Coupled modeling of storm surge and coastal inundation: A case study in New York City during Hurricane Sandy

Jie Yin1,2, Ning Lin2, and Dapeng Yu3
1School of Geographic Sciences, East China Normal University, Shanghai, China, 2Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA, 3Centre for Hydrological and Ecosystem Science, Department of Geography, Loughborough University, Loughborough, UK

Abstract In this paper, we describe a new method of modeling coastal inundation arising from storm surge by coupling a widely used storm surge model (ADCIRC) and an urban flood inundation model (FloodMap). This is the first time the coupling of such models is implemented and tested using real events. The method offers a flexible and efficient procedure for applying detailed ADCIRC storm surge modeling results along the coastal boundary (with typical resolution of ~100 m) to FloodMap for fine resolution inundation modeling (<5 m). The coastal inundation during Hurricane Sandy was simulated at both the city (New York City) and subregional (lower Manhattan) scales with various resolutions. Results obtained from the ADCIRC and coupled ADCIRC-FloodMap simulations were compared with the recorded (high water marks) and derived (inundation extent based on the planar method) data from FEMA. At the city scale, coupled ADCIRC-FloodMap modeling demonstrates improved prediction over ADCIRC modeling alone for both the extent and depth of inundation. The advantage of the coupled model is further illustrated in the subregional modeling, using a mesh resolution of 3 m which is substantially finer than the inland mesh resolution used by ADCIRC (>70 m). In further testing, we explored the effects of mesh resolution and roughness specification. Results agree with previous studies that fine resolution is essential for capturing intricate flow paths and connectivity in urban topography. While the specification of roughness is more challenging for urban environments, it may be empirically optimized. The successful coupling of ADCIRC and FloodMap models for fine resolution coastal inundation modeling unlocks the potential for undertaking large numbers of probabilistically based synthetic surge events for street-level risk analysis.

1. Introduction

Coastal flooding is among the most frequent and devastating natural hazards, resulting in significant casualties, losses, and impacts in many coastal low-lying areas around the world. Coastal cities are particularly vulnerable to storm surge induced flooding given their highly developed economy and high-density population [Hanson et al., 2011; Hallegatte et al., 2013]. Rapid urbanization in coastal zones is set to continue in the foreseeable future [Hallegatte et al., 2013; Jongman et al., 2012], and hence to increase social and economic exposure. Anthropogenic changes such as land subsidence, reduced permeability, and wetland loss further exacerbate the hydrologic regime of flood responses in coastal cities [e.g., Wu et al., 2012; Yin et al., 2015]. Moreover, it is widely recognized that climatic change has significant negative impacts on many regions through raising sea level and possibly enhancing storminess, which in turn amplify coastal flood events [IPCC, 2013, 2014]. For example, due to the combined effects of sea level rise and storm climatology change, the 10 year flood height in New York Harbor has increased by nearly 0.72 ± 0.25 m since the mid-1800s [Talke et al., 2014], and the present 100 year flood is projected to occur every 3–20 years by the end of the 21st century [Lin et al., 2012].

Recent storm flood catastrophes in the United States (2005, 2012), China (2006, 2013), and the United Kingdom (2013) have attracted growing attention to the increasing threat and its enhanced consequences in flood-prone coastal cities [Jongman et al., 2012; Aerts et al., 2014]. New York City (NYC) is one of the most exposed and vulnerable coastal megacities [Aerts et al., 2014]. On 29 October 2012, Hurricane Sandy made landfall near Atlantic City with a “worst-case” storm track and unusually large size, resulting in a catastrophic storm surge in NYC. The storm surge peak coincided with an abnormal astronomical high tide (up to 15 cm higher than a normal high tide) along the New Jersey and New York coastlines, resulting in storm tides...
reaching record levels over NYC and extensive flooding in lower Manhattan, the Brooklyn/Queens waterfront, Jamaica Bay, and Staten Island [Blake et al., 2013]. Hurricane Sandy claimed about 45 lives in NYC, mostly due to drowning, and caused more than $20 billion in flood-associated damages (approximately 40% of the total U.S. loss in this event) to the city’s building and infrastructure, making it the worst disaster in NYC history [NYC Office of Emergency Management, 2014]. Afterward, a $20 billion NYC Storm Flood Protection Plan was proposed by the City Council to implement a series of resilience measures, including construction of flood defense system along the coast. The quality of storm-tide inundation estimates has a significant effect on damage/loss estimation [Schubert et al., 2015] and thus coastal cites’ flood management. Accurate and efficient forecast/nowcast/hindcast of storm flood inundation (e.g., extent, depth, and velocity) is needed to inform decision making in the process of flood risk management.

Considerable progress has been made during past decades in the development and application of methodologies for simulating both storm tides and overland inundation by means of numerical modeling and/or GIS-based analysis. A number of approaches to modeling coastal flooding exist. The simplest, static approach assumes that the land area is horizontally submerged if its elevation is below the maximum water surface elevation and if it is hydraulically connected to flooded areas or open water [Brown, 2006; Poulter and Halpin, 2008]. Due to the algorithmic simplicity and computational efficiency of this so-called planar method, it has been widely applied in a large number of coastal environments to quickly delineate flood extent at fine resolutions [NPCC, 2010; Lichter and Felsenstein, 2012; Aerts et al., 2013, 2014; Lloyd et al., 2015]. However, a key shortcoming of this approach is its physical oversimplification of flood routing over land, which tends to be biased toward overestimation in inundation depths and areas; establishing hydraulic connectivity is not trivial due to the dynamic nature of flooding, especially for topographically flat regions [Bates et al., 2005; Gallien et al., 2011; Ramirez et al., 2016]. Second, two-dimensional (2-D) or three-dimensional (3-D) hydrodynamic models (e.g., POM, SLOSH, ADCIRC, SELFE, and MIKE 21) were directly applied to replicate coastal storm surge flooding from the open sea to coastal regions by incorporating land areas into the computational domain [Wang et al., 2012; Forbes et al., 2014]. For example, the second New York Panel of Climate Change (NPCC2) has undertaken hydrodynamic modeling of NYC coastal flooding with ADCIRC utilizing an unstructured mesh with a 70 m resolution for land surface [Orton et al., 2015]. Moreover, fine resolution (or subgrid) waterfront inundation models (3–5 m) nested within storm surge models have been developed to investigate the street-level inundation process [Blumberg et al., 2015]. Although these nested models have demonstrated higher accuracy in inundation prediction at local (community or street) scales, the massive computational cost and complexity involved in the data preprocessing of urban topography make them less practical for widespread applications and hamper their use in simulating probabilistically based synthetic events ($10^3$ to $10^6$).

