The winter and early spring weather in the New York City metropolitan region is highly influenced by extratropical storm systems, and the storm surge associated with these systems is one of the main factors contributing to inundation of coastal areas. This study demonstrates the predictive capability of an established statistical relationship between the "storm maximum" storm surge associated with an extratropical storm system and the "average maximum" significant wave height during that storm. Data from publicly available retrospective forecasts of sea level pressure and wave heights, along with a regression equation for storm surge, were used to predict the storm-maximum storm surge for 41 storms in the New York metropolitan region during the period from February 2005 to December 2008. The statistical storm-surge estimates were compared with the surge values predicted by NOAA's extratropical storm-surge model and NOAA's operational surge forecast, which includes an error correction, and with water gauge observations taken at the Battery, located at the southern tip of Manhattan Island, New York. The mean difference between the statistical surge prediction and the observed values is shown to be smaller than the difference between NOAA's deterministic surge prediction and the observed surge at the 95% significance level and to be statistically indistinguishable from the difference between NOAA's operational surge forecast and the observed values of surge. These statistical estimates can be used as part of a system for predicting coastal flooding.

1. Introduction

The New York City, New York, metropolitan region is particularly vulnerable to the damage caused by the frequent extratropical low pressure systems that occur during autumn, winter, and spring months. In association with the storms, severe conditions such as high winds, heavy rain, blizzards, very low temperatures, and storm surge can prevail from several hours up to a few days. The Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR4; Pachauri et al. 2007) estimates that over the next century global sea level is likely to rise between 18 and 59 cm, and this rise will extend the zone of impact from storms, storm surge, and storm waves farther inland. The potential for property damage and loss of life due to storm surge and flooding necessitates accurate predictions of high water levels associated with storm conditions.

Storm surge is commonly defined as the difference between the predicted astronomical tide and the observed water level associated with a particular phenomenon such as a storm. The predicted astronomical tide is the daily change in water level produced by the gravitational interactions of the earth, moon, and sun; it is calculated by performing complex harmonic analysis of observations of water level at a given location. The
characteristics of storm surge depend on, for example, tides, wind stress, atmospheric pressure, wave height, transport of water by waves and swell, and effects of coastline configuration and bathymetry. Predictions of storm surge have been made using both dynamical and statistical methods. The earliest efforts at dynamical modeling were hampered by the lack of meteorological observations over the water and by the oversimplifications needed to make the dynamics computationally tractable. This led researchers to develop empirical and statistical relationships between wind and pressure fields and water-level changes based on relationships derived from simplified theory and equations of motion. An early method for forecasting the maximum storm surge based on connections between significant wave height and storm surge in the Boston, Massachusetts, region was presented by Tancredo (1958). He used the Sverdrup–Munk method as revised by Bretschneider (1951) to compute significant wave heights from equations that relate wave heights to wind data and a regression equation to forecast storm surge from these computed significant wave heights. Another statistical method, based on linearized two-dimensional hydrodynamic equations, was developed by Harris (1962) and consisted of a regression equation that related the surge at a specific location and time to a “meteorological factor” selected according to the type of observation and the location of the observation station. Harris and Angelo (1963) tested the model using past data from Buffalo, New York, and Toledo, Ohio. They concluded that the prediction obtained with this approach was equivalent or superior to a prediction based on the direct integration of the hydrodynamic equations and using the same data.

A method using statistical relationships derived purely from observations was developed by Pore et al. (1974), who related meteorological data to the observed storm surge. The technique is based on data for 13 winter seasons, from 1956 to 1969, and data from eight East Coast water gauge stations for the storms that caused surge $\geq 2$ ft (61 cm). Atmospheric surface pressure values, with appropriate time lags, were considered as possible predictors of the storm surge. The surge equation for New York City involved predictors at eight grid points with time lags ranging from 0 to 6 h. A test of this method using data from a series of past storms revealed that the time of the peak surge values was determined correctly but that the forecasts were generally too low. This low bias was adjusted by multiplying the forecast surge by the reciprocal of the correlation coefficient between the storm surge and the predictors. This statistical method to forecast storm surge became operational shortly after its development, and it continues to be used today to guide the official National Oceanic and Atmospheric Administration (NOAA) surge forecasts.

