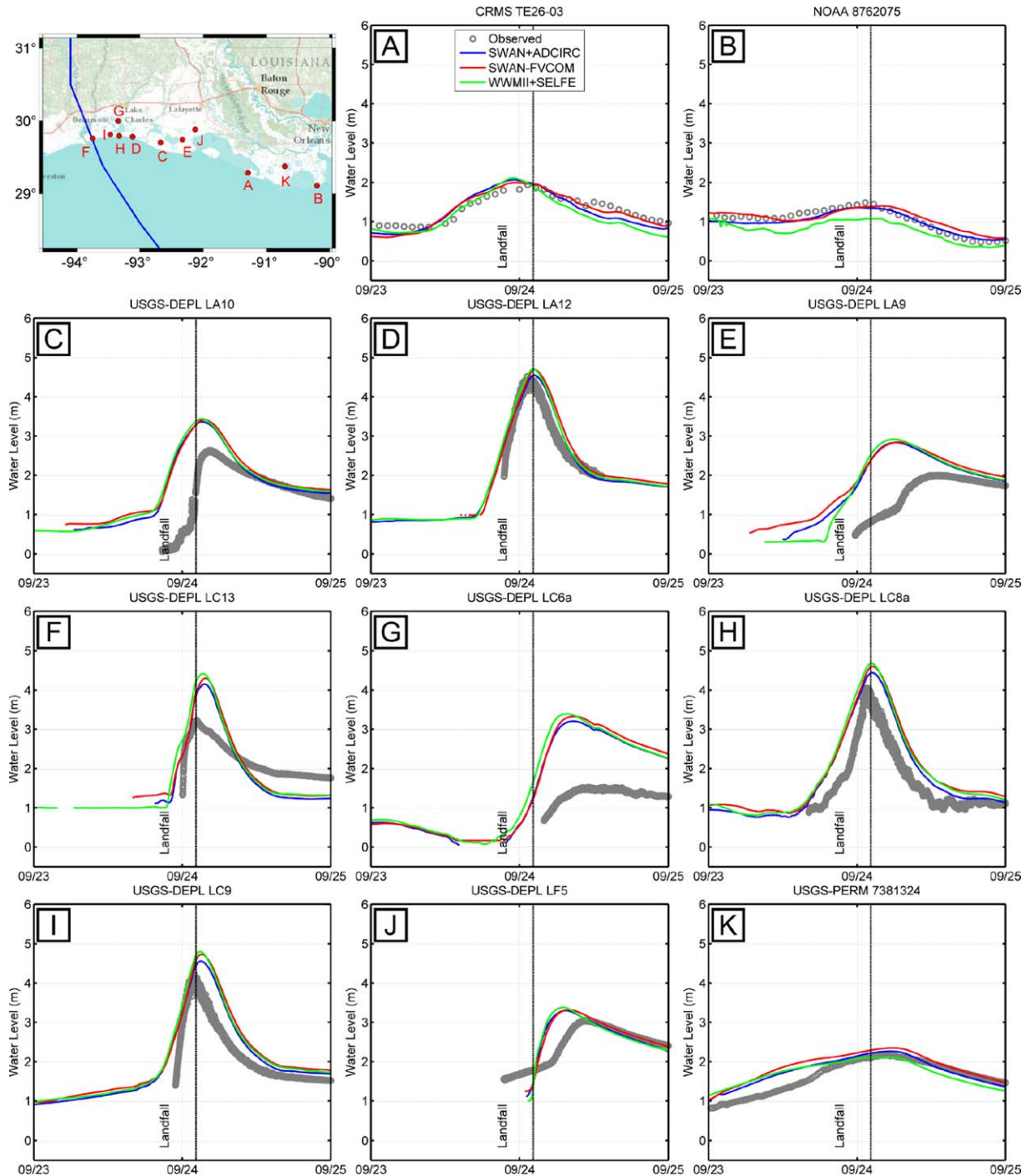
Figure 15. Observed and coupled wave and circulation modeled water level time histories (date in 2005) for Rita at select stations.

supporting information. As mentioned previously in the comparison of water level time series, the unstructured models perform similarly, and is confirmed here by very similar extents of inundation in the area around Sabine Lake at the border of Texas and Louisiana. The influence of waves was very minimal in this area as shown here by the slight variation in inundation lines. In contrast to Sabine Lake, the slighter relief of the West Bank area of New Orleans featured more apparent differences between the unstructured mesh models. In comparison to ADCIRC and

FVCOM, which were very similar in terms of inundation extents, SELFE noticeably inundated less area. The contribution of waves is slightly apparent here for SWAN+ADCIRC and SWAN-FVCOM, but far more apparent for WWMII+SELFE. Differences are accentuated in this area due to the low-lying topography typical of southern Louisiana's marshes. The small topographic gradient in the region makes the models much more sensitive to wetting and drying, which means that the difference in lines of inundation between models is subject not only to
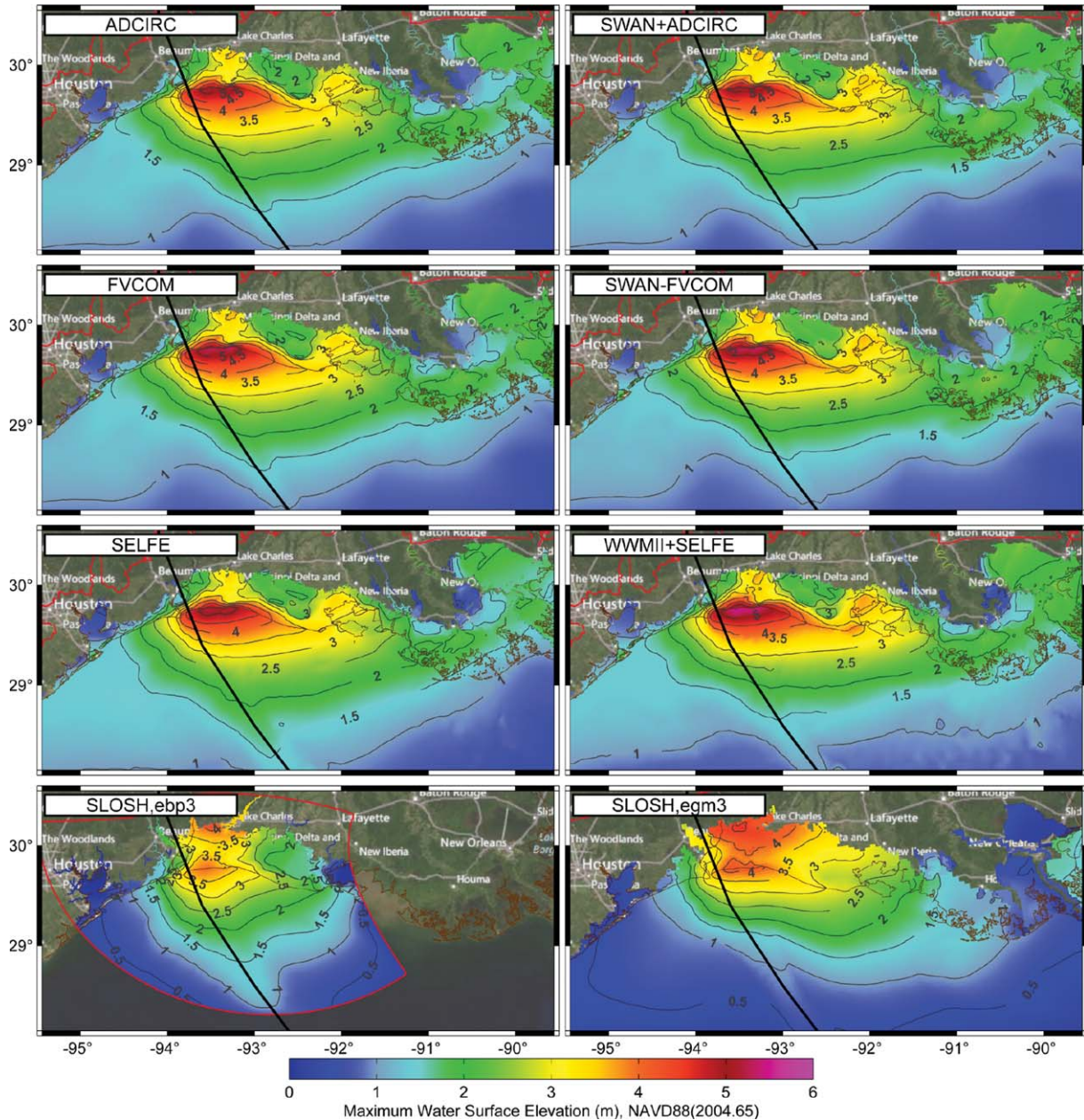
Figure 16. Contour plots of Rita maximum water surface elevation in meters NAVD88 (2004.65). Also shown are the hurricane track (black) and domain boundary (red).

water levels, but also to the models internal wetting and drying scheme.

[59] Inundation extents for the SLOSH simulations differ depending on local or Gulf scale meshes. SLOSH's overprediction of peak water levels is evident in the much further inland penetration of surge by all SLOSH models as compared to the unstructured models. Also notable is the coarseness of the *egm3* mesh as compared to the *ebp3* and unstructured inundation extent lines. The higher resolution of the *ebp3* mesh in comparison to the *egm3* mesh allows for a more accurate representation of the coastal topography as seen by the more defined channels and rivers in the *ebp3* inundation extents. This overprediction of water levels by SLOSH, as shown earlier in Figure 17, is likely

caused by the SLOSH's internal friction formulation, which tended to underdissipate inland surge, when the water level time series were examined.

5. Hurricane Ike (2008)

[60] Ike entered the Gulf of Mexico at 2030 UTC 9 September 2008 after making landfall in the Cuban state of Pinar del Rio. Ike proceeded on a northwest track with its wind field broadening, with tropical storm force and hurricane force winds extending 445 and 185 km, respectively, from the storm's center at 1800 UTC 10 September 2008. Ike's intensity fluctuated until 0000 UTC 12 September 2008 when peak winds (10 min averaged OWI winds based
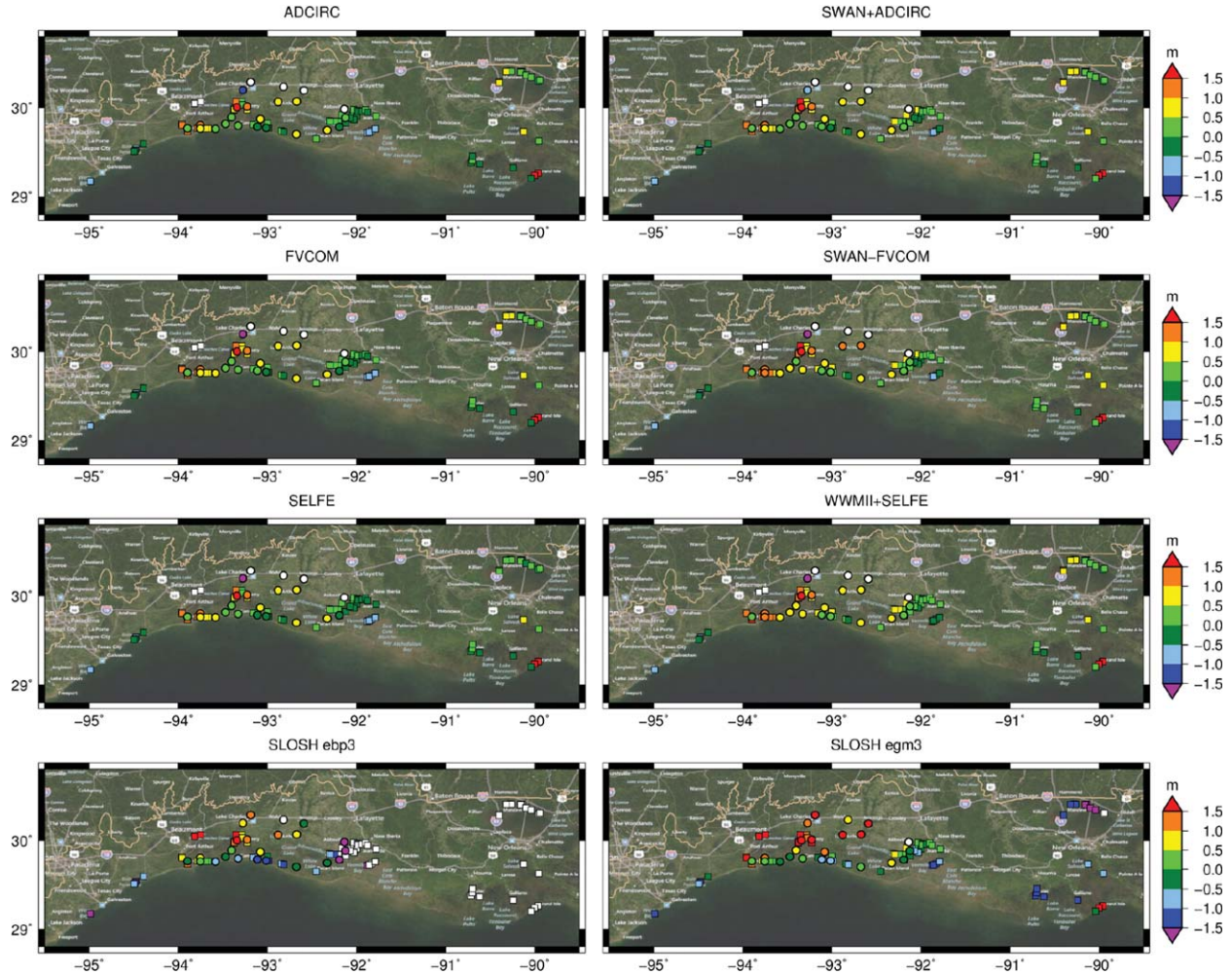
Figure 17. Locations of Hurricane Rita HWMs (circles) and hydrographs (squares) along the Northwest Gulf coast. The points are color-coded to show the errors between measured and modeled peak water levels. Green points indicate matches within 0.5 m and white points indicate locations that were never wetted by the model.

on H*WIND/IOKA) were roughly 40 m/s at which point Ike began a steady but moderate strengthening process that occurred until landfall. At 0700 UTC 12 September 2008 (Figure 19a), approximately 24 h prior to landfall, Ike's wind field aligns with the large-scale LATEX coastal geography. This results in southeasterly winds over New Orleans and the marshes of southeastern Louisiana, easterly winds over the marshes and lakes of southern Louisiana, and northeasterly winds across the Texas coast. Progressing 12 h ahead to 1900 UTC 12 September (Figure 19b), similar directionality is seen across the LATEX shelf and coast with wind velocities decreasing over southeastern Louisiana and increasing in southwestern Louisiana and Texas. These moderate shore-parallel winds generated the shore-parallel current that resulted in a geostrophic setup known as the "forerunner" surge that inundated coastal areas with up to 2 m of water prior to landfall [*Kennedy et al.*, 2011; *Hope et al.*, 2013]. At approximately 1900 UTC 12 September 2008, Ike departed from its northwest track moving north-northwest toward Galveston Island. Winds remained largely shore parallel until immediately before landfall (0700 UTC 13 September 2008, Figure 19c) when winds shifted to directly onshore (southeas-

terly) to the northeast of landfall and directly offshore (northwesterly) to the southwest of landfall. Peak winds (10 min averaged OWI winds based on H*WIND/IOKA) at landfall were roughly 40 m/s over the Bolivar Peninsula in Texas. At 1300 UTC 13 September 2008, 6 h after landfall, winds over Galveston Bay and the areas to the northwest of landfall were still subject to 20 m/s southwesterly and southerly winds, hindering the recession of storm surge out of Galveston Bay and the surrounding flood plains.