More recently, reduced-complexity hydraulic models with simplified 2-D solutions have been increasingly used for the treatment of storm surge flooding because they can capture the dominant physics of coastal flood process at different scales by directly using the available GIS-based topographic datasets such as LiDAR Digital Elevation Model (DEM) and are consequently transferable to any coastal location [Bates et al., 2005; Ramirez et al., 2016]. The simplified approaches based on raster grids have been proved to perform as well as full 2-D models for predicting coastal flood, but at a substantially reduced computational cost [Bates et al., 2005]. These models should be forced by the boundary conditions along the coast as they do not simulate the storm-tide process in the ocean. A number of studies have used recorded or simulated tidal curves at a limited number of gauge stations to drive simplified 2-D models (e.g., LISFLOOD-FP), without accounting for detailed spatial variability in storm-tide levels [Bates et al., 2005; Smith et al., 2012; Skinner et al., 2015; Ramirez et al., 2016], which storm surge models typically provide. To date, fully coupling storm surge models and simplified flood routing techniques, which overcome the aforementioned limitations, have yet to be comprehensively applied or tested for coastal cities’ flood modeling. Brown et al. [2007] described the coupling of a coarse-resolution (8 km by 8 km) storm surge model (Dutch Continental Shelf Model) with a 2-D hydraulic model (Delft-FLS) for coastal flood modeling, with a focus on evaluating modeling uncertainties using hypothetical scenarios. However, the use of hypothetical events limits the scope of evaluating model performance and cross-model comparison.

In this paper, we apply a combined 2-D numerical simulation approach, whereby a storm-tide model (ADCIRC) and a simplified 2-D hydrodynamic model (FloodMap) [Yu and Lane, 2006a, 2006b, 2011] run in sequence, to derive the flood inundation in NYC during Hurricane Sandy. The main objectives of this study
are three: (1) to develop an accurate and efficient approach for the dynamic modeling of storm induced flood inundation at both the city level and street level; (2) to investigate the performance and sensitivity of the coupled ADCIRC-FloodMap model; and (3) to compare estimated inundation (area and depth) from different sources (i.e., FEMA static, ADCIRC, and ADCIRC-FloodMap coupled modeling). The rest of this article is organized as follows: section 2 describes the modeling methodology, including the storm-tide modeling, the inundation modeling, model coupling, observational data, and evaluation methods; section 3 presents the results and discussions; and section 4 summarizes the findings.

2. Methodology

2.1. Storm-Tide Modeling

The regional scale storm-tide modeling was performed using the two-dimensional, depth-integrated implementation of the ADCIRC model. ADCIRC is a finite element model developed by [Luetich et al., 1992] and [Westerink et al., 1992] for simulating hydrodynamic circulations along shelves and coasts and within estuaries. ADCIRC has been used extensively to study storm surge in various coastal regions [e.g., Westerink et al., 2008; Sheng et al., 2010; Lin et al., 2010a; Dietrich et al., 2011; Zachry et al., 2013], including the New York area [Colle et al., 2008; Lin et al., 2010b; Orton et al., 2015]. The model fully describes the complex physical processes associated with storm surge and allows the use of an unstructured grid with comparably high resolution near the coast and inland (e.g., <100 m) and much coarser resolution in the deep ocean (e.g., >50 km). In this study, we used a comprehensive numerical grid (Region II) generated by FEMA [FEMA, 2014a]. The grid covers the entire US East Coast and Gulf Coast with high resolution (<100 m; ~70 m for land surface) concentrated in the Delaware Bay, NJ; the Hudson River Valley up to Troy, NJ; New York City; Long Island Sound; and Long Island. The mesh has been calibrated and validated against various historical tropical and extratropical events on the Northeast Coast [FEMA, 2014b] and used by the NPCC2 for modeling NYC coastal flooding [Orton et al., 2015]. In the model calibration and validation study undertaken by the FEMA [2014b] which evaluated 218 storm surge events, in most cases, more than 75% of the modeled peak surge levels are within 0.46 m of the measured peaks, but in a few cases less than 70% of the comparisons are within 0.46 m. For all the storms they evaluated, the average difference between the modeled and observed peaks is 0.05 m, with the absolute average difference being 0.32 m. The accuracy of peak surge level prediction depends on both the quality of the mesh used, and the information about the storm modeled (e.g., accuracy of the wind field).

The ADCIRC storm-tide modeling was driven by the surface wind, sea level pressure, and tidal forcing. Tidal forcing was applied at the open boundary by eight tidal constituents (K1, K2, M2, N2, O1, P1, Q1, and S2). We applied a parametric approach to simulate the wind and pressure associated with Hurricane Sandy, given the storm characteristics (center location, maximum wind speed, minimum central pressure, and radius of maximum wind) obtained from the Extended Best Track database [Demuth et al., 2006; updated]. As in previous studies [Lin et al., 2012; Lin and Emanuel, 2015], the sea level pressure was estimated from a simple parametric model [Holland, 1980], and the surface wind was estimated by fitting the wind velocity to a symmetrical hurricane wind profile and adding a fraction (0.55 at 20° cyclonically) of the storm translation velocity to account for the asymmetry of the wind field induced by the surface background wind [Lin and Chavas, 2012]. However, here we applied a newly developed, complete wind profile constructed by mathematically merging theoretical solutions for the radial wind structure at the top of the boundary layer in the inner ascending and outer descending regions of the storm [Chavas et al., 2015]. Theoretically based, this wind profile compares well with various satellite and surface observations [Chavas et al., 2015; Chavas and Lin, 2016]. In this study, we also find that this complete wind profile performs better in estimating Sandy’s extensive storm tides than the previous wind model that focused mainly on the inner ascending region of the storm [Emanuel and Rotunno, 2011].