DeGaetano (2008) more recently investigated another statistical approach and used a published series of strong East Coast winter storms during the October–April storm season to correlate the occurrence of storms with extreme surge events at three water gauge stations in the New York metropolitan region. Extreme-surge events were defined based on hourly values that exceeded either the 99th or the 99.9th percentile from October through April over the period of record. Analysis revealed that less than 7% (24%) of the extreme-surge events defined using the 99th (99.9th) percentile threshold occurred in association with these storms. DeGaetano (2008) concluded that even though the occurrence of strong storms and extreme surge is correlated the predictive capability of this correlation is limited. He suggested a possible explanation for this limitation associated with the fact that the storms used in the analysis represent coastwide conditions, whereas the extreme-surge events at the water gauge stations used to detect surge were caused only by those storms that directly affect the metropolitan region.

Current dynamical models include the NOAA “ET-SURGE” extratropical storm-surge model developed by Kim et al. (1996), the Advanced Circulation (ADCIRC) model for coastal ocean hydrodynamics of Luettich et al. (1992), and the Estuarine Coastal and Ocean Model (ECOM) of Blumberg and Mellor (1987). ET-SURGE is NOAA’s deterministic, real-time forecast model for extratropical storm-surge information (Ji et al. 2010). It is the extratropical version of the Sea, Lake, and Overland Surfage model (Jelesnianski et al. 1992). The ADCIRC model solves a set of steady-state, barotropic equations and was developed to simulate wind-driven and tidal circulation in coastal waters. The specifics of the ADCIRC grid enable the simulation of flooding of coastal areas above sea level during storm events (Colle et al. 2008). ECOM is the model currently run by the New York Harbor Observing and Prediction System and uses a technique that solves separate equations for the fast, barotropic, external waves and slow, baroclinic, internal waves. Despite the considerable improvement of models and availability of observations, modeling of storm surge and prediction of coastal flooding remain problematic issues. For example, the National Weather Service (NWS) extratropical storm–related flood warnings that were issued to coastal residents during the period 2002–06 had a false-alarm rate of 85% (Colle et al. 2008), indicating the necessity of a better understanding of the complexities determining storm surge.
2. Data and methods

The availability of accurate weather forecasts out to 5–10 days or longer raises the issue of surge forecasts at longer lead times than is current practice. As a result of larger computers and more sophisticated models, numerical weather prediction (NWP) skill has improved markedly over time (Simmons and Hollingsworth 2002), and NWP centers are generally issuing forecasts out to 10 days or longer. The National Centers for Environmental Prediction (NCEP), for example, now issues forecasts out to 16 days. The advent of more reliable seasonal and climate forecasts offers the opportunity for storm-surge forecasts on those longer time scales as well. The IPCC AR4 contains an entire section devoted to regional downscaling of coarse-resolution simulations of future climate (Pachauri et al. 2007, chapter 11), making it possible to drive storm-surge models on regional scales. In addition to characterizations of regional storm strength and frequency in a future climate, characterizing storm surge and inundations would be of interest to regional planners, as it would guide mitigation measures, such as the construction of storm walls.

A new observationally based statistical method for storm surge is evaluated here. It is based on a regression relation between storm surge and significant wave heights established in a study by Salmun et al. (2009, hereinafter S09), who used a focused regional approach to investigate properties of coastal storm systems (winds, precipitation, and waves) and the resulting regional impacts (erosion, storm surge, flooding, and wind damage). The procedure for evaluation presented here uses a set of known storm events and uses time series of forecast significant wave heights along with the regression equation to compute a forecast storm surge. The statistical forecast is then evaluated against dynamically predicted storm surge and is compared with the storm surge calculated from water-level observations at the Battery, located at the southern tip of Manhattan Island, New York. This study seeks to establish the predictive value of the regression equation of S09. Following this introduction, the statistical method and the forecast products used are presented. A discussion of the results and of the comparisons of these results with the observed storm-surge values and with the storm-surge forecast by NOAA is presented in section 3, followed by a summary and conclusions.