5.1. Regional Effects (24 h Before Landfall)

[61] A day before Ike made landfall, the eye of the storm was positioned at the 2000 m depth contour of the continental slope about 470 km southeast of Galveston, TX. Winds (Figure 19a) were generally easterly along the Louisiana Coast (∼20 m/s) and northeasterly along the Texas coast (∼10 m/s). As shown in Figure 20, east of the Mississippi River, each unstructured model featured elevated water levels (1–2 m) from Lake Pontchartrain to Breton and Chandeleur Sounds. Water levels were similar for ADCIRC and FVCOM, but SELFE had approximately 0.5 m higher water levels in this region. SWAN+ADCIRC and
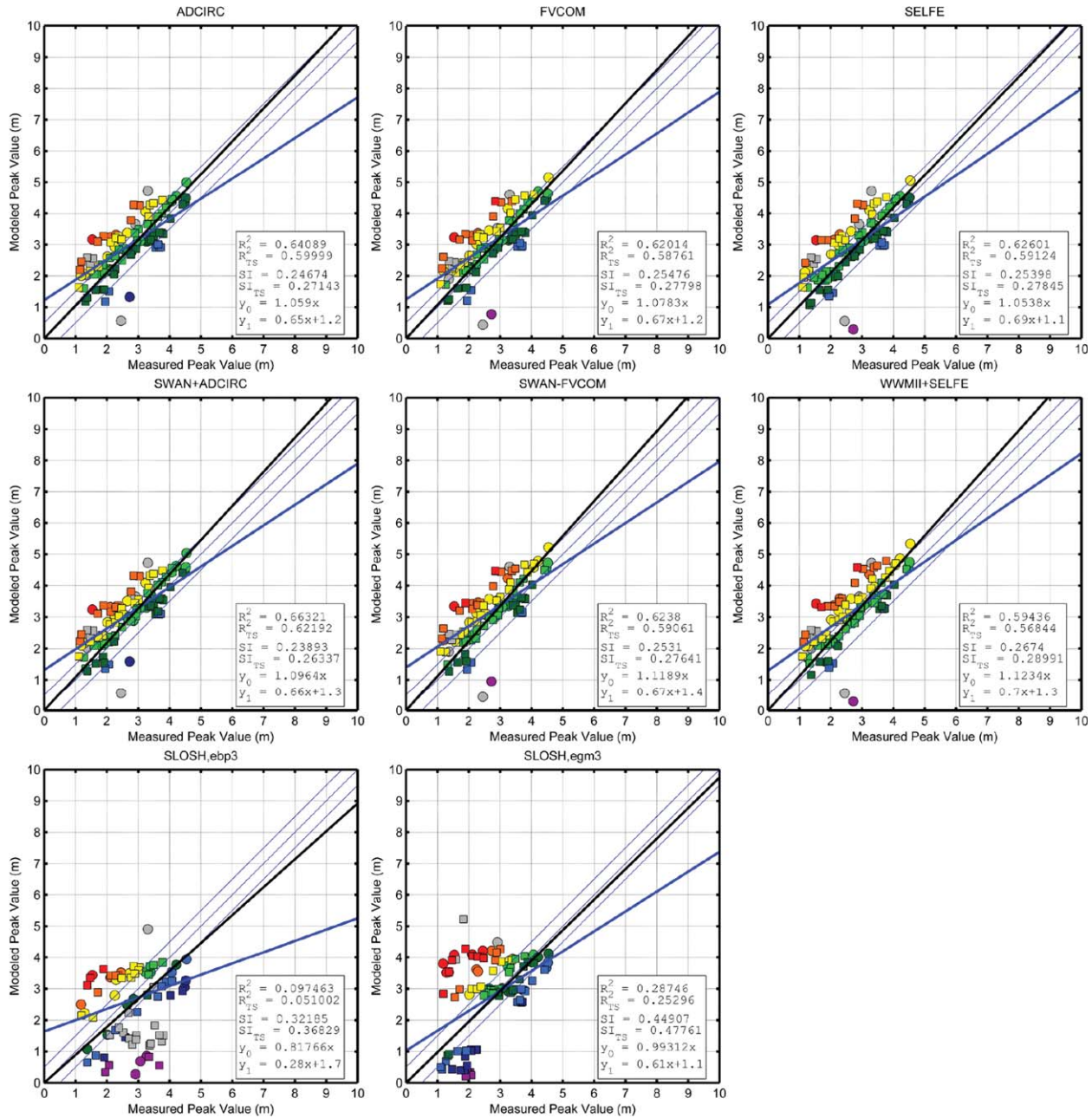
Figure 18. Scatter plots of Rita HWMs (circles) and peak hydrograph levels (squares) at 107 stations (Figure 25). Red, orange, yellow, and light green points indicate overprediction by the model; dark green, blue, dark blue, and purple points indicate underprediction. Green points indicate a match within 0.5 m and gray points indicate locations that were never wetted by the model. Metrics shown are for the TS method, unless they have the subscript "C" which refers to "Wet-Only" method. The thick blue line and the thick black line represent the y_1 and y_0 best fit lines, respectively.

SWAN-FVCOM were also similar, while WWMII+ SELFE had approximately 0.5 m higher water levels in this region. Waves contributed to a slight increase (<0.5 m) of water levels in the marshes and back bays. Currents in this area were similar and generally mild for the unstructured models (Figure 21). SELFE and WWMII+SELFE did generate a slightly higher current at the northern end of Lake Borgne than ADCIRC or FVCOM. This region was outside of the *egl3* domain. SLOSH *egm3* had elevated levels (~1.5 m) in Caernarvon Marsh, Biloxi Marsh, and Lake

Borgne, but did not have elevated water levels in Lake Pontchartrain.

[62] Between the Mississippi River Delta and Galveston (Figure 20), TX, ADCIRC, and FVCOM had water surface elevations of roughly 1 m, whereas SELFE had water surface elevations closer to 1.5 m. Currents for the unstructured models were very similar for ADCIRC and FVCOM, but slightly different for SELFE. In general, the high currents were restricted to the top of the shelf for ADCIRC and FVCOM, but SELFE's high currents extended out over

Table 7. Rita Simulation HWM Error Statistics[a]

	Model	R^2	E_{RMS}	$\overline{E}$	B_{MN}	σ	SI	MAE	E_{NORM}	Dry	Out	Wet
HWM (TS)	ADCIRC	0.6	0.786	0.347	0.134	0.708	0.271	0.578	0.286	7	0	100
	FVCOM	0.588	0.823	0.396	0.152	0.725	0.278	0.614	0.299	7	0	100
	SELFE	0.591	0.788	0.312	0.12	0.727	0.278	0.557	0.286	7	0	100
	SWAN+ADCIRC	0.622	0.821	0.454	0.175	0.687	0.263	0.61	0.299	7	0	100
	SWAN-FVCOM	0.591	0.886	0.519	0.2	0.721	0.276	0.676	0.322	7	0	100
	WWMII+SELFE	0.568	0.914	0.518	0.199	0.756	0.29	0.675	0.332	7	0	100
	SLOSH,ebp3	0.051	1.308	−0.37	−0.13	1.048	0.368	1.106	0.441	16	33	58
	SLOSH,egm3	0.253	1.252	0.169	0.065	1.246	0.478	1.026	0.455	3	0	104
HWM (wet-only)	ADCIRC	0.641	0.749	0.33	0.125	0.653	0.247	0.54	0.269	7	0	100
	FVCOM	0.62	0.797	0.391	0.149	0.674	0.255	0.585	0.286	7	0	100
	SELFE	0.626	0.752	0.294	0.111	0.672	0.254	0.517	0.27	7	0	100
	SWAN+ADCIRC	0.663	0.789	0.445	0.169	0.632	0.239	0.574	0.283	7	0	100
	SWAN-FVCOM	0.624	0.866	0.524	0.199	0.67	0.253	0.651	0.311	7	0	100
	WWMII+SELFE	0.594	0.892	0.514	0.195	0.708	0.267	0.645	0.32	7	0	100
	SLOSH,ebp3	0.097	1.214	−0.124	−0.045	0.893	0.322	1.009	0.416	16	33	58
	SLOSH,egm3	0.287	1.194	0.104	0.04	1.178	0.449	0.985	0.432	3	0	104

[a]The rows listed as Topo-Substitution (TS) use the station bathymetry at stations where it is not wetted by the model, while the rows listed "wet-only" omit any station that is not wetted by the model from the analysis.

the shelf break (Figure 21). This is illustrated best by the narrow high current field between the Bird's Foot and the shelf break for ADCIRC and FVCOM and the wide current field over the shelf break for SELFE. Coupled wave and circulation models were very similar to their without waves simulation. The coupled wave and circulation models each featured ~0.5 m greater water levels in the marshes and backwaters around Terrebonne and Barataria Bays than their without waves counterparts. The currents for the wave and circulation models were similar to their without waves counterparts. For SLOSH *egl3*, the water levels were around 0.5 m near Galveston, which is about 0.5 m less than FVCOM and ADCIRC, and 1 m less than SELFE. For SLOSH *egm3*, water levels were less than for the unstructured models but greater than for SLOSH *egl3*. In addition, SLOSH *egm3* had some elevated water levels near Terrebonne Marsh.

5.2. Regional Effects (12 h Before Landfall)

[63] Roughly 12 h before landfall, the eye of Ike crossed onto the top of the edge of the continental shelf near the 300 m contour and roughly 230 km from Galveston. Winds (Figure 19b) ranged from southeasterly at the Bird's Foot (15–20 m/s), to easterly at Vermillion Bay (~20 m/s), to northeasterly at Galveston (15–20 m/s). East of the Mississippi River, results for the unstructured models were relatively similar (Figure 22). Water levels for ADCIRC and FVCOM were 1.5–2 m in Lake Pontchartrain, Lake Borgne, Caernarvon Marsh, and Biloxi Marsh, while water levels were ~2 m in these areas for SELFE. Water levels for the unstructured wave and circulation models were slightly higher than their without waves counterparts. Currents in this region for this slow time scale process were very small and similar in all the unstructured models (Figure 23). SLOSH *egm3* had much lower water levels in this area than the unstructured models and did not have high water levels in Lake Pontchartrain, but did have elevated water levels (~1.5 m) in Lake Borgne, Caernarvon Marsh, and Biloxi Marsh. The area east of the Mississippi River is outside of the *egl3* domain.

[64] Between Isles Dernieres, LA, and the Mississippi Bird's Foot, the unstructured models had high water levels in the back bays and marshes, whereas the SLOSH *egm3* did not. The wave and circulation models had about 0.5 m higher water levels in this area than their without waves counterparts due to wave action. The unstructured model water levels in this area were similar, except that SELFE was slightly higher. Currents in this region were similar for ADCIRC and FVCOM, but slightly different for SELFE. ADCIRC and FVCOM had a strong narrow and long current field by the Mississippi Bird's Foot, but SELFE had a wider, weaker, and shorter current field in comparison (Figure 22).

[65] All the unstructured models experienced relatively similar water levels along the coast between Corpus Christi and the Atchafalaya Delta. Water levels were 2–2.25 m between Galveston and Marsh Island, LA. The only notable difference is that SELFE was slightly higher than ADCIRC and FVCOM and had a slightly wider area affected. The unstructured model current fields were similar with the exception that the high current field for SELFE was a little bit wider than ADCIRC and FVCOM. The coupled wave and circulation models had mildly higher water levels in this region than their without waves counterparts. SLOSH *egm3* had lower water levels than the unstructured models with water levels around 1 m near Galveston, TX and 1.5 m near Marsh Island. The increased water levels for the unstructured runs were spread out from Terrebonne Bay past Corpus Christi and centered around Calcasieu Shipping Channel, whereas the increased water levels for the SLOSH *egm3* model were focused between Matagorda Bay and the Atchafalaya Delta and focused around Marsh Island. SLOSH *egl3* did not experience water levels higher than 0.5 m.

5.3. Regional Effects (At Landfall)

[66] Ike made landfall at the eastern edge of Galveston Island. Winds (Figure 19c) were generally southerly or southeasterly along the coast east of landfall and northwesterly or westerly along the coast west of landfall. OWI wind speeds (10 min averaged based on H*WIND/IOKA) ranged from 15 m/s at Corpus Christi to 35–40 m/s just east of
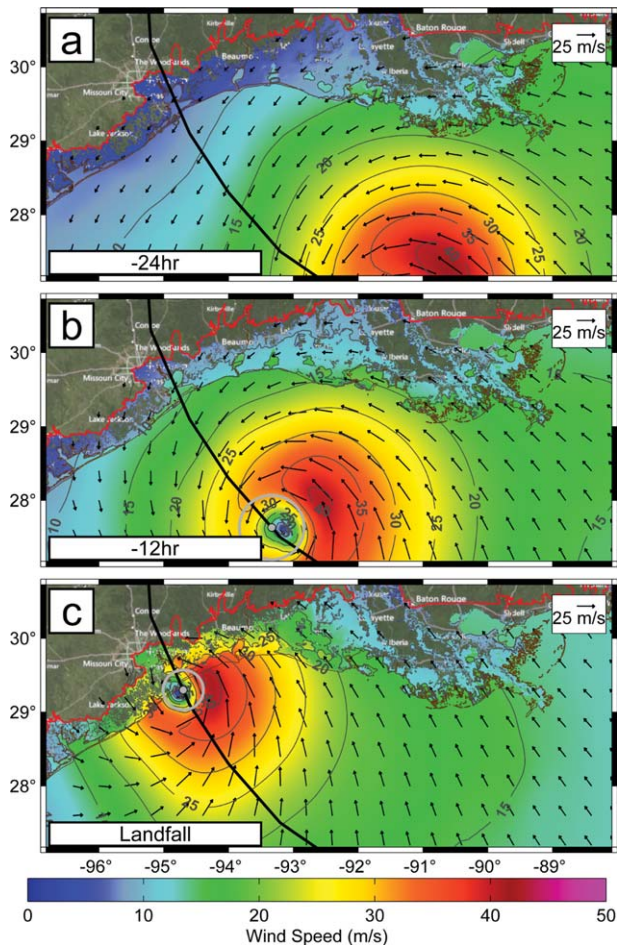
Figure 19. Contour plots of Ike 10 min averaged wind speed magnitudes and vectors in meters per second on (a) 0700 UTC 12 September 2008, roughly 24 h before landfall, (b) 1900 UTC 12 September 2008, roughly 12 h before landfall, and (c) 0700 UTC 13 September 2008, roughly at landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and ULLR domain boundary (red).

Galveston, to 20 m/s at Vermillion Bay, and 15 m/s at the Bird's Foot. East of the Mississippi River water levels were similar for the unstructured models (Figure 24), with the coupled wave and circulation models having slightly higher water levels than their without waves counterparts. The unstructured models had water levels of 1.5–2 m in Lake Pontchartrain and Caernarvon Marsh, whereas the SLOSH *egm3* model had water levels of 0.5 m in Lake Pontchartrain and 1−1.5 m in Caernarvon Marsh. This area was outside of the *egl3* domain.

[67] West of the Atchafalaya Delta and the Mississippi River Bird's Foot, the water levels for the unstructured models were similar. Water levels along the coast ranged from 1 to 1.5 m adjacent to Corpus Christi, ~5 m just east of Galveston, ~2 m near Marsh Island, and ~1 m near Barataria Bay. As shown in Figure 25, the currents for the unstructured models were relatively similar except for between the Atchafalaya Delta and the Bird's Foot, where ADCIRC and FVCOM had higher currents than SELFE,

and adjacent to Corpus Christi, where SELFE had a wider high current field than ADCIRC and FVCOM. The SLOSH *egm3* water levels were slightly lower near Galveston (~4 m) than for the unstructured models and were lower near Corpus Christi (~0.75 m) and Barataria Bay (~0.75 m). Water levels for the SLOSH *egl3* model were much less than for the SLOSH *egm3* model. The peak surge for the SLOSH *egl3* model was ~3 m.

5.4. Observation Data

[68] Shown in Figure 26 is the spatial distribution of stations, where observational data was collected during Ike. This data was originally assembled for analysis by *Hope et al.* [2013]. The same classification and examination process was applied to Ike's data as was to Rita's data (See Figure S9 of the supporting information). In total 81 wave parameter time series, 579 water level time series, and 243 high water marks were used in analysis. Denser spatial coverage of observation data was obtained for Ike as compared to Rita. The increased number of permanent stations, the use of quick-deploy gauges prestorm, and the fortification of existing gauges allowed for a much larger number of observations to be obtained.

5.5. Waves

[69] As shown in Figure 27, modeled significant wave heights were compared to measured data in the Gulf of Mexico and on the LATEX shelf. Similar to Rita, at most stations, SWAN+ADCIRC and SWAN-FVCOM performed comparably and accurately except during highly dynamic periods such as the passing of peak winds. Also similar to Rita, WWMII+SELFE tended to underpredict significant wave heights as compared to SWAN+ADCIRC and SWAN-FVCOM; however, the deterioration of performance when moving away from the storm center as was seen in Rita by WWMII+SELFE was not seen for Ike. Overall, qualitatively all models performed well, and statistically (Table 8) all models performed comparably with the exception that SWAN+ADCIRC performed slightly better in some categories. The differences, however, were slight and can likely be attributed to the different approaches to wind drag used by the wave models. See Figure S10 of the supporting information for a comparison of wave characteristic performance categorized by data source.