2.2. Flood Inundation Modeling

Inland flood inundation was undertaken using the 2-D flood inundation module of FloodMap [Yu and Lane, 2006a, 2006b], which is a 1-D/2-D-coupled research code for flood inundation modeling. The model is the basis of (i) a subgrid treatment approach [Yu and Lane, 2006b] which utilizes subgrid fine topography in a simulation with a coarser mesh and (ii) its parallel version FloodMap-Parallel [Yu, 2010]. The model has been tested and verified with a range of boundary conditions and in a number of environments [e.g., Yu, 2005;
version of FloodMap is used in the inertial model to calculate the time step: 

\[
\frac{\partial q}{\partial t} + gh_i (h + z) + \frac{g n^2 q^2}{R^{1/3} h} = 0
\]

where \( q \) is the flow per unit width, \( g \) is the acceleration due to gravity, \( R \) is the hydraulic radius, \( z \) is the bed elevation, \( h \) is the water depth, and \( n \) is the Manning’s roughness coefficient. \( R \) can be approximated with \( h \) for wide and shallow flows. Discretizing the equation with respect to time produces the Courant-Freidrich-Levy Condition (FCFL) approach described in Taylor et al. [2007; Lane et al., 2007, 2008; Yu and Lane, 2011], with a more recent addition of an urban surface water flood modeling component [Yu and Coulthard, 2015; Yin et al., 2013, 2016a, 2016b]. Depending on the quality of topography and flow boundary conditions, the published performance of FloodMap in terms of the extent prediction varies but the \( Fit \) statistic is consistently within a range of mid 60% and lower 80%, suggesting a good level of predictive skill.

FloodMap assumes that the floodplain is protected by an embankment (river/coast) that essentially acts as a continuous, broad-crested weir through which flow exchange occurs between the channel/sea and floodplain. The model solves the one-dimensional river flow and two-dimensional floodplain flows simultaneously in a raster environment. The one-dimensional in-channel model solves the full Saint-Venant equations for unsteady open-channel flow using the Preissmann scheme based upon the fixed bed model of Abbott and Basco [1989]. For coastal flood inundation modeling in NYC, as the river flow simulation is not applied here, the 1-D module is not used. Instead, time series of sea level along the coastal line as predicted by ADCIRC is used as input to drive the 2-D inundation modeling.

The 2-D flood inundation model (FloodMap-Inertial) takes the same structure as the inertial model of Bates et al. [2010], but with a slightly different approach to the calculation of the time step. Neglecting the convective acceleration term in the Saint-Venant equation, the momentum equation becomes

\[
\frac{\partial q}{\partial t} + \frac{g h_i (h + z)}{x} + \frac{g n^2 q^2}{h_i^{1/3}} = 0
\]

where one of the \( q_i \) in the friction term can be replaced by \( q_{i+\Delta t} \), resulting in the explicit expression of the flow at the next time step:

\[
q_{i+\Delta t} = q_t - \frac{g h_i \Delta t}{(1 + g h_i \Delta t \Delta m^2 q_t / h_i^{10/3})} \left( \frac{\partial h_i (h + z)}{\partial x} \right)
\]

The flows in the \( x \) and \( y \) directions are decoupled and take the same form. Discharge is evaluated at the cell edges and depth at the center. To maintain model stability and minimize numerical diffusion, the Forward Courant-Freidrich-Levy Condition (FCFL) approach described in Yu and Lane [2011] for the diffusion-based version of FloodMap is used in the inertial model to calculate the time step:

\[
\Delta t = \min \left( \frac{w d_i d_j}{d_i^{1/6} (S_i)^{1/2} + d_j^{1/6} (S_j)^{1/2}} \right)
\]

where \( w \) is the cell size, \( d_i \) and \( d_j \) are the effective water depths; \( S_i \) and \( S_j \) are water surface slopes; and \( i \) and \( j \) are the indices for the flow direction in the \( x \) and \( y \) directions, respectively. The effective water depth is defined as the difference between the higher water surface elevation and the higher bed elevation of two cells that exchange water. The minimum time step that satisfies the FCFL condition for all the wet cells is used as the global time step for this iteration. This approach does not require the back calculation of Courant number as the time step is calculated based on the CFL condition that satisfies every wet grid cell for the current iteration. As the FCFL condition is not strictly the right stability criteria for an inertial system, this scheme still may not guarantee a stable solution, and thus may still produce unrealistic wave propagation. The universal time step calculated with FCFL may need to be scaled further by a coefficient, the value of which ranges between 0 exclusive and 1 inclusive. A scaling factor of 0.5–0.8 was found to give stable solution to all the simulations carried out in this study, and a scaling factor of 0.7 was used in all the simulations undertaken.

### 2.3. Model Coupling

FloodMap required two types of datasets for a simulation: floodplain topography and flow boundary conditions. In terms of the floodplain topography, a NYC DEM constructed from LiDAR point cloud is available.
This is provided as a “bare earth” topography; features such as buildings and trees are removed, with a horizontal resolution of 30 cm and a vertical accuracy of ±10–20 cm. The DEM is used as a base map and can be further reduced to coarser resolutions. Building representation in 2-D flood inundation model is an active field of research as buildings represent blockage to flow and may act as storage [e.g., Yu and Lane, 2006b; Fewtrell et al., 2008; Neal et al., 2011]. Schubert and Sanders [2012] applied a 2-D flood model to evaluate the impact of four building treatment approaches including building resistance, building block, building hole, and building porosity. Results demonstrated that all four methods are capable of high predictive skill for flood extent and stream flow, but the best method for a particular application will depend on available data, computing resources, time constraints, and the specific modeling objectives. Considering the size of the simulation domain and focus of the paper, we did not evaluate building treatment in our analysis.