2. Data and methods

The method used in this paper follows the work of S09. To aid the description of the method used here, we begin with a summary of the relevant details of S09. The flow diagram of Fig. 1 outlines both methods: that used by S09 and the one used here. The section of the flow diagram above the dotted line corresponds to the early work, and the section below the dotted line corresponds to the current work. In S09, sea level pressure data from National Data Buoy Center (NDBC) stations in the New York City metropolitan area are used to identify East Coast cool-weather storms (ECCSs) based on the times at which the pressure was below a statistically determined threshold. This method is depicted in the progression from left to right across the top of the flow diagram.

The list of storms compiled in this manner, along with wind and wave data collected at the NDBC station and the observationally based storm surge at the Battery, was used to compute storm-composite values of wind and wave fields and observed storm surge. This step of the method is depicted in the center of the upper half of the diagram. In most instances the storm composite is a simple average of field values over the duration of the storm. The exception is the storm-composite storm surge, defined here as the maximum value of the surge attained during the storm period, hereinafter referred to as “storm maximum” storm surge (SSMAX). The storm-composite values were used to perform a regression analysis to determine the best storm-composite predictors of SSMAX at the Battery. This step of the method is depicted in the bottom sequence of the upper half of the flow diagram.

The observed storm-surge data were calculated using water-level data at the Battery for the period 1959–2007 obtained from NOAA (available online at http://tidesandcurrents.noaa.gov/). The values of storm surge were computed as the difference between the observed water levels at the water gauge and the NOAA-predicted astronomical tide levels. The average seasonal cycle in mean sea level caused by changes in ocean variables was removed, and the resulting time series was corrected for a small (on the order of millimeters) sea level rise during the period of observation. Details of the computation of storm surge from observations at the Battery can be found in Colle et al. (2008).

The regression equation constructed to calculate SSMAX at the Battery using the storm-composite significant wave height measured at NDBC station 44025 (denoted by SSMAX_{44025} and H_{44025}, respectively) was

$$SSMAX_{44025} = 0.2055H_{44025} - 0.0851,$$

with root-mean-square (RMS) error of 0.167 m. The fields used as part of the regression analysis reported in S09 were the storm composites of minimum pressure, pressure tendency, wind speed, wind direction, wind gustiness, significant wave height, wave direction, dominant wave period, and storm duration. The regression analysis revealed that SSMAX estimated using the significant
wave height as the sole predictor is statistically equivalent to $SSMAX$ estimated using any other combination of predictors. In addition, S09 reported that the regression analysis performed using observations taken at other NDBC stations in the region showed that the best estimate of observed $SSMAX$ at the Battery was obtained when using predictors that were based on data from NDBC station 44025. A map of the study area, indicating the locations of NDBC station 44025 and the Battery, is provided in Fig. 2.

To establish the predictive value of the regression equation obtained by S09, a series of retrospective forecasts of $SSMAX$ was performed using forecast sea level pressure fields, forecasts of significant wave heights, and the regression equation to compute $SSMAX$. The NOAA ET-SURGE standard forecast of surge was used for comparison, and both the NOAA and the statistical forecasts were compared against storm surge computed from observations of water level at the Battery. This procedure is depicted in the bottom half of the flow diagram in Fig. 1. Retrospective forecasts of sea level pressure were used to verify the existence of candidate events in the forecast record. Forecasts from the NCEP Environmental Modeling Center’s North American Mesoscale (NAM) operational model runs (model information is available online at http://www.emc.ncep.noaa.gov/modelinfo/) were obtained from NOAA’s National Operational Model Archive and Distribution System (NOMADS; see http://nomads.ncdc.noaa.gov/). The choice of NAM forecasts [instead of the NCEP Global Forecast System (GFS) used as driving data for the wave model referred to below] for the purpose of eliminating storms not in the forecast record was motivated by the report of Charles and Colle (2009) that the GFS forecasts outperform the NAM forecasts in terms of
predicting the minimum pressure during storm events. The more conservative criterion resulted in the elimination of one extra storm event from the list of test cases. NAM sea level pressure values are available at 3-h time intervals, and point values at the location of interest (40.25°N, 73.17°W) were calculated using a bilinear interpolation. Any storm that was not forecast was eliminated from the final list of test storms. In addition, storm starting times and ending times were adjusted based on forecast sea level pressure. Results of SSMAX forecasts of 41 storm events selected using the method described here during the period February 2005–December 2008 are presented in this study. This is depicted by the sequence at the top of the bottom half of the flow diagram.