5.6. Water Level Time Series

[70] Circulation model water level time series were compared to observed water levels at coastal locations (Figure 28). Location A, in Lake Pontchartrain experienced a gradual rise in water due to the steady southeasterly winds in the region. Notable at station A is the lack of surge in SLOSH *egm3* and the presence of a tidal signal in Pass Manchac. The tidal signal is an example of the kind of potential error that can occur by superimposing tidal signals, as explained in section 2.5. In this case, the *ULLR* mesh is too coarse to resolve the Pass and thus receives no tidal signal, whereas the *egm3* is even coarser than the *ULLR* mesh and therefore receives the ADCIRC tidal signal from Lake Pontchartrain instead of Pass Manchac because the Lake Pontchartrain node is closer to the centroid of that particular *egm3* mesh element. At other stations closer to the coast and located in southern Louisiana and Alabama (B, C, D, E, and F) results vary. At station E in
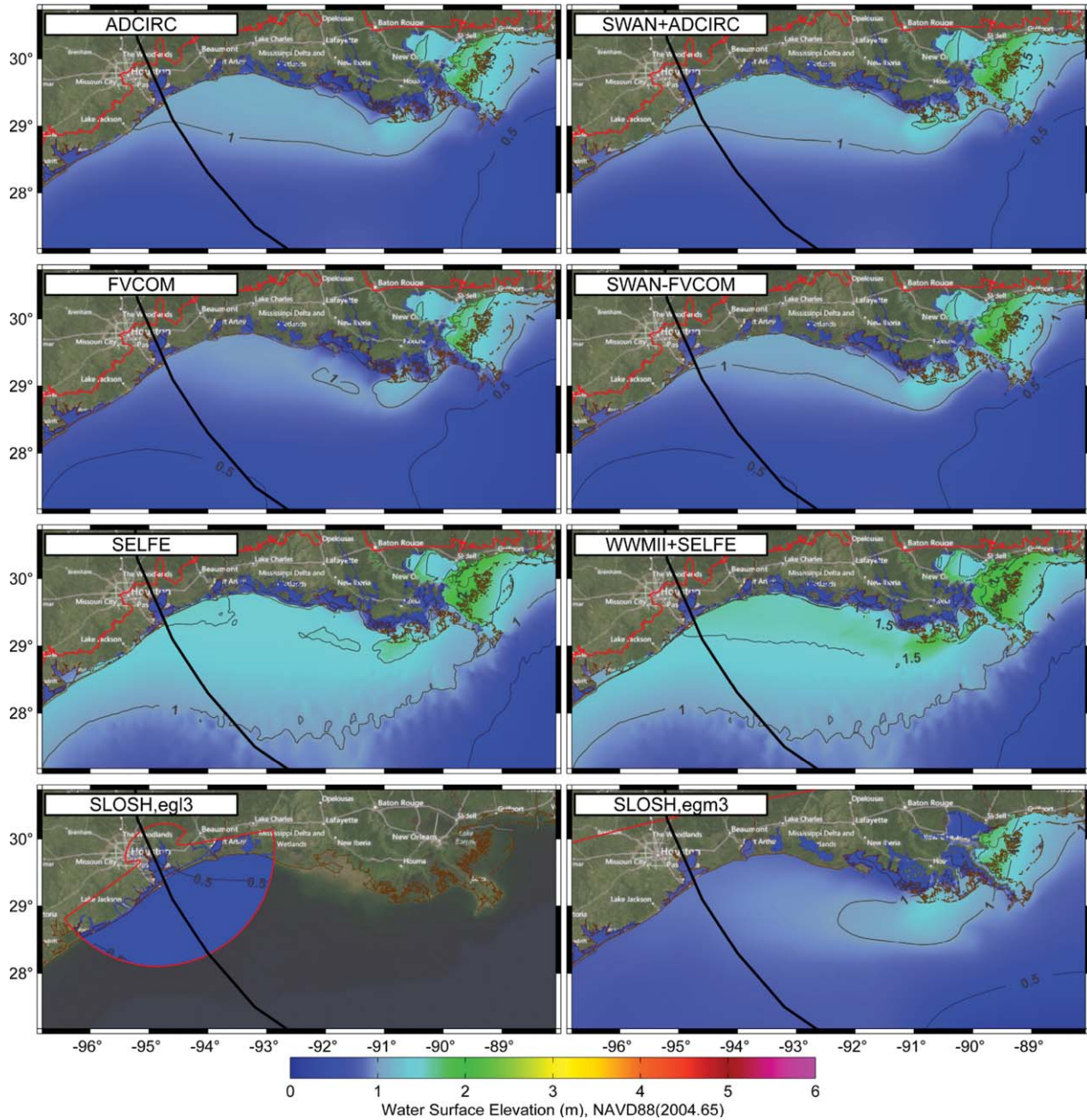
Figure 20. Contour plots of Ike water surface elevation in meters NAVD88 (2004.65) on 0700 UTC 12 September 2008, roughly 24 h before landfall. Also shown are the hurricane track (black) and domain boundary (red).

Terrebonne Bay, all models accurately captured tides and peak surge. Just over 100 km away in Vermillion Bay at station C, all models underpredicted water levels during the storm, indicating a regional deficiency across all models, likely a misrepresentation of bathymetry/topography associated with too coarse of local resolution. Near landfall (stations H, I, J, and K) all models performed acceptably with the exception of SLOSH *egl3*. SLOSH *egl3* underpredicted both the peak water levels as well as the prestorm "forerunner" surge levels. Additionally, SLOSH *egm3* failed to capture the "forerunner" surge but accurately modeled peak water levels.

[71] Coupled circulation and wave model water level time series were also compared to observed water levels (Figure 29). At all stations, all models produced acceptable results for both the "forerunner" surge and peak water level with few exceptions. One of these exceptions is the overprediction of water levels by all models at location F, north of Calcasieu Lake. As shown by *Kerr et al.* [2013b], this area is identified as one where model overprediction is common due to inaccurate representation of conveyance due to poor local resolution by the *ULLR*.

[72] The statistical performance of modeled water level time series are presented in Table 8. For the R^2, the unstructured models performed similarly and accurately, with SWAN+ADCIRC having the highest at 0.759; whereas the SLOSH performance was considerably lower with the largest being 0.498 for SLOSH *egm3*. SWAN+ADCIRC
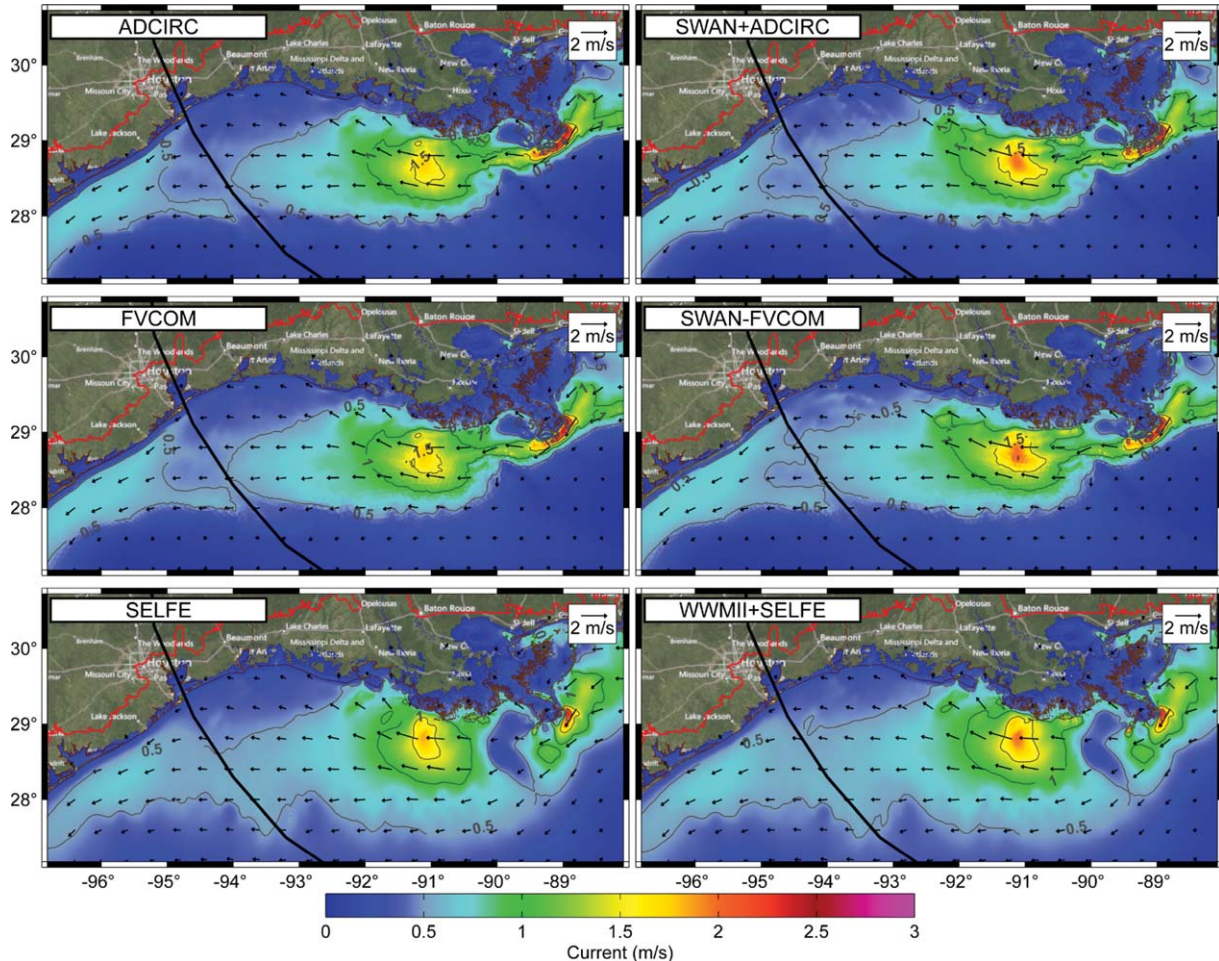
5155

Figure 21. Contour plots of Ike depth-averaged water velocity magnitudes and vectors in meters per second on 0700 UTC 12 September 2008, roughly 24 h before landfall. Also shown are the hurricane track (black) and domain boundary (red).

also has the lowest E_{RMS} as well as sharing the lowest MAE with SWAN-FVCOM. The large negative B_{MN} values for SLOSH models are due to the consistent underprediction of prestorm water levels. Figure S11 of the supporting information for a comparison of water level performance categorized by data source, which graphically illustrates the uniformity of underprediction by SLOSH and the relative similarity of the unstructured models for all data sources.

5.7. High Water Marks

[73] As shown in Figure 30, the maximum water levels during Ike are very similar for the unstructured models but are highly variable for SLOSH. Maximum water levels were compared to measured still water high water marks (HWM) spatially (Figure 31), correlatively (Figure 32), and statistically (Table 9). The spatial distribution of HWMs in Figure 31 clearly points out the significant underprediction of water levels for SLOSH in southeastern Louisiana and the overprediction of water levels in the area of Grand and White Lakes in southwestern Louisiana. Despite overall underprediction, the SLOSH $egm3$ model overpredicts HWMs in the area to the east of Galveston Bay, the area of maximum inundation. In general, SLOSH underpredicted along the coast and near landfall, but overpredicted inland

adjacent to the track. SWAN-FVCOM had the most accurate slope of best fit with SWAN+ADCIRC providing the best R^2. Similar to previous results, all unstructured models perform comparably and accurately, with all model B_{MN} values within 10 cm and $\overline{E}$ values less than 40 cm, which can be considered high skill level for a storm with a maximum surge of over 5 m. The high correlation of the SLOSH $egl3$ run is misleading because the good linear relationship between modeled and measured surge values is way too low and consistently underpredicted.

5.8. Inundation Extents

[74] For Ike, geographic snapshots of the maximum extent of inundation for each of the unstructured mesh models and SLOSH simulations can be found in Figures S14 and S15 of the supporting information. Inundation extents were very similar for each of the unstructured mesh models in the Texas topography. Some notable differences occur in Chambers County where the contribution of waves is more significant as evident by the increased inundation for SWAN+ADCIRC in comparison to ADCIRC. In contrast, the slighter relief of the West Bank of southern Louisiana led to more apparent differences in inundation extents for the unstructured mesh models, although still not overly
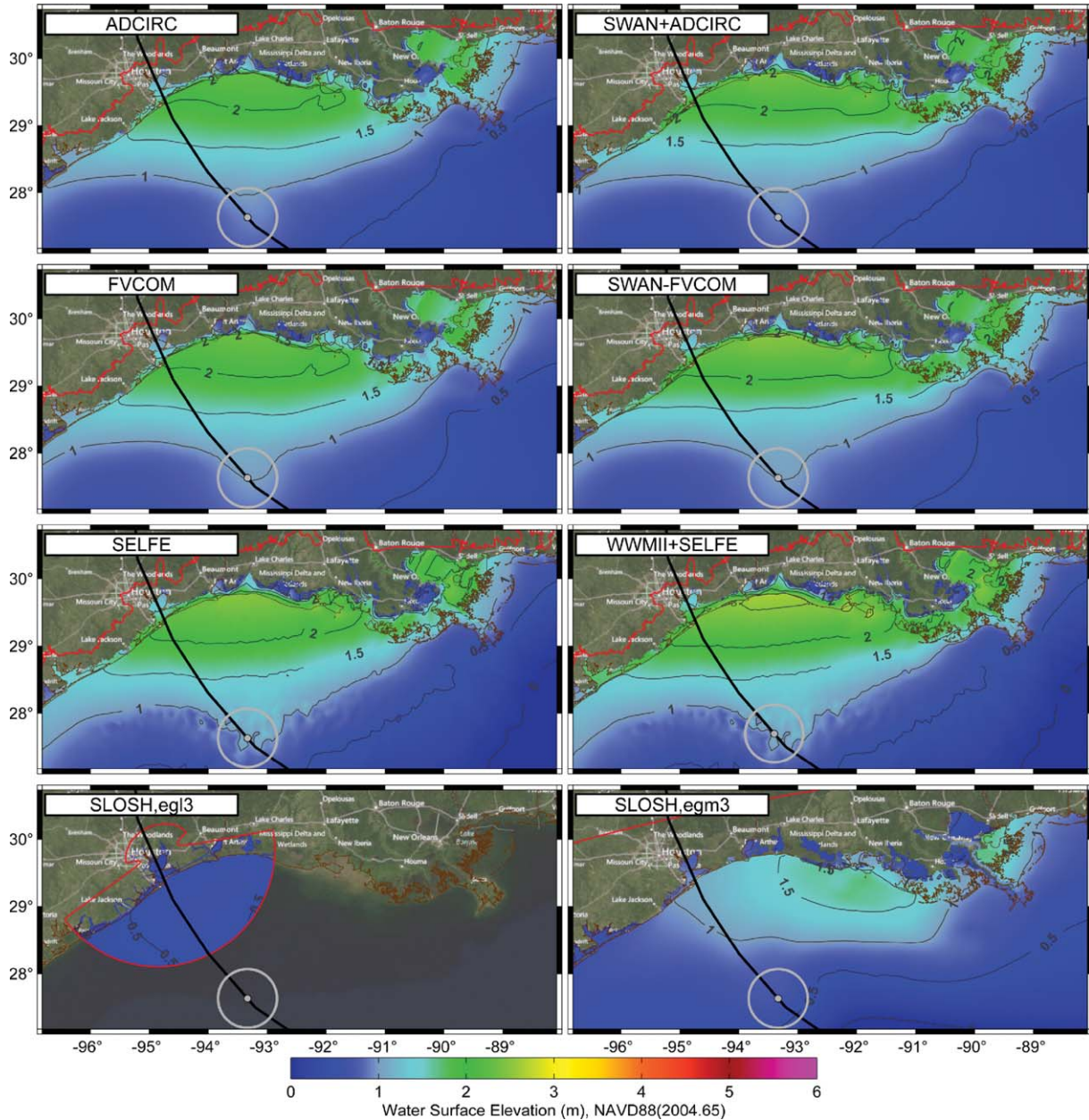
Figure 22. Contour plots of Ike water surface elevation in meters NAVD88 (2004.65) on 1900 UTC 12 September 2008, roughly 12 h before landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and domain boundary (red).

dramatic. The inundation extents for ADCIRC and FVCOM are very similar, whereas the inundation extent for SELFE is noticeably less; the same is true for the coupled wave models. In both Texas and Louisiana, waves contributed to additional inundation not simulated by the models simulated without waves. It was expected that the steeper topography of Texas (in comparison to Louisiana) in conjunction with the landfall of Ike in Texas would lead to a greater contribution of waves to water levels in Texas. While this was the case, as demonstrated by *Kerr et al.* [2013b], the steeper topography of Texas did not lead to a significantly greater inundation extent in comparison to the inundation experienced by the milder relief of Louisiana. This is because the horizontal components of the steeper to-

pography and larger water level increase in Texas balanced relatively evenly with the milder relief and smaller water level increase in Louisiana.

[75] The inundation extents for the SLOSH simulations differ depending on local or Gulf scale meshes. The coarseness of the *egm3* mesh is readily apparent, as seen from its box-like pattern edge in comparison to the visually smoother ADCIRC inundation extent. The *egl3* mesh inundation extent changes from smooth to coarse moving from the Houston Ship Channel to its northeast boundary. It is readily seen that the quality of SLOSH's inundation extent decreases away from the center of the mesh as a result of coarser resolution, boundary effects, and domain coverage. To resolve this inadequacy, SLOSH uses overlapping
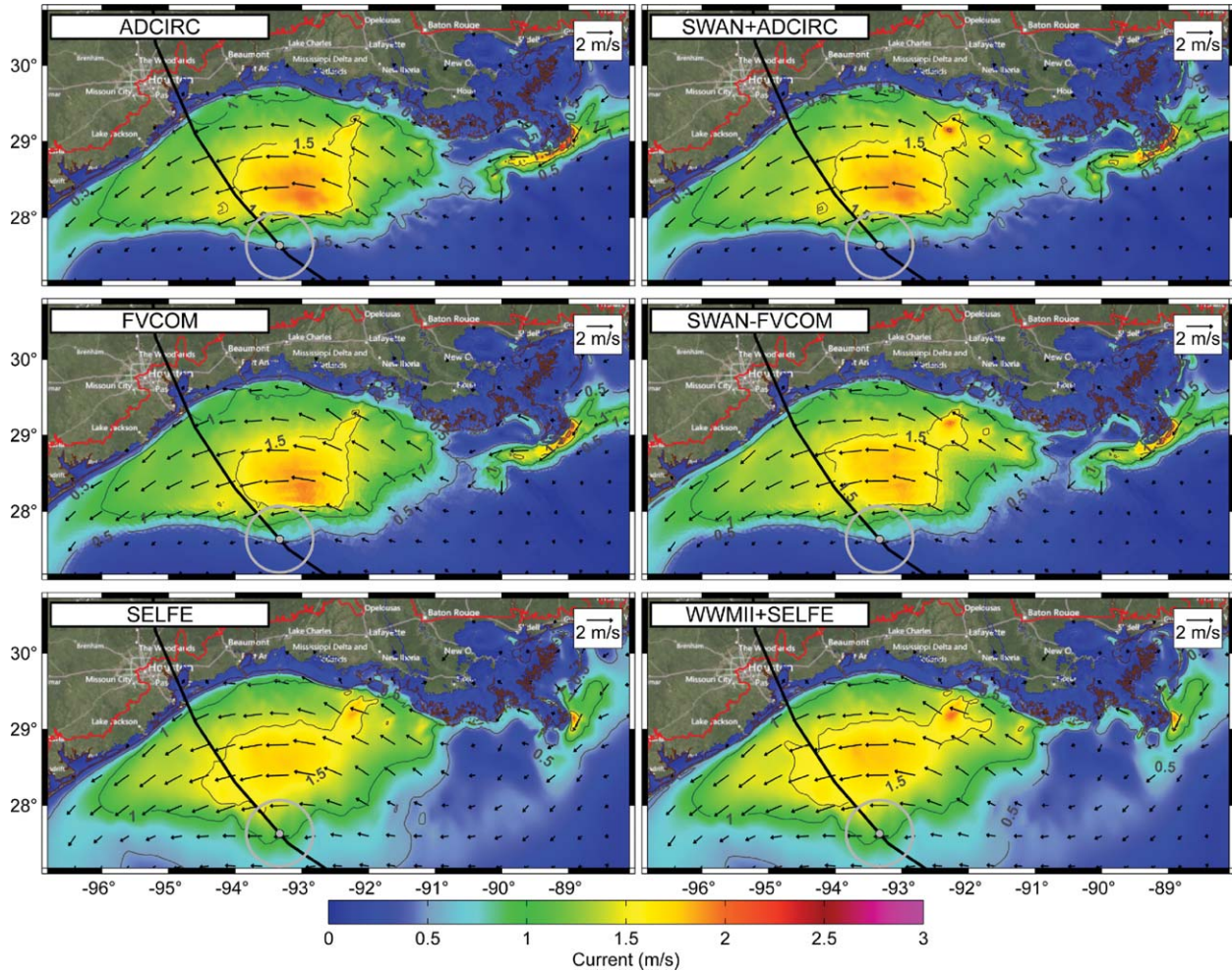
Figure 23. Contour plots of Ike depth-averaged water velocity magnitudes and vectors in meters per second on 1900 UTC 12 September 2008, roughly 12 h before landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and domain boundary (red).

domains so that higher resolution and domain coverage is preserved along the coastal floodplain; this methodology is not necessary for unstructured mesh models. In comparison to the ADCIRC inundation extent, the *egm3* inundation extent was greater and less realistic; whereas the *egl3* near the Houston Ship Channel was more similar and in some cases better resolved with regard to channels and tributaries, but got poorer further away the center of the mesh.