The 15 min stage hydrographs at points every 300 m along the NYC floodplain-sea boundary were derived from the ADCIRC model run. An automatic procedure based on Python script was developed to generate spatial and temporal grids at the flood-sea boundary, which are used as inputs to FloodMap as flow boundary conditions. This procedure allows multiple simulations to be set up and processed relatively quickly for a coupled simulation by integrating ADCIRC and FloodMap. In this case study, simulated 48 h storm-tide hydrographs during Hurricane Sandy were applied, starting at 7:00 UTC on 29 October and including the rapidly rising phase shortly before inundation occurred and the falling limb when flood gradually receded. Inundation extents and water depths obtained from (i) ADCIRC and (ii) coupled ADCIRC-FloodMap solutions are compared, using USGS high water marks (HWMs) and FEMA authoritative flood extents as validation data.

2.4. Observed Flood Data
Modeled outputs were compared against observed flood level and inundation extent during Hurricane Sandy. Time series of tidal records at NOAA gauge stations in NYC (i.e., Battery and Bergen Point) were collected to validate the storm-tide model. FloodMap modeled maximum flood water heights were directly compared against the highest water levels obtained from USGS HWMs, which were deployed along the NYC coast prior to storm landfall. The HWMs collected by sensors were initially referenced to permanent objects and subsequently adjusted using survey-grade GNSS equipment. The reference points and HWMs were surveyed to a vertical accuracy of 7.9 cm at the 95% confidence level and 3.05 m horizontally [McCallum et al., 2013]. A complete picture of observed flood extent is not available. As an alternative, the FEMA Modeling Task Force [2013] created a high-resolution Sandy flood extent for NYC by interpolating a water surface elevation from the field verified HWMs and subsequently subtracting the best available DEM. The FEMA flood map was widely accepted as the most reliable existing estimate of Sandy-induced inundation extent [e.g., Wang et al., 2014; Ramirez et al., 2016] and was therefore used as the main reference for inundation observation in this study.

2.5. Inundation Change Detection
Three statistics are used to assess the degree of matching and variation between model predictions, including the inundation area, Fit statistic (F), and Root-Mean-Square Errors (RMSEs). In each simulation, the F and RMSE measures are calculated against the FEMA flood extent and HWMs respectively. The F statistic is commonly used for assessing the goodness of agreement between model predicted inundation extent and the reference [Bates and De Roo, 2000; Horritt and Bates, 2001]. It is defined as follows:

\[
F = \frac{A_o}{A_r + A_o - A_o}
\]

where \( A_o \) is the referenced wet areas, \( A_r \) is the predicted wet areas, and \( A_o \) is the overlap of \( A_r \) and \( A_o \). \( F \) varies between 1 for a perfect fit and 0 when no overlap exists.

The RMSE is particularly suitable for the comparison of water levels (or depths) between two paired results, on a cell-by-cell (point-to-point) basis in the case of flood inundation modeling. The metric can be defined by

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=0}^{n} (d_i^p - d_i^o)^2}{n}}
\]

where \( d_i^p \) and \( d_i^o \) are the predicted and referenced water levels (or depths), respectively; \( i \) is the index of the wet cells; and \( n \) is the total number of wet cells in the prediction and observation. This metric is useful for...
evaluating the overall agreement/discrepancy of water depth between predicted and observed/reference water depths.

3. Results and Discussions

Figure 1 shows the time series of the observed versus ADCIRC-modeled water levels (storm tide) at the NYC tide gauge stations (i.e., Battery and Bergen Point). The predicted water levels (phase and amplitude) mostly agree well with the observations. The timing and stage of peak values are characterized by a high degree of consistency between the hindcast simulation and the observations. It is also noted that there are discrepancies in the model results before and after the peak, likely induced by the simple nature of our parametric hurricane wind profile for modeling Sandy, which had significant extratropical and hybrid features [Galarneau et al., 2013]. The deviations presented here resemble SLOSH predictions of Sandy at NY stations, using also parametric hurricane wind modeling, by Forbes et al. [2014]. Nevertheless, the storm-tide simulation provides adequate inputs for driving the inundation modeling, as the flood inundation depends predominantly on the peak of the water level. We perform both city-scale and street-scale modeling of the Sandy-induced flood inundation driven by the ADCIRC-modeled storm tide.

3.1. City-Scale Model Validation

At the city-scale, mid and southern NYC is selected as the study area including lower Manhattan, Jamaica Bay, and Staten Island (we excluded the north part of the city from our analysis as the impact in this area is minor). The maximum inundation estimated by FEMA and ADCIRC and their differences are presented in Figure 2. The validation process yields a low $F$ value of 0.57, revealing a significant spatial discrepancy. Compared to the results of the FEMA static mapping based on the USGS HWMs, the ADCIRC model underestimates the maximum flood extent, particularly in the extensive flat and low-lying area of southern Queens and eastern Brooklyn around Jamaica Bay. In terms of flood depth, the ADCIRC's model results are substantially lower (by up to 90 cm) than FEMA’s flood elevations at most open estuarine land areas (i.e., around Jamaica Bay and Lower New York Bay), but slightly higher (by mostly less than 30 cm) along the western shore of Staten Island, Upper Bay, and lower Manhattan. Apart from the potential errors of the wind field modeling as mentioned above, spatial interpolation of topographic dataset into the relatively low resolution (≈70 m in the land area of NYC) ADCIRC model grid scale contributed to the inconsistency as spatial interpolation from a finer grid to a coarser one tends to smooth the topographic surface [Yu and Lane, 2006a], remove hydraulic connectivity, and slow down flow propagation in urban environment [Ozdemir et al., 2013].

