Detailed modeling of recent severe storm tides in estuaries of the New York City region

Philip Orton,[1] Nickitas Georgas,[1] Alan Blumberg,[1] and Julie Pullen[1]

Received 20 May 2012; revised 6 August 2012; accepted 15 August 2012; published 25 September 2012.

[1] Detailed simulations, comparisons with observations, and model sensitivity experiments are presented for the August 2011 tropical cyclone Irene and a March 2010 nor'easter that affected the New York City (NYC) metropolitan area. These storms brought strong winds, heavy rainfall, and the fourth and seventh highest gauged storm tides (total water level), respectively, at the Battery, NYC. To dissect the storm tides and examine the role of various physical processes in controlling total water level, a series of model experiments was performed where one process was omitted for each experiment, and results were studied for eight different tide stations. Neglecting remote meteorological forcing (beyond ~250 km) led to typical reductions of 7–17% in peak storm tide, neglecting water density variations led to typical reductions of 1–13%, neglecting a parameterization that accounts for enhanced wind drag due to wave steepness led to typical reductions of 3–12%, and neglecting atmospheric pressure gradient forcing led to typical reductions of 3–11%. Neglecting freshwater inputs to the model domain led to reductions of 2% at the Battery and 9% at Piermont, 14 km up the Hudson River from NYC. Few storm surge modeling studies or operational forecasting systems incorporate the "estuary effects" of freshwater flows and water density variations, yet joint omission of these processes for Irene leads to a low-bias in storm tide for NYC sites like La Guardia and Newark Airports (9%) and the Battery (7%), as well as nearby vulnerable sites like the Indian Point nuclear plant (23%).

Citation: Orton, P., N. Georgas, A. Blumberg, and J. Pullen (2012), Detailed modeling of recent severe storm tides in estuaries of the New York City region, *J. Geophys. Res.*, *117*, C09030, doi:10.1029/2012JC008220.

1. Introduction

[2] Coastal storms are among the world's most costly disasters, with strong winds, floodwater inundation, and coastal erosion capable of damaging and disabling infrastructure. New York City (NYC) is highly vulnerable to storm surge flooding, with some subways and highway tunnels at risk of being shut down after coastal flood elevations only ~2.25 m above local mean sea level (LMSL) [*Colle et al.*, 2008]. NYC was struck by hurricanes three times from 1700 to 1900, with storm tides (total water levels) estimated to be roughly 2.5–3.6 m LMSL [*Scileppi and Donnelly*, 2007]. More frequent, moderate storms can bring storm tides of 2.0–2.5 m that can still threaten infrastructure, such as an extra-tropical "nor'easter" storm that flooded and shut down the subway system for several days in 1992. With a daily gross metropolitan product of about $3.5 billion in 2010, even these moderate flooding events can prove extremely costly.

[3] Storm surge modeling is an important tool for operational flood forecasting systems and for flood hazard assessments, particularly for understanding risks from these moderate to severe storms that have occurred only a few times, if any, in the historical record. In the past, highly simplified storm surge models were favored over more detailed models, and today, most storm surge modeling is still performed with two-dimensional (2D) hydrodynamic models [*Resio and Westerink*, 2008]. This reliance on simplified models occurs in part because of incomplete understanding of the exact role and mechanisms of many factors that contribute to coastal water elevations. It also occurs because storm surge risk assessment and probabilistic forecasts can require dozens or even hundreds of storm simulations, and therefore having short model run times is paramount.

[4] A recent study [*Weisberg and Zheng*, 2008] concluded that increasing model dimensionality to three dimensions (3D) can improve forecasts versus a 2D model in some cases, because the near-bottom velocity is a more appropriate determinant of bed stress than a depth-integrated velocity. However, 2D models typically utilize adjusted bed friction parameters (Manning's n) that attempt to account for the difference between a depth-averaged and a near-bed velocity [e.g., *Bunya et al.*, 2010], a factor that was not assessed in that paper. An important distinction, then, is that

[1]Center for Maritime Systems, Stevens Institute of Technology, Hoboken, New Jersey, USA.

Corresponding author: P. Orton, Center for Maritime Systems, Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ 07030, USA. (philip.orton@stevens.edu)

0148-0227/12/2012JC008220

the 2D models cannot resolve the influences of density variations on storm tide propagation through stratification and its impact on turbulence production.

[5] Another example of a physical process that is often neglected in storm surge modeling, yet has rarely been evaluated, is storm-driven rainfall and resulting flooding in tidal rivers and estuaries. The worst storm surges often come from storms that also bring heavy precipitation, and many population centers are located on estuaries where freshwater flow and storm tides merge. The standard assumption with FEMA flood zone mapping [e.g., *Divoky et al.*, 2005; *Federal Emergency Management Agency* (*FEMA*), 2012; *Niedoroda et al.*, 2010], with flood hazard assessment studies [e.g., *Lin et al.*, 2012], and with operational warning systems [e.g., *Colle et al.*, 2008; *Glahn et al.*, 2009], is that that storm-driven freshwater flow (rain and rivers) can be neglected when modeling a storm surge. This assumption relies on the argument that the storm comes and goes rapidly, and any rainfall-driven river flood arrives later [e.g., *Divoky et al.*, 2005; *Glahn et al.*, 2009]. However, extratropical storm surges in the NYC region in many cases build more slowly and last for multiple days [*Colle et al.*, 2010], so it is useful to quantitatively evaluate this assumption. Moreover, with storm tides of only 2.0–2.5 m LMSL threatening vital infrastructure around NYC, a realistic 0.3 m local rainfall during a tropical cyclone (e.g., Hurricane Floyd [*Colle*, 2003]) may also have a non-negligible influence on storm tide elevations.

[6] Given the vulnerability of the NYC region, it is imperative to quantify the importance of all physical processes and modeling choices in coastal flood prediction for this study area. Here, this goal is addressed for the region's estuaries, tidal straits, and tidal rivers, with storm tide modeling experiments using a highly detailed, well-validated, operational coastal ocean modeling system. Our research strategy is to perform several model experiments, removing various model components to quantify their influence on water levels in hindcasts of two of the region's worst recorded storm tides. These components are (a) freshwater inputs from rivers, sewers, and direct rain-on-water (b) water density variations, (c) nesting the model inside a larger-scale ocean model, (d) atmospheric pressure load forcing, and (e) wave steepness or age. Below, sections include (2) methods, providing a description of the ocean model, background on the two storms studied, experimental design and model evaluation procedures, (3) results, in terms of both model evaluation and the findings of the experiments, (4) discussion, an analysis of each of the experiments and the role of processes listed above, and (5) a summary with primary conclusions.

2. Methods

2.1. Numerical Model

[7] The Stevens Institute Estuarine and Coastal Ocean hydrodynamic Model (sECOM) [e.g., *Blumberg and Georgas*, 2008] is a variant of the Princeton Ocean Model (POM) [*Blumberg and Mellor*, 1987] and its shallow-water derivative model ECOMSED [*Blumberg et al.*, 1999]. It is a free-surface, hydrostatic, primitive equation model, with terrain-following (sigma) vertical coordinates, set on an orthogonal, curvilinear Arakawa C-grid. A major benefit of using sECOM for this study is that it has been demonstrated to provide highly accurate storm tide, velocity, temperature and salinity forecasts throughout the NY Bight and its waterways [*Di Liberto et al.*, 2011; *Georgas and Blumberg*, 2010] as part of the NY Harbor Observing and Prediction System, NYHOPS [*Bruno and Blumberg*, 2004]. The operational forecast model runs incorporate forecasts of river flow from the Northeast River Forecast Center, so NYHOPS is a complete hydrological and oceanic flood forecasting system for the tidal rivers within its domain (e.g., the Hudson River). Flood forecasts and a system of flood warnings via email or text message are available online, through the associated Storm Surge Warning System (http://stevens.edu/SSWS).

[8] For this study, the sECOM model was utilized with typical grid and model settings used for NYHOPS, except for a few exceptions that are outlined at the end of this section. An orthogonal-curvilinear grid covers the continental shelf and inland tidal regions from Maryland to Cape Cod (Figure 1). Horizontal resolution varies from around 4 km at the open ocean boundary to around 50 m in many parts of the waterways surrounding NYC. Vertical resolution is 10 terrain-following (sigma) levels. Vertical turbulent mixing is parameterized with the Mellor-Yamada 2.5 closure. The bed roughness is 0.001 m, and the minimum drag coefficient is 0.003 as determined for this region by *Blumberg and Pritchard* [1997]. The impact of wave orbital velocities on bed stress is also included via the parameterization of *Grant and Madsen* [1979]. Surface gravity waves are modeled in sECOM using a National Oceanic and Atmospheric Administration (NOAA) Great Lakes Environmental Research Laboratory (GLERL) wave model, a parametric model of wind-wave growth, propagation and frictional decay [*Donelan*, 1977], as described elsewhere [*Georgas et al.*, 2007; *Georgas*, 2010].

[9] The model is forced at the offshore boundary with clamped elevation boundary conditions (BC) that superimpose (1) sub-tidal sea levels using forecasts at each BC grid point from a large-scale operational 2D model of the NW Atlantic run by NOAA every 6 h with 1-h resolution output (ET-Surge, http://www.nws.noaa.gov/mdl/etsurge/) [*Chen et al.*, 1993], (2) a climatological cross-shelf slope ranging from 0.11 to 0.13 m (maximal at the coast) that is important for capturing average shelf currents [*Blumberg et al.*, 1999], and (3) tides based on seven major tidal constituents and two over-tides [*Georgas and Blumberg*, 2010]. A spatially constant bias correction is added to #1, based on low-pass filtered ET-Surge forecast error at coastal stations [*Georgas*, 2010].

[10] Meteorological forcing comes from 12-km, 3-h resolution NAM WRF forecasts for wind, atmospheric pressure, and other variables, run every 6 h [*Janjic et al.*, 2001]. Hydrological forcing includes freshwater and heat inputs from 93 major tributaries that are observed with gauges and 426 estimated water and heat sources including power plants, wastewater treatment plants, and minor tributaries [*Georgas*, 2010; *Georgas and Blumberg*, 2010]. This freshwater is injected to the model grid as a volume flux boundary condition.

[11] Changes from the typical NYHOPS setup used through 2011 include (1) switching on the atmospheric pressure gradient terms in the model's momentum equations, (2) use of a wave steepness dependent formulation for sea surface wind drag [*Taylor and Yelland*, 2001], with air

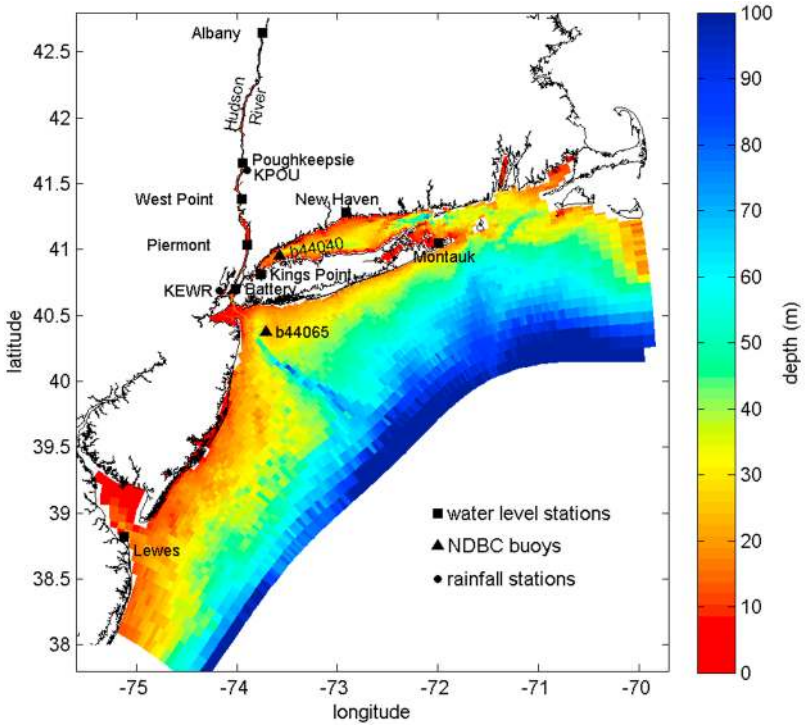

Figure 1. The NYHOPS domain used for sECOM model runs, showing bathymetry in meters. Several sites that are discussed in the paper are shown; including nine water elevation gauges, buoy 44065 in the apex of NY Bight, buoy 44040 in Long Island Sound, and two rainfall stations (KEWR and KPOU).

density computed using the ideal gas law with local air temperature, pressure and humidity, (3) inclusion of direct precipitation on water in the model domain based on observations at several airport sites, and (4) doubled bed drag coefficient values in the Hudson River north of Poughkeepsie, more appropriate for the gravel beds and sandy beds with sand waves found in that area [*Nitsche et al.*, 2007]. Regarding #2 above, NYHOPS normally uses the wind speed dependent drag formulation of *Large and Pond* [1981] with constant air density of 1.2 kg m^{-3}. The influence of changes 1–3 on storm tide predictions are evaluated later in this paper.