6. Model Scalability/Execution Speed

[76] Operational forecasting requires that simulations be run and the results be output and analyzed within a short time frame so that the appropriate agencies and ultimately the public can be well informed prior to the storm's arrival. To evaluate the (in)feasibility, execution performance, and computer capacity implementation requirements of each of the models, a timing study was performed for a 4 day period of Hurricane Ike (2008). The 4 day period, 3 days before, and 1 day after landfall, was selected to represent an expected lead time to be simulated ahead of an approaching hurricane. It also happens that a 4 day period

is the maximum duration simulation that SLOSH can run. Each of the models were run on the Texas Advanced Computing Center's (TACC) (http://www.tacc.utexas.edu) Ranger (Table 10) with and without 30 min output; where output was the same output listed in Table 4. ADCIRC, SELFE, FVCOM, SWAN+ADCIRC, and WWMII+SELFE were clocked on Ranger's AMD Opteron Quad-Core processors (4 per node) using computational cores of: 64, 128, 256, 512, 1024, and 2048. Unstructured models were all run on the *ULLR* mesh. This is contrast to the tidal and hindcast simulations, where SELFE was run on the extended mesh, *ULLR-EC2001*. The National Weather Service's operational forecast model, SLOSH, is currently not parallelized, so it was run in serial on the *egl3* and *egm3* meshes and its timings along with those of the unstructured models are plotted in Figure 33. A serial ADCIRC simulation was also performed and plotted. SWAN-FVCOM was not included in this analysis on account of the infeasibility of benchmarking its loose coupling method on identical processors.

[77] These results illustrate each model's scalability, which is the ability for a model to improve in speed with
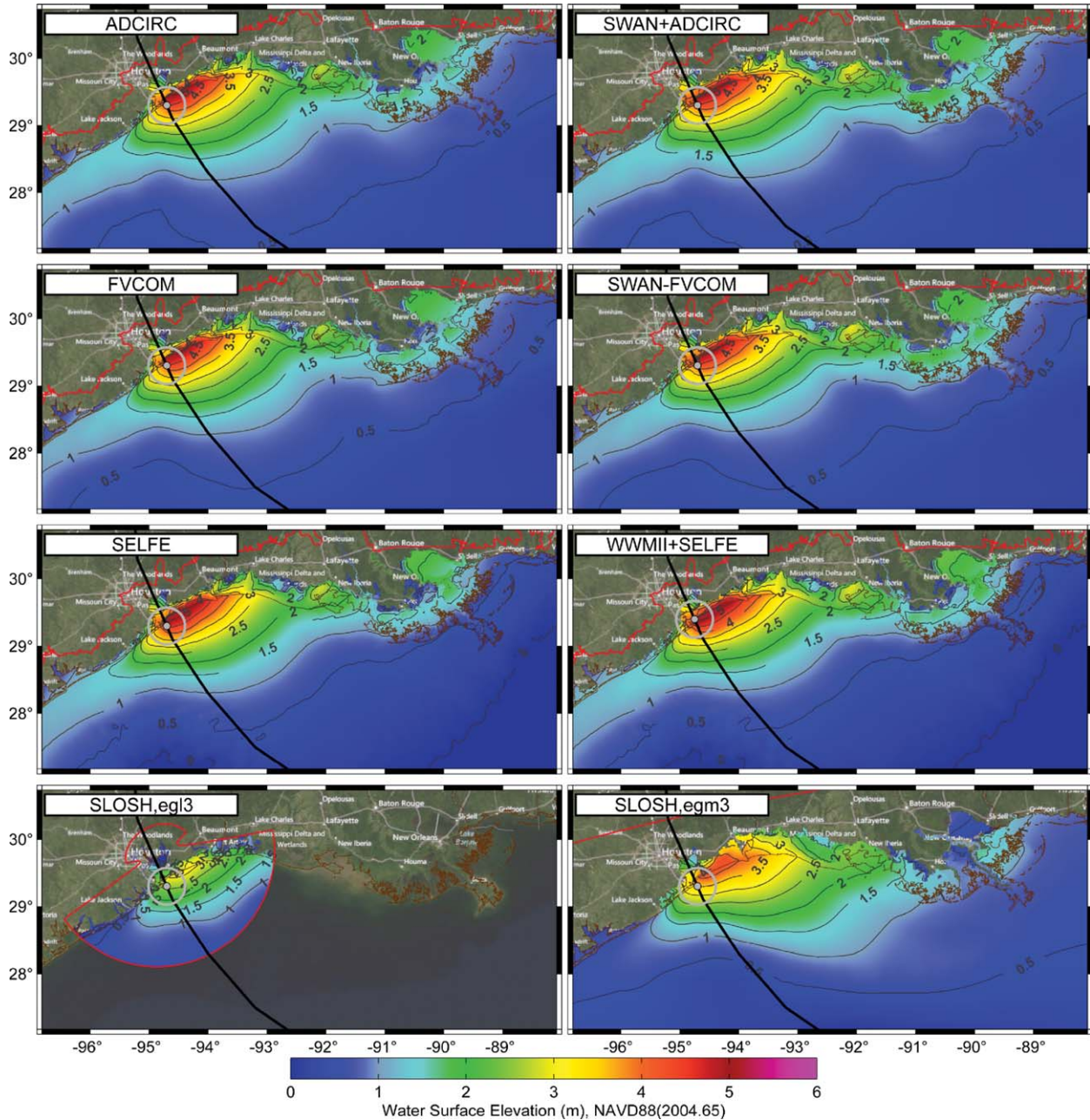
Figure 24. Contour plots of Ike water surface elevation in meters NAVD88 (2004.65) on 0700 UTC 13 September 2008, roughly at landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and domain boundary (red).

increasing numbers of computational cores. If a model is not very scalable, it may not be able to achieve the necessary execution speed for operational forecasting, no matter how many computational cores are used. If a model can meet the necessary execution speed, then the number of computational cores used to meet that represents the computer capacity needed for operational forecasting. ADCIRC, FVCOM, SWAN+ADCIRC, and WWMII+SELFE showed scalability, while SELFE did not. While benchmarking studies by *Tanaka et al.* [2011] and *Dietrich et al.* [2012b] illustrate the strong scalability of ADCIRC and SWAN+ADCIRC, this study is the first to present a comparison of benchmarking for multiple storm surge models.

[78] Model time steps, selected for optimum performance and accuracy, were 4 s for ADCIRC, 1 s for FVCOM, 120 s for SELFE, 600 s for SWAN, 600 s for WWMII, and 8 s for SLOSH (both meshes). An explicit temporal discretization was used for ADCIRC, FVCOM, and SLOSH, while SELFE used a semi-implicit discretization, and the wave models, SWAN and WWMII, used implicit formulations. Despite stability advantages, implicit methods can be more costly than explicit methods and their scalability tends to reach local communication limitations for fewer cores than explicit methods [*Tanaka et al.*, 2011]. At 64 processors, the ADCIRC and FVCOM simulations took 45.6 and 435 min, respectively. The ADCIRC simulations
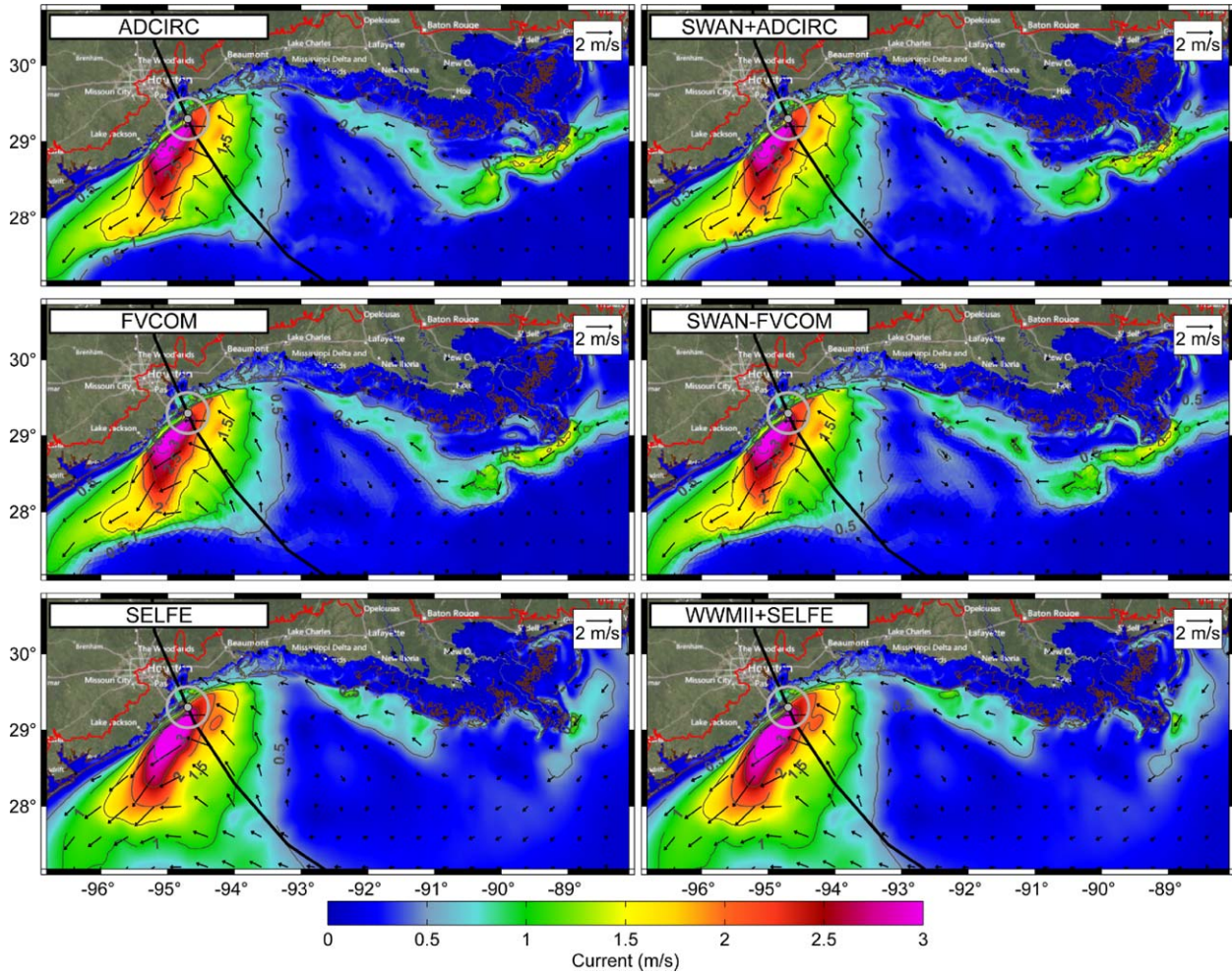
Figure 25. Contour plots of Ike depth-averaged water velocity magnitudes and vectors in meters per second on 0700 UTC 13 September 2008, roughly at landfall. The gray dot and circle identify the center and radius of maximum winds of the storm. Also shown are the hurricane track (black) and domain boundary (red).

Figure 26. Locations of water level and wave observation stations for Ike categorized by data source. Also shown are the ULLR domain boundary (black line) and the track of Ike (green line).

Figure 27. Observed and modeled significant wave height time histories (Date in 2008) for Ike at select stations.

were roughly 10 times faster than the FVCOM simulations, irrespective of number of processors. This difference can be explained by the difference in time steps, time stepping algorithm, the difference in number of equations b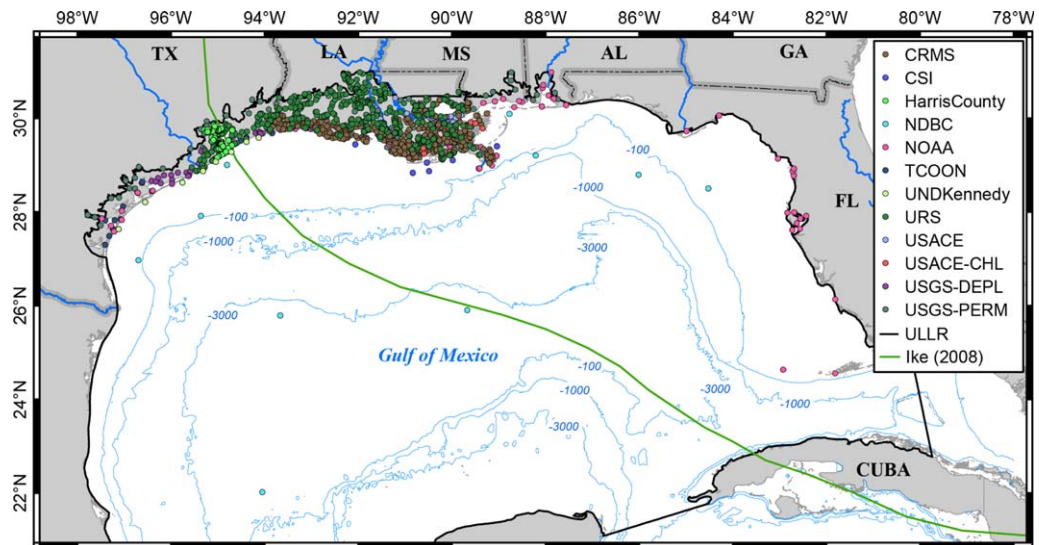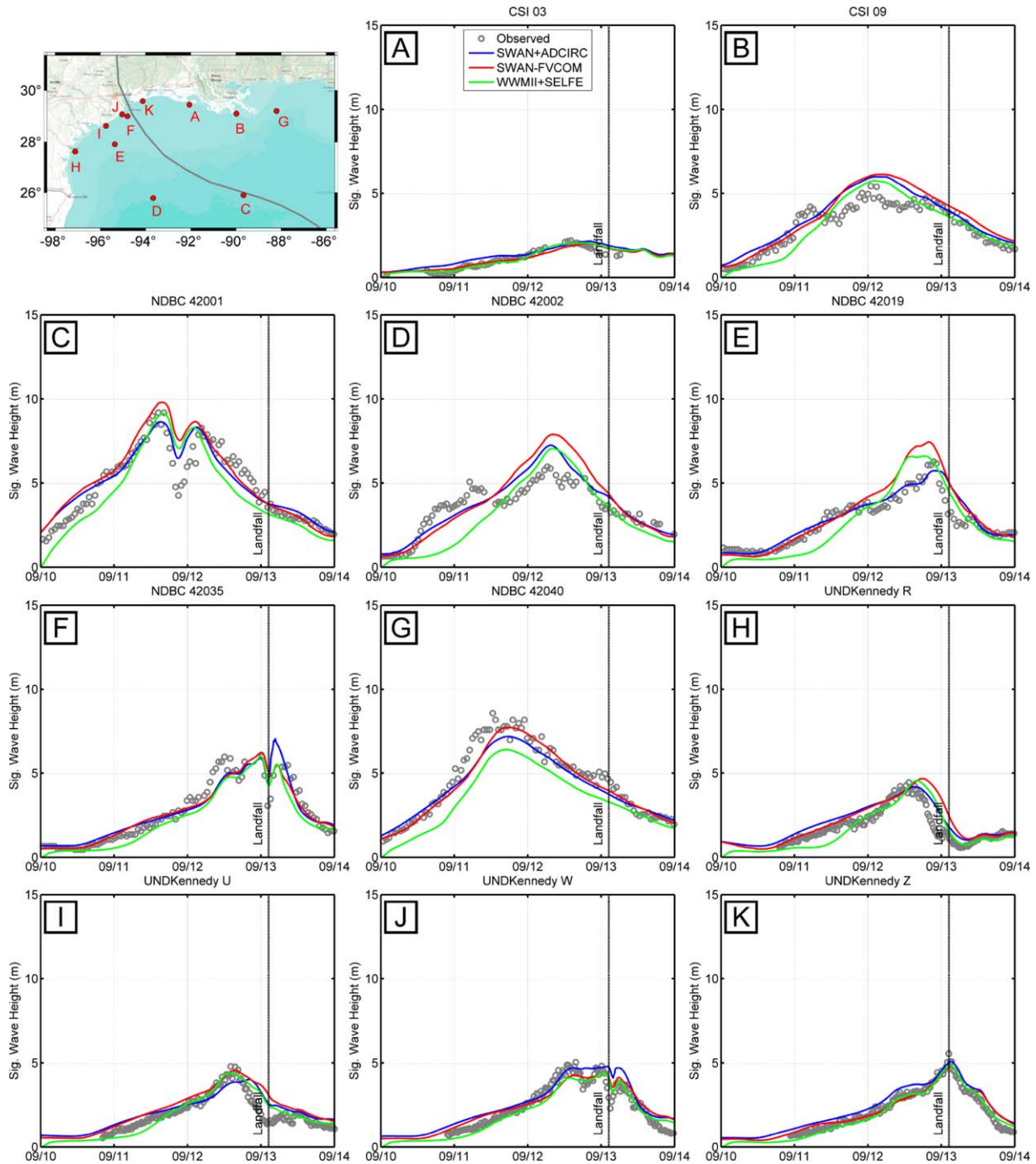eing solved between the finite element method and the finite volume method, and other differences in numerical and logical implementations.