The FloodMap model is applied to the study area at a resolution of 45 m by spatially averaging the elevation of the original 0.3 m LiDAR DEM. In order to evaluate the model sensitivity to floodplain friction, a number of Manning’s $n$ values (0.01–0.1 at a 0.01 interval) are used in this analysis. This range of plausible roughness values may be, in the higher end, close to the range of $n$ values for which Manning’s relationship was originally formulated (e.g., as in Yu and Lane [2006a]). Two examples of the model
Figure 2. The USGS measured maximum inundation estimated by (a) FEMA and (b) the ADCIRC simulated maximum inundation, and (c) their difference (FEMA-ADCIRC) at city scale during Hurricane Sandy.
outputs and differences (FloodMap versus FEMA) are illustrated in Figure 3, which shows the simulations with \( n = 0.01 \) and \( n = 0.1 \) for Sandy. The results demonstrate an improved spatial agreement with FEMA’s prediction, with \( F \) values ranging from 0.72 for \( n = 0.01 \) to 0.70 for \( n = 0.1 \). It might be counterintuitive that a simulation with \( n = 0.01 \) generates a higher \( F \) statistic than a simulation with \( n = 0.1 \), although the difference is marginal, especially when we consider the use of a digital terrain model which does not contain building heights, hence should have a higher than normal roughness value to compensate building blockage effect in urban areas, e.g., in Manhattan. However, we consider roughness used herein as a means to parameterization which may compensate the effects of simplified process representation (e.g., dispersion terms associated with secondary circulation, diffusion terms associated with turbulence) and topographic coarsening [Yu and Lane, 2006a; Lane, 2005], and in this study the effect of building blockage and the inaccuracy in FEMA’s flood mapping. Future analysis could evaluate the use of distributed roughness according to land use and different building treatment methods [e.g., Schubert and Sanders, 2012], which may further improve the results and offer more insight into the role of roughness in flood inundation modeling.

The close range of \( F \) statistics for simulations with various roughness specifications suggest that changing roughness does not result in significant changes in the maximum inundation area (Figure 3), particularly in terrains with high coastal gradients (e.g., Staten Island), which limit further inland propagation of flood waves. Similar results were obtained in a study by Yu and Lane [2006a, 2006b] which highlighted the limitation of using point at a time inundation extent to distinguish simulations. However, plotting the difference between the observed and predicted maximum inundation depth reveals the spatial sensitivity of the model to friction (Figure 3). To evaluate the model performance against observation data, high-accuracy temporal flood extent and depth observation data are needed to better evaluate a model’s performance. The
spatial and temporal sensitivity of the model to roughness will be further explored in the street-scale modeling, focusing on temporal F statistic and overall depth variation (RMSE).

A significant discrepancy between the FEMA inundation map and modeling is noticed for the John F. Kennedy (JFK) International Airport area: it was completely flooded in the FEMA map while largely dry in both ADCIRC and FloodMap modeling results. Although the Hurricane Sandy flood extent produced by FEMA is regarded as the authoritative reference for model benchmarking, possible uncertainties and errors can be introduced from spatial interpolation of limited and unevenly distributed monitoring points (65 HWMs throughout NYC, in many areas spaced over distances of several kilometers). According to an online report (http://www.georgetownclimate.org/resources/jfk-airport-runway-13r-31l-rehabilitation-john-f-kennedy-international-airport-new-york-ci), the storm water rose onto the southern safety area (e.g., the shoulder and erosion surfaces) of JFK’s bay runway during Sandy but did not reach the primary areas of the airport. However, due to the lack of monitoring points in JFK, widespread inundation has been mistakenly modeled by FEMA, revealing that the static approach is subject to overestimation in wide/flat area applications. In addition, many local inland streets, which were not hydraulically connected to the flood, have been wrongly classified as inundated area because FEMA’s planar method considered neither mass conservation nor hydraulic connectivity and surface roughness. The likely overestimated nature of FEMA flood map is thought to contribute to the relatively low 57% F value for ADCIRC and 70%–72% for the coupled solution. It is expected that if the apparent errors (e.g., the JFK area) are removed from the FEMA flood map, the F values are likely to improve for both models. However, the exact extent of the wrongly classified areas in the FEMA flood map is unknown, preventing a convincing modification of the FEMA flood map.

To evaluate the modeling with direct observations, a comparison of HWM observations and the corresponding model results is shown in Figure 4. The quality of the HWMs was verified by USGS into six categories based on vertical uncertainty [Koenig et al., 2016], including excellent (within ±0.15 cm), good (within ±3 cm), fair (within ±6 cm), poor (within ±12 cm), very poor (more than ±12 cm), and at least this high (ALTH). Thirteen percent of the HWMs in NYC were classified as poor, and the rest were rated as fair (22%), good (32%), and excellent (27%). The quality of the individual HWMs is shown in the comparison (Figures 2 and 3). Similar to the comparison of modeled and observed extent (Figure 3), validation using HWMs shown in Figure 4 also reveals an overall underestimation for both models. The ADCIRC modeling predicted 44 HWMs (out of 65 in total) being flooded, with a vertical accuracy (RMSE) of 0.5 m when dry points are excluded and of 0.54 m when all points are included. Ground elevation was used as the water level of a predicted dry HWM point. The greatest differences between the observed and modeled peaks occurred mainly at locations close to the waterfront. These isolated elevation points at the land/sea boundary might have been assigned inaccurate surface features and elevations in the modeling, as the elevation data may contain a mixture of land and sea in the relatively coarse grid cells. The discrepancies can also be caused by interpolating the elevation data to a relatively coarse terrain mesh. This significant underestimation also
confirms the visual inspection in Figure 4. A larger number of HWM points (50 for \( n = 0.01 \) and 49 for \( n = 0.1 \)) were predicted to be wet in the FloodMap simulations. The FloodMap modeling has an overall RMSE of less than 0.3 m without the dry points, increasing from 0.26 m (\( n = 0.01 \)) to 0.29 m (\( n = 0.1 \)). When dry points are included, RMSE increases to approximately 0.4 m (i.e., 0.41 for \( n = 0.01 \) and 0.42 for \( n = 0.1 \)). This result was satisfactory at the city scale, indicating that, with an improved topographic resolution (45 m), FloodMap improved inland water height prediction on the basis of ADCIRC storm-tide simulation (>70 m). We also note that although the overall vertical accuracy of the reference points and HWMs is less than 0.079 m at the 95% confidence level for the majority of the HWMs surveyed by USGS during Hurricane Sandy [McCallum et al., 2013], the absolute error and uncertainty of individual point data can be much larger. In particular, the uncertainty of HMW estimate postflood can be significantly high as noted in a previous study by Neal et al. [2009]. This is more associated with the challenges in determining whether the perceived high water mark evidence (e.g., seed line or debris) represents the highest water mark reached at the site, rather than the survey accuracy. However, as the accuracy of the HWMs (less than \( \pm 12 \) cm) used herein is much lower than the typical vertical accuracy of LiDAR data (\( \pm 15 \) cm) and the quality control undertaken by the USGS, we consider that the HWMs offer a robust dataset for model comparison.