2.2. Tropical Cyclone Irene, August 2011

[12] Tropical cyclone (TC) Irene made landfall three times in the United States, first on the Outer Banks of North Carolina on 27 August, then at 0935 UTC 28 August near Atlantic City, New Jersey, with a minimum surface pressure of 959 hPa and maximum sustained winds of 31 m s^{-1} estimated by the National Hurricane Center (NHC). The third landfall of Irene was on 1300 UTC 28 August in the Coney Island neighborhood of NYC as a tropical storm with a pressure of 965 hPa and estimated maximum sustained winds of 28 m s^{-1} [*Avila and Cangialosi*, 2011], traveling 41 km h^{-1} (11.4 m s^{-1}). Irene arrived at NYC two days before a spring tide, though fortuitously on that day's lower high tide that had a predicted range of 1.6 m at the Battery (versus 1.8 m for the later higher-high tide that day and annual minima and maxima of 1.1 and 2.3 m). The storm caused the fourth highest gauged storm tide (total water level, since 1920) at the Battery in New York Harbor, and tenth highest (since 1931) at Kings Point in far western Long Island Sound (see also section 2.4), which is near the former tide station called Willet's Point. After its final landfall, the storm continued northward into interior New England, causing significant inland flooding. All storm tide rankings in this paper were provided by Nelson Vaz from the NOAA National Weather Service Upton office and are based on NOAA tide gauge data.

[13] At the height of the storm at NOAA-NDBC buoy 44065 in the apex of NY Bight (Figure 1), between 0400 and 1300 UTC 28 August, observed surface winds were 17–22 m s^{-1} from generally east and then southeast directions, with atmospheric pressure falling to a minimum of 968 hPa at 1250 UTC (Figures 2–4). Winds then rotated around to blow at 13–20 m s^{-1} from the southwest and eventually the west-northwest for the remainder of that day. Waves at buoy 44065 peaked at 8.0 m with an average period of 15 s, while waves in western Long Island Sound at buoy 44040 peaked at 2.2 m with an average period of 7 s (Figure 5). Total rainfall measured in NYC, to the west in New Jersey, and to the north up the Hudson Valley was from 0.10 to 0.25 m (e.g., KEWR and KPOU in Figure 4e), with smaller amounts (0.05–0.18 m) to the east on Long Island. August 2011 was a wet month even before Irene, including another heavy, record-setting precipitation event that preceded Irene by two weeks, and Irene's passage resulted in severe and often record-setting river flooding. The freshwater flow rate entering the tidal portion of the Hudson at Troy peaked on 29 August at 5100 m^3 s^{-1} (data from the U.S. Geological Survey, http://water.usgs.gov/), and was a 70-year flow event. On Rondout and Esopus Creeks, tributaries that enter the Hudson near Kingston, peak flows of 1034 and 714 m^3 s^{-1}

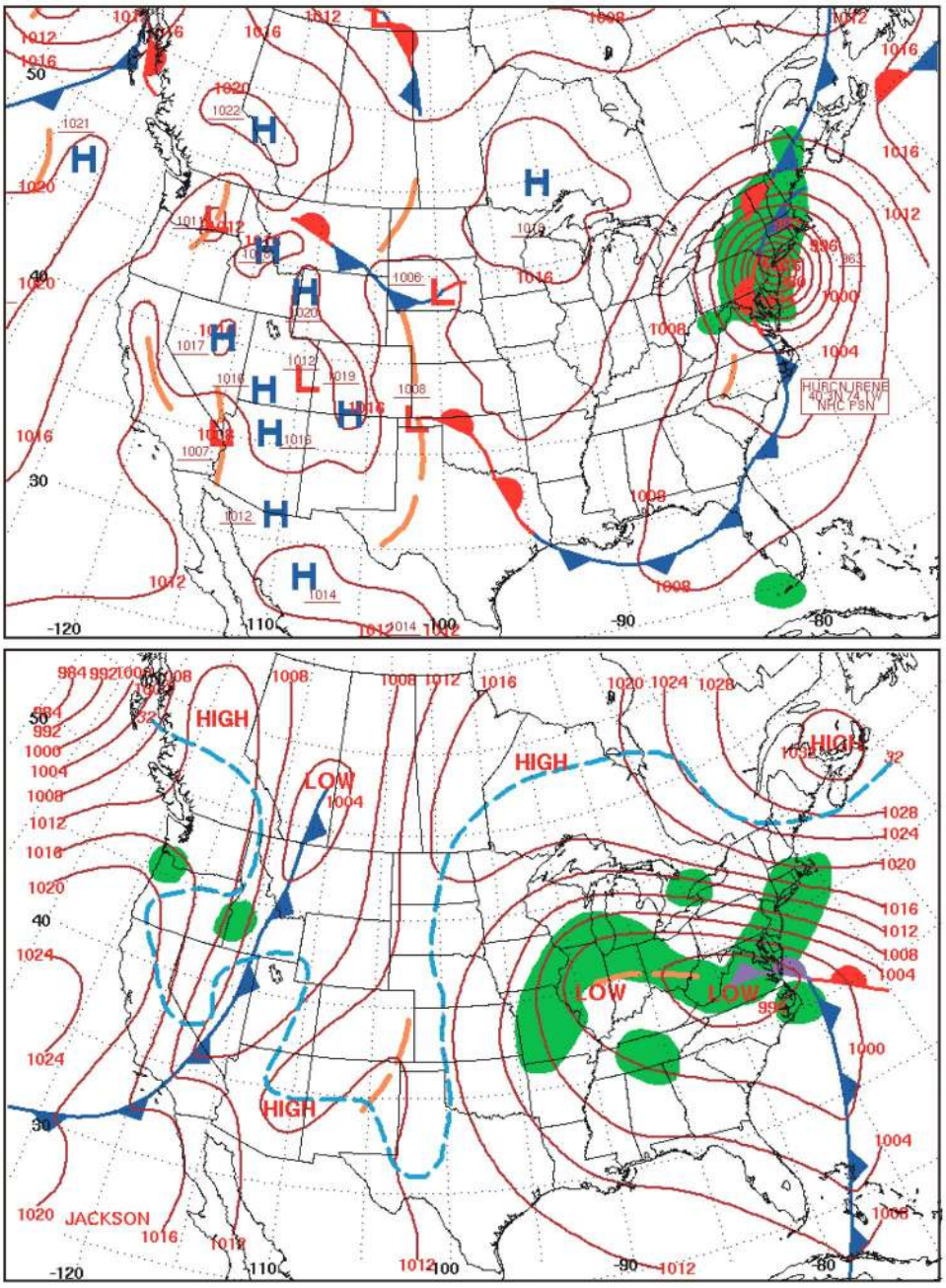

Figure 2. Observation-based analyses of surface atmospheric conditions for (top) 1100 UTC, 28 August 2011, the morning that TC Irene made landfall on Long Island, and (bottom) 1200 UTC, 13 March 2010, the morning of the day that the nor'easter storm tide peaked. Minimum pressure shown for Irene is 976 hPa, and for the nor'easter is 990 hPa. Credit: NOAA Central Library U.S. Daily Weather Maps Project.

occurred late on 28 August and preliminary estimates are that they were 80-year flow events [*Suro*, 2011].

2.3. Nor'easter, March 2010

[14] A severe extra-tropical nor'easter storm affected the U.S. Northeast from 12 to 15 March 2010, with sustained east-northeast winds from 15 to 20 m s^{-1} at buoy 44065 and across NY Bight (Figures 2, 6, and 7) causing the seventh highest gauged storm tide at the Battery and eleventh highest at Kings Point. The nor'easter came during a period with a relatively small 1.3 m tide range at the Battery. For the NYC region, Hudson Valley, and surrounding waterways, up to 0.10 m of total rainfall fell on top of saturated soil and a large amount of remaining snow from antecedent storm events, leading to moderate river flooding across the region. Elevated parts of the Hudson River's watershed had snow on the ground on 11 March, with 0.05 to 0.15 m of snow-water-equivalent (data from the National Operational Hydrologic Remote Sensing Center). Rainfall in coastal New Jersey and

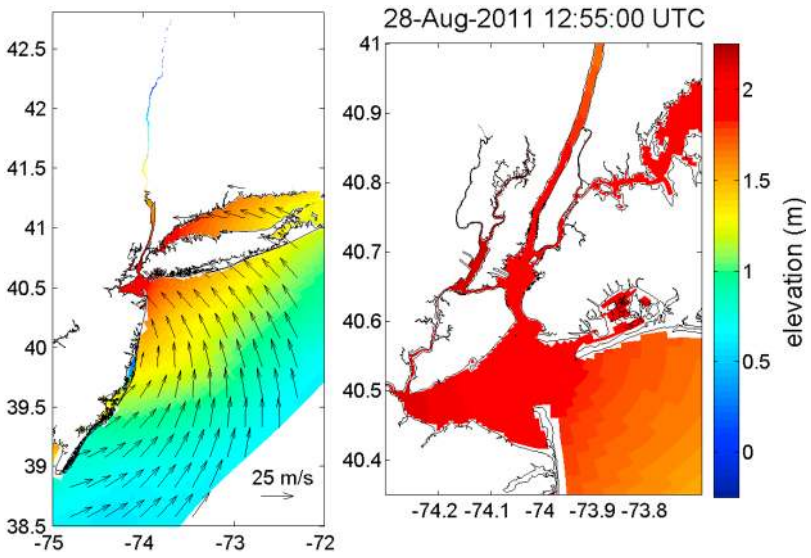

Figure 3. TC Irene maps of modeled 10 m wind velocity and total water elevation (shaded; in meters relative to mean sea level at the Battery) at the time of maximum storm tide at the Battery. A zoom-in to NYC is shown on the right.

Long Island was even more intense, with storm rainfall totals from 0.08 to 0.16 m.