In direct analogy to the procedure used in S09 and depicted in the center of the lower half of the diagram, storm-composite significant wave heights were computed based on retrospective forecasts from NOAA’s “WAVEWATCH III” (WWIII) operational wave model for each storm event (retrospective forecast data were obtained from the Marine Modeling and Analysis Branch server online at ftp://polar.ncep.noaa.gov/pub/history/waves) and were used as the predictor for SSMAX at the Battery. The model output includes point data at locations of the NDBC data stations.

For comparison against the statistical SSMAX prediction, storm-maximum storm surge was computed from NOAA ET-SURGE predictions. ET-SURGE is forced by basin-scale surface winds and sea level pressure from the GFS, and the predicted surge is added to the predicted astronomical tide and to a model error-correction term to produce a prediction of water level at the Battery. (Details of NOAA’s ET-SURGE model are available online at http://www.nws.noaa.gov/mdl/etsurge.) Archived surge predictions at the Battery produced by the ET-SURGE model were obtained from NOAA (A. A. Taylor 2009, personal communication). The NOAA operational storm-surge forecast consists of the ET-SURGE output and an error correction, computed as the 5-day running mean of the previous 5 days’ error of ET-SURGE output. The analysis presented here will include ET-SURGE estimates with and without the error correction, which was computed using the archive data.

The regression equation used to produce the statistical SSMAX forecasts was slightly modified from the regression equation derived in S09. The study reported in S09 derived the regression equation using observations at hourly intervals for the period 1991–2007. WWIII retrospective forecasts are available at 3-hourly intervals and include data for 2008. Therefore, retrospective predictions of SSMAX at the Battery were computed according to a slightly modified regression equation derived from observations at NDBC station 44025 for the period 1991–2008, using 3-hourly data. The equation used in the analysis presented here is

\[
SSMAX_{44025} = 0.1961H_{44025} - 0.0412,
\]
with RMS error of 0.145 m.

The values of SSMAX computed as described above from the statistical forecasts, the NOAA ET-SURGE model output, the NOAA operational forecasts, and the observations at the Battery are compared, and the results are presented in the next section.

3. Discussion of results

Each of the 41 ECCS test events was assigned a storm identifier number (ID). The beginning and end times/dates and duration in hours corresponding to each ID can be found in Table 1 (the storm list). The focus of the discussion presented here is the analysis of 12-h-lead-time forecasts, and results for 24- and 48-h-lead-time forecasts will be briefly summarized for comparison. Figures 3a and 3b show the SSMAX for each predicted ECCS event. The black bars in Fig. 3a represent the SSMAX statistical estimates using our method ("STAT FCST"), and the dark-gray bars in both panels represent the observed SSMAX at the Battery ("OBS"). The error bars in Fig. 3a, represented by the white portion of the dark-gray bars, correspond to the RMS error of the regression, that is, the RMS error associated with the regression estimate of SSMAX when the observed significant wave height is the predictor. The white bars in Fig. 3b represent the SSMAX
computed from ET-SURGE model output, and the light-gray bars represent SSMAX computed from the anomaly-corrected NOAA operational surge forecast.