[79] SLOSH's *egl3* mesh (45,792 cells) took 13 min 41 s and SLOSH's *egm3* mesh (185,736 cells) took 65 min 15 s. The *egm3* mesh, at 4.1 times the size of the *egl3* mesh, took

4.8 times longer to run; which suggests that SLOSH scaled fairly well serially and did not exceed memory limits. In contrast to the *egm3* mesh, the *ULLR* mesh has 417,642 nodes or 2.24 times *egm3*'s size. For ADCIRC, the serial simulation took 2452 min or 37.6 times longer than the *egm3* simulation. If the *egm3* timing is scaled to match the time step used by ADCIRC and the number of nodes in the *ULLR* mesh, then the SLOSH simulation would be expected to take 293.4 min or 12% the runtime of ADCIRC. Some reasoning for this dramatic difference can

Table 8. Ike Simulation Wave Characteristic Time Series and Water Level Hydrograph Error Statistics[a]

	Model	R^2	E_{RMS}	$\overline{E}$	B_{MN}	σ	SI	MAE	E_{NORM}	Pts
Wave direction	SWAN+ADCIRC		42.956	−4.813		37.483	0.218	25.649		8
	SWAN-FVCOM		49.103	5.13		42.226	0.252	34.841		8
	WWMII+SELFE		47.708	−2.439		43.176	0.267	31.724		8
Significant wave height	SWAN+ADCIRC	0.836	0.523	0.234	0.423	0.373	0.305	0.43	0.483	25
	SWAN-FVCOM	0.792	0.581	0.19	0.322	0.444	0.256	0.462	0.432	25
	WWMII+SELFE	0.798	0.669	−0.209	0.281	0.486	0.363	0.564	0.533	25
Mean period	SWAN+ADCIRC	0.577	11.392	−5.161	0.111	8.441	0.514	7.571	0.365	19
	SWAN-FVCOM	0.451	13.37	−7.751	−0.369	10.148	0.997	8.426	0.524	19
	WWMII+SELFE	0.357	13.589	−7.821	−0.268	10.272	0.882	8.462	0.45	19
Peak period	SWAN+ADCIRC	0.597	2.239	1.024	0.228	1.824	0.264	1.622	0.363	27
	SWAN-FVCOM	0.516	3.408	−1.101	−0.203	2.76	0.603	2.542	0.704	27
	WWMII+SELFE	0.476	3.155	−0.367	0.1	2.644	0.314	2.31	0.416	27
Water level (wet-only)	ADCIRC	0.759	0.37	−0.143	−0.033	0.174	0.155	0.32	0.32	500
	FVCOM	0.713	0.375	−0.152	−0.045	0.174	0.149	0.323	0.307	494
	SELFE	0.734	0.395	−0.089	0.011	0.211	0.205	0.338	0.368	501
	SLOSH,*egl3*	0.397	1.006	−0.81	−0.51	0.436	0.326	0.855	0.59	79
	SLOSH,*egm3*	0.498	0.638	−0.322	−0.188	0.277	0.244	0.556	0.509	443
	SWAN+ADCIRC	0.77	0.35	−0.013	0.079	0.182	0.167	0.298	0.326	503
	SWAN-FVCOM	0.744	0.351	−0.005	0.087	0.178	0.157	0.298	0.321	501
	WWMII+SELFE	0.724	0.399	0.006	0.095	0.232	0.227	0.335	0.387	503
Water level (TS)	ADCIRC	0.706	0.658	0.127	0.21	0.221	0.217	0.605	0.621	570
	FVCOM	0.682	0.667	0.114	0.198	0.23	0.224	0.612	0.627	570
	SELFE	0.676	0.667	0.189	0.27	0.24	0.244	0.611	0.646	570
	SLOSH,*egl3*	0.303	1.031	−0.518	−0.382	0.481	0.388	0.874	0.678	103
	SLOSH,*egm3*	0.456	0.975	0	0.114	0.322	0.311	0.895	0.906	570
	SWAN+ADCIRC	0.728	0.636	0.217	0.297	0.232	0.236	0.581	0.616	570
	SWAN-FVCOM	0.708	0.641	0.21	0.288	0.24	0.241	0.585	0.619	570
	WWMII+SELFE	0.679	0.671	0.254	0.332	0.26	0.27	0.609	0.66	570

[a]The rows listed as Topo-Substitution (TS) use the station bathymetry at stations where it is not wetted by the model, while the rows listed "Wet-Only" omit any station that is not wetted by the model from the analysis. Some statistics are not applicable for wave direction.

be attributed to: (1) SLOSH's simplification of the Shallow Water Equation which includes the omission of nonlinear advection terms, tidal forcing, and horizontal viscosity; (2) use of single precision in SLOSH instead of double; (3) general differences in logical and numerical implementations, such as wet-dry and bottom friction schemes; and (4) the efficiency of a structured system versus an unstructured system.

[80] The run times for the coupled wave and circulation models were considerably more expensive than just the circulation model. Between 64 and 256 processors, the SWAN+ADCIRC clock time was 11 times ADCIRC's clock time. Both showed strong similarity and scalability up until 512 processors when the scalability for ADCIRC by itself started to bottom out for this relatively small grid. ADCIRC scales on a much higher number of cores for larger meshes [*Tanaka et al.*, 2011; *Dietrich et al.*, 2012b]. The SELFE runtimes appear to be poorly influenced by the number of processors and performed best for a lower core count, suggesting this plot would benefit from additional SELFE timings at lower core counts. This is in contrast to WWMII+SELFE which does show some scalability and improved timings for larger processor counts. Regardless of model, the addition of wave coupling can be expected to significantly increase simulation times.

[81] If the run time for SLOSH's *egm3* model is used as a guideline for other models to meet a reasonable operational forecast speed, the necessary computer capacity for each model can be identified by finding the number of pro-

cessors needed to match or exceed that clock time. The required computer capacity in terms of computational cores on Ranger was 44 for ADCIRC, 64 for SELFE, 600 for SWAN+ADCIRC, and 850 for FVCOM. All three of the unstructured circulation models were able to meet the required run time, but only one of the coupled wave and circulation models, SWAN+ADCIRC was able to do so. It is apparent that SLOSH is significantly less computationally costly than the unstructured models; however, a consideration must be made for the higher accuracy provided by the unstructured models, and run times and scalability are expected to improve for each of these models as computational power increases, and better parallelization methods are applied. Low computational cost and quick execution speed are important factors, considering that operational forecasting is not limited to a single simulation, but rather an ensemble of probable storm deviations. The Probabilistic Hurricane Storm Surge (P-SURGE) model, used by the NWS, runs a high number of simulations to account for the uncertainty in the forecast of the hurricane's track, intensity, and size. Simulations are performed on each of the local-scale meshes for each of the probable variations in storm characteristics. A single unstructured model mesh, with its inherent ability to provide much more extensive coverage along the coast than a structured mesh, could be used to perform all these simulations instead of the high number of overlapping meshes along the coastline currently performed with SLOSH. The unstructured mesh model would still need a computer cluster with a higher
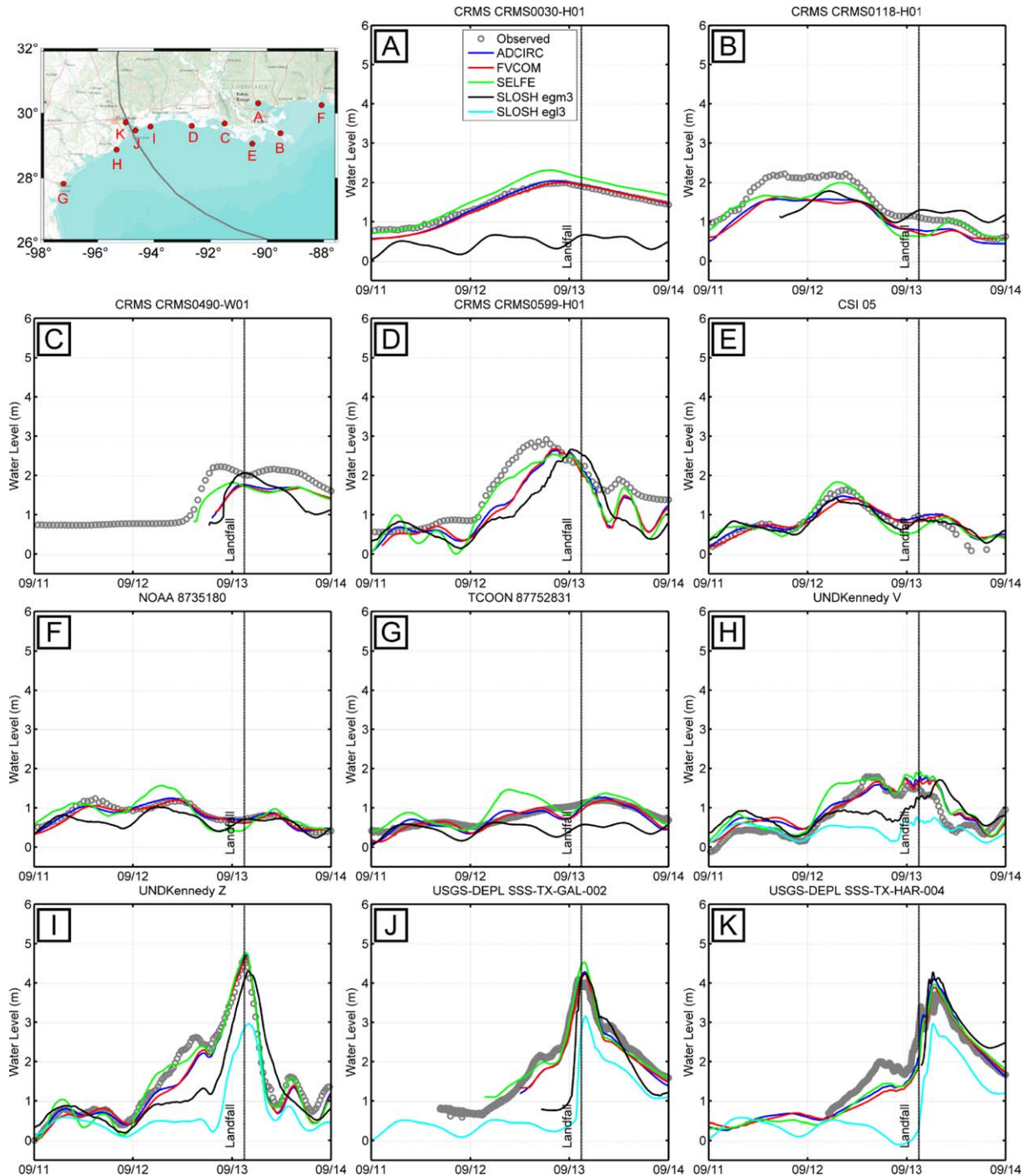
Figure 28. Observed and circulation (without waves) modeled water level time histories (date in 2008) for Ike at select stations.

number of cores necessary to perform similarly with SLOSH, but that number could be divided by the number of additional mesh simulations that would not need to be executed.

7. Discussion

7.1. SLOSH Domain Sensitivity

[82] This study presents the results of simulations using SLOSH for Hurricanes Rita (2005) and Ike (2008) on local

scale (*egl3* for Ike and *ebp3* for Rita) and Gulf scale (*egm3*) meshes. The purpose behind using and comparing the results of two separate size meshes, one local and one Gulf, is to address SLOSH's ability to capture regional processes, and determine how regional processes affect SLOSH's solution in comparison to observed data. In all cases, the Gulf scale mesh produced higher and more accurate water levels than the local-scale meshes. These results match well with the findings by *Blain et al.* [1994] and *Morey et al.* [2006] who suggested that larger domains are important for
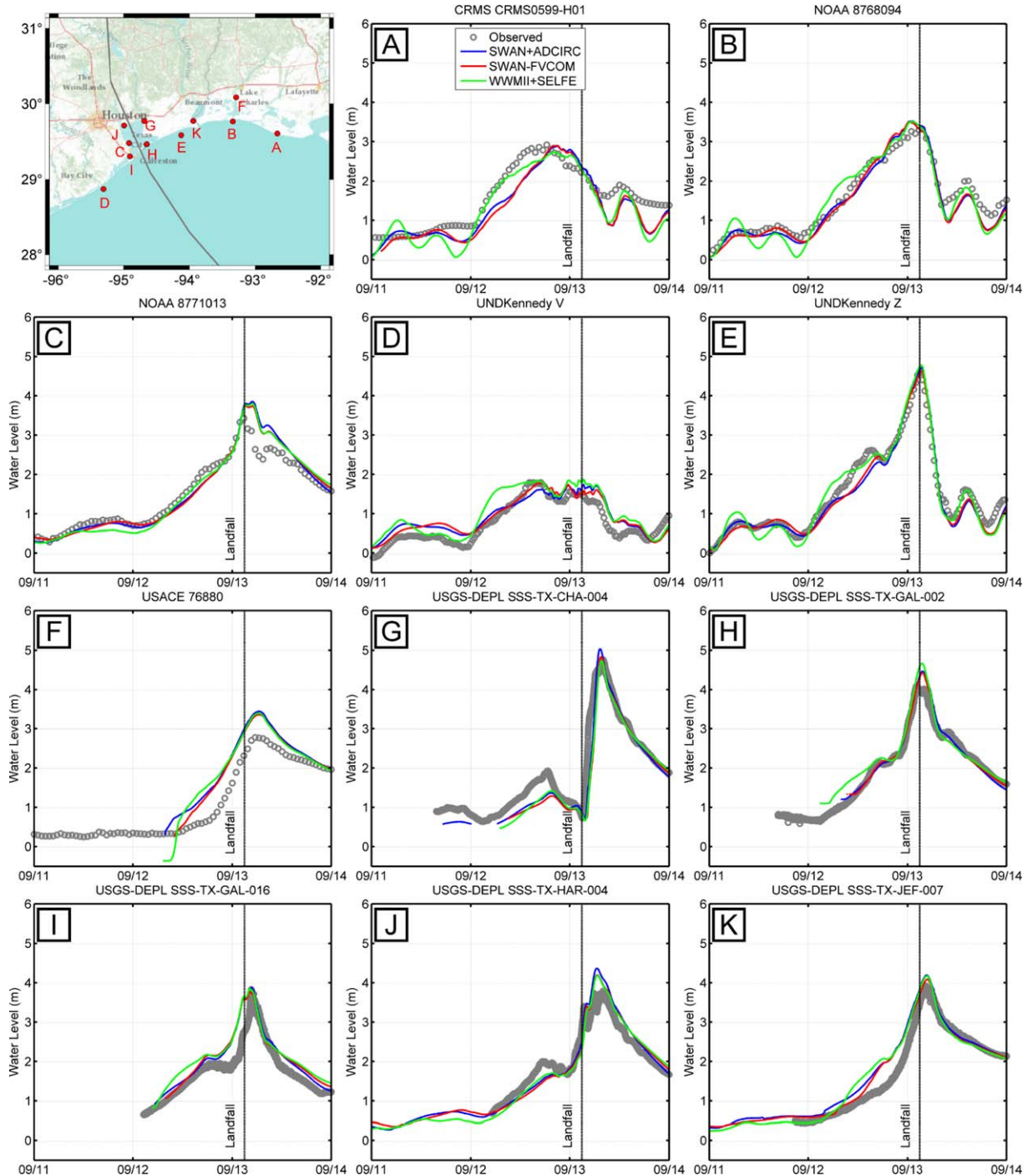
Figure 29. Observed and coupled wave and circulation modeled water level time histories (date in 2008) for Ike at select stations.

correctly capturing processes that may be influenced by a smaller domain boundary.

7.2. Inter-Model Skill Comparison

[83] Results from the harmonic analyses demonstrated similar skill of all the unstructured circulation models at simulating tides on the *ULLR* mesh. While these models showed relatively good agreement with observed harmonics, the performance of these models can be further improved in the inland coastal environment through higher

resolution as recommended by *Kerr et al.* [2013b]. While moderate resolution of the *ULLR* is appropriate for the open coast along the Gulf, higher resolution in the bays, channels, rivers, and around islands, is necessary to correctly capture tidal propagation and attenuation. The SLOSH model was not included in the tidal validation component of the study due to the fact that tides cannot be directly simulated in the model.