3.2. Street-Level Model Evaluation
Lower Manhattan is chosen as the test case for model validation at street level as it is physically and socio-economically vulnerable to coastal flooding and because Sandy wreaked havoc in the area. The local-scale maximum inundation and the difference derived by FEMA and ADCIRC (the same as city-scale modeling) during Hurricane Sandy are presented in Figure 5. The general pattern of flood extent when compared with FEMA’s result shows an \( F \) value of 0.7. However, similar to what has been reported at city scale, inundation depth is overestimated in the ADCIRC simulation for the majority of coast areas. Furthermore, there are three significant areas of discrepancies: (1) the New York City Supreme Court building, (2) the World Trade Center, and (3) the intersection of Canal Street and West Broadway and its adjacent region (Figure 5). Location 1, located within an isolated inland point, is spuriously predicted to be wet in FEMA’s map; dynamic modeling can avoid such problems. Moreover, the Ground Zero construction pit (Location 2) in the Financial District, which is the lowest part of downtown Manhattan, is estimated to be dry in FEMA’s map, but it was reported by various media outlets to be flooded during the storm with photographic and video records (e.g., ArchitechMagazine available at http://www.architectmagazine.com/article/flooding-at-the-world-trade-center-site_o). This flooding has been faithfully captured by the dynamic model. Conversely, Location 3 is estimated to be wet by FEMA but dry by ADCIRC, because a narrow and low-lying route-way at the corner of Canal and Hudson streets is “smeared out” in the coarse topography used by ADCIRC, thus restricting the flow of flood water to Location 3.
To arrive at a better understanding of how sensitive the maximum inundation is to model resolution and friction at finer resolutions (i.e., street scale), the flood event is then simulated with FloodMap using three different spatial resolutions (3, 6, and 12 m) and varying Manning’s $n$ values (0.01–0.1 at a 0.01 interval). Results suggest that in terms of flood extent, the FloodMap model performs well when compared with FEMA’s prediction (Table 1). The difference of maximum inundation between simulations is not readily discernable with $F$ values ranging from 0.81 (12 m resolution) to 0.83 (3 m). Higher resolution DEMs used in the model did not significantly improve the overall accuracy. This can be explained by the presence of lateral topographic confinement present in the floodplain and the fact that the storm water is restricted to coastal low-lying areas. The FloodMap simulation captures the flood conditions in Locations 1 and 2, similar to the ADCIRC modeling, but FloodMap also faithfully predicted the flooding in Location 3 (Figure 6c). With the increase in grid resolution, flood extent and depth reduce slightly because detailed surface features represented by high-resolution DEM may block specific flow pathways connected to Locations 2 and 3. These results suggest that the sensitivity of inland flooding to grid resolution is more pronounced than that of inundation at the waterfront close to the ocean-land boundary where forcing inputs are imposed.

### Table 1. FloodMap Model Accuracy for Maximum Inundation of Each Simulation at Street Level

| Accuracy Metrics | 3 m $n = 0.01$ | 3 m $n = 0.1$ | 6 m $n = 0.01$ | 6 m $n = 0.1$ | 12 m $n = 0.01$ | 12 m $n = 0.1$ |
|------------------|---------------|---------------|--------------|--------------|--------------|--------------|
| $F$ statistic    | 0.83          | 0.83          | 0.82         | 0.82         | 0.82         | 0.81         |
| RMSE             | 0.12          | 0.12          | 0.10         | 0.10         | 0.19         | 0.20         |

To arrive at a better understanding of how sensitive the maximum inundation is to model resolution and friction at finer resolutions (i.e., street scale), the flood event is then simulated with FloodMap using three different spatial resolutions (3, 6, and 12 m) and varying Manning’s $n$ values (0.01–0.1 at a 0.01 interval). Results suggest that in terms of flood extent, the FloodMap model performs well when compared with FEMA’s prediction (Table 1). The difference of maximum inundation between simulations is not readily discernable with $F$ values ranging from 0.81 (12 m resolution) to 0.83 (3 m). Higher resolution DEMs used in the model did not significantly improve the overall accuracy. This can be explained by the presence of lateral topographic confinement present in the floodplain and the fact that the storm water is restricted to coastal low-lying areas. The FloodMap simulation captures the flood conditions in Locations 1 and 2, similar to the ADCIRC modeling, but FloodMap also faithfully predicted the flooding in Location 3 (Figure 6c). With the increase in grid resolution, flood extent and depth reduce slightly because detailed surface features represented by high-resolution DEM may block specific flow pathways connected to Locations 2 and 3. These results suggest that the sensitivity of inland flooding to grid resolution is more pronounced than that of inundation at the waterfront close to the ocean-land boundary where forcing inputs are imposed.

**Figure 6.** A comparison of HWM observations and the corresponding FloodMap model results (peak values) at street level: (a) 3 m $n = 0.01$, (b) 6 m $n = 0.01$, (c) 12 m $n = 0.01$, (d) 3 m $n = 0.1$, (e) 6 m $n = 0.1$, and (f) 12 m $n = 0.1$. 

**Table 1.** FloodMap Model Accuracy for Maximum Inundation of Each Simulation at Street Level

| Accuracy Metrics | 3 m $n = 0.01$ | 3 m $n = 0.1$ | 6 m $n = 0.01$ | 6 m $n = 0.1$ | 12 m $n = 0.01$ | 12 m $n = 0.1$ |
|------------------|---------------|---------------|--------------|--------------|--------------|--------------|
| $F$ statistic    | 0.83          | 0.83          | 0.82         | 0.82         | 0.82         | 0.81         |
| RMSE             | 0.12          | 0.12          | 0.10         | 0.10         | 0.19         | 0.20         |
In terms of flood water level, HWMs in lower Manhattan (19) are utilized to compare the model results with the observed highest water levels (Figure 6). The statistical results (RMSE) for representative comparisons are presented in Table 1. In general, the model provides good predictions of flood height with a RMSE around 0.1–0.2 m, commensurate with the typical accuracy of LiDAR data, and it improves over the ADCIRC model (RMSE = 0.29 m for the same area). Similar to what we found in the city-scale modeling, the maximum extent does not distinguish between simulations with varying spatial resolutions and friction values. However, a coarser mesh of 12 m resolution results (Figures 6c and 6f) in an overestimated water depth at a HWM point close to the wetting front further inland (circled in Figure 5), with a difference of more than 60 cm (circled in red Figure 6). The uncertainty in the observed water level at this HWM, which was