The error in the SSMAX estimate using the regression equation based on the predicted significant wave heights can be thought of as having a contribution due to the significant wave height forecast error and a contribution due to the regression equation estimate itself. Assessment of the SSMAX prediction error relative to the error in an SSMAX estimate using observed significant wave heights leaves us with a measure of the error due to the predictive nature of the significant wave heights. An SSMAX forecast inside the error bars is one for which using the predicted significant wave heights as an SSMAX predictor is statistically indistinguishable from using the observed significant wave heights.
Figure 3 shows that the observed SSMAX at the Battery (the dark-gray bars common to both panels in the figure) is always positive and ranges from 0.1 to about 0.92 m, whereas the black bars in Fig. 3a range from 0.17 to 0.83 m. Examination of the statistical forecast series in relation to the error bars around the observed series shows that 66% of the points lie inside the error bars. Based on the discussion above, this indicates that in two-thirds or more of the cases using predicted significant wave heights does not have a negative impact on the statistical estimates of SSMAX. Figure 3b shows that the values of SSMAX from the ET-SURGE model output are negative for four of the test cases and range from −0.28 to 0.72 m, whereas the NOAA operational forecast of SSMAX is negative in only one case and ranges from −0.19 to 0.84 m. The anomaly correction itself (not shown) ranges from −0.13 to 0.28, indicating the variability in the error of the ET-SURGE model output.

In Fig. 4, the differences between the statistical and observed SSMAX are shown alongside the differences between the ET-SURGE model output and observed SSMAX and the differences between NOAA’s operational forecast and the observed SSMAX. The black bars represent the error of the statistical estimates of SSMAX using significant wave height, the white bars represent the error of the ET-SURGE model output of SSMAX, and the light-gray bars represent the error of NOAA’s operational forecast of SSMAX. The mean and standard deviation of the difference between estimates of SSMAX using our statistical method and observations are 0.0534 and 0.1591, respectively; those of the error in the ET-SURGE model output of SSMAX are −0.2477 and 0.1186, respectively; and those of the error in NOAA’s operational forecast are −0.1459 and 0.1151, respectively. These metrics are summarized in Table 2 along with those for the 24- and 48-h-lead-time forecasts.

Typical sea level pressure fields from the National Aeronautics and Space Administration Modern-Era Retrospective Analysis for Research and Applications (Bosilovich 2008) are shown in Figs. 5a and 5b for storms for which the statistical SSMAX errors are respectively small and large. Those events for which the errors in SSMAX forecast were smaller (Fig. 5a) were stronger and their centers passed over or close to NDBC station 44025. The storms for which the errors in SSMAX were larger (Fig. 5b) were weaker and their centers passed over or close to NDBC station 44025.

### Table 2: Metrics for the error associated with the statistical estimate of SSMAX (STAT), the NOAA ET-Surge model output estimate of SSMAX (NOAA ET), and the NOAA operational forecast of SSMAX (NOAA ETANOM) at the Battery for the 12-, 24- and 48-h-lead-time forecasts.

<table>
<thead>
<tr>
<th>Lead time</th>
<th>Statistic</th>
<th>STAT − Obs</th>
<th>NOAA ET − Obs</th>
<th>NOAA ETANOM − Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 h</td>
<td>Mean (m)</td>
<td>0.0534</td>
<td>−0.2477</td>
<td>−0.1459</td>
</tr>
<tr>
<td></td>
<td>Std dev</td>
<td>0.1591</td>
<td>0.1186</td>
<td>0.1151</td>
</tr>
<tr>
<td>24 h</td>
<td>Mean (m)</td>
<td>0.0927</td>
<td>−0.2346</td>
<td>−0.121</td>
</tr>
<tr>
<td></td>
<td>Std dev</td>
<td>0.1597</td>
<td>0.1266</td>
<td>0.126</td>
</tr>
<tr>
<td>48 h</td>
<td>Mean (m)</td>
<td>0.0418</td>
<td>−0.2713</td>
<td>−0.137</td>
</tr>
<tr>
<td></td>
<td>Std dev</td>
<td>0.1341</td>
<td>0.1346</td>
<td>0.1474</td>
</tr>
</tbody>
</table>
larger (Fig. 5b) were weaker and passed farther away. This suggests that the regression relation more easily captures the behavior of the surge during stronger and closer storms, possibly because of a more robust physical relationship between wave heights and surge during stronger events and possibly because of the reduced sampling error inherent in the Eulerian nature of our technique.