[15] This cool-season storm fit the classic definition of a nor'easter – a cyclone forming within 167 km of the U.S. East Coast and tracking north or northeastward up the coast [*Huschke*, 1959]. A low pressure system off the South Carolina coast on 12 March caused light to moderate rain across the entire region through the day. The system drifted north and slowly strengthened, and rainfall rates became heavy overnight 12 March and through the evening of 13 March. Measured rain at Poughkeepsie (KPOU; Figure 7e) and Albany along the Hudson north of NYC was below 0.02 m in this storm, but according to radar-based estimates from the Northeast Regional Climate Center, 0.05–0.10 m fell in the mountains running along the west side of the Hudson and feeding into rivers like the Mohawk. The combination of rainfall and snowmelt led to a gradual increase in freshwater input to the Hudson from its two main tributaries at Troy, the Mohawk and upper Hudson, through 12 and 13 March (Figure 7f). Unlike the case with Irene, direct rainfall onto the Hudson was minimal in this storm and is not shown in the last panel of this figure. A high-pressure system centered over the Canadian Maritimes also strengthened and a very strong pressure gradient developed between the two pressure systems overnight 12 March, leading to very strong winds through the day 13 March over a more than 1000 km wind fetch at sea. Winds at buoy 44065 in the apex of NY Bight peaked from 1700 to 2400 UTC 13 March at 15–20 m s^{-1} (Figures 6 and 7) with gusts to 25 m s^{-1}. Wave heights peaked at 7.0 m with a 10 s dominant period at 0500 UTC on 14 March (Figure 8). The wind and rains weakened the night of 13–14 March, though the surface low remained over the Delmarva Peninsula (east of Chesapeake Bay). Weakened easterly winds and some rainfall continued 14–15 March and the system finally moved out to sea by mid-day on 16 March.

2.4. Total Water Elevation: Storm Tide

[16] In this paper, total water level or elevation, also known as the storm tide, is the focus. Many studies focus on storm surge, and either subtract modeled tides or use harmonic analysis to remove tidal fluctuations. However, the modeled and observed tide in the inland waterways is strongly non-stationary, with significant phase and amplitude changes during the storm, so it is very difficult to remove the tidal signal. One cause of this non-stationarity is the high freshwater flow during both these events, which is especially noticeable at Albany – there the tide was completely wiped out during TC Irene. Tidal ranges are commonly reduced or even eliminated entirely in tidal rivers when there are large freshwater inputs [e.g., *Georgas*, 2012; *Kukulka and Jay*, 2003]. Moreover, the tidal range across this region varies from 1–3 m, so it represents a large contributor to variations of total water elevation, which is the quantity of primary interest in flood studies.

[17] Peak water elevations relative to the Battery mean sea level (MSL) during TC Irene at the Battery, Kings Point, West Point and Albany were 2.11 m (Figure 3), 2.57 m, 2.10 m, and 4.53 m, respectively (subtract 0.06 m to get levels relative to NAVD88). This was the highest water level at the Battery since the December 1992 nor'easter, which peaked 0.06 m higher. As a result, some low-lying infrastructure was briefly closed due to flooding, including sections of major waterfront highways, some electrical sub-stations, and portions of the city's steam power system. The Hoboken PATH train station entrances would have taken in water as they did in 1992, but better advance warning led to preparations that minimized flooding to the underground tunnels.

[18] Peak water elevations relative to MSL during the March 2010 nor'easter at the Battery, Kings Point, West Point and Albany were 1.91 m (Figure 6), 2.58 m, 1.55 m, and 2.25 m, respectively. The peak water elevations at Kings Point were higher than the other three locations and identical

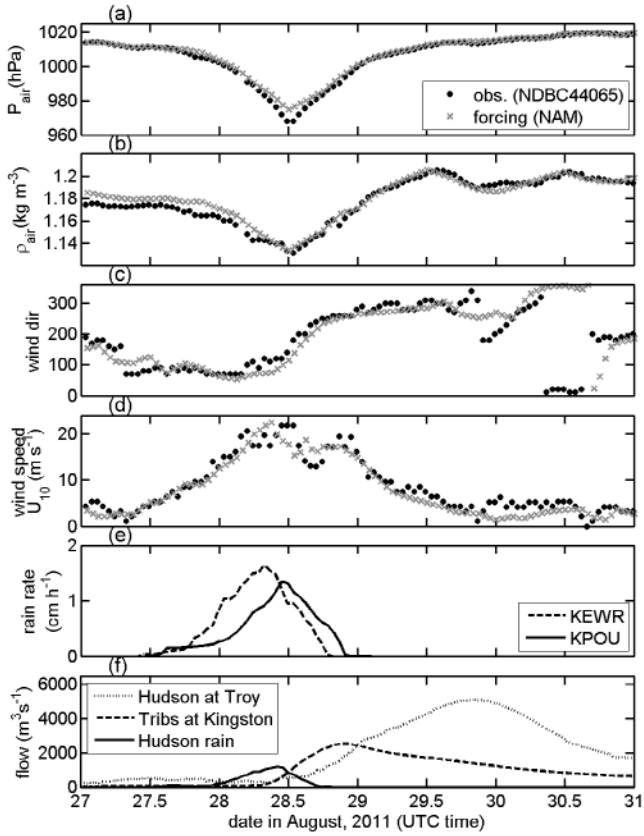

Figure 4. TC Irene meteorological (NDBC buoy 44065) and hydrological forcing time series, showing (a) observed atmospheric pressure compared with the NAM wind-forcing that was applied to the ocean model, (b) likewise for air density, (c) likewise for 10 m wind direction, (d) likewise for 10 m wind speed, (e) rain rates at two meteorological stations identified in Figure 1, and (f) preliminary measured freshwater inputs to the Hudson River at Troy from tributaries near Kingston (25 km north of Poughkeepsie), and from direct rain on water along the Hudson below Troy. Troy is a town just north of Albany, and the northernmost point of tidal propagation in the Hudson River due to a dam.

to those for Irene, due to the 0.46 m higher astronomical tide there and a ~0.2 m larger storm surge. The latter was likely due to the steady strong winds blowing along Long Island Sound.

2.5. Experimental Design

[19] The control model run for each storm included wind stress, freshwater inputs, atmospheric pressure, offshore sea level, and wave boundary conditions as prescribed in section 2.1. Cold start runs were utilized, starting five days before the onset of increasing winds. Initial temperature and salinity fields were specified based on fields from the operational modeling system, for each storm run's start date. The contributions of several physical processes and modeling components to water elevation predictions were tested using experiments where the relevant model components were individually removed, as shown in Table 1. The result when each modeling component was removed was then compared to the control run to measure the contribution of that component to water elevation. The processes that were removed include freshwater inputs from rivers, sewers and direct rain-on-water (NODIS), atmospheric pressure forcing over the NYHOPS domain (NOPRE), remote wind and pressure forcing beyond the NYHOPS continental shelf domain (NOETS), the influence of wave steepness and air density variations on wind stress (NOTAY), and water heterogeneity (HOMOG), which eliminates stratification and baroclinic forcing.

2.6. Quantifying Model Performance

[20] Various measures of model performance are utilized by different researchers, so multiple metrics were tabulated in evaluating the agreement between observations and the control model runs for this study. Each storm's control run was evaluated for a five-day period surrounding the peak water elevation, from 27 to 31 August 2011 for TC Irene,

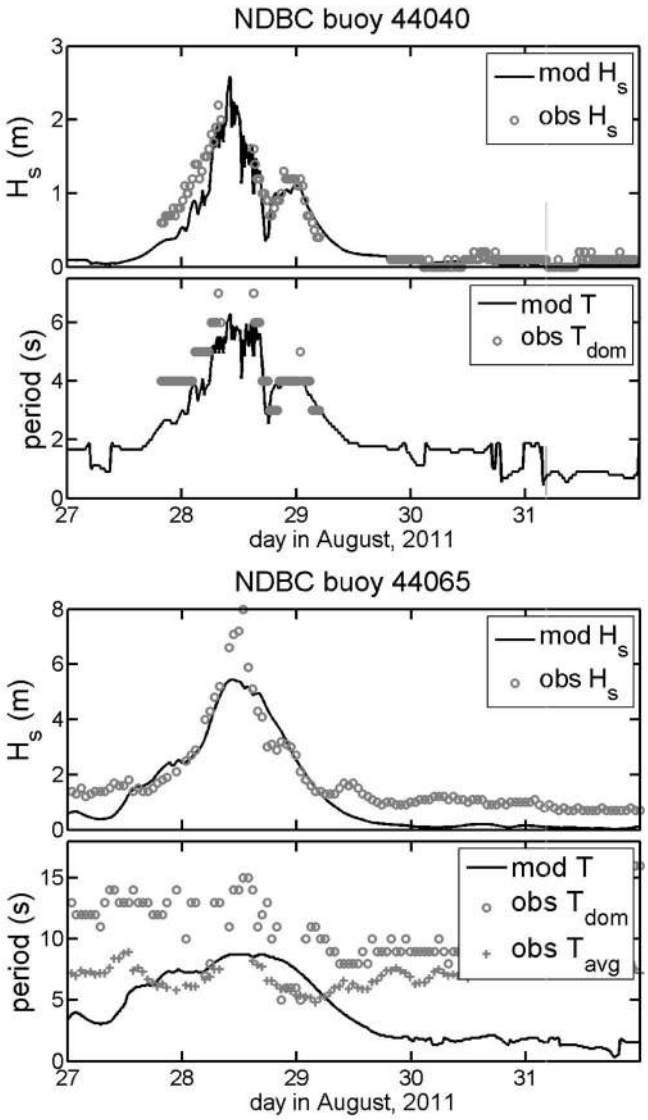

Figure 5. TC Irene observed versus modeled (control run) significant wave heights (H_s), average wave periods (T_{avg}; where available), and dominant wave periods (T_{dom}) at the two buoy stations.

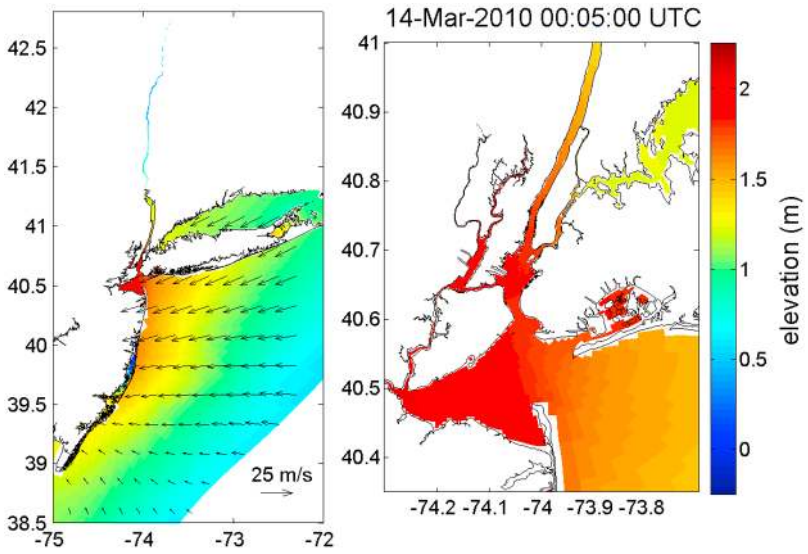

Figure 6. March 2010 nor'easter maps of modeled 10 m wind velocity and total water elevation (shaded; as Figure 3) at the time of maximum storm tide at the Battery. A zoom-in to the NYC area is shown on the right.

and from 12 to 16 March 2010 for the nor'easter. The root-mean squared-error (RMSE) was computed as

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}(\eta_{model} - \eta_{obs})^2\right]^{1/2} \quad (1)$$

[21] Here, N is the number of data points, and η_{model} and η_{obs} are the modeled and observed water elevations, respectively. The mean error (ME) is defined as

$$ME = \frac{1}{N}\sum_{i=1}^{N}(\eta_{model} - \eta_{obs}) \quad (2)$$

[22] And lastly, the Skill is defined as

$$Skill = 1 - \frac{\sum_{i=1}^{N}|\eta_{model} - \eta_{obs}|^2}{\sum_{i=1}^{N}\left(|\eta_{model} - \overline{\eta_{obs}}| + |\eta_{obs} - \overline{\eta_{obs}}|^2\right)} \quad (3)$$

[23] Perfect agreement gives a Skill of 1.0, while a complete disagreement gives a Skill of 0.0 [*Warner et al.*, 2005; *Willmott*, 1981]. Last, "Peak" is the peak elevation error, defined as the maximum modeled water elevation minus the maximum observed elevation.