[84] Like the tidal study, the results of the Rita and Ike hindcasts were also similar for the unstructured circulation
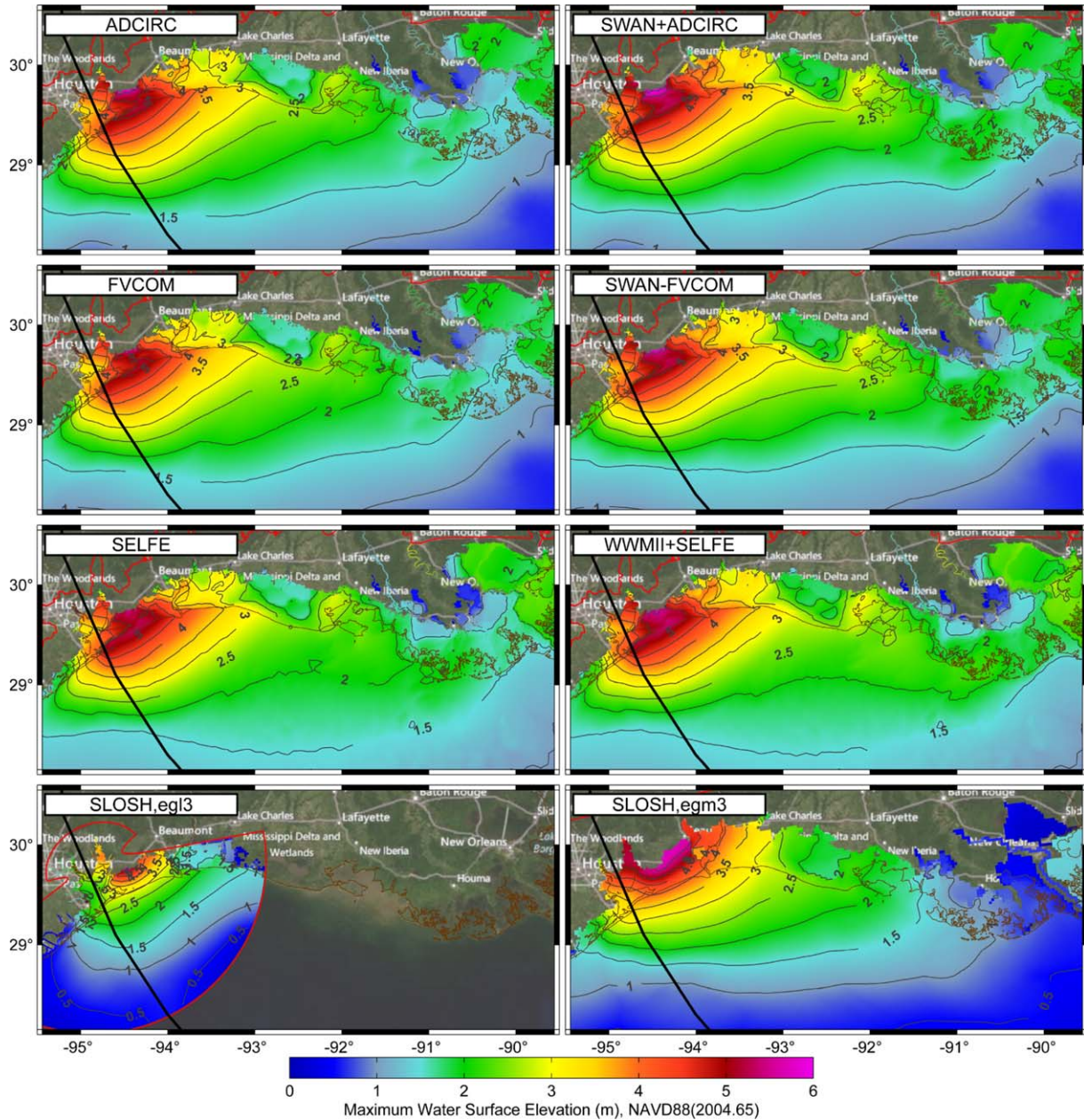
Figure 30. Contour plots of Ike maximum water surface elevation in meters NAVD88 (2004.65). Also shown are the hurricane track (black) and domain boundary (red).

models. And while the unstructured models were similar, it was apparent that SELFE did differ slightly from ADCIRC and FVCOM, but that the differences were generally small. They included a somewhat different velocity field particularly along the LATEX shelf and a slightly higher "forerunner" level with earlier arrivals of high water. This study demonstrated that these three unstructured models can produce similar results when run using the same mesh, identical forcing, and identical friction.

[85] When coupled with wave models, these unstructured models also produced similar results. Each of the wave models used identical wave parameters and winds. The only differences between the wave models were the wind drag law implementation and the circulation data, such as the currents and water levels that were transmitted

to the wave model from the unstructured circulation model. All the models performed extremely well for significant wave heights and mean period. There was some dropoff in accuracy for wave direction and mean peak period for each of the models, with SWAN+ADCIRC generally producing the best results, but overall no coupled wave and circulation model dramatically out-performed or underperformed relative to the other models.

[86] Whereas the results of the unstructured models were very similar, this was not the case when compared to the SLOSH results. There was a significant reduction in accuracy when SLOSH was used. The unstructured circulation models and SLOSH were run using the same wind stress and pressure forcing, but despite this and despite using a Gulf scale domain, the SLOSH runs consistently
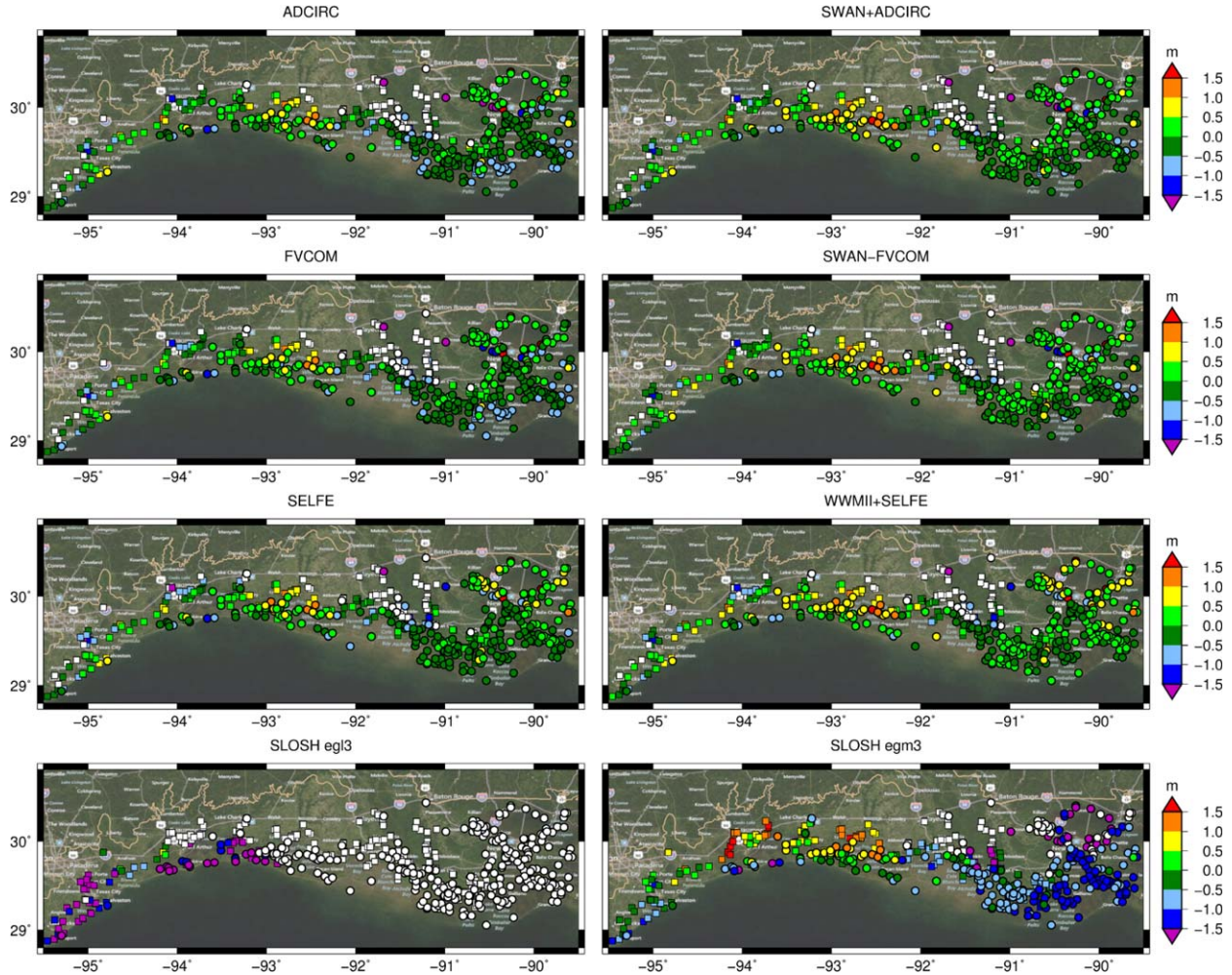
Figure 31. Locations of Ike HWMs (circles) and hydrographs (squares) along the Northwest Gulf coast. The points are color-coded to show the errors between measured and modeled peak water levels. Green points indicate matches within 0.5 m and white points indicate locations that were never wetted by the model.

underperformed in comparison to ADCIRC, FVCOM, and SELFE. For Rita water levels, the ADCIRC R^2 and B_{MN} were 0.759 and 0.044, respectively; whereas the R^2 and B_{MN} were 0.522 and -0.055, respectively, for SLOSH *egm3*. The differences for Ike are even more dramatic than for Rita, due to early arrival of high water from the forerunner. For Ike water levels, the ADCIRC R^2 and B_{MN} were 0.759 and -0.033, respectively; whereas the R^2 and B_{MN} were 0.313 and -0.188, respectively, for SLOSH *egm3*. The poor performance of the SLOSH *egm3* is visually more stunning for high water marks as seen in Figures 17, 18, 31, and 32 where the SLOSH high water marks clearly under perform in comparison to the unstructured models.

[87] Besides execution speed, the four primary ways SLOSH differs inherently from the unstructured models ADCIRC, FVCOM, and SELFE are: (1) nonlinear advection, (2) tides, (3) resolution, and (4) friction. The inclusion of wind waves is currently being addressed and therefore is not included in this list. The first item on the list is nonlinear advection, which is not included in SLOSH, because its contribution was considered negligible by *Jelesnianski et al.* [1992]. As *Kerr et al.* [2013b] demonstrated, the geo-strophic setup that preceded Ike was influenced by advection and the inclusion of these nonlinear terms improves model solutions for the forerunner and shelf waves.

[88] The omission of tidal physics is also a difference between SLOSH and the unstructured models, and the remedial method of superimposing tidal signals onto SLOSH water levels has drawbacks. While the time scale of tidal ebb and flow is not a major concern, the tide-surge interaction is entirely missing. This interaction can be strong and the simple addition of tides to the surge-induced water levels is often inadequate [*Horsburgh and Wilson*, 2007; *Rego and Li*, 2010b]. In addition, a host of resolution problems detract from the accuracy of the superpositioning of signals. If the resolution of the SLOSH mesh is too coarse to reflect the intertidal environment or if the tidal database is too coarse, there is the potential for unrealistic tidal signals in channels, wetlands, and other intertidal systems.

[89] As this study and previous studies have identified, Gulf scale meshes are necessary to capture Gulf scale physics. It was clear that the Gulf scale *egm3* model better captures Gulf scale physics than the local scale *ebp3* and *egl3* meshes. While the *ebp3* and *egl3* local-scale meshes
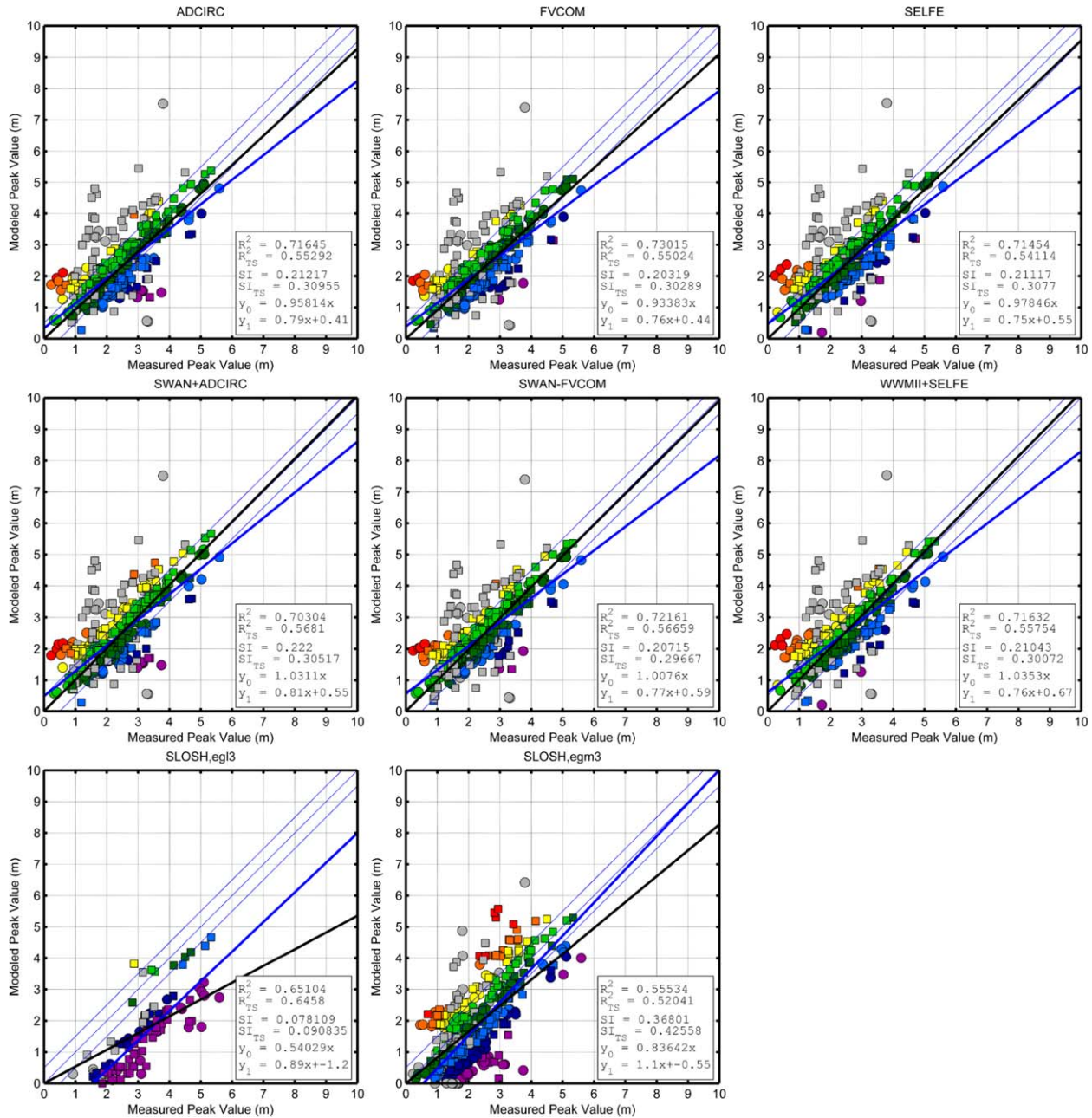
Figure 32. Scatter plots of Ike HWMs (circles) and peak hydrograph levels (squares) for 606 stations. Red, orange, yellow, and light green points indicate overprediction by the model; dark green, blue, dark blue, and purple points indicate underprediction. Green points indicate a match within 0.5 m and gray points indicate locations that were never wetted by the model. Metrics shown are for the TS method, unless they have the subscript "C" which refers to "wet-Only" method. The thick blue line and the thick black line represent the y_1 and y_0 best fit lines, respectively.

have fine resolution similar to the *ULLR*, the *egm3* does not and is 5–10 times coarser in these areas. *Kerr et al.* [2013b] found that the resolution provided by the *ULLR* was successful in capturing Gulf scale physics, but it was inadequate at simulating processes at the floodplain and channel scale; and that higher-resolution models such as *SL18TX33* are more accurate for inundation. This deficiency inherent to coarser resolution in inland areas is also seen in some of the SLOSH *egm3* results. Ultimately, the development of a high-resolution Gulf scale structured

SLOSH mesh would lead to unnecessary and excessive resolution in the Gulf and necessitate a parallelized framework for SLOSH due to the increased number of nodes. This highlights the advantage of unstructured meshes in that they can resolve the coastal and inland areas without unnecessarily over-resolving the Gulf.