The statistics of Table 2 indicate that in general the statistical SSMAX estimates tend to slightly underpredict or overpredict the observed SSMAX on average and that the NOAA SSMAX forecasts, with and without the anomaly correction, tend to underpredict the observed SSMAX. A series of statistical tests revealed that the error in NOAA’s ET-SURGE SSMAX is greater than the error of the statistical SSMAX estimates at greater than the 95% significance level while the error in the NOAA operational forecast of SSMAX is statistically indistinguishable from the error in the statistical SSMAX forecasts. The statistical comparisons for the 24- and 48-h-lead-time forecasts show the same pattern. Comparison among the statistical SSMAX forecasts at 12-, 24- and 48-h lead times shows that the 12- and 24-h-lead-time forecast errors are statistically indistinguishable and that both are statistically smaller than the 48-h-lead-time forecast errors at the 95% significance level.

A qualitative analysis of those ECCS events for which the error in the statistical SSMAX prediction was large resulted in a distinction between those events for which the SSMAX error was mainly due to errors in forecast significant wave heights and those events for which the SSMAX error was mainly due to the failure of the regression relation. In general, underpredictions of SSMAX (storm ID 3, 5, 14, 38, and 40) are attributable to errors in forecast significant wave heights whereas overpredictions of SSMAX (storm ID 10, 20, 21, 29, and 30) are attributable to failure of the regression relation.

4. Summary and conclusions

A new statistical method for predicting storm-maximum storm surge related to East Coast cool-weather storms was presented here, demonstrating the predictive capability of an established regression relation between storm “average maximum” significant wave height and storm-maximum surge. The statistical method was tested by performing a series of retrospective forecasts during the period from February 2005 to December 2008, using existing operational forecasts of surface pressure from NOAA’s North American Mesoscale weather forecast model, operational forecasts of wave height from NOAA’s WAVEWATCH III model, and the regression relation established based on observations of sea level pressure and significant wave heights from the NDBC station closest to New York Harbor. The statistical storm-maximum surge prediction was compared with NOAA ET-SURGE model output, with the operational surge forecast from NOAA, and with the water-level observations taken at the Battery in New York City. A distinction was made between events for which the error in the statistical prediction was due to errors in the predicted wave heights and events for which the error was due to a failure in the regression relation.

The results of the evaluation of our method for 12-, 24- and 48-h-lead-time forecasts showed that the mean error is smaller than the mean error of the ET-SURGE model forecasts with 95% confidence and is statistically indistinguishable from the NOAA operational forecast,
which is the ET-SURGE output with an anomaly correction. An advantage that the statistical method offers relates to the limitations of the error-correction technique used by NOAA. The lead time of NOAA’s operational storm-surge forecast is limited to the time span over which the anomaly correction can be assumed to be constant. In the context of weather predictions that go out for 10 days or more, operational seasonal predictions, and longer-term climate forecasts, it would be highly desirable to explore corrections to the dynamical storm-surge forecast that do not require constancy of these corrections throughout the forecast.

The method presented here establishes a robust prediction of the storm-maximum storm surge associated with a particular forecast storm that could provide valuable information as an element of operationally issued storm warnings. The results lend confidence to the usefulness of our statistical technique as an element of an improved error-correction method for use in operational surge forecasts. This application of our method would require a quantitative characterization of the error of the statistical prediction in terms of the nature, location, and strength of the storms.

Acknowledgments. Partial support for this work was provided by the CUNY Research Foundation through PSC-CUNY Award 68640-00 37, and by New York Sea Grant under Project R/CCP-14. Author KW holds an Alfred Harcourt Foundation Fellowship and a Catalyst Scholarship funded by the National Science Foundation through Grant 40104-00 01. We appreciate the detailed critical comments from three anonymous reviewers, which led to a much improved paper.

REFERENCES


Unauthenticated | Downloaded 06/15/22 11:27 AM UTC