3. Results

3.1. Control Run Evaluations

[24] sECOM control run water elevation predictions are evaluated for eight stations in Figure 9 (Irene) and Figure 10 (nor'easter), as well as in Table 2. The RMSE is low, from 0.07–0.24 m, relative to the storm tide elevations of 1.5–4.5 m. The Skill is high, from 0.97 to 1.00. Peak water elevation is captured within 0.20 m at all locations for the nor'easter, and all but two for Irene – at West Point and Piermont.

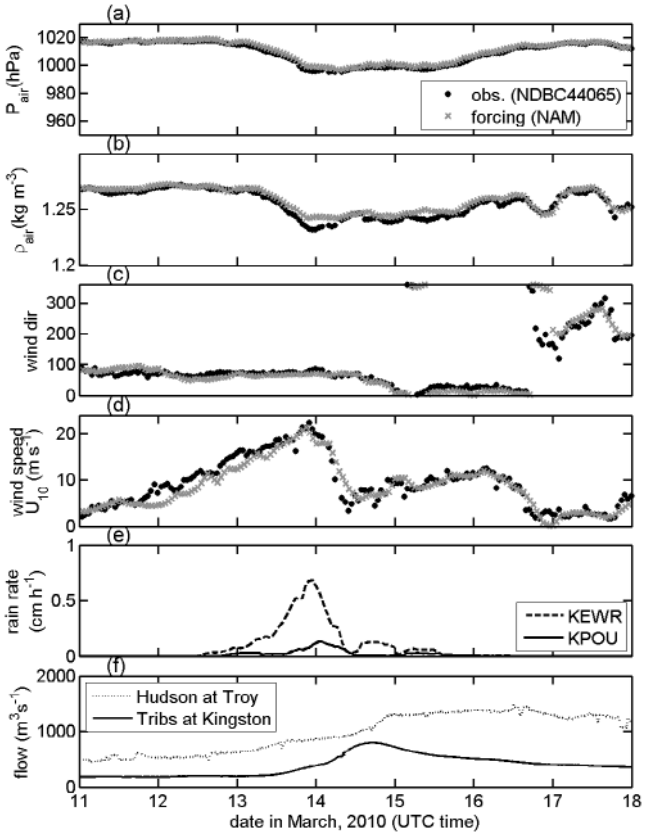

Figure 7. Nor'easter meteorological and hydrological forcing time series, showing (a) observed atmospheric pressure compared with the NAM wind-forcing that was applied to the ocean model, (b) likewise for air density, (c) likewise for 10 m wind direction, (d) likewise for 10 m wind speed, (e) rain rates at two meteorological stations identified in Figure 1, and (f) measured freshwater inputs to the Hudson River at Troy and from tributaries near Kingston.

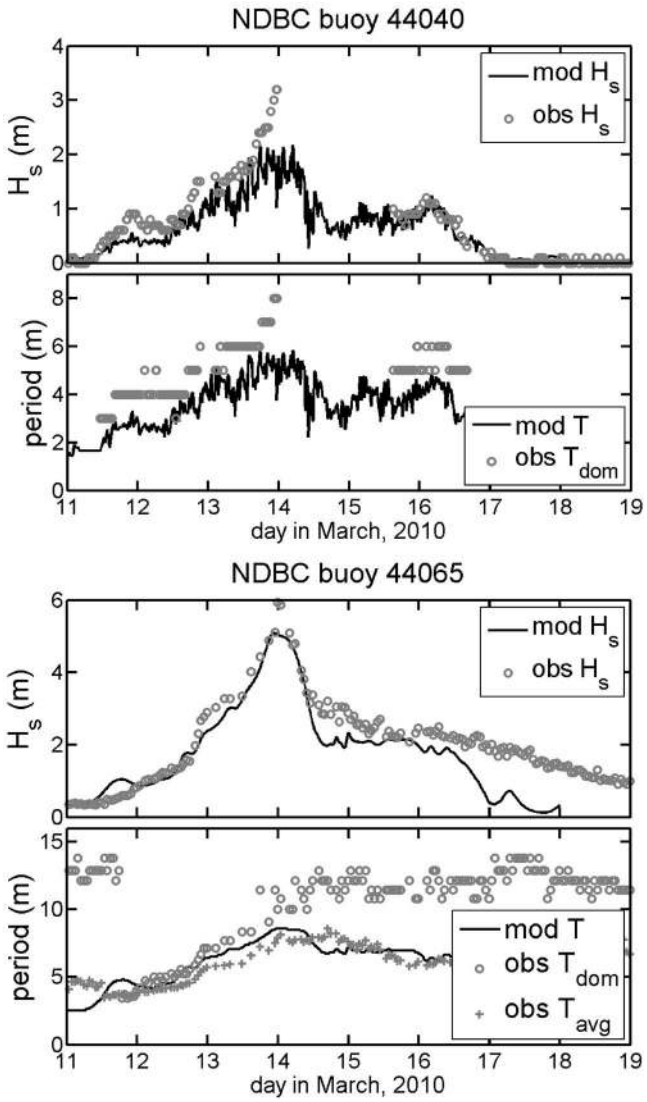

Figure 8. Nor'easter observed versus modeled (control run) significant wave heights (H_s), average wave periods (T_{avg}; where available), and dominant wave periods (T_{dom}) at the two buoy stations.

[25] The errors in modeled water elevation for TC Irene control runs appear to result mainly from errors in the NAM winds that are used as forcing, and this may illustrate the limits of a 12 km resolution mesoscale model for simulating a tropical cyclone event. For example, modeled elevation at The Battery, Piermont, and West Point started tracking too low, relative to observations, around 0300 UTC 28 August near low tide (Figure 9). From August day 27.95–28.20, wind speed was underestimated with NAM by an average of 13%, and wind direction averaged 80° (ENE) versus observed direction of 100° (ESE). Prior surge modeling has shown that the storm surge for NYC is controlled most strongly by wind speed from 130°, which blows directly into New York Harbor [*Lin et al.*, 2010]. Accordingly, cross-correlation analyses between TC Irene modeled surge error and NAM wind speed error on that 130° axis (southeast-to-northwest) suggest that the errors in storm surge were controlled to a large extent by the wind errors. The maximum correlation coefficients (r^2) at Battery and West Point of 0.79 and 0.82 come at lag times of 50 and 230 min, respectively. These lag times match very well with observed tide propagation times or expected shallow water wave propagation speeds over distances of 40 and 130 km. Along with these wind errors, NAM barometric pressure for TC Irene did not drop as low as observed (Figure 4a). A future goal is to set up model runs for our hindcasts and real-time forecasts that utilize forcing from a high-resolution hurricane model.

[26] sECOM wave predictions are compared with observations in Figure 5 (Irene) and Figure 8 (nor'easter), focusing on buoy 44040 at 18 m depth in western Long Island Sound and a buoy 44065 at 50 m depth on the open continental shelf (both shown in Figure 1). For the stormy periods that matter most for the surge predictions, the model and observations show moderate agreement on wave height and period at these two stations (typically within 30% - note that the observed "average" wave period is most appropriate to compare with the model). The relevance of the average versus the dominant wave period for various aspects of storm surge physics is discussed in section 4.5.

3.2. Overview of Experimental Results

[27] Examining time series of the experiment results for three stations during Irene (Figure 11), one sees moderate decreases of over 0.20 m (or ~10%) at some or all sites when neglecting each process. For the NOETS run omitting remote forcing of water elevation (ET-Surge model nesting), decreases in the peak elevation were from 0.21 to 0.25 m. Removing the dependence of wind stress on wave steepness and air density (NOTAY) and removing the atmospheric pressure gradient forcing (NOPRE) also have strong effects at these three stations, with decreases in elevation of 0.11–0.17 and 0.17–0.25 m, respectively. Omitting freshwater inputs (NODIS) reduces peak elevation by 0.29 m at West Point, though only by 0.03–0.04 m at the Battery and Kings Point. Omitting water density variations (HOMOG) reduces peak elevation by 0.13, 0.25, and 0.20 m at the Battery, West Point, and Kings Point, respectively.

[28] A similar pattern is seen in the experimental results for the nor'easter, with moderate decreases at some or all sites when neglecting each process (Figure 12). Remote forcing accounts for 0.13–0.20 m of peak elevation, a similar

Table 1. Model Experiments

Name	Description	Purpose: To Test What?
INCLU	control run	for comparisons to cases below
NOETS	no ET-Surge sub-tidal elevation BC	the importance of remote atmospheric effects
NODIS	no freshwater discharges/inputs	the importance of rivers, rainfall
HOMOG	homogeneous run	the role of stratification and baroclinicity
NOTAY	no Taylor-Yelland - uses Large-Pond	the role of wave steepness for wind stress
NOPRE	no atmospheric pressure forcing	the role of atmospheric pressure gradients

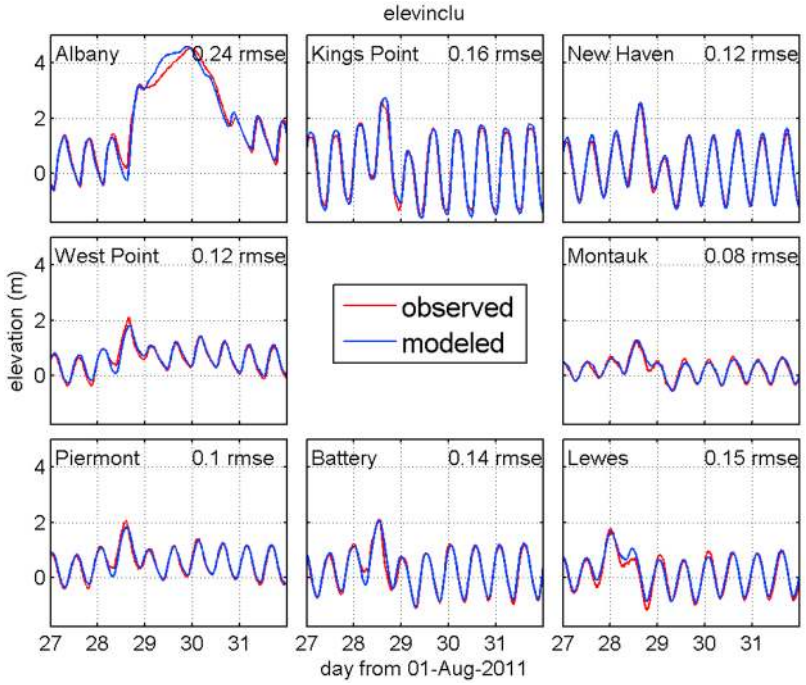

Figure 9. Model evaluation for the Irene control run–water elevation at all sites. All elevations are relative to mean sea level at the Battery.

percentage of the total water elevations as with Irene. The influence of the wave steepness parameterization (NOTAY) is slightly larger than it was with Irene, accounting for 0.17–0.22 m of peak elevation, while the influence of atmospheric pressure (NOPRE) is smaller, accounting for 0.07–0.11 m of peak elevation. Omitting freshwater inputs (NODIS) had a smaller effect than Irene, reducing peak elevation by ~0.15 m at West Point, and again only by centimeters at the