Figure 7. Time series of inundation area, $F$ statistic, and depth RMSE for FloodMap simulations with different resolutions (left column for 3 m, middle column for 6 m, and right column for 12 m) and roughness. Simulations with $n = 0.01$ are used as the reference for $F$ statistic and depth RMSE.
estimated based on mudline on a building window, is relatively low and the quality is rated as “good” (within ±3 cm) according to USGS. The greater than average difference at this point compared to the differences for the same point with the 3 m (Figures 6a and 6d) and 6 m (Figures 6c and 6e) meshes suggests that the larger discrepancy is likely due to the mesh coarsening effect which is amplified at the wetting front.

In addition to flood extents and maximum inundation, the time series of the total inundation area, \( F \) statistic, and depth RMSE (with the \( n = 0.01 \) simulation as reference) are also investigated (Figure 7) to evaluate model sensitivity to mesh resolution and roughness. Comparison of the inundation areas leads to three observations. First, the inundated area is synchronized with the storm tide and the model generally captured the timing of overland flow. Second, the curves display a consistent shape during the rising phase, but there are clear deviations in the falling limb. Furthermore, with the increase of resolution, the extent of peak inundation increases slightly, and the model’s sensitivity to roughness becomes more pronounced. The findings are more evident in the \( F \) statistic plots when inundated areas are compared with a reference simulation (i.e., \( n = 0.01 \)) through time. According to Yu and Lane (2006a), the coarser resolutions may result in a smoother floodplain, reduce the effects of cell blockage and simplify the representation of wetting processes, indicating that DEM resolution has a major effect on the inundation time series results. Depth RMSE is further calculated to distinguish the temporal evolution of flood depth between each simulation. Results show significant variations in the temporal pattern of agreement throughout the simulations (\( F \) statistic and RMSE in Figure 7), suggesting the model’s strong sensitivity to both resolution and roughness.

Overall, for both the city-scale and street-scale modeling, the coupled ADCIRC-FloodMap solution demonstrates that when evaluated against the FEMA flood extent, the model is insensitive to roughness and mesh resolution. For example, in the city-scale, \( F \) values range from 0.72 for \( n = 0.01–0.70 \) for \( n = 0.1 \), and in street-scale modeling, \( F \) values range from 0.81 (12 m resolution) to 0.83 (3 m). However, when the temporal comparison of extent and water depth is undertaken, the model is shown to be sensitive to both roughness and mesh resolution (Figure 7).

4. Conclusions

This study demonstrated the integration of a widely used storm surge model (ADCIRC) and a 2-D hydrodynamic flood model (FloodMap) for flood inundation modeling in the highly urban environment of New York City. The integration is facilitated by an automated procedure of postprocessing ADCIRC outputs as inputs for inundation modeling. This coupling allows urban inundation to be modeled at both the city-scale the street-scale resolutions.

We cross-compare the predictions of ADCIRC and FloodMap for inland inundation using USGS HWMs observation and an inundation map derived by FEMA from the HWMs using a static planar method. This was undertaken at both the city and subregional (i.e., lower Manhattan) scales. City-scale simulations demonstrate an improved performance by FloodMap over ADCIRC, with a \( F \) statistic value increasing from 57% to 70%–72%. This is likely due to the slightly finer grid resolution used by FloodMap. The advantage of coupling ADCIRC and FloodMap is further demonstrated in the subregional modeling of flood inundation at the street-level scale. The computational grid used by the ADCIRC surge model typically covers large coastal areas and extends to the deep ocean, and thus the grid resolution is limited for inland areas. However, by integrating ADCIRC and FloodMap we are able to achieve much finer resolutions (~3 m), allowing inundation dynamics to be simulated at the street level. The prediction of subregional and street-level inundation is required for various flood-related risk analysis, including, in particular, impact analysis on critical infrastructure (e.g., traffic disruption and its cascading impacts on emergency response) [Yin et al., 2016a; Yu et al., 2015].

References

Abbott, M. B., and D. R. Basco (1989), Computational Fluid Dynamics: An Introduction for Engineers, Longman, London, U. K.

Aerts, J. C., N. Lin, W. J. Botzen, K. Emanuel, and H. de Moel (2013), Low-probability flood risk modeling for New York City, Risk Anal., 33, 772–788.

Aerts, J. C., W. J. Botzen, K. Emanuel, N. Lin, H. de Moel, and E. O. Michel-Kerjan (2014), Evaluating flood resilience strategies for coastal mega-cities, Science, 344, 473–475.

Bates, P. D., and A. P. J. De Roo (2000), A simple raster-based model for flood inundation simulation, J. Hydrol., 236, 54–77.

Acknowledgments

This paper was supported by the National Science Foundation of the United States (grant EAR-1520683), the National Natural Science Foundation of China (grant 41201550), and Engineering and Physical Science Research Council of the United Kingdom (EP/N005961/1). The research used open source data available in the public domain. Specifically, LiDAR data were acquired from the New York City government data portal (https://data.cityofnewyork.us). ADCIRC simulation used mesh available from http://adcirc.org/products/grids. Higher watermarks can be obtained from USGS at http://water.usgs.gov/floods. Static FEMA flood map for Hurricane Sandy is available from https://www.arcgis.com/home/item.html?id=307dd5224994d44a33d7296a5da5ea0. Results generated are available upon request. We also would like to thank Paul Bates at Bristol University, Jorge Ramirez at Colorado State University, and an anonymous reviewer for their thorough review and constructive comments which improved the paper’s quality.
Bates, P. D., R. J. Dawson, J. W. Hall, M. S. Horritt, R. J. Nicholls, J. Wicks, and M. A. A. M. Hassan (2005), Simplified two-dimensional numerical modelling of coastal flooding and example applications, Coastal Eng., 52, 793–810.