[90] The last difference to discuss between SLOSH and the unstructured models is the implementation of bottom friction. ADCIRC, FVCOM, and SELFE use a spatially varying quadratic Manning's *n* formulation; whereas

Table 9. Ike Simulation HWM Error Statistics[a]

	Model	R^2	E_{RMS}	$\overline{E}$	B_{MN}	σ	SI	MAE	E_{NORM}	Dry	Out	Wet
HWM (TS)	ADCIRC	0.553	0.649	−0.027	−0.013	0.649	0.31	0.439	0.286	75	0	531
	FVCOM	0.55	0.638	−0.073	−0.035	0.635	0.303	0.436	0.281	80	0	526
	SELFE	0.541	0.645	0.037	0.018	0.645	0.308	0.435	0.284	76	0	530
	SWAN-FVCOM	0.567	0.63	0.104	0.049	0.622	0.297	0.42	0.278	76	0	530
	SWAN+ADCIRC	0.568	0.655	0.146	0.07	0.639	0.305	0.441	0.289	68	0	538
	WWMII+SELFE	0.558	0.653	0.173	0.083	0.63	0.301	0.44	0.288	74	0	532
	SLOSH,*egl3*	0.646	1.698	−1.551	−0.498	0.283	0.091	1.582	0.517	16	504	86
	SLOSH,*egm3*	0.52	0.987	−0.424	−0.203	0.892	0.426	0.85	0.435	70	0	536
HWM (wet-only)	ADCIRC	0.716	0.486	−0.102	−0.049	0.445	0.212	0.358	0.214	75	0	531
	FVCOM	0.73	0.474	−0.129	−0.061	0.426	0.203	0.353	0.209	80	0	526
	SELFE	0.715	0.472	−0.025	−0.012	0.441	0.211	0.351	0.208	76	0	530
	SWAN-FVCOM	0.722	0.47	0.067	0.032	0.435	0.207	0.34	0.206	76	0	530
	SWAN+ADCIRC	0.703	0.502	0.089	0.042	0.466	0.222	0.365	0.221	68	0	538
	WWMII+SELFE	0.716	0.486	0.128	0.061	0.44	0.21	0.357	0.215	74	0	532
	SLOSH,*egl3*	0.651	1.781	−1.646	−0.501	0.257	0.078	1.673	0.52	16	504	86
	SLOSH,*egm3*	0.555	0.961	−0.459	−0.213	0.794	0.368	0.837	0.412	70	0	536

[a]The rows listed as Topo-Substitution (TS) use the station bathymetry at stations where it is not wetted by the model, while the rows listed "wet-only" omit any station that is not wetted by the model from the analysis.

SLOSH is internally parameterized with a mild depth-dependent linear Ekman-based friction formulation [*Platzman*, 1963]. Friction is an extremely important part of capturing hurricane physics and has been the focus of numerous studies. *Bunya et al.* [2010] and *Dietrich et al.* [2011b] are examples of recent hurricane validation studies that identified significant improvements in model response through the use of a spatially varying bottom friction coefficient classified by land-use and vegetation types for the coastal floodplains. Previous studies [*Dietrich et al.*, 2011b; *Kerr et al.*, 2013b; *Hope et al.*, 2013] noted the importance of classifying the bottom friction coefficient of the LATEX shelf (muddy) differently than the Florida shelf (sandy), because material-type influences friction. Because SLOSH's friction formulation does not take into account land-use or bottom material, is linear in nature, and mildly depends on depth, it did not develop the strong geostrophic setup that occurred during Ike, even when using the same wind stress as the unstructured models, because it was overdamped on the shelf on account of its bottom friction. Based on comparisons of water level time series, SLOSH's internal frictional formulation is also responsible for a lack of overland dissipation, which led to much higher levels further inland than what was observed or simulated by the unstructured mesh models.

8. Conclusions

[91] As part of the SURA-led U.S. IOOS-funded Coastal and Ocean Modeling Testbed's objective to evaluate and improve models already in operational use as well as facilitating the transition of additional models to operational use, a Gulf of Mexico region specific performance evaluation of coastal inundation, surge, and wave models was executed through the assessment of tidal harmonics, hindcasts of Hurricanes Rita (2005) and Ike (2008), and model execution speed. Three unstructured coastal and ocean models, ADCIRC, FVCOM, and SELFE, were simulated on a new common Gulf scale mesh (ULLR) created specifically for the Testbed, with identical frictional parameterization and

forcing for the tidal harmonic analyses and hurricane hindcasts. Due to the large time steps used by SELFE and the need to specify boundary conditions, SELFE and WWMII+SELFE used an extended mesh (*ULLR-EC2001*) for the tidal and hindcast simulations but not for the benchmarking. NOAA's official operational forecast storm surge model, SLOSH was also included in the hindcast studies and was simulated on both local and Gulf scale meshes.

[92] A comparative tidal harmonic analysis using ADCIRC, FVCOM, and SELFE was conducted using 59 NOAA observation stations in the Gulf of Mexico, a 75 day water level time series, and 38 tidal constituents for decomposition. Using the O_1, K_1, P_1, Q_1, M_2, S_2, N_2, and K_2 constituents for comparison, the results and statistical analyses demonstrated that each of the models performed with the same relative level of accuracy. The R^2 values for the full set of these tidal harmonic constituent amplitudes were 0.81, 0.81, and 0.79, for ADCIRC, FVCOM, and SELFE, respectively. The SLOSH model was not included in the tidal validation component of the study due to the fact that tides are not directly simulated in the model.

[93] For this study, two Gulf of Mexico hurricanes were selected for analysis: Hurricanes Rita (2005) and Ike (2008). These storms were selected for their large-scale impacts on the LATEX coastline, unique individual characteristics, and wealth of recorded wave and water level data. The primary quantitative differences between Ike and Rita

Table 10. Specifications of the Compute Nodes on the TACC Ranger Machine

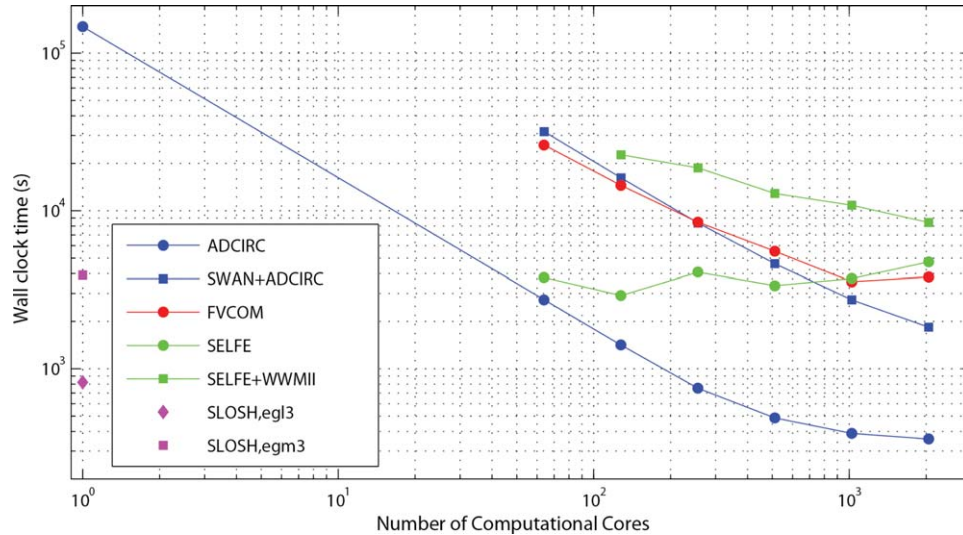| Node	Sun Blade ×6420
CPU	4 Quad-core AMD Opteron 8356
Core	AMD Opteron 8356
Frequency	2.3 GHz
Architecture	AMD K10 (Barcelona)
L1-Cache	64+64 KB per core
L2-Cache	512KB per core
L3-Cache	2MB on die shared

Figure 33. Benchmarking results for ADCIRC (blue circles), SWAN+ADCIRC (blue squares), FVCOM (red circles), SELFE (green circles), WWMII+SELFE (green squares), SLOSH *egl3* (magenta diamond), and SLOSH *egm3* (magenta square) on TACC's Ranger.

were that Ike was a larger storm and Rita was a more intense storm, but also of importance are the slight differences in track and landfall. The combination of Ike's size and track were critical for the development of the "forerunner" that preceded it. Ike's steady track across the Gulf combined with its unusually large wind field created a strong shore-parallel shelf current driven by shore-parallel winds in the 36 h before landfall. This current caused a geostrophic setup known as a "forerunner" surge to be created along the LATEX coast filling coastal lakes and bays reaching up to 2 m before Ike made landfall [*Kennedy et al.*, 2011].

[94] To validate these hindcasts an extensive and quality controlled database of high water marks, water level time series and wave characteristic time series was gathered and prepared through cleaning and filtering of stations from a wide variety of sources for each of the hindcasts and made freely available along with each model simulation's Input and Output as part of the Testbed's cyberinfrastructure mandate. The three unstructured models, ADCIRC, FVCOM, and SELFE all showed similar skill in modeling the two hurricanes. Tightly coupled SWAN+ADCIRC and WWMII+SELFE along with FVCOM loosely coupled with SWAN also validated with similar skill for the hurricane hindcasts. The HWM R^2 values ranged from 0.56 to 0.62 for all the unstructured mesh model runs, including Ike and Rita. Whereas the moderate resolution of the *ULLR* mesh performs well in the open Gulf, accuracy in inland areas can be further increased with the use of higher-resolution meshes. *Kerr et al.* [2013b] found that TS HWM R^2 increased from 0.568 for the *ULLR* mesh to 0.7 for the *SL18TX33* mesh and that TS water level R^2 increased from 0.713 for the *ULLR* mesh to 0.765 for the *SL18TX33*. Aside from water levels and inundation, waves in themselves are significant for navigation, erosion, and forces on structures, onshore and offshore. While waves are a minor contributor to water levels in most areas of Louisiana and Texas, they become more significant in steep-sloped regions.

[95] SLOSH was simulated for both hurricanes on both local and Gulf scale meshes using the same data-assimilated OWI meteorological forcing based on H*wind/IOKA used by the unstructured models. SLOSH's local meshes failed to capture regional processes such as Ike's "forerunner." SLOSH's model response is sensitive to domain size because regional and Gulf scale processes are not correctly modeled by the smaller meshes. SLOSH does have the ability to account for regional processes like geostrophic setup, but since its friction is too high on the continental shelf, it could not develop the strong shore-parallel currents that powered the "forerunner" preceding Ike. The geostrophic setup created by the SLOSH model for Ike was much less than what the unstructured models produced and underpredicted observation data indicating that SLOSH's model was overdamped on the LATEX shelf.

[96] SLOSH performed poorly in comparison to the unstructured models for Hurricanes Rita and Ike. In general, the HWM *MAE* was 2–3 times greater for the SLOSH simulations than for the unstructured mesh model simulations. The unstructured models have an intrinsic advantage relating to mesh resolution that allows for coarser resolution in the Gulf and higher resolution at the coastal floodplain and channel scale that can't be achieved with SLOSH's structured meshes. The unstructured models also have a better implementation of bottom friction and additional physics built in into the model such as tides and nonlinear advection that contribute to overall water levels and currents during storms. The transition from structured to unstructured represents the current state of evolution of coastal and ocean modeling, but it is expected that these models can be significantly improved in their efficiency with higher order methods [*Wirasaet et al.*, 2010].

[97] The unstructured models were benchmarked on TACC's Ranger using 64–2048 computational cores in order to assess and compare execution speed and scalability for a prototypical operational simulation. Ranger uses AMD Opteron Quad-Core Processors (4 per node). SLOSH

in its serial framework was also benchmarked and used to represent the speed necessary for the unstructured models to meet operational implementation requirements. Each of the unstructured models was capable of matching or beating the timing of the SLOSH *egm3* simulation, but a higher number of computational cores are needed in comparison to the single core used by SLOSH. It was also apparent that SELFE was not scalable and that it could benefit from effective parallelization. The coupling of wave models to storm surge models does significantly increase timings and computational cost. For example, in order for ADCIRC and SWAN+ADCIRC to meet the same timing as SLOSH on Ranger, they would require 44 and 600 cores, respectively. While SLOSH is significantly less computationally costly than the unstructured models, the unstructured mesh models deliver a significantly higher accuracy in water level and inundation. Run times and scalability are expected to improve for each of these models as computational power increases and more efficient high-order algorithms and parallelization methods are developed and applied.

References

Atkinson, J., J. Westerink, and R. Luettich Jr. (2004a), Two-dimensional dispersion analyses of finite element approximations to the shallow water equations, *Int. J. Numer. Methods Fluids*, *45*(7), 715–749.

Atkinson, J., J. Westerink, and J. Hervouet (2004b), Similarities between the quasi-bubble and the generalized wave continuity equation solutions to the shallow water equations, *Int. J. Numer. Methods Fluids*, *45*(7), 689–714.

Battjes, J., and J. Janssen (1978), Energy loss and setup due to breaking of random waves, paper presented at the 16th International Conference on Coastal Engineering, Hamburg, Germany.

Battjes, J., and M. Stive (1985), Calibration and verification of a dissipation model for random breaking waves, *J. Geophys. Res.*, *90*(C5), 9159–9167.

Bertin, X., A. Oliveira, and A. B. Fortunato (2009), Simulating morphodynamics with unstructured grids: Description and validation of a modeling system for coastal applications, *Ocean Modell.*, *28*(1), 75–87.

Bertin, X., N. Bruneau, J.-F. Breilh, A. B. Fortunato, and M. Karpytchev (2012), Importance of wave age and resonance in storm surges: The case Xynthia, Bay of Biscay, *Ocean Modell.*, *42*, 16–30, doi:10.1016/j.ocemod.2011.11.001.

Blain, C., J. Westerink, and R. Luettich Jr. (1994), The influence of domain size on the response characteristics of a hurricane storm surge model, *J. Geophys. Res.*, *99*(C9), 18,467–18,479.

Booij, N., R. Ris, and L. Holthuijsen (1999), A third-generation wave model for coastal regions: 1. Model description and validation, *J. Geophys. Res.*, *104*(C4), 7649–7666.

Brovchenko, I., V. Maderich, and K. Terletska (2011), Numerical simulations of 3D structure of currents in the region of deep canyons on the east coast of the Black Sea, *Int. J. Comput. Civil Struct. Eng.*, *7*(2), 47–53.

Bunya, S., et al. (2010), A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part I: Model development and validation, *Mon. Weather Rev.*, *138*(2), 345–377.

Burla, M., A. M. Baptista, Y. Zhang, and S. Frolov (2010), Seasonal and interannual variability of the Columbia River plume: A perspective enabled by multiyear simulation databases, *J. Geophys. Res.*, *115*, C00B16, doi:10.1029/2008JC004964.

Cardone, V., and A. Cox (2009), Tropical cyclone wind field forcing for surge models: Critical issues and sensitivities, *Nat. hazards*, *51*(1), 29–47.

Chen, C., H. Liu, and R. C. Beardsley (2003), An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: Application to coastal ocean and estuaries, *J. Atmos. Oceanic Technol.*, *20*(1), 159–186.

Chen, C., R. C. Beardsley, and G. Cowles (2006a), An unstructured grid, finite-volume coastal ocean model: FVCOM user manual, School for Marine Science and Technology/University of Massachusetts Dartmouth (SMAST/UMASSD).

Chen, C., R. C. Beardsley, and G. Cowles (2006b), An unstructured grid, finite-volume coastal ocean model (FVCOM) system, Special issue entitled "Advances in computational oceanography," *Oceanography*, *19*(1), 78–89.

Chen, C., H. Huang, R. C. Beardsley, H. Liu, Q. Xu, and G. Cowles (2007), A finite volume numerical approach for coastal ocean circulation studies: Comparisons with finite difference models, *J. Geophys. Res.*, *112*, C03018, doi:10.1029/2006JC003485.

Chen, C., J. Qi, C. Li, R. C. Beardsley, H. Lin, R. Walker, and K. Gates (2008), Complexity of the flooding/drying process in an estuarine tidal-creek salt-marsh system: An application of FVCOM, *J. Geophys. Res.*, *113*, C07052, doi:10.1029/2007JC004328.

Chen, C., G. Gao, J. Qi, A. Proshutinsky, R. C. Beardsley, Z. Kowalik, H. Lin, and G. Cowles (2009), A new high-resolution unstructured grid finite volume Arctic Ocean model (AO-FVCOM): An application for tidal studies, *J. Geophys. Res.*, *114*, C08017, doi:10.1029/2008JC004941.

Chen, C., H. Huang, R. C. Beardsley, Q. Xu, R. Limeburner, G. W. Cowles, Y. Sun, J. Qi, and H. Lin (2011), Tidal dynamics in the Gulf of Maine and New England shelf: An application of FVCOM, *J. Geophys. Res.*, *116*, C12010, doi:10.1029/2011JC007054.

Cho, K.-H., H. V. Wang, J. Shen, A. Valle-Levinson, and Y. cheng Teng (2012), A modeling study on the response of Chesapeake Bay to hurricane events of Floyd and Isabel, *Ocean Modell.*, *49*50, 22–46, doi:10.1016/j.ocemod.2012.02.005.

Cox, A. T., J. Greenwood, V. Cardone, and V. Swail (1995), An interactive objective kinematic analysis system, in *Fourth International Workshop on Wave Hindcasting and Forecasting*, pp. 109–118, Atmos. Environ. Serv., Banff, Alberta, Canada.

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*(1), 63–88.

Dietrich, J., et al. (2010), A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part II: Synoptic description and analysis of hurricanes Katrina and Rita, *Mon. Weather Rev.*, *138*(2), 378–404.

Dietrich, J., M. Zijlema, J. Westerink, L. Holthuijsen, C. Dawson, R. Luettich Jr., R. Jensen, J. Smith, G. Stelling, and G. Stone (2011a), Modeling hurricane waves and storm surge using integrally-coupled, scalable computations, *Coastal Eng.*, *58*(1), 45–65.

Dietrich, J., et al. (2011b), Hurricane Gustav (2008) waves and storm surge: Hindcast, synoptic analysis, and validation in southern Louisiana, *Mon. Weather Rev.*, *139*(8), 2488–2522.

Dietrich, J., et al. (2012a), Surface trajectories of oil transport along the northern coastline of the Gulf of Mexico, *Cont. Shelf Res.*, *41*(1), 17–47, doi:10.1016/j.csr.2012.03.015.