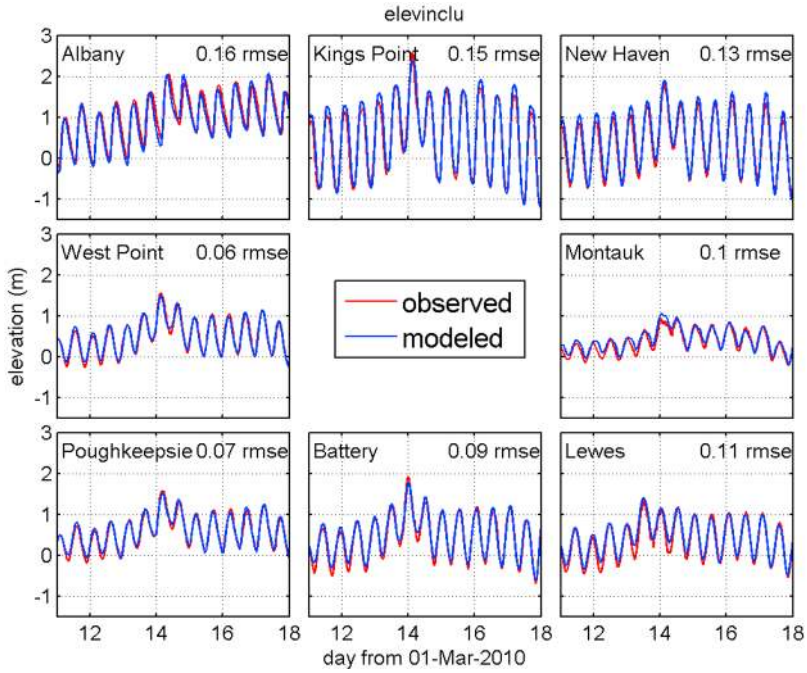

Figure 10. Model evaluation for the nor'easter control run–water elevation at all sites. All elevations are relative to mean sea level at the Battery. The Piermont panel has been replaced by Poughkeepsie (bottom left), relative to Figure 9, because Poughkeepsie data were only available for the nor'easter (and likewise, Piermont data for Irene).

Table 2. Statistical Assessments of the Model Control Run, Versus Observations, as Defined in Section 2.6

Station	RMSE	ME	Peak	Skill
		TC Irene		
Albany	0.24	0.01	0.04	0.99
Kings Point	0.16	0.04	0.16	1.00
New Haven	0.12	0.02	0.06	1.00
West Point	0.12	0.04	−0.31	0.98
Montauk	0.08	0.01	−0.01	0.99
Piermont	0.1	0.02	−0.24	0.99
Battery	0.14	0.03	−0.06	0.99
Lewes	0.15	0.07	−0.08	0.99
		Nor'easter		
Albany	0.16	−0.07	0.03	0.98
Kings Point	0.15	0.08	−0.19	0.99
New Haven	0.13	0.09	0.11	0.99
West Point	0.06	0.01	−0.07	0.99
Montauk	0.1	0.07	0.1	0.97
Poughkeepsie	0.07	0.02	−0.06	0.99
Battery	0.09	0.04	−0.12	0.99
Lewes	0.11	0.07	0.05	0.98

Battery and Kings Point. Omitting water density variations (HOMOG) reduces peak elevation by 0.05, 0.14 and 0.12 m at the Battery, West Point and Kings Point, respectively.

[29] Histograms summarize results for all stations, with percent changes in peak water elevation in Figures 13 and 14. Following similarly from the time series figures described above, neglecting remote atmospheric forcing (NOETS) leads to decreases of up to 16% during Irene and up to 17% during the nor'easter, with the highest values at Montauk due to its location halfway toward the eastern edge of the model domain. The influence of freshwater inputs is large at the upriver Hudson stations, with NODIS decreases of 15 and 82% at West Point and Albany (respectively) for Irene and 9 and 36% at these sites for the nor'easter. It is also noteworthy that it is only 1–2% of total elevation at the Battery and Kings Point over both storms. The HOMOG runs omitting water density variations stand out as having everywhere >5% decrease in elevation during Irene (excepting river-stage dominated Albany), yet a smaller change for the nor'easter. The decreases in peak elevation are notably 9–13% at West Point for the two storms, 5–7% at Kings Point and even 2–5% at the Battery, all within the NY Metropolitan area. The NOTAY experiment led to a 4–8% decrease for Irene, and 8–11% for the nor'easter. The NOPRE experiment led to a decrease in peak elevation of 9–11% for Irene, 3–6% for the nor'easter.

[30] Spatial map views and an examination of joint omission of two processes for TC Irene are shown in Figure 15. The left panel shows that the freshwater influence (NODIS) is near zero at the ocean and increases with distance upriver into enclosed inland waterways like the Hudson, with maximal values at the model's inland boundary at the head-of-tide (Albany/Troy). The center panel shows that the water density (HOMOG) influence for that storm is 4–5% near the ocean and increases for inland tidal waterways, increasing to maximum values of 13% at the head of the salt intrusion (West Point). Omission of the various processes is a nearly linear process, as joint omission of both density variations

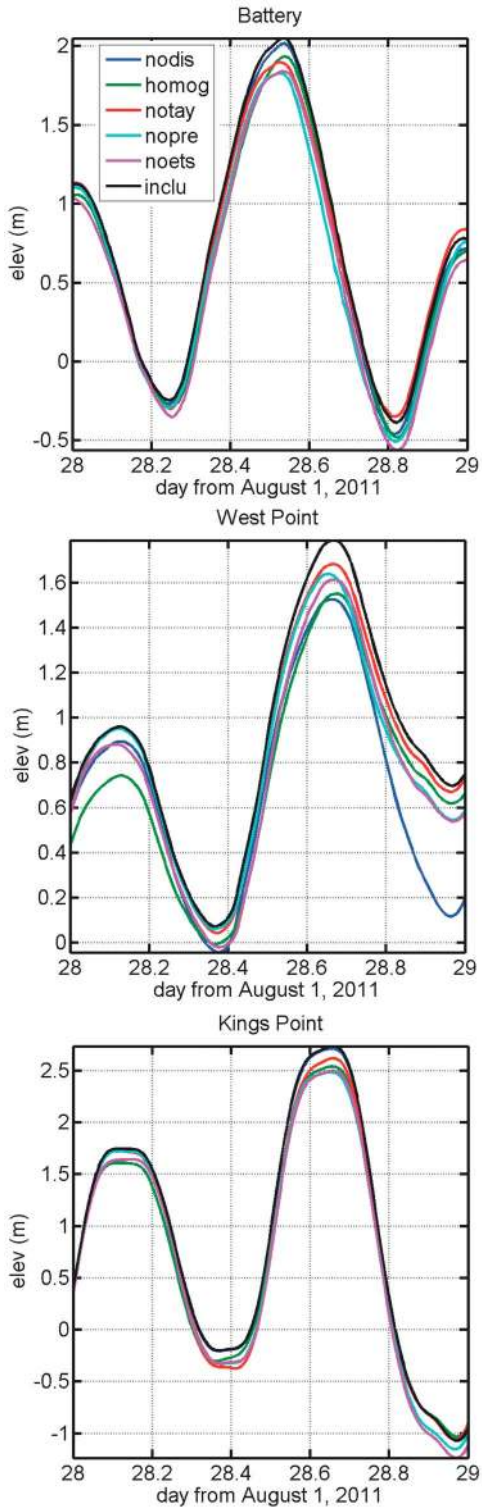

Figure 11. Irene experiment results plots in time series form for the Battery, West Point, and Kings Point, showing the peak water elevation for the control run with all processes modeled (INCLU), as well as experimental runs with various processes omitted, as identified in Table 1.

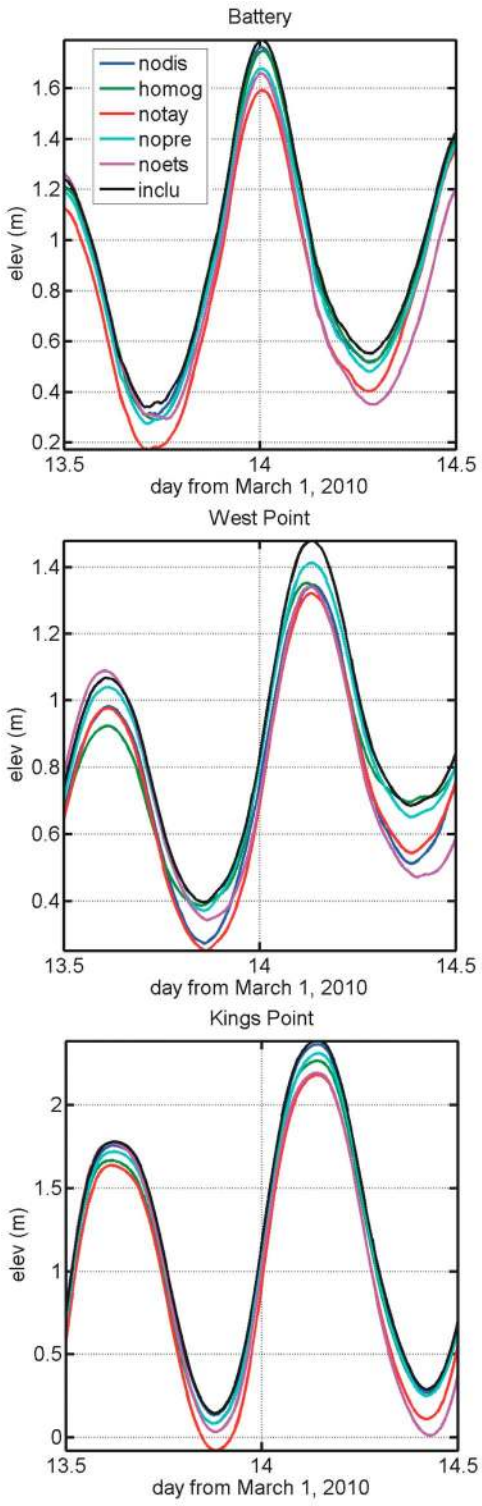

Figure 12. Nor'easter experiment results plots in time series form, as Figure 11.

and freshwater results in a percentage decrease (Figure 15, right) that is just below the addition of the experimental results shown in the left and center panels. This result is also borne out when all the experimental processes are jointly applied, with a HOMOG, NODIS, NOPRE, NOTAY, NOETS model run giving a reduction of 35% at Battery, relative to a sum of all the experimental results of 37%. Similar linear behavior is seen for West Point and Albany.

4. Discussion

[31] Below, separate sections address each set of experimental results and put them into context of prior studies. Surprisingly, for a tropical cyclone and extra-tropical nor'easter, the two storms had some similarities. Both storms were very large and had similar influences from remote atmospheric forcing beyond the NYHOPS domain (section 4.1 below). The two storms had different hydrometerological and hydrological characteristics, and this, along with the timing of freshwater relative to surge, is discussed in detail below in section 4.2. The atmospheric pressure minimum and pressure gradient for Irene were more intense than the nor'easter, and had a stronger influence on sea level (section 4.4). Irene came in late August, when the coastal surface ocean was warm and stratified, whereas the nor'easter came in March when waters were more uniformly cool (section 4.3).

4.1. Influence of Domain Size on Water Elevations

[32] At the time of peak storm tide at New York Harbor (Figures 3 and 6), the wind patterns fit expectations for these two different types of storms – winds during a nor'easter typically last longer but blow softer, and Earth's rotation deflects the net water flux (and thus, the surge) toward the right of the wind, a process called Ekman transport. On the other hand, the strongest winds around tropical cyclones drive a net water flux more directly downwind at the time of landfall. Irene's winds in NY Bight just prior to the time of peak water elevation at the Battery were from the southeast at 22 m s^{-1} (Figure 4). However, prior to that time, the winds were much more like a nor'easter, blowing at 15–20 m s^{-1} for 10 h from the east and pushing up sea level in New York Bight and Long Island Sound. The nor'easter's winds blew onshore at that rate for about 24 h. Irene was unusually large for a tropical cyclone, nearly 1000 km in diameter at the time of its final landfall, giving it some similarities to a nor'easter – both storms had a wind fetch that went eastward of the model domain (upwind).