Bates, P. D., M. Horritt, and T. Feuvrillé (2010), A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, J. Hydrol., 387(1–2), 33–45.

Blake, E. S., T. B. Kimberlain, R. J. Berg, J. P. Cangialosi, and J. L. Beven (2013), Tropical Cyclone Report, Hurricane Sandy (AL182012) 22–29 October 2012, pp. 1–300, Natl. Hurricane Cent., Miami, Fla.

Blumberg, A. N., Georgas, L. Yin, T. Harrington, and P. Orton (2015), Street scale modeling of storm surge inundation along the New Jersey Hudson River waterfront, J. Atmos. Oceanic Technol., 32(2), 1486–1497, doi:10.1175/JTECH-D-14-02213.1.

Brown, B. J. (2006), Modelling future landscape change on coastal floodplains using a rule-based GIS. Environ. Model. Software, 21, 1479–1490.

Brown, J. D., T. Spencer, and I. Moeller (2007), Modeling storm surge flooding of an urban area with particular reference to modeling uncertainties: A case study of Canvey Island, United Kingdom, Water Resour. Res., 43, W06402, doi:10.1029/2005WR004597.

Chavas, D. R., and N. Lin (2016), A model for the complete radial structure of the tropical cyclone wind field. Part II: Wind field variability, J. Atmos. Sci., 72(9), 150709110433002, doi:10.1175/JAS-D-15-0014.1.

Chavas, D. R., N. Lin, and K. A. Emanuel (2015), A model for the complete radial structure of the tropical cyclone wind field. Part I: Comparison with observed structure, J. Atmos. Sci., 72(9), 3647–3662.

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.

Demuth, J. L., M. DeMaria, and J. A. Knaff (2006), Improvement of advanced microwave sounding unit tropical cyclone intensity and size estimation algorithms, J. Appl. Meteorol. Climatol., 45(11), 1573–1581.

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

Emanuel, K. A., and R. Rotunno (2011), Self-stratification of tropical cyclone outflow. Part I: Implications for storm structure, J. Atmos. Sci., 68(10), 2236–2249.

FEMA (2014a), Region II Coastal Storm Surge Project: Mesh Development, Washington, D. C.

FEMA (2014b), Region II Coastal Storm Surge Project: Model Calibration and Validation, Washington, D. C.

FEMA Modeling Task Force (2013), Hurricane Sandy Impact Analysis, FEMA. [Available at http://www.arcgis.com/home/item.html?id=307dde52299946a4a4a33d7296a5da5ea0l].

Feuvrillé, T. P. D. Bates, M. Horritt, and N. Hunter (2008), Evaluating the effect of scale in flood inundation modelling in urban environments, Hydrol. Processes, 22, 5107–5118, doi:10.1002/hyp.7148.

Forbes, C., J. Rhome, C. Mattocks, and A. Taylor (2014), Predicting the storm surge threat of Hurricane Sandy with the National Weather Service SLOSH Model, J. Mar. Sci. Eng., 2, 437–476.

Galanteau, T., A. D. Christopher, and A. S. Melvyn (2013), Intensification of Hurricane Sandy (2012) through extratropical warm core slcon, Mon. Weather Rev., 141, 4296–4321, doi:10.1175/MWR-D-13-00181.1.

Gallen, T. W., J. E. Schubert, and B. F. Sanders (2011), Predicting tidal flooding of urbanized embayments: A modeling framework and data requirements, Coastal Eng., 58, 567–577.

Hallegraeff, S. C. Green, R. J. Nicholls, and J. Corfee-Morlot (2013), Future flood losses in major coastal cities, Nat. Clim. Change, 3, 802–806.

Hanson, S., R. Nicholls, N. Ranger, S. Hallegraeff, J. Corfee-Morlot, C. Herweijer, and J. Chateau (2011), A global ranking of port cities with high exposure to climate extremes, Clim. Change, 104, 89–111.

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

Horritt, M. S., and P. D. Bates (2001), Effects of spatial resolution on a raster based model of flood flow, J. Hydrol., 253, 239–249.

IPCC (2013), Climate Change 2013: The Physical Science Basis, Cambridge Univ. Press, Cambridge, U. K.

IPCC (2014), Climate Change 2014: Impacts, Adaptation and Vulnerability, Cambridge Univ. Press, Cambridge, U. K.

Jongman, B., P. J. Ward, and J. C. Aerts (2016), Global exposure to river and coastal flooding: Long term trends and changes, Global Environ. Change, 22, 823–835.

Koenig, T. A., et al. (2016), Identifying and Preserving High-Water Mark Data, Techniques and Methods, Book 3, chap. A24, 47 pp., U.S. Geol. Surv., Reston, Va., doi:10.3133/tm3A24.

Lane, S. N. (2005), Roughness – time for a re-evaluation?, Earth Surf. Proc. and Land., 30(2), 251–253.

Lane, S. N., V. Tayefi, S. C. Reid, D. Yu, and R. J. Hardy (2007), Interactions between sediment delivery, channel change, climate change and flood risk in a temperate upland environment, Earth Surf. Proc. Land., 32(3), 429–446.

Lin, N., D. Chavas, J. J. Westerink, and N. W. Scheffner (1992), ADCCIRC: An advanced three-dimensional circulation model for shelves, coasts, and estuaries. Report 1: Theory and methodology of ADCCIRC-2DDI and ADCCIRC-3DL, Dredging Res. Program Tech. Rep. DRP-92-6, Coastal Eng. Res. Cent., Vicksburg, Miss.

McCallum, B. E., S. M. Wicklein, R. G. Reiser, R. Buscioiano, J. Morrison, R. J. Verdi, A. J. Painter, and A. J. Gotvald (2013), Monitoring storm tide and flooding from Hurricane Sandy along the Atlantic coast of the United States, October 2012, U.S. Geol. Surv. Open File Rep., 2013–043, 42 pp.

Neal, J. C., P. D. Bates, T. J. Feuvrillé, N. M. Hunter, M. D. Wilson, and M. S. Horritt (2009), Distributed whole city water level measurement from the Carlisle 2005 urban flood event and comparison with hydraulic model simulations, J. Hydrol., 368, 42–55.