Dietrich, J., S. Tanaka, J. Westerink, C. Dawson, R. Luettich, M. Zijlema, L. Holthuijsen, J. Smith, L. Westerink, and H. Westerink (2012b), Performance of the unstructured-mesh, SWAN+ADCIRC model in computing hurricane waves and surge, *J. Sci. Comput.*, *52*(2), 468–497, doi:10.1007/s10915-011-9555-6.

Dietrich, J., et al. (2012c), Limiters for spectral propagation velocities in SWAN, *Ocean Modell.*, *70*, 85–102, doi:10.1016/j.ocemod.2012.11.005.

DiMarco, S. F., and R. O. Reid (1998), Characterization of the principal tidal current constituents on the Texas-Louisiana shelf, *J. Geophys. Res.*, *103*(C2), 3093–3109, doi:10.1029/97JC03289.

Ebersole, B., J. Westerink, D. Resio, and R. Dean (2007), Performance evaluation of the New Orleans and southeast Louisiana hurricane protection system, volume IV–the storm, Final report of the interagency performance evaluation task force, technical report, 263 pp., U.S. Army Corps of Eng., Washington, D. C.

Egbert, G. D., and S. Y. Erofeeva (2002), Efficient inverse modeling of barotropic ocean tides, *J. Atmos. Oceanic Technol.*, *19*(2), 183–204.

Egbert, G. D., A. F. Bennett, and M. G. Foreman (1994), TOPEX/Poseidon tides estimated using a global inverse model, *J. Geophys. Res.*, *99*(C12), 24,821–24,852.

FEMA (2009), Flood Insurance Study: Southeastern Parishes, Louisiana: Offshore Water Levels and Waves, U.S. Army Corps of Engineers, New Orleans, La.

Fleming, J., C. Fulcher, R. Luettich, B. Estrade, G. Allen, and H. Winer (2008), A real time storm surge forecasting system using ADCIRC, in *Estuarine and Coastal Modeling X*, edited by M. Spaulding, pp. 373–392, Am. Soc. Civ. Eng., Reston, Va.

Forbes, C., and J. Rhome (2012), An automated operational storm surge prediction system for the national hurricane center, in *Estuarine and Coastal Modeling*, edited by M. Spaulding, pp. 213–229, Am. Soc. Civ. Eng., Reston, Va.

Foreman, M., R. Walters, and R. Henry (1993), A model for simulating currents in eastern Juan de Fuca Strait and the southern Strait of Georgia, in *Proceedings of OCEANS'93. Engineering in Harmony with Ocean*, Oceanic Engineering Society, IEEE Council on Oceanic Engineering, Victoria, B. C., *C12*, 28,629–28,651, doi:10.1109/OCEANS.1993.325987.

Gao, G., C. Chen, J. Qi, and R. C. Beardsley (2011), An unstructured-grid, finite-volume sea ice model: Development, validation, and application, *J. Geophys. Res.*, *116*, C00D04, doi:10.1029/2010JC006688.

Glahn, B., A. Taylor, N. Nurkowski, and W. Shaffer (2009), The role of the SLOSH model in national weather service storm surge forecasting, *Natl. Weather Dig.*, *33*(1), 3–14.

Gouillon, F., S. L. Morey, D. S. Dukhovskoy, and J. J. O'Brien (2010), Forced tidal response in the Gulf of Mexico, *J. Geophys. Res.*, *115*, C10050, doi:10.1029/2010JC006122.

Holland, G. J. (1980), An analytic model of the wind and pressure profiles in hurricanes, *Mon. Weather Rev.*, *108*(8), 1212–1218.

Hope, M. E., et al. (2013), Hindcast and validation of Hurricane Ike (2008) waves, forerunner, and storm surge, *J. Geophys. Res. Oceans*, in Press.

Horsburgh, K., and C. Wilson (2007), Tide-surge interaction and its role in the distribution of surge residuals in the North Sea, *J. Geophys. Res.*, *112*, C08003, doi:10.1029/2006JC004033.

Huang, H., C. Chen, G. W. Cowles, C. D. Winant, R. C. Beardsley, K. S. Hedstrom, and D. B. Haidvogel (2008), FVCOM validation experiments: Comparisons with roms for three idealized barotropic test problems, *J. Geophys. Res.*, *113*, C07042, doi:10.1029/2007JC004557.

Huang, Y., R. H. Weisberg, and L. Zheng (2010), Coupling of surge and waves for an Ivan-like hurricane impacting the Tampa Bay, Florida region, *J. Geophys. Res.*, *115*, C12009, doi:10.1029/2009JC006090.

Huang, Y., R. H. Weisberg, L. Zheng, and M. Zijlema (2013), Gulf of Mexico hurricane wave simulations using SWAN: Bulk formula based drag coefficient sensitivity for Hurricane Ike, *J. Geophys. Res. Oceans*, doi:10.1002/jgrc.20283, in press.

Jelesnianski, C. P. (1966), Numerical computations of storm surges without bottom stress, *Mon. Weather Rev.*, *94*(6), 379–394.

Jelesnianski, C. P. (1967), Numerical computations of storm surges with bottom stress, *Mon. Weather Rev.*, *95*(11), 740–756.

Jelesnianski, C. P., J. Chen, and W. A. Shaffer (1992), SLOSH: Sea, Lake, and Overland Surges from Hurricanes, U.S. Dep. of Commer., Natl. Oceanic and Atmos. Admin., Natl. Weather Serv., Silver Spring, Md.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*(3), 437–471.

Kennedy, A. B., U. Gravois, B. C. Zachry, J. J. Westerink, M. E. Hope, J. C. Dietrich, M. D. Powell, A. T. Cox, R. A. Luettich Jr., and R. G. Dean (2011), Origin of the Hurricane Ike forerunner surge, *Geophys. Res. Lett.*, *38*, L08608, doi:10.1029/2011GL047090.

Kerr, P., et al. (2013a), Surge generation mechanisms in the lower Mississippi River and discharge dependency, *J. Waterway Port Coastal Ocean Eng.*, *139*(4), 326–335, doi:10.1061/(ASCE)WW.1943–5460.000018.

Kerr, P. C., R. Martyr, A. Donahue, M. Hope, J. Westerink, R. Luettich Jr., J. Dietrich, C. Dawson, and H. Westerink (2013b), 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*, doi:10.1002/jgrc.20305, in press.

Knabb, R. D., D. P. Brown, and J. Rhome (2006), Tropical cyclone report, Hurricane Rita, 18–26 Sep 2005, report, Natl. Hurricane Cent., Miami.

Kolar, R., W. Gray, J. Westerink, and R. Luettich Jr. (1994), Shallow water modeling in spherical coordinates: Equation formulation, numerical implementation, and application, *J. Hydraul. Res.*, *32*(1), 3–24.

Lai, Z., C. Chen, G. W. Cowles, and R. C. Beardsley (2010a), A nonhydrostatic version of FVCOM: 1. Validation experiments, *J. Geophys. Res.*, *115*, C11010, doi:10.1029/2009JC005525.

Lai, Z., C. Chen, G. W. Cowles, and R. C. Beardsley (2010b), A nonhydrostatic version of FVCOM: 2. Mechanistic study of tidally generated non-

linear internal waves in Massachusetts Bay, *J. Geophys. Res.*, *115*, C12049, doi:10.1029/2010JC006331.

Le Provost, C., M. Genco, F. Lyard, P. Vincent, and P. Canceil (1994), Spectroscopy of the world ocean tides from a finite element hydrodynamic model, *J. Geophys. Res.*, *99*(C12), 24,777–24,797.

Le Provost, C., A. Bennett, and D. Cartwright (1995), Ocean tides for and from TOPEX/Poseidon, *Science*, *267*, 639–639.

Longuet-Higgins, M. S., and R. Stewart (1964), Radiation stresses in water waves: a physical discussion, with applications, *Deep Sea Res. Oceanogr. Abstr.*, *11*(4), 529–562.

Luettich, R., Jr., J. Westerink, 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, *Dredging Research Program Tech. Rep. DRP-92-6*, 137 pp., U.S. Army Engineers Waterways Experiment Station, Vicksburg, Miss.

Luettich, R. A., and J. J. Westerink (2004), Formulation and Numerical Implementation of the 2D/3D ADCIRC Finite Element Model Version 44. XX. http://www.unc.edu/ims/adcirc/publications/2004/2004_Luettich.pdf.

Lyard, F., F. Lefevre, T. Letellier, and O. Francis (2006), Modelling the global ocean tides: Modern insights from FES2004, *Ocean Dyn.*, *56*(5), 394–415.

Martyr, R., et al. (2013), Simulating hurricane storm surge in the lower Mississippi River under varying flow conditions, *J. Hydraul. Eng.*, *139*(5), 492–501, doi:10.1061/(ASCE)HY.1943–7900.0000699.

McGee, B. D. (2006), Hurricane Rita Surge Data, Southwestern Louisiana and Southeastern Texas, September to November 2005, *Data Ser. 220*, U.S. Dep. of the Interior, U.S. Geol. Surv., Reston, Va.

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, doi:10.1029/2006GL027021.

Mousavi, M., J. Irish, A. Frey, F. Olivera, and B. Edge (2010), Global warming and hurricanes: The potential impact of hurricane intensification and sea level rise on coastal flooding, *Clim. Change*, *104*, 575–597.

Mukai, A., J. Westerink, R. Luettich Jr., and D. Mark (2002), Eastcoast 2001, a tidal constituent database for western North Atlantic, Gulf of Mexico, and Caribbean Sea, *Tech. Rep. ERDC/CHL TR-02-24*, 201 pp., US Army Eng. Res. and Dev. Cent., Coastal and Hydraulics Lab., Vicksburg, Miss.

Ning, L., K. Emanuel, M. Openheimer, and E. Vanmarcke (2012), Physically based assessment of hurricane surge threat under climate change, *Nat. Clim. Change*, *2*(6), 462–467, doi:10.1038/nclimate1389.

Pawlowicz, R., B. Beardsley, and S. Lentz (2002), Classical tidal harmonic analysis including error estimates in Matlab using T_TIDE, *Comput. Geosci.*, *28*(8), 929–937.

Pinto, L., A. Fortunato, Y. Zhang, A. Oliveira, and F. Sancho (2012), Development and validation of a three-dimensional morphodynamic modelling system for non-cohesive sediments, *Ocean Modell.*, *5758*, 1–14, doi:10.1016/j.ocemod.2012.08.005.

Platzman, G. W. (1963), The dynamica prediction of wind tides on Lake Erie, *Meterol. Monogr.*, *4*(26), 44.

Powell, M. (2006), Drag coefficient distribution and wind speed dependence in tropical cyclones, final report to the National Oceanic and Atmospheric Administration (NOAA) Joint Hurricane Testbed (JHT) program, technical report, Atl. Oceanogr. and Meteorol. Lab., Miami, Fla.

Powell, M. D., S. H. Houston, and T. A. Reinhold (1996), Hurricane Andrew's landfall in South Florida. Part I: Standardizing measurements for documentation of surface wind fields, *Weather Forecasting*, *11*(3), 304–328.

Powell, M. D., S. H. Houston, L. R. Amat, and N. Morisseau-Leroy (1998), The HRD real-time hurricane wind analysis system, *J. Wind Eng. Ind. Aerodyn.*, *77*, 53–64.

Powell, M. D., P. J. Vickery, and T. A. Reinhold (2003), Reduced drag coefficient for high wind speeds in tropical cyclones, *Nature*, *422*(6929), 279–283.

Powell, M. D., et al. (2010), Reconstruction of Hurricane Katrina's wind fields for storm surge and wave hindcasting, *Ocean Eng.*, *37*(1), 26–36.

Rego, J. L., and C. Li (2010a), Storm surge propagation in Galveston Bay during Hurricane Ike, *J. Mar. Syst.*, *82*(4), 265–279.

Rego, J. L., and C. Li (2010b), Nonlinear terms in storm surge predictions: Effect of tide and shelf geometry with case study from Hurricane Rita, *J. Geophys. Res.*, *115*, C06020, doi:10.1029/2009JC005285.

Reid, R. (1990), *Tides and Storm Surges, Handbook of Coastal and Ocean Engineering, Volume 1—Wave Phenomena and Coastal Structures*, Gulf Publ., Houston, Tex.

Ris, R., L. Holthuijsen, and N. Booij (1999), A third-generation wave model for coastal regions: 2. Verification, *J. Geophys. Res.*, *104*(C4), 7667–7681.

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*(2), 366–389.

Roland, A., A. Cucco, C. Ferrarin, T.-W. Hsu, J.-M. Liau, S.-H. Ou, G. Umgiesser, and U. Zanke (2009), On the development and verification of a 2-D coupled wave-current model on unstructured meshes, *J. Mar. Syst.*, *78*, S244–S254.

Roland, A., Y. J. Zhang, H. V. Wang, Y. Meng, Y.-C. Teng, V. Maderich, I. Brovchenko, M. Dutour-Sikiric, and U. Zanke (2012), A fully coupled 3D wave-current interaction model on unstructured grids, *J. Geophys. Res.*, *117*, C00J33, doi:10.1029/2012JC007952.

Schureman, P. (1958), Manual of harmonic analysis and prediction of tides, *Special Publication 98*, U.S. Dept. of Commer., Coast and Geodetic Surv., Washington, D. C.

Tanaka, S., S. Bunya, J. Westerink, C. Dawson, and R. Luettich (2011), Scalability of an unstructured grid continuous galerkin based hurricane storm surge model, *J. Sci. Comput.*, *46*(3), 329–358.

Taylor, A., and B. Glahn (2008), Probablistic guidance for hurricane storm surge, Preprints, 19th Conference on Probability and Statistics, New Orleans, LA, Amer. Meteor. Soc., 7.4.

Thompson, E. F., and V. J. Cardone (1996), Practical modeling of hurricane surface wind fields, *J. Waterway Port Coastal Ocean Eng.*, *122*(4), 195–205.

Tolman, H. (1999), User Manual and System Documentation of WAVEWATCH-III Version 1.18., NOAA/NWS/NCEP/OMB, *Tech. Note 166*, 110 pp. [Available at http://polar.ncep.noaa.gov/mmab/papers/tn166/OMB_166.pdf.]

Uhlhorn, E. W., P. G. Black, J. L. Franklin, M. Goodberlet, J. Carswell, and A. S. Goldstein (2007), Hurricane surface wind measurements from an operational stepped frequency microwave radiometer, *Mon. Weather Rev.*, *135*(9), 3070–3085.

USACE (2009), Hydraulics and hydrology appendix, Louisiana coastal protection and restoration final technical report, U.S. Army Corps of Eng., New Orleans, La.

Vickery, P. J., D. Wadhera, M. D. Powell, and Y. Chen (2009), A hurricane boundary layer and wind field model for use in engineering applications, *J. Appl. Meteorol. Climatol.*, *48*(2), 381–405.

Wahr, J. M. (1981), Body tides on an elliptical, rotating, elastic and ocean-less earth, *Geophys. J. R. Astron. Soc.*, *64*(3), 677–703.

Weisberg, R. H., and L. Zheng (2006a), Circulation of Tampa Bay driven by buoyancy, tides, and winds, as simulated using a finite volume coastal ocean model, *J. Geophys. Res.*, *111*, C01005, doi:10.1029/2005JC003067.

Weisberg, R. H., and L. Zheng (2006b), Hurricane storm surge simulations for Tampa Bay, *Estuaries Coasts*, *29*(6), 899–913.

Weisberg, R. H., and L. Zheng (2008), Hurricane storm surge simulations comparing three-dimensional with two-dimensional formulations based on an ivan-like storm over the Tampa Bay, Florida region, *J. Geophys. Res.*, *113*, C12001, doi:10.1029/2008JC005115.

Westerink, J. J., R. A. Luettich, J. C. Feyen, J. H. Atkinson, C. Dawson, H. J. Roberts, M. D. Powell, J. P. Dunion, E. J. Kubatko, and H. Pourtaheri (2008), A basin-to channel-scale unstructured grid hurricane storm surge model applied to southern Louisiana, *Mon. Weather Rev.*, *136*(3), 833–864.

Wirasaet, D., S. Tanaka, E. J. Kubatko, J. J. Westerink, and C. Dawson (2010), A performance comparison of nodal discontinuous galerkin methods on triangles and quadrilaterals, *Int. J. Numer. Methods Fluids*, *64*(10-12), 1336–1362, doi:10.1002/fld.2376.

Wu, J. (1982), Wind-stress coefficients over sea surface from breeze to hurricane, *J. Geophys. Res.*, *87*(C12), 9704–9706.

Yanenko, N. (1972), *The Method of Fractional Steps*, 160 pp., Springer, New York.

Yang, Z., and T. Khangaonkar (2008), Modeling of salt intrusion, intertidal mixing, and circulation in a braided estuary, *J. Coastal Res.*, *52*, 171–180.

Zhang, Y., and A. M. Baptista (2008), Selfe: A semi-implicit eulerian–lagrangian finite-element model for cross-scale ocean circulation, *Ocean Modell.*, *21*(3), 71–96.

Zheng, L., and R. H. Weisberg (2010), Rookery Bay and Naples Bay circulation simulations: Applications to tides and fresh water inflow regulation, *Ecol. Modell.*, *221*(7), 986–996, doi:10.1016/j.ecolmodel.2009.01.024.

Zheng, L., and R. H. Weisberg (2012), Modeling the west Florida coastal ocean by downscaling from the deep ocean, across the continental shelf and into the estuaries, *Ocean Modell.*, *48*, 10–29, doi:10.1016/j.ocemod.2012.02.002.

Zijlema, M. (2010), Computation of wind-wave spectra in coastal waters with SWAN on unstructured grids, *Coastal Eng.*, *57*(3), 267–277.