[33] The results of the NOETS experiment suggest that the similarities in wind fetch led to similar importance of remote forcing for both storms. This experiment gauged the importance of remote forcing of sea level beyond the NYHOPS grid area shown in Figure 1. Typical reductions in peak elevation were from 7 to 16% during Irene (exempting Albany, dominated by river flow) and from 9 to 17% during the nor'easter. The reduction in peak elevation was larger for Montauk (16 and 17%), which is located much closer than the other stations to the eastern model domain boundary and thus had a shorter wind fetch for these two storms. Likewise, it was smaller for Lewes (7 and 2%), the southernmost site with the largest fetch. A few prior studies have shown the importance of remote forcing for hurricanes [*Blain et al.*, 1994] and nor'easters [*Shen and Gong*, 2009], and our results further underscore the importance of having either a very large grid area or model nesting.

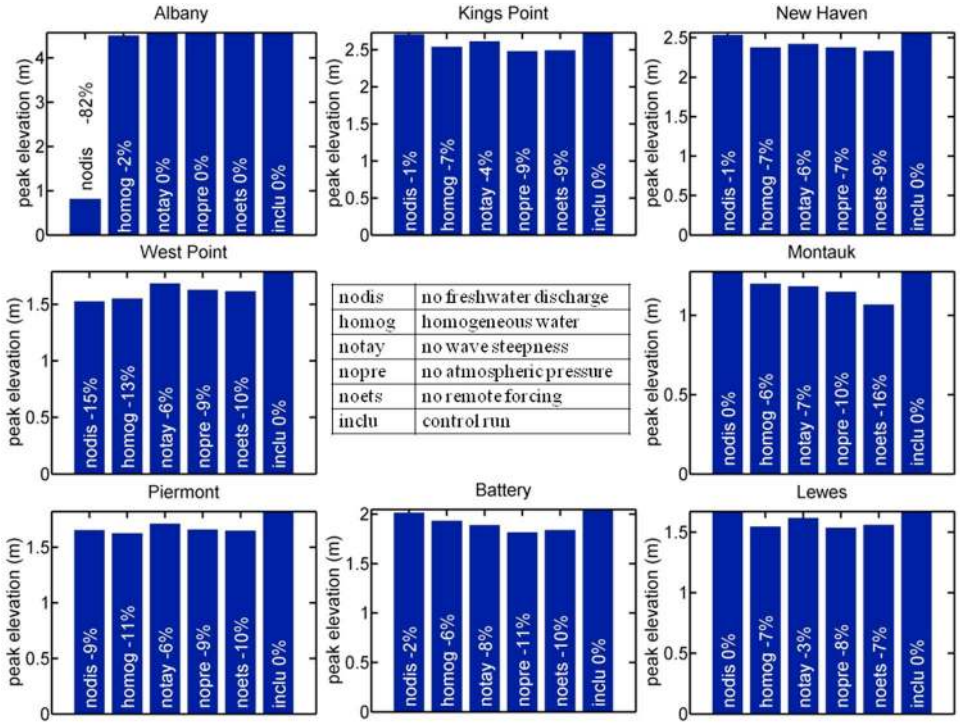

Figure 13. Irene experimental results summary, comparing INCLU peak elevation (right of each plot) with the peak elevation for all experimental runs. Percent change is also shown as text for each experiment.

4.2. Influence of Freshwater Inputs

[34] As mentioned in section 1, little work has been done to quantify the importance of rainfall-driven flooding relative to coastal storm surge for inland tidal waterways. The NODIS experiment shows that freshwater flooding at sites like West Point on the Hudson during both Irene and the nor'easter contributed substantially to the total water

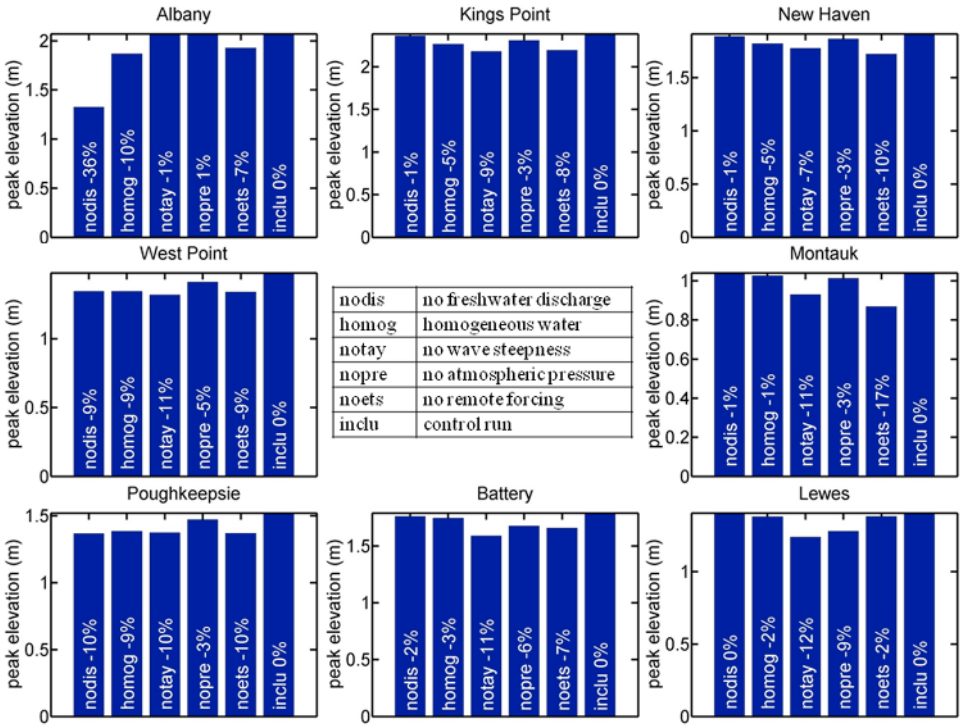

Figure 14. Nor'easter experimental results summary, comparing INCLU peak elevation (right of each plot) with the peak elevation for all experimental runs. Percent change is also shown as text for each experiment.

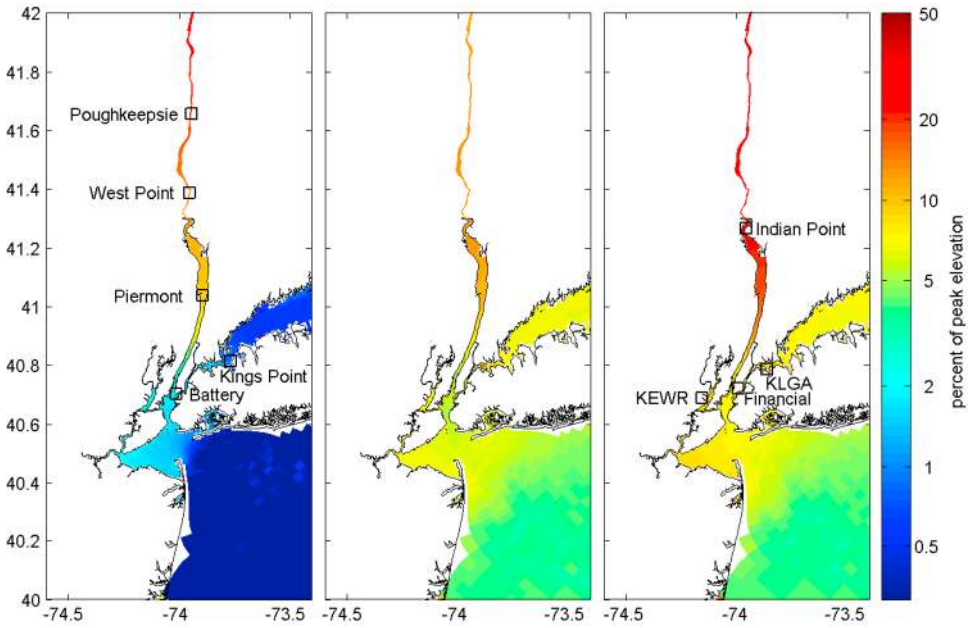

Figure 15. Percent reduction in peak water elevation when (left) freshwater inputs are omitted, (middle) density variations are omitted, and (right) both freshwater and density variations are omitted from model runs for TC Irene. Tide stations are added to the left panel, and a few examples of infrastructure in the right panel.

elevation (Figures 13 and 14), and therefore merged with coastal storm surge from other forcing, negating the common assumption of independence. The model results for Irene show that the impact of freshwater inputs on peak water elevation was only 1–2% around NYC, but increased rapidly upriver to 11% near the Indian Point nuclear facility and 15% at West Point (Figure 15, left). The tide gauge at Poughkeepsie failed during the storm, so this paper doesn't highlight that city in spite of extensive street flooding, but the model results suggest a freshwater contribution of 18% there along the banks of the Hudson.

[35] Looking closer at the timing of freshwater and storm tide during Irene, the Hudson's tributary inputs that enter the Hudson far to the south of Albany were already large at the time of peak elevation at West Point, 1605 UTC 28 August. The estimated sum of USGS measured flows at Rondout Creek, Esopus Creek and Wallkill River entering the Hudson at Kingston was 1850 $m^3 s^{-1}$ at that time (Figure 4f), and the total freshwater inflow rate estimated by NYHOPS for tributary inputs south of Albany and north of West Point totaled 3600 $m^3 s^{-1}$ (neglecting rainfall on water). Transit time of the freshwater flood is important, and if it propagates at roughly the shallow water wave speed (square root of the gravitational acceleration times depth, $\sim$9 m s^{-1}), then there is a $\sim$2 h transit time for the 60 km distance from Kingston to West Point. Two hours earlier, at 1405 UTC, the flow rate was 1350 $m^3 s^{-1}$.

[36] Direct rainfall-on-water during Irene produced a total inflow rate (distributed along the Hudson) that peaked at 1200 $m^3 s^{-1}$ (Figure 4f), and this is paired with river inflows for the NODIS experiment. Though not addressed in the results figures, this process was also assessed with an experiment – rain falling directly onto the Hudson River during Irene raised peak water elevations at West Point and the Battery by 4 cm and 1 cm, or 2% and 0.5% of total elevation. The small influence at the Battery is likely because rainfall rates are small, O(1 cm h^{-1}), and because rain doesn't occur at this rate everywhere at once, there is time for the added water volume to spread out to sea.

[37] It is possible that Irene's rainfall-driven river flooding began earlier than would be the case with a typical hurricane, as it was an unusually large tropical cyclone with a diameter close to 1000 km at the time of landfall at Long Island. The outer rainbands preceded the storm center by 15.5 h (e.g., Figure 4). A large-diameter, slow-moving storm allows more time for river flooding to rise before the peak storm surge arrives, and at 41 km h^{-1}, Irene's speed was somewhat slow compared with hurricanes that have struck this region. Based on NOAA Atlantic Ocean hurricane track maps summarized in *Landsea et al.* [2004], storm speeds for hurricanes that historically struck Long Island are Hurricane Donna at $\sim$51 km h^{-1}, the New England hurricane of 1893 at $\sim$45 km h^{-1}, and the hurricane of 1938 at a swift $\sim$110 km h^{-1}.

[38] The relative timing of freshwater and storm tide for the nor'easter was markedly different than with Irene, with high flows preceding the storm due to seasonal snowmelt and with lower peak flows. By the time of peak water elevation at West Point, 0255 UTC 14 March 14 2010, fresh water flows had ramped up at Kingston, nearly doubling from 220 to 400 $m^3 s^{-1}$. Flows into the Hudson at both Kingston and Troy were already elevated before the storm due to rain in the day in the mountains prior to the peak surge, as well as heavy snow events in February and subsequent warm springtime weather. These hydrological conditions are likely to be common before spring nor'easters, and some of the worst storm surges in the U.S. Northeast have occurred in late winter or early spring – most notably, the Ash Wednesday storm of 6–7 March 1962 (sixth worst measured elevation at

the Battery), and the nor'easter of 29 March 1984 (ninth worst), so these results have broad applicability.

[39] These results for both storms demonstrate that freshwater flow can be important to storm-driven water elevations along the Hudson. The freshwater input time series demonstrate that tributaries south of Albany and direct rainfall on water play a role with their moderate but early freshwater flood contributions. Both storms had different characteristics in terms of flood timing and magnitude relative to the peak wind-driven surge. Yet, some of the region's worst storms have much less rainfall (e.g., the 1992 nor'easter) and thus freshwater can at times have a negligible effect. Therefore, no simple approach (e.g., using seasonal average hydrological inputs) is apparent for accurately accounting for these variable effects. Although adding rainfall and hydrology requires a great deal of model set-up time and can add one or more new dimensions to a risk-assessment analysis, it is not computationally expensive for a single model run.

4.3. Influence of Water Heterogeneity

[40] As mentioned in section 1, many storm surge studies omit the influence of water density variations on storm tide propagation, either by using constant temperature and salinity or by using 2D models. Our experiments demonstrate that neglecting water density variations in this region can lead to moderate underestimation of peak water elevation. The reductions in peak water elevation for runs with homogeneous water (HOMOG) were moderate for inland waterways, notably 9–13% at West Point for the two storms, 5–7% at Kings Point and even 3–6% at the Battery. At the sites just inside embayments near the coast during Irene, HOMOG showed a 6–7% decrease in peak elevation (Lewes, Montauk), whereas during the nor'easter there were smaller decreases of 1–2% at these stations. Decreases along the open coasts of Long Island and New Jersey are 4–5% for Irene (Figure 15, middle), and ∼1% during the nor'easter, indicating that even coastal ocean water density variations were somewhat important during Irene and can have different importance for different storms.

[41] These results are not surprising in light of the fact that tides often show summertime amplitude increases in shallow shelf seas, an effect attributed to increased stratification and an associated decrease in turbulence [*Müller*, 2012]. Looking closer at stratification during Irene, modeled stratification in the Hudson at the time of peak elevation at the Battery (1300 UTC 18 August) was 4 kg m^{-3} over a total depth of 20 m at the Verrazano Narrows (latitude 40.6) and 6 kg m^{-3} over a total depth of 12 m in the Hudson up to the Tappan Zee Bridge (latitude 41.15), yet the surge had been building for 13 h by this time. The storm's strong winds that built up the surge in the 12 h prior to landfall blew across, not along, the Hudson, so had very little fetch to cause turbulent mixing and waves.

[42] Rutgers University glider observations of stratification offshore from Atlantic City (latitude 39.37°) and across the continental shelf show that the water column did not become unstratified there as Irene approached, except at inshore locations inside the 20 m contour (S. Glenn, unpublished data, 2012). Our modeled estimates of shelf stratification on the same transect agree qualitatively with these observations and similarly show stratification only being eliminated inside the 25 m contour.

[43] These results strongly suggest that stratification was not eliminated prior to the time of peak flood elevation in the Hudson estuary, and they also show a moderate influence of water heterogeneity on peak water elevations (6–13%) in the estuarine regions surrounding NYC. Therefore, for modeling storm surge propagation into the area's estuaries and tidal rivers, in particular, there are clear benefits to using a 3D model. This conclusion is also likely transferable to other estuaries and coastal regions with substantial freshwater influence and stratification. While modeling storm surges with 3D versus 2D takes additional time, about four times more with sECOM for example, other processes with a similar percentage influence on water elevations (e.g., set up of mean sea level by breaking waves [*Bunya et al.*, 2010]) can also take substantial computing time but are increasingly being included.

4.4. Influence of Atmospheric Pressure Gradients

[44] The effect of atmospheric pressure gradients on sea level, also referred to as atmospheric pressure load forcing, is well-known to be important for hurricanes. It was shown to be important for predicting water elevation in two recent studies of moderate nor'easter events in the Chesapeake Bay, accounting for up to 32% of sub-tidal sea level fluctuations [*Gong et al.*, 2009; *Salas-Monreal and Valle-Levinson*, 2008]. Incorporating its effect in a surge model is a simple choice because atmospheric forcing typically includes barometric pressure, models typically already have the ocean effect coded, and it requires no additional computational expense.

[45] Here, our NOPRE experiments showed that omitting atmospheric pressure gradient forcing within the NYHOPS domain caused a decrease in peak elevation of 7–11% for Irene, 3–9% for the nor'easter. The stronger effect during Irene was simply caused by a greater atmospheric pressure drop during the storm, which had a central minimum pressure of 965 hPa at landfall on Long Island and a much stronger spatial gradient in atmospheric pressure (closely spaced isobars in Figure 2). Also, the experimental result may underestimate the decrease in peak elevation caused by pressure for Irene, as the comparison of modeled and observed atmospheric pressure shown in Figure 4a shows the model overestimates the pressure by ∼20% near the time of the pressure minimum. These results corroborate the finding that local atmospheric pressure gradients are an important contributor to storm tides, even for nor'easters. As of spring 2012, after a long evaluation period, atmospheric pressure gradient terms have now been enabled in the sECOM model's momentum equations for the operational forecasting in NYHOPS, allowing local atmospheric pressure gradients to influence water elevations and currents.

4.5. Influence of Wave Steepness

[46] Increasingly, storm surge modeling groups are using sea surface drag formulations with wave steepness [e.g., *Warner et al.*, 2010] or wave age dependence [e.g., *Bertin et al.*, 2012]. For winds from 12 to 23 m s^{-1}, this is a physically justified approach that is based on the results of many field studies over the past two decades [*Taylor and Yelland*, 2001]. The standard and widely accepted

COARE 3.0 formulation [*Fairall et al.*, 2003] gives drag coefficients that are roughly 30% higher versus *Large and Pond* [1981] in this wind speed range, so the similarly enhanced estimates prescribed using wave steepness or wave age are consistent with the changing consensus on what constitutes an appropriate drag coefficient.

[47] Here, results of the NOTAY experiment measured the impact of replacing the Taylor-Yelland wind drag parameterization with the Large-Pond parameterization that is typically used with a constant air density. The specific goal was to examine the influence that explicit representation of wave steepness can have on storm surge predictions, and the air density change was of minor importance. Also note that the ET-Surge model within which our model is nested does not include these effects, so we are only measuring the local effect within our domain. The result for NOTAY was a 3–8% decrease in peak elevation for Irene and a 7–12% decrease for the nor'easter. The influence of the use of Taylor-Yelland in both storms was actually very similar, and the difference between the two sets of results was mainly driven by the difference in air density – utilizing a constant density instead of a variable one for Irene led to an increase of 1.5% in surge, as air density was lower than 1.2 kg m^{-3} (Figure 4). The opposite was true for the nor'easter, which had a higher density (Figure 7).

[48] The wave steepness wind stress (Taylor-Yelland) is better used with an average wave period, because it is more sensitive to the small-scale, young wavefield, than the long-period swell. The wave-current bed stress, on the other hand, depends more on the dominant wave period, which reflects the long-period swell waves that have influence deeper in the water column. The modeled wave periods at the peak of the storms in NY Bight are in good agreement with the average wave period, and therefore the Taylor-Yelland sea surface roughness should be accurate. On the other hand, the wave-current bed stress is likely underestimated, as a result of poor agreement between the model (8 s wave period at peak) and the observations (15 s dominant period).

[49] Parameterized wave-related enhancement of wind stress can be computationally inexpensive, and generally should be incorporated to improve model accuracy. A major goal in our future model development is to couple a spectral wave model to sECOM, seeking to improve wave period predictions and also capture more detailed air-sea interaction processes. We are also evaluating various options for wind drag parameterization, including wave age formulations and the COARE parameterization, the latter of which only relies on wind speed, not on modeled wave characteristics.

5. Summary and Conclusions

[50] A study of two recent severe storm surge events was conducted, focusing on the NYC region and surrounding estuarine waterways. The atmospheric characteristics of two recent severe storms were described, as well as the water level response of the coastal ocean and hydrologically affected inland waterways. These two storms, tropical cyclone Irene and the March 2010 nor'easter, caused the highest storm tides around NYC in twenty years.

[51] A series of model experiments was performed to examine the role of various physical processes in controlling total water elevation, with one process omitted for each experiment. Each process showed importance of at least 10% for at least some tide gauge stations, which is striking because some processes are routinely incorporated in surge modeling (e.g., atmospheric pressure, remote forcing) and some others are nearly always ignored (e.g., water density variations and storm-driven freshwater inputs). The influence of omitting remote forcing led to typical reductions in elevation from 7 to 17% for both Irene and the nor'easter. Experiments that neglected explicit accounting for the role of wave steepness for wind stress led to modest decreases of 4–11% for both storms. Neglecting the influence of atmospheric pressure gradient forcing on water elevation led to reductions in peak elevation of 9–11% for Irene, and 3–6% for the nor'easter, which had a much weaker pressure drop and spatial gradient. The influence of freshwater inputs was large at the upriver Hudson stations, with decreases of 15 and 82% at West Point and Albany (respectively) for Irene and 9 and 36% at these sites for the nor'easter. Experiments omitting water density variations showed decreases of up to 13%. Here, the largest decreases were found deep inside the Hudson River estuary (e.g., Piermont and West Point), likely due to estuary stratification causing a reduction in turbulence production. Also, decreases were larger for Irene than for the nor'easter, likely due to the increased thermal ocean stratification on the continental shelf in late summer.

[52] Several groups studying risk and forecasting surges for the NYC region run 2D barotropic, homogeneous models, with no freshwater inputs [e.g., *Colle et al.*, 2008; *Glahn et al.*, 2009; *Lin et al.*, 2012], and this includes the current FEMA coastal flooding study that is being used to update flood zone maps for the New Jersey coast, NYC and NY Harbor, and the Hudson up to Albany [*FEMA*, 2012]. However, our results demonstrate that joint omission of these processes led to a low-bias in storm tide for TC Irene at NYC sites like La Guardia Airport (9%; Figure 15, right) and Manhattan's financial district near the Battery (7%), as well as nearby vulnerable sites like the Indian Point nuclear plant (23%), or Newark Airport and the Port of Newark-Elizabeth (9%).

[53] It is particularly noteworthy that the errors induced by omitting freshwater inputs and water density variations are negative biases. Omitting these processes during Irene leads to biases of 0.15–0.48 m at the sites listed above. Considering the low elevation grade of land in many of the NYC region's flood zones and the presence of critical underground electrical infrastructure and transportation tunnels with access points only a few decimeters above the storm flood levels of Irene, this level of bias should be avoided. Also, Indian Point is vulnerable at a high 4.5 m MSL flood height, according to the Vice President of Operations [*Public Broadcasting System*, 2012]. Relative to the maximum known historic water level of ~3.2 m MSL for a hurricane that struck NYC at low tide in 1821 [*Scileppi and Donnelly*, 2007], and adding a 1.6 m mean tidal range and a ~0.5 m local sea level rise since then, one can see how a low-bias of several decimeters in modeled water elevation could be important for risk assessment. Due to the large population within 50 km, it is important to accurately estimate the return period for a flood with this 4.5 MSL height, and correctly adapt our coastal defenses to present or future dangers.

[54] In conclusion, storm surge modelers should move beyond asking whether these processes should be considered

at all, and instead ask what technique and complexity should be used for incorporating them. On one extreme, statistical approaches can be utilized to estimate the increases in water levels caused by each process, such as by developing a multiplicative "correction" based on a small number of higher-complexity model runs. On the other extreme, there are detailed and dynamically modeled approaches, such as what was utilized in this study. The specific needs of different modeling efforts factor strongly into these decisions. Risk assessment efforts have substantial uncertainty in other areas of their analysis, such as with development of a storm climatology for a region with few historical hurricanes such as New York City. In such cases, it may be more reasonable to use simplified approaches for making corrections for missing physical processes. Operational storm surge forecast systems, on the other hand, can benefit from more complete dynamical modeling or having the most detailed model inputs available. Whatever the approach, many cities worldwide are located inside estuaries and tidal rivers, and future modeling efforts will need to account for freshwater inputs and water density variations to ensure that we do not chronically underestimate flood forecasts and risks.

[55] **Acknowledgments.** The authors would like to thank Vivien Gornitz, Radley Horton, and two anonymous reviewers for their valuable comments that strengthened the paper, as well as Nelson Vaz at NOAA-NWS in Upton, NY, for his helpful interactions while we were performing this research. Research was funded by the National Ocean and Atmospheric Administration's RISA Program, award NA10OAR4310212, for the "Consortium for Climate Risk in the Urban Northeast (CCRUN)."

References

Avila, L. A., and J. Cangialosi (2011), Tropical cyclone report: Hurricane Irene, *Rep. AL092011*, 45 pp., Natl. Hurricane Cent., Miami, Fla.

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, doi:10.1029/94JC01348.

Blumberg, A. F., and N. Georgas (2008), Quantifying uncertainty in estuarine and coastal ocean circulation modeling, *J. Hydraul. Eng.*, *134*(4), 403–415, doi:10.1061/(ASCE)0733-9429(2008)134:4(403).

Blumberg, A. F., and G. L. Mellor (1987), A description of a three-dimensional coastal ocean circulation model, in *Three-Dimensional Coastal Ocean Models*, *Coastal Estuarine Sci.*, vol. 4, edited by N. S. Heaps, pp. 1–16, AGU, Washington, D. C., doi:10.1029/CO004p0001.

Blumberg, A. F., and D. W. Pritchard (1997), Estimates of the transport through the East River, New York, *J. Geophys. Res.*, *102*, 5685–5703, doi:10.1029/96JC03416.

Blumberg, A. F., L. A. Khan, and J. St. John (1999), Three-dimensional hydrodynamic model of New York Harbor region, *J. Hydraul. Eng.*, *125*(8), 799–816, doi:10.1061/(ASCE)0733-9429(1999)125:8(799).

Bruno, M. S., and A. F. Blumberg (2004), An urban ocean observatory for the maritime community, *Sea Technol.*, *45*, 27–32.

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, doi:10.1175/2009MWR2906.1.

Chen, J., W. Shaffer, and S. Kim (1993), A forecast model for extra-tropical storm surge, *Adv. Hydrosci. Eng.*, *1*, 1437–1444.

Colle, B. A. (2003), Numerical simulations of the extratropical transition of Floyd (1999): Structural evolution and responsible mechanisms for the heavy rainfall over the northeast United States, *Mon. Weather Rev.*, *131*(12), 2905–2926, doi:10.1175/1520-0493(2003)131<2905:NSOTET>2.0.CO;2.

Colle, B. A., F. Buonaiuto, M. J. Bowman, R. E. Wilson, R. Flood, R. Hunter, A. Mintz, and D. Hill (2008), New York City's vulnerability to coastal flooding, *Bull. Am. Meteorol. Soc.*, *89*(6), 829–841, doi:10.1175/2007BAMS2401.1.

Colle, B. A., K. Rojowsky, and F. Buonaito (2010), New York City storm surges: Climatology and an analysis of the wind and cyclone evolution, *J. Appl. Meteorol. Climatol.*, *49*(1), 85–100, doi:10.1175/2009JAMC2189.1.

Di Liberto, T., B. A. Colle, N. Georgas, A. F. Blumberg, and A. A. Taylor (2011), Verification of a multi-model storm surge ensemble around New York City and Long Island for the cool season, *Weather Forecast.*, *26*, 922–939, doi:10.1175/WAF-D-10-05055.1

Divoky, D., R. Battalo, R. Dean, I. Collins, D. Hatheway, and N. Scheffner (2005), Storm meteorology: FEMA coastal flood hazard analysis and mapping guidelines focused study report, 31 pp., Fed. Emerg. Manage. Agency, Washington, D. C.

Donelan, M. (1977), A simple numerical model for wave and wind stress prediction, 28 pp., Natl. Water Res. Inst., Burlington, Ont., Canada.

Fairall, C. W., E. F. Bradley, J. E. Hare, A. A. Grachev, and J. B. Edson (2003), Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm, *J. Clim.*, *16*(4), 571–591, doi:10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2.

Federal Emergency Management Agency (FEMA) (2012), Redefinition of the coastal flood hazard zones in FEMA region II: Analysis of the coastal storm surge flood frequencies, Fairfax, Va.

Georgas, N. (2010), Establishing confidence in marine forecast systems: The design of a high fidelity marine forecast model for the NY/NJ harbor estuary and its adjoining coastal waters, PhD thesis, 272 pp., Stevens Inst. of Technol., Hoboken, N. J.

Georgas, N. (2012), Large seasonal modulation of tides due to ice cover friction in a midlatitude estuary, *J. Phys. Oceanogr.*, *42*(3), 352–369, doi:10.1175/JPO-D-11-063.1.

Georgas, N., and A. F. Blumberg (2010), Establishing confidence in marine forecast systems: The design and skill assessment of the New York Harbor Observation and Prediction System, version 3 (NYHOPS v3), paper presented at Eleventh International Conference in Estuarine and Coastal Modeling, Am. Soc. of Civ. Eng., Seattle, Wash.

Georgas, N., A. Blumberg, and T. Herrington (2007), An operational coastal wave forecasting model for New Jersey and Long Island waters, *Shore Beach*, *75*(2), 30–35.

Glahn, B., A. Taylor, N. Kurkowski, and W. A. Shaffer (2009), The role of the SLOSH model in National Weather Service storm surge forecasting, *Natl. Weather Dig.*, *33*(1), 4–14.

Gong, W., J. Shen, K. H. Cho, and H. V. Wang (2009), A numerical model study of barotropic subtidal water exchange between estuary and subestuaries (tributaries) in the Chesapeake Bay during northeaster events, *Ocean Modell.*, *26*(3–4), 170–189, doi:10.1016/j.ocemod.2008.09.005.

Grant, W. D., and O. S. Madsen (1979), Combined wave and current interaction with a rough bottom, *J. Geophys. Res.*, *84*(C4), 1797–1808, doi:10.1029/JC084iC04p01797.

Huschke, R. (Ed.) (1959), *Glossary of Meteorology*, 638 pp., Am. Meteorol. Soc., Boston, Mass.

Janjic, Z. I., J. Gerrity Jr., and S. Nickovic (2001), An alternative approach to nonhydrostatic modeling, *Mon. Weather Rev.*, *129*(5), 1164–1178, doi:10.1175/1520-0493(2001)129<1164:AAATNM>2.0.CO;2.

Kukulka, T., and D. A. Jay (2003), Impacts of Columbia River discharge on salmonid habitat: 1. A nonstationary fluvial tide model, *J. Geophys. Res.*, *108*(C9), 3293, doi:10.1029/2002JC001382.

Landsea, C. W., C. Anderson, N. Charles, G. Clark, J. Dunion, J. Fernandez-Partagas, P. Hungerford, C. Neumann, and M. Zimmer (2004), The Atlantic hurricane database re-analysis project: Documentation for the 1851–1910 alterations and additions to the HURDAT database, in *Hurricanes and Typhoons: Past, Present and Future*, edited by R. J. Murnane and K. Liu, pp. 177–221, Columbia Univ. Press, New York.

Large, W. G., and S. Pond (1981), Open ocean momentum flux measurements in moderate to strong winds, *J. Phys. Oceanogr.*, *11*, 324–336, doi:10.1175/1520-0485(1981)011<0324:OOMFMI>2.0.CO;2.

Lin, N., K. Emanuel, J. Smith, and E. Vanmarcke (2010), Risk assessment of hurricane storm surge for New York City, *J. Geophys. Res.*, *115*, D18121, doi:10.1029/2009JD013630.

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

Müller, M. (2012), The influence of changing stratification conditions on barotropic tidal transport and its implications for seasonal and secular changes of tides, *Cont. Shelf Res.*, doi:10.1016/j.csr.2012.07.003, in press.

Niedoroda, A., D. Resio, G. Toro, D. Divoky, H. Das, and C. Reed (2010), Analysis of the coastal Mississippi storm surge hazard, *Ocean Eng.*, *37*(1), 82–90, doi:10.1016/j.oceaneng.2009.08.019.

Nitsche, F. O., W. B. Ryan, S. M. Carbotte, R. E. Bell, A. Slagle, C. Bertinado, R. Flood, T. Kenna, and C. McHugh (2007), Regional patterns and local variations of sediment distribution in the Hudson River estuary, *Estuarine Coastal Shelf Sci.*, *71*(1–2), 259–277, doi:10.1016/j.ecss.2006.07.021.

Public Broadcasting System (2012), Joe Pollock: Inside the Indian Point nuclear power plant, Frontline, 17 Jan.

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

Salas-Monreal, D., and A. Valle-Levinson (2008), Sea-level slopes and volume fluxes produced by atmospheric forcing in estuaries: Chesapeake Bay case study, *J. Coastal Res.*, *24*, 208–217, doi:10.2112/06-0632.1.

Scileppi, E., and J. P. Donnelly (2007), Sedimentary evidence of hurricane strikes in western Long Island, New York, *Geochem. Geophys. Geosyst.*, *8*, Q06011, doi:10.1029/2006GC001463.

Shen, J., and W. Gong (2009), Influence of model domain size, wind directions and Ekman transport on storm surge development inside the Chesapeake Bay: A case study of extratropical cyclone Ernesto, 2006, *J. Mar. Syst.*, *75*(1–2), 198–215, doi:10.1016/j.jmarsys.2008.09.001.

Suro, T. P. (2011), Historical flood peaks and peaks during the flood of Aug. 28–29, 2011, at selected U.S. Geological Survey streamgages in New York, provisional data, 4 pp., U.S. Geol. Surv. N. Y. Water Sci. Cent., New York.

Taylor, P. K., and M. J. Yelland (2001), The dependence of sea surface roughness on the height and steepness of the waves, *J. Phys. Oceanogr.*, *31*(2), 572–590, doi:10.1175/1520-0485(2001)031<0572:TDOSSR>2.0.CO;2.

Warner, J. C., W. R. Geyer, and J. A. Lerczak (2005), Numerical modeling of an estuary: A comprehensive skill assessment, *J. Geophys. Res.*, *110*, C05001, doi:10.1029/2004JC002691.

Warner, J. C., B. Armstrong, R. He, and J. B. Zambon (2010), Development of a coupled ocean-atmosphere-wave-sediment transport (COAWST) modeling system, *Ocean Modell.*, *35*(3), 230–244, doi:10.1016/j.ocemod.2010.07.010.

Weisberg, R., 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.

Willmott, C. J. (1981), On the validation of models, *Phys. Geogr.*, *2*(2), 184–194.